Universität Stuttgart

Institut für Nachrichtenvermittlung und Datenverarbeitung

Prof. Dr.-Ing. habil. P. J. Kühn

56. Bericht über verkehrstheoretische Arbeiten

VERGLEICHENDE UNTERSUCHUNGEN AN ATM-KOPPELNETZSTRUKTUREN

von

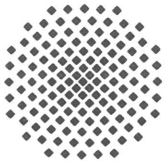
Thomas Theimer

1994

© 1994 Institut für Nachrichtenvermittlung und Datenverarbeitung Universität Stuttgart

Druck: E.Kurz & Co., Druckerei + Reprografie GmbH., Stuttgart

ISBN 3-922403-66-2



University of Stuttgart

Institute of Communications Switching and Data Technics

Prof. Dr.-Ing. habil. P. J. Kühn

56th Report on Studies in Congestion Theory

Comparative Studies on

ATM Switching Network Structures

by

Thomas Theimer

1994

Summary

Asynchronous Transfer Mode (ATM) is the underlying transport mechanism for future high-speed communication networks in the local area as well as in the wide area. ATM is a packet-oriented transport technique based on short packets of fixed size called cells. ATM networks are connection-oriented by definition. A virtual connection has to be set up before any data can be transferred across the network. During the connection setup phase, a path is selected by the network that all cells of the connection will follow. On each link along this path, a logical channel number is assigned to the connection and carried in the cell header. Cells of many connections can be multiplexed onto the same physical transmission link using the asynchronous time division (ATD) multiplexing scheme. Cells are switched using translation tables associating each incoming logical channel number with an outgoing port and a new logical channel number. Although the tables are controlled by software, it is mandatory to implement the actual cell transport in hardware using dedicated high-speed switching fabrics.

A plethora of different architectures for ATM switching networks have been proposed in the literature. One of the objectives of this thesis is to provide a systematic description of ATM switch fabrics by identifying a common set of key characteristics defining each architecture. The second part of the thesis is devoted to the performance evaluation and comparison of major characteristics affecting switch performance.

Chapter 1 introduces the fundamental concepts of B-ISDN specified in various recommendations of CCITT ([35] - [46]). These recommendations set the framework for the design of an ATM switching system. Chapter 1 describes the protocol reference model for B-ISDN as well as the concept of virtual channels and virtual paths, which has been introduced to simplify the operation of ATM networks and to increase their flexibility. Future networks must also provide new capabilities such as point-to-multipoint connections and multi-connection calls, requiring an evolution of existing signalling protocols for ISDN.

ATM networks offer many advantages including the shared use of buffers and transmission links. However, this flexibility comes at the expense of additional traffic control schemes which have to guarantee that each connection receives the requested grade of service in terms of cell loss and delay. During connection establishment the network invokes a connection admission

control function which decides whether the connection can be accepted. This decision is based on a traffic descriptor supplied by the calling party, which is used to estimate the grade of service resulting from an admission of the new connection. When a connection has been accepted, the network must verify that the actual traffic entering the network is conforming with the traffic descriptor declared by the user. Hence, a usage parameter control function is provided at each network access point, which will discard or delay cells violating the traffic parameters agreed. The switches will also support two traffic classes with different cell loss requirements indicated by a cell loss priority bit in the cell header.

A systematic classification and description of ATM switch architectures is given in chapter 2. After summarizing the basic concepts and functions of an ATM switching system, various requirements are defined which have to be taken into account when designing a switch. The remainder of chapter 2 provides a classification of the architectures according to various criteria including the placement and organisation of internal buffers, the cell loss priority schemes, the topology of the switch fabric and its internal operation. Throughout the chapter, performance aspects are already considered in the discussion of possible alternatives.

Performance evaluation of communication systems usually requires a simplified model of the system incorporating its structure and the relevant characteristics of its internal operation. In addition, the external traffic injected into the system has to be defined using suitable traffic models. Chapter 3 starts by introducing a hierarchical modelling approach that is commonly used to characterise the behaviour of traffic sources in ATM networks. Hierarchical models are based on a decomposition of the source behaviour into several layers corresponding to different time scales. Some typical source models are reviewed, and examples are given for models of each layer. Chapter 3 also presents the basic concepts of those analytical methods and simulation techniques which are used extensively throughout the following chapters. Numerous analytical and simulation tools have been developed allowing detailed performance studies of various switch architectures.

ATM switch fabrics are usually multistage networks composed of many single stage switching elements. The overall performance of a switch results from the collective behaviour of all its components. Hence, not only the switching elements, but also the characteristics of the switch fabric as a whole must be optimised to obtain a high performance switch. The performance of single-stage switching elements is discussed in chapter 4. In particular, this chapter compares the performance of different buffering concepts and arbitration strategies. An analytical performance evaluation is presented where possible, but in many cases simulation is the only way to assess the performance.

Aspects such as the location and organisation of cell buffers are of prime importance when designing an ATM switching element. In addition, technological limitations like the maximum clock frequency, the available chip area and the maximum pin count must also be taken into

account. Each architecture has different requirements concerning these parameters, calling for a tradeoff between performance and implementation complexity.

Switching elements with output buffers are known to provide the best throughput and delay performance. On the other hand, this architecture requires large buffers and a high memory bandwidth, so that many other solutions have been proposed. Among these solutions, the following two architectures have been identified as potential alternatives:

- Switching elements with shared buffers minimise the memory size, at the expense of a very high buffer access speed and of additional buffer management schemes.
- Input buffers with random access also allow for a reduction of the memory size while maintaining a moderate access speed. However, in order to provide an adequate performance they need very complex control schemes to select the cells to be transmitted.

An additional aspect is the choice of an appropriate arbitration strategy. An arbitration is required, e.g., in case of a switching element with input buffers when two cells are destined for the same output port. Aiming at a reduction of the memory size, it is preferable to serve the longest queue first, even if this leads to a slightly higher cell delay variation compared to a selection of the oldest cell (FIFO policy).

Chapter 5 concentrates on a cell-level analysis of multistage ATM switch fabrics. Assuming an external Bernoulli arrival process, traditional approaches are based on the hypothesis that internal cell arrivals can be modelled by a Bernoulli process as well. However, detailed studies have revealed that the internal traffic streams display a complex correlation structure which can have a significant impact on the queueing performance. A three-state Markov Modulated model is proposed which is sufficiently accurate to capture the main characteristics of the internal cell arrival processes. This model still allows an analytical performance analysis of an output buffer, yielding the state probabilities and the delay distribution in each stage of the switch.

It is shown in chapter 5 that a Bernoulli process cannot be used as an input traffic model for dimensioning the buffers at internal stages. This is particularly important for switching elements with only few output ports, where the Bernoulli process may lead to a substantial underestimation of the required buffer size. To simplify the buffer dimensioning process, a worst-case model is proposed which accurately approximates the internal traffic after an infinite number of stages. The model is derived from the departure process of a slotted M/D/1 queue and does not depend on the number of input/output ports. An approximation for the delay distribution of the entire switch fabric can be computed from a convolution of the individual delay distributions assuming independence between the delay in successive stages. However, this approximation is not very accurate because of the positive correlation of the delay in neighbouring stages of the switch.

Chapter 6 investigates the effect of various input traffic characteristics on the performance of single- and multistage ATM switches. Using a superposition of several on-off sources, it is shown that the burst-level behaviour of the sources is very important for the switch performance. The cell loss probability depends not only on the peak cell rate and the average burst length, higher moments of the burst length distribution must be taken into account as well. Due to the connection-oriented operation of ATM networks, the burst-level behaviour of a connection remains almost unchanged throughout the entire network. Assuming that connection admission control is based on the burst-level characteristics of a connection, it can be shown that buffer dimensioning using an external Bernoulli arrival process is sufficient.

A comparison of various architectures for single-stage switches with bursty input traffic has revealed some notable effects. It was observed that switching elements with input buffers or shared buffers may in some cases have a significantly higher cell loss probability than output buffered switches. This effect is due to the buffer sharing of cells destined for different output ports, which occurs in both input and shared buffer switches. In case of a short-term overload situation, the buffer is flooded with cells destined for a single output, thus leading to substantial cell loss for all outputs. The impact of such an overload situation, which is unavoidable if statistical multiplexing is applied, must be restricted to the overloaded output port by employing suitable buffer management schemes. A simple yet effective scheme is a limitation of the length of each logical output queue, preventing a single output from hogging the entire buffer.

The overall performance of an ATM switch fabric depends on its cell-level characteristics and on its connection-level behaviour. Chapter 7 presents a discussion of various issues such as the topology of the switching network and the path selection strategy. Using a three-stage Clos network as an example, it is possible to derive conditions under which the switch fabric is strictly nonblocking. Without internal speedup, the switch can only be made nonblocking by applying some restrictions on the minimum or maximum bandwidth of connections. On the other hand, the switching network is always nonblocking if the internal links operate at three times the speed of an external input/output link.

In case of an internally blocking switch fabric, the performance is influenced heavily by the path selection strategy invoked at each connection setup. The optimum path selection strategy depends mainly on the traffic mix. If all connections have the same bit rate requirements (homogeneous traffic), a uniform distribution of connections across all links should be preferred. ATM networks, however, are designed to integrate all traffic types with widely different bit rate requirements. In this scenario the free capacity should be concentrated on as few links as possible, in order to reduce the blocking probability for high bit rate connections. This can be achieved by the well-known sequential hunting scheme with homing, which provides the best performance among the strategies compared. A possible extension of this scheme would be a separation of low and high bit rate connections to reduce their mutual interference.

Inhaltsverzeichnis

Inhaltsverzeichnis v

Abkürzungen xii

Formelzeichen xiv

1 Einleitung 1

1.1 Architektur des Breitband-ISDN 2

1.1.1 Übermittlungsverfahren 2

1.1.1.1 Zeitmultiplextechniken 2

1.1.1.2 Vermittlungsprinzipien 3

1.1.1.3 Eigenschaften des ATM-Übermittlungsverfahrens 4

1.1.2 Zellformat 5

1.1.3 Konzept der Virtuellen Pfade 6

1.1.4 Protokoll-Referenzmodell 9

1.1.4.1 Bitübertragungsschicht 10

1.1.4.2 ATM-Schicht 12

1.1.4.3 ATM-Anpassungsschicht 12

1.1.5 Signalisierung 14

1.1.6 Verkehrssteuerungsverfahren 14

1.1.6.1 Verbindungsannahme 15

1.1.6.2 Quellflußkontrolle 16

1.1.6.3 Prioritäten 17

1.2 Übersicht über die Arbeit 17

2 Architektur von ATM-Koppelnetzen **20**

2.1 Einführung 20

2.2 Anforderungen an eine ATM-Koppelanordnung 23

2.2.1 Funktionalität 23

2.2.2 Verkehrsleistung 24

2.2.3 Herstellung und Betrieb 25

2.3 Klassifizierung der ATM-Koppelnetze 26

2.3.1 Anordnung und Organisation der Puffer 26

2.3.1.1 Ausgangspuffer 28

2.3.1.1.1 Prinzip 28

2.3.1.1.2 Realisierung 29

2.3.1.2 Zentralspeicher 31

2.3.1.2.1 Prinzip 31

2.3.1.2.2 Speicherverwaltung 32

2.3.1.2.3 Realisierung 33

2.3.1.3 Eingangspuffer 36

2.3.1.3.1 Prinzip 36

2.3.1.3.2 Puffer mit FIFO-Abfertigung 36

2.3.1.3.3 Puffer mit wahlfreiem Zugriff 38

2.3.1.4 Verteilte Speicher 42

2.3.1.4.1 Prinzip 42

2.3.1.4.2 Arbitrierungsstrategien 43

2.3.1.4.3 Implementierung 44

2.3.1.5 Kombinierte Pufferkonzepte 44

2.3.1.5.1 Eingangs- und Ausgangspuffer 44

2.3.1.5.2 Eingangs- und Zentralspeicher 46

2.3.1.5.3 Eingangspuffer mit verteilten Speichern 46

2.3.1.5.4 Verteilte Speicher mit Ausgangspuffern 47
2.3.2 Realisierung von Verlustprioritäten 48
2.3.2.1 Wegetrennung 49
2.3.2.2 Getrennte Puffer 49
2.3.2.3 Schwellwertverfahren 50
2.3.2.4 Verdrängungssystem 51
2.3.3 Steuerungsverfahren 52
2.3.3.1 Selbststeuernder Betrieb 52
2.3.3.2 Tabellengesteuerter Betrieb 53
2.3.4 Topologie des Koppelnetzes 53
2.3.4.1 Einpfad-Netzwerke 55
2.3.4.2 Mehrpfad-Netzwerke 58
2.3.4.2.1 Erweiterte Banyan-Netzwerke 58
2.3.4.2.2 Batcher-Banyan-Netzwerk 60
2.3.4.2.3 Clos-Netzwerk 64
2.3.5 Verbindungskonzepte 67
2.3.5.1 Verbindungsorientierter Betrieb 67
2.3.5.2 Verbindungsloser Betrieb 67

3 Verkehrsmodelle und Analyseverfahren **69**

3.1 Modellierung von Verkehrsquellen 70
3.1.1 Verbindungsebene 71
3.1.2 Dialog-Ebene 72
3.1.3 Burst-Ebene 72
3.1.4 Zellebene 73
3.2 Modellierung des Verkehrsflusses 75
3.3 Methoden zur Leistungsuntersuchung 76
3.3.1 Simulationstechniken 76
3.3.2 Analytische Verfahren 78

4 Leistungsuntersuchung einstufiger ATM-Koppelanordnungen 82

4.1 Ausgangspuffer 82

4.2 Zentralspeicher 88

4.2.1 Analyse der Speicherbelegung 88

4.2.2 Speicherverwaltung 92

4.3 Eingangspuffer 94

4.3.1 Eingangspuffer mit FIFO-Abfertigung 95

4.3.2 Eingangspuffer mit wahlfreiem Lesezugriff 100

4.4 Verteilte Puffer 102

4.5 Eingangspuffer mit zusätzlichen Ausgangspuffern 104

4.6 Vergleich und Bewertung der Architekturen 106

5 Analyse mehrstufiger Koppelnetze 108

5.1 Eigenschaften des internen Verkehrs 109

5.2 Modellierung des Ausgangsprozesses 111

5.2.1 Verteilung der Freiperiode 111

5.2.2 Verteilung der Betriebsperiode 111

5.2.2.1 Approximation durch eine geometrische Verteilung 112

5.2.2.2 Approximation durch eine hypergeometrische Verteilung . . . 112

5.2.3 Modellierung durch einen Markov-Modulierten Bernoulli-Prozeß 113

5.2.3.1 Modell 1 114

5.2.3.1.1 Modellbeschreibung 114

5.2.3.1.2 Korrelationsverhalten 115

5.2.3.2 Modell 2 116

5.2.3.2.1 Modellbeschreibung 116

5.2.3.2.2 Korrelationsverhalten 117

5.3 Ankunftsprozeß des Ausgangspuffers 119

5.3.1 Korrelation der Gruppenankünfte 119
5.3.2 Modellierung des Ankunftsprozesses 121
5.4 Analyse des Koppelnetzes 122
5.5 Ergebnisse 123
5.5.1 Verteilung der Warteschlangenlänge 123
5.5.2 Verteilung der Durchlaufzeit 125

6 Verfeinerte Modellierung des Quellverhaltens **128**
6.1 Ein Quellmodell zur gemeinsamen Modellierung von Zell- und Burst-Ebene . . 128
6.2 Modellierung von virtuellen Verbindungen 130
6.3 Einfluß von stoßartigem Verkehr auf das Verhalten eines Ausgangspuffers . . . 131
6.4 Vergleich von einstufigen Anordnungen 136
6.5 Einfluß der virtuellen Verbindungen 139

7 Untersuchung von Koppelnetzstrukturen auf Verbindungsebene **141**
7.1 Intern blockierungsfreie ATM-Koppelnetze 142
7.1.1 Netzwerke ohne interne Geschwindigkeitserhöhung 145
7.1.1.1 Durchschaltevermittlung 145
7.1.1.2 Begrenzte Maximalbitrate 145
7.1.1.3 Begrenzte Mindestbitrate 146
7.1.2 Netzwerke mit interner Geschwindigkeitserhöhung 147
7.1.2.1 Koppelnetze ohne interne Aufweitung 147
7.1.2.2 Koppelnetze mit interner Aufweitung 147
7.2 Wegesuche in mehrstufigen ATM-Koppelnetzen 148
7.2.1 Strategien zur Wegesuche 148
7.2.1.1 Stufenweise Wegesuchverfahren 148
7.2.1.2 Weitspannende Wegesuchverfahren 150
7.2.2 Modellierung 150

7.2.2.1 Systemmodell 150
7.2.2.2 Modellierung der Verkehrsquellen 151
7.2.3 Leistungsvergleich 152
7.2.3.1 Homogenes Verkehrsgemisch 152
7.2.3.2 Inhomogenes Verkehrsgemisch 154

8 Zusammenfassung und Ausblick **156**

Literaturverzeichnis **160**

A Analyse des zeitdiskreten $D^{[X]}/GI/1$-Systems **177**
A.1 Modellbeschreibung 177
A.2 Analyse des Systems mit unbegrenztem Wartespeicher 179
A.2.1 Zustandswahrscheinlichkeiten nach einem Bedienende 179
A.2.1.1 Bearbeitungsreihenfolge AF 182
A.2.1.2 Bearbeitungsreihenfolge DF 183
A.2.2 Zeitgemittelte Zustandswahrscheinlichkeiten 183
A.2.3 Zustandswahrscheinlichkeiten vor der Ankunft einer Anforderung . . . 186
A.2.3.1 Bearbeitungsreihenfolge AF 187
A.2.3.2 Bearbeitungsreihenfolge DF 188
A.2.4 Wartezeitverteilung 188
A.2.5 Verteilung der Betriebsperiode 190
A.3 Analyse des Systems mit begrenztem Wartespeicher 193
A.3.1 Zustandswahrscheinlichkeiten nach einem Bedienende 193
A.3.2 Zeitgemittelte Zustandswahrscheinlichkeiten 195
A.3.2.1 Bearbeitungsreihenfolge AF 195
A.3.2.2 Bearbeitungsreihenfolge DF 196
A.3.3 Zustandswahrscheinlichkeiten vor der Ankunft einer Anforderung . . . 198

A.3.3.1 Bearbeitungsreihenfolge AF 198

A.3.3.2 Bearbeitungsreihenfolge DF 199

A.3.4 Verteilung der Betriebsperiode 199

A.3.4.1 Zeitdiskrete Phasenverteilungen 200

A.3.4.2 Bearbeitungsreihenfolge AF 201

A.3.4.3 Bearbeitungsreihenfolge DF 202

B Verteilung der Durchlaufzeit im $D^{[X]}/D/1$-System mit zufälliger Abfertigung und konstanter Bedienzeit $H = 1$ **203**

C Der Markov-Modulierte Bernoulli-Prozeß **206**

C.1 Definition 206

C.2 Charakteristische Größen 206

C.3 Analyse des Korrelationsverhaltens 207

Abkürzungen

AAL	ATM Adaptation Layer
AF	Arrival First
ATD	Asynchronous Time Division Multiplexing
ATM	Asynchronous Transfer Mode
B-ISDN	Broadband Integrated Services Digital Network
CAC	Connection Admission Control
CBR	Constant Bit Rate
CC	Crossconnect
CCITT	Comité Consultatif International Téléphonique et Télégraphique
CLP	Cell Loss Priority
CS	Convergence Sublayer
DF	Departure First
DBMAP	Discrete-Time Batch Markovian Arrival Process
DMAP	Discrete-Time Markovian Arrival Process
FIFO	First In First Out (Abfertigung in der Reihenfolge des Eintreffens)
GFC	Generic Flow Control
GMDP	Generally Modulated Deterministic Process
HDTV	High Definition Television
HEC	Header Error Control
HOL	Head of Line (vorderste Position einer Warteschlange)
HVSt	Hauptvermittlungsstelle
ISDN	Integrated Services Digital Network
LIFO	Last In First Out (Abfertigung in der umgekehrten Reihenfolge des Eintreffens)
MMBP	Markov-Modulierter Bernoulli-Prozeß
MMPP	Markov-Modulierter Poisson-Prozeß
NNI	Network Node Interface
NPC	Network Parameter Control
OVSt	Ortsvermittlungsstelle
PA	Protokollabschluß
PHY	Physical Layer
PM	Physical Medium (Sublayer)
PT	Payload Type
SAR	Segmentation and Reassembly (Sublayer)
SDH	Synchronous Digital Hierarchy
SIRO	Service In Random Order (Abfertigung in zufälliger Reihenfolge)
STD	Synchronous Time Division Multiplexing
STM	Synchronous Transfer Mode

TC	Transmission Convergence (Sublayer)
UNI	User Network Interface (Benutzer-Netz-Schnittstelle)
UPC	Usage Parameter Control
VBR	Variable Bit Rate
VC	Virtual Channel
VCC	Virtual Channel Connection
VCI	Virtual Channel Identifier
VP	Virtual Path
VPI	Virtual Path Identifier
VPC	Virtual Path Connection

Formelzeichen

A	Verkehrsangebot
B	Dauer einer Betriebsperiode (Busy Period)
b_k	Verteilung von B
$b(z)$	Erzeugende Funktion von b_k
H	Bedienzeit einer Anforderung
h_k	Verteilung von H
$h(z)$	Erzeugende Funktion von h_k
Ξ	Anzahl der Anforderungen in einer Gruppenankunft
ξ_k	Verteilung von ξ
$\xi(z)$	Erzeugende Funktion von ξ_k
V	Anzahl der während einer Bedienzeit eintreffenden Anforderungen
v_k	Verteilung von V
$v(z)$	Erzeugende Funktion von v_k
W	Wartezeit einer Anforderung
w_k	Verteilung von W
$w(z)$	Erzeugende Funktion von w_k
X	Anzahl der Anforderungen im System unmittelbar nach einem Bedienende
x_k	Verteilung von X
$x(z)$	Erzeugende Funktion von x_k
Y	Zustand des Systems im zeitlichen Mittel
y_k	zeitgemittelte Zustandswahrscheinlichkeiten
$y(z)$	Erzeugende Funktion von y_k
Z	Zustand des Systems vor der Ankunft einer Anforderung
z_k	Verteilung von Z
$z(z)$	Erzeugende Funktion von z_k

Kapitel 1

Einleitung

Der technologische Fortschritt der letzten Jahre auf den Gebieten der optischen Übertragungstechnik, der Halbleitertechnik sowie der Informationsverarbeitung ist die treibende Kraft für die Entwicklung neuer Breitband-Dienste und der entsprechenden Kommunikations-Infrastruktur. Einerseits wird durch den technischen Fortschritt die Nachfrage nach Breitband-Diensten verstärkt, andererseits sind auch die technischen Möglichkeiten zur Realisierung leistungsfähiger Kommunikationsnetze vorhanden.

Aus heutiger Sicht ist vor allem ein steigender Bedarf an Hochgeschwindigkeits-Datendiensten zu verzeichnen, der zunächst im lokalen und regionalen Bereich durch Kommunikationsnetze wie FDDI (Fiber Distributed Data Interface [177]) und DQDB (Distributed Queue Dual Bus [182]) befriedigt werden soll. Die Standardisierung von FDDI und DQDB ist abgeschlossen, und entsprechende Produkte sind bereits auf dem Markt oder kurz vor der Markteinführung.

Im Weitverkehrsbereich wird ebenfalls intensiv an einem weltweit einheitlichen Breitbandnetz gearbeitet, dessen Einsatzmöglichkeiten jedoch über die reine Datenkommunikation weit hinaus gehen sollen. Das Breitband-ISDN ist als universelles Kommunikationsnetz der Zukunft konzipiert, in dem alle bekannten und auch alle neu hinzukommenden Dienste integriert werden können.

Eine hohe Flexibilität ist eine der wichtigsten Anforderungen an ein solches Netz, da die Eigenschaften zukünftiger Dienste nur schwer vorhersehbar sind, und auch die heute existierenden Anwendungen bereits sehr unterschiedliche Anforderungen bezüglich Bandbreite und Dienstgüte stellen. Aus diesem Grund wird im Breitband-ISDN ein Übermittlungsverfahren verwendet, das als *Asynchronous Transfer Mode* (ATM) bekannt ist. Im Gegensatz zum Schmalband-ISDN arbeitet das ATM-basierte Breitband-ISDN paketorientiert, wobei alle Informationen in Form von kurzen Blöcken mit fester Länge transportiert werden. Auf diese Weise können Dienste mit nahezu beliebiger Datenrate unterstützt werden, wobei die Bitrate auch innerhalb einer Verbindung schwanken kann.

Die Realisierung eines breitbandigen Kommunikationsnetzes auf der Basis von ATM erfordert neben einer geeigneten Übertragungstechnik auch leistungsfähige vermittlungstechnische Einrichtungen, die hochratige Datenströme bis zu etwa 600 Mbit/s vermitteln können. Mit ersten Prototypen auf der Basis von ATM-ähnlichen Übermittlungsverfahren wurde bereits 1984 experimentiert [98],[206], und seit der offiziellen Festlegung des CCITT auf das ATM-Prinzip im Jahr 1988 arbeiten viele Forschungslabors und Hersteller von Vermittlungssystemen an entsprechenden Entwicklungen. Diese werden auch von der Europäischen Gemeinschaft im Rahmen von Forschungsprogrammen wie RACE (Research and Development in Advanced Communication Technologies in Europe) und ESPRIT (European Strategic Programme for Research and Development in Information Technology) gefördert.

Das Ziel dieser Arbeit ist eine systematische Klassifizierung der verschiedenen ATM-Koppelnetzarchitekturen auf Grund ihrer wesentlichen Eigenschaften sowie eine Untersuchung der Verkehrsleistung der wichtigsten Systemvarianten. Der Vergleich unterschiedlicher Realisierungen und die Auswahl der für einen bestimmten Anwendungsfall am besten geeigneten Lösung setzen jedoch eine genaue Kenntnis der Rahmenbedingungen voraus, unter denen eine ATM-Koppeleinrichtung betrieben wird. Es folgt daher zunächst eine kurze Beschreibung der Architektur des Breitband-ISDN, wobei insbesondere auf solche Aspekte eingegangen wird, die für die Realisierung der Vermittlungsknoten von Bedeutung sind.

1.1 Architektur des Breitband-ISDN

1.1.1 Übermittlungsverfahren

Das Übermittlungsverfahren bildet die Grundlage für die Kommunikation in einem Nachrichtennetz und kennzeichnet sowohl die Übertragung als auch die Vermittlung von Informationen. Bevor auf die wichtigsten Eigenschaften des ATM-Übermittlungsverfahrens eingegangen wird, soll zunächst ein Überblick über diejenigen Vermittlungsprinzipien und Multiplextechniken gegeben werden, die prinzipiell für den Einsatz im Breitband-ISDN geeignet erscheinen.

1.1.1.1 Zeitmultiplextechniken

Die Multiplextechnik beschreibt die gemeinsame Nutzung eines Kanals durch mehrere Kommunikationsbeziehungen. Typische Multiplextechniken sind Raummultiplex, Frequenz- oder Wellenlängenmultiplex, Zeitmultiplex und Kodemultiplex. Der oben genannte Kanal kann selbst wiederum Bestandteil eines Multiplex-Signals höherer Ordnung sein, wodurch sich sehr komplexe Multiplex-Hierarchien aufbauen lassen.

Im Zeitalter der Digitaltechnik hat vor allem die Zeitmultiplextechnik große Bedeutung erlangt. Sie beruht darauf, daß die einzelnen Informationsströme zeitlich ineinander verschachtelt werden. Man unterscheidet hierbei zwischen synchronem Zeitmultiplex (Synchronous Time Division Multiplexing, STD) und asynchronem Zeitmultiplex (Asynchronous Time Division Multiplexing, ATD). Im synchronen Zeitmultiplex wird der Kanal in periodisch wiederkehrende Rahmen unterteilt, welche wiederum aus einer festen Anzahl von Zeitschlitzen bestehen. Jeder Kommunikationsbeziehung sind im allgemeinen mehrere Zeitschlitze fest zugeordnet, wobei die Identifizierung der Zeitschlitze ausschließlich auf Grund ihrer Position innerhalb des Rahmens erfolgt. Dieses Prinzip wird daher im Englischen auch „Position Multiplexing" genannt.

Die asynchrone Zeitmultiplextechnik ist dadurch gekennzeichnet, daß die Informationsblöcke unterschiedlicher Kommunikationsbeziehungen in beliebiger Reihenfolge (asynchron) aneinander gereiht werden können. Jeder Informationsblock enthält entweder die vollständige Zieladresse des Empfängers oder eine logische Kanalnummer, welche die Zuordnung zu einer Kommunikationsbeziehung ermöglicht. Die asynchrone Zeitmultiplextechnik wird daher in der Literatur auch als „Adreßmultiplex" oder „Label Multiplexing" bezeichnet.

Die Länge dér Informationsblöcke im asynchronen Zeitmultiplex kann sowohl fest als auch variabel sein. Obwohl die Abkürzung ATD von einigen Autoren nur im Zusammenhang mit einer festen Blocklänge gebraucht wird, soll dieser Begriff hier in seiner allgemeinen Form belassen werden. Informationsblöcke mit variabler Länge bieten zwar prinzipiell die größtmögliche Flexibilität, erfordern aber auch einen höheren Aufwand zur Erkennung der Blockgrenzen sowie zur Speicherverwaltung in den Netzknoten. In vielen Hochgeschwindigkeits-Kommunikationsnetzen wurde daher eine feste Blocklänge gewählt.

1.1.1.2 Vermittlungsprinzipien

Das Vermittlungsprinzip kennzeichnet den Transport von Nachrichtenblöcken durch ein Kommunikationsnetz. Es besteht daher eine enge Beziehung zwischen dem Vermittlungsprinzip und der zugrundeliegenden Multiplextechnik. Die beiden grundlegenden Vermittlungsprinzipien sind die Durchschaltevermittlung und die Paketvermittlung. Darüber hinaus existieren verschiedene Varianten dieser Verfahren, wie die Mehrkanaldurchschaltung, die schnelle Durchschaltevermittlung und die schnelle Paketvermittlung.

Die Durchschaltevermittlung ist dadurch gekennzeichnet, daß zwischen kommunizierenden Endgeräten ein durchgehender Kanal mit fester Bandbreite geschaltet wird. Für die Dauer der Verbindung steht dieser Kanal den Kommunikationspartnern zur ausschließlichen Nutzung zur Verfügung, so daß eine gegenseitige Beeinflussung bestehender Verbindungen ausgeschlossen ist.

In der Paketvermittlungstechnik werden alle Datenströme in einzelne Pakete unterteilt, die abschnittsweise auf Grund von Steuerinformationen im Paketkopf übertragen werden. Falls der entsprechende abgehende Kanal momentan belegt ist, müssen Pakete kurzzeitig zwischengespeichert werden, so daß zufällige Verzögerungen auftreten können. Da hier mehrere Kommunikationsbeziehungen um die Bandbreite eines Kanals konkurrieren, ist eine gegenseitige Beeinflussung nicht zu vermeiden.

In einem paketvermittelnden Netz ist sowohl eine verbindungsorientierte als auch eine verbindungslose Kommunikation möglich. Die verbindungslose Kommunikation erfolgt mit Hilfe von Paketen, welche die vollständige Adresse des Empfängers enthalten (sogenannte Datagramme). Datagramme werden völlig unabhängig voneinander durch das Netz transportiert und können ihr Ziel auch auf unterschiedlichen Wegen erreichen, so daß die Reihenfolge der Pakete am Empfänger nicht garantiert werden kann.

Die verbindungsorientierte Kommunikation erfordert dagegen den Aufbau einer virtuellen Verbindung als Voraussetzung für den Transfer von Paketen. Während des Verbindungsaufbaus wird der Weg der virtuellen Verbindung festgelegt, die erforderlichen Betriebsmittel werden bereitgestellt, und auf jedem Abschnitt entlang dieses Weges wird der Verbindung eine logische Kanalnummer zugewiesen. Die Zuordnung zwischen den einzelnen Verbindungsabschnitten erfolgt mit Hilfe von Tabellen, die gleichzeitig auch den entsprechenden Zielausgang der Pakete enthalten. Da alle Pakete einer virtuellen Verbindung auf dem selben Weg transportiert werden, sind keine Reihenfolgefehler möglich. Darüber hinaus können zusätzliche Protokollfunktionen, wie abschnittsweise Flußsteuerung und Fehlerbehandlung, unterstützt werden.

1.1.1.3 Eigenschaften des ATM-Übermittlungsverfahrens

Das ATM-Übermittlungsverfahren basiert auf der asynchronen Zeitmultiplextechnik. Alle Informationen werden in Form von kurzen Blöcken mit fester Länge übertragen, die in der Terminologie des CCITT als *Zellen* bezeichnet werden. Jede Zelle besteht aus einem Informationsfeld mit einer Länge von 48 Bytes und aus einem 5 Byte langen Steuerfeld, dem sogenannten Zellkopf. Das Informationsfeld steht im wesentlichen für Benutzerdaten zur Verfügung, wogegen der Zellkopf für netzinterne Steuerfunktionen benötigt wird.

Entsprechend der asynchronen Zeitmultiplextechnik erfolgt auch die Vermittlung der Zellen in einem ATM-Netz nach dem Prinzip der Paketvermittlung. Dabei wurde eine verbindungsorientierte Betriebsart auf der Basis von logischen Kanalnummern gewählt. Die logische Kanalnummer ist neben anderen Steuerinformationen im Zellkopf enthalten (siehe auch 1.1.2).

Im Unterschied zur konventionellen Paketvermittlungstechnik (beispielsweise nach dem X.25-Standard) werden in einem ATM-Netz nur sehr einfache abschnittsweise Transportprotokolle

verwendet. Wegen der geringen Bitfehlerwahrscheinlichkeiten auf optischen Übertragungsstrecken kann auf eine abschnittsweise Fehlerbehandlung völlig verzichtet werden. Nur der Zellkopf wird durch einen Mechanismus zur Fehlerkorrektur geschützt, um die fälschliche Zuordnung einer gestörten Zelle zu einer anderen Verbindung zu vermeiden.

Das ATM-Übermittlungsverfahren wurde vor allem wegen seiner Flexibilität als Grundlage für das Breitband-ISDN gewählt [37]. Die paket-orientierte Vermittlungstechnik besitzt keine festen Kanalstrukturen und unterstützt daher nahezu beliebige Bitraten bis zur maximalen Kapazität einer Leitung. Auf diese Weise lassen sich Dienste mit völlig verschiedenen Eigenschaften und Bandbreite-Anforderungen in einem einheitlichen Netz integrieren. Die asynchrone Zeitmultiplextechnik bringt jedoch auch neue Probleme mit sich, wie zum Beispiel eine variable Zellverzögerung und Zellverluste durch Pufferüberlauf in den Vermittlungsknoten. Diese Effekte beeinflussen die Dienstgüte und müssen daher bei bestimmten Anwendungen durch entsprechende Maßnahmen kompensiert werden (siehe 1.1.4.3).

1.1.2 Zellformat

Die Festlegung des genauen Zellformats erfordert eine Abwägung verschiedener Aspekte. Zellen mit kurzem Informationsfeld haben Vorteile bezüglich der Verzögerung der Nutzdaten, wobei insbesondere die für das Auffüllen der Zellen benötigte Zeitdauer bei Diensten mit geringer Bitrate (beispielsweise 64 kbit/s) ein kritischer Parameter ist. Auch der Speicherplatz, der für das Zwischenspeichern der Zellen in den Netzknoten vorgesehen werden muß, nimmt proportional mit der Zellgröße zu. Andererseits weisen kurze Zellen ein ungünstigeres Verhältnis zwischen Steuer- und Nutzdaten auf und erfordern außerdem eine höhere Verarbeitungsgeschwindigkeit zur Auswertung der Steuerinformationen. Die Festlegung der Größe des Informationsfeldes auf 48 Bytes sowie des Zellkopfes auf 5 Bytes ist daher ein Kompromiß zwischen den genannten Vor- und Nachteilen.

Die Struktur des Zellkopfes ist in der Empfehlung I.361 detailliert beschrieben [42]. Die Aufteilung der einzelnen Felder hängt davon ab, ob sich die Zelle an der Benutzer-Netz-Schnittstelle (User Network Interface, UNI) oder an einer netzinternen Schnittstelle (Network Node Interface, NNI) befindet. Bild 1.1 zeigt die genaue Aufteilung des Zellkopfes an beiden Schnittstellen.

Der wichtigste Bestandteil des Zellkopfes ist die logische Kanalnummer, die hier aus zwei Teilen besteht: einer Kennung für den virtuellen Kanal (Virtual Channel Identifier, VCI) und einer Kennung für den virtuellen Pfad (Virtual Path Identifier, VPI). Ein virtueller Pfad ist ein logisches Bündel mehrerer virtueller Verbindungen, die gemeinsam auf einem Verbindungsabschnitt transportiert und als einheitlicher Zellstrom behandelt werden. Eine ausführliche Beschreibung des Konzeptes der virtuellen Pfade folgt in Abschnitt 1.1.3.

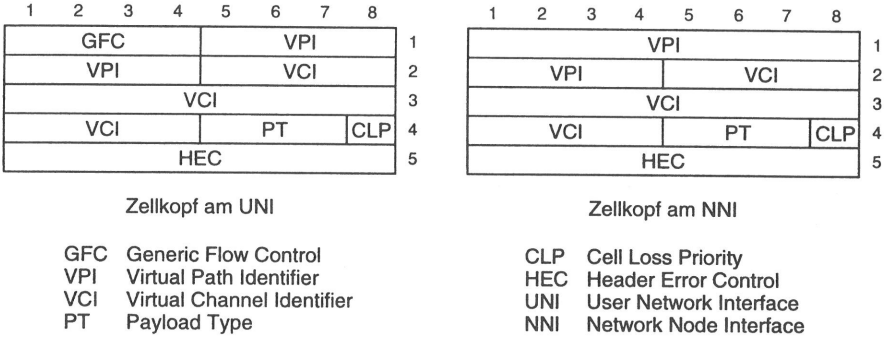


Bild 1.1: Struktur des Kopffeldes einer ATM-Zelle

Die Länge des VCI-Feldes beträgt an beiden Schnittstellen 16 Bit, das VPI-Feld umfaßt am NNI 12 Bit sowie am UNI 8 Bit. An der Benutzer-Netz-Schnittstelle ist ein zusätzliches, 4 Bit großes Feld vorhanden, das durch eine Flußsteuerung (Generic Flow Control, GFC) den Anschluß mehrerer Endgeräte an eine Schnittstelle ermöglichen soll (siehe auch 1.1.4.2).

Mit Hilfe des PT-Feldes (Payload Type, PT) kann zwischen Zellen mit Benutzerdaten und Zellen mit netzinternen Daten unterschieden werden. Beispielsweise tauschen auch die Endgeräte innerhalb einer virtuellen Verbindung bestimmte Daten aus, die für ATM-spezifische Betriebs- und Wartungsaufgaben benötigt werden und nicht für den Benutzer bestimmt sind. Weitere Funktionen des PT-Feldes sind die Anzeige von Überlastsituationen an die Endgeräte und die Unterstützung einer schnellen Betriebsmittelverwaltung [42].

Das CLP-Bit (Cell Loss Priority, CLP) kennzeichnet Zellen, die bei einer Überlastsituation des Netzes bevorzugt verworfen werden. Dieses Bit wird von den Endgeräten verwendet, um Zellen mit geringerer Bedeutung zu markieren. Es kann aber auch von netzinternen Funktionen wie der Quellflußkontrolle (siehe 1.1.6.2) verändert werden.

Sämtliche Steuerinformationen des Zellkopfes werden schließlich durch das Hinzufügen von 8 Prüfbits (Header Error Control, HEC) gegen Bitfehler gesichert. Der verwendete Kode ermöglicht die Korrektur von Einzel-Bitfehlern sowie die Erkennung bestimmter Mehrfach-Bitfehler. Das HEC-Feld wird außerdem zur Erkennung der Zellgrenzen benutzt (siehe auch 1.1.4.1).

1.1.3 Konzept der Virtuellen Pfade

Ein virtueller Pfad bietet die Möglichkeit, mehrere virtuelle Verbindungen zusammenzufassen und als einheitlichen Zellstrom zu behandeln [36], [95]. Die einzelnen Verbindungen in einem

virtuellen Pfad besitzen einen gemeinsamen VPI und unterscheiden sich nur durch ihr VCI-Feld. Andererseits können virtuelle Verbindungen in unterschiedlichen virtuellen Pfaden auch identische VCI-Felder aufweisen, für die eindeutige Zuordnung einer Zelle zu einer virtuellen Verbindung werden daher sowohl VCI als auch VPI benötigt.

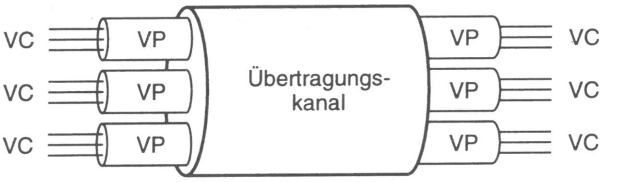


Bild 1.2: Beziehung zwischen virtuellem Kanal (VC), virtuellem Pfad (VP) und dem zugrundeliegenden Übertragungskanal

Wie in Bild 1.2 dargestellt, umfaßt ein physikalischer Übertragungskanal im allgemeinen mehrere virtuelle Pfade, die wiederum verschiedene virtuelle Kanäle enthalten können. In der Terminologie des CCITT wird dabei zwischen dem Konzept der virtuellen Pfade bzw. Kanäle und den zugehörigen *Verbindungen* und *Verbindungsabschnitten* unterschieden. Die folgenden Begriffe sind in der Empfehlung I.113 definiert [36]:

- Ein *VC-Verbindungsabschnitt* (Virtual Channel Link, VC-Link) ermöglicht den Transport von ATM-Zellen von einem Punkt, an dem das VCI-Feld gefüllt oder übersetzt wird, bis zu dem Punkt, an dem der VCI erneut übersetzt oder entfernt wird. Auf einem VC-Verbindungsabschnitt ist also jeder VC-Verbindung ein fester VCI zugeordnet.
- Eine *VC-Verbindung* (Virtual Channel Connection, VCC) besteht aus einer Folge von VC-Verbindungsabschnitten und erstreckt sich zwischen zwei Endpunkten, an denen ein Übergang zur ATM-Anpassungsschicht (AAL, siehe 1.1.4) stattfindet. Die Reihenfolge der Zellen einer VC-Verbindung wird vom Netz nicht verändert.
- Ein *VP-Verbindungsabschnitt* (Virtual Path Link, VP-Link) ermöglicht den Transport von ATM-Zellen von einem Punkt, an dem das VPI-Feld gefüllt oder übersetzt wird, bis zu dem Punkt, an dem der VPI erneut übersetzt oder entfernt wird. Auf einem VP-Verbindungsabschnitt ist also jeder VP-Verbindung ein fester VPI zugeordnet.
- Eine *VP-Verbindung* (Virtual Path Connection, VPC) besteht aus einer Folge von VP-Verbindungsabschnitten und erstreckt sich zwischen zwei Endpunkten, an denen das VCI-Feld verändert (gefüllt, übersetzt oder entfernt) wird. Innerhalb einer VP-Verbindung wird der VCI also nicht verändert.

Der Zusammenhang zwischen VC-Verbindung, VC-Verbindungsabschnitt, VP-Verbindung sowie VP-Verbindungsabschnitt ist in Bild 1.3 veranschaulicht. Jede virtuelle (VC-)Verbindung setzt sich aus mehreren (VC-)Verbindungsabschnitten zusammen, die jeweils durch

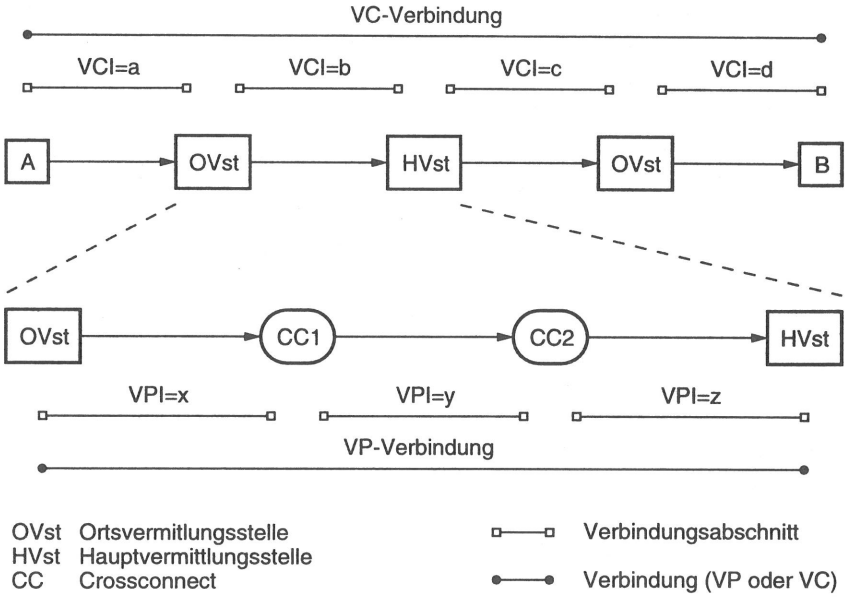


Bild 1.3: Hierarchische Beziehung zwischen virtuellem Kanal und virtuellem Pfad

einen eigenen VCI gekennzeichnet sind. Jeder VC-Verbindungsabschnitt verwendet eine VP-Verbindung, um die Zellen zwischen seinen Endpunkten zu transportieren. Außerdem kann eine VP-Verbindung mehrere VC-Verbindungen enthalten.

Bild 1.3 zeigt das Beispiel einer typischen (VC-)Verbindung, die über eine Ortsvermittlungsstelle (OVSt), eine Hauptvermittlungsstelle (HVSt) und eine weitere Ortsvermittlungsstelle führt. Die Vermittlungsstellen sind jeweils durch eine VP-Verbindung miteinander verbunden, die von den einzelnen VC-Verbindungsabschnitten benutzt werden kann. Diese VP-Verbindung kann eine direkte physikalische Leitung sein, sie kann aber auch aus mehreren VP-Verbindungsabschnitten bestehen. Die in Bild 1.3 dargestellte VP-Verbindung führt über mehrere VP-Verbindungsabschnitte, die jeweils durch einen VP-Vermittlungsknoten (auch VP-Crossconnect genannt) miteinander verbunden sind. Aus übertragungstechnischer Sicht kann ein VP-Verbindungsabschnitt noch weiter unterteilt werden, die genaue Strukturierung ist hier jedoch nicht von Bedeutung.

Eine VP-Vermittlung schließt einen VP-Verbindungsabschnitt ab und muß daher das VPI-Feld übersetzen. Die Funktionsweise einer VP-Vermittlung unterscheidet sich nur dadurch von einer VC-Vermittlung, daß ausschließlich das VPI-Feld ausgewertet wird und somit kleinere Übersetzungstabellen verwendet werden können. Außerdem benötigt ein VP-Vermitt-

lungsknoten keine Teilnehmer-bezogenen Signalisierfunktionen, da die virtuellen Pfade durch das Netz verwaltet werden. Im Vergleich zu einem VC-Vermittlungsknoten ist die Komplexität einer VP-Vermittlungsstelle daher wesentlich geringer.

Virtuelle Pfade verbinden nicht nur die Netzknoten untereinander, sie können auch zwischen Benutzern und Netzknoten oder direkt zwischen zwei Benutzern eingerichtet werden. Der Aufbau einer VP-Verbindung erfolgt entweder auf Antrag durch den Netzbetreiber oder durch entsprechende Signalisier- bzw. Netzmanagementprozeduren, die jedoch noch nicht genauer spezifiziert sind.

Das Konzept der virtuellen Pfade hat eine Reihe verschiedener Anwendungen. Beispielsweise vereinfachen netzinterne virtuelle Pfade vor allem den Betrieb und die Steuerung des Netzes, da eine VP-Verbindung wie ein physikalischer Übertragungskanal zwischen zwei VC-Vermittlungen angesehen werden kann. Beim Aufbau einer VC-Verbindung sind daher nur die Endpunkte der virtuellen Pfade beteiligt, wodurch der Verbindungsaufbau stark beschleunigt wird. Auch die Flexibilität eines solchen Netzes ist wesentlich größer, da virtuelle Pfade bei Bedarf beliebig geändert werden können. Insbesondere bei Leitungsausfällen ist auf diese Weise eine schnelle Umkonfigurierung des Netzes auf VP-Ebene möglich (Protection Switching). Weitere Anwendungsbeispiele sind die Kopplung von lokalen Netzen oder die Anbindung der Teilnehmer an spezielle Server für verbindungslose Dienste.

1.1.4 Protokoll-Referenzmodell

In modernen Kommunikationssystemen werden die verschiedenen Protokollfunktionen meist hierarchisch gegliedert und in sogenannte *Schichten* unterteilt. Die Funktionen der Schichten und die Beziehungen zwischen den einzelnen Schichten werden dabei in einem Protokoll-Referenzmodell definiert. Ein bekanntes Beispiel eines solchen Protokollmodells ist das 7-Schichten-Modell der ISO [107].

Das Protokoll-Referenzmodell des Breitband-ISDN ist ebenfalls in mehrere Schichten unterteilt. Wie bereits beim Schmalband-ISDN kennt das Protokollmodell außerdem verschiedene Ebenen, welche die Trennung der Nutzdatenübertragung von den Signalisier- und Managementfunktionen widerspiegeln. Bild 1.4 zeigt die vom CCITT standardisierte Architektur des Protokoll-Referenzmodells für das Breitband-ISDN [40].

Die Benutzerebene enthält alle Protokollfunktionen, die für den Transfer von Benutzerdaten benötigt werden. Der Aufbau von virtuellen Verbindungen sowie deren Überwachung erfolgt mit Hilfe der Steuerebene, welche sämtliche Signalisierprotokolle enthält. Die Verwaltung der einzelnen Schichten beider Ebenen wird schließlich von der Management-Ebene durchgeführt, welche in Funktionen des Ebenen-Managements und des Schichten-Managements unterteilt

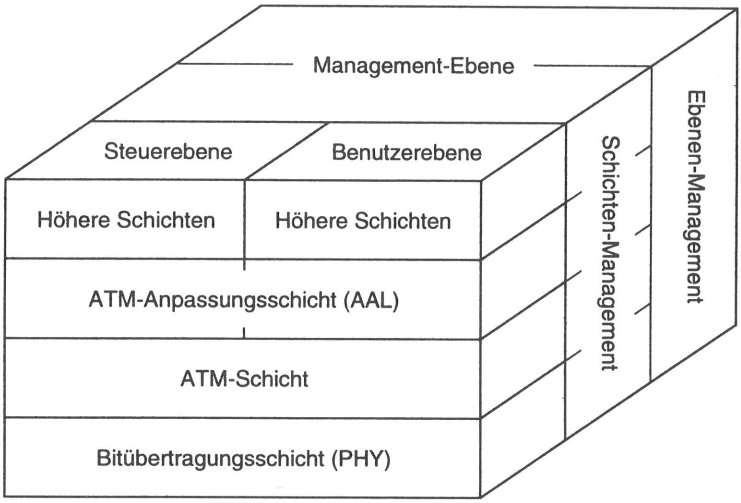


Bild 1.4: Protokoll-Referenzmodell des Breitband-ISDN

ist. Das Schichten-Management verwaltet die Betriebsmittel und Parameter aller Schichten und ist daher entsprechend strukturiert. Das Ebenen-Management ist dagegen für die Koordination zwischen den Ebenen und die Verwaltung gemeinsamer Daten verantwortlich.

Bislang sind nur die Funktionen der untersten drei Schichten festgelegt, die höheren Protokollschichten sind noch nicht genau definiert. Sowohl die Bitübertragungsschicht als auch die ATM-Schicht werden von Steuer- und Benutzerebene gemeinsam benutzt, alle höheren Protokollschichten ab der ATM-Anpassungsschicht sind von der jeweiligen Anwendung abhängig. Eine Übersicht über die Funktionalität der einzelnen Schichten sowie deren weitere Unterteilung in Subschichten ist in Bild 1.5 dargestellt. Die Funktionen der einzelnen Schichten werden, soweit bekannt, in den folgenden Abschnitten genauer beschrieben.

1.1.4.1 Bitübertragungsschicht

Die Bitübertragungsschicht ist unterteilt in die PM-Subschicht (Physical Medium Sublayer, PM) und die TC-Subschicht (Transmission Convergence Sublayer, TC). Die PM-Subschicht enthält nur die vom physikalischen Übertragungsmedium abhängigen Funktionen wie Bit-Synchronisation, Leitungskodierung und, sofern notwendig, elektrisch-optische Wandlung. Das physikalische Medium ist in den meisten Fällen eine Glasfaser, es sind jedoch auch andere Lösungen wie Koaxialkabel möglich. Die PM-Subschicht stellt die eigentliche Bitübertragungsfunktion bereit und entspricht somit der Bitübertragungsschicht des ISO-Protokollmodells.

Bild 1.5: Strukturierung und Funktionalität der Schichten des Protokoll-Referenzmodells

Zwischen der ATM-Schicht und der PM-Subschicht befindet sich die TC-Subschicht, in welcher der Zellstrom an das verwendete Übertragungssystem angepaßt wird. Für die Benutzer-Netz-Schnittstelle wurden bisher zwei unterschiedliche Übertragungssysteme standardisiert: ein Rahmen-orientiertes System auf der Basis der Synchronen Digitalen Hierarchie (SDH) und ein Zell-orientiertes Format [46]. Beide Übertragungssysteme arbeiten mit einer Gesamt-Bitrate von 155.520 Mbit/s oder 622.080 Mbit/s und weisen einen Overhead-Anteil von 1/27 der Gesamtkapazität auf. Dieser Overhead wird für übertragungstechnische Betriebs- und Wartungsfunktionen benötigt.

Unabhängig vom verwendeten Übertragungssystem enthält die TC-Subschicht weitere Funktionen wie

- *Fehlerkorrektur:* In Senderichtung wird die Fehlerkorrektur des Zellkopfes erzeugt und in das HEC-Feld eingefügt. In Empfangsrichtung erfolgt eine Überprüfung der Fehlerkorrektur, wobei Zellen mit nicht korrigierbaren Fehlern verworfen werden.
- *Zell-Synchronisation:* Die Erkennung der Zellgrenzen erfolgt ebenfalls mit Hilfe der Fehlerkorrektur. Dabei wird das empfangene Signal bitweise nach einem Zellkopf mit gültiger Fehlerkorrektur durchsucht. Falls ein solcher Zellkopf gefunden wird, befin-

det sich an dieser Stelle mit hoher Wahrscheinlichkeit der Anfang einer Zelle. Die Überprüfung wird nun Zelle für Zelle fortgesetzt, und nach weiteren 6 gültigen HEC-Feldern kann von einer erfolgreichen Zellsynchronisation ausgegangen werden.

- *Entkopplung der Zellraten:* Zur Anpassung zwischen der Zellrate der ATM-Schicht und der Nutzlast des Übertragungssystems werden Leerzellen (Idle Cells) eingefügt und wieder entfernt. Diese Leerzellen sind durch einen speziellen Zellkopf gekennzeichnet und werden nicht an die ATM-Schicht weitergegeben.

1.1.4.2 ATM-Schicht

Die ATM-Schicht ermöglicht die abschnittsweise Übertragung von ATM-Zellen und ist sowohl vom jeweiligen Dienst als auch vom physikalischen Übertragungsmedium unabhängig. Die Aufgaben der ATM-Schicht beziehen sich ausschließlich auf den Zellkopf und umfassen VC- sowie VP-bezogene Funktionen.

An den Endpunkten einer virtuellen Verbindung der ATM-Schicht muß der Zellkopf erzeugt (Senderichtung) bzw. entfernt werden (Empfangsrichtung), da zwischen ATM-Schicht und AAL-Schicht nur das Informationsfeld der Zellen ausgetauscht wird. Hier erfolgt auch die Zuordnung zwischen einer virtuellen Verbindung und dem entsprechenden Wert des VCI/VPI-Feldes. Die ATM-Schicht führt allerdings keine Fehlerkorrektur durch, das HEC-Feld wird nur von der Bitübertragungsschicht berechnet und überprüft.

Die Unterteilung der logischen Kanalnummer in VCI und VPI spiegelt sich auch in der funktionalen Gliederung der ATM-Schicht wider. Das Multiplexen verschiedener Verbindungen sowie die Übersetzung der logischen Kanalnummer in den Netzknoten erfolgt sowohl auf VC-Ebene als auch auf VP-Ebene. Abhängig von der Art des Vermittlungsknotens sind entweder beide Ebenen oder nur VP-bezogene Funktionen vorhanden (siehe 1.1.3).

Zu den Aufgaben der ATM-Schicht gehört auch die Überwachung der Zellströme aller Verbindungen am Netzzugang sowie an bestimmten Netzübergängen mit Hilfe der Quellflußkontrolle. Eine ausführlichere Beschreibung der Quellflußkontrolle folgt in Abschnitt 1.1.6.2. Für Benutzer-Netz-Schnittstellen mit einer Punkt-zu-Mehrpunkt-Konfiguration ist außerdem eine Flußsteuerung mit Hilfe des GFC-Feldes vorgesehen. Bislang konnte jedoch keine Einigung über die Einzelheiten des entsprechenden Protokolls erzielt werden, so daß diese Funktion noch nicht genauer spezifiziert werden kann.

1.1.4.3 ATM-Anpassungsschicht

Die ATM-Anpassungsschicht (ATM Adaptation Layer, AAL) bildet die Protokoll-Dateneinheiten der höheren Schichten auf das Informationsfeld der Zellen ab und erweitert die Funktionalität der ATM-Schicht entsprechend den Erfordernissen der höheren Protokollschichten.

Die einzelnen Aufgaben sind daher vom jeweiligen Dienst abhängig und beziehen sich ausschließlich auf das Informationsfeld der Zellen. Die Funktionen der ATM-Anpassungsschicht sind üblicherweise in den Endsystemen implementiert, sie können sich aber auch an den Übergängen zu anderen Netzen (z.B. Schmalband-ISDN) befinden.

Um die Anzahl der verschiedenen Anpassungsprotokolle zu begrenzen, wurde vom CCITT eine Unterteilung in vier AAL-Dienstklassen mit unterschiedlichen Eigenschaften vorgeschlagen [43]. Tabelle 1.1 zeigt die Einteilung der Dienstklassen auf Grund der Kriterien

- Zeitbeziehung zwischen Sender und Empfänger,
- Schwankung der Bitrate und
- Verbindungstyp.

Typische Anwendungsbeispiele für die einzelnen Dienstklassen sind:

Klasse A: Emulation von durchschaltevermittelten Kanälen
Klasse B: Video- und Audiodienste mit variabler Bitrate
Klasse C: Verbindungsorientierte Datenkommunikation
Klasse D: Verbindungslose Datenkommunikation

Diese Dienstklassen werden durch insgesamt 5 verschiedene AAL-Protokolltypen abgedeckt, die in der CCITT-Empfehlung I.363 ausführlich beschrieben sind [44]. Jeder dieser Protokolltypen ist wiederum in zwei Subschichten unterteilt:

- In der *SAR-Subschicht* (Segmentation and Reassembly, SAR) findet die Umsetzung von Protokoll-Dateneinheiten der höheren Schichten in ATM-Zellen und umgekehrt statt. Dabei wird ein Teil des Informationsfeldes (je nach Protokolltyp zwischen 1 und 4 Bytes) für AAL-interne Funktionen benötigt, der restliche Teil steht den höheren Schichten zur Verfügung.
- Die *CS-Subschicht* (Convergence Sublayer, CS) enthält verschiedene Anpassungsfunktionen, welche die Funktionalität der ATM-Schicht erweitern und sehr stark von der jeweiligen Dienstklasse bzw. Anwendung abhängen. Beispiele für mögliche Funktionen sind der Ausgleich von Verzögerungsschwankungen, die Behandlung von Zellverlusten, eine Fehlerkorrektur des Informationsfeldes und eine Flußkontrolle.

	Klasse A	Klasse B	Klasse C	Klasse D
Zeitbeziehung	vorhanden		nicht vorhanden	
Bitrate	konstant	variabel		
Verbindungstyp	verbindungsorientiert			verbindungslos

Tabelle 1.1: Dienstklassen der ATM-Anpassungsschicht nach CCITT

1.1.5 Signalisierung

Ähnlich wie im Schmalband-ISDN ist die Verbindungssteuerung im B-ISDN logisch von der Nutzdatenübertragung getrennt. Der Aufbau, die Überwachung und das Auslösen virtueller Verbindungen erfolgt mit Hilfe spezieller Signalisierverbindungen zwischen den Endgeräten und den entsprechenden Instanzen des Netzes. Ein Teilnehmer kann auch über mehrere Signalisier-Instanzen verfügen, die jeweils über eine eigene virtuelle Verbindung mit dem Netz kommunizieren.

Signalisierverbindungen können nach Bedarf eingerichtet werden und sind auf der ATM-Schicht nicht von anderen virtuellen Verbindungen eines Benutzers unterscheidbar. Für den Aufbau und die Überwachung der Signalisierverbindungen wird daher ein zusätzliches sogenanntes „Meta-Signalisierungsprotokoll" benötigt. Die Meta-Signalisierung basiert auf einer vordefinierten virtuellen Verbindung (Metasignalling Virtual Channel), für die an allen Schnittstellen ein einheitlicher VCI/VPI reserviert ist. Diese Verbindung wird im „Broadcast"-Modus betrieben, so daß die zugehörigen Zellen von allen Endgeräten einer Schnittstelle empfangen werden können.

Die Funktionen der Signalisierung sind Teil der bereits in 1.1.4 erwähnten Steuerebene. Die Einzelheiten des Signalisierprotokolls sind noch nicht festgelegt, gegenüber dem im Schmalband-ISDN verwendeten D-Kanal Protokoll müssen jedoch verschiedene Erweiterungen vorgenommen werden. Das Breitband-ISDN stellt neue Anforderungen an die Funktionalität der Signalisierung, beispielsweise müssen auch Punkt-zu-Mehrpunkt-Verbindungen und Multimedia-Dienste unterstützt werden. Das B-ISDN soll außerdem die Grundlage für zukünftige Entwicklungen wie das „Persönliche Kommunikationsnetz" (Personal Communication Network, PCN) bilden, in dem jeder Teilnehmer eine weltweit gültige persönliche Rufnummer erhält. Um die Einführung des B-ISDN nicht zu verzögern, werden die Signalisierprotokolle daher, beginnend mit den wichtigsten Grundfunktionen, stufenweise erweitert.

1.1.6 Verkehrssteuerungsverfahren

Die Flexibilität der Bandbreitenzuteilung in ATM-Netzen ermöglicht eine effiziente Nutzung der zur Verfügung stehenden Übertragungskapazitäten. Im Gegensatz zu durchschaltevermittelnden Netzen können sich bestehende Verbindungen jedoch gegenseitig beeinflussen, so daß zusätzliche Steuer- und Kontrollfunktionen erforderlich sind. Unter dem Oberbegriff „Verkehrssteuerungsverfahren" werden verschiedene Mechanismen und Funktionen zusammengefaßt, deren Hauptaufgabe darin besteht, die Einhaltung bestimmter Netzgüteparameter zu garantierten [12], [109], [180]. Andererseits soll auch eine möglichst hohe Netzauslastung erzielt werden, wodurch die Dienstgüte der Verbindungen unweigerlich verschlechtert wird.

Die beiden Ziele stehen also in einem gewissen Widerspruch zueinander, der Einhaltung der Netzgüte wird momentan jedoch die größere Bedeutung zugemessen.

Die wichtigsten Verkehrssteuerungsverfahren sind die Verbindungsannahme, die Quellflußkontrolle und die Prioritätssteuerung [45]. Weitere Mechanismen sind ebenfalls in der Empfehlung I.371 beschrieben, deren Verwendung ist dem Netzbetreiber jedoch freigestellt.

1.1.6.1 Verbindungsannahme

Die Verbindungsannahme (Connection Admission Control, CAC) ist eine Funktion der Verbindungssteuerung, die sowohl während des Verbindungsaufbaus als auch bei einer Änderung der Verkehrsparameter benötigt wird. Das Netz muß unter Berücksichtigung der bereits bestehenden Verbindungen darüber entscheiden, ob eine neue Verbindung unter Einhaltung der geforderten Dienstgüte aufgebaut werden kann [45]. Diese Funktion bezieht sich sowohl auf VC- als auch auf VP-Verbindungen und muß auf jedem Verbindungsabschnitt ausgeführt werden.

Als Grundlage der Entscheidung werden sowohl die Verkehrsparameter als auch die geforderten Dienstgüteparameter der Verbindung benötigt und müssen daher beim Verbindungsaufbau angegeben werden. Die Verkehrsparameter beschreiben die charakteristischen Verkehrseigenschaften einer Verbindung, wie zum Beispiel die Spitzenbitrate oder die mittlere Bitrate. Die Bandbreitenzuteilung ist umso effizienter, je genauer die Verkehrseigenschaften einer Verbindung spezifiziert werden.

Dabei treten allerdings verschiedene Probleme auf: Selbst einfache Verkehrsparameter wie die mittlere Bitrate sind zu Beginn einer Verbindung oft noch unbekannt, da sie von der Aktivität des Benutzers abhängen. Dies trifft insbesondere auf interaktive Dienste zu, bei denen der Datenstrom oft durch längere (Denk-)Pausen unterbrochen ist. Darüber hinaus müssen die Verkehrsparameter auch kontrollierbar sein, da die Einhaltung der angegebenen Parameter vom Netz überwacht werden muß (siehe 1.1.6.2). Diese Schwierigkeiten haben dazu geführt, daß zunächst nur die Spitzenbitrate als einziger Verkehrsparameter standardisiert wurde [45].

Die Bandbreitenzuteilung auf Grund der Spitzenbitrate ist vergleichsweise einfach, sofern davon ausgegangen wird, daß eine exakte Überwachung der Spitzenbitrate durch die Quellflußkontrolle möglich ist. In diesem Fall muß lediglich darauf geachtet werden, daß die Summe der Spitzenbitraten aller Verbindungen unter der Auslastungsgrenze der ATM-Schicht bleibt. Dabei ist jedoch zu beachten, daß die Auslastungsgrenze einer Leitung wegen der asynchronen Zeitmultiplextechnik deutlich unter 100% liegt.

Die tatsächliche Auslastung des Netzes auf Zellebene wird allerdings noch wesentlich geringer sein, da viele Verbindungen nicht ständig die volle Spitzenbitrate beanspruchen. Eine

bessere Netzauslastung ist nur durch Ausnutzen der statistischen Schwankungen des momentanen Bandbreitebedarfs aller Verbindungen möglich (statistisches Multiplexen). Dies setzt jedoch die Kenntnis weiterer Verkehrsparameter zusätzlich zur Spitzenbitrate voraus, um das statistische Verhalten der Verkehrsquellen abschätzen zu können.

1.1.6.2 Quellflußkontrolle

Die Quellflußkontrolle (Usage Parameter Control, UPC) hat die Aufgabe, die Einhaltung aller angegebenen Verkehrsparameter zu überwachen und bei Überschreitungen geeignete Maßnahmen zu ergreifen [45]. Mögliche Maßnahmen sind beispielsweise das sofortige Verwerfen von Zellen, das Markieren von Zellen durch Setzen des CLP-Bits oder ein Abbau der betreffenden Verbindung. Eine zusätzliche Funktion der Quellflußkontrolle ist die Überprüfung der Gültigkeit des VCI/VPI-Feldes aller Zellen.

Alle virtuellen Verbindungen an der Benutzer-Netz-Schnittstelle einschließlich der Signalisierverbindungen müssen der Quellflußkontrolle unterworfen werden. Um eine Beeinflussung der Verkehrsparameter durch das Multiplexen mit anderen Verbindungen zu vermeiden, sollte die Quellflußkontrolle möglichst nahe am Endgerät erfolgen, aus Sicherheitsgründen muß die Funktion jedoch unter der Kontrolle des Netzbetreibers bleiben. Die Überwachung von VC-Verbindungen erfolgt daher am Endpunkt des ersten VC-Verbindungsabschnitts, der sich normalerweise in der Ortsvermittlungsstelle befindet. Analog dazu müssen VP-Verbindungen am Endpunkt des ersten VP-Verbindungsabschnitts kontrolliert werden. Darüber hinaus ist eine Überwachung auch an den Übergängen zwischen Netzen unterschiedlicher Betreiber oder an anderen Stellen innerhalb des Netzes möglich (Network Parameter Control, NPC).

Da die Quellflußkontrolle für jede Verbindung durchgeführt wird, muß der Mechanismus einfach und leicht zu implementieren sein. Trotzdem sollte eine Verletzung der Verkehrsparameter möglichst schnell und mit hoher Genauigkeit erkannt werden. Andererseits darf der Zellstrom bei Einhaltung der Parameter nicht verändert werden, da die Quellflußkontrolle sonst die Netzgüte beeinflußt. Die Dimensionierung wird zusätzlich durch statistische Verzögerungsschwankungen auf dem Weg der Zellen vom Endgerät bis zur Überwachung erschwert. Diese Schwankungen sind auf das Multiplexen mit anderen Verbindungen, aber auch auf teilnehmerseitige Einrichtungen wie lokale Netze und Nebenstellenanlagen zurückzuführen und können die Verkehrsparameter verändern. Aus diesem Grund muß der Teilnehmer die maximale Verzögerungsschwankung spezifizieren, die zwischen Endgerät und Benutzer-Netz-Schnittstelle auftreten kann.

In der Literatur wurde eine Vielzahl verschiedener Mechanismen zur Implementierung der Quellflußkontrolle vorgeschlagen. Eine Zusammenstellung der wichtigsten Verfahren sowie deren Leistungsfähigkeit ist in [175] zu finden. Bislang wurde kein spezielles Verfahren zur Quellflußkontrolle standardisiert, der sogenannte „Leaky Bucket"-Mechanismus wird in [45]

jedoch als Beispiel für die Überwachung der Spitzenbitrate angegeben. Dieser Mechanismus besteht im wesentlichen aus einem Zähler, der durch jede Zellankunft inkrementiert und entsprechend der Spitzenzellrate dekrementiert wird. Der Zähler besitzt eine obere Grenze, bei deren Erreichen alle ankommenden Zellen entweder verworfen oder markiert werden. Durch geeignete Wahl der Zählergrenze kann das Verfahren somit gewisse Verzögerungsschwankungen tolerieren.

1.1.6.3 Prioritäten

Das Prioritätsbit (CLP-Bit) im Kopf der Zellen ermöglicht der ATM-Schicht eine direkte Unterstützung von zwei Übermittlungsdiensten mit unterschiedlichen Zellverlustwahrscheinlichkeiten. Auf diese Weise kann eine höhere Netzauslastung erzielt werden als bei einer einheitlichen Behandlung aller Zellen, die eine Auslegung des Netzes und der Verkehrssteuerungsverfahren entsprechend den Anforderungen des anspruchsvollsten Dienstes voraussetzt.

Die Zuweisung der Priorität kann sowohl auf Verbindungsebene als auch für jede einzelne Zelle erfolgen. Beispielsweise verwenden Signalisierverbindungen aus Gründen der Zuverlässigkeit ausschließlich Zellen mit hoher Priorität. Bestimmte Videokodierverfahren dagegen basieren auf einer Zerlegung des Bildinhalts in mehrere Komponenten mit unterschiedlicher Bedeutung für die Bildqualität (Layered Coding), so daß hier eine Übertragung der weniger wichtigen Komponenten mit niedriger Priorität denkbar ist.

Da für beide Verkehrsklassen eine bestimmte Netzgüte garantiert wird, müssen die Verkehrsparameter beider Klassen bei der Verbindungsannahme und der Quellflußkontrolle berücksichtigt werden. In der Literatur wird zwar auch die Einbindung von verbindungslosen Datendiensten als Verkehrsklasse mit niedriger Priorität ohne garantierte Netzgüte vorgeschlagen, dieses Konzept wird gegenwärtig jedoch nicht vom CCITT unterstützt [45].

Die Vermittlungsknoten des Breitband-ISDN müssen über spezielle Mechanismen des Pufferzugangs verfügen, um die unterschiedlichen Verlustwahrscheinlichkeiten der beiden Verkehrsklassen zu ermöglichen. Für eine detaillierte Beschreibung verschiedener Möglichkeiten zur Implementierung von Verlustprioritäten wird auf Abschnitt 2.3.2 verwiesen.

1.2 Übersicht über die Arbeit

Die vorliegende Arbeit beschäftigt sich mit der Architektur sowie der Leistungsfähigkeit von ATM-Koppelanordnungen. In Kapitel 2 werden zunächst die grundlegenden Funktionen sowie die Anforderungen an eine ATM-Koppelanordnung beschrieben. Den Hauptteil des Kapitels bildet eine systematische Klassifizierung der einzelnen Architekturen, wobei insbesondere

diejenigen Eigenschaften berücksichtigt werden, welche die Verkehrsleistung der Anordnung beeinflussen. Die Klassifizierung erfolgt nach Kriterien wie der Anordnung und Organisation der Pufferspeicher, dem Steuerungsverfahren sowie der Topologie des Koppelnetzes.

Kapitel 3 enthält eine Zusammenfassung der wichtigsten verkehrstheoretischen Grundlagen, die im Zusammenhang mit der Leistungsuntersuchung von ATM-Koppelnetzen von Bedeutung sind. Ausgehend von einer hierarchischen Modellierung des Verkehrs werden verschiedene Quellmodelle zur Beschreibung der Verkehrsströme in ATM-Netzen vorgestellt. Wegen des verbindungsorientierten Betriebs von ATM-Netzen muß auch der Modellierung des Verkehrsflusses besondere Beachtung geschenkt werden. Um quantitative Aussagen über die Verkehrsleistung eines Systems zu gewinnen, werden entsprechende Methoden zur Leistungsuntersuchung benötigt, wobei in dieser Arbeit sowohl simulative als auch analytische Methoden Verwendung finden.

Die Verkehrsleistung von verschiedenen einstufigen Koppelanordnungen wird in Kapitel 4 untersucht, wobei insbesondere die unterschiedlichen Möglichkeiten zur Pufferorganisation verglichen werden. Für einige der betrachteten Varianten werden geeignete Analyseverfahren vorgestellt, in vielen Fällen ist jedoch nur eine simulative Untersuchung möglich. Den Abschluß dieses Kapitels bildet ein Vergleich sowie eine Bewertung der einzelnen Architekturen.

Kapitel 5 befaßt sich mit der Analyse von mehrstufigen Koppelnetzen, wobei insbesondere die Eigenschaften des internen Verkehrs auf Zellebene sowie dessen Modellierung im Vordergrund stehen. Um eine analytische Behandlung zu ermöglichen, muß der interne Verkehr durch ein geeignetes Ersatzmodell approximiert werden, welches eine hinreichend genaue Nachbildung der wesentlichen Eigenschaften erlaubt. Am Beispiel von Koppelelementen mit Ausgangspuffern wird ein Analyseverfahren vorgestellt, welches eine approximative Bestimmung der Verlustwahrscheinlichkeit sowie der Durchlaufzeitverteilung ermöglicht.

Mit Hilfe einer verfeinerten Modellierung der Verkehrsquellen soll in Kapitel 6 der Zusammenhang zwischen den Eigenschaften des Quellverkehrs und dem Verhalten der Netzelemente untersucht werden. Zunächst wird der prinzipielle Einfluß verschiedener Parameter auf das Verhalten von Koppelelementen mit Ausgangspuffern gezeigt. Es folgt ein Vergleich von mehreren einstufigen Anordnungen, um die Sensitivität der einzelnen Varianten gegenüber der Quellcharakteristik zu bestimmen. Weiterhin werden die Auswirkungen einer Vernachlässigung des verbindungsorientierten Betriebs auf die Ergebnisse der Leistungsuntersuchung diskutiert.

Die Verkehrsleistung von Koppelnetzen mit verbindungsorientiertem Betrieb hängt außer von den Verkehrsparametern auf Zell- und Burst-Ebene ganz wesentlich von Eigenschaften der Verbindungsebene (wie der Blockierwahrscheinlichkeit für virtuelle Verbindungen) ab. In Kapitel 7 werden daher die Eigenschaften verschiedener Koppelnetzstrukturen auf Verbindungsebene untersucht. Am Beispiel des Clos-Netzwerks wird zunächst die Realisierung einer

intern blockierungsfreien Koppelanordnung beschrieben. Der zweite Teil des Kapitels befaßt sich mit der internen Wegesuche, die ebenfalls einen großen Einfluß auf die Blockierwahrscheinlichkeit hat. Anhand einer ausgewählten Koppelnetztopologie wird die Leistungsfähigkeit verschiedener Wegesuchverfahren verglichen. Dabei wird insbesondere der Fall eines inhomogenen Verkehrsgemischs betrachtet, welches aus Verbindungen mit unterschiedlichen Bandbreiteanforderungen besteht.

Kapitel 2

Architektur von ATM-Koppelnetzen

2.1 Einführung

Die Koppelanordnung ist derjenige Bestandteil einer ATM-Vermittlungsstelle, der für die eigentliche Vermittlung der Zellen verantwortlich ist. Abhängig vom jeweiligen Anwendungsfall umfaßt eine Vermittlungsstelle jedoch eine Vielzahl weiterer Funktionsblöcke. Bild 2.1 zeigt eine Gliederung der verschiedenen Funktionen, die sich am Protokoll-Referenzmodell für das Breitband-ISDN orientiert.

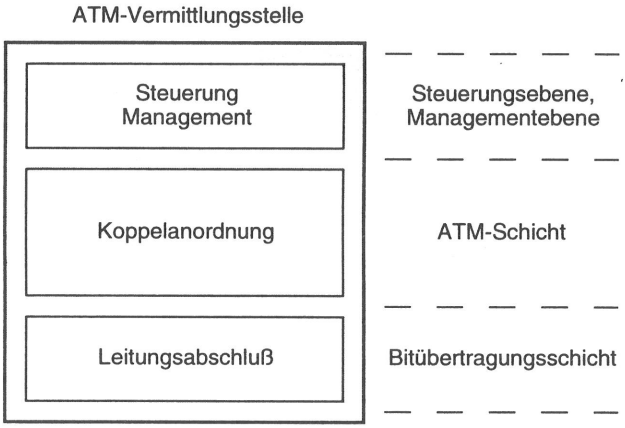


Bild 2.1: Funktionsblöcke einer ATM-Vermittlungsstelle

Der Leitungsabschluß bildet die Schnittstelle zwischen den externen Verbindungsleitungen (oder den Teilnehmeranschlußleitungen) und der eigentlichen Koppelanordnung. Es handelt sich dabei im wesentlichen um übertragungstechnische Funktionen, die der Bitübertra-

gungsschicht zuzurechnen sind. Obwohl der Leitungsabschluß oft genauso aufwendig wie die Koppelanordnung selbst ist, wird im folgenden nicht näher darauf eingegangen.

Für die Abwicklung des Verbindungsauf- und abbaus sowie für verschiedene übergeordnete Steuerfunktionen, Aufgaben des Netzmanagements und der Wartung wird in jeder Vermittlungsstelle auch eine Steuerung benötigt. Deren Funktionen sind der Steuerungs- bzw. Managementebene des Protokoll-Referenzmodells zugeordnet. Verbindungsbezogene Aufgaben der Steuerung sind beispielsweise:

- *Verkehrslenkung (Routing)*: Auf Grund der Adresse (Rufnummer) des gerufenen Teilnehmers wird die Folge der Vermittlungsstellen bestimmt, über die eine Verbindung geführt wird.
- *Interne Wegesuche*: Bei vielen Koppelnetztypen kann ein Ausgang über mehrere interne Wege erreicht werden, die sich durch ihre Zieladresse unterscheiden. In diesem Fall muß bereits beim Verbindungsaufbau einer der Wege ausgewählt werden, wozu ein interner Wegesuchalgorithmus benötigt wird.
- *VCI/VPI-Verwaltung*: Während des Verbindungsaufbaus muß ein freier VCI/VPI gesucht und der Verbindung zugeordnet werden.
- *Verbindungsannahme*: Unter Berücksichtigung der Verbindungsparameter muß geprüft werden, ob die Verbindung ohne Verletzung der Dienstgüte-Kriterien des Netzes angenommen werden kann.
- *Einstellen der Koppelanordnung*: Vor dem Austausch von Nutzdaten benötigt die Koppelanordnung bestimmte Steuerinformationen, wie zum Beispiel die interne Adresse des Ausgangs, zu dem die Zellen der Verbindung vermittelt werden sollen.

Die Realisierung der Steuerungsfunktionen erfolgt üblicherweise in Form von sehr umfangreichen Software-Systemen. Die Architektur der Steuerung sowie deren Umsetzung in ein entsprechendes Rechnerkonzept ist für diese Arbeit jedoch von untergeordneter Bedeutung. Außerdem werden nur solche Steuerungsaufgaben behandelt, die sich unmittelbar auf die Koppelanordnung beziehen. Dazu gehören insbesondere verschiedene Verfahren zur internen Wegesuche und -auswahl sowie das Einstellen der Koppelanordnung.

Die Aufgaben der Koppelanordnung beziehen sich ausschließlich auf die ATM-Schicht. Hier erfolgt die räumliche Vermittlung der Zellen und die Übersetzung des VCI/VPI-Feldes im Zellkopf. Die Koppelanordnung umfaßt im allgemeinen neben dem eigentlichen Koppelnetz auch verschiedene Anpassungsfunktionen, die in Bild 2.2 unter dem Begriff „Protokollabschluß“ zusammengefaßt sind. Der Protokollabschluß kann die folgenden Funktionen enthalten:

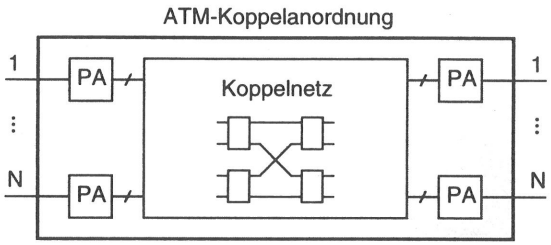


Bild 2.2: Aufteilung der Koppelanordnung in Protokollabschluß (PA) und Koppelnetz

- *Umsetzung in das interne Datenformat*: Viele Koppelanordnungen verwenden ein internes Datenformat, das sich von der genormten Zellstruktur meist durch zusätzliche Felder unterscheidet. Den Zellen werden interne Steuerinformationen wie Zieladresse und Zeitstempel hinzugefügt, die in abgehender Richtung wieder entfernt werden müssen. Teilweise erfolgt auch eine Zerlegung der Zellen in kleinere Segmente, die am Ausgang wieder zusammengefaßt werden müssen [94], [172].
- *Anpassung der Taktfrequenz*: Die interne Taktfrequenz des Koppelnetzes kann sowohl höher als auch niedriger als die externe Datenrate sein. Eine höhere Taktfrequenz wird verwendet, um den Overhead der zusätzlichen Steuerinformationen auszugleichen oder um die interne Last zu reduzieren. Technologische Gründe können auch eine Verringerung der Taktrate notwendig machen, beispielsweise um den Anschluß von Leitungen mit 600 Mbit/s an ein Koppelnetz zu ermöglichen, das nur für 150 Mbit/s ausgelegt ist. Eine interne Aufweitung oder Parallelisierung kann ebenfalls zu einer niedrigeren Taktfrequenz führen.
- *Reihenfolgesicherung*: Bei einigen Koppelnetzen kann die Zellreihenfolge auf Grund der internen Arbeitsweise gestört werden, so daß besondere Maßnahmen zur Regeneration der Reihenfolge erforderlich sind. Dies kann mit Hilfe eines Zeitstempels realisiert werden, der zum Ausgleich der unterschiedlichen Durchlaufzeiten verwendet wird (siehe [94]).

Die Vermittlung einer ATM-Zelle umfaßt sowohl den räumlichen Transport zum entsprechenden Ausgang als auch die Umsetzung des VCI/VPI-Feldes in einen auf dem abgehenden Verbindungsabschnitt gültigen Wert. Diese Umsetzung geschieht mit Hilfe einer Übersetzungstabelle, welche die Zuordnung zwischen den gültigen VCI/VPI-Feldern enthält (siehe Bild 2.3). Gleichzeitig sind meist auch interne Steuerinformationen wie die Zieladresse der Zellen enthalten. Diese Daten werden von der Steuerung bei jedem Verbindungsaufbau eingetragen und nach dem Auslösen einer Verbindung wieder gelöscht.

Die Tabelle kann beispielsweise als Speicher implementiert werden, der direkt durch das VCI/VPI-Feld adressiert wird. Der Speicherbedarf einer derartigen Lösung ist jedoch enorm,

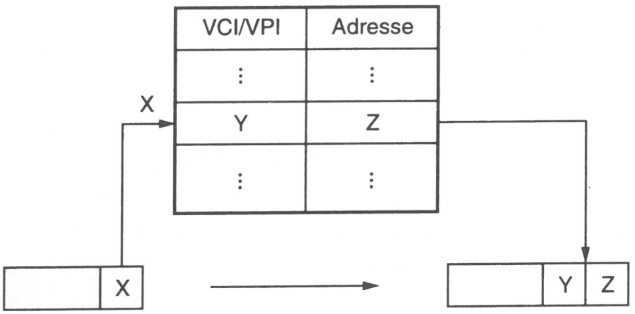


Bild 2.3: Übersetzungstabelle für VCI und VPI

da die Größe des VCI/VPI-Feldes 24 Bit (UNI) bzw. 28 Bit (NNI) beträgt. Um alle möglichen Werte zu erfassen, würden daher für jede Leitung 16 Millionen (UNI) bzw. 256 Millionen (NNI) Speicherworte benötigt. Eine Reduzierung des Speicherbedarfs ist nur durch eine Begrenzung der Anzahl gleichzeitig aktiver Verbindungen möglich. Dies kann entweder durch eine feste Auswahl einer bestimmten Untermenge oder durch die Verwendung eines inhaltsadressierbaren Speichers erfolgen. Die erste Lösung ist etwas unflexibel, und die zweite Lösung ist immer noch relativ aufwendig, da inhaltsadressierbare Speicher dieser Größenordnung nur schwer zu realisieren sind.

Der Schwerpunkt dieses Kapitels liegt auf der Architektur des Koppelnetzes, das den Kern jeder Koppelanordnung bildet. Im folgenden Abschnitt werden zunächst die wichtigsten Anforderungen diskutiert, welche die Grundlage für den Entwurf des Koppelnetzes darstellen. Darauf folgt der eigentliche Hauptteil des Kapitels, eine systematische Klassifizierung der einzelnen Architekturen nach verschiedenen Kriterien. In diesem Zusammenhang werden auch die Vor- und Nachteile der unterschiedlichen Lösungsansätze jeweils kurz diskutiert.

2.2 Anforderungen an eine ATM-Koppelanordnung

An eine ATM-Koppelanordnung werden im Hinblick auf Funktionalität, Verkehrsleistung und Betrieb die unterschiedlichsten Anforderungen gestellt.

2.2.1 Funktionalität

Die wichtigste Aufgabe einer ATM-Koppelanordnung ist das Vermitteln von Zellen zu den entsprechenden Ausgängen. Dabei muß beachtet werden, daß die Reihenfolge der Zellen einer virtuellen Verbindung durch das ATM-Netz nicht verändert werden darf [38]. Falls die

Reihenfolge der Zellen durch das Koppelnetz gestört werden kann, müssen Reihenfolgefehler durch entsprechende Maßnahmen zur Reihenfolgesicherung innerhalb der Protokollanpassung behoben werden.

Neben der einfachen Vermittlungsfunktion von Punkt-zu-Punkt-Verbindungen sollte die Koppelanordnung auch Punkt-zu-Mehrpunkt-Verbindungen unterstützen. Dazu müssen die Zellen der betreffenden Verbindungen kopiert und zu einer definierten Untermenge der Ausgänge vermittelt werden. Diese auch als Multicast-Betrieb bezeichnete Funktion ist vor allem für Konferenzdienste sowie bei der gezielten Verteilung von Audio- und Videosignalen von großer Bedeutung.

Aus funktionaler Sicht ist die erforderliche Geschwindigkeit eine der härtesten Randbedingungen. Bei einer Übertragungsrate von 622 Mbit/s treffen an jedem Eingang des Koppelnetzes bis zu 1.4 Millionen Zellen pro Sekunde ein. Diese müssen in Realzeit bearbeitet und mit möglichst geringer Verzögerung zum entsprechenden Zielausgang vermittelt werden. Aus Geschwindigkeitsgründen werden daher alle Funktionen, die unmittelbar mit dem Transport der Zellen zusammenhängen, in Hardware realisiert. Traditionelle Paketvermittlungssysteme, die auf einer teilweisen oder vollständigen Software-Lösung basieren, sind für die ATM-Vermittlung völlig ungeeignet. Ein ATM-Koppelnetz weist üblicherweise auch eine verteilte Architektur mit einem sehr hohen Grad an Parallelverarbeitung auf. Nur so können Koppelanordnungen in einer Größenordnung von mehreren Tausend Eingängen realisiert werden.

2.2.2 Verkehrsleistung

Die Verkehrsgüte von ATM-Koppelnetzen bezieht sich sowohl auf den Transport der Zellen durch das Koppelnetz als auch auf den Verbindungsaufbau. Die wesentlichen Parameter der Verkehrsgüte sind die Zellverlustwahrscheinlichkeit, die Verteilung der Durchlaufzeit von Zellen durch die Koppelanordnung und die Wahrscheinlichkeit, daß eine Verbindung aus Mangel an internen Ressourcen abgewiesen werden muß. Eine vorläufige Definition verschiedener Verkehrsgüteparameter der ATM-Schicht ist auch in einem Entwurf der CCITT-Empfehlung I.35B enthalten [35]. Bei vorgegebener Netzbelastung dürfen diese Parameter bestimmte Grenzwerte nicht überschreiten.

Die Zellverlustwahrscheinlichkeit ist eines der wichtigsten Kriterien für die Verkehrsgüte. Sie ist definiert als Verhältnis aus der Anzahl verlorener Zellen zu der Anzahl der ingesamt übertragenen Zellen und kann entweder auf jede einzelne Verbindung bezogen oder über alle Verbindungen einer Leitung gemittelt werden. Die Wahrscheinlichkeit für das Auftreten von Zellverlusten innerhalb des Koppelnetzes sollte nicht wesentlich größer als die Wahrscheinlichkeit sein, daß eine Zelle durch Bitfehler während der Übertragung gestört wird. Ein genauer

Grenzwert ist zwar noch nicht festgelegt [35], es wird jedoch über Verlustwahrscheinlichkeiten zwischen 10^{-8} und 10^{-12} diskutiert. Diese Grenzwerte müssen durch eine entsprechende Dimensionierung und durch wirksame Verkehrssteuerungsmechanismen sichergestellt werden. Dazu gehört auch die Einführung von zwei Verlustprioritäten, die durch ein Bit innerhalb des Zellkopfes unterschieden werden [42]. Durch entsprechende Verwaltung der Pufferspeicher können den Prioritätsklassen unterschiedliche Verlustwahrscheinlichkeiten zugeordnet werden. Eine ausführliche Beschreibung verschiedener Möglichkeiten zur Realisierung von Verlustprioritäten erfolgt in Abschnitt 2.3.2.

Anders als bei der Durchschaltevermittlung ist die Durchlaufzeit von Zellen durch das Koppelnetz hier zufälligen Schwankungen unterworfen. Diese auch „Jitter" genannten Schwankungen bereiten sowohl bei der Übertragung von synchronen Signalen als auch bei der Überwachung der Spitzenbitrate durch die Quellflußkontrolle gewisse Schwierigkeiten. Jeder Netzknoten muß daher eine obere Grenze für die Schwankungen der Durchlaufzeit garantieren.

Sowohl Zellverluste als auch Durchlaufzeitschwankungen hängen jedoch ganz entscheidend von den Eigenschaften des angebotenen Verkehrs ab. Obwohl die Verkehrsparameter jeder Verbindung durch entsprechende Mechanismen der Quellflußkontrolle überwacht werden, sind dennoch gewisse Toleranzen der Verkehrsparameter möglich, so daß das Koppelnetz möglichst robust gegenüber der Verkehrscharakteristik der Verbindungen sein sollte. Dazu gehört auch die Robustheit gegenüber einer schiefen Last durch ungleiche Belastung von Eingangs- oder Ausgangsleitungen. Außerdem sollte die Koppelanordnung weder im Normalbetrieb noch unter Überlast eine systembedingte Unfairneß aufweisen, die beispielsweise zu einer schlechteren Dienstgüte einzelner Verbindungen führt, obwohl die Verkehrsgüte im Mittel über alle Verbindungen eingehalten wird.

2.2.3 Herstellung und Betrieb

Aus der Sicht des Herstellers bzw. Betreibers einer Vermittlungsstelle ist unter anderem der Implementierungsaufwand des Koppelnetzes (und damit dessen Kosten) von Bedeutung. Um die Anzahl der verwendeten Baugruppen und folglich auch die Herstellungskosten zu senken, ist eine möglichst gute Integrierbarkeit der einzelnen Komponenten anzustreben. Eine hohe Integrationsdichte wirkt sich auch auf die Zuverlässigkeit des Systems sehr günstig aus, da weniger fehleranfällige Komponenten vorhanden sind und vor allem auch weniger Verbindungen zwischen den einzelnen Baugruppen benötigt werden. Trotzdem muß die Fehlertoleranz auch bei der Architektur eines Vermittlungssystems berücksichtigt werden, um die Anforderungen der Netzbetreiber an die Verfügbarkeit des Systems zu erfüllen.

Grundsätzlich sollte jedes Vermittlungssystem möglichst unempfindlich gegen den Ausfall einzelner Komponenten sein. Insbesondere solche Baugruppen, die von zentraler Bedeutung

für die Funktion des Systems sind, werden daher meist durch entsprechende Redundanz gegen Ausfälle geschützt. Darüber hinaus werden einerseits Mechanismen zur schnellen und zuverlässigen Erkennung sowie Lokalisierung von Fehlern benötigt, und andererseits müssen die Auswirkungen eines erkannten Fehlers durch kurzfristige Maßnahmen, beispielsweise durch Umlenken des Verkehrs, so gering wie möglich gehalten werden.

Flexibilität in Bezug auf Einsatz und Erweiterbarkeit des Systems ist eine weitere Anforderung. Das Systemkonzept sollte für so unterschiedliche Anwendungen wie Konzentratoren, elektronische Schaltverteiler (Crossconnects), Ortsvermittlungsstellen und Fernvermittlungsstellen geeignet sein. Für das Koppelnetz bedeutet dies vor allem, daß es durch entsprechende Dimensionierung an die verschiedenen Anwendungsfälle anpaßbar sein sollte.

Zu einem flexiblen Systemkonzept gehört auch eine gute Erweiterbarkeit. Der Ausbau des Koppelnetzes muß im laufenden Betrieb und mit möglichst geringen Auswirkungen auf bestehende Verbindungen möglich sein. Hierfür sind Umkehrgruppierungen (siehe 2.3.4.2.3) besonders geeignet, die nur durch Hinzufügen neuer Baugruppen und ohne Änderung der bestehenden Konfiguration erweitert werden können. Dagegen sollten solche Topologien vermieden werden, die im Falle einer Erweiterung teilweise oder vollständig geändert werden müssen.

2.3 Klassifizierung der ATM-Koppelnetze

Eine systematische Klassifizierung von ATM-Koppelnetzen kann nach vielen unterschiedlichen Gesichtspunkten erfolgen. In der Literatur häufig verwendete Kriterien sind Topologie, Steuerungsverfahren und Plazierung der Pufferspeicher. In dieser Arbeit sollen die wesentlichen Merkmale, die eine Koppelanordnung charakterisieren, als Grundlage der Klassifizierung dienen. Der Schwerpunkt liegt dabei vor allem auf solchen Eigenschaften, die die Verkehrsleistung beeinflussen.

2.3.1 Anordnung und Organisation der Puffer

Wegen des statistischen Charakters der Zellankünfte sind Kollisionen durch gleichzeitig eintreffende Zellen mit gleichem Ziel in der asynchronen Zeitmultiplextechnik unvermeidlich. Diese Kollisionen müssen durch kurzzeitiges Zwischenspeichern einzelner Zellen aufgelöst werden. Anordnung und Organisation der Pufferspeicher sind eines der wichtigsten Merkmale eines ATM-Koppelnetzes, da dessen Verkehrsleistung ganz wesentlich von der Art der Kollisionsauflösung abhängt. Je nach Art der Koppelelemente können zwei große Klassen von Koppelnetzen unterschieden werden:

- *Ungepufferte Koppelnetze* bestehen aus Koppelelementen, die keine internen Zellpuffer enthalten. Letztere befinden sich nur an den Eingängen und/oder Ausgängen der Anordnung. Koppelelemente ohne interne Puffer haben den Vorteil einer sehr geringen Durchlaufzeit, da die Zellen nur wenige Bittakte verzögert werden müssen, um die Steuerinformation auszuwerten und das Koppelelement entsprechend einzustellen. Auch die Komplexität solcher Koppelelemente ist sehr gering, für die Realisierung werden nur wenige Gatterfunktionen benötigt [158]. Um einen möglichst hohen Durchsatz zu erzielen, sollte ein ungepuffertes Koppelnetz allerdings intern blockierungsfrei sein, wie beispielsweise das in Abschnitt 2.3.4.2.2 beschriebene Batcher-Banyan Netzwerk. Wegen der Abnehmerblockierung werden jedoch selbst bei einem intern blockierungsfreien Koppelnetz externe Pufferspeicher benötigt.

- In einem *gepufferten Koppelnetz* werden interne Kollisionen durch Zwischenspeichern von Zellen in den Koppelelementen gelöst, das Koppelnetz braucht also nicht intern blockierungsfrei zu sein. In jedem Koppelelement werden die Zellen vor dem erneuten Aussenden meist vollständig gespeichert. Prinzipiell besteht jedoch auch die Möglichkeit, Zellen nur beim tatsächlichen Auftreten von Kollisionen zu speichern, um deren Durchlaufzeit zu verringern. Diese auch „Virtual Cut Through" genannte Methode [116] wurde in [32] im Zusammenhang mit Koppelnetzen untersucht.

  Wegen der internen Speicher sind gepufferte Koppelelemente relativ komplex. Da die für die Puffer benötigten Speicherelemente einen großen Teil der Chipfläche beanspruchen [194], ist die Puffergröße einer der begrenzenden Faktoren bei einer Ein-Chip-Realisierung des Koppelelementes. Falls eine derartige Realisierung angestrebt wird, muß daher die Puffergröße so klein wie möglich gehalten werden. Mögliche Maßnahmen sind eine geeignete Pufferorganisation (beispielsweise als Zentralspeicher, siehe 2.3.1.2), eine Verringerung der internen Last [33] und eine Flußsteuerung zwischen benachbarten Koppelelementen.

  Durch einen Rückstaumechanismus können interne Zellverluste selbst bei kleiner Puffergröße vollständig vermieden werden. Falls ein Puffer keine freien Plätze mehr zur Verfügung hat, wird dies den betroffenen Koppelelementen der vorherigen Stufe sofort mitgeteilt, so daß keine Zellen mehr gesendet werden. Dies kann entweder durch eine zusätzliche Signalleitung in Rückrichtung oder, bei bestimmten Topologien, durch Markieren der in Gegenrichtung übertragenen Zellen erfolgen. Eine derartige Flußsteuerung hat allerdings den Nachteil, daß sich der Rückstau über mehrere Stufen hinweg bis zum Eingang des Koppelnetzes fortpflanzen kann, wobei auch völlig unbeteiligte Verbindungen beeinträchtigt werden. Der Einsatz eines derartigen Verfahrens erfordert daher eine sorgfältige Abwägung zwischen der zu erwartenden Puffereinsparung und der gleichzeitigen Verschlechterung der Verkehrseigenschaften.

Die Pufferspeicher können unter anderem an den Eingängen oder Ausgängen eines Koppelelementes angeordnet werden. Die wesentlichen Eigenschaften dieser Alternativen wurden bereits in [96] und [111] untersucht und verglichen. Eine Beschreibung der wichtigsten Möglichkeiten zur Plazierung und Organisation der Puffer findet sich auch in verschiedenen Übersichtsartikeln zum Thema Koppelnetze für ATM [1], [50], [155], [174], [202], [207], [218]. Das folgende Klassifikationsschema soll darauf aufbauend eine Einordnung aller bekannten ATM-Koppelnetztypen ermöglichen. Es ist daher sowohl auf die internen Puffer in einem Koppelelement als auch auf die externen Speicher von ungepufferten Koppelnetzen anwendbar. Der Einfachheit halber wird zwar meist nur auf gepufferte Koppelelemente Bezug genommen, dies soll jedoch keine Einschränkung der Allgemeinheit darstellen.

Die einzelnen Varianten werden im folgenden zusammen mit ihren besonderen Eigenschaften sowie Vor- und Nachteilen ausführlich beschrieben. Außer auf die Plazierung wird dabei auch auf die interne Organisation der Puffer sowie auf Auswahlverfahren zur Kollisionsauflösung eingegangen. Soweit möglich werden zu jeder Variante einige Beispiele von bereits realisierten Systemen angegeben.

2.3.1.1 Ausgangspuffer

2.3.1.1.1 Prinzip

Das Konzept der Ausgangspuffer beruht darauf, daß die ankommenden Zellen auf Grund ihrer Zielinformation unmittelbar dem entsprechenden Ausgang zugeleitet werden. Da gleichzeitig bis zu N Zellen für einen Ausgang bestimmt sein können, verfügt jeder Ausgang über einen eigenen Pufferspeicher zur Kollisionsauflösung, siehe Bild 2.4. Koppelelemente mit Ausgangspuffer sind also dadurch gekennzeichnet, daß die räumliche Vermittlungsfunktion *vor* dem Zwischenspeichern der Zellen ausgeführt wird.

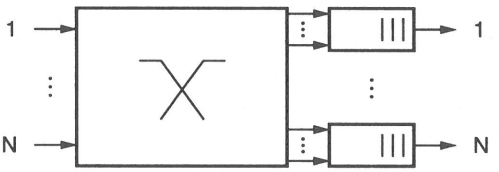


Bild 2.4: Schematische Darstellung eines Koppelelementes mit Ausgangspuffer

Ein Koppelelement mit Ausgangspuffern weist nahezu ideale Verkehrseigenschaften auf, da nur die unvermeidbare Abnehmerblockierung auftritt [96],[111]. Jeder Ausgangspuffer verhält sich wie ein idealer statistischer Multiplexer, der unabhängig von allen anderen Ausgängen arbeitet. Bei unbegrenzter Puffergröße kann daher ein maximaler Durchsatz von 100% erreicht

werden [96], [111]. Außerdem erhält man die kleinstmöglichen Durchlaufzeitschwankungen, da die Puffer (um die Zellreihenfolge einzuhalten) üblicherweise nach dem FIFO-Prinzip abgearbeitet werden. Die Pufferverwaltung ist daher sehr einfach und benötigt keine Koordination zwischen den einzelnen Puffern.

2.3.1.1.2 Realisierung

Speicherzugang: Um interne Konflikte beim Einschreiben von Zellen in einen Ausgangspuffer zu vermeiden, muß jeder Eingang über einen eigenen Zugang zu allen Puffern verfügen. Außerdem muß die Geschwindigkeit des Pufferzugriffs so hoch sein, daß in einer Zellperiode eine Zelle gelesen und bis zu N Zellen eingeschrieben werden können. Die Bandbreite des Speicherzugriffs ist daher eine der wesentlichen technologischen Randbedingungen, welche die Anzahl N der Eingänge begrenzt. Ein 16×16 Koppelelement, dessen Eingänge mit 622 Mbit/s betrieben werden, muß beispielsweise eine Speicherbandbreite von 10 Gbit/s aufweisen (die Begriffe Bandbreite und Bitrate werden in der Literatur oft synonym verwendet).

Um dennoch mit heute üblichen Zugriffszeiten auszukommen, muß der Speicherzugang intern sehr stark parallelisiert werden. Eine derartige Realisierung mit Hilfe des Bit-Slice Prinzips ist in [15], [108] und [196] beschrieben. Eine andere Möglichkeit der Parallelisierung ist die Aufteilung des Puffers in mehrere physikalisch getrennte Speicher, welche zyklisch beschrieben und gelesen werden. Die Speicherverwaltung erfolgt wie bei einem Ringpuffer, so daß sich die Anordnung logisch wie ein FIFO-Speicher verhält [10].

Eine weitere Möglichkeit zur Reduzierung der Zugriffszeit ist eine Begrenzung der gleichzeitig möglichen Speicherzugriffe auf weniger als N [222]. Insbesondere für große N ist die Wahrscheinlichkeit für das gleichzeitige Eintreffen von N Zellen, die alle für den selben Ausgang bestimmt sind, sehr gering. Für den Grenzfall $N \rightarrow \infty$ ist beispielsweise die Wahrscheinlichkeit, daß selbst bei einem Angebot von 100% mehr als 14 Zellen gleichzeitig gespeichert werden müssen, kleiner als 10^{-10} [222]. Aus dieser Begrenzung ergeben sich zwar zusätzliche Zellverluste, deren Auftrittswahrscheinlichkeit kann aber durch entsprechende Dimensionierung der Anzahl der Zugriffe pro Zelltakt wesentlich kleiner als die Verlustwahrscheinlichkeit durch Pufferüberlauf gemacht werden.

Räumliche Vermittlungsfunktion: Die Verbindung zwischen den Eingängen und den Ausgangspuffern kann auf unterschiedliche Arten realisiert werden [2], [10], [64], [81], [108], [196], [222]. Allen Varianten ist jedoch gemeinsam, daß sie aus Geschwindigkeitsgründen keine zentrale Steuerung besitzen, sondern selbständig die Zielinformation der Zellen auswerten. Häufig wird eine interne Bus- oder Ringstruktur verwendet, mit deren Hilfe die ankommenden Zellen zunächst an alle Ausgänge verteilt werden [15], [196], [81] (siehe Bild 2.5). Jeder Ausgang besitzt ein Adreßfilter, das nur die für den betreffenden Ausgang bestimmten Zellen

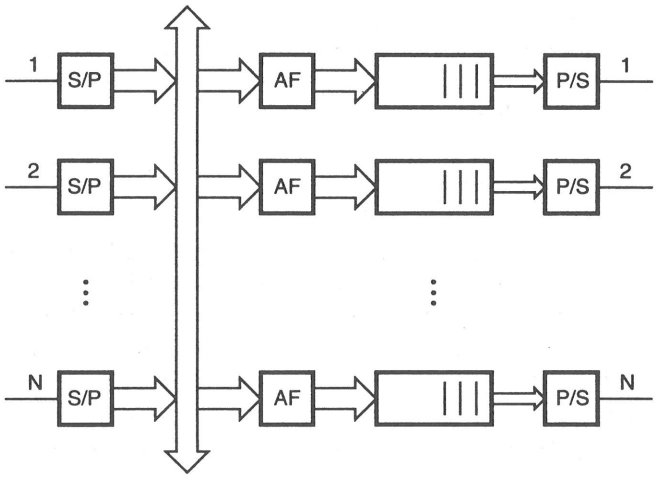


Bild 2.5: Struktur eines Koppelelementes mit Ausgangspuffern auf der Basis eines internen Zeitmultiplex-Bussystems [196]

an den Ausgangspuffer weiterleitet. Auf diese Weise ist auch eine Multicast-Funktion sehr einfach zu realisieren, indem ankommende Zellen von mehreren Adreßfiltern durchgelassen werden.

Die an den Eingängen ankommenden Zellströme werden im synchronen Zeitmultiplex auf das interne Bussystem geschaltet. Dessen Kapazität muß mindestens so groß sein wie die Summe der Bitraten aller Eingangsleitungen, andernfalls treten bereits hier Zellverluste auf. Im einfachsten Fall ist jedem Eingang ein Zeitschlitz zugeordnet, so daß insgesamt N Zeitschlitze erforderlich sind. Ähnlich wie beim Speicherzugang muß folglich auch das interne Bussystem sehr stark parallelisiert werden [65], [64], [108], [196]. Aus funktionaler Sicht verhält sich ein internes Ringsystem sehr ähnlich wie ein Bus, hat aber den Vorteil, daß jeder Knoten am Ring mit nur einem Folgeknoten verbunden ist, so daß geringere Ansprüche an die Treiberleistung gestellt werden als bei einem Bus [81].

Ein wichtiger Aspekt ist die Zugriffsreihenfolge der Eingänge auf den Zeitmultiplex-Bus. Falls, wie in [64] beschrieben, jedem Eingang ein fester Zeitschlitz zugeordnet ist, werden die Eingänge ungleich behandelt. Derjenige Eingang, der auf den ersten Zeitschlitz zugreifen kann, wird immer geringere Zellverluste erleiden als alle anderen Eingänge (siehe auch Ab-

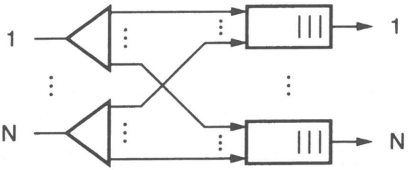


Bild 2.6: Realisierung mit Hilfe einer Demultiplexer-Anordnung

schnitt 4.1). Die Zuordnung der Eingänge zu den Zeitschlitzen auf dem internen Bus sollte daher zyklisch wechseln, um eine faire Behandlung aller Eingangsleitungen sicherzustellen.

Um den Nachteil einer internen Geschwindigkeitserhöhung bei der räumlichen Vermittlung zu vermeiden, kann jeder Eingang auch über einen eigenen Demultiplexer mit den Ausgangspuffern verbunden werden (vgl. Bild 2.6). Der Demultiplexer kann beispielsweise als baumartige Struktur [2] oder als Bus mit Adreßfiltern [222] realisiert werden, wobei hier allerdings N Adreßfilter pro Ausgang benötigt werden. Die in [222] beschriebene Lösung („Knockout Switch") enthält zusätzlich vor jedem Ausgangspuffer einen Konzentrator, um wie oben erwähnt die Anzahl der gleichzeitigen Pufferzugriffe zu verringern.

2.3.1.2 Zentralspeicher

2.3.1.2.1 Prinzip

Das Prinzip der Vermittlung mit Hilfe eines zentralen Speichers ist sowohl in der synchronen als auch in der asynchronen Zeitmultiplextechnik gebräuchlich. Wie in Bild 2.7 gezeigt, werden die eintreffenden Zellen zunächst in einen gemeinsamen Speicher geschrieben. Intern ist jedem Ausgang eine Warteschlange zugeordnet, aus der die Zellen in der Reihenfolge ihres Eintreffens entnommen werden. Das Verkehrsverhalten in Bezug auf die Durchlaufzeit der Zellen und den maximalen Durchsatz kommt daher einem Koppelelement mit Ausgangspuffern sehr nahe.

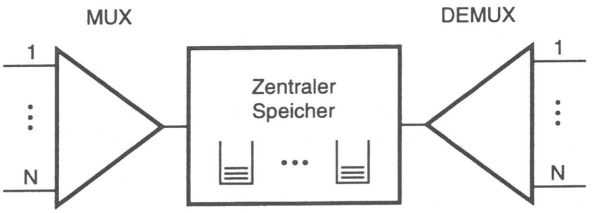


Bild 2.7: Prinzip eines Koppelelementes mit zentralem Speicher

2.3.1.2.2 Speicherverwaltung

Die interne Verwaltung des Speichers basiert auf logischen Ausgangswarteschlangen, in welche die ankommenden Zellen auf Grund ihrer Zieladresse eingereiht werden. Diese Warteschlangen können entweder in Form von eigenen FIFO-Adreßspeichern realisiert werden, welche die Speicheradressen der zu einem Ausgang gehörigen Zellen enthalten [10], [20], [137] oder mit Hilfe von Zeigern, die zusammen mit den Zellen im gemeinsamen Speicher abgelegt sind [14], [82], [124]. Letztere Variante entspricht einer linear verketteten Liste, die von der Steuerlogik verwaltet wird. Neben den Ausgangswarteschlangen wird auch eine Liste der freien Speicherstellen benötigt, aus der die Adressen für das Einschreiben neuer Zellen entnommen werden können. Freigewordene Adressen müssen nach dem Auslesen der Zellen aus dem gemeinsamen Speicher wieder an diese Liste zurückgegeben werden.

Der insgesamt zur Verfügung stehende Speicherplatz kann auf verschiedene Arten verwaltet werden [51], [110], [215]. Eine naheliegende Lösung ist die gemeinsame Nutzung des Speichers durch alle Ausgänge, wobei jeder Ausgang den gesamten Speicher belegen kann. Dieses Prinzip wird in [110] auch als „Complete Sharing" bezeichnet und hat den Vorteil, daß sich die einzelnen Ausgangswarteschlangen gegenseitig aushelfen können. Der Speicher kann auf diese Weise sehr effizient genutzt werden, so daß insgesamt wesentlich weniger Speicherplatz pro Ausgang benötigt wird als bei physikalisch getrennten Ausgangspuffern. Ausgehend von einer Zellverlustwahrscheinlichkeit von 10^{-12} ist der Speicherbedarf bei einem 16×16 Koppelelement um den Faktor 7.5 geringer als bei getrennten Ausgangspuffern [28]. Dies ist insbesondere dann von großer Bedeutung, wenn eine Realisierung des Koppelelementes in Form eines einzigen hochintegrierten Schaltkreises angestrebt wird. In diesem Fall wird ein großer Teil der Chipfläche durch den Speicher belegt, so daß ein möglichst geringer Speicherbedarf anzustreben ist. Aus diesem Grund basieren viele der bisher bekannten Ein-Chip Lösungen auf dem Prinzip des zentralen Speichers [78], [94].

Die gemeinsame Speichernutzung hat jedoch den Nachteil, daß bei einseitiger Überlast eine Ausgangswarteschlange alle anderen Zellen aus dem Speicher verdrängen kann, so daß auch die anderen Ausgänge sehr hohe Verluste erleiden. Kurzzeitige Überlastsituationen können beispielsweise beim statistischen Multiplexen von stark stoßartigem Verkehr auftreten. Um diesen Effekt zu vermeiden, muß die Belegung des Speichers durch einzelne Ausgänge begrenzt werden [51], [110], [184]:

- Die Länge jeder Ausgangswarteschlange kann so reduziert werden, daß ein einzelner Ausgang nur einen Teil des gesamten Speichers belegen darf [51], [110], [137], [184]. Dadurch wird erreicht, daß der restliche Teil des Speichers für die anderen Ausgangswarteschlangen zur Verfügung steht, so daß bei einer einzelnen Überlast die Zellverluste der gering belasteten Ausgänge nur wenig zunehmen. Falls jedoch mehrere Ausgänge in einer Überlastsituation sind, kann der Speicher dennoch blockiert werden.

Auf eine interessante Eigenschaft einer derartigen Begrenzung wird in [137] hingewiesen. Dort wird gezeigt, daß durch eine entsprechende Dimensionierung der Ausgangswarteschlangen auch die Verlustwahrscheinlichkeit bei symmetrischer Last reduziert werden kann.

- Für jeden Ausgang kann ein kleiner Teil des Speichers reserviert werden, den andere Ausgangswarteschlangen nicht belegen können. Der Rest des Speichers wird von allen Ausgangswarteschlangen gemeinsam benutzt [51], [110]. Auf diese Weise wird in jedem Fall verhindert, daß ein oder mehrere Ausgänge den gesamten Speicher blockieren. In einer Überlastsituation steht jedem gering belasteten Ausgang wenigstens der reservierte Teil des Speichers zur Verfügung, so daß dessen Verluste nur in geringem Ausmaß zunehmen.
- Durch eine Kombination der oben genannten Ansätze läßt sich auch eine gerechtere Belegung des gemeinsam benutzten Speicherbereichs erzielen, so daß dieser nur im Fall einer Überlast von mehreren Ausgängen blockiert werden kann [51], [110].
- In [215] wird ein neues Verfahren zur Speicherverwaltung vorgeschlagen, welches unempfindlich gegen kurzzeitige Überlastsituationen ist und trotzdem den Vorteil einer gemeinsamen Speichernutzung besitzt. Es unterscheidet sich vom „Complete Sharing"-Verfahren nur dadurch, daß beim Auftreten von Zellverlusten nicht in jedem Fall die ankommende Zelle verworfen wird, sondern eine Zelle aus der längsten Ausgangswarteschlange. Auf diese Weise wird erreicht, daß kein Ausgang den Speicher überfluten kann, und es ergibt sich selbst bei symmetrischer Last eine geringere Verlustwahrscheinlichkeit als beim „Complete Sharing"-Verfahren (siehe 4.2). Der einzige Nachteil dieses Verfahrens ist der Implementierungsaufwand, der erforderlich ist, um die längste Ausgangswarteschlange zu suchen und aus dieser eine Zelle zu verwerfen.
- Eine gegenseitige Beeinflussung der Ausgangswarteschlangen kann allerdings nur dadurch vermieden werden, daß für jeden Ausgang ein eigener Speicherbereich zur exklusiven Nutzung bereitgestellt wird. Die Anordnung ist dann völlig äquivalent zu einem Koppelelement mit getrennten Ausgangspuffern und benötigt auch genauso viel Speicherplatz. Eine Reduktion des Speicherplatzbedarfs ist nur durch ganz oder teilweise gemeinsame Nutzung des Speichers durch mehrere Ausgänge zu erzielen.

2.3.1.2.3 Realisierung

Bei der Realisierung von Koppelelementen mit zentralem Speicher müssen zwei wichtige Aspekte beachtet werden:

- Die Anforderungen an die Speicherbandbreite sind noch höher als bei einem Ausgangspuffer, da pro Zelltakt $2N$ Speicherzugriffe erfolgen. Auch hier ist deshalb eine starke interne Parallelisierung und eine geschickte Organisation des Speichers notwendig.

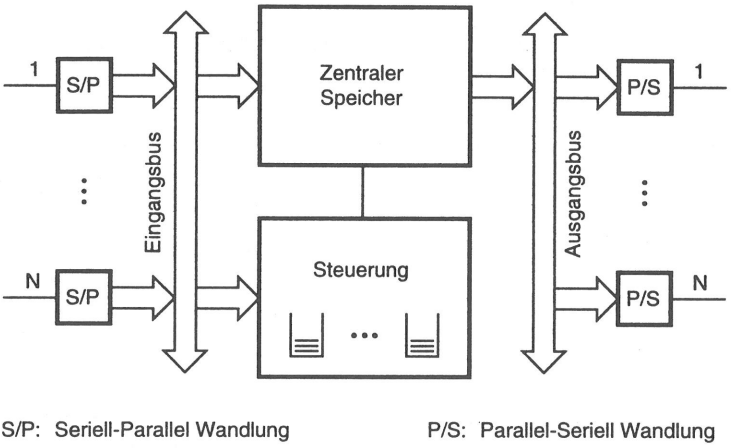


Bild 2.8: Realisierung des Zentralspeichers mit internem Bus

- Es wird eine zentrale Steuerung benötigt, die in einem Zelltakt sowohl N ankommende als auch N abgehende Zellen verarbeiten kann. Den ankommenden Zellen muß jeweils eine freie Speicheradresse zugewiesen werden, und die abgehenden Zellen sind aus den Ausgangswarteschlangen zu entnehmen und an die entsprechenden Ausgänge weiterzuleiten. Häufig werden daher auch die internen Abläufe der Steuerung durch die Anwendung von entsprechenden Pipeline-Konzepten parallelisiert.

Ähnlich wie bei Koppelelementen mit Ausgangspuffern basieren mehrere Realisierungen der Zentralspeicher-Architektur auf einem internen Bussystem [10], [14], [78], [79], [113], [124], [192], [219] (siehe Bild 2.8). Auch hier wird jedem Eingang ein Zeitschlitz auf dem Eingangsbus fest oder variabel zugeteilt, wobei die gleichen Fairneßprobleme wie bei Koppelelementen mit Ausgangspuffern auftreten. Um sicherzustellen, daß kein Eingang bevorzugt behandelt wird, sollte die Zugriffsreihenfolge daher zyklisch wechseln [192]. Die zentrale Steuerung stellt die Speicheradressen bereit, in welche die ankommenden Zellen geschrieben werden, und verwaltet die internen Ausgangswarteschlangen. Auf dem Ausgangsbus ist jedem Ausgang ein fester Zeitschlitz zugeordnet. Die Steuerung koordiniert das Auslesen der Zellen aus dem Speicher derart, daß die einem bestimmten Ausgang zugehörigen Zellen zum richtigen Zeitpunkt auf dem Bus erscheinen.

Der Grad der internen Parallelisierung hängt von der Größe des Koppelelementes und von technologischen Randbedingungen wie der Speicherzykluszeit ab. Bei einer Speicherzykluszeit von 10 ns ist eine Busbreite von mindestens 1/4 Zelle erforderlich, um ein 16×16 Koppelelement bei 622 Mbit/s zu realisieren [10]. In [113] wird sogar jede Zelle als ein 424 Bit breites Speicherwort parallel in den Speicher geschrieben.

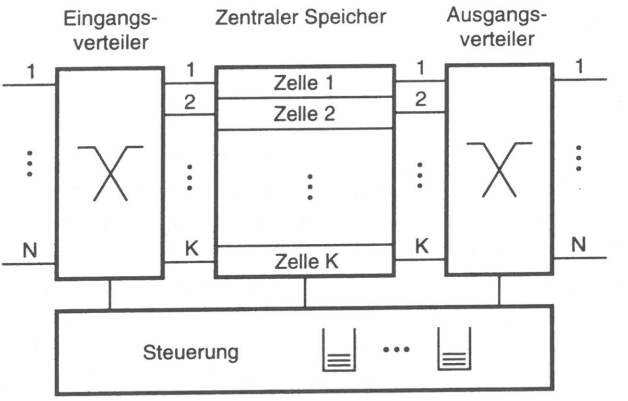


Bild 2.9: Realisierung des Speicherzugangs mit Hilfe eine Schaltmatrix

In [20] und [186] ist ein anderes Realisierungskonzept beschrieben, das durch eine geschickte Organisation des Zentralspeichers ohne interne Geschwindigkeitserhöhung auskommt. Wie in Bild 2.9 dargestellt, ist der Speicher in K unabhängige Zeilen unterteilt, wobei jede Zeile dem Speicherplatz einer Zelle entspricht und als Schieberegister implementiert ist (im allgemeinen ist $K > N$). Die ankommenden Zellen werden, unter Kontrolle der Steuerung, von einer Schaltmatrix auf die freien Speicherplätze (Zeilen) des Speichers verteilt. Da jeder Eingang einen eigenen Zugang zum Speicher besitzt, können bis zu N Zellen parallel und ohne Geschwindigkeitserhöhung in den Speicher eingeschrieben werden. Gleichzeitig mit dem Einschreiben können bis zu N Zellen ausgelesen und über eine weitere Schaltmatrix den entsprechenden Ausgängen zugeleitet werden.

Die Breite des Speicherzugriffs hängt zunächst nur von der Puffergröße K ab, diese steht jedoch indirekt auch mit der Anzahl der Ausgänge N in Zusammenhang. Wegen der Komplexität der Eingangs- und Ausgangsverteiler ist das Konzept vor allem für kleinere Koppelelemente mit sehr hohen Geschwindigkeitsanforderungen geeignet, die nur wenige Speicherplätze benötigen.

Neben den bisher beschriebenen Architekturen existieren weitere Realisierungen, wie beispielsweise das im Rahmen des „Prelude"-Experiments entworfene Koppelelement [206] oder das auf einem Batcher-Banyan Netzwerk basierende „Starlite"-Konzept [98]. Beide Prototypen sind Vorläufer der heutigen Systemarchitekturen und entsprechen daher nicht den gegenwärtigen CCITT-Spezifikationen. Sie werden aus diesem Grund hier nicht genauer erläutert.

2.3.1.3 Eingangspuffer

2.3.1.3.1 Prinzip

In einem Koppelelement mit Eingangspuffern ist jeder Eingangsleitung eine eigene Warteschlange zugeordnet, in der alle ankommenden Zellen gespeichert werden (siehe Abbildung 2.10). Die Steuerung des Koppelelementes wertet die Zielinformation der gespeicherten Zellen aus und koordiniert deren Transport zu den entsprechenden Ausgängen. Im Unterschied zu einem ausgangsgepufferten Koppelelement findet die räumliche Vermittlung also erst *nach* dem Speichern der Zellen statt. Die Eingangspuffer sind mit den Ausgängen des Koppelelementes durch eine intern blockierungsfreie Koppelanordnung verbunden, die im einfachsten Fall aus einer Schaltmatrix besteht [194]. Das Konzept der Eingangspuffer ist jedoch auch auf mehrstufige Anordnungen übertragbar, wobei die Verbindung zwischen Eingangspuffern und Ausgängen durch eine ungepufferte mehrstufige Koppelanordnung erfolgt.

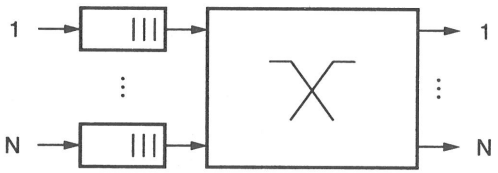


Bild 2.10: Koppelelement mit Eingangspuffern

Die Eingangspuffer können entweder nach dem FIFO-Prinzip abgearbeitet werden oder einen wahlfreien Zugriff auf beliebige Speicherplätze ermöglichen. Beide Varianten werden im folgenden ausführlich diskutiert. Eine Zusammenstellung verschiedener Möglichkeiten zur Organisation der Eingangspuffer ist auch in [164] zu finden.

2.3.1.3.2 Puffer mit FIFO-Abfertigung

Bei einer Abfertigung nach dem FIFO-Prinzip enthält der vorderste Warteplatz des Puffers jeweils diejenige Zelle, die momentan vermittelt werden soll. Falls mehrere Zellen (aus unterschiedlichen Puffern) für einen Ausgang bestimmt sind, muß die Steuerung des Koppelelementes die Kollision durch Auswahl eines Puffers auflösen. Diese Auswahl (Arbitrierung) kann entweder unabhängig vom Zustand der Puffer erfolgen, oder bestimmte Puffer-spezifische Kriterien einbeziehen:

1. Zustandsunabhängige Arbitrierungsstrategien

   (a) *Zufällige Auswahl* [105],[111],[174]: Eine der Zellen wird zufällig ausgewählt. Falls k Zellen mit gleicher Zieladresse vorhanden sind, wird jede dieser Zellen mit Wahrscheinlichkeit $1/k$ übertragen. Dies garantiert zwar eine faire Behandlung aller Eingangspuffer, wegen der schwierigen Implementierung ist diese Strategie jedoch eher von theoretischem Interesse.

   (b) *Feste Priorisierung* [11],[111]: Die Abfrage der Eingangspuffer erfolgt in einer festen Reihenfolge. Dadurch werden bestimmte Eingänge bevorzugt abgefertigt, was beispielsweise zu ungleichen Wartezeiten und Verlustwahrscheinlichkeiten führen kann [111]. Diese Strategie sollte daher nicht angewendet werden.

   (c) *Zyklische Priorisierung* [11],[132]: Die Abfrage der Eingangspuffer erfolgt ebenfalls sequentiell, allerdings wird der Startpunkt der Abfrage in jedem Zyklus um eine Position verschoben. Diese Strategie ist ebenfalls leicht zu implementieren und garantiert eine faire Behandlung der Eingangspuffer.

2. Zustandsabhängige Arbitrierungsverfahren

   (a) *Wartezeitabhängige Auswahl*: Die Zelle mit der längsten Wartezeit wird gewählt. Dabei kann entweder die gesamte Wartezeit im Eingangspuffer [174] oder nur die Wartezeit an der vordersten Stelle der Warteschlange [105] berücksichtigt werden. In Abschnitt 4.3 wird gezeigt, daß mit einer wartezeitabhängigen Arbitrierung die Schwankungen der Durchlaufzeit minimiert werden können. Allerdings ist dazu ein hoher Implementierungsaufwand erforderlich, da die Wartezeit der Zellen benötigt wird.

   (b) *Auswahl der längsten Warteschlange* [80],[111],[174]: Das Bevorzugen der längsten Warteschlange (Englisch „Longest Queue First", LQF) führt dazu, daß die Länge der einzelnen Warteschlangen möglichst ausgeglichen wird. Dadurch läßt sich zwar die Verlustwahrscheinlichkeit minimieren, gleichzeitig steigen aber die Schwankungen der Durchlaufzeit etwas an (siehe Abschnitt 4.3). Die Implementierung dieser Strategie erfordert nur einen geringen Zusatzaufwand, da ein Zähler für die Warteschlangenlänge der Eingangspuffer in den meisten Fällen ohnehin vorhanden ist.

Beide Strategien kommen jedoch nicht immer zu einem eindeutigen Ergebnis, da die Wartezeit oder die Puffergröße auch identische Werte aufweisen können. Um zu einem eindeutigen Ergebnis zu gelangen, wird eine zusätzliche Strategie benötigt. Da diese Zweitstrategie nur selten zum Tragen kommt, ist deren Einfluß kaum messbar, so daß hier auch einfache Verfahren wie eine feste Priorisierung eingesetzt werden können.

Für eine genauere Untersuchung der Eigenschaften und der Leistungsfähigkeit dieser Arbitrierungsstrategien wird auf Kapitel 4 verwiesen.

Die FIFO-Abfertigung der Puffer am Eingang des Koppelelementes hat jedoch einen wesentlichen Nachteil: falls insgesamt k Zellen aus verschiedenen Eingangspuffern zu dem selben Ausgang vermittelt werden sollen, bleiben $k-1$ Ausgänge unbenutzt, obwohl in den Warteschlangen auch andere Zellen vorhanden sein können, deren Zielausgänge frei sind. Dieser Blockierungseffekt (im Englischen auch „Head-of-Line Blocking" genannt) führt zu einer Reduzierung des maximalen Durchsatzes auf rund 58% bei einem 16×16 Koppelelement ([111], siehe auch Abschnitt 4.3).

In der Literatur sind daher verschiedene Maßnahmen zur Leistungssteigerung von Koppelanordnungen mit Eingangspuffern vorgeschlagen worden. Zu diesen Maßnahmen zählen unter anderem die Verwendung von Pufferspeichern mit wahlfreiem Zugriff (siehe unten), der Einsatz zusätzlicher Ausgangspuffer (siehe Abschnitt 2.3.1.5.1) und die Möglichkeit mehrfacher Lesezugriffe auf die Eingangspuffer. Letztere Variante erfordert eine Erhöhung der internen Zugriffsgeschwindigkeit, so daß aus einem FIFO-Eingangspuffer mehrere aufeinanderfolgende Zellen ausgelesen werden können, falls diese für bisher unbenutzte Ausgänge des Koppelelementes bestimmt sind. Auf diese Weise läßt sich bereits mit zwei möglichen Zugriffen pro Eingangspuffer ein nahezu ideales Verhalten erreichen [178].

2.3.1.3.3 Puffer mit wahlfreiem Zugriff

Eine häufig verwendete Möglichkeit zur Leistungssteigerung von Koppelelementen mit Eingangspuffern ist eine vom FIFO-Prinzip abweichende Abfertigungsdisziplin. Dies erfordert allerdings eine wesentlich aufwendigere Pufferorganisation, die ein Auslesen von Zellen an (nahezu) beliebiger Position innerhalb der Warteschlange erlaubt. Außerdem muß beachtet werden, daß die Reihenfolge von Zellen einer Verbindung nicht gestört wird. Eine Abweichung vom FIFO-Prinzip sollte daher auf solche Fälle beschränkt bleiben, in denen der oben erwähnte Blockierungseffekt zu einer Durchsatzreduzierung führen würde.

Trotz dieser Einschränkung erlauben Eingangspuffer mit wahlfreiem Zugriff eine Vielzahl verschiedener Strategien zur Auswahl der zu übertragenden Zellen. Bild 2.11 zeigt eine Klassifizierung der wichtigsten Verfahren, die von einfachen Strategien wie sequentiellem Absuchen bis zu aufwendigen Verfahren unter Einsatz von Neuronalen Netzen reichen. Im folgenden werden diese Strategien zusammen mit ihren Eigenschaften jeweils kurz beschrieben.

1. Sequentielles Absuchen [7],[96],[139]

   Diese auch „Bypass Queueing" genannte Strategie [32], [191] ist eine einfache Erweiterung der FIFO-Abfertigung. Zunächst werden nur die Zellen an vorderster Stelle

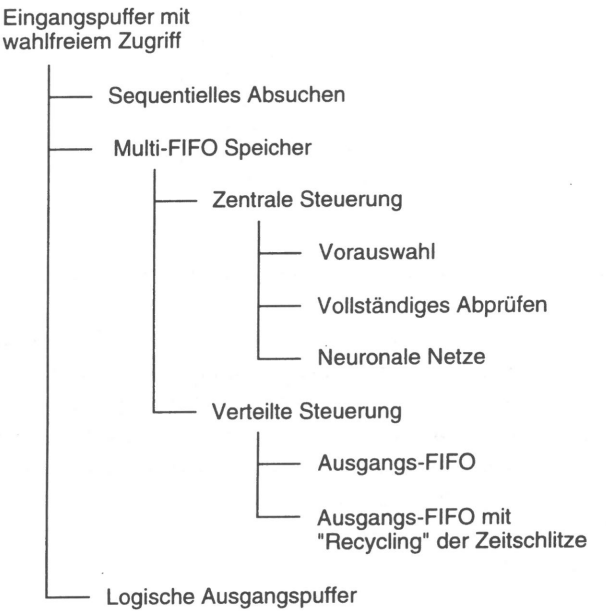


Bild 2.11: Arbitrierungsstrategien für Eingangspuffer mit wahlfreiem Zugriff

der Puffer bei der Auswahl berücksichtigt. Falls danach noch Ausgänge unbenutzt geblieben sind, werden in den bislang blockierten Puffern schrittweise auch nachfolgende Warteplätze zur Auswahl herangezogen. Dies kann bis zu einer maximalen Suchtiefe (in [96] auch Fenstergröße genannt) oder bis zum Ende der Warteschlangen fortgesetzt werden. In [96] wird gezeigt, daß mit einer Suchtiefe von 8 Warteplätzen der maximale Durchsatz bei einem 16×16 Koppelelement auf 88% gesteigert werden kann. Eine verteilte Implementierung eines derartigen Verfahrens ist in [6] und [7] beschrieben.

2. Multi-FIFO-Eingangspuffer

Die gegenseitige Behinderung von Zellen, die für unterschiedliche Ausgänge bestimmt sind, kann auch dadurch verringert werden, daß jeder Eingangspuffer in N Teilbereiche aufgeteilt wird, die jeweils einem Ausgang zugeordnet sind [203]. Jeder dieser Teilbereiche ist als FIFO-Speicher organisiert, der gesamte Speicherplatz eines Eingangspuffers steht jedoch allen Teilbereichen zur Verfügung. Die Speicherverwaltung der Eingangspuffer kann entweder mit Hilfe von verketteten Listen erfolgen [203], oder durch zusätzliche FIFO-Speicher, in denen nur die Speicheradresse der Zellen im gemeinsamen physikalischen Eingangspuffer gespeichert wird.

	1	2	3	4
1	1	1	0	1
2	0	0	0	0
3	1	0	1	0
4	0	0	1	0

Bild 2.12: Ein Beispiel für die Zustandsmatrix eines 4×4 Koppelelementes

Aufgabe der Steuerung des Koppelelementes ist es nun, durch geschickte Auswahl von Zellen einen möglichst hohen Durchsatz zu erzielen. Diese Auswahl wird zusätzlich durch die Randbedingung erschwert, daß aus jedem Eingangspuffer nur eine Zelle ausgelesen werden kann. Die Zuteilung kann entweder durch eine zentrale Steuerung erfolgen, die über ein Abbild des gesamten Systemzustands verfügt, oder durch unabhängige, auf die Ausgänge verteilte Steuerfunktionen. Durch die Kenntnis des globalen Systemzustands ist eine zentrale Steuerung zwar leistungsfähiger, aber auch wesentlich aufwendiger zu implementieren.

(a) Multi-FIFO-Eingangspuffer mit zentraler Steuerung

Der globale Zustand der Puffer kann durch eine Matrix dargestellt werden, deren Elemente x_{ij} $(1 \leq i, j \leq N)$ entweder '0' oder '1' sind. Falls am Eingang i mindestens eine für Ausgang j bestimmte Zelle vorhanden ist, wird x_{ij} der Wert '1' zugeordnet, sonst der Wert '0'. Bild 2.12 zeigt als Beispiel die Zustandsmatrix eines 4×4 Koppelelementes. Die Auswahl der Zellen kann nun als Optimierungsproblem aufgefaßt werden: es ist diejenige Kombination zu bestimmen, bei der am meisten Zellen übertragen werden, wobei als Randbedingung in jeder Spalte und in jeder Zeile der Matrix höchstens eine Zelle gewählt werden darf.

- Durch *vollständiges Abprüfen* aller möglichen Zuordnungen der Ausgänge zu den Eingangspuffern kann diejenige Kombination gefunden werden, die die Anzahl der übertragenen Zellen und somit auch den Durchsatz maximiert [133]. Dieses Vorgehen ist jedoch, vor allem bei größeren Koppelelementen, sehr zeitaufwendig und dürfte daher nur von theoretischer Bedeutung sein. Immerhin kann mit Hilfe dieser Strategie aber ein maximaler Durchsatz von 100% erreicht werden [133].
- *Neuronale Netze* werden in [3], [210] und [211] als ein vielversprechender Ansatz zur Lösung des Optimierungsproblems vorgeschlagen. Beide Konzepte beruhen auf einer Abbildung der Zustandsmatrix auf ein Neuronales Netz, dessen Neuronen den Elementen der Matrix entsprechen. Die Gewichte der

Verbindungen zwischen den Neuronen werden so eingestellt, daß unter Einhaltung der Randbedingungen möglichst viele Zellen ausgewählt werden. Mit dem in [211] beschriebenen Neuronalen Netz läßt sich ein maximaler Durchsatz von 97% erreichen, wobei die dort angegebene Hardware-Realisierung für 8×8 Koppelelemente innerhalb von 200 ns ein Ergebnis liefert.

- Durch geschickte Vorauswahl von Zellen kann die Blockierwahrscheinlichkeit ebenfalls günstig beeinflußt werden. In [133] ist ein Verfahren beschrieben, das zunächst alle Spalten der Zustandsmatrix auswählt, in denen genau eine '1' vorhanden ist, da diese Ausgänge andernfalls mit Sicherheit blockiert wären. Anschließend wird eine Zweitstrategie verwendet, die beispielsweise die Warteschlangenlänge oder die Wartezeit der Zellen berücksichtigt. Auf diese Weise kann ebenfalls ein maximaler Durchsatz von nahezu 100% erreicht werden [133]. Eine mögliche Erweiterung dieses Verfahrens besteht darin, die Vorauswahl der Zellen auch auf alle Spalten auszudehnen, in denen genau eine '1' vorhanden ist, da diese Eingänge sonst ebenfalls blockiert sind.

(b) Multi-FIFO-Eingangspuffer mit verteilter Steuerung

- In [162] ist ein verteiltes Steuerungskonzept beschrieben, bei dem jedem Ausgang eine eigene Steuerung zugeordnet ist, welche die Übertragung der Zellen von verschiedenen Eingängen zu diesem Ausgang koordiniert. Die Zuteilung erfolgt in zwei Phasen: in einer ersten Phase sendet jeder Eingang an diejenigen Ausgangs-Steuerblöcke eine Anforderung, deren zugehöriger FIFO-Speicher mindestens eine Zelle enthält. In der zweiten Phase werden diese Anforderungen in eine logische FIFO-Ausgangswarteschlange eingereiht und der daraus resultierende Zeitpunkt der Übertragung den entsprechenden Eingängen mitgeteilt. Es ist zu beachten, daß dieses Verfahren keine globale FIFO-Strategie implementiert, da die Zellen erst bei Erreichen des vordersten Warteplatzes in die Ausgangswarteschlange eingetragen werden.
  Da die logischen Ausgangswarteschlangen unabhängig voneinander arbeiten, können verschiedene Ausgänge einem Eingangspuffer den selben Übertragungszeitpunkt zuteilen. Jeder Eingang muß daher zunächst prüfen, ob die ihm zugeteilten Sendezeitpunkte noch frei sind und eventuell unbefriedigte Anforderungen im folgenden Zyklus wiederholen. Durch die fehlende Koordination zwischen den Ausgängen bleiben bereits zugeteilte Sendezeitpunkte an einigen Ausgängen unbenutzt, so daß der maximale Durchsatz bei diesem Verfahren auf 76% beschränkt ist [162].
- Um den Durchsatz weiter zu steigern, können diese unbenutzten Sendezeitpunkte anderen Eingangspuffern zur Verfügung gestellt werden [112]. Zu diesem Zweck muß der Eingangspuffer, bei dem die Sendekollision auftritt, der entsprechenden Ausgangssteuerung mitteilen, daß der betreffende Sen-

dezeitpunkt nicht genutzt werden kann. Die Steuerung versucht dann, den Sendezeitpunkt erneut zu vergeben, wobei allerdings ausgeschlossen ist, daß dieser Zeitpunkt dem selben Eingang sofort wieder zugeteilt wird. Auf diese Weise steigt die Wahrscheinlichkeit, daß ein möglicher Sendezeitpunkt auch tatsächlich benutzt wird, und der maximale Durchsatz erhöht sich auf 92% [112].

3. Logische Ausgangspuffer

   Ein Steuerungskonzept, das dem der Multi-FIFO-Eingangspuffer mit verteilter Steuerung sehr ähnlich ist, kann mit Hilfe von logischen Ausgangspuffern realisiert werden [132],[148],[149]. Die Zellen werden zwar physikalisch weiterhin an den Eingängen gespeichert, die Verwaltung der Eingangspuffer erfolgt jedoch durch FIFO-Adreßspeicher an den Ausgängen. Bei jeder Zellankunft wird die Speicheradresse und die Nummer des jeweiligen Eingangspuffers in den Adreßspeicher eingetragen, welcher dem Zielausgang der Zelle zugeordnet ist. Auf diese Weise werden die Zellen in der Reihenfolge ihres Eintreffens am Koppelelement aus den Eingangspuffern ausgelesen, das Verfahren wird daher auch als globale FIFO-Strategie bezeichnet.

   Allerdings ist auch hier eine Koordination zwischen den Zugriffen auf die Eingangspuffer erforderlich, da die Anzahl der gleichzeitigen Zugriffe beschränkt ist. Falls nur ein Zugriff je Eingangspuffer erlaubt ist, treten die gleichen Zugriffskonflikte wie bei einem Multi-FIFO-Eingangspuffer mit verteilter Steuerung auf. Eine deutliche Verbesserung kann durch zwei gleichzeitige Lesezugriffe erzielt werden [132],[178], und bereits bei drei möglichen Zugriffen je Eingang kann kaum noch ein Unterschied zu einem Koppelelement mit realen Ausgangspuffern festgestellt werden [149].

2.3.1.4 Verteilte Speicher

2.3.1.4.1 Prinzip

Bild 2.13 zeigt das Prinzip eines Koppelelements mit verteilten Speichern, die an den Kreuzungspunkten einer matrix-artigen Struktur angeordnet sind. Die Vermittlung der an einem Eingang ankommenden Zellen erfolgt durch Auswahl desjenigen Puffers, der der entsprechenden Ausgangsleitung zugeordnet ist. Die Auswahl wird häufig mit Hilfe von Adreßfiltern realisiert, die entweder direkt das VCI/VPI-Feld oder die interne Zieladresse auswerten. Auf diese Weise ist auch ein Multicast-Betrieb sehr leicht zu realisieren.

Alle Puffer einer Spalte sind mit der gemeinsamen Ausgangsleitung verbunden, so daß hier ein Arbitrierungsverfahren zur Kollisionsauflösung benötigt wird (siehe unten). Jede Spalte kann auch als verteilter Ausgangspuffer interpretiert werden, dessen Teilpuffer je einem Eingang fest zugeordnet sind. Durch entsprechende Arbitrierungsverfahren kann auch hier die

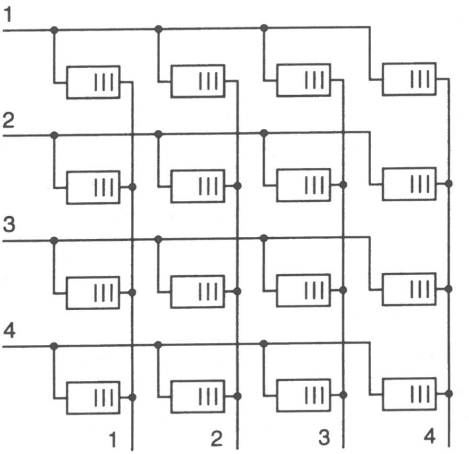


Bild 2.13: Koppelelement mit verteilten Speichern ($N = 4$)

Leistungsfähigkeit eines Koppelelementes mit Ausgangspuffern erreicht werden, wobei insgesamt jedoch mehr Speicherplatz benötigt wird, da kein Aushelfen zwischen den einzelnen Teilpuffern möglich ist. Ein weiterer Nachteil ist die große Anzahl einzelner FIFO-Speicher, die jeweils getrennt verwaltet werden müssen. Da für ein $N \times N$ Koppelelement N^2 Puffer benötigt werden, ist dieses Konzept nur für kleine N sinnvoll.

2.3.1.4.2 Arbitrierungsstrategien

Die Arbitrierungsstrategie regelt den Zugriff der zu einem Ausgang gehörigen Puffer auf die gemeinsame Ausgangsleitung. Ähnlich wie bei Koppelelementen mit FIFO-Eingangspuffern (siehe 2.3.1.3.2) können dabei Auswahlverfahren wie

- Zufällige Auswahl
- Feste Priorisierung
- Zyklische Priorisierung
- Auswahl der längsten Warteschlange

verwendet werden. Darüber hinaus ist hier auch eine globale FIFO-Strategie möglich [188], wobei der Arbiter die Zellen in der Reihenfolge ihres Eintreffens an den Eingängen auswählt. Für einen Vergleich dieser Verfahren wird auf Abschnitt 4.4 sowie auf [97] verwiesen.

2.3.1.4.3 Implementierung

Durch eine verteilte Anordnung der Ausgangspuffer kann eine interne Geschwindigkeitserhöhung auf Kosten des benötigten Speicherplatzes völlig vermieden werden. Schreibzugriffe erfolgen mit der Bitrate der Eingangsleitung, und Lesezugriffe müssen mit der Geschwindigkeit der Ausgangsleitung durchgeführt werden. Durch eine asynchrone Implementierung sind sogar unterschiedliche Bitraten an Ein- und Ausgang möglich [188]. Wegen der geringen Geschwindigkeitsanforderungen und der einfachen Struktur eignet sich das Konzept daher insbesondere für sehr hohe Bitraten und ermöglicht eine große Flexibilität in Bezug auf die Konfiguration des Koppelelementes.

Aus der Literatur sind verschiedene Implementierungsvorschläge für Koppelelemente mit verteilten Speichern bekannt [115], [131], [161], [188], [200]. Ein naheliegender Ansatz besteht darin, den FIFO-Puffer eines Kreuzungspunktes zusammen mit der erforderlichen Steuer- und Auswahllogik auf einem Chip zu integrieren [115]. Aus funktionaler Sicht ist dieses Bauelement dann einem Koppelpunkt aus der klassischen Vermittlungstechnik sehr ähnlich. Wegen der Anzahl der benötigten Bauelemente ist diese Architektur jedoch nur für kleinere Anordnungen geeignet.

In einem nächsten Schritt können, analog zu dem in [117] beschriebenen Konzept, die zu einem Ausgang gehörigen Puffer gemeinsam mit dem Arbiter auf einem Chip zusammengefaßt werden. Durch eine Parallelschaltung von N solcher Komponenten mit je M Eingängen erhält man ein $M \times N$ Koppelelement. Für kleine N kann schließlich auch ein ganzes Koppelelement auf einem Chip realisiert werden. Die Implementierung eines 2×2 Koppelelementes mit verteilten Puffern ist in [131] beschrieben.

2.3.1.5 Kombinierte Pufferkonzepte

Durch Kombination verschiedener Konzepte zur Pufferanordnung wird häufig versucht, die Vorteile der einzelnen Varianten zu vereinigen und gleichzeitig deren Nachteile zu vermeiden. Die Leistungsfähigkeit eines Koppelelementes mit Ausgangspuffern soll möglichst ohne die Verwendung von schnellen und somit teuren Speichern erreicht werden.

2.3.1.5.1 Eingangs- und Ausgangspuffer

Bild 2.14 zeigt das Konzept eines Koppelelementes, das neben den Ausgangspuffern über zusätzliche Eingangspuffer verfügt. Im allgemeinsten Fall können in einem Zelltakt aus jedem Eingangspuffer bis zu m_i Zellen gelesen und in jeden Ausgangspuffer bis zu m_o Zellen geschrieben werden. Zunächst soll jedoch nur der Sonderfall $m_i = 1$, d.h. Eingangspuffer mit Einfachzugriff, betrachtet werden.

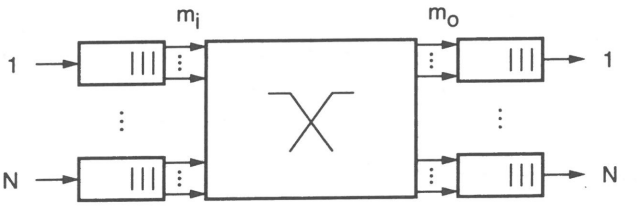


Bild 2.14: Koppelelement mit Eingangs- und Ausgangspuffern

Durch die Kombination von Eingangs- und Ausgangspuffern kann sowohl die Größe als auch die Zugriffsgeschwindigkeit der Ausgangspuffer reduziert werden. Die Speicherbandbreite läßt sich sehr einfach dadurch senken, daß in einem Zelltakt weniger als N Schreibzugriffe auf einen Ausgangspuffer zugelassen werden ($m_o < N$). Falls mehr als m_o Zellen gleichzeitig für einen Ausgang bestimmt sind, werden die restlichen Zellen in den Eingangspuffern zurückgehalten.

Die Realisierung einer derartigen Koppelnetzarchitektur ist beispielsweise in [83] und [170] beschrieben. In [48] und [163] wird gezeigt, daß bereits mit $m_o = 2$ Schreibzugriffen pro Zelltakt ein maximaler Durchsatz von 88% erreicht werden kann. Eine Erhöhung auf $m_o = 4$ läßt den Maximaldurchsatz sogar auf 99% ansteigen, so daß die Verkehrsleistung kaum noch von dem Fall $m_o = N$ (reine Ausgangspuffer) zu unterscheiden ist. Die Dimensionierung der Eingangspuffer für $m_o = 4$ ergibt, daß selbst bei einem Angebot von 90% nur 8 Pufferplätze benötigt werden, um eine Zellverlustwahrscheinlichkeit von 10^{-10} zu erreichen [138].

Auch durch eine Verringerung der Größe der aufwendigen Ausgangspuffer auf Kosten der langsameren (und billigeren) Eingangspuffer kann der Implementierungsaufwand gesenkt werden [103], [105]. Zu diesem Zweck wird innerhalb des Koppelelementes ein Rückstaumechanismus verwendet, der Zellverluste an den Ausgangspuffern verhindert. Falls ein Ausgangspuffer keine Zellen mehr aufnehmen kann, werden alle für diesen Ausgang bestimmten Zellen in den Eingangspuffern zurückgehalten. Da die Eingangspuffer meist nach dem FIFO-Prinzip organisiert sind, kann dieser Rückstaumechanismus zu einem ähnlichen Blockierungseffekt wie bei Koppelelementen mit reinen Eingangspuffern führen [168]. Je nach Größe der Ausgangspuffer wird der Rückstaumechanismus hier jedoch nur selten wirksam, so daß sich der maximale Durchsatz durch eine entsprechende Dimensionierung der Ausgangspuffer beeinflussen läßt [103], [105], [168]. Bei einer Puffergröße von 10 Zellen ergibt sich beispielsweise ein Maximaldurchsatz von 86% [103].

Es sollte jedoch beachtet werden, daß bei einer Puffergröße von S Zellen bis zu $Min(S, N)$ gleichzeitige Schreibzugriffe auf einen Ausgangspuffer erfolgen können. Es liegt daher nahe, auch hier eine Begrenzung der Speicherbandbreite auf beispielsweise $m_o = 4$ vorzunehmen [168], [170], da dies wie oben gezeigt den maximalen Durchsatz nur unwesentlich beeinflußt. Eine andere Erweiterungsmöglichkeit besteht darin, die Anzahl der Lesezugriffe pro

Eingangspuffer zu erhöhen. Dies ist vor allem bei Systemen mit kleinen Ausgangspuffern von Vorteil, weil dadurch die durch den Rückstaumechanismus hervorgerufene Blockierung der Eingangspuffer reduziert werden kann.

2.3.1.5.2 Eingangs- und Zentralspeicher

Der Speicherplatzbedarf von Koppelelementen mit Eingangs- und Ausgangspuffern läßt sich auch dadurch reduzieren, daß die Ausgangspuffer zu einem gemeinsamen, zentralen Speicher zusammengefaßt werden [104]. Ein interner Rückstaumechanismus verhindert hier ebenfalls, daß Zellen durch ein Überlaufen des zentralen Speichers verloren gehen. Auf diese Weise kann zwar die Größe des aufwendigen Speichers nochmals verringert werden, eine Reduktion der Zugriffszeit ist, anders als bei getrennten Ausgangspuffern, jedoch nicht möglich. Darüber hinaus treten die bereits in Abschnitt 2.3.1.2 diskutierten Fairness-Probleme bei der Verwaltung des Zentralspeichers gleichermaßen auf.

2.3.1.5.3 Eingangspuffer mit verteilten Speichern

In [91] wird eine verteilte Anordnung der Ausgangspuffer vorgeschlagen, um eine interne Geschwindigkeitserhöhung ganz zu vermeiden. Wie in Bild 2.15 dargestellt, verfügt jeder Ausgang über je einen Pufferplatz für jeden Eingang, so daß insgesamt N^2 Pufferplätze zusätzlich zu den Eingangspuffern vorhanden sind. Ein Rückstaumechanismus sorgt auch hier dafür, daß eine Zelle nur dann aus ihrem Eingangspuffer entnommen werden kann, wenn der dem Zielausgang zugehörige Pufferplatz frei ist.

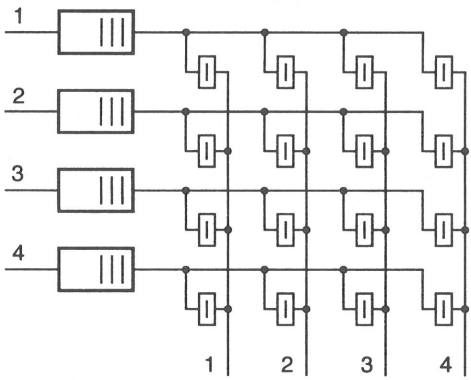


Bild 2.15: Eingangspuffer mit zusätzlichen verteilten Speichern

Der maximale Durchsatz einer derartigen Anordnung hängt sehr stark von dem Arbitrierungsverfahren ab, das den Zugriff der verteilten Puffer auf die gemeinsame Ausgangsleitung

festlegt. Durch bevorzugte Auswahl derjenigen Puffer, die bereits eine Blockierung des entsprechenden Eingangspuffers verursachen, kann bei einem 16×16 Koppelelement ein maximaler Durchsatz von rund 88% erreicht werden [91]. Durch eine Vergrößerung der verteilten Speicher läßt sich der Maximaldurchsatz auch noch weiter steigern.

2.3.1.5.4 Verteilte Speicher mit Ausgangspuffern

Ausgangspuffer mit reduzierter Zugriffsgeschwindigkeit lassen sich auch durch die Kombination mit verteilten Speichern realisieren [68],[117]. Wie in Bild 2.16 dargestellt, verfügt jeder Ausgang über M zusätzliche FIFO-Speicher, in denen die am entsprechenden Eingang ankommenden und für diesen Ausgang bestimmten Zellen zunächst gespeichert werden. In einem Zelltakt können bis zu m Zellen aus diesen Speichern ausgelesen und in den Ausgangspuffer übertragen werden, wobei sich bereits für $m = 3$ ein nahezu ideales Verkehrsverhalten ergibt. Da die verteilten Speicher mit einer um den Faktor m erhöhten Taktrate ausgelesen werden, brauchen diese nur wenige Zellen groß zu sein [117]. Bei einer Takterhöhung um den Faktor $m = 12$ genügt sogar eine Speichergröße von einer Zelle pro Puffer, um eine Zellverlustwahrscheinlichkeit von 10^{-10} zu garantieren [68].

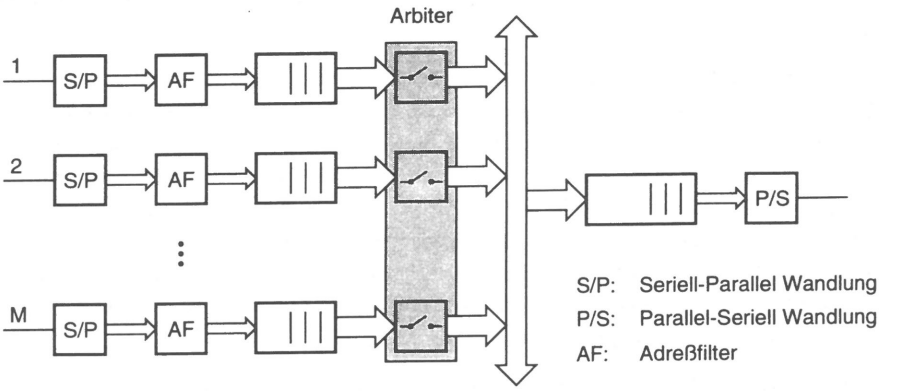


Bild 2.16: Kombination von Ausgangspuffer und verteilten Speichern

Die Vorteile der verteilten Speicher in Bezug auf Modularität und Entkopplung der Bitraten von Ein- und Ausgängen werden auf diese Weise mit der Verkehrsleistung eines Ausgangspuffers vereint. Jeder Ausgangsmodul ist völlig unabhängig von anderen Ausgängen und kann auf einem Chip integriert werden [117]. Durch paralleles Anordnen von N solcher Bauelemente lassen sich Koppelelemente mit M Eingängen und N Ausgängen realisieren.

Die in [68] und [117] vorgeschlagenen Architekturen basieren beide auf diesem Konzept. Sie unterscheiden sich im wesentlichen durch das Zugriffsprotokoll, die Takterhöhung und die

Realisierung des internen Bussystems. In [68] werden die eingangsseitigen Puffer in einer festen Reihenfolge abgefragt, wodurch eine gewisse Unfairneß hervorgerufen wird, da die ersten m Puffer nie einen Zellverlust erleiden. Um auch für den zuletzt abgefragten Puffer eine akzeptable Verlustwahrscheinlichkeit zu erhalten, wird daher eine relativ hohe interne Taktrate benötigt.

In [117] wird zwar auch eine zyklische Arbitrierungsstrategie verwendet. Die Abfrage beginnt jedoch nicht an einer festen Stelle, sondern wird bei demjenigen Puffer fortgesetzt, aus dem zuletzt eine Zelle entnommen wurde. Dadurch erreicht man eine faire Behandlung aller Eingänge, und es genügt eine interne Takterhöhung um den Faktor 3.

2.3.2 Realisierung von Verlustprioritäten

Die Bereitstellung von zwei Übermittlungsdiensten auf der ATM-Schicht, die sich durch ihre garantierte Zellverlustwahrscheinlichkeit unterscheiden, erfordert besondere Maßnahmen in Bezug auf Steuerung und Implementierung der Netzkomponenten. Die beiden Verkehrsklassen können entweder bereits beim Netzzugang getrennt werden und das Netz auf verschiedenen Wegen durchlaufen, oder innerhalb der Netzkomponenten unterschiedlich behandelt werden.

Tabelle 2.1 zeigt eine Übersicht über die im folgenden beschriebenen Möglichkeiten zur Realisierung von Verlustprioritäten. Außer durch getrennte Wege läßt sich die Zellverlustwahrscheinlichkeit auch durch eine entsprechende Pufferverwaltung beeinflussen, wobei die Priorität der Zellen beim Pufferzugang und/oder beim Auslesen aus dem Puffer berücksichtigt wird. Falls eine Zelle mit hoher Priorität andere niederpriore Zellen während des Speicherzugangs oder der Bedienung verdrängen kann, spricht man auch von verdrängenden Prioritäten. Ein weiteres Unterscheidungskriterium ist die Einhaltung der Zellreihenfolge, die bei einigen Verfahren nicht garantiert werden kann. In diesen Fällen muß daher eine Reihenfolgesicherung in der Anpassungsschicht vorgesehen werden.

Alle in Abschnitt 2.3.1 beschriebenen Architekturen können, unabhängig von der Anordnung der Speicher, um einen Mechanismus zur Unterstützung von Verlustprioritäten erweitert

Verfahren	Ort der Priorisierung	Hardware-Modifikation	Reihenfolge garantiert	Verdrängung
Wegetrennung	Netzzugang	nein	nein	nein
Getrennte Puffer	Bedienung	ja	nein	nein
Schwellwertverfahren	Pufferzugang	ja	ja	nein
Verdrängungssystem	Pufferzugang	ja	ja	ja

Tabelle 2.1: Ansätze zur Realisierung von Verlustprioritäten und deren Eigenschaften

werden. Bei Koppelelementen mit Zentralspeicher sollte der Mechanismus allerdings auf die logischen Ausgangswarteschlangen angewendet werden, um eine gegenseitige Beeinflussung unterschiedlicher Ausgänge möglichst zu vermeiden.

2.3.2.1 Wegetrennung

Die Zellen der beiden Prioritätsklassen können bereits am Netzzugang getrennt und auf unterschiedlichen Wegen vermittelt werden [128], [143]. Jede netzinterne Leitung wird ausschließlich für Zellen einer bestimmten Priorität verwendet, so daß die Verlustwahrscheinlichkeit der beiden Klassen über die Auslastung der jeweiligen Leitungen beeinflußt werden kann. Zellen mit hoher Priorität sind dabei Leitungen mit geringer Auslastung zugeordnet.

Die Auswahl der Wege für beide Prioritätsklassen erfolgt bereits beim Verbindungsaufbau. Dadurch wird zwar die Wegesuche innerhalb des Netzes etwas komplexer, es sind jedoch keine Änderungen an der Implementierung der Koppelnetze erforderlich. Allerdings kann die Zellreihenfolge gestört werden, falls innerhalb einer virtuellen Verbindung beide Prioritätsklassen benutzt werden. In diesem Fall muß daher eine Reihenfolgesicherung in der Anpassungsschicht vorgesehen werden.

2.3.2.2 Getrennte Puffer

Eine Trennung der beiden Prioritäten kann auch koppelnetzintern durch eine Aufteilung der Puffer in zwei separate Warteschlangen erfolgen, die jeweils einer Klasse fest zugeordnet sind (siehe Bild 2.17). Zellen mit hoher Priorität werden von der Bedieneinheit immer zuerst abgefertigt und besitzen daher sowohl eine kürzere Durchlaufzeit als auch eine geringere Verlustwahrscheinlichkeit. Aus diesem Grund ist jedoch auch hier ein gegenseitiges Überholen von Zellen einer virtuellen Verbindung möglich, falls diese zu unterschiedlichen Prioritätsklassen gehören.

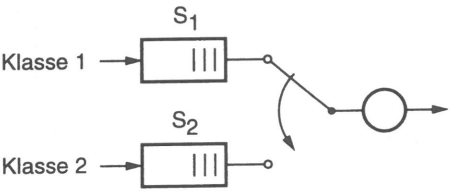


Bild 2.17: Getrennte Warteschlangen mit bevorzugter Bedienung von Zellen mit hoher Priorität

Bei Koppelelementen mit Eingangspuffern oder verteilten Speichern kann es vorkommen, daß eine an vorderster Stelle der Warteschlange stehende Zelle auf Grund des Arbitrierungsverfahrens nicht sofort bedient wird. In diesem Spezialfall kann eine niederpriore Zelle auch

durch eine Zelle mit hoher Priorität aus dem Arbitrierungsprozeß verdrängt werden. In allen anderen Fällen ist wegen der getrennter Puffer keine Verdrängung möglich.

Bei einer Realisierung mit getrennten Puffern können Zellen mit hoher Priorität allerdings auch dann abgewiesen werden, wenn noch Zellen mit niedriger Priorität gespeichert sind. Abhängig vom Verhältnis zwischen dem Verkehrsaufkommen der beiden Klassen kann sogar die Verlustwahrscheinlichkeit für Zellen mit niedriger Priorität kleiner sein als für Zellen mit hoher Priorität. Ein weiterer Nachteil ist der erhöhte Speicherplatzbedarf für die Implementierung der beiden Puffer.

2.3.2.3 Schwellwertverfahren

Schwellwertverfahren beruhen darauf, daß sich alle Zellen in einem gemeinsamen Speicher der Größe S befinden. Die Annahme von Zellen einer Prioritätsklasse ist jedoch vom Erreichen eines Grenzwertes S_i abhängig, der sich entweder auf die Gesamtzahl der im Puffer gespeicherten Zellen oder nur auf Zellen der Klasse $i, i = 1, 2$ beziehen kann (Zellen hoher Priorität werden im folgenden auch Klasse 1 genannt). In beiden Fällen werden jedoch keine bereits gespeicherten Zellen verdrängt.

In [128] wird unter der Bezeichnung „Partial Buffer Sharing“ ein Schwellwertverfahren vorgeschlagen, bei dem Zellen mit hoher Priorität den ganzen Puffer belegen können ($S_1 = S$), wogegen Zellen mit niedriger Priorität nur dann angenommen werden, wenn insgesamt weniger als S_2 Zellen im Puffer gespeichert sind (siehe Bild 2.18). Ein Teilbereich des Puffers, bestehend aus $S - S_2$ Warteplätzen, ist also für Zellen hoher Priorität reserviert.

Ein anderes Schwellwertverfahren wird in [143] vorgeschlagen. Wie in Bild 2.19 dargestellt, ist der Puffer hier in zwei Bereiche unterteilt, wobei S_1 Warteplätze für Klasse 1 sowie

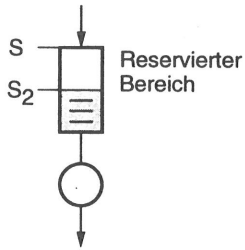


Bild 2.18: Schwellwertverfahren mit teilweiser Reservierung des Puffers [128]

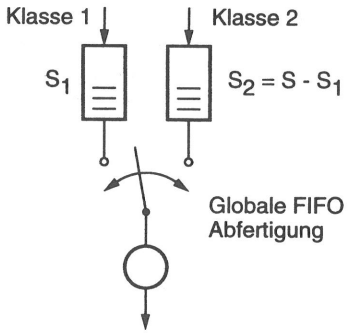


Bild 2.19: Schwellwertverfahren mit vollständiger Aufteilung des Puffers [143]

$S_2 = S - S_1$ Warteplätze für Klasse 2 zur Verfügung stehen. Diese Aufteilung hat eine gewisse Ähnlichkeit mit der oben beschriebenen Realisierung durch getrennte Puffer, die Zellen befinden sich jedoch in einem gemeinsamen physikalischen Speicher, der ebenfalls streng nach dem FIFO-Prinzip abgearbeitet wird. Die beiden Prioritätsklassen besitzen daher eine nahezu identische Verteilung der Durchlaufzeit und unterscheiden sich im wesentlichen durch ihre Zellverlustwahrscheinlichkeit.

Der zusätzliche Implementierungsaufwand der Schwellwertverfahren ist sehr gering. Für jede Verkehrsklasse wird lediglich ein Register für den Grenzwert, ein Zähler für den aktuellen Füllstand sowie die entsprechende Vergleichslogik benötigt.

2.3.2.4 Verdrängungssystem

Das Verdrängungssystem [195], [128], [143] basiert ebenfalls auf einem gemeinsamen FIFO-Puffer, in dem alle Zellen gespeichert werden. Zellen mit niedriger Priorität gehen sofort verloren, falls bei deren Ankunft keine freien Plätze mehr vorhanden sind. Zellen mit hoher Priorität können dagegen eine niederpriore Zelle (falls vorhanden) aus dem Puffer verdrängen. Auf diese Weise kann eine Zelle mit hoher Priorität nur dann verloren gehen, wenn keine Zellen mit niedriger Priorität mehr vorhanden sind.

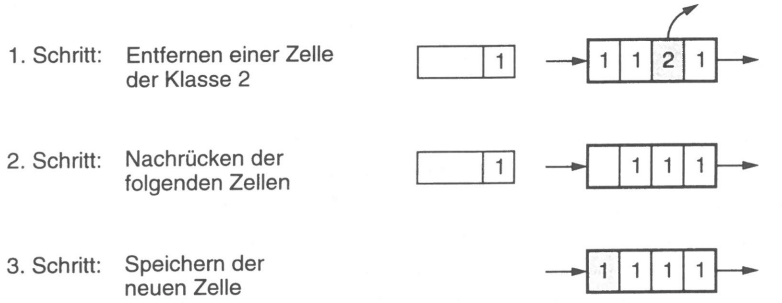


Bild 2.20: Verdrängung einer Zelle mit niedriger Priorität

Im Fall einer Verdrängung muß eine Zelle mit niedriger Priorität aus dem Puffer entfernt werden. Dabei wird meist die zuerst („First In First Drop", FIFD) oder die zuletzt („Last In First Drop", LIFD) gespeicherte niederpriore Zelle ausgewählt. Die verdrängte Zelle darf jedoch nicht einfach durch die neu angekommene Zelle ersetzt werden, da sonst die Reihenfolge der Zellen gestört werden kann. Vielmehr müssen, wie in Bild 2.20 dargestellt, alle nachfolgenden Zellen zunächst um eine Position verschoben werden. Anschließend wird die neue Zelle am Ende der Warteschlange angehängt. Die Pufferverwaltung in einem Verdrängungssystem ist

daher sehr aufwendig und kann nur mit Hilfe von mehreren verketteten Listen realisiert werden [8]. In den meisten Implementierungen von ATM-Koppelnetzen mit Verlustprioritäten wird aus diesem Grund das wesentlich einfachere Schwellwertverfahren bevorzugt [10], [124], zumal sich dessen Leistungsfähigkeit nur unwesentlich von einem Verdrängungssystem unterscheidet [128].

2.3.3 Steuerungsverfahren

Für die Vermittlung der Zellen benötigt jedes Koppelnetz interne Steuerinformationen, die aus dem VCI/VPI-Feld im Zellkopf abgeleitet werden. Diese Auswertung kann entweder nur am Eingang oder in jeder Stufe des Koppelnetzes erfolgen. Abhängig vom Ort der Auswertung unterscheidet man daher selbststeuernde und tabellengesteuerte Koppelnetze [155], [218].

2.3.3.1 Selbststeuernder Betrieb

Hier wird jede Zelle um einen internen Steuer- oder Adreßkopf erweitert, der die Steuerinformationen für alle Stufen des Koppelnetzes enthält. Jeder Stufe ist ein bestimmtes Feld des Adreßkopfes zugeordnet, welches beispielsweise die interne Adresse des jeweiligen Ausgangs angibt [64], [82], [94], [99], [181]. Die Zellen steuern den Vermittlungsvorgang also selbständig durch ihren Adreßkopf, dieses Steuerungsverfahren wird deshalb als „Selbststeuernder Betrieb“ bezeichnet.

Der Vorteil des selbststeuernden Betriebs besteht darin, daß die sehr speicheraufwendigen Übersetzungstabellen, aus denen die Steuerinformation entnommen wird, nur einmal an jedem Eingang des Koppelnetzes vorhanden sind. Da der Steuerkopf jedoch von den Zellen mitgeführt wird, muß die dafür benötigte Übertragungskapazität durch entsprechende Maßnahmen wie zum Beispiel eine interne Takterhöhung oder eine interne Lastreduktion bereitgestellt werden. Die Länge des Adreßkopfes begrenzt außerdem die Anzahl der Stufen des Koppelnetzes und muß daher für den geplanten Maximalausbau ausgelegt werden.

Die Unterstützung von Punkt-zu-Mehrpunkt-Verbindungen ist in einem selbststeuernden Koppelnetz vergleichsweise aufwendig. Da die für das Kopieren der Zellen benötigten Steuerinformationen im Adreßkopf meist nicht vollständig Platz finden, müssen besondere Adressierungsverfahren verwendet werden [94]. Ein weiteres Problem ist die Bestimmung sowie Erzeugung des VCI/VPI-Feldes von Multicast-Zellen. Um eine eindeutige Zuordnung zu ermöglichen, muß intern ein VCI/VPI gewählt werden, der auf allen betroffenen Ausgängen unbenutzt ist. Außerdem wird an jedem Ausgang eine weitere Übersetzungstabelle benötigt, um für jede kopierte Zelle einen eigenen VCI/VPI zu erzeugen.

2.3.3.2 Tabellengesteuerter Betrieb

Die Vermittlung der Zellen kann auch durch Übersetzungstabellen in den einzelnen Koppelelementen gesteuert werden, die für jede aktive Verbindung die benötigte Steuerinformation enthalten. Dies kann beispielsweise die interne Adresse des jeweiligen Zielausgangs sein, die unmittelbar an die Steuerung des Koppelelementes weitergegeben wird. Falls das Koppelelement durch Adreßfilter gesteuert wird, genügt auch ein einziges Bit zur Kennzeichnung derjenigen Verbindungen, die vom Adreßfilter durchgelassen werden müssen.

Der tabellengesteuerte Betrieb erfordert nicht nur einen höheren Aufwand zur Implementierung der Übersetzungstabellen und der zugehörigen Logik zur Überprüfung des Zellkopfes, sondern auch einen hohen Kommunikationsaufwand für deren Einstellung. Jede neue Verbindung muß in sämtliche Tabellen auf ihrem Weg durch das Koppelnetz eingetragen werden, so daß für jeden Verbindungsauf- und abbau eine Vielzahl interner Meldungen auszutauschen sind. Da die Zellen hier jedoch ohne zusätzlichen Adreßkopf auskommen, könnte die entsprechende Übertragungskapazität für die interne Kommunikation verwendet werden.

Der wesentliche Vorteil des tabellengesteuerten Betriebs liegt aber in seiner hohen Flexibilität. Da die Stufenzahl nicht durch die Länge eines internen Adreßkopfes begrenzt ist, kann das Koppelnetz beliebig konfiguriert und erweitert werden. Auch die Vermittlung von Punkt-zu-Mehrpunkt-Verbindungen wird wesentlich einfacher, da alle benötigten Informationen lokal in der Übersetzungstabelle abgelegt werden können. Koppelelemente, die durch Adreßfilter gesteuert werden, kommen sogar ohne jeglichen Zusatzaufwand zur Multicast-Adressierung aus. Eine Multicast-Verbindung wird einfach in die Adreßfilter aller Ausgänge eingetragen, zu denen die Verbindung vermittelt werden soll.

Der tabellengesteuerte Betrieb eignet sich vor allem für Koppelnetze mit einem verteilten Steuerungskonzept, bei dem die Koppelelemente auch bestimmte Funktionen des Verbindungsaufbaus übernehmen. Falls die Koppelelemente genügend Intelligenz enthalten, um eine verteilte Steuerung der Wegesuche zu ermöglichen, können auch die Übersetzungstabellen lokal verwaltet werden. Dennoch ist der Aufwand für Implementierung und Steuerung der Tabellen so hoch, daß nahezu alle bekannten Prototypen für ATM-Koppelnetze selbststeuernd betrieben werden.

2.3.4 Topologie des Koppelnetzes

ATM-Koppelnetze sind meist mehrstufige Anordnungen, die durch entsprechendes Zusammenschalten vieler Koppelelemente entstehen. Ein ATM-Koppelelement ist also die kleinste funktionelle Einheit des Koppelnetzes. Das einfachste Koppelnetz besteht aus genau einem Koppelelement, welches mindestens 2 Ein- und Ausgänge besitzen muß, um eine Vermittlung

zu ermöglichen. Die Gesamtheit aller Koppelelemente, die strukturell an gleicher Stelle des Koppelnetzes liegen, bezeichnet man auch als Koppelstufe.

Die Folge der Zwischenleitungen und Koppelelemente, über die ein Ausgang erreicht werden kann, wird Weg oder Pfad genannt. Die Länge eines Weges ist definiert als die Anzahl der Koppelelemente, die auf diesem Weg durchlaufen werden. Ein Koppelnetz ist *vollständig*, wenn jeder Ausgang von jedem Eingang aus über mindestens einen Weg erreichbar ist. Falls alle Wege die gleiche Länge besitzen, ist das Koppelnetz auch *homogen*.

Die Topologie des Koppelnetzes wird durch die Lage der Koppelelemente sowie durch die Zwischenleitungsführung beschrieben. Man unterscheidet dabei zwischen streng blockierungsfreien, bedingt blockierungsfreien und nicht blockierungsfreien Strukturen. Eine *streng blockierungsfreie Struktur* wird durch einen Graphen beschrieben, der für jede Verbindung eines freien Eingangs mit einem freien Ausgang mindestens einen freien Weg enthält. Eine *bedingt blockierungsfreie Struktur* enthält zwar nicht immer einen freien Weg, durch Umordnen der bereits bestehenden Verbindungen kann jedoch in jedem Fall ein freier Weg geschaffen werden. Falls beide Eigenschaften nicht erfüllt sind, handelt es sich um eine nicht blockierungsfreie Struktur.

Diese Definitionen haben ihren Ursprung in der klassischen Vermittlungstechnik, in der die Durchschaltevermittlung vorherrschend war. In der ATM-Technik muß der Begriff der Blockierungsfreiheit jedoch wesentlich stärker differenziert werden. Man unterscheidet hier zunächst zwischen Zell- und Verbindungsebene. Auf *Zellebene* ist ein Koppelnetz genau dann blockierungsfrei, wenn es von allen Zellen, die für unterschiedliche Ausgänge bestimmt sind (keine Abnehmerblockierung), ohne interne Kollisionen durchlaufen wird. Die Blockierungsfreiheit ist hier keine rein strukturelle Eigenschaft mehr, da auch die Steuerung der einzelnen Koppelelemente eine Rolle spielt. Außerdem ist die Unterscheidung zwischen streng und bedingt blockierungsfrei auf Zellebene nicht mehr sinnvoll, da ein Umordnen der Wege wegen der verteilten Steuerung weder möglich noch notwendig ist.

Um eine Definition der Blockierungsfreiheit auf Verbindungsebene zu ermöglichen, wird der Begriff der „Zulässigen Verbindung“ benötigt. Eine Verbindung ist genau dann zulässig, wenn auf der betreffenden Eingangs- und Ausgangsleitung des Koppelnetzes genügend freie Kapazität für die Annahme der Verbindung vorhanden ist. Ein ATM-Koppelnetz ist auf *Verbindungsebene* nur dann streng blockierungsfrei, wenn jede zulässige Verbindung unabhängig von anderen bestehenden Verbindungen angenommen werden kann. Wird das Auftreten möglicher Blockierungen durch ein entsprechendes Wegesuchverfahren verhindert, nennt man das Koppelnetz im weiteren Sinne blockierungsfrei. Ein bedingt blockierungsfreies Koppelnetz liegt dann vor, wenn jede zulässige Verbindung zwar angenommen werden kann, dabei aber bestehende Verbindungen umgeordnet werden müssen.

Die Eigenschaft der Blockierungsfreiheit hängt folglich nicht nur von der Topologie des Koppelnetzes ab, sondern auch von Steuerungsfunktionen wie Verbindungsannahme und Wege-

suche. Nach rein topologischen Gesichtspunkten ist daher eine Klassifizierung von ATM-Koppelnetzen hinsichtlich der Blockierungsfreiheit nicht möglich. Die Anzahl der Pfade zwischen den Ein- und Ausgängen eines Netzwerks ist dagegen ein eindeutiges topologisches Kriterium, mit dessen Hilfe zwischen Einpfad- und Mehrpfad-Netzwerken unterschieden werden kann. In einem Einpfad-Netzwerk kann jeder Ausgang von jedem Eingang über genau einen Weg erreicht werden. Falls einige oder alle Ausgänge über mehrere Wege erreichbar sind, handelt es sich um ein Mehrpfad-Netzwerk.

2.3.4.1 Einpfad-Netzwerke

Das Einpfad-Netzwerk ist die einfachste Möglichkeit, um eine vollständige und homogene Koppelanordnung zu realisieren. Die Grundlage des Einpfad-Netzwerkes ist eine Baumstruktur, mit deren Hilfe ein einzelner Eingang N Ausgänge erreichen kann. Bild 2.21 zeigt als Beispiel einen binären Baum, der den Ausgangspunkt für die Konstruktion eines Einpfad-Netzwerkes bildet. Durch schrittweises Hinzufügen von Koppelelementen entsteht aus diesem Baum zunächst ein 4×8, dann ein 6×8 und schließlich ein 8×8 Koppelnetz. Es kann leicht nachgeprüft werden, daß jeder Ausgang von allen Eingängen aus über eine Baumstruktur erreichbar ist; es existiert also genau ein Weg von jedem Eingang zu jedem Ausgang.

Die zur Klasse der Einpfad-Netzwerke gehörenden Koppelnetze werden in der Literatur auch als „Banyan-Netzwerke" bezeichnet [50], [169], [174], [207]. Der Begriff des Banyan-Netzwerks geht auf eine Veröffentlichung von Goke und Lipovski [86] zurück, in der die Struktur dieser Netzwerke mit Hilfe von graphentheoretischen Methoden beschrieben wurde. Im folgenden werden die Begriffe „Einpfad-Netzwerk" und „Banyan-Netzwerk" synonym verwendet.

Ein allgemeines Einpfad-Netzwerk setzt sich aus Koppelelementen mit a Eingängen und b Ausgängen zusammen. Im folgenden wird jedoch ausschließlich der Fall $a = b$ betrachtet, da der allgemeine Fall nur von geringer praktischer Bedeutung ist. Ein aus n Stufen bestehendes Koppelnetz hat $N = b^n$ Ein- und Ausgänge. Jede Stufe enthält N/b Koppelelemente, insgesamt werden also $(N/b) \log_b N$ Koppelelemente benötigt. Die Komplexität eines derartigen Netzwerks ist daher proportional zu $N \log_b N$. In [50] konnte gezeigt werden, daß die minimale Komplexität eines vollständigen Koppelnetzes ebenfalls proportional zu $N \log_b N$ ist.

Die Struktur eines Einpfad-Netzwerks ist sehr regelmäßig und außerdem modular. Von praktischer Bedeutung ist insbesondere die Möglichkeit der rekursiven Konstruktion, wodurch eine sehr gute Erweiterbarkeit gewährleistet ist. Bild 2.22 zeigt das Prinzip dieser rekursiven Erweiterung für den Sonderfall $b = 2$. Beginnend mit einem einzelnen Koppelelement lassen sich auf diese Weise Netzwerke mit beliebiger Stufenzahl schrittweise erzeugen.

Wegen der impliziten Baumstruktur eignen sich Einpfadnetzwerke besonders gut für einen selbststeuernden Betrieb. Die Adresse einer Zelle entspricht hier einfach der Nummer des

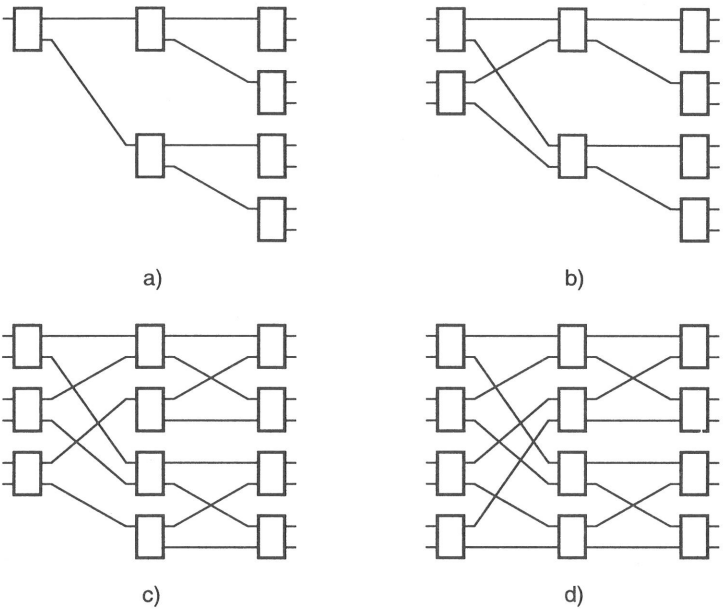


Bild 2.21: Schrittweise Erweiterung eines binären Baumes zu einem Einpfad-Netzwerk

Zielausgangs in einem Zahlensystem zur Basis b, wobei angenommen wird, daß die Ausgänge von 0 bis $N-1$ numeriert sind. Für den Spezialfall $b=2$ ergibt sich die Adresse also als Binärdarstellung der Ausgangsnummer. Die interne Zieladresse ist völlig unabhängig vom jeweiligen Eingang, so daß die Steuerung des Koppelnetzes sehr einfach ist. Diese Eigenschaft ist auch der Grund dafür, daß Banyan-Netzwerke häufig als Verbindungsnetzwerke für Parallelrechner [193] und später auch als ATM-Koppelnetze eingesetzt wurden [85], [169].

Aus der Literatur sind viele verschiedene Einpfad-Netzwerke mit unterschiedlicher Zwischenleitungsführung bekannt. Als Beispiele sollen hier nur das Flip-Netzwerk [17], das Omega-Netzwerk [135], das Baseline-Netzwerk [216] und das SW-Banyan-Netzwerk [86] genannt werden. Obwohl jedes dieser Netzwerke auf den ersten Blick eine andere Struktur besitzt, sind dennoch alle genannten Netzwerke topologisch äquivalent, das heißt sie können ohne das Auftrennen von Zwischenleitungen nur durch Umordnen der Koppelelemente einer Stufe ineinander übergeführt werden [216].

Da zwischen jedem Eingang und Ausgang genau ein interner Weg existiert, sind Einpfad-Netzwerke im allgemeinen weder auf Zellebene noch auf Verbindungsebene blockierungsfrei. Auf Zellebene führt dies zu häufigen Kollisionen, die die Leistungsfähigkeit des Koppelnetzes beeinträchtigen, und auf Verbindungsebene resultiert daraus eine sehr hohe Blockierwahr-

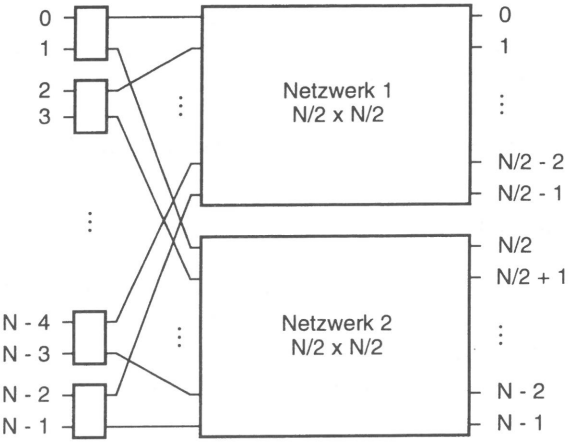


Bild 2.22: Rekursive Konstruktion von Einpfad-Netzwerken

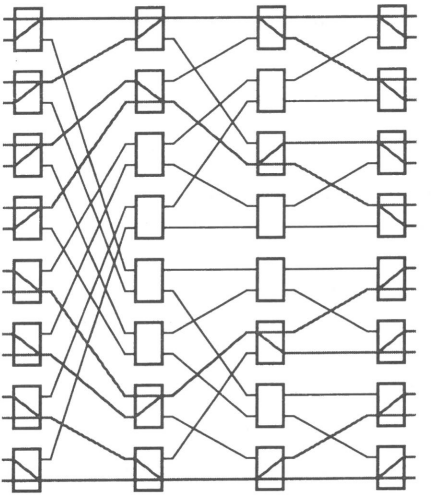


Bild 2.23: Interne Blockierung von Einpfad-Netzwerken

scheinlichkeit. Dies wird insbesondere an dem in Bild 2.23 dargestellten Beispiel deutlich, in dem sich vier Verbindungen eine interne Leitung teilen müssen. In diesem Sonderfall kann jede Eingangsleitung nur bis zu einem Viertel ihrer Kapazität belastet werden.

Unter bestimmten Bedingungen können Banyan-Netzwerke auf Zellebene allerdings auch intern blockierungsfrei sein. Dies ist beispielsweise dann der Fall, wenn die Zellen an den

Eingängen nach ihrer Zieladresse sortiert werden. Falls die Zieladressen eine streng monoton fallende oder steigende Folge bilden und zwischen benutzten Eingängen keine Lücken bleiben, kann das Netzwerk ohne interne Kollisionen durchlaufen werden [50], [120]. Diese Eigenschaft kann dazu verwendet werden, um mit Hilfe eines zusätzlichen Sortierers ein intern blockierungsfreies Koppelnetz zu realisieren (siehe Abschnitt 2.3.4.2.2).

2.3.4.2 Mehrpfad-Netzwerke

Mehrpfad-Netzwerke besitzen eine Reihe von Vorteilen gegenüber Einpfad-Netzwerken. Neben einer erhöhten Fehlertoleranz weisen Mehrpfad-Netzwerke im allgemeinen eine geringere Blockierwahrscheinlichkeit und damit einen insgesamt höheren Durchsatz auf. Für die Auswahl zwischen den einzelnen Wegen ist allerdings auch ein größerer Steuerungsaufwand erforderlich (siehe auch Kapitel 7).

Im folgenden werden einige der am häufigsten verwendeten Topologien für Mehrpfad-Netzwerke vorgestellt. In der Literatur existiert eine Vielzahl von weiteren Koppelnetzarchitekturen, eine erschöpfende Behandlung ist jedoch im Rahmen dieser Arbeit nicht möglich.

2.3.4.2.1 Erweiterte Banyan-Netzwerke

Durch das Hinzufügen zusätzlicher Stufen kann ein Banyan-Netzwerk auf einfache Weise zu einem Mehrpfad-Netzwerk erweitert werden. Jede zusätzliche Stufe vergrößert dabei die Anzahl der unterschiedlichen Wege um den Faktor b (falls $b \times b$-Koppelelemente verwendet werden). Ein bereits aus der klassischen Vermittlungstechnik bekanntes Beispiel für ein derartiges Koppelnetz ist das in Bild 2.24 dargestellte Beneš-Netzwerk [18]. Dieses Koppelnetz besitzt eine symmetrische Topologie, in der jedes Koppelelement der mittleren Stufe von jedem Eingang aus erreichbar ist. Die Anzahl der unterschiedlichen Wege ist daher gleich der Anzahl der Koppelelemente in der mittleren Stufe.

Ähnlich wie das Banyan-Netzwerk (siehe Bild 2.22) kann auch ein Beneš-Netzwerk rekursiv aus kleineren Strukturen zusammengesetzt werden. Im Unterschied zum Banyan-Netzwerk wird dabei sowohl vor als auch nach den Teilnetzwerken je eine zusätzliche Stufe benötigt, wobei die Zwischenleitungsführung den Symmetrie-Eigenschaften entspricht.

Für den Fall der Durchschaltevermittlung kann gezeigt werden, daß das Beneš-Netzwerk bedingt blockierungsfrei ist. Dazu wird allerdings ein komplexer Algorithmus für das Umordnen der bestehenden Verbindungen benötigt, der nur mit einer zentralen Steuerung realisierbar ist. In einem ATM-Koppelnetz kann die Auswahl der internen Wege entweder auf Verbindungsebene oder für jede Zelle erfolgen (siehe Abschnitt 2.3.5). Der oben genannte Algorithmus eignet sich wegen seiner Komplexität jedoch nicht für eine Auswahl der Wege auf

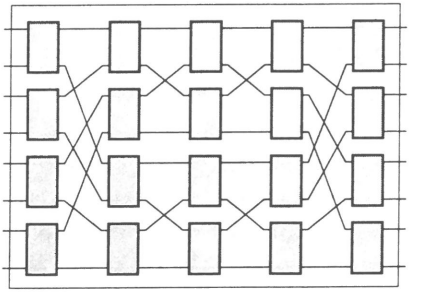


Bild 2.24: Topologie des Beneš-Netzwerks

Zellebene, hier wird meist eine zufällige Aufteilung verwendet. Durch die zufällige Verteilung des Verkehrs auf alle Koppelelemente der mittleren Stufe kann außerdem eine Beeinträchtigung der Leistungsfähigkeit durch ungünstige Verkehrsmuster (siehe Bild 2.23) vermieden werden.

Die Anzahl der internen Wege kann auch durch eine Aufteilung des Verkehrs auf m parallele Koppelnetzebenen erhöht werden. Wie in Bild 2.25 gezeigt, erfolgt die Aufteilung durch einen geeigneten Demultiplexer, der auch hier entweder auf Zellebene oder auf Verbindungsebene arbeiten kann. Außerdem werden zusätzliche Ausgangspuffer benötigt, da an jedem Koppelnetzausgang bis zu m Zellen gleichzeitig eintreffen können.

Durch die Aufweitung auf m parallele Ebenen erhöht sich die interne Gesamtkapazität um den Faktor m. Bei einer gleichmäßigen Aufteilung des Verkehrs auf alle Ebenen sinkt außerdem die Belastung jeder einzelnen Ebene um den Faktor $1/m$, so daß für die Realisierung der Ebenen auch Netzwerke mit einem Maximaldurchsatz deutlich unter 100% geeignet sind. In Kapitel 7 wird beispielsweise gezeigt, daß für den Fall $N = 64$ bereits zwei parallele Banyan-

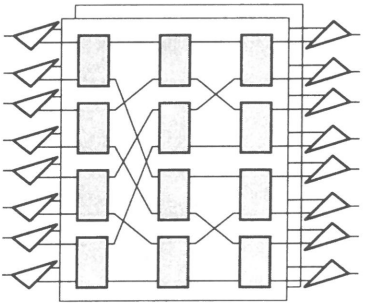


Bild 2.25: Koppelnetz mit mehreren internen Ebenen

Netzwerke genügen, um ein auf Verbindungsebene nahezu blockierungsfreies Koppelnetz zu realisieren. Diese Anordnung ist einem Beneš-Netzwerk mit 64 Eingängen, welches etwa die gleiche Anzahl von Koppelelementen benötigt, bezüglich des maximalen Durchsatzes weit überlegen.

Die Sensitivität von Banyan-Netzwerken gegenüber bestimmten Verkehrsmustern läßt sich allerdings auch durch mehrere parallele Ebenen nur unwesentlich beeinflussen. Aus diesem Grund erscheint die Kombination einer internen Parallelisierung mit einem zusätzlichen Verteilnetzwerk durchaus sinvoll, um die Vorteile beider Ansätze zu nutzen.

2.3.4.2.2 Batcher-Banyan-Netzwerk

Das Batcher-Banyan-Netzwerk basiert auf zwei grundlegenden Komponenten, aus denen sich auch dessen Bezeichnung ableitet:

- Einem Sortiernetzwerk, welches die Zellen nach ihren Zieladressen sortiert. Eine mögliche Realisierung eines solchen Sortierers ist das in [17] beschriebene Batcher-Netzwerk, das in allen bisher bekannten Implementierungen verwendet wird.
- Einem Banyan-Netzwerk, welches die sortierten Zellen zum entsprechenden Ausgang transportiert.

Die Verwendung eines Batcher-Banyan-Netzwerks als Koppelnetz wurde erstmals in [98] vorgeschlagen. Dieses Konzept ist in der Vermittlungstechnik auf großes Interesse gestoßen, und in der Zwischenzeit wurden verschiedene Prototypen auf der Basis des Batcher-Banyan-Netzwerks realisiert [6], [22], [62], [83], [166]. Eine Übersicht über die verschiedenen Varianten dieser Architektur ist in [101] enthalten.

Sortiernetzwerke als Koppelnetz: Ein Sortiernetzwerk hat die Aufgabe, die an seinen Eingängen anliegenden numerischen Daten (in diesem Fall die Zieladressen der Zellen) so zu ordnen, daß eine monoton fallende oder steigende Folge entsteht. Falls mehrere gleiche Zieladressen vorhanden sind (Abnehmerblockierung), erscheinen diese nebeneinander an benachbarten Ausgängen des Sortierers. Alle benutzten Zellen werden durch ein vorangestelltes Aktivitäts-Bit (höchstwertiges Bit) gekennzeichnet und somit automatisch von den unbenutzten Zellen getrennt.

In der Literatur existieren verschiedene Vorschläge zur Realisierung eines derartigen Netzwerks (siehe [100] und [101] sowie die dort angegebenen Literaturstellen), meist wird jedoch das Konzept des Bitonischen Sortierers nach Batcher [17] verwendet. Da die Einzelheiten der Realisierung eines Sortiernetzwerkes für die Funktion des Koppelnetzes unerheblich sind, wird im folgenden nicht weiter darauf eingegangen.

Unter bestimmten Bedingungen erfüllt das Sortiernetzwerk bereits die Funktion eines Koppelnetzes. Wenn jede Zieladresse an den Eingängen des Sortierers genau einmal vorhanden ist, werden die Zellen durch den Sortiervorgang zum korrekten Ausgang transportiert (siehe Bild 2.26a). Dies setzt allerdings voraus, daß keiner der Eingänge unbenutzt bleibt und keine Abnehmerblockierung auftritt. Beide Voraussetzungen sind im Betrieb einer Koppelanordnung jedoch selten erfüllt, so daß ein Sortierer allein noch nicht als Koppelnetz verwendbar ist.

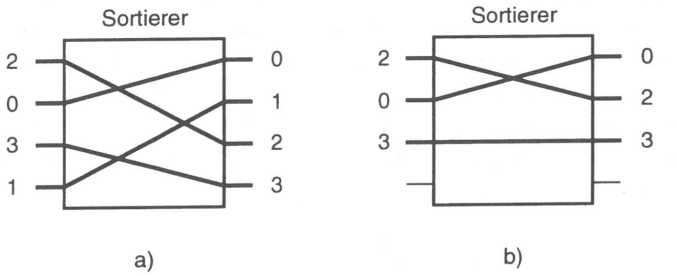


Bild 2.26: Beispiele für die Funktion eines Sortiernetzwerkes

Die Auflösung der Abnehmerblockierung erfolgt durch spezielle Arbitrierungsverfahren, die im folgenden Abschnitt genauer beschrieben werden. Falls außerdem einige Eingänge unbenutzt sind, ergibt sich die in Bild 2.26b dargestellte Situation: Durch das Fehlen von einigen Adressen stimmt die Zieladresse der Zellen nicht mehr mit der Nummer des jeweiligen Ausgangs überein. Es wird daher ein zusätzliches Verteilnetzwerk benötigt, welches die Zellen schließlich zum korrekten Ausgang transportiert.

Für diesen Zweck kann beispielsweise ein ungepuffertes Banyan-Netzwerk verwendet werden, das unter diesen speziellen Voraussetzungen ebenfalls ohne interne Blockierungen durchlaufen werden kann (siehe auch [50], [100] und [101]). Bild 2.27 veranschaulicht die Blockierungsfreiheit der gesamten Anordnung anhand des oben gezeigten Beispiels.

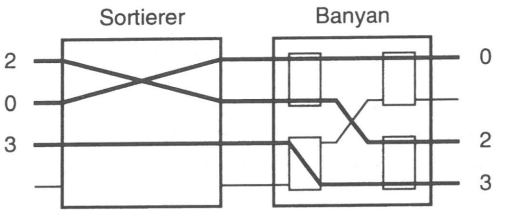


Bild 2.27: Intern blockierungsfreies 4×4 Batcher-Banyan-Netzwerk

Die Implementierung eines Batcher-Banyan-Netzwerks stellt sehr hohe Anforderungen an die Synchronisierung bzw. Taktverteilung innerhalb des Systems. Da alle Komponenten des Netzwerks sowohl Zell- als auch Bit-synchron arbeiten, müssen sämtliche Verbindungsleitungen exakt gleich lang sein. Eine sehr kompakte Realisierung ermöglicht die in [62] und [101] beschriebene dreidimensionale Aufbautechnik.

Arbitrierungsverfahren: Mit Hilfe eines entsprechenden Arbitrierungsverfahrens muß sichergestellt werden, daß jede Zieladresse an den Eingängen des Banyan-Netzwerks höchstens einmal auftritt. Dies kann wiederum unter Ausnutzung der besonderen Eigenschaften des Sortierers erfolgen: Falls einige Zieladressen mehrfach vorhanden sind, werden diese zu benachbarten Ausgängen des Sortierers gelenkt (siehe Bild 2.28). Durch einen Vergleich der Zieladressen jeweils benachbarter Zellen können derartige Konflikte daher an den Ausgängen des Sortierers leicht erkannt werden.

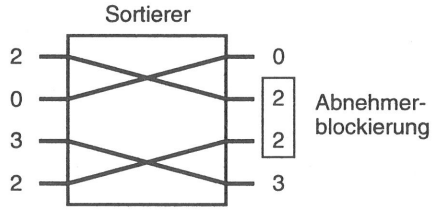


Bild 2.28: Erkennung von Konflikten am Ausgang des Sortierers

Die eigentliche Auflösung der Abnehmerblockierung kann auf zwei verschiedene Arten erfolgen. Ein möglicher Ansatz basiert auf dem Rücklauf derjenigen Zellen, die nicht sofort zu den Ausgängen vermittelt werden können [83], [98]. Alle Zellen durchlaufen zu diesem Zweck ein weiteres Sortiernetzwerk, wobei überzählige Zellen mit mehrfach vorhandenen Adressen zuvor markiert werden (siehe Bild 2.29). Am Ausgang des zweiten Sortierers sind die Zellen nun in zwei Gruppen aufgeteilt: Eine Gruppe von Zellen, die ohne Konflikte durch das Banyan-Netzwerk vermittelt werden können, und eine zweite Gruppe, die für den Rücklauf bestimmt ist. Diese Zellen werden in je einem Puffer pro Rückleitung zwischengespeichert und im nächsten Zelltakt wieder an zusätzlichen Eingängen des Koppelnetzes eingespeist.

Um die Reihenfolge der Zellen nicht zu verändern, müssen die zurückgekoppelten Zellen bei der Arbitrierung bevorzugt werden. Dazu wird an die Zieladresse ein zusätzliches Prioritätsfeld angehängt, welches bei der Sortierung ebenfalls berücksichtigt wird. Dieses Prioritätsfeld wird bei jedem Rücklauf inkrementiert und entspricht somit dem „Alter" einer Zelle. Im Fall einer Abnehmerblockierung wird die Zelle mit dem höchsten „Alter" ausgewählt, wodurch für jeden Ausgang eine logische FIFO-Warteschlange nachgebildet wird.

Als Folge des Rücklaufs ist allerdings eine starke interne Aufweitung der Koppelanordnung unumgänglich, da die Anzahl der Ein- und Ausgänge der beiden Sortierer und des Markierers um die Anzahl der Rückleitungen erhöht werden muß. Die Rückleitungen bilden mit ihren Puffern einen zentralen Speicher, in welchem nicht sofort vermittelte Zellen zwischengespeichert werden. Die Dimensionierung der Rückleitungen kann daher ähnlich wie die Dimensionierung der Speichergröße in einem Koppelelement mit Zentralspeicher erfolgen. Bei einem 16×16 Koppelnetz werden beispielsweise $M = 40$ Rückleitungen benötigt, um bei einer Last von 80% eine Zellverlustwahrscheinlichkeit von 10^{-6} zu erreichen [83].

Die zweite Art der Kollisionsauflösung basiert auf einer Speicherung der Zellen an den Eingängen des Koppelnetzes [6], [99], [167]. Um festzustellen, welche Zellen das Koppelnetz ohne Blockierung durchlaufen können, werden zunächst kurze Testzellen, die nur aus der jeweiligen Zieladresse bestehen, durch den Sortierer geschickt. Ähnlich wie oben werden hier die Kollisionen erkannt und erfolglose Zellen entsprechend markiert. Schließlich wird das Ergebnis der Arbitrierung den Eingangspuffern in Form einer positiven oder negativen Quittung mitgeteilt. Diese Mitteilung kann entweder durch eine getrennte Signalleitung in Rückrichtung erfolgen [6], [99] oder mit Hilfe eines weiteren Durchlaufs der erfolgreichen Testzellen durch das gesamte Batcher-Banyan-Netzwerk [99], [167] übertragen werden. Im Anschluß an die Arbitrierungsphase dürfen nur diejenigen Eingangspuffer eine Zelle senden, die eine positive Quittung erhalten haben.

Die Kollisionsauflösung mit Hilfe von Eingangspuffern kommt zwar ohne interne Aufweitung aus, für den Arbitrierungsvorgang wird jedoch eine gewisse Zeit benötigt, während der das Koppelnetz nicht zur Nutzdatenübertragung zur Verfügung steht. Dieser „Overhead“ muß daher durch geeignete Maßnahmen ausgeglichen werden. Wegen der Eingangspuffer ist der Maximaldurchsatz einer derartigen Koppelanordnung außerdem auf etwa 58% begrenzt.

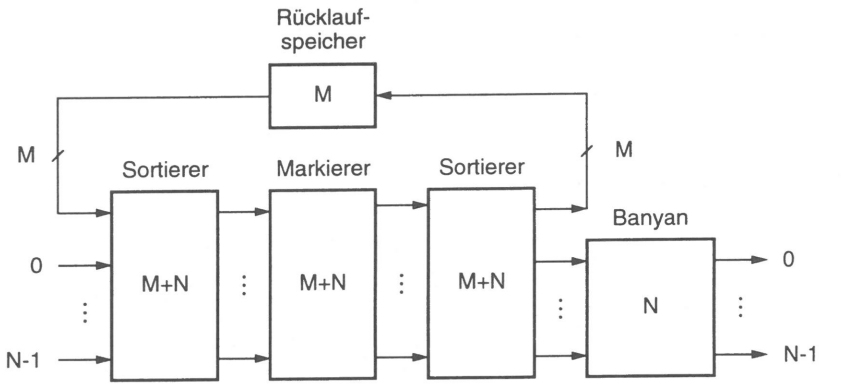


Bild 2.29: Batcher-Banyan-Netzwerk mit Rücklauf

Erweiterungen: Die Nachteile der oben beschriebenen Konzepte für den Einsatz von Batcher-Banyan-Netzwerken in Koppelnetzen haben zu verschiedenen Verbesserungsvorschlägen geführt, die entweder eine Reduzierung des Implementierungsaufwands oder eine Erhöhung der Leistungsfähigkeit zum Ziel haben.

- Die Anzahl der Rückleitungen in einem Netzwerk mit Rücklauf kann stark reduziert werden, wenn jeder Ausgang mehrere Zellen aufnehmen kann. Zu diesem Zweck werden außer zusätzlichen Ausgangspuffern auch m parallele Banyan-Netzwerke benötigt, die zusammen bis zu m Zellen je Ausgang konfliktfrei transportieren können [83]. Durch entsprechende Wahl von m erreicht man, daß die Zellen im wesentlichen nur noch an den Ausgängen gespeichert werden, so daß eine geringe Anzahl von Rückleitungen ausreichend ist. Für $m = 4$ genügt beispielsweise eine Aufweitung um 10%, damit die Verlustwahrscheinlichkeit durch den Rücklauf selbst bei einem Angebot von 100% kleiner als 10^{-10} bleibt [83]. Eine Implementierung völlig ohne Rücklauf ($m = 8$) ist in [166] beschrieben.
- Durch mehrere parallele Batcher-Banyan-Netzwerke läßt sich auch die Leistungsfähigkeit von Anordnungen mit Eingangspuffern deutlich verbessern [170]. Mit Hilfe von m parallelen Ebenen können in jedem Zelltakt bis zu m Zellen zu einem Ausgang transportiert werden, wobei auch hier zusätzliche Ausgangspuffer notwendig sind. Eine derartige Anordnung verhält sich wie ein Koppelelement mit kombinierten Eingangs- und Ausgangspuffern, das ohne internen Rückstaumechanismus betrieben wird (siehe Abschnitt 2.3.1.5.1).
- Die Einführung von logischen Leitungsbündeln („Trunk Groups“) führt insbesondere bei Batcher-Banyan-Netzwerken mit Eingangspuffern zu einer weiteren Leistungssteigerung [83], [167]. Dabei werden mehrere Ausgangsleitungen zu einem Bündel zusammengefaßt und gemeinsam adressiert. Die Zieladresse bezieht sich dabei nicht auf eine bestimmte Ausgangsleitung, sondern auf das Bündel als eine logische Einheit. Auf diese Weise können ebenfalls mehrere Zellen mit gleicher Zieladresse zu einem Ausgangsbündel vermittelt werden.

2.3.4.2.3 Clos-Netzwerk

Clos-Netzwerke [57] sind mehrstufige Strukturen mit einfacher Zwischenleitungsführung, die bereits aus der klassischen Vermittlungstechnik bekannt sind. Bild 2.30 zeigt die Topologie und die strukturellen Parameter eines 3-stufigen Clos-Netzwerks, wie es häufig auch in ATM-Koppelnetzen eingesetzt wird [81], [114], [141], [181], [189], [196]. Die Anordnung setzt sich im allgemeinen aus Koppelelementen unterschiedlicher Größe zusammen, an deren Stelle jedoch auch intern blockierungsfreie, mehrstufige Gruppierungen (z.B. Batcher-Banyan-Netzwerke) verwendet werden können.

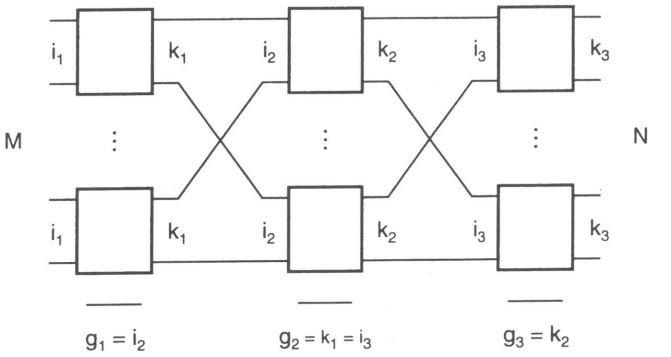


Bild 2.30: Struktur eines 3-stufigen Clos-Netzwerks

Um eine niedrige Blockierwahrscheinlichkeit zu erhalten, werden Clos-Netzwerke meist mit einer internen Aufweitung betrieben ($i_1 < k_1$). Für den Fall der Durchschaltevermittlung ist eine 3-stufige Anordnung sogar streng blockierungsfrei, wenn $k_1 \geq i_1 + k_3 - 1$ gewählt wird (eine ähnliche Bedingung gilt auch für 5-stufige Clos-Netzwerke) [57]. Diese Beziehungen können zwar nicht unverändert auf ATM-Koppelnetze übertragen werden, unter bestimmten Voraussetzungen lassen sich jedoch auch hier Bedingungen für die Blockierungsfreiheit 3-stufiger Clos-Netzwerke auf Verbindungsebene ableiten (siehe Abschnitt 7.1).

Für die Realisierung von Clos-Netzwerken mit interner Aufweitung werden eigentlich Koppelelemente mit einer ungleichen Anzahl von Ein- und Ausgängen benötigt. Dies kann zwar auch unter Verwendung von symmetrischen Koppelelementen erreicht werden, indem einige Eingänge oder Ausgänge unbeschaltet bleiben, diese Möglichkeiten sind jedoch nicht sehr wirtschaftlich. Hier bietet sich vielmehr eine weitere Realisierungsmöglichkeit auf der Basis einer Umkehrgruppierung an (siehe Bild 2.31). Umkehrgruppierungen sind Anordnungen, bei denen alle Eingangs- und Ausgangsleitungen an einer Seite des Koppelnetzes angeschlossen sind. Dies setzt allerdings eine symmetrische Struktur voraus, bei der $i_3 = k_1$ und $k_3 = i_1$ gilt.

Die Darstellung in Bild 2.31 ist so zu interpretieren, daß jede eingezeichnete Verbindung aus zwei physikalischen Leitungen besteht (je einer Leitung in Hin- und Rückrichtung). Dies kann sehr einfach realisiert werden, indem jeder Richtung jeweils ein Teil der Eingänge (Ausgänge) eines Koppelelementes zugeordnet wird. Jedes Koppelelement wird also in zwei logisch getrennte Bereiche unterteilt, die je einer Richtung zugehörig sind. Dieses Konzept hat den Vorteil, daß eine interne Aufweitung auch mit symmetrischen $b \times b$-Koppelelementen sehr einfach realisierbar ist, sofern die Bedingung $i_1 + k_1 = b$ erfüllt ist. Aus Gründen der Übersichtlichkeit werden jedoch auch Umkehrgruppierungen oft in gestreckter Form gezeichnet.

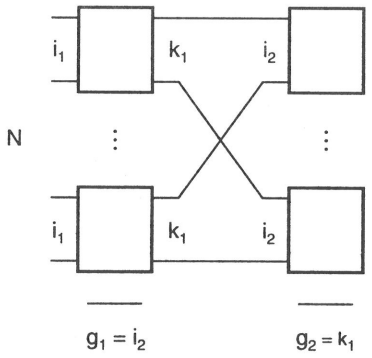


Bild 2.31: Realisierung eines Clos-Netzwerks als Umkehrgruppierung

Jeder Weg durch eine Umkehrgruppierung führt zunächst in Vorwärtsrichtung bis zur sogenannten Umkehrstufe, die normalerweise der letzten Stufe der Anordnung entspricht. Hier wechselt der Weg innerhalb eines Koppelelementes von der Vorwärtsrichtung in die Rückrichtung, durchläuft alle Stufen nochmals in umgekehrter Richtung und erreicht schließlich den gewünschten Ausgang. Darüber hinaus besteht bei bestimmten Beziehungen zwischen Eingang und Ausgang auch die Möglichkeit von Kurzwegen. Ein Kurzweg kehrt bereits vor der letzten Stufe um und belegt auf diese Weise weniger interne Leitungen. Dies ist beispielsweise dann möglich, wenn sowohl Eingangs- als auch Ausgangsleitung am selben Koppelelement der ersten Stufe angeschlossen sind.

Umkehrgruppierungen sind sehr flexibel und leicht erweiterbar. Ein Koppelnetz, das für eine bestimmte Maximalgröße ausgelegt wurde, braucht nicht von Anfang an voll ausgebaut zu sein. In der kleinsten Ausbaustufe genügt ein einzelnes Koppelelement, dessen Verbindungsleitungen zur zweiten Stufe zunächst unbeschaltet bleiben. Für eine Erweiterung werden Koppelelemente in der ersten sowie in der zweiten Stufe hinzugefügt, wobei beide Stufen auch hier noch nicht voll ausgebaut werden müssen. Auf diese Weise kann das Koppelnetz nach Bedarf schrittweise vergrößert werden. Von besonderer Bedeutung ist dabei, daß die Erweiterung ohne Änderungen an der bestehenden Konfiguration erfolgt. Es werden lediglich zusätzliche Koppelelemente und Zwischenleitungen installiert, so daß der Ausbau ohne die Beeinträchtigung aktiver Verbindungen möglich ist.

Das Konzept der Umkehrgruppierung ist natürlich nicht auf Clos-Netzwerke beschränkt, sondern auch auf andere Topologien mit entsprechenden Symmetrie-Eigenschaften anwendbar. Als Beispiel sei hier nur das wesentlich komplexere Koppelnetz in [94] genannt.

2.3.5 Verbindungskonzepte

Obwohl das Breitband-ISDN auf der ATM-Schicht prinzipiell verbindungsorientiert arbeitet, können innerhalb des Koppelnetzes auch andere Betriebsarten Verwendung finden, die jedoch nach außen hin durch entsprechende Anpassungsfunktionen verborgen werden. ATM-Koppelnetze können daher sowohl verbindungsorientiert als auch verbindungslos betrieben werden [155], [218].

2.3.5.1 Verbindungsorientierter Betrieb

In einem verbindungsorientierten Koppelnetz wird, ähnlich wie im ATM-Netz, der interne Weg einer Verbindung bereits während des Verbindungsaufbaus festgelegt. Alle Zellen einer Verbindung durchlaufen das Koppelnetz auf diesem Weg, so daß die Reihenfolge der Zellen nicht gestört werden kann (sofern innerhalb der Koppelelemente keine Reihenfolgefehler auftreten). Die Steuerung eines verbindungsorientierten Koppelnetzes ist jedoch sehr aufwendig, da die Ressourcen jeder internen Leitung getrennt verwaltet werden müssen. Während des Verbindungsaufbaus muß beispielsweise für jede interne Leitung entlang des Weges geprüft werden, ob genügend freie Bandbreite zur Annahme der Verbindung vorhanden ist. Insbesondere bei größeren Koppelanordnungen ist es daher sehr schwierig, die für den Verbindungsaufbau geltenden Realzeitanforderungen zu erfüllen. Um eine möglichst hohe Auslastung zu erzielen, wird außerdem ein leistungsfähiges Wegesuchverfahren zur Auswahl des internen Weges benötigt (siehe Kapitel 7).

2.3.5.2 Verbindungsloser Betrieb

Ein verbindungslos betriebenes Koppelnetz ist dadurch gekennzeichnet, daß die von den Zellen einer Verbindung benutzten internen Leitungen und Koppelelemente nicht während des Verbindungsaufbaus festgelegt werden. Vielmehr erfolgt eine zufällige oder lastabhängige Verteilung der Zellen auf alle möglichen Wege. Ein derartiges Koppelnetz erhält man beispielsweise mit Hilfe eines zusätzlichen Verteilnetzwerkes, das die ankommenden Zellen zufällig auf alle Eingänge des eigentlichen Koppelnetzes verteilt [213]. Die Verteilung der Zellen kann aber auch durch die zufällige Auswahl einer Leitung aus einem Leitungsbündel innerhalb des Koppelnetzes realisiert werden [33], [34], [94]. Eine weitere Möglichkeit ist die Bestimmung von Teilen des internen Adreßkopfes durch einen Pseudo-Zufallszahlengenerator.

Wegen der lokalen Bedeutung des VCI/VPI-Feldes setzen verbindungslose Koppelnetze einen selbststeuernden Betrieb voraus. Das VCI/VPI-Feld kennzeichnet jede Verbindung nur auf einer bestimmten Leitung eindeutig, so daß Verbindungen auf anderen Leitungen durchaus identische VCI/VPI-Werte aufweisen können. Bedingt durch die zufällige Verteilung der

Zellen kann die Auswertung und Übersetzung von VCI/VPI daher nur am Eingang eines verbindungslosen Koppelnetzes erfolgen. Aus dem gleichen Grund ist auch ein Multicast-Betrieb nur sehr schwierig zu realisieren.

Eine weitere Schwierigkeit stellen Reihenfolgefehler dar, die in einem gepufferten Koppelnetz nicht völlig auszuschließen sind, da sich einzelne Zellen einer Verbindung auf Grund von stark unterschiedlichen Durchlaufzeiten innerhalb des Koppelnetzes überholen können. In diesem Fall werden daher spezielle Mechanismen zur Reihenfolgesicherung benötigt, die meist mit Hilfe eines Zeitstempels arbeiten [94].

Die Steuerung eines verbindungslosen Koppelnetzes ist dagegen sehr einfach, da eine Entscheidung über die Verbindungsannahme nur für die jeweilige Ausgangsleitung getroffen werden muß. Auf Verbindungsebene ist das Koppelnetz daher intern blockierungsfrei. Die Auffächerung der Zellen einer Verbindung hat außerdem eine gewisse Glättung des internen Verkehrs zur Folge, wodurch beispielsweise der stoßartige Charakter einzelner Verbindungen zunächst stark abgeschwächt wird. Da die Zellen einer Verbindung jedoch den gleichen Zielausgang besitzen, wird die Verkehrscharakteristik am Ausgang nahezu unverändert wiederhergestellt. Durch entsprechende Auslegung insbesondere der hinteren Stufen des Koppelnetzes muß daher sichergestellt sein, daß die Verkehrsgüte auf Zellebene in jedem Fall eingehalten wird.

Kapitel 3

Verkehrsmodelle und Analyseverfahren

Die Modellierung von Kommunikationssystemen im Sinne der Leistungsuntersuchung ist ein wichtiges Hilfsmittel sowohl in der Entwurfsphase als auch für spätere Dimensionierungs- und Optimierungsaufgaben. Dies trifft insbesondere auf den Entwurf von ATM-Koppelnetzen zu, die durch einen engen Zusammenhang zwischen Architektur und Verkehrsleistung gekennzeichnet sind. Die Bewertung verschiedener Systemalternativen erfordert daher eine möglichst wirklichkeitsnahe Modellierung, um deren Leistungsfähigkeit beurteilen und vergleichen zu können.

Mit Hilfe einer verkehrstheoretischen Modellierung sollen quantitative Aussagen über die Verkehrsleistung eines Systems gewonnen werden. Zu diesem Zweck wird das reale System durch ein abstraktes Modell nachgebildet, welches die wesentlichen Systemeigenschaften enthält. Dieses abstrakte Modell beschreibt nicht nur die Struktur des Systems und dessen Komponenten, sondern auch das Ablaufgeschehen im System und die Beziehungen zwischen den Komponenten. Darüber hinaus wird auch eine Beschreibung des Verkehrs benötigt, durch den das System belastet wird. Im Zusammenhang mit ATM-Koppelnetzen handelt es sich dabei um Quellmodelle, welche beispielsweise den Ankunftsprozeß der Zellen am Eingang der Koppelanordnung bestimmen. Auch der Verkehrsfluß innerhalb des Netzwerks hängt mit der Erzeugung des Verkehrs zusammen und muß durch geeignete Mechanismen nachgebildet werden.

In diesem Kapitel werden einige grundlegende Aspekte der Modellierung von ATM-Koppelnetzen behandelt. Neben der Modellierung der Verkehrsquellen und des internen Verkehrsflusses wird auch auf verschiedene Methoden zur Leistungsuntersuchung eingegangen, welche die Grundlage der Untersuchungen in den folgenden Kapiteln bilden.

3.1 Modellierung von Verkehrsquellen

Die Modellierung der Verkehrsquellen erfordert eine möglichst genaue Kenntnis der Kommunikationsdienste und des Teilnehmerverhaltens. Das Teilnehmerverhalten hat beispielsweise einen großen Einfluß auf die Verbindungsdauer und die Häufigkeit der Verbindungswünsche. Erschwerend kommt hinzu, daß diese Parameter meist auch zeitlichen Schwankungen unterworfen sind, die von der Tageszeit, dem Wochentag und teilweise sogar von der Jahreszeit abhängen. Aus diesem Grund werden Kommunikationsnetze meist nach der Spitzenbelastung dimensioniert, die in repräsentativen Hauptverkehrsstunden auftritt.

Die Eigenschaften des von den Endgeräten erzeugten Zellstroms hängen außer vom Teilnehmerverhalten ganz wesentlich vom jeweiligen Kommunikationsdienst ab. Man unterscheidet hier meist zwischen Diensten mit konstanter Bitrate (Constant Bit Rate, CBR) und mit variabler Bitrate (Variable Bit Rate, VBR). Während die Modellierung von Verbindungen mit konstanter Bitrate keine Schwierigkeiten bereitet, ist eine exakte und realistische Charakterisierung von Verkehrsquellen mit variabler Bitrate im allgemeinen sehr aufwendig. Hier werden oft hierarchische Quellmodelle verwendet, die das statistische Verhalten einer Verkehrsquelle auf mehreren Ebenen beschreiben [76],[183].

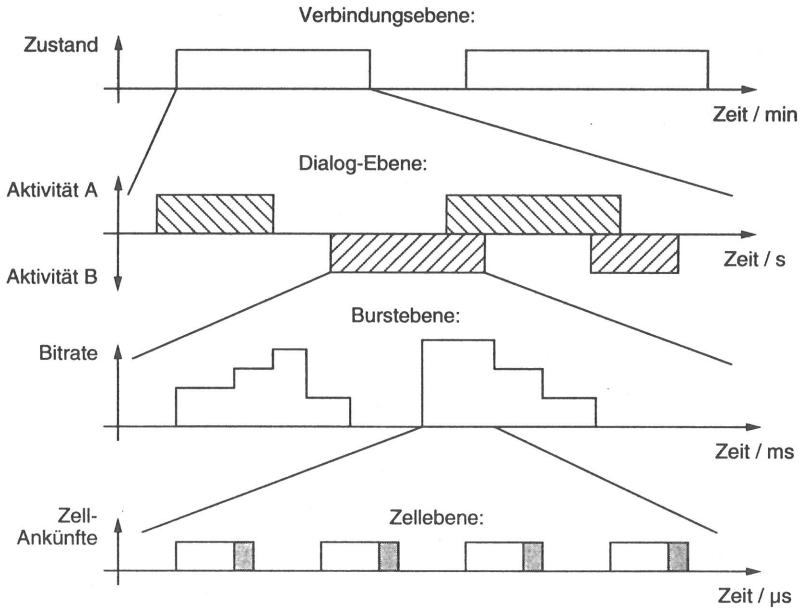


Bild 3.1: Hierarchische Modellierung des Quellverhaltens

Bild 3.1 zeigt die verschiedenen Ebenen zur Modellierung des Quellverhaltens in einem ATM-Netz [175], wobei als oberste Ebene die Verbindungsebene gewählt wurde. Die einzelnen Ebenen erfassen bestimmte charakteristische Eigenschaften der Quellen, die teilweise vom Ablauf der Kommunikation und teilweise von internen Funktionen wie dem Kodierverfahren abhängig sind. Die hierarchische Struktur des Modells spiegelt sich auch in den Beziehungen zwischen den Ebenen wider: Der Zustand der Quelle auf einer Ebene bestimmt das Verhalten und die Aktivität auf den darunterliegenden Ebenen.

Die zeitlichen Abläufe der einzelnen Ebenen reichen von wenigen Mikrosekunden auf Zellebene bis hin zu mehreren Stunden auf Verbindungsebene. Dies entspricht einer zeitlichen Dynamik in der Größenordnung von neun Zehnerpotenzen, die bei einer gleichzeitigen Modellierung aller Ebenen zu großen Problemen führt. Bei der Systemanalyse beschränkt man sich daher meist auf die Modellierung von ein oder zwei Ebenen, wobei der Zustand der darüberliegenden Ebenen „eingefroren“ wird.

Die Aktivität der Quelle innerhalb einer Ebene wird häufig durch zustandsgesteuerte Modelle beschrieben. Jeder Zustand ist gekennzeichnet durch die Verteilung der Zustandsdauer und durch die Übergangswahrscheinlichkeiten, von denen der Folgezustand abhängt. Die Markov-Kette ist ein derartiges Zustandsmodell, in welchem die Dauer eines Zustands entweder negativ-exponentiell (zeitkontinuierlich) oder geometrisch (zeitdiskret) verteilt ist. Folglich genügt die Angabe des Mittelwertes zur vollständigen Charakterisierung der Zustandsdauer. Die Übergänge zwischen den Zuständen werden durch eine Übergangsmatrix $\mathbf{P} = [p_{ij}]$ beschrieben, wobei p_{ij} die Wahrscheinlichkeit für einen Übergang vom Zustand i in den Zustand j ist.

Im folgenden werden die einzelnen Ebenen zusammen mit geeigneten Modellen zur Nachbildung ihrer statistischen Eigenschaften genauer beschrieben.

3.1.1 Verbindungsebene

Die Verbindungsebene ist üblicherweise die oberste Hierarchiestufe bei der Modellierung von Verkehrsquellen. Wie in Bild 3.1 dargestellt, existieren auf Verbindungsebene meist nur zwei Zustände: ein aktiver Zustand, in dem eine virtuelle Verbindung besteht, und ein passiver Zustand. Die Dauer der Zustände hängt sehr stark vom jeweiligen Dienst ab. Typische Verbindungsdauern liegen beispielsweise bei Telefongesprächen im Minutenbereich und bei Videokonferenzen eher im Stundenbereich.

Für die Modellierung der Verbindungsebene genügt daher ein einfaches Quellmodell mit zwei Zuständen. Die Verbindungsdauer wie auch die Dauer der Pause zwischen zwei Verbindungen wird im allgemeinen durch eine negativ-exponentielle Verteilung charakterisiert, das Quellverhalten kann also durch eine Markov-Kette mit zwei Zuständen modelliert werden.

3.1.2 Dialog-Ebene

Die Dialog-Ebene beschreibt die Aktivität der Quelle innerhalb einer virtuellen Verbindung, wobei insbesondere die Abhängigkeiten zwischen den Kommunikationspartnern bei interaktiven Diensten berücksichtigt werden. Ein typisches Beispiel ist das Telefongespräch, welches besonders ausgeprägte Abhängigkeiten aufweist. Bei einer normalen Konversation spricht meist nur einer der beiden Teilnehmer, während sein Gegenüber zuhört und schweigt. Es können jedoch auch Zeitintervalle auftreten, in denen beide Teilnehmer gleichzeitig sprechen oder gleichzeitig schweigen.

Die Abhängigkeiten einer wechselseitigen Kommunikation sind vor allem dann von großer Bedeutung, wenn beide Kommunikationsrichtungen über ein gemeinsames physikalisches Medium übertragen werden, wie zum Beispiel ein lokales Netz mit Bus- oder Ringstruktur. In einem öffentlichen Weitverkehrsnetz wie dem Breitband-ISDN werden jedoch im Normalfall für Hin- und Rückrichtung völlig getrennte Übertragungskanäle benutzt, so daß die gegenseitigen Abhängigkeiten nur eine untergeordnete Rolle spielen. Hier ist lediglich der Wechsel zwischen Phasen der Aktivität und den dazwischenliegenden Pausen relevant.

Bei der Übertragung von Videosignalen mit variabler Bitrate ist ebenfalls eine Aktivitätsschwankung auf Dialog-Ebene zu beobachten, welche auf die Abfolge von Szenen mit unterschiedlicher Aktivität zurückzuführen ist. Darüber hinaus wird bei Szenenwechseln kurzzeitig eine sehr hohe Bitrate benötigt, da der Bildinhalt vollständig neu aufgebaut werden muß.

Die Dynamik der Dialog-Ebene liegt im Sekundenbereich, bei der Videokommunikation sind auch Szenendauern bis zu mehreren Minuten möglich. Falls die Aktivität nur zwischen zwei diskreten Zuständen schwankt, können hier ähnliche Modelle wie auf Verbindungsebene verwendet werden. Insbesondere für die Nachbildung von Videoquellen sind allerdings wesentlich komplexere Modelle notwendig, die an dieser Stelle jedoch nicht genauer beschrieben werden sollen. Der interessierte Leser wird auf [26] und [187] verwiesen.

3.1.3 Burst-Ebene

Ein sogenannter „Burst“ ist eine Folge von Zellen, die mit konstanter Rate von der Quelle ausgesendet werden. Ein solcher Burst entsteht zum Beispiel bei der Segmentierung eines größeren Datenpakets in mehrere Zellen, die anschließend mit der zulässigen Spitzenbitrate übertragen werden. Bei der Sprachübertragung führt der Einsatz einer Pausenunterdrückung ebenfalls zu ausgeprägten Bursts, die durch kurze Pausen zwischen Silben oder Wörtern unterbrochen sind. Auch bei der Bewegtbild-Kommunikation können Bursts beobachtet werden, die auf eine zeilen- oder bildweise Kodierung des Videosignals zurückzuführen sind. Die Zellrate eines Bursts ist hier vom Bildinhalt abhängig und ändert sich bei einer Bildfrequenz

von 25 Hz periodisch alle 40 ms. In allen genannten Beispielen besteht zwischen den Zellen eines Bursts auch ein gewisser inhaltlicher Zusammenhang, der durch die Zuordnung zu einer gemeinsamen übergeordneten Dateneinheit (z.B. Paket, Silbe oder Bild) hervorgerufen wird.

Die statistischen Eigenschaften der Burst-Ebene werden durch Parameter wie der mittleren Burst-Dauer und der Zellrate innerhalb eines Bursts beschrieben. Ein typisches Modell der Burst-Ebene ist beispielsweise das sogenannte „Fluid Flow"-Modell [4],[150], welches im wesentlichen aus einer Markov-Kette besteht, deren Zustand die momentane Ankunftsrate bestimmt. Die Besonderheit dieses Modells besteht darin, daß der Zellstrom als Flüssigkeit modelliert wird, die mit einer entsprechenden Rate fließt. Der diskrete Charakter der Zellankünfte wird in einem „Fluid Flow"-Modell also nicht berücksichtigt.

Auch auf der Burst-Ebene wird oft der Sonderfall einer Markov-Kette mit nur zwei Zuständen verwendet: einem aktiven Zustand und einem Pausenzustand, in dem keine Ankünfte erzeugt werden. Die im folgenden Abschnitt beschriebene „Sporadische Quelle" basiert ebenfalls auf einem derartigen Modell.

3.1.4 Zellebene

Die Zellebene beschreibt schließlich den Ankunftsprozeß der Zellen innerhalb eines Bursts, wobei es sich um zeitliche Abläufe im Bereich von Mikrosekunden handelt. Die Modellierung auf Zellebene erfolgt häufig mit Hilfe von sogenannten Erneuerungsprozessen, die durch statistisch voneinander unabhängige Ankunftsabstände gekennzeichnet sind. Ein stationärer Erneuerungsprozeß hat außerdem die Eigenschaft, daß alle Ankunftsabstände dieselbe Verteilungsfunktion besitzen (Independent and Identically Distributed, IID).

Zur Klasse der Erneuerungsprozesse gehören beispielsweise der Poisson-Prozeß und der Bernoulli-Prozeß. Beide Prozesse haben außerdem die Eigenschaft der Gedächtnisfreiheit: Die Wahrscheinlichkeit dafür, daß innerhalb eines beliebigen Zeitintervalls $[t, t+u]$ ein Ankunftsereignis auftritt, hängt nur von der Länge u des Intervalls ab und nicht vom Abstand zur vorhergehenden Ankunft. Auf Grund dieser Eigenschaft eignet sich der Poisson-Prozeß wie auch der Bernoulli-Prozeß besonders gut für analytische Untersuchungen.

Bei einem *Poisson-Prozeß* besitzt der Abstand T_A zwischen zwei Ankunftsereignissen eine negativ-exponentielle Verteilungsfunktion:

$$\mathrm{P}\{T_A \leq t\} = 1 - e^{-\lambda t}, \quad t \geq 0, \tag{3.1}$$

wobei λ die Ankunftsrate und $\mathrm{E}[T_A] = 1/\lambda$ der mittlere Ankunftsabstand ist. Der Poisson-Prozeß ist allerdings ein zeitkontinuierlicher Prozeß und eignet sich daher weniger für die Modellierung von Zellankünften in einem ATM-Netz. Hier wird meist der Bernoulli-Prozeß verwendet, der auch als zeitdiskrete Version des Poisson-Prozesses aufgefaßt werden kann.

Ein zeitdiskreter Prozeß ist dadurch gekennzeichnet, daß Ereignisse nur zu diskreten Zeitpunkten $k \cdot \Delta T$ auftreten können. Die kleinste Zeiteinheit ΔT, auf die alle Zeitangaben bezogen werden, entspricht in einem ATM-Netz der Übertragungsdauer einer Zelle. Bei einem *Bernoulli-Prozeß* ist die Wahrscheinlichkeit q für das Eintreffen einer Ankunft zu jedem Zeitpunkt gleich groß, unabhängig von den Ankünften zu anderen Zeitpunkten. Der Bernoulli-Prozeß kann auch als eine Folge von unabhängigen Bernoulli-Versuchen interpretiert werden, wobei jeweils mit Wahrscheinlichkeit q ein „Erfolg" (also eine Zellankunft) eintritt. Aus dieser Definition läßt sich die diskrete Wahrscheinlichkeitsverteilung des Ankunftsabstands unmittelbar ableiten:

$$\mathrm{P}\{T_A = k \cdot \Delta T\} = q\,(1-q)^{k-1}, \quad k \geq 1, \tag{3.2}$$

wobei hier die Ankunftsrate durch $q/\Delta T$ und der mittlere Ankunftsabstand durch $\mathrm{E}[T_A] = \Delta T/q$ gegeben ist. Die Ankunftsabstände besitzen also eine geometrische Verteilung, die um eine Zeiteinheit verschoben ist.

Für die Modellierung von Quellen mit konstanter Bitrate wird ein weiterer Erneuerungsprozeß verwendet, der einen konstanten Abstand zwischen zwei Zellankünften aufweist. Da der Ankunftsabstand nicht mehr zufallsabhängig ist, spricht man hier auch von einem *Deterministischen Prozeß*. Der Ankunftsabstand besitzt natürlich eine deterministische Verteilung:

$$P\{T_A = k \cdot \Delta T\} = \begin{cases} 1, & k = d, \\ 0, & \text{sonst}, \end{cases} \tag{3.3}$$

wobei d der (konstante) Abstand zwischen zwei Ankünften ist.

Einige Quellmodelle erlauben auch die gleichzeitige Modellierung von Burst- und Zellebene. Ein Beispiel dafür ist der allgemein modulierte deterministische Prozeß (General Modulated Deterministic Process, GMDP) [130], der auf Burst-Ebene durch einen endlichen Zustandsautomaten gesteuert wird. In jedem Zustand werden Zellen mit einem konstanten, aber zustandsabhängigen Abstand d_i erzeugt. Anders als bei einer Markov-Kette kann die Anzahl der in einem Zustand erzeugten Zellen hier jedoch durch eine beliebige, zustandsindividuelle Verteilung beschrieben werden, wobei auch Pausenzustände ohne Ankünfte möglich sind. Ein Beispiel für den Ankunftsprozeß eines GMDP mit zwei aktiven Zuständen und einem Pausenzustand ist in Bild 3.2 dargestellt.

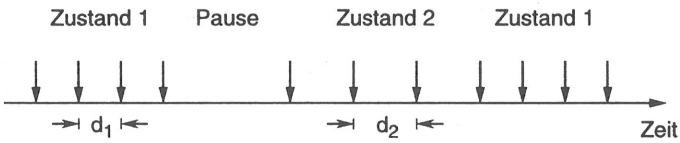


Bild 3.2: Allgemein modulierter deterministischer Prozeß mit drei Zuständen

Der GMDP ist kein Erneuerungsprozeß, da die Ankunftsabstände wegen der Modulation auf Burst-Ebene meist hochgradig korreliert sind. Falls die Anzahl der Zellen in einem Burst geometrisch verteilt ist, kann das Korrelationsverhalten analytisch berechnet werden [23], für den allgemeinen Fall ist dies jedoch sehr aufwendig. Die Verteilung des Ankunftsabstands sowie deren Momente sind in [204] hergeleitet.

Die „Sporadische Quelle“ (auch Burst-Silence-Quelle oder Talkspurt-Silence-Quelle genannt) ist ein Sonderfall des GMDP mit nur zwei Zuständen: einem aktiven Zustand und einem Pausenzustand. Die Anzahl der Zellen im aktiven Zustand ist geometrisch verteilt, wodurch der resultierende Ankunftsprozeß zu einem Erneuerungsprozeß wird. Die Pausendauer kann je nach Anwendungsfall entweder geometrisch oder negativ-exponentiell verteilt sein. Eine ausführlichere Beschreibung der statistischen Eigenschaften dieses Modells ist in [175] zu finden.

Ursprünglich wurde die Sporadische Quelle (unter der Bezeichnung Talkspurt-Silence-Quelle) als Modell für die paketierte Sprachübertragung mit Pausenunterdrückung verwendet [61]. Als eines der einfachsten Modelle zur gleichzeitigen Nachbildung von Burst- und Zellebene wird die Sporadische Quelle jedoch auch für die verschiedensten Untersuchungen in ATM-Netzen eingesetzt (siehe Kapitel 6).

3.2 Modellierung des Verkehrsflusses

Der Verkehrsfluß in einem Koppelnetz hängt sowohl von der Belastung der einzelnen Eingänge als auch von der Verteilung des angebotenen Verkehrs auf die Ausgänge ab. Meist wird eine vollkommen symmetrische Belastung aller Eingänge sowie eine gleichmäßige Aufteilung des Verkehrs auf alle Ausgänge angenommen (uniform traffic). Falls diese Bedingungen nicht erfüllt sind, spricht man auch von einer „Verkehrsschiefe“. Da jede Eingangs- bzw. Ausgangsleitung nicht über 100% hinaus belastet werden kann, führt eine Verkehrsschiefe zwangsläufig zu einer Reduzierung des gesamten Durchsatzes. Obwohl in der Literatur auch verschiedene Studien bezüglich der Verkehrsschiefe durchgeführt wurden (siehe beispielsweise [16], [52], [106] und [144]), wird in dieser Arbeit im wesentlichen nur der Fall einer symmetrischen und gleichmäßigen Belastung aller Ein- und Ausgänge untersucht.

Bei der internen Verzweigung des Verkehrs muß auch der verbindungsorientierte Charakter von ATM-Netzen berücksichtigt werden, der dazu führt, daß alle Zellen einer virtuellen Verbindung die gleiche Zieladresse aufweisen. Diese Tatsache ist insbesondere bei solchen Koppelnetzen von Bedeutung, die intern ebenfalls verbindungsorientiert arbeiten. In einer Simulation kann der verbindungsorientierte Betrieb leicht nachgebildet werden, bei analytischen Untersuchungen wird jedoch meist von einer zufälligen Verzweigung ausgegangen,

wobei die Verzweigungswahrscheinlichkeit dem Verkehrsanteil in der jeweiligen Richtung entspricht.

Auf Grund der zufälligen Verzweigung gehen jedoch die charakteristischen Verkehrseigenschaften einer virtuellen Verbindung nahezu vollständig verloren, so daß insbesondere auf Burst-Ebene signifikante Abweichungen auftreten können ([174], siehe auch Kapitel 6). Bei Koppelelementen mit Ausgangspuffern kann dieser Effekt teilweise vermieden werden, indem angenommen wird, daß sich die Quellen der entsprechenden virtuellen Verbindungen unmittelbar am Eingang des Puffers befinden. Dadurch wird zwar der Einfluß aller vorhergehenden Koppelelemente auf die Verkehrscharakteristik vernachlässigt, in den meisten Fällen liefert diese Näherung jedoch durchaus brauchbare Ergebnisse [129]. Dies ist darauf zurückzuführen, daß die Verkehreigenschaften auf Burst-Ebene innerhalb des Netzes weitgehend erhalten bleiben.

3.3 Methoden zur Leistungsuntersuchung

Die Verkehrsleistung von ATM-Koppelnetzen wird durch Parameter wie die Zellverlustwahrscheinlichkeit und den Maximaldurchsatz bestimmt. Diese Größen können sowohl durch Messungen an einem realen System als auch mit Hilfe von geeigneten verkehrstheoretischen Methoden ermittelt werden. Da in der Entwurfsphase naturgemäß noch kein reales System vorhanden ist, muß hier auf verkehrstheoretische Methoden zurückgegriffen werden. Neben der ereignisorientierten Simulationstechnik können in bestimmten Fällen auch analytische Berechnungsverfahren eingesetzt werden. Im folgenden wird ein kurzer Überblick über die in der vorliegenden Arbeit verwendeten analytischen und simulativen Methoden gegeben, für eine ausführlichere Behandlung wird auf die entsprechende Literatur verwiesen (siehe z.B. [121], [134] und [90]).

3.3.1 Simulationstechniken

Die ereignisgesteuerte Systemsimulation ist eine sehr verbreitete Methode zur Leistungsuntersuchung von Kommunikationssystemen. Sie basiert auf einem Modell, welches das zu untersuchende System auf einem Rechner nachbildet. Dabei werden die strukturellen Systemkomponenten wie Zellen, Warteschlangen, Koppelelemente und Zwischenleitungen auf entsprechende Datenstrukturen abgebildet. Funktionale Abläufe wie der Transport von Zellen durch das Koppelnetz werden von systemspezifischen Prozeduren durchgeführt, welche die internen Datenstrukturen entsprechend manipulieren.

Der Ablauf der Simulation wird ausschließlich durch Ereignisse wie die Ankunft einer Zelle oder die Bedienung einer Warteschlange gesteuert. Die Bearbeitung dieser Ereignisse erfolgt

in der zeitlichen Reihenfolge ihres Eintretens, wobei die Zeit zwischen zwei Ereignissen übersprungen wird. Das jeweils nächste Ereignis wird aus einer Ereignis-Liste entnommen, in der zukünftige Ereignisse, geordnet nach ihrem Ereigniszeitpunkt, gespeichert sind. Diese Ereignis-Liste wird auch als „Kalender" bezeichnet und stellt das Kernstück der zentralen Ablaufsteuerung dar.

Nachdem ein Ereignis aus dem Kalender entnommen wurde, müssen die dem Ereignistyp entsprechenden Abläufe ausgeführt werden. Jedes Ereignis kann außerdem weitere Folge-Ereignisse auslösen, die ebenfalls in den Kalender eingetragen werden. Bei der Bestimmung des Ereigniszeitpunkts werden die statistischen Eigenschaften zeitbehafteter Vorgänge durch die Verwendung von Pseudo-Zufallszahlen berücksichtigt.

Auch die Erzeugung des angebotenen Verkehrs basiert auf Ereignissen, welche den Ankunftszeitpunkt der Zellen angeben. Der Kalender muß daher für jede Quelle das jeweils nächste Ankunftsereignis enthalten, welches entsprechend der Charakteristik des Quellmodells vorausgeplant wird. Um die Simulation in Gang zu setzen, müssen diese Ereignisse bereits bei der Initialisierung des Systems in den Kalender aufgenommen werden.

Während der Simulation werden die gewünschten Meßwerte erfaßt und zur späteren Auswertung gespeichert. Um auch die Aussagesicherheit der gewonnenen Ergebnisse beurteilen zu können, wird jeder Simulationslauf in mehrere Teiltests unterteilt, die getrennt ausgewertet werden. Aus den Einzelergebnissen der Teiltests lassen sich anschließend sogenannte Vertrauensintervalle bestimmen, welche den Toleranzbereich angeben, in dem der tatsächliche Wert mit einer vorgegebenen Wahrscheinlichkeit liegt. Bei einer stationären Systemsimulation wird außerdem ein Warmlauf benötigt, der das System in einen eingeschwungenen Zustand bringt.

Der Rechenzeitbedarf einer Simulation hängt im wesentlichen von der Anzahl der zu verarbeitenden Ereignisse und deren Bearbeitungsaufwand ab. Bei der Untersuchung von ATM-Netzen werden häufig sehr viele Ereignisse benötigt, um die zeitliche Dynamik der Quellen hinreichend genau zu erfassen. Beispielsweise ist der Rechenzeitbedarf bei einer Simulation auf Burst- und Zellebene in etwa proportional zur mittleren Burstlänge. Darüber hinaus bereitet auch die direkte Messung von Wahrscheinlichkeiten im Bereich von 10^{-10} große Schwierigkeiten, da in diesem Fall mindestens 10^{12} Ereignisse beobachtet werden müssen, um eine brauchbare Aussagesicherheit zu erhalten.

Je nach Systemmodell können auf heute üblichen (skalaren) Rechnern bis zu etwa 10^8 Ereignisse mit vertretbarem Zeitaufwand (Rechenzeit $<$ 24 Stunden) bewältigt werden. Auch die um eine Größenordnung höhere Rechenleistung moderner Rechner mit RISC-Architektur bringt hier keinen entscheidenden Vorteil. Es wird daher häufig versucht, die Grenzen der Anwendbarkeit simulativer Verfahren durch andere Maßnahmen zu erweitern. Zu diesen Maßnahmen zählt unter anderem die parallele Simulationstechnik auf der Basis von Transputer-

Netzwerken [71],[72],[171] und anderen parallelen Rechnerarchitekturen, die bei Koppelnetzen wegen der systembedingten Parallelität durchaus effizient sein kann. Bei sehr einfachen Modellen können auch spezielle Verfahren zur Simulation seltener Ereignisse eingesetzt werden [214], die bisher bekannten Verfahren sind jedoch weder für komplexere Quellmodelle noch für Koppelnetze geeignet.

Auch wenn mit den oben genannten Methoden eine direkte Messung von Wahrscheinlichkeiten im interessierenden Bereich nicht möglich ist, können in bestimmten Fällen zumindest Schätzwerte mit Hilfe einer Extrapolation gewonnen werden [19]. Zu diesem Zweck wird eine bekannte oder vermutete Abhängigkeit der gesuchten Größe von einem Parameter ausgenutzt. Diese Abhängigkeit wird zunächst in einem Bereich mit genügend hoher Aussagesicherheit gemessen und anschließend auf den interessierenden Bereich extrapoliert. Ein typisches Beispiel ist hier die Abhängigkeit der Verlustwahrscheinlichkeit von der Puffergröße, die sich unter bestimmten Bedingungen einem geometrischen Kurvenverlauf asymptotisch annähert [122]. Insbesondere bei der gleichzeitigen Untersuchung auf Burst- und Zellebene ist jedoch Vorsicht geboten, da hier ein Abknicken der Kurven nicht ausgeschlossen werden kann [126] (siehe auch Kapitel 6).

3.3.2 Analytische Verfahren

Auf Grund der oben genannten Beschränkungen der Simulationstechnik haben analytische Verfahren insbesondere bei der Leistungsuntersuchung von ATM-Netzen große Bedeutung erlangt. Wahrscheinlichkeiten in der Größenordnung von 10^{-10} können analytisch ohne Schwierigkeiten berechnet werden, eine ausreichende numerische Stabilität der verwendeten Algorithmen vorausgesetzt. Außerdem benötigen analytische Lösungsverfahren im Normalfall wesentlich weniger Rechenzeit als eine Simulation, so daß auch umfangreiche Parameterstudien möglich sind.

Um ein Modell der verkehrstheoretischen Analyse zugänglich zu machen, sind allerdings gewisse Einschränkungen erforderlich. Beispielsweise können komplexere Abläufe wie Protokollmechanismen und Arbitrierungsstrategien meist nur näherungsweise berücksichtigt werden. Im allgemeinen lassen sich Auswahlverfahren genauso wie die Verkehrsverzweigung nur als zufallsabhängige Entscheidungen modellieren. Die beteiligten stochastischen Prozesse müssen ebenfalls bestimmte Eigenschaften aufweisen, wobei in den meisten Fällen nur gedächtnisfreie Prozesse oder Prozesse mit einer gedächtnisfreien Verteilung der Zustandsdauer analytisch behandelt werden können.

Die Analyse von ATM-Koppelnetzen auf Zellebene erfordert eine Modellierung der Pufferspeicher in den Koppelelementen und eine Berücksichtigung von internen Abläufen wie Arbitrierungsverfahren und Rückstaumechanismen. Dies führt auf komplexe zeitdiskrete Warteschlangennetze mit endlichen Pufferspeichern, die im allgemeinen nicht exakt berechnet

werden können. Um die Komplexität und die Größe des Zustandsraumes zu reduzieren, wird das Modell daher in kleinere Einheiten wie Koppelelemente oder Pufferspeicher zerlegt (Dekomposition), welche getrennt untersucht werden können. Dabei werden gewisse Annahmen bezüglich der Unabhängigkeit der einzelnen Komponenten getroffen, die in der Realität jedoch nur teilweise zutreffen. Darüber hinaus müssen auch die Koppelnetz-internen Ankunftsprozesse geeignet approximiert werden, da eine exakte Modellierung meist nicht möglich ist (siehe auch Kapitel 5). Zur Validierung dieser Näherungsannahmen wird daher in den meisten Fällen dennoch eine Simulation benötigt, wobei die Überprüfung allerdings in einem auch der Simulation zugänglichen Bereich erfolgen kann.

Das Grundmodell für die Analyse von ATM-Koppelnetzen ist ein einstufiges Bediensystem mit einer Bedieneinheit und einer endlichen Warteschlangenlänge. Zur Klassifikation derartiger Modelle wird häufig die Kendallsche Notation verwendet, welche den Typ von Ankunfts- und Bedienprozeß, die Anzahl der Bedieneinheiten sowie die Warteschlangenlänge (falls endlich) angibt. In der Kendallschen Notation besitzt dieses Modell die Kurzbezeichnung „$G^{[X]}$/G/1-S“, wobei „G“ für einen allgemeinen (General) Prozeß steht, das hochgestellte „[X]“ auf Gruppenankünfte hinweist und S die Größe des Wartespeichers *ohne* Bedieneinheit ist. Die Abfertigungsdisziplin der Warteschlange ist FIFO, sofern keine anderen Angaben gemacht werden.

Eine Besonderheit zeitdiskreter Bediensysteme ist die Behandlung von gleichzeitig auftretenden Ankunfts- und Bedienende-Ereignissen. Prinzipiell existieren zwei Möglichkeiten zur Festlegung der Reihenfolge, in der diese Ereignisse bearbeitet werden:

AF: Das Ankunftsereignis wird vor dem Bedienende bearbeitet (Arrival First, AF).
DF: Das Bedienende wird vor dem Ankunftsereignis bearbeitet. (Departure First, DF).

Bei Ankunft einer Zelle in einem AF-System befindet sich die gerade bediente Zelle noch in der Bedieneinheit, auch wenn ihre Bedienung zum gleichen Zeitpunkt endet. Dies kann zu einem Zellverlust führen, falls das System vor der Zellankunft voll belegt ist. In einem DF-System könnte die ankommende Zelle jedoch gespeichert werden, wenn gleichzeitig ein Bedienende auftritt und somit bereits ein freier Warteplatz vorhanden ist. Die Berücksichtigung der Bearbeitungsreihenfolge bei der Analyse von Bediensystemen ist in Anhang A ausführlich beschrieben.

Eine exakte Analyse von Bediensystemen mit allgemeinen Ankunfts- und Bedienprozessen ist jedoch nicht möglich, so daß weitere Einschränkungen bezüglich dieser Prozesse notwendig sind. Das allgemeinste bisher bekannte Modell dieser Art, für das geeignete Analysemethoden zur Verfügung stehen [24],[25], ist ein System des Typs „DBMAP/GI/1-S“. Der Ankunftsprozeß dieses Modells ist ein zeitdiskreter, doppelt-stochastischer Prozeß mit Gruppenankünften, der von einer Markov-Kette gesteuert wird (Discrete-Time Batch Markovian Arrival Process, DBMAP). Die Verteilung der Gruppengröße hängt dabei vom jeweiligen

Zustand der Markov-Kette ab, so daß auch korrelierte Gruppenankünfte modelliert werden können. Die Bedienzeiten besitzen keine gegenseitigen Abhängigkeiten und sind auch vom Zustand des Systems und von den Anforderungen unabhängig (General Independent, GI).

Die Analyse dieses Modells erfolgt mit den Methoden der Matrixgeometrie, die eine numerische Berechnung der Zustandswahrscheinlichkeiten sowie aller charakteristischen Größen erlauben [24],[25]. Je nach Komplexität des DBMAP kann die numerische Berechnung jedoch ebenfalls sehr aufwendig sein, da entweder sehr große Gleichungssysteme gelöst werden müssen, oder aber iterative numerische Algorithmen zum Einsatz kommen. Bei der Leistungsuntersuchung von Koppelnetzen wird daher meist auf einfachere, gedächtnisfreie Ankunftsprozesse wie den Bernoulli-Prozeß zurückgegriffen (siehe Kapitel 4).

Die Berechnung der Zustandwahrscheinlichkeiten basiert bei den oben beschriebenen Systemtypen meist auf der Methode der „Eingebetteten Markov-Kette". Dabei wird der Zustand des Systems zu speziellen Zeitpunkten betrachtet, an denen die zukünftige Entwicklung nur vom gegenwärtigen Zustand abhängig ist, nicht jedoch von der Vergangenheit. Auf Grund der Eigenschaften des Ankunftsprozesses (Gedächtnisfreiheit oder gedächtnisfreie Zustandsdauern) lassen sich die Übergangswahrscheinlichkeiten für Übergänge zwischen diesen Zuständen bestimmen, so daß die Zustandswahrscheinlichkeiten zu den „eingebetteten" Zeitpunkten berechnet werden können. Als Zeitpunkt für die Einbettung der Markov-Kette eignet sich insbesondere das Bedienende, da hier die gedächtnisbehaftete Bedienphase abgeschlossen ist. Die Einzelheiten dieser Methode sollen an dieser Stelle jedoch nicht genauer erläutert werden, für eine ausführliche Beschreibung wird auf die einschlägige Fachliteratur verwiesen [90],[121].

Bei der Behandlung von zeitdiskreten Modellen und Verteilungen wird häufig das Hilfsmittel der „Erzeugenden Funktion" verwendet. Die Erzeugende Funktion entspricht einer Transformation der Verteilung und bietet ähnliche Anwendungsmöglichkeiten wie die Laplace-Stieltjes-Transformation bei kontinuierlichen Verteilungsfunktionen. Die Erzeugende Funktion $f(z)$ der Verteilung einer diskreten Zufallsvariable X ist wie folgt definiert:

$$f(z) = \sum_{k=0}^{\infty} \mathrm{P}\{X = k\}\, z^k, \quad |z| \leq 1. \tag{3.4}$$

Eine ähnliche Transformationsmethode wird unter der Bezeichnung „z-Transformation" auch in der Systemtheorie verwendet. Die Erzeugende Funktion geht aus der z-Transformierten hervor, indem z durch z^{-1} ersetzt wird. Beide Transformationsmethoden sind daher äquivalent. Im folgenden sollen die wichtigsten Eigenschaften der Erzeugenden Funktion kurz zusammengefaßt werden.

Der Wert der Erzeugenden Funktion für $z = 1$ entspricht der Summe aller Elemente der Verteilung. Es gilt daher in jedem Fall

$$f(1) = 1. \tag{3.5}$$

Die ersten beiden gewöhnlichen Momente einer Verteilung erhält man aus den ersten beiden Ableitungen von $f(z)$ an der Stelle $z = 1$:

$$\mathrm{E}[X] = \frac{d}{dz} f(z)\Big|_{z=1} = f'(1), \tag{3.6}$$

$$\mathrm{E}[X^2] = \left(\frac{d^2}{dz^2} f(z) + \frac{d}{dz} f(z)\right)\Big|_{z=1} = f''(1) + f'(1). \tag{3.7}$$

Auch die einzelnen Elemente der Verteilung können durch wiederholtes Ableiten aus der Erzeugenden Funktion gewonnen werden:

$$\mathrm{P}\{X = k\} = \frac{1}{k!} \frac{d^k}{dz^k} f(z)\Big|_{z=0}, \quad k \geq 0. \tag{3.8}$$

Eine wichtige Eigenschaft ist die Transformation der Faltung von zwei Verteilungen in das Produkt der entsprechenden Erzeugenden Funktionen. Es seien $f(z)$ und $g(z)$ die Erzeugenden Funktionen der Verteilung von zwei unabhängigen Zufallsvariablen. Die Verteilung der Summe aus beiden Zufallsvariablen besitzt folglich die Erzeugende Funktion $f(z) \cdot g(z)$.

Weitere Eigenschaften der Erzeugenden Funktionen und deren Anwendung können beispielsweise in [121] nachgelesen werden.

Kapitel 4

Leistungsuntersuchung einstufiger ATM-Koppelanordnungen

Dieses Kapitel beschäftigt sich mit der Leistungsuntersuchung von einstufigen Koppelanordnungen auf Zellebene. Es enthält eine Beschreibung der wichtigsten Verfahren zur verkehrstheoretischen Analyse von einstufigen Anordnungen sowie einen Vergleich ihrer verkehrsabhängigen Eigenschaften. Dabei werden insbesondere die verschiedenen Möglichkeiten für die Anordnung und Organisation der Puffer untersucht und verglichen.

Für die Modellierung der Ankunftsprozesse wird in diesem und dem folgenden Kapitel ausschließlich der Bernoulli-Prozeß verwendet. Der Einfluß anderer Quellmodelle auf die Leistungsfähigkeit von Koppelnetzen wird in Kapitel 6 untersucht. Sofern keine anderen Angaben gemacht werden, soll auch eine zufällige Verzweigung des Verkehrs ohne Verkehrsschiefe vorausgesetzt werden. Diese Annahmen werden in der Literatur häufig getroffen, um eine analytische Behandlung zu ermöglichen.

Entsprechend dem getakteten Betrieb der Koppelnetze sind alle hier beschriebenen Modelle zeitdiskreter Natur. Die zugrundeliegende Zeiteinheit ΔT entspricht der Übertragungsdauer einer Zelle und wird auch als „Zeitschlitz" bezeichnet. Sofern nicht anders angegeben, wird in allen Modellen die Bearbeitungsreihenfolge DF (Departure First) verwendet.

4.1 Ausgangspuffer

Das verkehrstheoretische Modell eines Koppelelements mit Ausgangspuffern ist in Bild 4.1 dargestellt. Jeder Ausgangspuffer wird durch eine Warteschlange mit einer Bedieneinheit modelliert, wobei die Bedienzeit H konstant ist und der Übertragungsdauer ΔT einer Zelle entspricht. Die Warteschlange hat eine begrenzte Länge und kann S Zellen speichern, so daß die gesamte Puffergröße einschließlich der Bedieneinheit $S + 1$ Zellen beträgt.

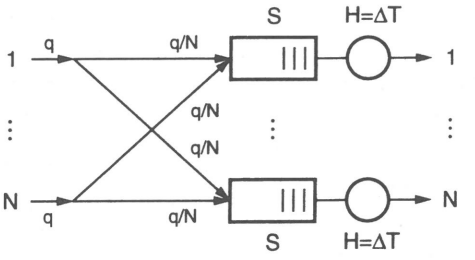


Bild 4.1: Modell eines Koppelelements mit Ausgangspuffern

Jeder an den Eingängen des Koppelelementes ankommende Zeitschlitz sei mit der Wahrscheinlichkeit q belegt. Die auf $1/\Delta T$ normierte Ankunftsrate λ der Zellen an einem Eingang ist ebenfalls gleich q. Jede ankommende Zelle ist gleichwahrscheinlich an einen der Ausgänge adressiert, so daß q/N die Wahrscheinlichkeit für das Eintreffen einer Zelle ist, die für den betrachteten Ausgangspuffer bestimmt ist. Da die Ankunftsprozesse der verschiedenen Eingänge voneinander unabhängig sind, besitzt die Anzahl Ξ der gleichzeitig an einem Ausgangspuffer eintreffenden Zellen eine Binomialverteilung. Wegen der Gedächtnisfreiheit des Bernoulli-Prozesses sind auch aufeinanderfolgende Gruppenankünfte statistisch voneinander unabhängig.

$$\xi_k = \mathrm{P}\{\Xi = k\} = \binom{N}{k} \left(\frac{q}{N}\right)^k \left(1 - \frac{q}{N}\right)^{N-k}, \quad k = 1, 2, \ldots N, \tag{4.1}$$

mit der Erzeugenden Funktion

$$\xi(z) = \sum_{k=0}^{N} \xi_k \, z^k = \left(1 - \frac{q}{N} + \frac{q}{N} z\right)^N. \tag{4.2}$$

Die n-te Ableitung von $\xi(z)$ kann mit Hilfe der folgenden Beziehung berechnet werden:

$$\frac{d^n}{dz^n}\xi(z) = \prod_{i=0}^{n-1}(N-i)\left(\frac{q}{N}\right)^n \left(1 - \frac{q}{N} + \frac{q}{N} z\right)^{N-n}, \quad 1 \le n \le N. \tag{4.3}$$

Daraus ergibt sich

$$\begin{aligned} \xi'(1) &= q, \\ \xi''(1) &= \frac{N-1}{N} q^2, \\ \xi'''(1) &= \frac{(N-1)(N-2)}{N^2} q^3. \end{aligned}$$

Die mittlere Gruppengröße ist durch $\mathrm{E}[\Xi] = q$ gegeben. Wegen der konstanten Bedienzeit $H = \Delta T$ ist das Verkehrsangebot A des Ausgangspuffers hier gleich der Ankunftswahrscheinlichkeit q. Da q eine Wahrscheinlichkeit ist, liegt das Angebot A zwischen 0 und 1.

Unter den oben angegebenen Voraussetzungen kann die Verkehrsleistung von Koppelelementen mit Ausgangspuffern durch eine exakte verkehrstheoretische Analyse bestimmt werden [92],[96],[111]. Wegen der symmetrischen Belastung und der zufälligen Verzweigung der Zellen sind die Verkehrseigenschaften aller Puffer in einem Koppelelement identisch, so daß die Untersuchung auf einen einzelnen Ausgangspuffer beschränkt werden kann.

Das Modell eines Ausgangspuffers ist ein Sonderfall des in Anhang A untersuchten Systems, welches eine beliebige Verteilung der Bedienzeit erlaubt. Bei unbegrenzter Warteschlangenlänge kann eine geschlossene Lösung für die Zustandsverteilung sowie die Verteilung der Wartezeit in Form von Erzeugenden Funktionen angegeben werden (dabei muß allerdings $A < 1$ vorausgesetzt werden, da der Ausgangspuffer sonst instabil ist). Es sei Y der Zustand des Systems aus der Sicht eines außenstehenden Beobachters und W die Wartezeit einer beliebigen Zelle. Aus den Gleichungen A.22 und A.34 des Anhangs A ergibt sich für den Sonderfall der konstanten Bedienzeit und bei FIFO-Abfertigung

$$y(z) = (1-A)\,(z-1)\,\frac{\xi(z)}{z-\xi(z)}, \tag{4.4}$$

$$w(z) = \frac{1-A}{\lambda}\,\frac{\xi(z)-1}{z-\xi(z)}, \tag{4.5}$$

wobei $y(z)$ bzw. $w(z)$ die Erzeugende Funktion der Verteilung von Y bzw. W ist. Daraus erhält man beispielsweise die mittlere Wartezeit $\mathrm{E}[W]$ zu

$$\mathrm{E}[W] = w'(1) = \frac{\xi''(1)}{2q(1-q)} = \frac{N-1}{N}\,\frac{A}{2(1-A)}, \quad A<1. \tag{4.6}$$

Ein Vergleich mit der mittleren Wartezeit $\mathrm{E}[W_{M/D/1}]$ in einem M/D/1-System weist auf einen interessanten Zusammenhang zwischen beiden Modellen hin:

$$\mathrm{E}[W_{M/D/1}] = \frac{A}{2(1-A)}, \quad A<1,$$

$$\Rightarrow \quad \mathrm{E}[W] = \frac{N-1}{N}\,\mathrm{E}[W_{M/D/1}].$$

Für $N \to \infty$ konvergiert die mittlere Wartezeit also gegen die mittlere Wartezeit des M/D/1-Systems. Dies hängt damit zusammen, daß die Binomialverteilung für große N gegen die Poisson-Verteilung konvergiert:

$$\lim_{N\to\infty} \binom{N}{k} \left(\frac{q}{N}\right)^k \left(1-\frac{q}{N}\right)^{N-k} = \mathrm{e}^{-q}\,\frac{q^k}{k!}.$$

Für die Berechnung des zweiten Moments der Wartezeit wird auch die zweite Ableitung von $w(z)$ an der Stelle $z=1$ benötigt:

$$w''(1) = \frac{N-1}{N}\,\frac{N-2}{N}\,\frac{q^2}{3(1-q)} + \left(\frac{N-1}{N}\right)^2 \frac{q^3}{2(1-q)^2}, \quad q<1. \tag{4.7}$$

Die mittlere Durchlaufzeit der Zellen (Summe aus Wartezeit und Bedienzeit) als Funktion des Angebots ist in Bild 4.2 für unterschiedliche Werte von N dargestellt. Bis zu einem Angebot von etwa 80% beträgt die mittlere Durchlaufzeit nur wenige Zeitschlitze, erst bei einem höheren Angebot steigen die Werte sehr schnell an. Die Abhängigkeit der mittleren Durchlaufzeit von der Größe des Koppelelements ist dagegen nicht sehr ausgeprägt, so daß der Grenzfall $N \rightarrow \infty$ bereits ab $N = 16$ eine sehr gute Näherung darstellt.

Die Berechnung der Verlustwahrscheinlichkeit in einem System mit endlicher Puffergröße $(S+1)$ ist etwas aufwendiger, da in diesem Fall keine geschlossene Lösung in Form von Erzeugenden Funktionen mehr möglich ist. Die Zustandswahrscheinlichkeiten können jedoch mit Hilfe von rekursiven numerischen Algorithmen exakt berechnet werden (siehe Anhang A.3). Wegen der konstanten Bedienzeit $H = \Delta T$ erhält man auch die Wartezeitverteilung sehr einfach aus der Zustandsverteilung bei Ankunft einer Anforderung.

Die folgenden Ergebnisse basieren auf der Annahme, daß die im gleichen Zeitschlitz eintreffenden Zellen in zufälliger Reihenfolge gespeichert werden (der Einfluß einer festen Reihenfolge wird später untersucht). In diesem Fall erhält man die in Bild 4.3 dargestellte Abhängigkeit der Zellverlustwahrscheinlichkeit von der Puffergröße. Es fällt zunächst der nahezu lineare Verlauf der Kurven auf, der wegen der logarithmischen Darstellung einer geometrisch abklingenden Asymptote entspricht. Die Verlustwahrscheinlichkeit wird also durch jeden zusätzlichen Pufferplatz um einen (näherungsweise) konstanten Faktor vermindert. Darüber hinaus zeigt sich hier ein wesentlich stärkerer Einfluß der Größe des Koppelelements (N) auf die Verlustwahrscheinlichkeit. Um bei einem Angebot von 80% eine bestimmte Verlustwahr-

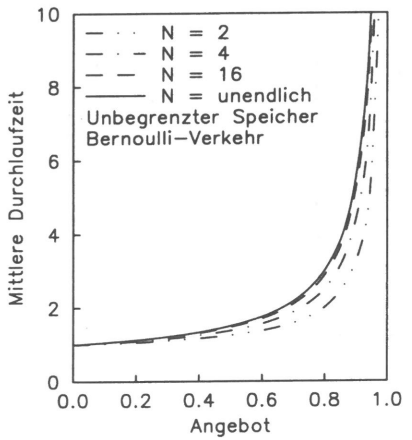


Bild 4.2: Mittlere Durchlaufzeit bei unbegrenzter Warteschlangenlänge

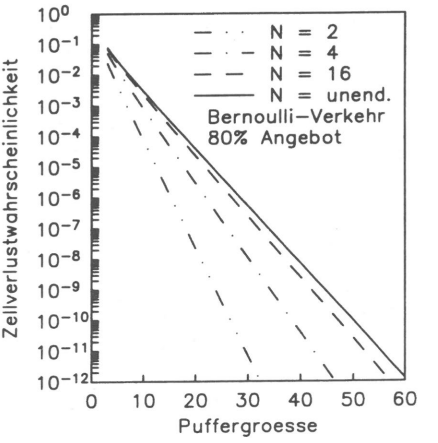


Bild 4.3: Verlustwahrscheinlichkeit in Abhängigkeit der Puffergröße bei 80% Angebot

scheinlichkeit zu erreichen, wird für $N \to \infty$ nahezu die doppelte Puffergröße benötigt wie für $N = 2$. Diese Zunahme des Pufferbedarfs hängt mit der Varianz der Gruppengröße Ξ zusammen, die mit zunehmendem N ebenfalls ansteigt:

$$\mathrm{VAR}[\Xi] = q(1 - \frac{q}{N}) \tag{4.8}$$

Ähnlich wie bei der mittleren Durchlaufzeit werden die Kurven auch hier durch den Fall $N \to \infty$ nach oben hin begrenzt.

Aus dem Zusammenhang zwischen Puffergröße, Angebot und Verlustwahrscheinlichkeit läßt sich diejenige Puffergröße ermitteln, die für das Erreichen einer vorgegebenen Verlustwahrscheinlichkeit mindestens benötigt wird. In Bild 4.4 ist die Abhängigkeit dieser Puffergröße vom Angebot anhand eines 16×16-Koppelelements dargestellt. Aus Bild 4.4 kann beispielsweise entnommen werden, daß bei einem Angebot von 80% und einer Verlustwahrscheinlichkeit von 10^{-10} mindestens 48 Pufferplätze vorgesehen werden müssen. Bei einer höheren Auslastung ist auch hier ein starker Anstieg der Kurven zu beobachten, wobei sich die erforderliche Puffergröße bei einer Erhöhung des Angebots von 80% auf 90% bereits verdoppelt. Aus diesem Grund wird die Auslastung der Puffer meist auf Werte um 80% begrenzt.

Die bisherigen Ergebnisse basieren auf der Annahme, daß gleichzeitig eintreffende Zellen in zufälliger Reihenfolge gespeichert werden. Mit Hilfe von Simulationsergebnissen konnte bestätigt werden, daß dies auch für den Fall der zyklischen Priorisierung eine sehr gute Näherung darstellt. Im folgenden soll nun der Fall untersucht werden, daß die Eingänge in einer

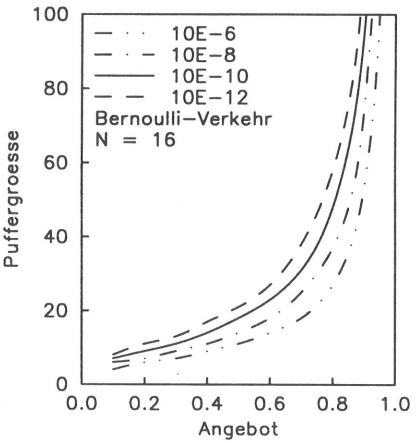


Bild 4.4: Dimensionierung der Puffergröße bei vorgegebener Verlustwahrscheinlichkeit

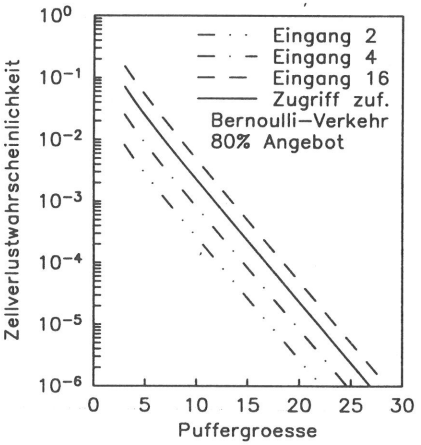


Bild 4.5: Verlustwahrscheinlichkeit bei fester Zugriffsreihenfolge der Eingänge ($N = 16$)

festen Reihenfolge auf die Ausgangspuffer zugreifen. Ohne Einschränkung der Allgemeinheit wird dabei angenommen, daß die Zugriffsreihenfolge mit der Numerierung der Eingänge übereinstimmt. Eine an Eingang i ankommende Zelle findet daher bereits alle gleichzeitig an den Eingängen 1 bis $i-1$ eingetroffenen Zellen mit gleichem Ziel im Puffer vor. Es sei Ψ^i die Anzahl der Zellen mit gleichem Ziel, die in einem Zeitschlitz an den Eingängen 1 bis $i-1$ eintreffen und daher vor der betrachteten Zelle gespeichert werden. Die Verteilung von Ψ^i ist ebenfalls eine Binomialverteilung:

$$\psi_k^i = \mathrm{P}\{\Psi^i = k\} = \binom{i-1}{k} \left(\frac{q}{N}\right)^k \left(1-\frac{q}{N}\right)^{i-1-k}, \quad 0 \le k \le i-1, \tag{4.9}$$

mit der Erzeugenden Funktion

$$\psi^i(z) = \sum_{k=0}^{i-1} \psi_k^i \, z^k = \left(1 - \frac{q}{N} + \frac{q}{N} z\right)^{i-1}. \tag{4.10}$$

Die Zugriffsreihenfolge beeinflußt nur die Verlustwahrscheinlichkeit sowie die Verteilung der Wartezeit bezogen auf die einzelnen Eingänge, die Zustandswahrscheinlichkeiten bleiben jedoch unverändert. Aus diesem Grund kann die Wahrscheinlichkeit z_k^*, daß der Puffer vor dem Eintreffen der gesamten Gruppenankunft k Zellen enthält, wie in Anhang A.3 angegeben berechnet werden. Analog zu Gleichung A.54 erhält man daraus die Wahrscheinlichkeit z_k^i, daß der Puffer unmittelbar vor dem Speichern einer Zelle von Eingang i bereits k Zellen enthält:

$$z_k^i = \begin{cases} \displaystyle\sum_{\nu=0}^{k} z_\nu^* \psi_{k-\nu}^i, & 0 \le k \le S, \\ \displaystyle\sum_{\nu=0}^{S+1} z_\nu^* \sum_{j=S+1-\nu}^{\infty} \psi_j^i, & k = S+1. \end{cases} \tag{4.11}$$

Die Verlustwahrscheinlichkeit einer vom Eingang i kommenden Zelle entspricht der Wahrscheinlichkeit z_{S+1}^i, daß der Puffer vor dem Eintreffen der Zelle bereits voll ist.

Bild 4.5 zeigt den Einfluß der Zugriffsreihenfolge auf die Verlustwahrscheinlichkeit bezogen auf einzelne Eingänge des Koppelelements, wobei zum Vergleich auch die Verlustwahrscheinlichkeit bei zufälligem Zugriff (durchgezogene Kurve) eingetragen ist. Mit den gewählten Parametern ist die Verlustwahrscheinlichkeit für Eingang 16 etwa um eine Größenordnung höher als für Eingang 2. Wegen der Bearbeitungsreihenfolge DF finden die an Eingang 1 eintreffenden Zellen immer einen freien Pufferplatz vor, so daß hier keine Verluste auftreten können.

Die Ergebnisse machen deutlich, daß der geringere Implementierungsaufwand einer festen Zugriffsreihenfolge mit einem höheren Speicherplatzbedarf bezahlt werden muß. Auch unter dem Gesichtspunkt der Fairneß sollte daher eine zyklisch wechselnde Zugriffsreihenfolge bevorzugt werden.

4.2 Zentralspeicher

Eine entsprechende Dimensionierung der Speichergröße vorausgesetzt, unterscheiden sich Koppelelemente mit Zentralspeicher bezüglich der Durchlaufzeit nicht von Koppelelementen mit Ausgangspuffern. Hier steht vielmehr die Untersuchung der Speicherbelegung im Vordergrund, wobei insbesondere der Einfluß der Speicherverwaltung berücksichtigt werden muß.

4.2.1 Analyse der Speicherbelegung

Für einfache Verfahren zur Verwaltung des zentralen Speichers kann eine approximative Analyse der Speicherbelegung durchgeführt werden. In diesem Abschnitt wird ein analytischer Ansatz vorgestellt, der sowohl bei einer Verwaltung nach dem „Complete Sharing"-Prinzip als auch bei einer Begrenzung der logischen Ausgangswarteschlangen anwendbar ist. Ein simulativer Vergleich mit weiteren Verfahren zur Speicherverwaltung erfolgt im nächsten Abschnitt.

Den Ausgangspunkt für die Bestimmung der Verteilung der Speicherbelegung bildet eine Analyse der logischen Ausgangswarteschlangen. Die Zustandswahrscheinlichkeiten der Ausgangswarteschlangen können mit den oben angegebenen Methoden zur Untersuchung von Koppelelementen mit Ausgangspuffern berechnet werden. Dabei wird zunächst davon ausgegangen, daß der Zentralspeicher nicht begrenzt ist, so daß Zellverluste nur an den logischen Ausgangswarteschlangen auftreten können.

Der Zustand des Zentralspeichers ist durch die Anzahl der insgesamt gespeicherten Zellen gegeben. Diese Anzahl entspricht auch der Summe aller Zellen in den logischen Ausgangswarteschlangen. Eine exakte Analyse erfordert die Bestimmung der gemeinsamen Zustandsverteilung aller Ausgangswarteschlangen und führt auf einen exponentiell mit N wachsenden Zustandsraum. In den meisten Fällen scheidet eine exakte Analyse daher aus Gründen der Komplexität aus.

Um eine approximative Bestimmung der Speicherbelegung zu ermöglichen, wird üblicherweise von einer Unabhängigkeit der einzelnen Ausgangswarteschlangen ausgegangen [96],[137]. In der Realität existieren jedoch deutliche Abhängigkeiten, die auf eine negative Korrelation zwischen den Ankünften der einzelnen Ausgangswarteschlangen zurückzuführen sind [73],[96]. Falls k der eintreffenden Zellen für einen Ausgang bestimmt sind, können nur noch $N - k$ Eingänge zu den Ankünften der anderen Ausgänge beitragen. Mit steigendem k sinkt daher die Anzahl der Ankünfte an den anderen Ausgangswarteschlangen, woraus sich, insbesondere für kleine N, eine negative Korrelation der Ankünfte und damit auch der Warteschlangenlängen ergibt.

Es seien Ξ_1 und Ξ_2 die Anzahl der Ankünfte an zwei beliebigen Ausgangswarteschlangen. Nach [73] ist die Erzeugende Funktion $\xi(z_1, z_2)$ der Verteilung $\mathrm{P}\{\Xi_1 = k_1, \Xi_2 = k_2\}$

$$\xi(z_1, z_2) = \left(1 - \frac{2q}{N} + \frac{q}{N} z_1 + \frac{q}{N} z_2\right)^N .$$

Daraus folgt für den Mittelwert des Produkts $\Xi_1 \Xi_2$

$$\mathrm{E}[\Xi_1 \Xi_2] = \frac{\partial^2}{\partial z_1 \partial z_2} f(z_1, z_2)|_{z_1=1, z_2=1} = \frac{N-1}{N} q^2,$$

und für die Kovarianz von Ξ_1 und Ξ_2

$$\mathrm{COV}[\Xi_1, \Xi_2] = \mathrm{E}[\Xi_1 \Xi_2] - \mathrm{E}[\Xi_1]\ \mathrm{E}[\Xi_2] = -\frac{q^2}{N}.$$

Mit Hilfe von Gleichung 4.8 für die Varianz von Ξ erhält man folgende Beziehung für den Korrelationskoeffizienten von Ξ_1 und Ξ_2:

$$\mathrm{CORR}[\Xi_1, \Xi_2] = \frac{\mathrm{COV}[\Xi_1, \Xi_2]}{\sqrt{\mathrm{VAR}[\Xi_1]\ \mathrm{VAR}[\Xi_2]}} = -\frac{q}{N-q}. \tag{4.12}$$

Der Korrelationskoeffizient hängt also sowohl von der Ankunftswahrscheinlichkeit q als auch von der Größe des Koppelelements N ab. Für $N \to \infty$ geht der Korrelationskoeffizient gegen 0, so daß auch die Abhängigkeiten zwischen den logischen Ausgangswarteschlangen verschwinden.

Falls diese Abhängigkeiten vernachlässigt werden, können die Zustandswahrscheinlichkeiten des Zentralspeichers aus einer $(N-1)$-fachen Faltung der Zustandsverteilungen der logischen Ausgangswarteschlangen berechnet werden [96],[137]. Wegen der negativen Korrelation führt diese Unabhängigkeitsannahme zu einer Abschätzung der Zustandsverteilung nach oben hin, so daß der benötigte Speicherplatz stets überschätzt wird ([73], siehe auch Bild 4.6). Der Fehler dieser Approximation ist umso kleiner, je geringer die Korrelationen zwischen den einzelnen Ausgängen sind. Für den Grenzfall $N \to \infty$ ist der Ansatz exakt, da die Ankünfte der einzelnen Ausgänge in diesem Fall unkorreliert sind.

In [73] wurde eine verbesserte Approximation vorgeschlagen, die auf einer Darstellung der Zustandsverteilung des Zentralspeichers als Gamma-Verteilung mit gleichem Mittelwert und gleicher Varianz basiert. Zur Bestimmung der Varianz wird die Kovarianz des Zustands von zwei beliebigen Ausgangswarteschlangen benötigt, die aus der gemeinsamen Zustandsverteilung beider Warteschlangen berechnet werden kann. Die in [73] angegebenen Ergebnisse zeigen eine gute Übereinstimmung mit den entsprechenden Simulationswerten.

Das im folgenden beschriebene Analyseverfahren basiert ebenfalls auf einer Unabhängigkeitsannahme bezüglich des Zustands der Ausgangswarteschlangen. Durch eine geeignete Wahl des Beobachtungszeitpunkts können die Korrelationen zwischen den Ankünften jedoch teilweise berücksichtigt werden. Zu diesem Zweck wird zunächst die Zustandsverteilung des

Zentralspeichers unmittelbar nach einem Bedienende bestimmt. Diese mit x_k^{ZS} bezeichnete Verteilung erhält man (näherungsweise) aus einer $(N-1)$-fachen Faltung der entsprechenden Zustandsverteilung x_i einer logischen Ausgangswarteschlange:

$$x_k^{ZS} = [\underbrace{x_i * x_i * \ldots * x_i}_{N-1 \text{ Faltungen}}]_k = [x_i^{(N)}]_k. \tag{4.13}$$

Dabei steht das Symbol „$*$“ für die Faltungsoperation, und $[x_i^{(N)}]_k$ bezeichnet das k-te Element der $(N-1)$-fachen Faltung der Verteilung von X mit sich selbst.

Aus der Verteilung x_k^{ZS} läßt sich die Zustandsverteilung des Zentralspeichers aus der Sicht einer ankommenden Zelle berechnen. Dazu werden die gleichzeitig an verschiedenen Eingängen eintreffenden Zellen als eine Gruppenankunft aufgefaßt, die zwischen 0 und N Anforderungen enthalten kann. Die Anzahl der Zellen in dieser Gruppenankunft wird mit Ξ^{ZS} bezeichnet und besitzt die folgende Binomialverteilung:

$$\xi_k^{ZS} = \mathrm{P}\{\Xi^{ZS} = k\} = \binom{N}{k} q^k \, (1-q)^{N-k}. \tag{4.14}$$

Die Zustandsverteilung z_k^{ZS*} des Zentralspeichers unmittelbar vor dem Eintreffen der Gruppenankunft kann durch eine ähnliche Überlegung wie in Anhang A.2.2 bestimmt werden. Für die Bearbeitungsreihenfolge DF erhält man

$$z_k^{ZS*} = \begin{cases} \dfrac{x_0^{ZS}}{\mathrm{E}[D]} \, \dfrac{1}{1-\xi_0^{ZS}}, & k = 0, \\ \dfrac{x_k^{ZS}}{\mathrm{E}[D]}, & k > 0, \end{cases} \tag{4.15}$$

mit

$$\mathrm{E}[D] = 1 + \frac{x_0^{ZS} \xi_0^{ZS}}{1-\xi_0^{ZS}}.$$

Zum Zeitpunkt des Speicherns einer Zelle können sich einige der gleichzeitig eingetroffenen Zellen bereits im Speicher befinden. Unter der Annahme, daß gleichzeitig ankommende Zellen in zufälliger Reihenfolge gespeichert werden, ergibt sich mit Hilfe von Gleichung A.24 folgende Beziehung für die Wahrscheinlichkeit ψ_k^{ZS}, daß bereits k der gleichzeitig ankommenden Zellen gespeichert sind:

$$\psi_k^{ZS} = \frac{1}{\lambda} \, (1 - \mathrm{P}\{\Xi^{ZS} \leq k\}). \tag{4.16}$$

Der Systemzustand Z^{ZS} aus der Sicht einer ankommenden Zelle ist die Summe aus Z^{ZS*} und der Anzahl derjenigen Zellen, die gleichzeitig eintreffen und bereits gespeichert sind. Da die Ankünfte vom Zustand des Systems unabhängig sind, kann die Verteilung von Z^{ZS} durch eine Faltung der entsprechenden Verteilungen berechnet werden.

$$z_k^{ZS} = \sum_{i=0}^{k} \psi_i^{ZS} \, z_{k-i}^{ZS*} \tag{4.17}$$

Bei einem begrenzten Zentralspeicher kann die Verlustwahrscheinlichkeit näherungsweise aus der komplementären Verteilungsfunktion von Z^{ZS} bestimmt werden. In einem Zentralspeicher der Größe S^{ZS} Zellen ist die Wahrscheinlichkeit für einen Zellverlust näherungsweise gleich der Wahrscheinlichkeit, daß eine ankommende Zelle mehr als $S^{ZS} - 1$ Zellen in einem Koppelelement mit unbegrenztem Speicher vorfindet.

$$P^{ZS}_{Verlust} = P\{\text{eine ankommende Zelle kann nicht gespeichert werden}\} \approx P\{Z^{ZS} > S^{ZS} - 1\}$$

Falls die Länge der logischen Ausgangswarteschlangen ebenfalls begrenzt ist, können auch hier Zellverluste auftreten. Die Wahrscheinlichkeit $P^{A}_{Verlust}$ für einen Zellverlust auf Grund einer Begrenzung der Ausgangswarteschlangen kann, unter Vernachlässigung von $P^{ZS}_{Verlust}$, wie bei einem Koppelelement mit Ausgangspuffern berechnet werden. Insgesamt ergibt sich damit für die Zellverlustwahrscheinlichkeit

$$P\{\text{eine ankommende Zelle geht verloren}\} = P^{ZS}_{Verlust} + (1 - P^{ZS}_{Verlust})\, P^{A}_{Verlust}$$

Bild 4.6 zeigt einen Vergleich der Analyseergebnisse mit den entsprechenden Simulationsdaten für ein 2×2- und ein 16×16-Koppelelement bei einem Angebot von $A = 0.8$. Die Verlustwahrscheinlichkeit ist hier über der auf die Anzahl der Ausgänge normierten Puffergröße aufgetragen. Deutlich ist die Abweichung zwischen Simulation und Analyse zu erkennen, die auf die Vernachlässigung der negativen Korrelation zwischen den einzelnen Ausgangswarteschlangen zurückzuführen ist. Die Analyse liefert eine konservative Abschätzung des Pufferbedarfs, wobei die erforderliche Speichergröße je nach Wahl der Parameter um bis zu 30% überschätzt wird. Die Ergebnisse zeigen außerdem, daß ein 16×16-Koppelelement bei Verlustwahrscheinlichkeiten kleiner als 10^{-2} wesentlich weniger Speicherplatz pro Ausgang benötigt als ein 2×2-Koppelelement. Dieser Effekt hängt damit zusammen, daß ein Aushelfen zwischen den einzelnen Ausgangswarteschlangen um so effektiver ist, je mehr Ausgänge vorhanden sind („Bündelungsgewinn").

In Bild 4.7 ist der Einfluß der Anzahl der Ausgänge N auf die benötigte Speichergröße (normiert pro Ausgang) bei einer Verlustwahrscheinlichkeit von 10^{-4} dargestellt. Auch hier ist die Differenz zwischen Analyse und Simulation deutlich zu erkennen. Überraschenderweise nimmt diese Differenz mit steigendem N nur sehr langsam ab, obwohl die Korrelation zwischen den Ankünften asymptotisch mit $1/N$ abfällt. Die Abnahme der Korrelation wird offensichtlich durch die höhere Anzahl der Faltungen teilweise kompensiert, so daß der Fehler insgesamt nahezu konstant bleibt. Eine Annäherung der analytischen Approximation an die Simulationsergebnisse ist erst für Werte von $N > 100$ zu beobachten [96].

Aus Bild 4.7 kann beispielsweise abgelesen werden, daß für $N = 16$ und bei einem Angebot von $A = 0.8$ bereits vier Pufferplätze pro Ausgang genügen, um eine Zellverlustwahrscheinlichkeit von 10^{-4} zu erreichen. Ein Koppelelement mit Ausgangspuffern benötigt bei gleicher Verlustwahrscheinlichkeit 17 Pufferplätze je Ausgang, der Speicherplatzbedarf des Zentralspeichers ist in diesem Fall also etwa um den Faktor 4 geringer. Wegen des steilen Abfalls der

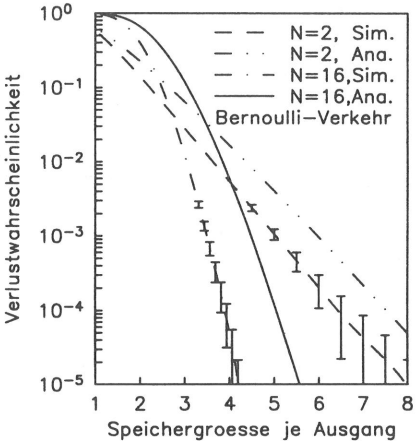


Bild 4.6: Verlustwahrscheinlichkeit über der Größe des Zentralspeichers (pro Ausgang) bei einem Angebot von 0.8

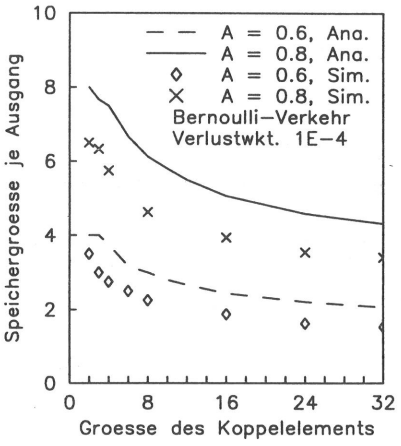


Bild 4.7: Dimensionierung der Speichergröße (pro Ausgang) am Beispiel einer Verlustwahrscheinlichkeit von 10^{-4}

Kurven in Bild 4.6 ist die Einsparung bei kleineren Verlustwahrscheinlichkeiten sogar noch höher (siehe auch Abschnitt 4.6).

Mit steigendem N konvergiert die erforderliche Speichergröße gegen die mittlere Warteschlangenlänge einer Ausgangswarteschlange [96], die bei einem Angebot von $A = 0.8$ beispielsweise 2.4 Zellen beträgt. Diese untere Grenze wird jedoch erst für sehr große N erreicht; für $N \leq 32$ ist der Speicherplatzbedarf deutlich höher.

4.2.2 Speicherverwaltung

Die Aufgabe der Speicherverwaltung besteht darin, eine Überflutung des Zentralspeichers durch einzelne Ausgangswarteschlangen zu verhindern und gleichzeitig eine möglichst effektive Ausnutzung des Speicherplatzes zu gewährleisten. Im folgenden soll die Leistungsfähigkeit der in Abschnitt 2.3.1.2 beschriebenen Strategien zur Verwaltung des Zentralspeichers untersucht werden.

Die Begrenzung der logischen Ausgangswarteschlangen ist ein sehr einfaches Mittel, um die gegenseitige Beeinflussung verschiedener Ausgangswarteschlangen zu reduzieren. Da sich diese Begrenzung auch auf die Verlustwahrscheinlichkeit bei symmetrischem Verkehr auswirkt, ist eine sorgfältige Dimensionierung der Ausgangswarteschlangen notwendig. Der Zusammenhang zwischen der Zellverlustwahrscheinlichkeit und der Länge der logischen Ausgangswarteschlangen S^A ist in Bild 4.8 dargestellt. Für kleine Werte von S^A sind die Verluste

im wesentlichen auf die Begrenzung der Ausgangswarteschlangen zurückzuführen, so daß die Verlustwahrscheinlichkeit mit steigendem S^A zunächst abnimmt. Für große Werte von S^A überwiegen dagegen die Verluste am Zentralspeicher selbst, die von der Größe des Speichers abhängig sind. Der Einfluß der Speichergröße ist in Bild 4.8 deutlich zu erkennen, die dargestellten Kurven beziehen sich auf einen Zentralspeicher von 60, 70 und 80 Zellen. Da die gezeigten Ergebnisse mit Hilfe der oben beschriebenen approximativen Analyse ermittelt wurden, ist die Verlustwahrscheinlichkeit des Zentralspeichers in der Realität etwas geringer. Der prinzipielle Verlauf der Kurven wird von der Analyse jedoch korrekt wiedergegeben.

In [137] wurde bereits darauf hingewiesen, daß die Verlustwahrscheinlichkeit durch eine geeignete Wahl von S^A auch bei symmetrischer Last minimiert werden kann. Die Kurve für $S^{ZS} = 60$ zeigt beispielsweise ein deutliches Minimum an der Stelle $S^A = 10$. Bei einer Verlustwahrscheinlichkeit von 10^{-10} ist dieser Effekt jedoch nicht mehr von Bedeutung, da das Minimum mit zunehmender Größe des Zentralspeichers immer schwächer wird. Das Minimum kann dennoch für die Dimensionierung der logischen Ausgangswarteschlangen verwendet werden, da die Verlustwahrscheinlichkeit für kleinere Werte von S^A sehr schnell ansteigt, und die gegenseitige Beeinflussung der Ausgangswarteschlangen für größere Werte von S^A ebenfalls zunimmt.

Um die Leistungsfähigkeit verschiedener Verfahren zur Speicherverwaltung beurteilen zu können, wird deren Verhalten bei der gezielten Überlastung eines einzelnen Ausgangs untersucht. Das Verkehrsangebot an diesen Ausgang, der auch als „Hot Spot" bezeichnet wird,

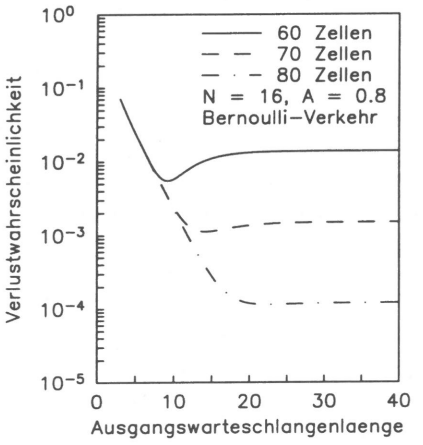


Bild 4.8: Einfluß von begrenzten logischen Ausgangswarteschlangen auf die Verlustwahrscheinlichkeit (Analyseergebnisse)

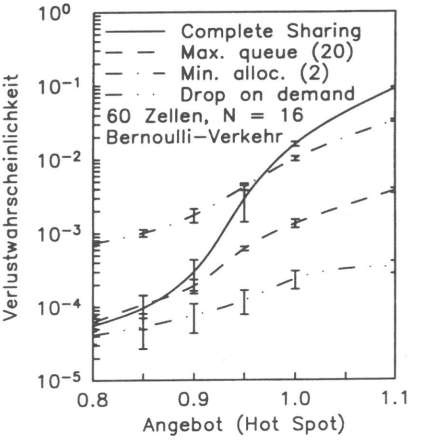


Bild 4.9: Auswirkung der „Hot Spot"-Belastung auf die Zellverluste unbeteiligter Ausgänge (Simulation)

besteht aus der symmetrischen Grundlast und einem zusätzlichen Verkehrsanteil, der von allen Eingängen ausschließlich an diesen Ausgang gesendet wird. Alle anderen Ausgänge werden nur mit der symmetrischen Grundlast beaufschlagt. Der Einfluß einer Überlastung des „Hot Spots“ auf die Zellverluste der anderen (unbeteiligten) Ausgänge ist in Bild 4.9 am Beispiel eines 16×16-Koppelelements mit einer Speichergröße von 60 Zellen dargestellt. Die Grundlast beträgt $A = 0.8$, so daß bei einem „Hot Spot“-Angebot von 0.8 alle Ausgänge gleich belastet sind.

Wie erwartet steigt bei einer Überlastung des „Hot Spots“ auch die Verlustwahrscheinlichkeit der unbeteiligten Ausgänge, wobei deutliche Unterschiede zwischen den einzelnen Verfahren zur Speicherverwaltung erkennbar sind. Der stärkste Anstieg ist bei einer Verwaltung nach dem „Complete Sharing“-Prinzip zu verzeichnen, da hier jeder Ausgang den gesamten Speicher überfluten kann. Ein nahezu optimales Verhalten zeigt dagegen die Strategie „Drop on Demand“, bei der jeweils eine Zelle aus der längsten Ausgangswarteschlange verworfen wird. Diese Strategie weist die geringste Empfindlichkeit gegenüber einer „Hot Spot“-Überlast auf und reduziert auch die Verlustwahrscheinlichkeit bei symmetrischem Verkehr [215].

Die Reservierung von Teilen des Speichers (Minimum Allocation) ohne eine Vergrößerung des Speicherplatzes hat bereits bei symmetrischer Last eine höhere Verlustwahrscheinlichkeit zur Folge. Auch wenn, wie in Bild 4.9 angenommen, für jeden Ausgang lediglich 2 Speicherplätze reserviert werden, reduziert sich dadurch der von allen Ausgängen nutzbare Speicherbereich um mehr als die Hälfte. Außerdem kann der gemeinsam genutzte Teil des Speichers weiterhin von einem Ausgang überflutet werden, so daß die Verluste der anderen Ausgänge ebenfalls stark ansteigen.

Die Begrenzung der logischen Ausgangswarteschlangen (Maximum Queue) stellt einen sehr guten Kompromiß zwischen Aufwand und Leistung dar. Diese Strategie kommt ohne jeglichen Zusatzaufwand aus und bringt dennoch eine deutliche Verbesserung gegenüber dem „Complete Sharing“-Prinzip. Wegen des hohen Aufwands zur Implementierung des „Drop on Demand“-Verfahrens erscheint die Begrenzung der Ausgangswarteschlangen daher am besten für den Einsatz in ATM-Koppelelementen geeignet.

4.3 Eingangspuffer

Eine analytische Untersuchung von Koppelelementen mit Eingangspuffern ist wesentlich schwieriger als bei Koppelelementen mit Ausgangspuffern. Die exakte Berücksichtigung des Arbitrierungsvorgangs führt meist zu sehr komplexen Modellen, deren Zustandsraum nicht mehr beherrschbar ist. Eine Analyse ist daher nur für wenige Sonderfälle möglich, im allgemeinen muß auf Simulationsergebnisse zurückgegriffen werden.

4.3.1 Eingangspuffer mit FIFO-Abfertigung

Das Arbitrierungsverfahren in einem Koppelelement mit FIFO-Eingangspuffern und einfachem Lesezugriff berücksichtigt nur die an vorderster Stelle der Warteschlange (Head of Line, HOL) stehenden Zellen. Falls k dieser Zellen an denselben Ausgang adressiert sind, wird entsprechend der Arbitrierungsstrategie eine der HOL-Zellen ausgewählt, und die verbleibenden $k-1$ Zellen warten bis zum nächsten Zeitschlitz. Die um einen Ausgang konkurrierenden HOL-Zellen bilden eine virtuelle Ausgangswarteschlange (siehe Bild 4.10), aus der in jedem Zeitschlitz höchstens eine Zelle entnommen und zum entsprechenden Ausgang transportiert werden kann. Die Verweildauer einer Zelle in dieser virtuellen Ausgangswarteschlange kann gleichzeitig als Verweildauer in der Bedieneinheit des Eingangspuffers aufgefaßt werden und entspricht somit der Bedienzeit H.

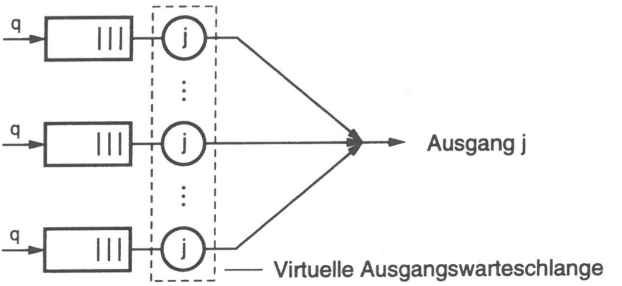


Bild 4.10: Virtuelle Ausgangswarteschlange in einem Koppelelement mit Eingangspuffern

Das Modell eines Eingangspuffers besteht also aus einer FIFO-Warteschlange mit einer Bedieneinheit, wobei die Bedienzeit der Zellen durch ihre Verweildauer in einer weiteren Warteschlange gegeben ist. Bei der Analyse des Eingangspuffers muß daher auch die virtuelle Ausgangswarteschlange untersucht werden, um die Verteilung der Bedienzeit $h_k = \mathrm{P}\{H = k\Delta T\}$ zu bestimmen. Zunächst soll jedoch der Eingangspuffer selbst betrachtet werden.

Unter der Voraussetzung, daß der Ankunftsprozeß des Eingangspuffers ein Bernoulli-Prozeß ist, können die in Anhang A gewonnenen Ergebnisse auch zur Analyse des Eingangspuffers herangezogen werden. Bei einem Bernoulli-Prozeß gilt für die Verteilung der in einem Zeitschlitz eintreffenden Zellen

$$\xi_k^E = \begin{cases} 1-q, & k=0, \\ q, & k=1. \end{cases} \tag{4.18}$$

Das hochgestellte „E“ soll dabei andeuten, daß es sich um den Ankunftsprozeß des Eingangspuffers handelt.

Im folgenden wird zunächst angenommen, daß die Warteschlangenlänge der Eingangspuffer nicht begrenzt ist. Mit $\xi^E(z) = 1 - q + qz$ ergibt sich aus Gleichung A.34 die Erzeugende Funktion $w(z)$ der Wartezeitverteilung zu

$$w(z) = (1 - qh'(1)) \frac{z - 1}{z - 1 + q(1 - h(z))}, \tag{4.19}$$

wobei $h(z)$ die Erzeugende Funktion der Bedienzeit mit Mittelwert $h'(1)$ ist. Im allgemeinen sind die Verweildauern aufeinanderfolgender Zellen in der virtuellen Ausgangswarteschlange (und damit die Bedienzeiten) zwar korreliert, wegen der Gedächtnisfreiheit des Ankunftsprozesses ist diese Korrelation für die Analyse des Modells jedoch nicht von Bedeutung.

Die mittlere Wartezeit einer Zelle im Eingangspuffer ist

$$\mathrm{E}[W] = \frac{qh''(1)}{2(1 - qh'(1))}, \tag{4.20}$$

wobei $h''(1)$ die zweite Ableitung von $h(z)$ an der Stelle $z = 1$ ist. $\mathrm{E}[W]$ hängt also nicht nur von der mittleren Bedienzeit, sondern auch vom zweiten Moment der Bedienzeitverteilung ab.

Für die Berechnung der Bedienzeitverteilung h_k ist eine genauere Betrachtung der virtuellen Ausgangswarteschlange erforderlich. Es sei Y_n^j die Anzahl der HOL-Zellen, die im Zeitschlitz n um den Ausgang j konkurrieren. Das Verhalten der virtuellen Ausgangswarteschlange wird durch folgende Beziehung beschrieben (vgl. Gleichung A.1):

$$Y_{n+1}^j = \max(0, Y_n^j + \Xi_n^j - 1),$$

wobei Ξ_n^j die Anzahl der für Ausgang j bestimmten Zellen ist, die im Zeitschlitz n erstmals an die vorderste Stelle eines Eingangspuffers rücken. Außer von Y_n^j hängt Ξ_n^j auch von der Belegung aller anderen virtuellen Ausgangswarteschlangen ab, da die entsprechenden Eingangspuffer keinen Beitrag zu Ξ_n^j liefern können. Für die Berechnung der Verteilung von Ξ_n^j wird daher die gemeinsame Zustandsverteilung aller virtuellen Ausgangswarteschlangen benötigt, so daß eine exakte Analyse wegen der Größe und Komplexität des Zustandsraumes auch hier ausscheidet. Für den Grenzfall $N \to \infty$ kann jedoch gezeigt werden, daß die Verteilung von Ξ_n^j gegen eine Poisson-Verteilung konvergiert [111]. Bei symmetrischer Last muß außerdem die Ankunftsrate der virtuellen Ausgangswarteschlangen gleich der Ankunftsrate an den Eingängen des Koppelelements sein. Es ergibt sich daher im Grenzfall $N \to \infty$ für die stationäre Verteilung von Ξ_n^j:

$$\mathrm{P}\{\Xi_n^j = k\} = \mathrm{e}^{-q} \frac{q^k}{k!}. \tag{4.21}$$

Mit Hilfe von Gleichung A.35 erhält man somit folgende Beziehung für die mittlere Verweildauer $\mathrm{E}[H]$ einer Zelle in der virtuellen Ausgangswarteschlange:

$$\mathrm{E}[H] = h'(1) = 1 + \frac{q}{2(1 - q)} = \frac{2 - q}{2(1 - q)}, \quad q < 1. \tag{4.22}$$

Da die Arbitrierungsstrategie keinen Einfluß auf die Zustandsverteilung der virtuellen Ausgangswarteschlange hat, ist die mittlere Warteschlangenlänge und (nach dem Gesetz von Little) auch die mittlere Verweildauer von der Arbitrierungsstrategie unabhängig. Für die Auslastung A^E des Eingangspuffers gilt daher in jedem Fall

$$A^E = q \cdot \mathrm{E}[H] = \frac{q(2-q)}{2(1-q)}, \quad q < 1. \tag{4.23}$$

Aus Gleichung 4.23 läßt sich entnehmen, daß bereits für $q = 2 - \sqrt{2}$ ein Angebot von $A^E = 1$ erreicht wird. Für den Grenzfall $N \to \infty$ beträgt der maximale Durchsatz somit etwa 0.58. Für kleine N kann der Sättigungsdurchsatz auch auf analytischem Weg berechnet werden [21], wobei sich die in Tabelle 4.1 enthaltenen Werte ergeben [111]. Der Maximaldurchsatz nimmt mit steigendem N sehr rasch ab und nähert sich dem Grenzwert von 0.58 für $N \to \infty$.

N	2	3	4	5	6	7	8
Maximaler Durchsatz	0.7500	0.6825	0.6553	0.6399	0.6302	0.6234	0.6184

Tabelle 4.1: Sättigungsdurchsatz von Koppelelementen mit Eingangspuffern [111]

Im folgenden soll die Verteilung der Verweildauer einer Zelle an der HOL-Position für verschiedene Arbitrierungsstrategien untersucht werden. Eine analytische Berechnung ist dabei mit vertretbarem Aufwand nur für solche Strategien möglich, die außer der virtuellen Ausgangswarteschlange keine weiteren Informationen über den Zustand der Eingangspuffer benötigen. Falls jedoch die Warteschlangenlänge oder die Wartezeit im gesamten Eingangspuffer als Kriterium dienen, scheidet eine analytische Untersuchung aus.

Die Verteilung der Verweildauer kann sehr einfach berechnet werden, wenn die Zellen in der Reihenfolge ihres Eintreffens an der HOL-Position ausgewählt werden. Diese mit HOL-FIFO bezeichnete Arbitrierungsstrategie entspricht einer Abarbeitung der virtuellen Ausgangswarteschlangen nach dem FIFO-Prinzip, so daß auf die in Abschnitt 4.1 gewonnenen Ergebnisse für die Wartezeitverteilung bei FIFO-Abfertigung zurückgegriffen werden kann. Unter Verwendung von Gleichung 4.5 erhält man für die Erzeugende Funktion $h(z)$ der Verweildauer in der virtuellen Ausgangswarteschlange

$$h(z) = \frac{1-q}{q} \, \frac{z\,(\mathrm{e}^{-q(1-z)} - 1)}{z - \mathrm{e}^{-q(1-z)}}, \tag{4.24}$$

wobei für $\xi(z)$ bereits die Erzeugende Funktion der Poisson-Verteilung eingesetzt wurde. Mit Hilfe von Gleichung 4.7 kann auch die zweite Ableitung von $h(z)$ an der Stelle $z = 1$ bestimmt werden:

$$h''(1) = q \, \frac{q^2 - 4q + 6}{6(1-q)^2}, \quad q < 1. \tag{4.25}$$

Nach Einsetzen der Gleichungen 4.22 und 4.25 in 4.20 ergibt sich somit die mittlere Wartezeit im Eingangspuffer für die Arbitrierungsstrategie HOL-FIFO zu

$$\mathrm{E}[W] = \frac{q^2}{6(1-q)} \left(1 + \frac{4}{(q-2)^2 - 2}\right), \quad q < 2 - \sqrt{2}. \tag{4.26}$$

Bild 4.11 zeigt einen Vergleich der analytisch berechneten mittleren Durchlaufzeit mit den Simulationsergebnissen eines 32×32-Koppelelements. Die geringfügige Abweichung zwischen Simulation und Analyse rührt daher, daß der maximale Durchsatz eines Koppelelementes mit 32 Eingängen etwas höher ist als für $N \to \infty$. Die Ergebnisse bestätigen, daß der maximale Durchsatz von Koppelelementen mit FIFO-Eingangspuffern und einfachem Lesezugriff für große N bei etwa 58% liegt.

Auch bei einer zufälligen Auswahl der HOL-Zellen läßt sich die Verteilung von H auf analytischem Weg bestimmen. In diesem Fall werden die Zellen in zufälliger Reihenfolge aus der virtuellen Ausgangswarteschlange entnommen, wobei jede Zelle mit der gleichen Wahrscheinlichkeit ausgewählt wird. Die Analyse der Wartezeitverteilung ist hier jedoch wesentlich aufwendiger als bei einer FIFO-Abfertigung, da die Wartezeit einer Zelle auch durch später eintreffende Zellen noch beeinflußt werden kann. Eine geschlossene Lösung dieses Problems ist bisher nicht bekannt, die Wartezeitverteilung bzw. die Verteilung der Verweildauer kann aber mit Hilfe des in Anhang B beschriebenen Verfahrens numerisch berechnet werden. In Anhang B wird außerdem gezeigt, daß eine zufällige Auswahl der HOL-Zellen zu wesentlich größeren Schwankungen der Bedienzeit führt als bei einer HOL-FIFO-Arbitrierung. Nach

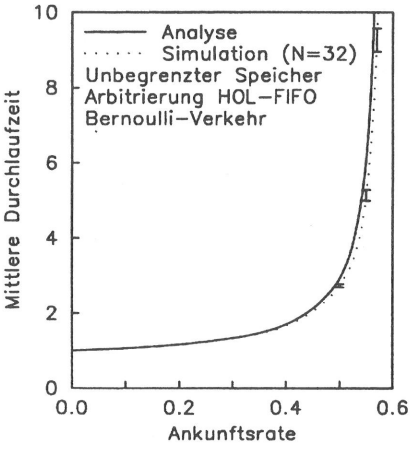


Bild 4.11: Vergleich von Analyse und Simulation der mittleren Durchlaufzeit bei einer HOL-FIFO-Arbitrierung

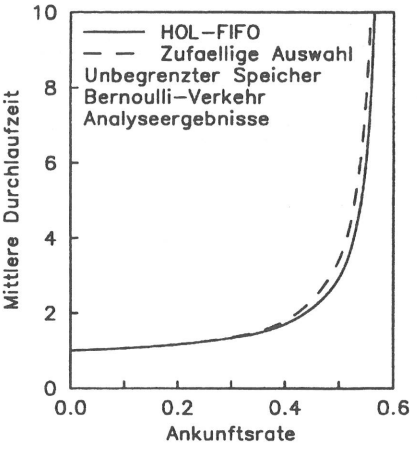


Bild 4.12: Vergleich der mittleren Durchlaufzeit bei einer zufälligen Auswahl und bei einer HOL-FIFO-Arbitrierung

Gleichung 4.20 ist daher auch die mittlere Durchlaufzeit der Zellen im Eingangspuffer bei zufälliger Auswahl etwas höher als bei der HOL-FIFO-Strategie (siehe Bild 4.12).

Obwohl sich die Arbitrierungsstrategie nur schwach auf die mittlere Durchlaufzeit auswirkt, treten bei der Zellverlustwahrscheinlichkeit wesentlich stärkere Unterschiede auf. Bild 4.13 zeigt die Abhängigkeit der Verlustwahrscheinlichkeit von der Puffergröße für beide Abfertigungsstrategien bei einem Angebot von 0.4 und 0.5. Die Ergebnisse zeigen, daß der Speicherbedarf bei einer zufälligen Auswahl um etwa 50% höher liegt als bei der HOL-FIFO-Strategie. In der gleichen Größenordnung bewegen sich auch die Unterschiede bei der Durchlaufzeitverteilung der Zellen. In Bild 4.13 sind außerdem einige Simulationsergebnisse in Form von Vertrauensintervallen eingetragen, die eine gute Übereinstimmung mit der Analyse zeigen. Die geringfügige Überschätzung der Verlustwahrscheinlichkeit durch die Analyse bei $A = 0.5$ ist auch hier darauf zurückzuführen, daß die Simulation für $N = 32$ durchgeführt wurde, während sich die Analyse auf den Fall $N \to \infty$ bezieht.

Mit Hilfe von zustandsabhängigen Arbitrierungsstrategien kann der Speicherbedarf noch wesentlich stärker reduziert werden. Die in Bild 4.14 dargestellten Simulationsergebnisse für die komplementäre Verteilungsfunktion der Warteschlangenlänge in einem 32×32-Koppelelement bestätigen, daß der Speicherbedarf durch die Auswahl der längsten Warteschlange (Longest Queue First, LQF) minimiert werden kann. Auch die Auswahl nach der Wartezeit im Eingangspuffer (FIFO) führt zu einer deutlichen Einsparung gegenüber einer zufälligen Auswahl und der Strategie HOL-FIFO.

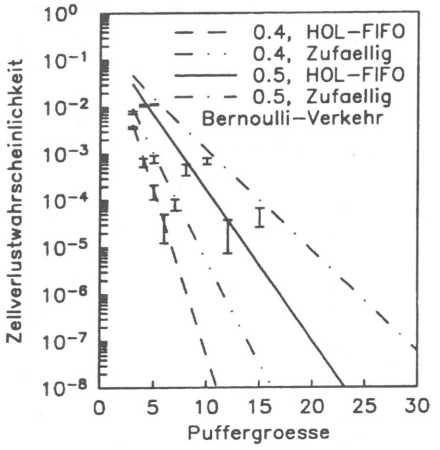


Bild 4.13: Zellverlustwahrscheinlichkeit bei HOL-FIFO-Arbitrierung und bei zufälliger Auswahl

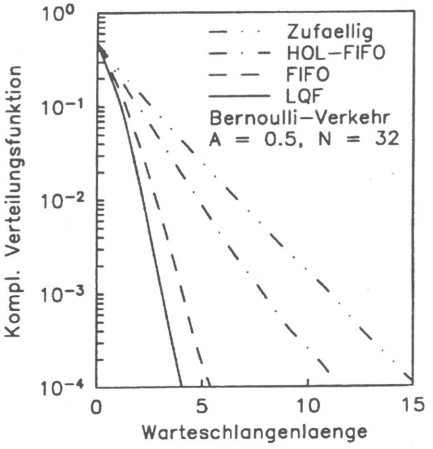


Bild 4.14: Einfluß der Arbitrierungsstrategie auf die Verteilung der Warteschlangenlänge (Simulationsergebnisse)

	FIFO	LQF	HOL-FIFO	Zuf. Auswahl
Mittlere Durchlaufzeit	2.367 ±0.021	2.231 ±0.008	2.727 ±0.041	3.099 ±0.059
Variationskoeffizient	0.542 ±0.003	0.565 ±0.002	0.774 ±0.011	0.930 ±0.012

Tabelle 4.2: Mittelwert und Variationskoeffizient der Durchlaufzeit für verschiedene Arbitrierungsstrategien ($A = 0.5$, $N = 32$)

Tabelle 4.2 zeigt den Einfluß der Arbitrierungsstrategie auf Mittelwert und Variationskoeffizient der Durchlaufzeit für ein Angebot von $A = 0.5$. Die Strategie LQF erweist sich auch bezüglich der mittleren Durchlaufzeit als optimal, lediglich der Variationskoeffizient ist hier etwas größer als bei einer wartezeitabhängigen Auswahl. Unter Berücksichtigung des Implementierungsaufwands erscheint daher die Strategie LQF am besten für ATM-Koppelelemente mit FIFO-Eingangspuffern geeignet.

4.3.2 Eingangspuffer mit wahlfreiem Lesezugriff

In diesem Abschnitt sollen die Möglichkeiten zur Durchsatzsteigerung am Beispiel von Koppelelementen mit wahlfreiem Zugriff (RAM-Eingangspuffer) und einem sequentiellen Auswahlverfahren untersucht werden. Bei einem sequentiellen Auswahlverfahren werden die Eingangspuffer, beginnend mit der HOL-Position, schrittweise nach Zellen durchsucht, die zu einem freien Ausgang übertragen werden können. Die Suchtiefe (Fenstergröße) soll nicht begrenzt sein, so daß der gesamte Puffer abgesucht werden kann.

Bild 4.15 zeigt die Simulationsergebnisse für die mittlere Durchlaufzeit der Zellen, wobei als Arbitrierungsstrategie stets eine wartezeitabhängige Auswahl (FIFO) verwendet wurde. Die RAM-Eingangspuffer wurden sowohl mit einfachem als auch mit zweifachem Lesezugriff untersucht. Ein zweifacher Lesezugriff bietet die Möglichkeit, aus einem Eingangspuffer gleichzeitig zwei Zellen zu entnehmen, die für unterschiedliche Ausgänge bestimmt sind. Zum Vergleich sind in Bild 4.15 auch die Ergebnisse für FIFO-Eingangspuffer und für Ausgangspuffer eingetragen.

Der wahlfreie Pufferzugriff mit unbegrenzter Fenstergröße ermöglicht bei unendlichen Eingangspuffern einen maximalen Durchsatz von 100%, da an der Sättigungsgrenze in jedem Puffer eine Zelle gefunden werden kann, die für einen freien Ausgang bestimmt ist. Im Vergleich zu Koppelelementen mit Ausgangspuffern treten jedoch unterhalb der Sättigungsgrenze ebenfalls längere Wartezeiten auf. Bei einem einfachen Lesezugriff ist die mittlere Wartezeit der Zellen für ein Angebot von $A = 0.8$ nahezu doppelt so groß wie in einem

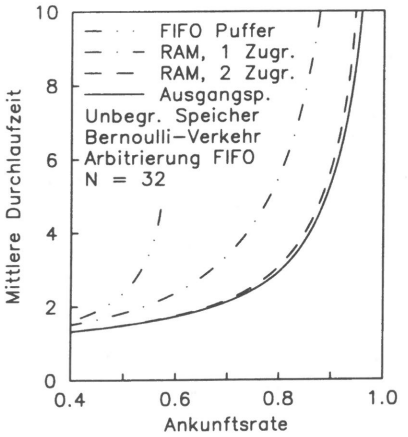


Bild 4.15: Mittlere Durchlaufzeit für Eingangspuffer mit wahlfreiem Zugriff

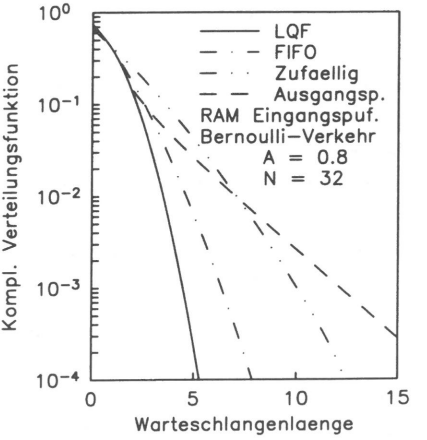


Bild 4.16: Komplementäre Verteilungsfunktion der Warteschlangenlänge für RAM-Eingangspuffer mit 2 Lesezugriffen

Koppelelement mit Ausgangspuffern. Wesentlich günstiger verhalten sich dagegen Eingangspuffer mit zweifachem Lesezugriff, deren mittlere Durchlaufzeit kaum von den Ergebnissen für Ausgangspuffer zu unterscheiden ist.

Bezüglich ihrer Verkehrsleistung sind Koppelelemente mit RAM-Eingangspuffern und zweifachem Lesezugriff sowie Koppelelemente mit Ausgangspuffern also durchaus vergleichbar. Die folgenden Ergebnisse zeigen, daß deren Speichbedarf im Vergleich zu Ausgangspuffern sogar wesentlich geringer ist. In Bild 4.16 ist die komplementäre Verteilungsfunktion der Warteschlangenlänge eines RAM-Eingangspuffers mit zweifachem Lesezugriff für verschiedene Arbitrierungsstrategien bei einem Angebot von $A = 0.8$ dargestellt. Falls die Arbitrierungsstrategie LQF verwendet wird, ist der Speicherbedarf des RAM-Eingangspuffers für die hier gewählten Parameter etwa um den Faktor 3 geringer als bei einem Ausgangspuffer.

Ein wesentlicher Nachteil von Eingangspuffern mit wahlfreiem Zugriff ist allerdings die aufwendigere Speicherverwaltung und die sehr komplexe Auswahl der Zellen. Die sequentielle Absuche erfordert eine zentrale Steuerung, die auf die Zieladressen aller gespeicherten Zellen zugreifen muß. Die Steuerung des Koppelelementes kann daher, insbesondere für große N, sehr schnell zu einem Engpaß werden. Ein Implementierungskonzept für ein derartiges Koppelelement mit 4 Eingängen ist in [151] und [194] beschrieben.

4.4 Verteilte Puffer

In einem Koppelelement mit verteilten Puffern werden die für einen Ausgang bestimmten Zellen in getrennten, dem jeweiligen Eingang zugeordneten Warteschlangen gespeichert. Die Leistungsfähigkeit der Anordnung hängt dabei ganz entscheidend von der verwendeten Arbitrierungsstrategie ab, mit deren Hilfe über die Auswahl der Zellen entschieden wird. Da eine Analyse der Puffer nur unter stark vereinfachten Voraussetzungen [87] bzw. für bestimmte Arbitrierungsstrategien [25] möglich ist, werden für die vergleichende Untersuchung ausschließlich Simulationsergebnisse herangezogen.

Im folgenden wird ein 16×16-Koppelelement mit verteilten Puffern betrachtet. Um die Unterschiede zwischen den einzelnen Arbitrierungsstrategien möglichst deutlich zu zeigen, wird ein vergleichsweise hohes Angebot von $A = 0.9$ gewählt. Die folgenden Arbitrierungsstrategien werden bezüglich der Verteilung der Durchlaufzeit und der Warteschlangenlänge verglichen:

- Zufällige Auswahl
- Zyklische Priorisierung
- Auswahl der „ältesten“ Zelle (FIFO)
- Auswahl der längsten Warteschlange (LQF)

Eine Abfrage der Puffer in fester Reihenfolge wurde nicht berücksichtigt, da diese Strategie zu sehr starken Unterschieden zwischen den einzelnen Eingängen führt [97]. Die Strategien FIFO und LQF liefern nicht immer ein eindeutiges Ergebnis, so daß hier eine Zweitstrategie zur Auswahl des Puffers benötigt wird. Als Zweitstrategie wurde sowohl eine feste als auch eine zyklische Priorisierung untersucht, wobei die Ergebnisse im Fall einer festen Priorisierung nur für den zuerst und für den zuletzt abgefragten Eingang eingezeichnet sind.

Die komplementäre Verteilungsfunktion der Durchlaufzeit ist in Bild 4.17 für die genannten Arbitrierungsstrategien dargestellt. Wie erwartet sind die Schwankungen der Durchlaufzeit bei einer wartezeitabhängigen Auswahl (FIFO) am geringsten, wobei die Zweitstrategie hier nur einen sehr schwachen Einfluß auf die Ergebnisse hat. Sowohl die zufällige Auswahl als auch die zyklische Priorisierung weisen ein deutlich ungünstigeres Verhalten auf.

Die rechte Hälfte von Bild 4.17 zeigt die Ergebnisse für die Strategie LQF, wobei die Kurven für die zufällige Auswahl und die FIFO-Arbitrierung (Zweitstrategie zyklisch) zu Vergleichszwecken ebenfalls eingezeichnet sind. Falls die Zellen immer aus der längsten Warteschlange entnommen werden, hat auch die Zweitstrategie einen starken Einfluß auf die Verteilung der Durchlaufzeit. Als Zweitstrategie für LQF kommt daher nur eine zyklische Priorisierung in Frage, um eine faire Abfertigung aller Puffer sicherzustellen. In diesem Fall ergibt sich bezüglich der Durchlaufzeitverteilung ein ähnliches Verhalten wie bei einer zufälligen Auswahl.

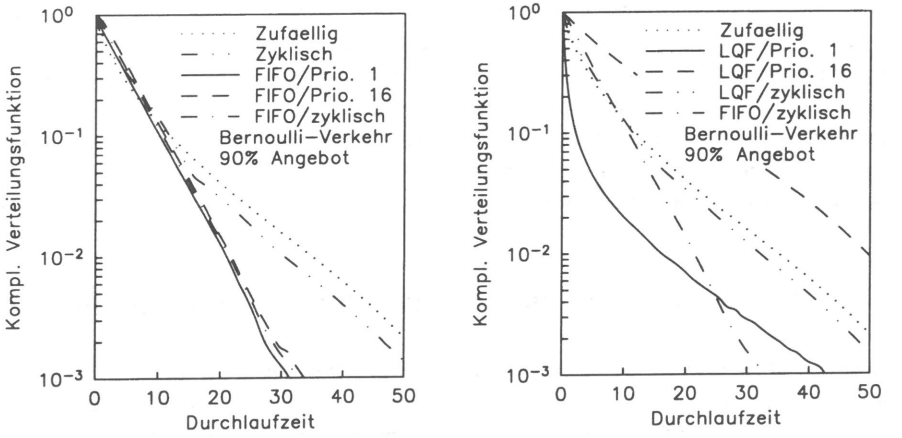


Bild 4.17: Einfluß der Arbitrierungsstrategie auf die komplementäre Verteilungsfunktion der Durchlaufzeit bei verteilten Puffern in einem 16×16-Koppelelement

Bild 4.18 zeigt die komplementäre Verteilungsfunktion der Warteschlangenlänge aus der Sicht einer ankommenden Zelle. Auch hier erweist sich eine Arbitrierung nach dem FIFO-Prinzip der zufälligen bzw. zyklischen Auswahl deutlich überlegen. Den geringsten Speicherbedarf erhält man jedoch mit der Strategie LQF, wobei aus den oben genannten Gründen eine zyklische Priorisierung als Zweitstrategie verwendet werden sollte.

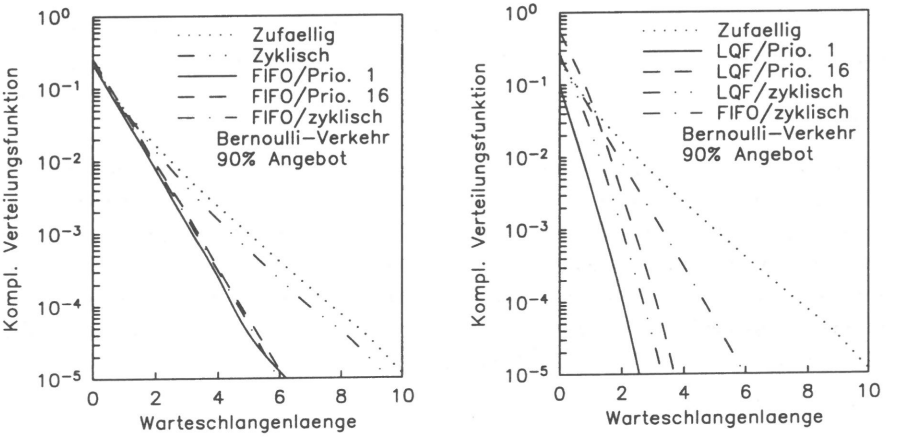


Bild 4.18: Komplementäre Verteilungsfunktion der Warteschlangenlänge für verschiedene Arbitrierungsstrategien

Zusammenfassend kann festgestellt werden, daß eine gleichzeitige Optimierung von Durchlaufzeit und Puffergröße mit den untersuchten Arbitrierungsstrategien nicht möglich ist. Die Strategie LQF weist den geringsten Speicherplatzbedarf auf und sollte immer dann verwendet werden, wenn die etwas größeren Durchlaufzeitschwankungen (verglichen mit dem FIFO-Prinzip) tolerierbar sind. Die FIFO-Strategie eignet sich dagegen für die Fälle, in denen es auf eine minimale Schwankung der Durchlaufzeit ankommt, wobei der Speicherplatzbedarf nur eine untergeordnete Rolle spielt.

4.5 Eingangspuffer mit zusätzlichen Ausgangspuffern

In diesem Abschnitt wird die Verkehrsleistung von Koppelelementen mit Eingangs- und Ausgangspuffern untersucht. Um den Einfluß einer internen Geschwindigkeitserhöhung zu zeigen, wird eine unbegrenzte Puffergröße aller Eingangs- und Ausgangspuffer vorausgesetzt. Es soll außerdem angenommen werden, daß die Eingangspuffer nur einen Lesezugriff pro Zeitschlitz erlauben und nach dem FIFO-Prinzip organisiert sind. Der maximale Durchsatz hängt in diesem Fall ausschließlich von der Anzahl der Schreibzugriffe ab, die gleichzeitig an einem Ausgangspuffer möglich sind.

Die Modellierung eines derartigen Koppelelements basiert auf der Untersuchung von Koppelelementen mit Eingangspuffern (siehe Abschnitt 4.3). Insbesondere wird auch hier das Konzept der virtuellen Ausgangswarteschlange verwendet, mit deren Hilfe der Arbitrierungsvorgang der HOL-Zellen modelliert werden kann. Eine virtuelle Ausgangswarteschlange besteht aus allen HOL-Zellen, die für denselben Ausgang bestimmt sind. Abhängig von der internen Geschwindigkeitserhöhung können aus dieser virtuellen Ausgangswarteschlange in einem Zeitschlitz bis zu m_o Zellen entnommen und in den entsprechenden Ausgangspuffer eingeschrieben werden (m_o ist die maximale Anzahl der gleichzeitigen Schreibzugriffe auf einen Ausgangspuffer).

Die Modellierung der virtuellen Ausgangswarteschlange führt auf ein Bediensystem mit insgesamt m_o Bedieneinheiten [163], [165]. Prinzipiell treten bei der Untersuchung der virtuellen Ausgangswarteschlange die gleichen Probleme auf wie bei reinen Eingangspuffern, so daß eine Analyse nur für den Grenzfall $N \rightarrow \infty$ möglich ist. Für Bediensysteme mit mehreren Bedieneinheiten und konstanter Bedienzeit kann außerdem keine geschlossene Lösung für die Erzeugende Funktion der Zustandsverteilung angegeben werden, die Bestimmung der Zustandswahrscheinlichkeiten muß daher auf numerischem Weg erfolgen.

Die analytische Berechnung des maximalen Durchsatzes für Koppelelemente mit Eingangs- und Ausgangspuffern wurde erstmals in [163] beschrieben. Ein ähnlicher Ansatz wird auch in [47] vorgestellt, dieser bezieht sich allerdings nur auf Systeme mit negativ-exponentiell verteilter Bedienzeit und kann hier nicht direkt angewendet werden. Das Verhalten der

Eingangspuffer ist in [138] analysiert, wobei die Verweildauer in der virtuellen Ausgangswarteschlange durch eine Phasenverteilung approximiert wird. In [31] wird auch der Einfluß von endlichen Ausgangspuffern auf den maximalen Durchsatz untersucht, das resultierende Modell besitzt jedoch einen zweidimensionalen Zustandsraum und ist bereits sehr aufwendig. Ein Vergleich von verschiedenen Arbitrierungsstrategien zur Auswahl der HOL-Zellen ist in [105] beschrieben, wobei ebenfalls begrenzte Ausgangspuffer (mit voller Geschwindigkeitserhöhung) verwendet werden.

Für Koppelelemente mit kleinem N ist eine analytische Untersuchung (ähnlich wie bei reinen Eingangspuffern) allerdings nicht möglich, hier muß wiederum die Simulationstechnik eingesetzt werden. Die Abhängigkeit des maximalen Durchsatzes von der Größe N des Koppelelements ist in Bild 4.19 für verschiedene Werte von m_o dargestellt. Der Fall $m_o = 1$ entspricht einem Koppelelement mit reinen FIFO-Eingangspuffern und wurde zu Vergleichszwecken eingezeichnet. Bild 4.19 zeigt, daß bereits mit $m_o = 2$ eine deutliche Leistungssteigerung erzielt werden kann, wobei der maximale Durchsatz für $N \rightarrow \infty$ bei etwa 88% liegt [31],[163]. Für $m_o = 4$ steigt der Maximaldurchsatz auf nahezu 100% an, die Anordnung verhält sich also praktisch wie ein Koppelelement mit Ausgangspuffern.

Die komplementäre Verteilungsfunktion der Warteschlangenlänge ist in Bild 4.20 für ein 16×16-Koppelelement mit $m_o = 2$ und $m_o = 4$ aufgezeichnet. Selbst bei einem Angebot von $A = 0.8$ und mit $m_o = 2$ benötigen die Eingangspuffer nur wenige Warteplätze, um eine sehr geringe Verlustwahrscheinlichkeit zu garantieren. Für den Fall $m_o = 4$ enthalten die Ein-

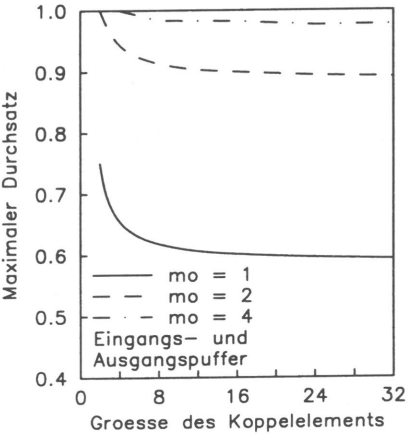


Bild 4.19: Maximaler Durchsatz bei Koppelelementen mit kombinierten Eingangs- und Ausgangspuffern

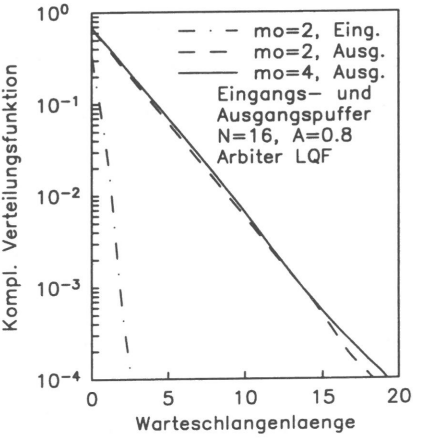


Bild 4.20: Komplementäre Verteilungsfunktion der Warteschlangenlänge für Eingangs- und Ausgangspuffer

gangspuffer während des gesamten Simulationslaufs nie mehr als 2 Zellen, die entsprechende Kurve ist daher in Bild 4.20 nicht eingezeichnet. Interessant ist auch, daß die Zustandsverteilung der Ausgangspuffer für $m_o = 2$ und $m_o = 4$ nahezu identisch ist. Der Vergleich mit reinen Ausgangspuffern ($m_o = 16$) zeigt ebenfalls eine gute Übereinstimmung, der Einfluß der Eingangspuffer kann also bei der Dimensionierung der Ausgangspuffer (zumindest für $N \geq 16$) vernachlässigt werden.

4.6 Vergleich und Bewertung der Architekturen

Die Ergebnisse in diesem Kapitel haben gezeigt, daß das Verhalten von Koppelelementen mit Ausgangspuffern bezüglich Maximaldurchsatz und Durchlaufzeitschwankungen auch von anderen Architekturen erreicht werden kann. Die einzelnen Varianten unterscheiden sich vor allem durch ihren Speicherbedarf, die Speicherbandbreite, die Speicherorganisation sowie durch die Komplexität der Steuerung.

Tabelle 4.3 enthält eine Zusammenfassung der wichtigsten Eigenschaften von einigen ausgewählten Architekturen für ein 16×16-Koppelelement. Bei der Auswahl wurde eine Optimierung hinsichtlich des Speicherbedarfs vorgenommen, als Arbitrierungsstrategie kommt daher stets das LQF-Prinzip mit einer zyklischen Priorisierung als Zweitstrategie zum Einsatz. Der Speicherbedarf wurde analytisch (Ausgangspuffer) bzw. mit Hilfe von Simulationsergebnissen für eine Verlustwahrscheinlichkeit von 10^{-4} ermittelt (ein Vergleich bei kleineren Verlustwahrscheinlichkeiten ist aus Gründen der Laufzeit der Simulationen schwierig). Die Zahlenwerte beziehen sich ausschließlich auf den Fall $N = 16$.

In Bezug auf den Speicherbedarf stellen Koppelelemente mit Zentralspeicher in jedem Fall die optimale Lösung dar. RAM-Eingangspuffer mit zwei Lesezugriffen schneiden hier ebenfalls

	Speicherbedarf pro Ausgang, Verlustwkt. 10^{-4}	Speicher-bandbreite	Speicher-organisation	Steuerungs-aufwand
Zentralspeicher	4 Zellen	$2NL$	RAM	hoch
RAM-Eingangspuffer	6 Zellen	$3L$	RAM	sehr hoch
Ausgangspuffer	17 Zellen	$(N+1)L$	FIFO	gering
Eingangs-/Ausgangspuffer	20 Zellen	$2L/3L$	FIFO	mittel
Verteilte Puffer	48 Zellen	$2L$	FIFO	mittel

Tabelle 4.3: Zusammenfassung der wichtigsten Eigenschaften ausgewählter Architekturen

sehr gut ab, alle anderen Varianten benötigen wesentlich mehr Speicherplatz. Die Speicherbandbreite ist als Vielfaches der Bitrate L einer Eingangsleitung angegeben und bezieht sich auf die Summe aller Schreib-/Lesezugriffe. Die höchsten Anforderungen an die Speicherbandbreite stellt ebenfalls der Zentralspeicher ($2NL$), gefolgt von Koppelelementen mit Ausgangspuffern ($(N+1)L$). Alle anderen Architekturen kommen mit zwei bzw. drei Speicherzugriffen pro Zelltakt aus, unabhängig von der Anzahl der Eingänge/Ausgänge.

Für den Implementierungsaufwand ist jedoch auch die Pufferorganisation von großer Bedeutung, wobei FIFO-Warteschlangen wesentlich effizienter realisiert werden können als Speicher mit wahlfreiem Zugriff (RAM). Der Steuerungsaufwand hängt sowohl von der Pufferorganisation als auch von weiteren Merkmalen wie der Arbitrierungsstrategie (falls vorhanden) und der Zugriffstiefe der Puffer ab. Koppelelemente mit RAM-Eingangspuffern, aus denen eine beliebige Zelle entnommen werden kann, benötigen mit Abstand den höchsten Steuerungsaufwand. Die Steuerung eines zentralen Speichers ist wegen der Verwaltung der freien Speicherplätze ebenfalls vergleichsweise aufwendig. Den geringsten Steuerungsaufwand besitzen Koppelelemente mit Ausgangspuffern, die keine Arbitrierung erfordern und mit einfachen FIFO-Puffern auskommen.

Tabelle 4.3 zeigt auch, daß eine Optimierung des Speicherbedarfs nur auf Kosten des Steuerungsaufwands möglich ist. Durch eine entsprechend komplexe Steuerung kann also Speicherplatz eingespart werden, wobei der Komplexität der Steuerung aus Geschwindigkeitsgründen jedoch Grenzen gesetzt sind. Für Anwendungen mit extremen Geschwindigkeitsanforderungen sind daher Architekturen mit einfacher Steuerung besser geeignet, wobei auch die Speicherbandbreite berücksichtigt werden muß.

Kapitel 5

Analyse mehrstufiger Koppelnetze

Die Untersuchung von mehrstufigen Koppelanordnungen auf Zellebene ist sowohl für die Bestimmung der Durchlaufzeitverteilung der Zellen als auch für die Dimensionierung der Puffer in den inneren Stufen des Koppelnetzes von Bedeutung. Eine exakte Analyse von mehrstufigen Koppelnetzen führt jedoch zu sehr komplexen Modellen mit mehrdimensionalen Zustandsräumen, die bereits für kleine Koppelnetze nicht mehr beherrschbar sind. Aus diesem Grund wird das Modell meist in kleinere Einheiten wie Koppelelemente oder Pufferspeicher zerlegt, die einzeln untersucht werden können (Dekomposition). Anschließend werden die gewünschten Ergebnisse (wie z.B. die Durchlaufzeitverteilung der Zellen) aus den Einzelergebnissen abgeleitet (Aggregation).

Sowohl die Dekomposition als auch die Aggregation erfordern allerdings bestimmte Näherungsannahmen bezüglich der Abhängigkeiten zwischen den einzelnen Komponenten des Modells. Für die Untersuchung eines internen Koppelelements müssen beispielsweise die Eigenschaften des Verkehrs bekannt sein, der an den Eingängen dieses Koppelelements anliegt. Da die Verkehrscharakteristik durch die Speicherung der Zellen in den Koppelelementen beeinflußt wird, können sich die Eigenschaften des internen Verkehrs von denen des ankommenden Verkehrs deutlich unterscheiden.

In der Literatur wird die Veränderung des Verkehrs durch die Koppelelemente bei der Analyse mehrstufiger Koppelanordnungen meist vernachlässigt. Es wird davon ausgegangen, daß der interne Verkehr die gleiche Charakteristik besitzt wie der an den Eingängen des Koppelnetzes ankommende Verkehr. Die meisten analytischen Untersuchungen basieren deshalb auf der Annahme, daß auch der interne Verkehr durch einen Bernoulli-Prozeß modelliert werden kann. Diese Annahme führt jedoch in bestimmten Fällen zu einer deutlichen Unterschätzung der Verlustwahrscheinlichkeit sowie der Durchlaufzeitschwankungen. Dieser Effekt ist zwar auch in der Literatur bereits beschrieben worden [125],[185], bisher sind jedoch keine zufriedenstellenden Modelle für eine analytische Behandlung bekannt.

In diesem Kapitel sollen die Eigenschaften des internen Verkehrs genauer untersucht und geeignete Ersatzmodelle gefunden werden. Der interne Verkehr entspricht im allgemeinen dem Ausgangsprozeß einer realen oder virtuellen Ausgangswarteschlange, so daß die Betrachtung von Koppelelementen mit Ausgangspuffern für die Bestimmung der prinzipiellen Verkehrscharakteristik völlig ausreichend ist. Um eine analytische Behandlung zu ermöglichen, muß dieser Ausgangsprozeß durch einfachere Modelle approximiert werden, welche die wesentlichen Eigenschaften wie das Korrelationsverhalten hinreichend genau wiedergeben. Am Beispiel eines Koppelelements mit Ausgangspuffern wird auch ein geeignetes Analyseverfahren zur Bestimmung der Verlustwahrscheinlichkeit sowie der Durchlaufzeitverteilung vorgestellt.

Für die im folgenden betrachteten Koppelnetze wurde die Topologie eines Banyan-Netzwerks gewählt, da dieses Netzwerk eine sehr einfache Struktur besitzt (siehe 2.3.4.1). Wegen der zufälligen Verzweigung des Verkehrs sind die Ergebnisse ohnehin von der jeweiligen Topologie des Koppelnetzes unabhängig, so daß die Wahl der Koppelnetztopologie hier nicht von Bedeutung ist.

5.1 Eigenschaften des internen Verkehrs

Die Eigenschaften des internen Verkehrs werden im folgenden für ein Koppelelement mit Ausgangspuffern untersucht. In diesem speziellen Fall ist der Ankunftsprozeß eines internen Koppelelements mit dem Ausgangsprozeß des Puffers in der vorhergehenden Stufe identisch. Ein Beispiel für den zeitlichen Ablauf des Ausgangsprozesses ist in Bild 5.1 zusammen mit dem „Belegungsgebirge“ des Ausgangspuffers dargestellt.

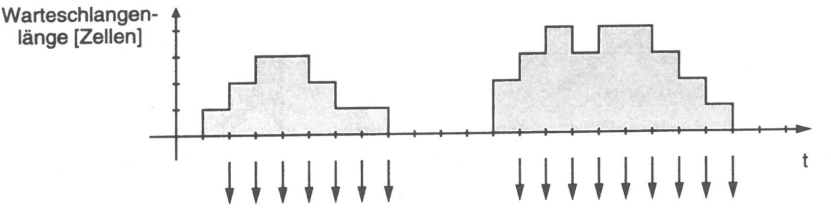


Bild 5.1: Ausgangsprozeß eines Koppelelementes mit Ausgangspuffern

Der Ausgangsprozeß ist durch einen Wechsel zwischen Betriebsperioden und Freiperioden gekennzeichnet. Solange sich Zellen in der Ausgangswarteschlange befinden, ist die Ausgangsleitung ohne Unterbrechung belegt (Betriebsperiode). In der anschließenden Freiperiode ist die Ausgangsleitung für mindestens eine Zeiteinheit frei. Wegen der konstanten Übertragungsdauer der Zellen verhält sich der Ausgangsprozeß wie ein deterministischer Prozeß, der durch einzelne Freiperioden unterbrochen ist. Bild 5.1 zeigt auch, daß der Ausgangsprozeß

eine stoßartige Charakteristik aufweist. Falls die Betriebs- und Freiperioden länger andauern, sind die Zellen des Ausgangsprozesses daher stark korreliert, so daß der Bernoulli-Prozeß nur eine grobe Näherung für den Ankunftsprozeß der folgenden Stufe darstellt.

Der interne Verkehr wird außerdem von jeder Stufe des Koppelnetzes beeinflußt, so daß die Eigenschaften des Ankunftsprozesses von Stufe zu Stufe verschieden sind. Da die Charakteristik des angebotenen Verkehrs einen direkten Einfluß auf das statistische Verhalten der Ausgangspuffer hat, unterscheiden sich die einzelnen Stufen auch bezüglich der Verteilung der Warteschlangenlänge und der Durchlaufzeit. Bild 5.2 zeigt beispielsweise die Verteilung der Warteschlangenlänge für die einzelnen Stufen eines Banyan-Netzwerks, welches aus 2×2-Koppelelementen aufgebaut ist. Als externer Ankunftsprozeß wurde ein Bernoulli-Prozeß gewählt, so daß die Ausgangspuffer in der ersten Stufe exakt analysiert werden können.

Die Ergebnisse zeigen, daß insbesondere in den ersten Stufen des Koppelnetzes eine starke Zunahme der Warteschlangenlänge zu beobachten ist. Diese Zunahme ist auf die Korrelationen des Ankunftsprozesses zurückzuführen, die von Stufe zu Stufe ebenfalls ansteigen (siehe Bild 5.3). Falls die Puffer auf der Basis eines Bernoulli-Ankunftsprozesses dimensioniert werden, führt dieser Effekt daher zu einer Überschreitung der zulässigen Verlustwahrscheinlichkeit in den hinteren Stufen des Koppelnetzes. Die Zunahme der mittleren Warteschlangenlänge und deren Varianz in einem mehrstufigen Koppelnetz wurde auch in [125] untersucht. Dort wird allerdings nur eine empirisch ermittelte Näherungsformel für die hinteren Stufen des Koppelnetzes angegeben.

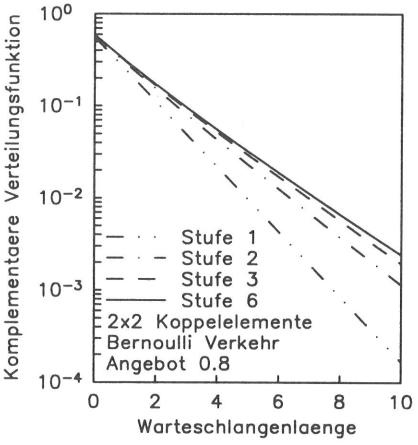


Bild 5.2: Komplementäre Verteilungsfunktion der Warteschlangenlänge in den einzelnen Stufen eines Banyan-Netzwerks

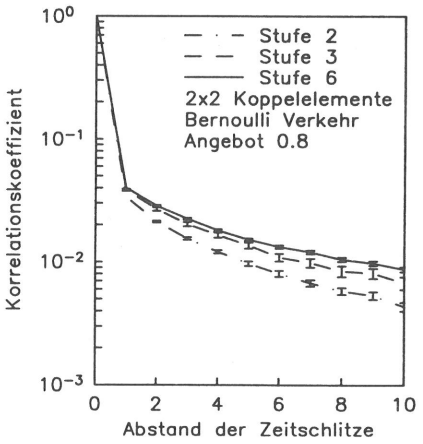


Bild 5.3: Korrelation der Zellankünfte am Eingang der Ausgangspuffer

5.2 Modellierung des Ausgangsprozesses

Falls der Ankunftsprozeß des Ausgangspuffers gedächtnisfrei ist, sind die einzelnen Betriebs- und Freiperioden voneinander unabhängig, so daß der Ausgangsprozeß durch ein Modell mit zwei Zuständen nachgebildet werden kann: einem Pausenzustand, welcher der Freiperiode entspricht, und einem aktiven Zustand, der die Betriebsperiode modelliert (vgl. Bild 5.1). Der Abstand der Zellen im aktiven Zustand ist konstant und entspricht der Übertragungsdauer einer ATM-Zelle. Bei gedächtnisbehafteten Ankunftsprozessen können die Betriebs- und Freiperioden zwar korreliert sein, im allgemeinen sind diese Korrelationen jedoch nicht sehr ausgeprägt und können vernachlässigt werden.

Zur Beschreibung des Ausgangsprozesses wird in jedem Fall die Verteilung der Dauer von Betriebs- und Freiperioden benötigt. Während die Verteilung der Freiperiode für gedächtnisfreie Ankunftsprozesse direkt angegeben werden kann, ist die Analyse der Betriebsperiode wesentlich aufwendiger. Auf der Grundlage einer Markov-Kette zur Beschreibung des Systemzustands kann die Verteilung der Betriebsperiode zwar in Form einer Phasenverteilung dargestellt werden, das resultierende Modell ist jedoch sehr komplex und daher nicht für analytische Untersuchungen geeignet. Um eine analytische Behandlung zu ermöglichen, wird der Ausgangsprozeß im folgenden durch wesentlich einfachere Modelle approximiert.

5.2.1 Verteilung der Freiperiode

Die Verteilung der Freiperiode soll durch eine geometrische Verteilung nachgebildet werden. Die Freiperiode besitzt tatsächlich eine geometrische Verteilung, falls der Ankunftsprozeß des betrachteten Ausgangspuffers gedächtnisfrei ist. Es sei I die Dauer einer Freiperiode mit Mittelwert $\mathrm{E}[I] = 1/\alpha$. Die Verteilung von I wird im folgenden durch eine geometrische Verteilung mit gleichem Mittelwert approximiert:

$$\mathrm{P}\{I = k\} = (1-\alpha)^{k-1}\alpha, \quad k \geq 1. \tag{5.1}$$

5.2.2 Verteilung der Betriebsperiode

Für die Approximation der Verteilung der Betriebsperiode werden im folgenden zwei verschiedene Ansätze entwickelt. In einem ersten Ansatz wird die Betriebsperiode ebenfalls durch eine geometrische Verteilung modelliert. Insbesondere bei hoher Auslastung ist die Varianz der Betriebsperiode jedoch wesentlich größer als die Varianz einer geometrischen Verteilung, so daß diese Näherung keine ausreichende Genauigkeit besitzt. In einem zweiten Ansatz wird daher eine hypergeometrische Verteilung verwendet, mit deren Hilfe auch die Varianz der Betriebsperiode angepaßt werden kann.

5.2.2.1 Approximation durch eine geometrische Verteilung

Eine geometrische Verteilung ermöglicht nur die Anpassung des Mittelwerts $\mathrm{E}[B]$ der Betriebsperiode, höhere Momente können nicht berücksichtigt werden. Die Verteilung der Betriebsperiode wird in diesem Fall durch die nachstehende Beziehung angenähert:

$$b_k = \mathrm{P}\{B = k\} = (1-\beta)^{k-1}\beta, \quad k \geq 1. \tag{5.2}$$

Der Parameter β der geometrischen Verteilung ist durch $\beta = 1/\mathrm{E}[B]$ gegeben.

Die Verteilung der Betriebsperiode kann jedoch nur bei einer geringen Auslastung mit guter Genauigkeit durch eine geometrische Verteilung approximiert werden. Bei einer Auslastung von $A = 0.8$ sind bereits starke Abweichungen zwischen beiden Verteilungen erkennbar (siehe auch Bild 5.7).

5.2.2.2 Approximation durch eine hypergeometrische Verteilung

Eine zweiphasige hypergeometrische Verteilung kann durch das in Bild 5.4 dargestellte Phasenmodell beschrieben werden. Das Modell besteht aus zwei geometrisch verteilten Phasen mit Mittelwert $\mathrm{E}[T_1] = 1/\beta_1$ bzw. $\mathrm{E}[T_2] = 1/\beta_2$, wobei mit Wahrscheinlichkeit p_1 bzw. $p_2 = 1 - p_1$ jeweils die Phase 1 bzw. 2 durchlaufen wird. Diese Verteilung kann auch als zeitdiskrete Version der hyperexponentiellen Verteilung aufgefaßt werden.

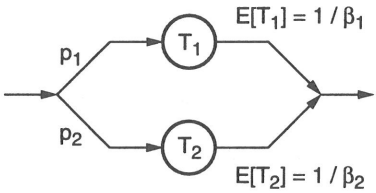


Bild 5.4: Phasenmodell der hypergeometrischen Verteilung

Entsprechend der Phasendarstellung in Bild 5.4 ist die hypergeometrische Verteilung eine gewichtete Summe aus zwei geometrischen Verteilungen:

$$b_k = \mathrm{P}\{B = k\} = p_1(1-\beta_1)^{k-1}\beta_1 + p_2(1-\beta_2)^{k-1}\beta_2. \tag{5.3}$$

Die ersten beiden Momente der Verteilung sind

$$\mathrm{E}[B] = \frac{p_1}{\beta_1} + \frac{p_2}{\beta_2}, \tag{5.4}$$

$$\mathrm{E}[B^2] = p_1 \frac{2-\beta_1}{\beta_1^2} + p_2 \frac{2-\beta_2}{\beta_2^2}. \tag{5.5}$$

Aus Gleichung 5.3 geht hervor, daß eine zweiphasige hypergeometrische Verteilung ingesamt drei freie Parameter besitzt. Es sollen jedoch nur die ersten beiden Momente angepaßt werden, zur eindeutigen Festlegung aller Parameter wird somit eine weitere Gleichung benötigt. Im folgenden soll angenommen werden, daß beide Teilphasen den gleichen Beitrag zum Mittelwert der hypergeometrischen Verteilung liefern (Symmetrie-Annahme). Die resultierende Verteilung bezeichnet man auch als symmetrische hypergeometrische Verteilung.

$$p_1 \mathrm{E}[T_1] = p_2 \mathrm{E}[T_2] = 0.5 \mathrm{E}[B]. \tag{5.6}$$

Die Parameter der Verteilung sollen nun so gewählt werden, daß die ersten beiden Momente der Betriebsperiode angepaßt werden können. Falls der Mittelwert $\mathrm{E}[B]$ und der Variationskoeffizient c_B der Verteilung der Betriebsperiode gegeben sind, erhält man die Parameter der symmetrischen hypergeometrischen Verteilung durch Umformung der Gleichungen 5.4 - 5.6. Mit $\beta = 1/\mathrm{E}[B]$ ergibt sich schließlich

$$p_1 = \frac{1}{2}\left(1 + \sqrt{\frac{c_B^2 + \beta - 1}{c_B^2 + \beta + 1}}\right), \tag{5.7}$$

$$p_2 = 1 - p_1, \tag{5.8}$$

$$\beta_1 = \frac{1}{\mathrm{E}[T_1]} = 2p_1\beta, \tag{5.9}$$

$$\beta_2 = \frac{1}{\mathrm{E}[T_2]} = 2p_2\beta. \tag{5.10}$$

Die Anpassung ist nur dann möglich, wenn $c_B^2 \geq 1 - \beta$ ist. Diese Bedingung ist jedoch immer erfüllt, da der Variationskoeffizient der Betriebsperiode stets größer als der Variationskoeffizient einer geometrischen Verteilung mit gleichem Mittelwert ist ($c^2 = 1 - \beta$).

5.2.3 Modellierung durch einen Markov-Modulierten Bernoulli-Prozeß

Ausgehend von den oben beschriebenen Ansätzen zur Approximation der Verteilung von Betriebs- und Freiperioden soll der resultierende stochastische Prozeß nun als ein Markov-Modulierter Bernoulli-Prozeß dargestellt werden. Diese Darstellung ist die Grundlage für die spätere Leistungsuntersuchung und ermöglicht auch eine analytische Berechnung des Korrelationsverhaltens.

Der Markov-Modulierte Bernoulli-Prozeß (MMBP) ist ein doppelt-stochastischer Prozeß, dessen momentane Ankunftsrate durch eine Markov-Kette gesteuert wird (siehe Anhang C). Der Prozeß basiert auf einer zeitdiskreten Markov-Kette mit n Zuständen, deren Verhalten durch die Übergangsmatrix $\mathbf{P} = [p_{ij}]$ bestimmt wird. Dabei ist p_{ij} die Wahrscheinlichkeit für einen Übergang vom Zustand i in den Zustand j. Die mittlere Dauer des Zustands i ist durch $1/(1 - p_{ii})$ gegeben.

Wenn sich die Markov-Kette in Zustand i befindet, ist die Wahrscheinlichkeit für ein Ankunftsereignis durch q_i gegeben. Innerhalb eines Zustands werden die Ankünfte also wie bei einem Bernoulli-Prozeß mit Ankunftsrate q_i erzeugt. Zur Vereinfachung der Schreibweise werden die Ankunftswahrscheinlichkeiten q_i zu dem Spaltenvektor $\mathbf{q} = (q_1, q_2, \ldots, q_n)^T$ zusammengefaßt. Der Markov-Modulierte Bernoulli-Prozeß ist somit durch die Übergangsmatrix $\mathbf{P}$ und den Vektor $\mathbf{q}$ vollständig definiert.

5.2.3.1 Modell 1

5.2.3.1.1 Modellbeschreibung

Falls sowohl die Freiperiode als auch die Betriebsperiode durch eine geometrische Verteilung approximiert werden, ist der resultierende stochastische Prozeß als MMBP mit zwei Zuständen darstellbar (siehe Bild 5.5). Ein äquivalentes Modell zur Approximation des Ausgangsprozesses ist auch in [27] beschrieben. Es soll angenommen werden, daß die Freiperiode dem Zustand 1 entspricht und die Betriebsperiode auf Zustand 2 abgebildet wird. Während der Freiperiode werden keine Zellen erzeugt ($q_1 = 0$), und in der Betriebsperiode folgen die Zellen unmittelbar aufeinander ($q_2 = 1$). Es gilt daher

$$\mathbf{q} = (0, 1)^T. \tag{5.11}$$

Für die Übergangsmatrix $\mathbf{P}$ der zugrundeliegenden Markov-Kette erhält man mit $\alpha = 1/\mathrm{E}[I]$ und $\beta = 1/\mathrm{E}[B]$

$$\mathbf{P} = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}. \tag{5.12}$$

Der stationäre Zustandsvektor $\boldsymbol{\pi} = (\pi_1, \pi_2)$ der Markov-Kette ist

$$\boldsymbol{\pi} = \frac{1}{\alpha+\beta}\,(\beta, \alpha) = \frac{1}{\mathrm{E}[B]+\mathrm{E}[I]}\,(\mathrm{E}[I], \mathrm{E}[B]). \tag{5.13}$$

Daraus ergibt sich die Ankunftsrate zu

$$\lambda = \boldsymbol{\pi}\,\mathbf{q} = \pi_1 = \frac{\alpha}{\alpha+\beta}. \tag{5.14}$$

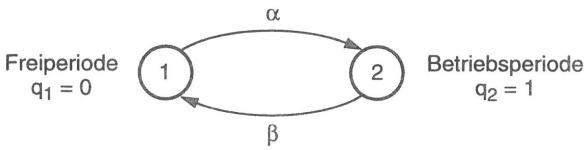


Bild 5.5: Modellierung des Ausgangsprozesses durch einen Markov-Modulierten Bernoulli-Prozeß mit zwei Zuständen

5.2.3.1.2 Korrelationsverhalten

Das Korrelationsverhalten des Ausgangsprozesses kann durch die Korrelation der Zellankünfte in unterschiedlichen Zeitschlitzen beschrieben werden. Es seien Ξ_j und Ξ_{j+k} jeweils die Anzahl der Ankünfte zu den Zeitpunkten j und $j+k$. Falls sich der Prozeß zum Zeitpunkt j im stationären Zustand befindet, hängt die Korrelation nur vom Abstand k der beiden Zeitpunkte ab.

Ein Maß für die Korrelation zwischen zwei Zufallsvariablen ist deren Kovarianz bzw. der zugehörige Korrelationskoeffizient. Die Kovarianz $\mathrm{COV}[\Xi_j, \Xi_{j+k}]$ ist wie folgt definiert:

$$\mathrm{COV}[\Xi_j, \Xi_{j+k}] = \mathrm{E}[(\Xi_j - \mathrm{E}[\Xi_j])(\Xi_{j+k} - \mathrm{E}[\Xi_{j+k}])].$$

Im stationären Zustand ist jedoch $\mathrm{E}[\Xi_j] = \mathrm{E}[\Xi_{j+k}] = \lambda$, es gilt daher

$$\begin{aligned} \mathrm{COV}[\Xi_j, \Xi_{j+k}] &= \mathrm{E}[\Xi_j \Xi_{j+k}] - \lambda^2 \\ &= \mathrm{P}\{\Xi_j = 1, \Xi_{j+k} = 1\} - \lambda^2. \end{aligned} \tag{5.15}$$

Der Korrelationskoeffizient $\mathrm{CORR}[\Xi_j, \Xi_{j+k}]$ ergibt sich aus der Kovarianz durch eine Normierung auf die Varianz $\mathrm{VAR}[\Xi]$:

$$\mathrm{CORR}[\Xi_j, \Xi_{j+k}] = \frac{\mathrm{COV}[\Xi_j, \Xi_{j+k}]}{\mathrm{VAR}[\Xi]} = \frac{\mathrm{COV}[\Xi_j, \Xi_{j+k}]}{\lambda(1-\lambda)}. \tag{5.16}$$

Dabei wurde vorausgesetzt, daß sich der Prozeß im stationären Zustand befindet, so daß $\mathrm{VAR}[\Xi_j] = \mathrm{VAR}[\Xi_{j+k}] = \lambda(1-\lambda)$ gilt.

In Anhang C wird gezeigt, daß die Kovarianz $\mathrm{COV}[\Xi_j, \Xi_{j+k}]$ für einen beliebigen MMBP in der folgenden Form dargestellt werden kann (siehe Gleichung C.14):

$$\mathrm{COV}[\Xi_j, \Xi_{j+k}] = \sum_{i=2}^{n} \mu_i^k a_i, \quad k \geq 1, \tag{5.17}$$

wobei $\mu_i, i > 1$, die von 1 verschiedenen Eigenwerte der Matrix $\mathbf{P}$ sind. Die Konstanten a_i ergeben sich aus den Eigenvektoren von $\mathbf{P}$, aus dem stationären Zustandsvektor $\boldsymbol{\pi}$ sowie aus $\mathbf{q}$ (vgl. C.12). Bei einem MMBP mit zwei Zuständen besitzt die Matrix $\mathbf{P}$ genau 2 Eigenwerte μ_1 und μ_2, wobei der erste Eigenwert bereits durch $\mu_1 = 1$ gegeben ist. Für die Kovarianz gilt daher

$$\mathrm{COV}[\Xi_j, \Xi_{j+k}] = \mu_2^k a_2, \quad k \geq 1. \tag{5.18}$$

Der zweite Eigenwert kann ebenfalls sofort angegeben werden, da die Summe der Eigenwerte gleich der Spur[1] einer Matrix ist:

$$\mu_2 = \mathrm{Spur}(\mathbf{P}) - \mu_1 = 1 - \alpha - \beta. \tag{5.19}$$

[1] Die Spur einer Matrix ist die Summe ihrer Diagonalelemente.

Die Konstante a_2 läßt sich zwar mit Hilfe der Eigenvektoren von $\mathbf{P}$ bestimmen, eine direkte Berechnung ist hier jedoch wesentlich einfacher. Aus Gleichung 5.18 folgt für die Kovarianz der Zellankünfte in zwei aufeinanderfolgenden Zeitschlitzen

$$\mathrm{COV}[\Xi_j, \Xi_{j+1}] = \mu_2 a_2. \tag{5.20}$$

Eine einfache Überlegung liefert für $\mathrm{COV}[\Xi_j, \Xi_{j+1}]$

$$\begin{aligned} \mathrm{COV}[\Xi_j, \Xi_{j+1}] &= \mathrm{P}\{\Xi_j = 1, \Xi_{j+1} = 1\} - \lambda^2 \\ &= \lambda(1 - \mathrm{P}\{\text{letzte Zelle der Betriebsperiode}\}) - \lambda^2 \\ &= \lambda(1 - \frac{1}{\mathrm{E}[B]}) - \lambda^2 \;=\; \lambda(1-\lambda)(1-\alpha-\beta). \end{aligned} \tag{5.21}$$

Es sollte beachtet werden, daß die Korrelation zwischen Zellankünften in aufeinanderfolgenden Zeitschlitzen nur von den Mittelwerten $\mathrm{E}[I]$ und $\mathrm{E}[B]$ abhängt, nicht jedoch von der Verteilung der Betriebs- und Freiperioden. Durch eine Anpassung dieser Mittelwerte ergibt sich daher in jedem Fall der korrekte Wert für $\mathrm{COV}[\Xi_j, \Xi_{j+1}]$.

Ein Vergleich mit Gleichung 5.20 liefert nun $a_2 = \lambda(1-\lambda)$, und man erhält

$$\mathrm{COV}[\Xi_j, \Xi_{j+k}] = \lambda(1-\lambda)(1-\alpha-\beta)^k, \tag{5.22}$$

$$\mathrm{CORR}[\Xi_j, \Xi_{j+k}] = (1-\alpha-\beta)^k, \qquad k \geq 0. \tag{5.23}$$

Die Korrelationsfolge dieses Ersatzmodells für den Ausgangsprozeß ist also durch eine einfache geometrische Reihe gegeben.

5.2.3.2 Modell 2

5.2.3.2.1 Modellbeschreibung

Um die Betriebsperiode durch eine hypergeometrische Verteilung zu approximieren, wird ein MMBP mit insgesamt drei Zuständen benötigt (vgl. Bild 5.6): einem Pausenzustand (Zustand 1) sowie zwei aktiven Zuständen (Zustände 2 und 3), die den beiden Phasen der hypergeometrischen Verteilung entsprechen. Nach einer Freiperiode wird entweder Zustand 2 oder Zustand 3 durchlaufen, wobei die Verzweigungswahrscheinlichkeiten durch p_1 und $p_2 = 1 - p_1$ gegeben sind. Die mittlere Dauer der beiden aktiven Zustände ist $1/\beta_1$ bzw. $1/\beta_2$, mit β_1 und β_2 aus Gleichung 5.9 und 5.10.

Die Übergangsmatrix $\mathbf{P}$ der zugrundeliegenden Markov-Kette ist daher

$$\mathbf{P} = \begin{pmatrix} 1-\alpha & p_1\alpha & p_2\alpha \\ \beta_1 & 1-\beta_1 & 0 \\ \beta_2 & 0 & 1-\beta_2 \end{pmatrix}. \tag{5.24}$$

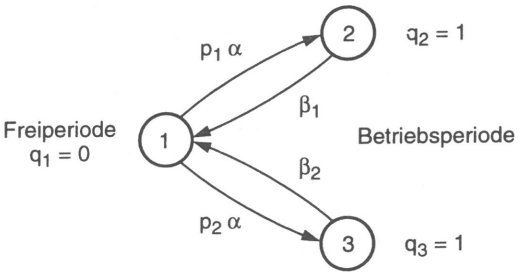


Bild 5.6: Modellierung des Ausgangsprozesses durch einen Markov-Modulierten Bernoulli-Prozeß mit drei Zuständen

Sowohl Zustand 2 als auch Zustand 3 sind aktive Zustände mit $q_2 = q_3 = 1$, der Vektor $\mathbf{q}$ ist somit

$$\mathbf{q} = (0, 1, 1)^T. \tag{5.25}$$

Für den stationären Zustandsvektor $\boldsymbol{\pi} = (\pi_1, \pi_2, \pi_3)$ der Markov-Kette ergibt sich hier

$$\boldsymbol{\pi} = \frac{1}{\alpha + \beta}\,(\beta, \frac{\alpha}{2}, \frac{\alpha}{2}) = \frac{1}{\mathrm{E}[B] + \mathrm{E}[I]}\,(\mathrm{E}[I], \frac{\mathrm{E}[B]}{2}, \frac{\mathrm{E}[B]}{2}). \tag{5.26}$$

Daraus berechnet sich die Ankunftsrate zu

$$\lambda = \boldsymbol{\pi}\,\mathbf{q} = \pi_1 + \pi_2 = \frac{\alpha}{\alpha + \beta}. \tag{5.27}$$

5.2.3.2.2 Korrelationsverhalten

Nach Gleichung 5.17 kann die Kovarianz der Zellankünfte mit einem Abstand von k Zeitschlitzen wie folgt berechnet werden:

$$\mathrm{COV}[\Xi_j, \Xi_{j+k}] = \mu_2^k a_2 + \mu_3^k a_3, \quad k \geq 1. \tag{5.28}$$

Dabei sind μ_2 und μ_3 die von 1 verschiedenen Eigenwerte der Übergangsmatrix $\mathbf{P}$. Bei einer 3×3-Matrix wird neben der Spur auch die Determinante zur Bestimmung der beiden unbekannten Eigenwerte benötigt. Allgemein gilt, daß die Determinante einer Matrix gleich dem Produkt ihrer Eigenwerte ist. Wegen $\mu_1 = 1$ gilt hier

$$1 + \mu_2 + \mu_3 = \mathrm{Spur}(\mathbf{P}), \tag{5.29}$$

$$\mu_2\mu_3 = \mathrm{Det}(\mathbf{P}). \tag{5.30}$$

Nach einigen Umformungen ergibt sich daraus

$$\mu_2 = 1 - \frac{\alpha}{2} - \beta + \sqrt{\left(\frac{\alpha}{2} - \beta\right)^2 - \beta_1\beta_2 + \alpha(p_1\beta_1 + p_2\beta_2)} \tag{5.31}$$

$$\mu_3 = 1 - \frac{\alpha}{2} - \beta - \sqrt{\left(\frac{\alpha}{2} - \beta\right)^2 - \beta_1\beta_2 + \alpha(p_1\beta_1 + p_2\beta_2)} \tag{5.32}$$

Die Berechnung der Konstanten a_2 und a_3 erfolgt am einfachsten unter Verwendung von $\mathrm{COV}[\Xi_j, \Xi_{j+1}]$ (siehe Gleichung 5.21) und $\mathrm{COV}[\Xi_j, \Xi_{j+2}]$. Eine explizite Berechnung von $\mathrm{COV}[\Xi_j, \Xi_{j+2}]$ liefert

$$\mathrm{COV}[\Xi_j, \Xi_{j+2}] = \lambda(1-\lambda)\left[(\alpha+\beta)\left(\alpha + \frac{(1-\beta_1)^2 + (1-\beta_2)^2}{2\beta}\right) - \frac{\alpha}{\beta}\right]. \tag{5.33}$$

Aus Gleichung 5.28 folgt außerdem

$$\begin{aligned} \mathrm{COV}[\Xi_j, \Xi_{j+1}] &= \mu_2 a_2 + \mu_3 a_3, \\ \mathrm{COV}[\Xi_j, \Xi_{j+2}] &= \mu_2^2 a_2 + \mu_3^2 a_3. \end{aligned}$$

Nach a_2 bzw. a_3 aufgelöst erhält man

$$a_2 = \frac{\mathrm{COV}[\Xi_j, \Xi_{j+2}] - \mu_3 \mathrm{COV}[\Xi_j, \Xi_{j+1}]}{\mu_2(\mu_2 - \mu_3)}, \tag{5.34}$$

$$a_3 = \frac{\mu_2 \mathrm{COV}[\Xi_j, \Xi_{j+1}] - \mathrm{COV}[\Xi_j, \Xi_{j+2}]}{\mu_3(\mu_2 - \mu_3)}. \tag{5.35}$$

Damit kann die Kovarianz- bzw. die Korrelationsfolge des Ausgangsprozesses auch für die hypergeometrische Approximation der Betriebsperiode analytisch berechnet werden.

Bild 5.7 zeigt einen Vergleich der beiden Modelle bezüglich der komplementären Verteilungsfunktion der Betriebsperiode für den Ausgangspuffer eines 2×2-Koppelelements der ersten Stufe. Die durchgezogene Kurve entspricht der nach Anhang A berechneten exakten Lösung,

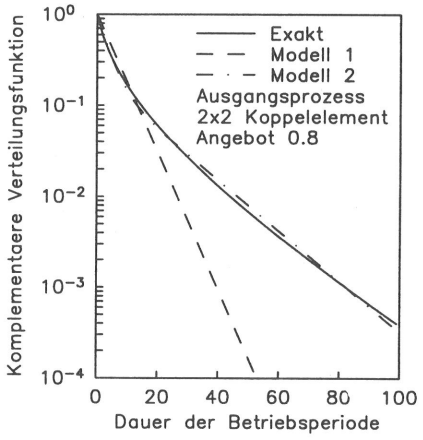


Bild 5.7: Komplementäre Verteilungsfunktion der Betriebsperiode eines Ausgangspuffers in Stufe 1 (Bernoulli-Ankunftsprozeß)

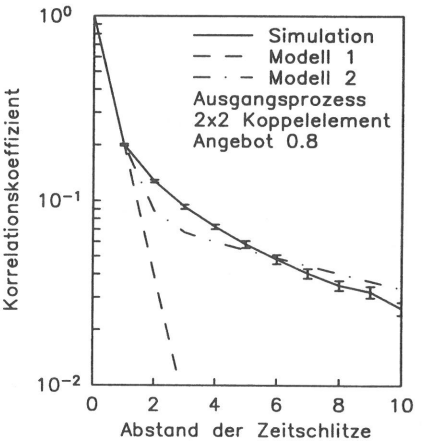


Bild 5.8: Korrelationsfolge des Ausgangsprozesses der ersten Stufe (Bernoulli-Ankunftsprozeß)

die anderen Kurven beziehen sich auf die beiden oben beschriebenen Modelle. Bei einem Verkehrsangebot von $A = 0.8$ weicht die Verteilung der Betriebsperiode bereits deutlich von einer geometrischen Verteilung mit gleichem Mittelwert (Modell 1) ab, während die hypergeometrische Verteilung mit der exakten Lösung nahezu übereinstimmt. Auch in [27] wurde bereits festgestellt, daß eine geometrische Verteilung nicht zur Modellierung der Betriebsperiode geeignet ist.

Deutliche Unterschiede zwischen den Modellen sind auch bei der in Bild 5.8 dargestellten Korrelationsfolge des Ausgangsprozesses zu erkennen. Die Approximation der Betriebsperiode durch eine geometrische Verteilung (Modell 1) führt zu einer starken Unterbewertung der Korrelation für Zellabstände größer als 1, so daß auch bei der Analyse der nachfolgenden Stufe entsprechende Ungenauigkeiten zu erwarten sind (siehe auch Bild 5.10). Die Korrelationsfolge der hypergeometrischen Approximation (Modell 2) zeigt ebenfalls gewisse Abweichungen von den Simulationsergebnissen, der prinzipielle Verlauf der Kurve wird jedoch wesentlich besser angenähert als bei Modell 1.

5.3 Ankunftsprozeß des Ausgangspuffers

Die an einem Eingang des Koppelelements eintreffenden Zellen sind gleichwahrscheinlich an einen der Ausgänge adressiert. Der Ankunftsprozeß des Koppelelements wird also zunächst in N Teilströme zerlegt, wobei jeder dieser Teilströme durch eine zufällige Auswahl von Zellen (jeweils mit Wahrscheinlichkeit $1/N$) aus dem Ankunftsprozeß hervorgeht. Die Zellankünfte an einem Ausgangspuffer sind wiederum die Überlagerung von insgesamt N solcher Teilströme und bilden daher einen Gruppenankunftsprozeß mit korrelierter Gruppengröße. Zunächst soll das Korrelationsverhalten der Gruppenankünfte aus der Korrelation des Ausgangsprozesses der vorherigen Stufe abgeleitet werden, und anschließend wird ein geeignetes Modell zur analytischen Behandlung des Gruppenankunftsprozesses beschrieben.

5.3.1 Korrelation der Gruppenankünfte

Die einzelnen Teilströme, in die der Ankunftsprozeß des Koppelelements zerlegt wird, können ebenfalls als MMBP modelliert werden. Wegen der zufälligen Verzweigung ergibt sich der Vektor $\mathbf{q}_{Split}$ eines Teilprozesses aus dem Vektor $\mathbf{q}$ des Ankunftsprozesses durch eine Multiplikation mit $1/N$, die Übergangsmatrix $\mathbf{P}$ der Markov-Kette bleibt jedoch unverändert. Unter Verwendung der Gleichungen C.5 und C.6 erhält man daher für einen Teilstrom

$$\begin{aligned} \mathrm{COV}_{Split}[\Xi_j, \Xi_{j+k}] &= \frac{1}{N^2}\,\mathrm{COV}[\Xi_j, \Xi_{j+k}], \\ \mathrm{VAR}_{Split}[\Xi] &= \frac{\lambda}{N}(1-\frac{\lambda}{N}) = \frac{1}{N^2}\frac{N-\lambda}{1-\lambda}\,\mathrm{VAR}[\Xi], \end{aligned}$$

wobei λ die Ankunftsrate der Zellen an den Eingängen des Koppelelements ist. Damit kann die Korrelationsfolge eines Teilprozesses sehr einfach aus der Korrelation des Ankunftsprozesses berechnet werden:

$$\mathrm{CORR}_{Split}[\Xi_j, \Xi_{j+k}] = \frac{1-\lambda}{N-\lambda}\,\mathrm{CORR}[\Xi_j, \Xi_{j+k}], \tag{5.36}$$

Diese Beziehung gilt unverändert auch für die zufällige Verzweigung von beliebigen zeitdiskreten Ankunftsprozessen mit Einzelankünften, da der Zusammenhang zwischen der Varianz bzw. Kovarianz der beiden Prozesse nicht von der Art des Ankunftsprozesses abhängt. Insbesondere ist Gleichung 5.36 daher auch auf den realen Ausgangsprozeß der vorhergehenden Stufe anwendbar.

Der Gruppenankunftsprozeß eines Ausgangspuffers ergibt sich aus der Überlagerung von N statistisch unabhängigen Teilprozessen. Es kann leicht gezeigt werden, daß die Korrelation der Gruppenankünfte mit der Korrelationsfolge eines einzelnen Teilprozesses identisch ist, so daß Gleichung 5.36 auch die Korrelation der Zellankünfte am Eingang eines Ausgangspuffers beschreibt.

Gleichung 5.36 bestätigt die Vermutung, daß die Korrelationen der Ankünfte durch die Aufspaltung in N Teilprozesse deutlich abgeschwächt werden. Die Abhängigkeit der Korrelation von der Größe des Koppelelements ist auch in Bild 5.9 zu erkennen, wobei die gezeigten Ergebnisse hier analytisch mit Hilfe der Approximation nach Modell 2 berechnet wurden. Mit zunehmendem Verzweigungsgrad werden die Korrelationen immer schwächer und können für große N schließlich vernachlässigt werden.

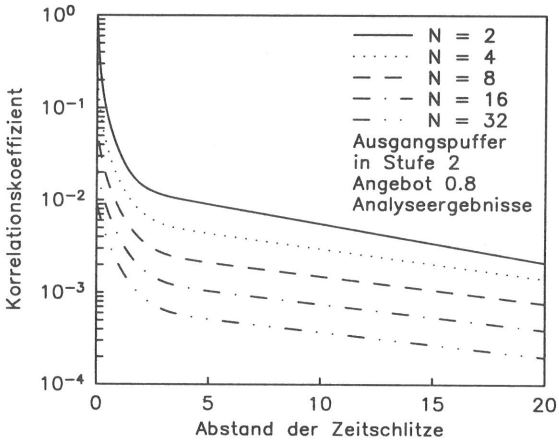


Bild 5.9: Korrelation der Gruppenankünfte eines Ausgangspuffers in Stufe 2

5.3.2 Modellierung des Ankunftsprozesses

Um eine analytische Behandlung des Ausgangspuffers zu ermöglichen, muß der Gruppenankunftsprozeß durch ein geeignetes Modell beschrieben werden. Ein solches Modell ist beispielsweise der sogenannte „Discrete-Time Batch Markovian Arrival Process" (D-BMAP) [24],[26], der ebenfalls zur Klasse der Markov-Ketten-gesteuerten Prozesse gehört.

Ähnlich wie der MMBP basiert der D-BMAP auf einer zeitdiskreten Markov-Kette mit n Zuständen. Die Ankunftsereignisse eines D-BMAP hängen direkt mit den Zustandsübergängen dieser Markov-Kette zusammen. Es sei $(d_k)_{ij}$ die Wahrscheinlichkeit für einen Übergang vom Zustand i in den Zustand j, der gleichzeitig eine Gruppenankunft von k Anforderungen auslöst. Zur Vereinfachung der Schreibweise werden die Übergangswahrscheinlichkeiten $(d_k)_{ij}$ zu den Matrizen $\mathbf{D}_k = [(d_k)_{ij}]$ zusammengefaßt, so daß die Matrix $\mathbf{D}_k$ alle Übergangswahrscheinlichkeiten enthält, die einer Gruppenankunft von k Anforderungen entsprechen. Die Übergangsmatrix der Markov-Kette ist daher $\mathbf{D} = \sum_{k=0}^{\infty} \mathbf{D}_k$.

Der Markov-Modulierte Bernoulli-Prozeß kann ebenfalls als D-BMAP dargestellt werden, wobei allerdings nur die Matrizen $\mathbf{D}_0$ und $\mathbf{D}_1$ von Null verschieden sind. Diese beiden Matrizen ergeben sich wie folgt aus den Parametern des MMBP:

$$\mathbf{D}_0 = (\mathbf{I} - \mathbf{Q})\,\mathbf{P}, \tag{5.37}$$

$$\mathbf{D}_1 = \mathbf{Q}\,\mathbf{P}, \tag{5.38}$$

mit $\mathbf{Q} = \mathrm{diag}(\mathbf{q})$.

Auch die Überlagerung von zwei statistisch unabhängigen D-BMAP-Prozessen ergibt wieder einen D-BMAP mit einem entsprechend größeren Zustandsraum. Die $n \times n$-Matrizen $\mathbf{D}_k^{(1)}$ und $\mathbf{D}_k^{(2)}$ sollen zwei unabhängige D-BMAP-Prozesse mit jeweils n Zuständen beschreiben. Die Überlagerung der beiden Prozesse führt auf einen D-BMAP mit $m = n^2$ Zuständen, dessen Übergangsmatrizen durch folgende Beziehung gegeben sind:

$$\mathbf{D}_k = \sum_{i=0}^{k} \mathbf{D}_i^{(1)} \otimes \mathbf{D}_{k-i}^{(2)}. \tag{5.39}$$

Dabei bezeichnet das Symbol „$\otimes$" das Kronecker-Produkt von zwei Matrizen. Gleichung 5.39 kann auch als eine verallgemeinerte Form der Faltung in Matrizenschreibweise aufgefaßt werden.

Ein Nachteil dieser formalen Vorgehensweise ist allerdings der exponentiell wachsende Zustandsraum des resultierenden Überlagerungsprozesses. Jede Überlagerung vergrößert den Zustandsraum um den Faktor n, so daß die unten beschriebene Analyse des Ausgangspuffers nur für kleine N möglich ist. Da in diesem Kapitel jedoch alle zu überlagernden Prozesse identisch sind, kann der Zustandsraum des resultierenden Ankunftsprozesses deutlich reduziert werden. Für eine eindeutige Beschreibung des Ankunftsprozesses auf der Basis von

Modell 2 genügt beispielsweise ein zweidimensionaler Zustandsraum, wobei jede Dimension die Anzahl der in Phase 1 bzw. 2 befindlichen Prozesse angibt. Auf diese Weise ergibt sich in einem $N \times N$-Koppelelement ein Zustandsraum mit $\frac{1}{2}(N+1)(N+2)$ Zuständen, der nur noch quadratisch mit N anwächst.

5.4 Analyse des Koppelnetzes

Mit Hilfe der oben beschriebenen Modellierung des internen Verkehrs können die einzelnen Stufen eines mehrstufigen Koppelnetzes nun schrittweise untersucht werden. Für die Modellierung des Ankunftsprozesses einer Stufe werden jeweils der Mittelwert von Betriebs- und Freiperiode sowie der Variationskoeffizient der Betriebsperiode in der vorhergehenden Stufe benötigt. Aus diesen Parametern kann der Gruppenankunftsprozeß des Ausgangspuffers nach Modell 1 oder Modell 2 abgeleitet und als D-BMAP dargestellt werden.

Die Modellierung des Ausgangspuffers führt somit auf ein D-BMAP/D/1-System mit endlicher Warteschlangenlänge, dessen Analyse in [24] und [26] beschrieben ist. Die Berechnung der Zustandswahrscheinlichkeiten des Ausgangspuffers beruht auch hier auf einer Markov-Kette zur Beschreibung des Systems, wobei der Systemzustand neben der Anzahl der gespeicherten Anforderungen auch den Zustand der Markov-Kette des Ankunftsprozesses umfaßt. Der Zustandsraum eines derartigen Modells ist daher wesentlich komplexer und auch größer als bei Systemen mit gedächtnisfreiem Ankunftsprozeß.

Die Zustandswahrscheinlichkeiten des Ausgangspuffers können mit Hilfe einer direkten numerischen Lösung des entsprechenden Gleichungssystems berechnet werden. Diese Vorgehensweise ist insbesondere bei Systemen mit relativ kleinem Zustandsraum durchaus effizient. Falls die Übergangsmatrix jedoch nicht mehr vollständig im Hauptspeicher des verwendeten Rechners Platz findet, steigt die zur Lösung des Gleichungssystems benötigte Rechenzeit sehr stark an. Dieser Nachteil kann mit dem in [26] angegebenen Verfahren vermieden werden, welches auf einer rekursiven Berechnung der Zustandswahrscheinlichkeiten beruht und ohne eine vollständige Speicherung der Übergangsmatrix auskommt. Die Einzelheiten dieses Analyseverfahrens sollen hier nicht wiederholt werden, der interessierte Leser sei auf [26] verwiesen.

Nachdem die Zustandswahrscheinlichkeiten sowie die charakteristischen Größen des Ausgangspuffers bekannt sind, kann auch die mittlere Dauer einer Freiperiode vergleichsweise einfach bestimmt werden. Außerdem gilt für die Auslastung ρ des Ausgangspuffers

$$\rho = \frac{\mathrm{E}[B]}{\mathrm{E}[I] + \mathrm{E}[B]}.$$

Die mittlere Betriebsperiode kann daher ebenfalls sofort angegeben werden. Die Berechnung des zweiten Moments der Betriebsperiode ist allerdings etwas aufwendiger, hier muß auf

die Darstellung in Form einer Phasenverteilung entsprechend Anhang A.3.4.1 zurückgegriffen werden. Zur Bestimmung des zweiten Moments muß die Übergangsmatrix der Phasenverteilung zudem invertiert werden, so daß auch hier ein erheblicher numerischer Aufwand erforderlich ist.

Mit Hilfe der Zustandswahrscheinlichkeiten läßt sich auch die Durchlaufzeitverteilung der Zellen in der betrachteten Stufe bestimmen. Falls die Durchlaufzeit einer Zelle statistisch unabhängig von der Durchlaufzeit in den anderen Stufen ist, erhält man die Verteilung der gesamten Durchlaufzeit näherungsweise aus einer Faltung der Durchlaufzeitverteilungen der einzelnen Stufen. Die Ergebnisse im folgenden Abschnitt zeigen jedoch, daß diese Näherungsannahme zu einer Unterschätzung der Durchlaufzeitschwankungen führt.

5.5 Ergebnisse

5.5.1 Verteilung der Warteschlangenlänge

Zunächst soll die Genauigkeit der beiden in Abschnitt 5.2 beschriebenen Modelle anhand der komplementären Verteilungsfunktion der Warteschlangenlänge in der zweiten Stufe eines Koppelnetzes untersucht werden. Wegen der deutlichen Zunahme der Warteschlangenlänge von der ersten zur zweiten Stufe werden dabei 2×2-Koppelelemente bei einem Angebot von 80% verwendet. Ein Vergleich der in Bild 5.10 dargestellten Ergebnisse zeigt, daß die Approximation der Betriebsperiode durch eine geometrische Verteilung (Modell 1) zu einer deutlichen Unterschätzung der Warteschlangenlänge führt. Dagegen stimmen die Ergebnisse von Modell 2 (hypergeometrische Verteilung der Betriebsperiode) nahezu exakt mit den Simulationsdaten überein. Die Unterschiede zwischen den beiden Modellen sind auf die wesentlich genauere Berücksichtigung des Korrelationsverhaltens bei Modell 2 zurückzuführen (siehe auch Bild 5.8). Für die weiteren Untersuchungen wird daher nur noch Modell 2 verwendet.

In Bild 5.11 ist neben den Ergebnissen für die ersten beiden Stufen auch die komplementäre Verteilungsfunktion der Warteschlangenlänge in der sechsten Stufe eingetragen. Auch hier liefert das Modell 2 noch sehr genaue Ergebnisse, die analytisch berechneten Werte liegen nur geringfügig unter den entsprechenden Simulationsergebnissen. Die approximative Analyse mit Hilfe von Modell 2 ist also durchaus für die Dimensionierung der Pufferspeicher in den hinteren Stufen eines mehrstufigen Koppelnetzes geeignet.

Die Ergebnisse in Bild 5.2 und 5.11 zeigen, daß sich die Verteilung der Warteschlangenlänge in den hinteren Stufen des Koppelnetzes kaum mehr ändert. Es stellt sich also die Frage nach einer „Grenzverteilung“, welche die Verteilung der Warteschlangenlänge in den hinteren Stufen nach oben hin begrenzt. Genauere Untersuchungen weisen darauf hin, daß der Ausgangsprozeß der internen Koppelelemente sehr gut durch den Ausgangsprozeß eines

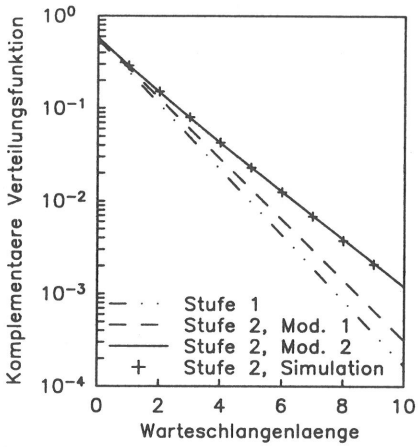


Bild 5.10: Vergleich der beiden analytischen Modelle mit Simulationsergebnissen (2×2-Koppelelemente, Angebot 80%)

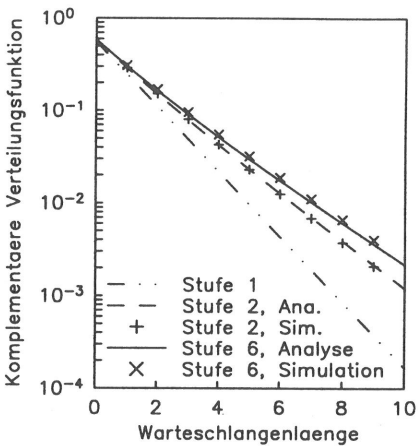


Bild 5.11: Analyse (Modell 2) und Simulation der Verteilung der Warteschlangenlänge in einem 6-stufigen Koppelnetz

getakteten M/D/1-S-Systems angenähert werden kann. Diese Approximation ist nicht nur auf 2×2-Koppelelemente beschränkt, sondern besitzt auch bei $N \times N$-Koppelelementen mit $N > 2$ eine hohe Genauigkeit. Für den Grenzfall $N \to \infty$ ist diese Näherung außerdem exakt, da hier die Korrelation der Gruppenankünfte verschwindet (siehe Gleichung 5.36), so daß die Ausgangspuffer in jeder Stufe durch ein M/D/1-S-System modelliert werden können. Der Ausgangsprozeß des M/D/1-S-Systems wird im folgenden auch als „Worst-Case"-Verkehr bezeichnet.

Ein Vergleich der unter Verwendung des „Worst-Case"-Verkehrs gewonnenen Ergebnisse mit den entsprechenden Verteilungen für koppelnetzinterne Stufen ist in Bild 5.12 dargestellt. Für $N = 2$ konnten die Ergebnisse des „Worst-Case"-Verkehrs analytisch mit Hilfe von Modell 2 berechnet werden, für $N > 2$ war jedoch nur eine simulative Untersuchung möglich. In allen Fällen ist eine sehr gute Übereinstimmung zwischen den mit einem „+" markierten Ergebnissen des „Worst-Case"-Verkehrs und der in einer hinteren Stufe des Koppelnetzes gemessenen Verteilung der Warteschlangenlänge zu verzeichnen.

Ein Beispiel für die Dimensionierung der internen Puffer ist in Tabelle 5.1 angegeben. Es handelt sich hierbei um diejenige Puffergröße, die benötigt wird, um bei einem Angebot von 80% eine Zellverlustwahrscheinlichkeit von 10^{-10} zu unterschreiten. Die Dimensionierung wurde sowohl bei Bernoulli-Verkehr als auch für den oben beschriebenen „Worst-Case"-Verkehr durchgeführt, wobei die Ergebnisse des „Worst-Case"-Verkehrs für $N > 2$ durch eine Extrapolation der entsprechenden Simulationsdaten gewonnen wurden.

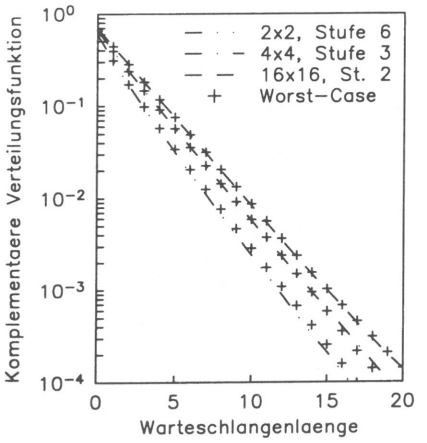


Bild 5.12: Verteilung der Warteschlangenlänge in koppelnetzinternen Stufen und bei „Worst-Case"-Verkehr (Angebot 80%)

N	Bernoulli-Verkehr	Worst-Case	Relative Zunahme
2	27	44	63%
4	39	49	26%
8	45	50	11%
16	48	51	6%
∞	51	51	0%

Tabelle 5.1: Dimensionierung der Puffergröße für Bernoulli-Verkehr und für den „Worst-Case"-Verkehr (Angebot 80%)

Die Berücksichtigung der internen Verkehrscharakteristik ist insbesondere für kleine N von sehr großer Bedeutung, da eine Dimensionierung auf der Basis des Bernoulli-Verkehrs hier zu einer deutlichen Unterschätzung des Pufferbedarfs führen würde. Ab $N = 16$ kann auch die Dimensionierung der Ausgangspuffer von internen Koppelelementen mit guter Näherung unter Verwendung des M/D/1-S-Systems erfolgen.

5.5.2 Verteilung der Durchlaufzeit

Die Verteilung der Durchlaufzeit in mehrstufigen Koppelnetzen wird für die Bestimmung der Durchlaufzeitschwankungen in ATM-Netzen benötigt. Um eine möglichst exakte Berechnung der Durchlaufzeitverteilung zu ermöglichen, muß sowohl das Verhalten der einzelnen Stufen wie auch die Korrelation der Durchlaufzeit in aufeinanderfolgenden Stufen bekannt sein. Aus den oben beschriebenen Ergebnissen geht bereits hervor, daß die Verteilung der Warteschlangenlänge und somit auch die Durchlaufzeit in den hinteren Stufen des Koppelnetzes durch die Analyse leicht unterschätzt wird.

Diese Tatsache wird auch durch die Gegenüberstellung von Analyse- und Simulationsergebnissen in Tabelle 5.2 bestätigt. Die Tabelle zeigt den Mittelwert sowie den Variationskoeffizienten der Durchlaufzeit für die einzelnen Stufen eines 6-stufigen Koppelnetzes, welches aus 2×2-Koppelelementen besteht. Abgesehen von der ersten Stufe liefert die Analyse etwas zu

	Mittelwert		Variationskoeffizient	
	Analyse	Simulation	Analyse	Simulation
Stufe 1	2.000	$2.000 \pm 2.9 \cdot 10^{-3}$	0.632	$0.632 \pm 2.1 \cdot 10^{-3}$
Stufe 2	2.187	$2.196 \pm 3.5 \cdot 10^{-3}$	0.708	$0.707 \pm 2.1 \cdot 10^{-3}$
Stufe 3	2.239	$2.260 \pm 3.7 \cdot 10^{-3}$	0.731	$0.732 \pm 3.4 \cdot 10^{-3}$
Stufe 4	2.254	$2.283 \pm 3.9 \cdot 10^{-3}$	0.738	$0.741 \pm 2.9 \cdot 10^{-3}$
Stufe 5	2.259	$2.292 \pm 5.3 \cdot 10^{-3}$	0.740	$0.745 \pm 3.8 \cdot 10^{-3}$
Stufe 6	2.260	$2.296 \pm 5.5 \cdot 10^{-3}$	0.740	$0.747 \pm 5.5 \cdot 10^{-3}$

Tabelle 5.2: Mittelwert und Variationskoeffizient der Durchlaufzeitverteilung bei einem Angebot von $A = 0.8$ (Vergleich von Analyse und Simulation für 2×2-Koppelelemente)

kleine Werte für die mittlere Durchlaufzeit sowie für den Variationskoeffizienten. Ein Vergleich mit den in [125] enthaltenen Ergebnissen zeigt eine gute Übereinstimmung, wobei die Genauigkeit der hier vorgestellten Analyse jedoch wesentlich höher ist als bei der in [125] beschriebenen empirischen Approximation.

Bild 5.13 zeigt die Durchlaufzeitverteilung für mehrstufige Koppelnetze mit unterschiedlicher Stufenzahl, die ebenfalls aus 2×2-Koppelelementen aufgebaut sind. Wie erwartet liegen die analytisch gewonnenen Werte unter den Simulationsergebnissen, wobei die Abweichungen hier deutlich größer sind als bei der Verteilung der Warteschlangenlänge in den einzelnen Stufen. Selbst bei einem zweistufigen Koppelnetz unterscheiden sich Analyse- und Simulationsergebnisse in ihrem asymptotischen Verhalten, obwohl die Analyse für die einzelnen Stufen jeweils sehr genaue Werte liefert. Dies läßt darauf schließen, daß die Durchlaufzeiten in aufeinanderfolgenden Stufen nicht statistisch unabhängig sind, sondern vielmehr eine leichte positive Korrelation aufweisen. Die Abweichung zwischen Simulation und Analyse nimmt außerdem von Stufe zu Stufe zu, da jede Faltung zu einer weiteren Vergrößerung des Fehlers beiträgt.

Das Auftreten einer positiven Korrelation kann mit Hilfe der stoßartigen Charakteristik des internen Verkehrs erklärt werden. Um den Zusammenhang zwischen Verkehrscharakteristik und Korrelation zu veranschaulichen, wird eine Zelle auf ihrem Weg durch das Koppelnetz beobachtet. Wenn diese Zelle in einer Stufe lange warten muß, dann ist auch die Dauer der Betriebsperiode bis zum Zeitpunkt der Übertragung der Zelle bereits sehr groß. Eine lange Betriebsperiode führt jedoch zu einem Anwachsen der Warteschlangen in der folgenden Stufe, so daß die betrachtete Zelle dort ebenfalls länger warten muß. Umgekehrt geht einer kurzen Wartezeit mit großer Wahrscheinlichkeit auch eine kurze Betriebsperiode voraus, so daß insgesamt eine positive Korrelation der Durchlaufzeiten in aufeinanderfolgenden Stufen vorliegt.

Ein ähnlicher Effekt wurde bereits in [129] beschrieben; dort wird auch eine approximative Analyse für den Fall einer geometrisch verteilten Aktivitätsdauer angegeben.

Bild 5.14 zeigt den Einfluß des Verkehrsangebots auf die Verteilung der Durchlaufzeit in einem 6-stufigen Koppelnetz. Ähnlich wie bei der Verteilung der Warteschlangenlänge führt eine Erhöhung des Angebots von 80% auf 90% zu einer Verdoppelung der Durchlaufzeitschwankungen. Während die Genauigkeit der Analyse bei einem Angebot von 50% noch zufriedenstellend ist, treten bei einem Angebot von 90% bereits größere Abweichungen auf. Die Zunahme des Fehlers hängt mit einer Zunahme der positiven Korrelation der Durchlaufzeiten zusammen, die wiederum auf die längere Dauer der Betriebsperioden zurückzuführen ist.

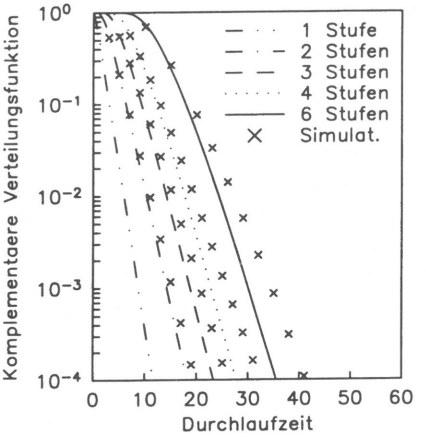


Bild 5.13: Verteilung der Durchlaufzeit in Koppelnetzen mit verschiedener Stufenzahl (2×2-Koppelelemente, Angebot 80%)

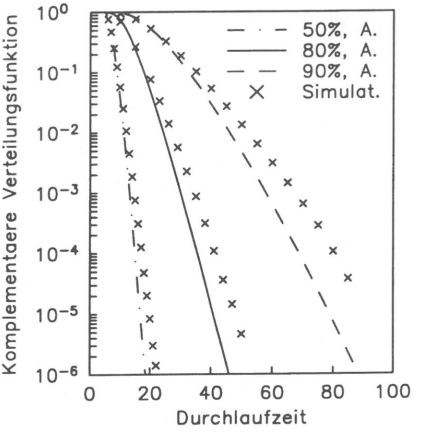


Bild 5.14: Verteilung der Durchlaufzeit in einem 6-stufigen Koppelnetz (2×2-Koppelelemente)

Kapitel 6

Verfeinerte Modellierung des Quellverhaltens

Für einen Vergleich der prinzipiellen Eigenschaften verschiedener Koppelnetzarchitekturen ist eine einfache Modellierung des Verkehrs völlig ausreichend. Die Dimensionierung der Netzkomponenten wie auch die Definition wirksamer Verkehrssteuerungsverfahren erfordert jedoch ein genaueres Verständnis dafür, welche Auswirkungen die Eigenschaften der Verkehrsquellen auf das Verhalten der Netzelemente haben. Neben der Modellierung der Quellen ist in diesem Zusammenhang auch die Berücksichtigung des verbindungsorientierten Betriebs von großer Bedeutung.

Beide Aspekte werden in diesem Kapitel am Beispiel der Sporadischen Quelle ausführlich untersucht. Nach einer Beschreibung des verwendeten Quellmodells wird zunächst der prinzipielle Einfluß verschiedener Parameter auf das Verhalten von Koppelelementen mit Ausgangspuffern diskutiert. Es folgt ein Vergleich von mehreren einstufigen Koppelanordnungen, wobei die Sensitivität der einzelnen Varianten gegenüber der Quellcharakteristik im Vordergrund steht. Schließlich werden auch die Auswirkungen einer Vernachlässigung des verbindungsorientierten Betriebs bei der Leistungsuntersuchung aufgezeigt.

6.1 Ein Quellmodell zur gemeinsamen Modellierung von Zell- und Burst-Ebene

Die bereits in Kapitel 3 eingeführte Sporadische Quelle eignet sich besonders für die gleichzeitige Modellierung von Verkehrsquellen auf Zell- und Burstebene. Mit Hilfe der Sporadischen Quelle kann insbesondere der Einfluß der Quellcharakteristik auf Burstebene auf das Verhalten der Netzkomponenten untersucht werden. In diesem Kapitel soll die zeitdiskrete Version

der Sporadischen Quelle verwendet werden, die auch als Sonderfall des GMDP (siehe Abschnitt 3.1.4) mit einem aktiven Zustand und einem Pausenzustand aufgefaßt werden kann. Das Verhalten der Sporadischen Quelle wird durch die folgenden Parameter beschrieben:

- Es sei X die Anzahl der Zellen, die in einer Aktivitätsphase gesendet werden. Sofern keine anderen Angaben gemacht werden, soll X eine geometrische Verteilung mit Mittelwert $\mathrm{E}[X]$ besitzen, wobei jede Aktivitätsphase mindestens eine Zelle enthält.
- Innerhalb einer Aktivitätsphase haben die Zellen einen konstanten Abstand von d Zeitschlitzen.
- Die Dauer der Pausenphase beträgt S Zeitschlitze und besitzt eine geometrische Verteilung mit Mittelwert $\mathrm{E}[S]$. Auch die Pausenphase ist mindestens einen Zeitschlitz lang.

Unter der Voraussetzung einer geometrisch verteilten Burst-Länge ist der resultierende Ankunftsprozeß ein Erneuerungsprozeß. Ein Beispiel für den Ankunftsprozeß einer Sporadischen Quelle ist in Bild 6.1 dargestellt.

Im folgenden sollen einige grundlegende Kenngrößen der Sporadischen Quelle abgeleitet werden. Die Wahrscheinlichkeit p_a, daß sich die Quelle im aktiven Zustand befindet, ist

$$p_a = \frac{d \cdot \mathrm{E}[X]}{d \cdot \mathrm{E}[X] + \mathrm{E}[S]}. \tag{6.1}$$

Daraus ergibt sich die Ankunftsrate λ zu

$$\lambda = p_a \frac{1}{d} = \frac{\mathrm{E}[X]}{d \cdot \mathrm{E}[X] + \mathrm{E}[S]}, \tag{6.2}$$

und der mittlere Ankunftsabstand $\mathrm{E}[T_A]$ ist

$$\mathrm{E}[T_A] = \frac{1}{\lambda} = d + \frac{\mathrm{E}[S]}{\mathrm{E}[X]} \tag{6.3}$$

Mit Hilfe der in [204] angegebenen Beziehungen können auch die höheren Momente des Ankunftsabstands berechnet werden.

Alle Untersuchungen in diesem Kapitel basieren auf den in Tabelle 6.1 angegebenen Quellparametern. Die einzelnen Parametersätze unterscheiden sich durch die Dauer von Aktivitäts- und Pausenphasen, in einigen Fällen wurde auch der Abstand d der Zellen in einer Aktivitätsphase variiert. Alle Parametersätze weisen jedoch eine gemeinsame Ankunftsrate von $\lambda = 0.02$ auf.

Bild 6.1: Beispiel für den Ankunfsprozeß der Zellen einer Sporadischen Quelle

E[X]	E[S]	d	λ
2	80	10	0.02
5	200	10	0.02
10	400	10	0.02
50	2000	10	0.02
50	1750	15	0.02
50	1500	20	0.02
500	20000	10	0.02

Tabelle 6.1: Die verwendeten Parametersätze der Sporadischen Quelle

6.2 Modellierung von virtuellen Verbindungen

Bei der Leistungsuntersuchung von ATM-Koppelnetzen wird meist von einer zufälligen Verzweigung des internen Verkehrs ausgegangen, wobei die Verzweigungswahrscheinlichkeiten den Verkehrsanteilen in der jeweiligen Richtung entsprechen. Die zufällige Verzweigung kann dabei entweder auf die Zellebene oder auf die Burst-Ebene bezogen werden. Im einen Fall wird für jede einzelne Zelle ein Ausgang zufällig bestimmt, im anderen Fall wird für alle Zellen einer Aktivitätsphase ein gemeinsames Ziel zufällig ausgewählt. Beide Ansätze stellen jedoch nur eine Näherung für den verbindungsorientierten Betrieb von ATM-Netzen dar. In Abschnitt 6.5 wird gezeigt, daß die Vernachlässigung des verbindungsorientierten Betriebs zu einer groben Fehleinschätzung des Verhaltens der Netzkomponenten führen kann.

In diesem Kapitel wird der verbindungsorientierte Charakter der ATM-Technik durch eine geeignete Modellierung des Verkehrsflusses vollständig berücksichtigt. Zu diesem Zweck ist jeder einzelnen Verkehrsquelle ein festes Ziel zugeordnet, zu dem alle Zellen dieser Quelle gesendet werden. Jede Zelle führt daher ein Adreßfeld mit sich, welches diese Zieladresse in einem speziellen, für eine schnelle Auswertung geeigneten Format enthält. Um eine gleichmäßige Belastung der gesamten Koppelanordnung zu erreichen, werden an jedem Eingang mehrere Verkehrsquellen mit unterschiedlichen Zielausgängen benötigt. Wie in Bild 6.2 dargestellt, sind diese Quellen zunächst mit einem Multiplexer verbunden, der zur Serialisierung des Ankunftsprozesses dient. Der Multiplexer wird als eine Warteschlange mit FIFO-Abfertigung modelliert, wobei durch die Dimensionierung der Puffergröße sichergestellt ist, daß keine oder nur verschwindend geringe Zellverluste auftreten. Der Ausgangsprozeß des Multiplexers ist bereits auf den internen Takt des Koppelnetzes synchronisiert und entspricht einer getakteten ATM-Leitung.

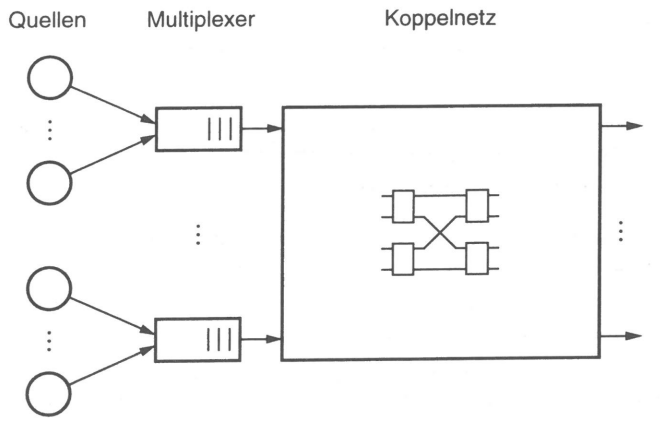


Bild 6.2: Simulationsmodell zur Modellierung von virtuellen Verbindungen

Bei den folgenden Untersuchungen befinden sich an jedem Eingang jeweils 32 Quellen, so daß die Eingangsleitungen ein Verkehrsangebot von $A = 0.64$ aufweisen. Die Zieladressen der Quellen an einem Eingang werden so gewählt, daß jeder Ausgang genau n-mal auftritt, wobei sich n aus der Division von 32 durch die Anzahl der Koppelnetzausgänge ergibt. Auf diese Weise können Koppelanordnungen mit bis zu 32 Ausgängen untersucht werden, wobei die Anzahl der Ausgänge allerdings eine Zweierpotenz sein muß.

6.3 Einfluß von stoßartigem Verkehr auf das Verhalten eines Ausgangspuffers

In diesem Abschnitt sollen zunächst einige grundlegende Untersuchungen an Koppelelementen mit Ausgangspuffern durchgeführt werden, um den Einfluß von stoßartigem Verkehr auf das Verhalten der Koppelanordnung zu zeigen. Aus verschiedenen Studien in der Literatur ist bereits bekannt, daß die Verlustwahrscheinlichkeit bei stoßartigem Verkehr um mehrere Größenordnungen höher sein kann als bei Bernoulli-Verkehr (siehe z.B. [27], [49], [156], [185]).

Bild 6.3 zeigt die Abhängigkeit der Zellverlustwahrscheinlichkeit von der Größe eines Ausgangspuffers für ein 16×16-Koppelelement, wobei der Verkehr aus der Überlagerung von 32 Sporadischen Quellen mit einer mittleren Burst-Länge von 50 Zellen und einer mittleren Pausendauer von 2000 Zeitschlitzen besteht. Die Simulationsergebnisse für die Zellverlustwahrscheinlichkeit (durchgezogene Kurve) zeigen einen charakteristischen Verlauf, der in zwei Bereiche unterteilt werden kann:

- Für kleine Werte der Puffergröße nimmt die Verlustwahrscheinlichkeit sehr schnell ab, wobei der Kurvenverlauf in diesem Bereich gut durch die Ergebnisse bei Bernoulli-Verkehr angenähert werden kann.
- Mit zunehmender Puffergröße knickt die Kurve plötzlich ab und geht in einen wesentlich flacheren Verlauf über. In diesem Bereich nimmt die Verlustwahrscheinlichkeit mit steigender Puffergröße nur sehr wenig ab.

Diese charakteristische Abhängigkeit der Verlustwahrscheinlichkeit von der Puffergröße ist auch in der Literatur bereits beobachtet und untersucht worden [13], [126], [127], [173]. Das scharfe Abknicken der Kurve ist vor allem auf das Auftreten von Überlastzuständen auf Burst-Ebene zurückzuführen [13]. Bei den hier gewählten Parametern beträgt der Zellabstand in der Aktivitätsphase einer Quelle $d = 10$ Zeitschlitze, so daß bei 10 aktiven Quellen ein momentanes Angebot von $A = 1$ erreicht wird. Falls mehr als 10 Quellen gleichzeitig im aktiven Zustand sind, befindet sich der Ausgangspuffer daher in einer Überlastsituation, wobei kurzzeitig sehr hohe Zellverluste auftreten.

Das Verhalten des Ausgangspuffers hängt sowohl von der Auftrittswahrscheinlichkeit einer derartigen Überlastsituation als auch von deren Dauer ab. Die Auftrittswahrscheinlichkeit einer Überlastsituation kann aus den Zustandswahrscheinlichkeiten der Quellen auf Burst-Ebene berechnet werden. Es sei p_i die Wahrscheinlichkeit, daß i von insgesamt N Sporadischen Quellen aktiv sind. Da die einzelnen Quellen voneinander unabhängig sind, ergibt sich für die Zustandsverteilung auf Burst-Ebene die folgende Binomialverteilung

$$p_i = \binom{N}{i} p_a^i \, (1 - p_a)^{N-i}, \tag{6.4}$$

wobei p_a die Wahrscheinlichkeit ist, daß eine einzelne Quelle aktiv ist (siehe Gleichung 6.2). Der Ausgangspuffer befindet sich in einer Überlast, wenn mehr als $i = d$ Quellen gleichzeitig aktiv sind. Die Wahrscheinlichkeit für das Auftreten einer Überlast erhält man folglich durch Aufsummieren aller Zustandswahrscheinlichkeiten für $i > d$.

Die Überlastwahrscheinlichkeit kann als obere Grenze für die Approximation der Verlustwahrscheinlichkeit auf Burst-Ebene verwendet werden. Dabei wird allerdings unabhängig von der Puffergröße angenommen, daß mit dem Auftreten einer Überlast alle ankommenden Zellen sofort verloren gehen. In einer Überlastsituation geht jedoch nur derjenige Teil der ankommenden Zellen verloren, der die Kapazität der Leitung übersteigt. Eine bessere Näherung erhält man daher durch eine genauere Berücksichtigung der Verluste in den einzelnen Überlastzuständen [145]:

$$\mathrm{P}\{\text{Zellverlust}\} \approx \frac{1}{\lambda} \sum_{i=d+1}^{N} \frac{i-d}{d} \, p_i. \tag{6.5}$$

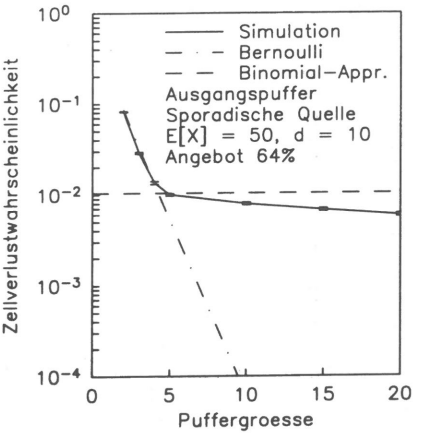


Bild 6.3: Verlustwahrscheinlichkeit eines Ausgangspuffers bei stoßartigem Verkehr

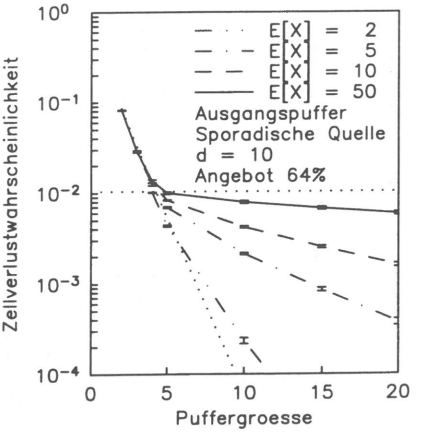


Bild 6.4: Einfluß der mittleren Burst-Länge auf die Zellverlustwahrscheinlichkeit

Diese aus der Binomialverteilung abgeleitete Näherung für die Verlustwahrscheinlichkeit ist in Bild 6.3 als gestrichelte, horizontale Linie eingezeichnet. Der Schnittpunkt dieser Linie mit der Kurve für Bernoulli-Verkehr gibt die Stelle an, an welcher der Kurvenverlauf abknickt. Für größere Pufferlängen ist die tatsächliche Verlustwahrscheinlichkeit jedoch geringer, da die Überlastsituation teilweise durch das Zwischenspeichern von Zellen aufgefangen werden kann. Die Steigung der Kurve hängt in diesem Bereich ganz wesentlich von der Dauer der Überlast ab, welche wiederum mit der Aktivitätsdauer der Quellen zusammenhängt.

Der Einfluß der mittleren Burst-Länge auf die Steigung der Kurven ist in Bild 6.4 deutlich zu erkennen. Während die Verlustwahrscheinlichkeit bei einer mittleren Burst-Länge von 2 Zellen nur geringfügig höher als bei Bernoulli-Verkehr (gepunktete Linie) ist, verläuft die Kurve bei einer mittleren Burst-Länge von 50 Zellen bereits nahezu horizontal. Mit zunehmender Burst-Länge nähert sich die Kurve immer mehr der ebenfalls gepunktet eingezeichneten Approximation nach Gleichung 6.5, da die Überlast kaum noch durch den Ausgangspuffer aufgefangen werden kann.

Die Zellverlustwahrscheinlichkeit hängt jedoch nicht nur von der mittleren Burst-Länge ab, auch die Verteilung der Burst-Länge hat einen großen Einfluß auf die Steigung im hinteren Teil der Kurve. Bild 6.5 zeigt die Verlustwahrscheinlichkeit für eine mittlere Burst-Länge von 5 Zellen, wobei die Anzahl der Zellen in einem Burst entweder konstant (Variationskoeffizient 0), geometrisch verteilt oder negativ-binomial verteilt ist. Für den Fall der negativ-binomial verteilten Burst-Länge wurde dabei ein Variationskoeffizient von $c = 2$ gewählt. Wie erwartet ergeben sich die geringsten Verluste bei einer konstanten Burst-Länge, während die

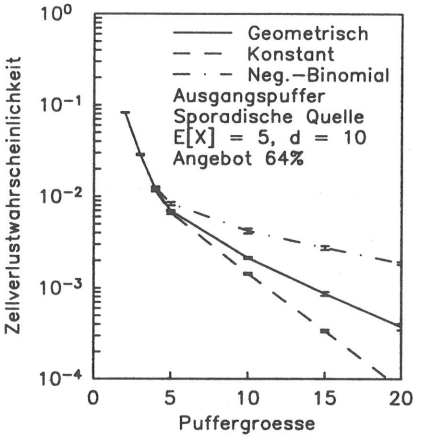


Bild 6.5: Einfluß der Verteilung der Burst-Länge auf die Verlustwahrscheinlichkeit

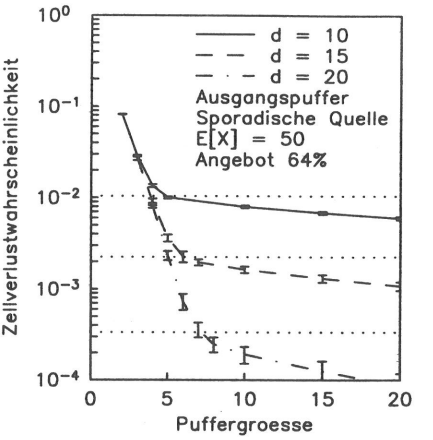


Bild 6.6: Abhängigkeit der Verlustwahrscheinlichkeit von der Spitzenbitrate der Quellen

Verlustwahrscheinlichkeit mit zunehmendem Variationskoeffizienten der Burst-Länge ebenfalls ansteigt.

In Bild 6.6 ist schließlich die Abhängigkeit der Verlustwahrscheinlichkeit von der Spitzenbitrate der Quellen dargestellt. Die mittlere Burst-Länge beträgt jeweils 50 Zellen, so daß die Abnahme der Spitzenbitrate durch eine Reduzierung der mittleren Pausendauer kompensiert werden muß, um die Ankunftsrate konstant zu halten (vgl. Tabelle 6.1). Eine Verringerung der Spitzenbitrate wirkt sich sowohl auf die Wahrscheinlichkeit für das Auftreten einer Überlast als auch auf deren Dauer aus. Beispielsweise müssen bei einem Zellabstand von $d = 20$ Zeitschlitzen bereits mehr als 20 Quellen gleichzeitig aktiv sein, um eine Überlast zu erzeugen, so daß die Überlastwahrscheinlichkeit entsprechend abnimmt. In Bild 6.6 ist die Abnahme der Überlastwahrscheinlichkeit an der Verschiebung des flachen Kurvenabschnitts hin zu kleineren Wahrscheinlichkeiten zu erkennen. Die kürzere Dauer der Überlastsituationen führt außerdem zu einem etwas steileren Kurvenverlauf, da die Überlast mit zunehmender Puffergröße eher aufgefangen werden kann.

Die in diesem Abschnitt vorgestellten Ergebnisse zeigen, daß die Charakteristik der Verkehrsquellen auf Burst-Ebene einen starken Einfluß auf das Verhalten der Ausgangspuffer hat. Die Verlustwahrscheinlichkeit hängt dabei nicht nur von der mittleren Zustandsdauer auf Burst-Ebene ab, auch die Verteilung der Burst-Länge muß berücksichtigt werden. Daraus ergeben sich einige Konsequenzen für die Verbindungsannahme (siehe 1.1.6.1) sowie die Pufferdimensionierung in ATM-Netzen [126]:

- Eine genaue Abschätzung der Verlustwahrscheinlichkeit erfordert eine möglichst genaue Kenntnis des Quellverhaltens auf Burst-Ebene. Neben der Zellrate in den einzelnen Zuständen wird dabei auch die Verteilung der Zustandsdauer benötigt. Im allgemeinen hängen diese Parameter jedoch sehr stark vom Verhalten der Benutzer ab und können deshalb beim Verbindungsaufbau noch nicht festgelegt werden. Hinzu kommt, daß bisher kein Verfahren zur Quellflußkontrolle bekannt ist, welches eine wirksame Überwachung dieser Parameter erlaubt. Aus diesem Grund ist die Verwendbarkeit von Verkehrsparametern wie der Verteilung der Burst-Länge oder deren Mittelwert fraglich.
- Die mittlere Bitrate ist der einzige Parameter, der in bestimmten Fällen außer der Spitzenbitrate angegeben werden kann. Mit Hilfe der Spitzenbitrate sowie der mittleren Bitrate können zumindest die Zustandswahrscheinlichkeiten einer virtuellen Sporadischen Quelle bestimmt werden, die als „Worst-Case"-Approximation für das tatsächliche Quellverhalten dienen kann. Aus Gleichung 6.5 erhält man auf diese Weise eine konservative Abschätzung der Zellverlustwahrscheinlichkeit (siehe auch Bild 6.3). Für den allgemeinen Fall müssen die Zustandswahrscheinlichkeiten p_i aus einer Faltung der Zustandsverteilungen der einzelnen Quellen berechnet werden, so daß dieses Verfahren auch als Faltungsansatz bezeichnet wird.
- Eine Verbindungsannahme auf der Basis des oben beschriebenen Faltungsansatzes stellt sicher, daß die Wahrscheinlichkeit für das Auftreten einer Überlast auf Burst-Ebene kleiner als die vorgegebene Zellverlustwahrscheinlichkeit ist. Der charakteristische Knick im Kurvenverlauf tritt daher erst unterhalb der vorgegebenen Zellverlustwahrscheinlichkeit auf und ist somit bedeutungslos. Oberhalb des Knicks kann der Kurvenverlauf jedoch sehr gut durch die Verlustwahrscheinlichkeit bei Bernoulli-Verkehr approximiert werden, so daß eine Dimensionierung der Puffer mit Hilfe des Bernoulli-Verkehrs unter den genannten Voraussetzungen durchaus gerechtfertigt ist. Wegen der starken Zunahme der Verlustwahrscheinlichkeit mit steigender Burst-Länge würde eine Berücksichtigung dieser Parameter bei der Pufferdimensionierung ohnehin zu sehr großen Pufferspeichern führen, die kaum realisierbar wären.

Aus der Literatur sind auch verschiedene Ansätze zur approximativen analytischen Behandlung eines Ausgangspuffers bzw. statistischen Multiplexers mit stoßartigem Eingangsverkehr bekannt (siehe z.B. [13], [102], [126], [127] und [173]). Da eine exakte Analyse des Puffers für den oben beschriebenen Eingangsverkehr wegen der Größe des benötigten Zustandsraums nur in wenigen Sonderfällen möglich ist [127], wurden zahlreiche Näherungen zur Bestimmung der Zellverlustwahrscheinlichkeit vorgeschlagen. Ein Vergleich von verschiedenen approximativen Ansätzen ist beispielsweise in [126] enthalten.

Um eine brauchbare Näherung für die Verlustwahrscheinlichkeit zu erhalten, muß das Verhalten der Quellen auf Burst-Ebene sowie das Auftreten von Überlastzuständen geeignet berücksichtigt werden. In [13] ist ein analytischer Ansatz beschrieben, der auf einer Appro-

ximation des Ankunftsprozesses durch einen Markov-Modulierten Poisson-Prozeß (MMPP) mit zwei Zuständen basiert. Dabei werden alle Überlastzustände der Quellen zu einem einzigen Zustand des MMPP zusammengefaßt, dessen Ankunftsrate der mittleren Zellrate aller Überlastzustände entspricht. Auf diese Weise kann zwar der charakteristische Knick des Kurvenverlaufs nachgebildet werden [13], die Steigung des flacheren Kurvenabschnitts stimmt jedoch nicht in jedem Fall mit dem tatsächlichen Verlauf überein. Diese Abweichung ist vor allem auf die Aggregation der Überlastzustände zu einem einzigen Zustand zurückzuführen.

Eine genauere Approximation des Quellverhaltens ermöglicht das sogenannte „Fluid Flow"-Modell [4],[53],[140], bei dem der diskrete Ankunftsprozeß der Zellen durch den kontinuierlichen Strom einer Flüssigkeit ersetzt wird. Die Zustandswechsel der Quellen auf Burst-Ebene werden dabei durch eine entsprechende Modulation des Flüssigkeitsstroms nachgebildet, und der Puffer entspricht einem Behälter mit endlichem Fassungsvermögen, der mit konstanter Rate geleert wird. Das „Fluid Flow"-Modell liefert eine sehr genaue Näherung für den flachen Kurvenabschnitt, dessen Verlauf im wesentlichen durch das Verhalten der Quellen auf Burst-Ebene bestimmt wird [126],[140],[173]. Für sehr kleine Puffer liefert das „Fluid Flow"-Modell zwar eine zu geringe Verlustwahrscheinlichkeit, in diesem Bereich können aber auch die Werte bei Bernoulli-Verkehr als gute Näherung verwendet werden (siehe Bild 6.3). Die Addition der Verlustwahrscheinlichkeiten bei Bernoulli-Verkehr und unter Verwendung des „Fluid Flow"-Modells ergibt daher eine sehr gute Approximation der Zellverlustwahrscheinlichkeit für beliebige Werte der Puffergröße [126],[173].

6.4 Vergleich von einstufigen Anordnungen

In Kapitel 4 wurden verschiedene Architekturen für einstufige ATM-Koppelanordnungen vorgestellt, die bei Bernoulli-Verkehr einen wesentlich geringeren Speicherplatzbedarf aufweisen als Koppelelemente mit Ausgangspuffern. Zu diesen Architekturen gehören beispielsweise Koppelelemente mit Zentralspeicher sowie Koppelelemente mit Eingangspuffern, die zwei Lesezugriffe pro Zeitschlitz auf beliebige Pufferplätze zulassen (siehe Abschnitt 4.3.2). Im folgenden soll nun das Verhalten dieser Koppelelemente bei stoßartigem Verkehr untersucht und mit der Leistungsfähigkeit von Koppelelementen mit Ausgangspuffern verglichen werden.

Zu diesem Zweck wurden die genannten Architekturen mit Eingangspuffer bzw. Zentralspeicher dem bereits im vorigen Abschnitt verwendeten Eingangsverkehr unterworfen. Wie in Bild 6.7 dargestellt, ist der Einfluß der mittleren Burst-Länge auf die Verlustwahrscheinlichkeit ähnlich wie bei einem Koppelelement mit Ausgangspuffern (vgl. Bild 6.4). Die mittlere Burst-Länge wirkt sich auch hier sehr stark auf die Steigung des Kurvenverlaufs aus, so daß der Speicherbedarf bei stoßartigem Verkehr entsprechend ansteigt.

Es stellt sich nun die Frage, ob der Vorteil des geringeren Speicherplatzbedarfs der hier betrachteten Architekturen gegenüber einem Koppelelement mit Ausgangspuffern auch bei

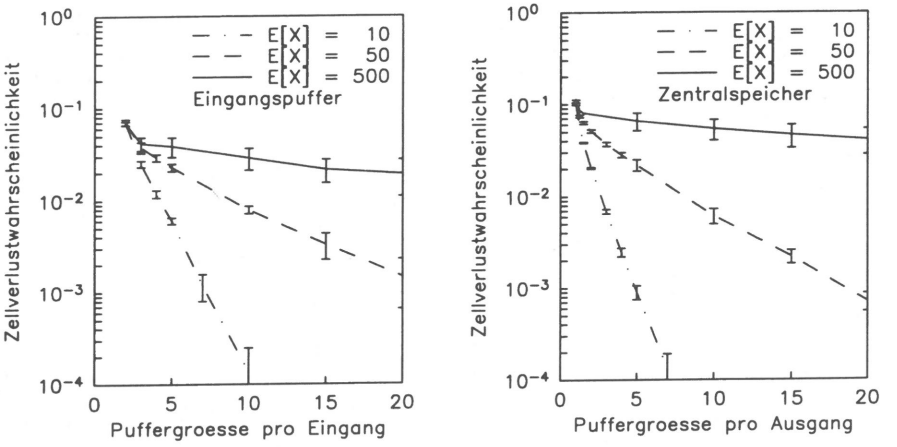


Bild 6.7: Einfluß der mittleren Burst-Länge auf die Verlustwahrscheinlichkeit von Koppelelementen mit Eingangspuffern (wahlfreier Zugriff) und mit Zentralspeicher

stoßartigem Verkehr erhalten bleibt. Bild 6.8 zeigt einen direkten Vergleich der drei Varianten bezüglich der Verlustwahrscheinlichkeit für eine mittlere Burst-Länge von 10 und 500 Zellen. Bei einer mittleren Burst-Länge von 10 Zellen schneiden Koppelelemente mit Eingangspuffer bzw. Zentralspeicher ebenfalls besser ab, da sich die gemeinsame Nutzung der Puffer durch mehrere Ausgänge hier günstig auswirkt. Insbesondere bei Verwendung eines Zentralspeichers können kurzzeitige Überlastsituationen wesentlich besser ausgeglichen werden als bei getrennten Ausgangspuffern, was zu einer deutlich schnelleren Abnahme der Verlustwahrscheinlichkeit über der Puffergröße (pro Ausgang) führt. In einem Koppelelement mit Eingangspuffern sind die für einen Ausgang bestimmten Bursts über mehrere Eingangspuffer verteilt, so daß auch hier eine kurzzeitige Überlast besser ausgeglichen werden kann.

Bei einer mittleren Burst-Länge von 500 Zellen führt dieser Vergleich jedoch zu einem völlig anderen Ergebnis. Hier ist die Verlustwahrscheinlichkeit von Koppelelementen mit Zentralspeicher bzw. Eingangspuffern deutlich größer als bei einem Ausgangspuffer. Dieses Verhalten hängt mit der Verwaltung der gemeinsam genutzten Puffer zusammen: Bei einer länger andauernden Überlastsituation an einem beliebigen Ausgang beginnen sich die Eingangspuffer (bzw. der Zentralspeicher) zu füllen, und es treten bei allen Verbindungen hohe Verluste auf (vgl. Abschnitt 4.2). Im Gegensatz dazu sind in einem Koppelelement mit Ausgangspuffern nur diejenigen Verbindungen von einer Überlast betroffen, die dem überlasteten Ausgang zugeordnet sind, so daß insgesamt wesentlich weniger Verluste auftreten.

Die höhere Verlustwahrscheinlichkeit von Koppelelementen mit Eingangspuffern bzw. Zentralspeicher kann auch zu Problemen bei der Verbindungsannahme führen, da die tatsächliche

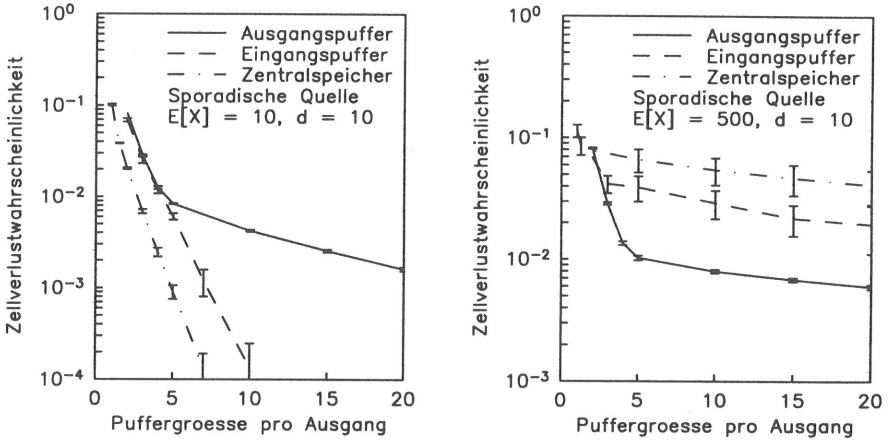


Bild 6.8: Vergleich verschiedener Architekturen (siehe Text) bei stoßartigem Verkehr mit einer mittleren Burst-Länge von 10 und 500 Zellen

Zellverlustwahrscheinlichkeit hier größer als die Überlastwahrscheinlichkeit eines einzelnen Ausgangs sein kann. Aus diesem Grund muß entweder der Algorithmus zur Verbindungsannahme angepaßt werden, oder aber die Auswirkung einer Überlast muß durch geeignete Maßnahmen auf den betroffenen Ausgang beschränkt werden.

Wie bereits in Abschnitt 4.2 für den Fall einer stationären Überlast beschrieben, kann die gegenseitige Beeinflussung der Ausgänge bei Koppelelementen mit Zentralspeicher durch eine entsprechende Verwaltung des Speichers stark reduziert werden. Auf diese Weise läßt sich, wie in Bild 6.9 dargestellt, auch das Verhalten bei stoßartigem Verkehr deutlich verbessern.

Bild 6.9 zeigt nochmals, daß die Verlustwahrscheinlichkeit bei einer Verwaltung des Zentralspeichers nach dem „Complete Sharing"-Prinzip ab einer bestimmten Burst-Länge höher ist als bei Koppelelementen mit Ausgangspuffern. Die geringsten Verluste treten auch hier bei dem Verfahren „Drop on Demand" (DOD) auf, dieses Verfahren ist wegen seiner Komplexität jedoch weniger für eine Hardware-Realisierung geeignet. Die Begrenzung der logischen Ausgangswarteschlangen (Maximum Queue, MQ) dürfte daher auch bei stoßartigem Verkehr am besten geeignet sein, um die Verlustwahrscheinlichkeit bei extrem langen Bursts zu reduzieren. Die in Bild 6.9 eingezeichnete Kurve bezieht sich auf einen Zentralspeicher mit 80 Warteplätzen, wobei die logischen Ausgangswarteschlangen auf 40 Zellen begrenzt sind.

Bei Koppelelementen mit Eingangspuffern ist eine entsprechende Speicherverwaltung nur dann möglich, wenn jeder Eingangspuffer ebenfalls in mehrere logische Ausgangspuffer unterteilt wird (Multi-FIFO-Eingangspuffer, vgl. Abschnitt 1.3.1.3).

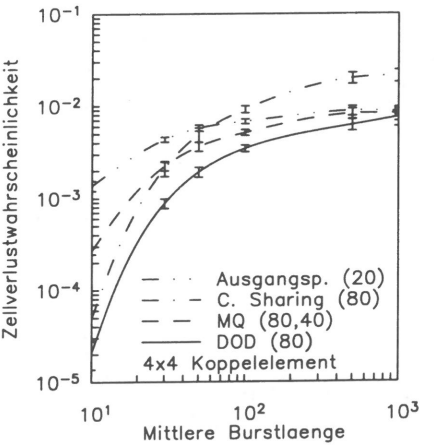


Bild 6.9: Einfluß der mittleren Burst-Länge auf die Verlustwahrscheinlichkeit für verschiedene Verfahren zur Speicherverwaltung

6.5 Einfluß der virtuellen Verbindungen

In diesem Abschnitt soll gezeigt werden, daß die Verkehrscharakteristik einer Quelle auf Burst-Ebene wegen des verbindungsorientierten Betriebs von ATM-Netzen auch innerhalb des Netzes weitgehend erhalten bleibt. Zu diesem Zweck wird ein mehrstufiges Banyan-Netzwerk mit verbindungsorientiertem Betrieb untersucht, dessen Eingangsverkehr aus der Überlagerung von mehreren Sporadischen Quellen besteht. Das Netzwerk ist aus 2×2-Koppelelementen mit Ausgangspuffern aufgebaut.

In Bild 6.10 ist die komplementäre Verteilungsfunktion der Warteschlangenlänge in den Stufen 1 und 5 des betrachteten Netzwerks für unterschiedliche Werte der mittleren Burst-Länge dargestellt (Angebot 0.64, Parameter siehe Tabelle 6.1). Die jeweils gepunktet eingezeichneten Ergebnisse für die Stufe 5 stimmen nahezu exakt mit den Werten für Stufe 1 überein, es kann also keine signifikante Glättung des Verkehrs durch die Puffer des Netzes festgestellt werden. Die Zwischenspeicherung der Zellen hat zwar gewisse Durchlaufzeitschwankungen zur Folge, diese haben jedoch keinen wesentlichen Einfluß auf die Modulation des Verkehrs auf Burst-Ebene.

Um eine approximative Analyse von netzinternen Koppelelementen auf Burst-Ebene zu ermöglichen, wird die Veränderung des Verkehrs durch vorhergehende Multiplex- und Vermittlungsvorgänge daher meist vernachlässigt. Dieses Vorgehen ist vor allem dann sinnvoll, wenn die Spitzenbitrate der Verbindungen deutlich kleiner als die Kapazität der Leitungen ist, und in jeder Stufe möglichst viele Verbindungen gemultiplext werden (siehe auch [129]).

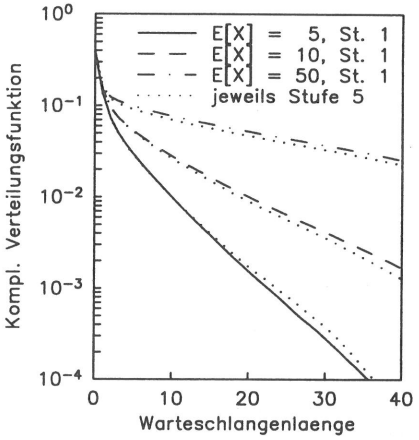


Bild 6.10: Verteilung der Warteschlangenlänge in den Stufen 1 und 5 bei Berücksichtigung von virtuellen Verbindungen

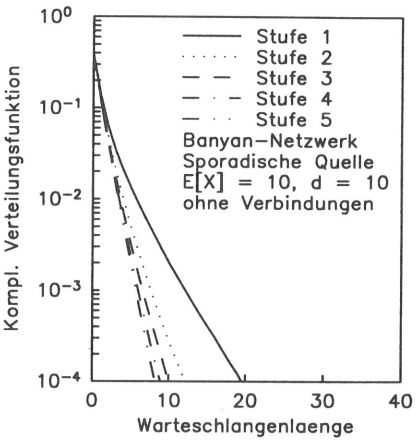


Bild 6.11: Verteilung der Warteschlangenlänge in den Stufen 1 bis 5 bei zufälliger Verzweigung der Zellen

Im Gegensatz zum verbindungsorientierten Betrieb hat eine zufällige Verzweigung der Zellen zur Folge, daß die Verkehrseigenschaften der Quelle bereits nach wenigen Stufen verloren gehen. Bild 6.11 zeigt die komplementäre Verteilungsfunktion der Warteschlangenlänge in den ersten 5 Stufen eines Banyan-Netzwerks mit zufälliger Verzweigung, wobei der Eingangsverkehr mit dem in Bild 6.10 verwendeten Verkehr bei einer mittleren Burst-Länge von $E[X] = 10$ Zellen völlig identisch ist. Verglichen mit der entsprechenden Kurve bei verbindungsorientiertem Betrieb ist hier eine starke Abnahme der Warteschlangenlänge zu verzeichnen, lediglich in der ersten Stufe hat die stoßartige Charakteristik des Eingangsverkehrs noch deutlich sichtbare Auswirkungen. Die Vernachlässigung des verbindungsorientierten Betriebs führt daher zu einer völlig falschen Einschätzung der Verlustwahrscheinlichkeit in den hinteren Stufen des Koppelnetzes.

Kapitel 7

Untersuchung von Koppelnetzstrukturen auf Verbindungsebene

In den Kapiteln 4, 5 und 6 wurden verschiedene Aspekte der Leistungsfähigkeit von ATM-Koppelanordnungen auf Zellebene untersucht. Die Anzahl der Verkehrsquellen und damit das Angebot auf Zellebene wurde dabei als konstant vorausgesetzt, Verbindungsauf- und abbauvorgänge blieben bisher also unberücksichtigt. Der Durchsatz eines Koppelnetzes hängt jedoch nicht nur von dessen Eigenschaften auf Zellebene ab, insbesondere bei Anordnungen mit verbindungsorientiertem Betrieb spielen auch andere Aspekte, wie z.B die Topologie des Koppelnetzes, das Verfahren zur internen Wegesuche sowie der Algorithmus zur Verbindungsannahme, eine große Rolle. In diesem Kapitel soll vor allem der Einfluß der Koppelnetztopologie und des Wegesuchverfahrens untersucht werden.

Koppelanordnungen mit verbindungsorientiertem Betrieb sind dadurch gekennzeichnet, daß der Weg aller Zellen einer Verbindung durch das Koppelnetz während des Verbindungsaufbaus festgelegt wird. Bei der Wegesuche muß für jede benutzte Zwischenleitung geprüft werden, ob die betreffende Verbindung unter Einhaltung der geforderten Dienstgüte angenommen werden kann (Verbindungsannahme, siehe Abschnitt 1.1.6.1). Diese Entscheidung basiert auf standardisierten Verkehrsparametern, welche die Spitzenbitrate und eventuell weitere Eigenschaften der Verbindung angeben [45].

In diesem Kapitel soll vereinfachend davon ausgegangen werden, daß jede Verbindung nur durch ihre Spitzenbitrate charakterisiert ist. Momentan sind ohnehin keine weiteren Verkehrsparameter festgelegt [45], so daß diese Annahme für die Einführungsphase von ATM-Netzen durchaus zutrifft. Falls in einer späteren Phase auch das statistische Multiplexen von Verbindungen möglich ist, kann dieser Verkehrsparameter auch als effektive Bandbreite einer

Verbindung interpretiert werden. Die effektive Bandbreite kennzeichnet diejenige Bandbreite, die bei der Annahme einer Verbindung auf der betreffenden Leitung reserviert werden muß, und liegt zwischen der mittleren Bitrate und der Spitzenbitrate. Obwohl die Verwendung einer effektiven Bandbreite für die Verbindungsannahme in ATM-Netzen umstritten ist, kann dieses Konzept dennoch als Grundlage für vergleichende Untersuchungen auf Verbindungsebene verwendet werden [59].

Um einen möglichst hohen Nutzdurchsatz zu erzielen, müssen Koppelnetze mit einer geringen internen Blockierwahrscheinlichkeit verwendet werden. Intern blockierungsfreie Koppelnetze besitzen natürlich die beste Verkehrsleistung auf Verbindungsebene, da hier jede zulässige Verbindung, unabhängig von anderen bereits bestehenden Verbindungen, angenommen werden kann (siehe auch 1.3.4). Die Blockierungsfreiheit einer Koppelanordnung hängt vor allem von ihren topologischen Eigenschaften und von den Verkehrsparametern der durchzuschaltenden Verbindungen ab. Falls eine Koppelanordnung auf Grund ihrer Topologie zwar Blockierungen aufweist, deren Auftreten jedoch durch ein geeignetes Wegesuchverfahren verhindert wird, ist die Koppelanordnung im weiteren Sinne blockierungsfrei.

Im folgenden Abschnitt wird zunächst eine Bedingung für die Blockierungsfreiheit von ATM-Koppelnetzen auf der Basis eines dreistufigen Clos-Netzwerks abgeleitet. Die Anwendung dieser Bedingung auf verschiedene Sonderfälle zeigt jedoch, daß die Realisierung einer intern blockierungsfreien ATM-Koppelanordnung sehr aufwendig sein kann. In der Praxis werden daher meist virtuell blockierungsfreie Koppelnetze eingesetzt, deren interne Blockierwahrscheinlichkeit gegenüber der Blockierwahrscheinlichkeit auf den ankommenden und abgehenden Leitungen vernachlässigt werden kann. Die Leistungsfähigkeit eines derartigen Koppelnetzes hängt sehr stark von dem verwendeten Wegesuchverfahren ab, da das Auftreten von Blockierungen durch eine geschickte Auswahl der Wege teilweise vermieden werden kann. In Abschnitt 7.2 wird daher die Leistungsfähigkeit verschiedener Wegesuchverfahren anhand einer ausgewählten Koppelnetztopologie verglichen.

7.1 Intern blockierungsfreie ATM-Koppelnetze

Die Theorie der blockierungsfreien Koppelnetze geht auf eine Arbeit von Clos [57] aus dem Jahr 1953 zurück. Später wurden die Ergebnisse von Clos auf Koppelnetze mit Mehrkanal-Durchschaltung [159] und zuletzt auch auf ATM-Koppelanordnungen [59], [153], [154], [181], [197] übertragen. Eine Bedingung für die Blockierungsfreiheit soll anhand des in Bild 7.1 dargestellten dreistufigen Clos-Netzwerks hergeleitet werden (siehe auch [59]). Die Topologie dieses Netzwerks wurde bereits in Abschnitt 1.3.4.2 ausführlich beschrieben, an dieser Stelle soll deshalb nicht näher darauf eingegangen werden. Eine Besonderheit des Netzwerks in Bild 7.1 ist die Möglichkeit einer internen Geschwindigkeitserhöhung, wobei die internen

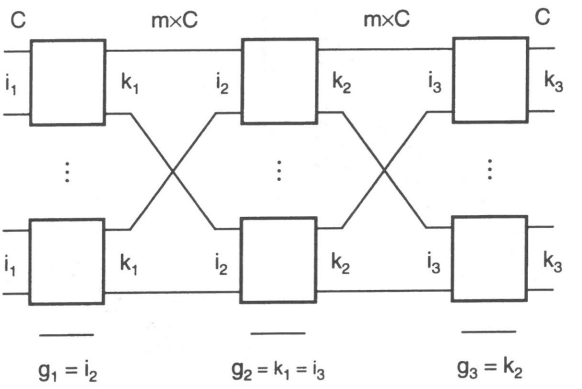


Bild 7.1: Topologie eines dreistufigen Clos-Netzwerks mit interner Geschwindigkeitserhöhung

Zwischenleitungen eine um den Faktor m höhere Verkehrskapazität besitzen als die externen Leitungen.

Für die weiteren Betrachtungen werden alle Verbindungsparameter auf die nutzbare Kapazität C einer externen Leitung des Koppelnetzes normiert. Es sei r die normierte Bandbreite einer Verbindung, mit $0 < r \leq 1$. In bestimmten Fällen kann auch eine Begrenzung von r auf das Intervall $(u, U]$ notwendig sein, wobei u die kleinste Bandbreite und U die maximal mögliche Bandbreite einer Verbindung angibt, so daß $0 \leq u < r \leq U \leq 1$ gilt.

Eine neue Verbindung kann nur dann angenommen werden, wenn die freie Bandbreite der betrachteten Leitung größer oder gleich r ist. Dabei ergibt sich die freie Bandbreite aus der Differenz der normierten Gesamtkapazität einer Leitung (1 bei externen sowie m bei internen Leitungen) und der bereits belegten Bandbreite. Unter diesen Voraussetzungen ist die in Bild 7.1 dargestellte Anordnung genau dann blockierungsfrei, wenn

$$g_2 > \max_{u<r\leq U} \left\lfloor \frac{i_1 - r}{\max(u, m-r)} \right\rfloor + \max_{u<r\leq U} \left\lfloor \frac{k_3 - r}{\max(u, m-r)} \right\rfloor \tag{7.1}$$

gilt. Die Herleitung dieser Bedingung basiert auf einer „Worst-Case"-Betrachtung [59], die in ähnlicher Form bereits für den Beweis der Blockierungsfreiheit in [57] verwendet wurde.

Für den Beweis von Gleichung 7.1 wird der Aufbau einer Verbindung mit der normierten Bandbreite r betrachtet. Die maximale Belastung der Eingangsleitung, auf der diese Verbindung ankommt, ist $1 - r$. Die restlichen $i_1 - 1$ Eingänge des Koppelelements der ersten Stufe, mit dem diese Eingangsleitung verbunden ist, sollen ebenfalls voll ausgelastet sein, so daß die Gesamtbelastung des Koppelelements durch $i_1 - r$ gegeben ist (siehe Bild 7.2). Im ungünstigsten Fall teilt sich der ankommende Verkehr gleichmäßig auf mehrere interne Zwischenleitungen auf, wobei die Belastung jeder Zwischenleitung mit $m - r$ gerade so groß

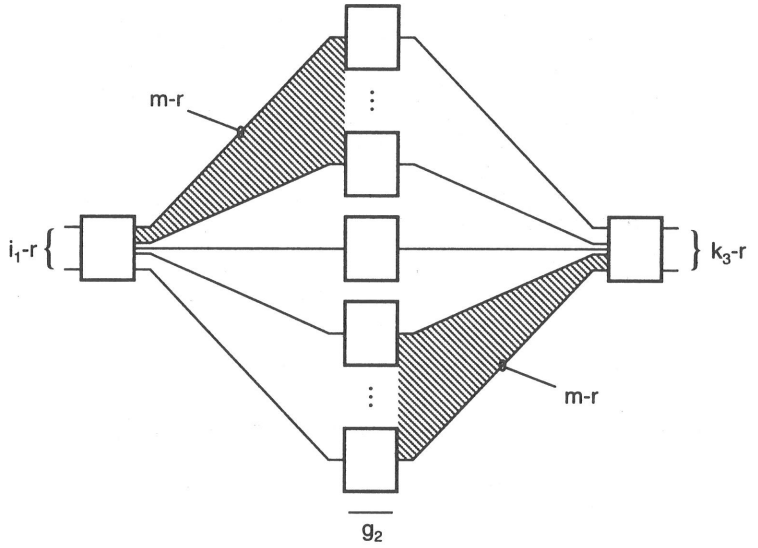


Bild 7.2: „Worst-Case"-Betrachtung zur Herleitung der Blockierungsfreiheit (Ausschnitt aus einem dreistufigen Clos-Netzwerk)

ist, daß die betrachtete Verbindung nicht mehr angenommen werden kann. In der geschilderten Situation sind daher bereits $\left\lfloor \frac{i_1 - r}{\max(u, m-r)} \right\rfloor$ Koppelelemente in Stufe 2 für die betreffende Verbindung blockiert. Dabei bezeichnet das Symbol $\lfloor X \rfloor$ die größte ganze Zahl, die kleiner oder gleich X ist. Außerdem ist bereits berücksichtigt, daß die Bandbreite einer Verbindung nie kleiner als u sein kann.

Eine ähnliche Betrachtung kann für das Koppelelement in Stufe 3 durchgeführt werden, über welches die betrachtete Verbindung das Koppelnetz verläßt. Von diesem Koppelelement aus können bis zu $\left\lfloor \frac{k_3 - r}{\max(u, m-r)} \right\rfloor$ Koppelelemente der zweiten Stufe nicht mehr erreichbar sein. Falls alle eingangs- und ausgangsseitig blockierten Koppelelemente in Stufe 2 jeweils verschieden sind, werden insgesamt

$$g_2 > \left\lfloor \frac{i_1 - r}{\max(u, m-r)} \right\rfloor + \left\lfloor \frac{k_3 - r}{\max(u, m-r)} \right\rfloor$$

Koppelelemente in der mittleren Stufe benötigt, um dennoch eine Verbindung der Bandbreite r aufbauen zu können. Das Koppelnetz ist somit blockierungsfrei, wenn diese Bedingung für alle zulässigen Werte von r erfüllt ist. Durch Bildung des Maximalwertes für $u < r \leq U$ erhält man daraus Gleichung 7.1.

Ausgehend von einem dreistufigen Netzwerk lassen sich auf diese Weise intern blockierungsfreie Clos-Netzwerke mit beliebiger ungerader Stufenzahl konstruieren. Zu diesem Zweck wird

jedes Koppelelement der mittleren Stufe jeweils durch ein intern blockierungsfreies, dreistufiges Clos-Netzwerk ersetzt, auf das Gleichung 7.1 angewendet werden kann. Durch schrittweises Ersetzen erhält man auf diese Weise beliebige intern blockierungsfreie Clos-Netzwerke mit ungerader Stufenzahl.

Es sollen nun verschiedene Sonderfälle von Gleichung 7.1 diskutiert werden, die von besonderer praktischer Bedeutung sind. Dabei werden Koppelanordnungen mit und ohne interne Geschwindigkeitserhöhung getrennt behandelt.

7.1.1 Netzwerke ohne interne Geschwindigkeitserhöhung

Zunächst soll der Fall betrachtet werden, daß die internen Zwischenleitungen die gleiche Verkehrskapazität wie die externen Leitungen besitzen. Es wird daher $m = 1$ vorausgesetzt.

7.1.1.1 Durchschaltevermittlung

Für den Fall der Durchschaltevermittlung belegt jede einzelne Verbindung die gesamte Bandbreite einer Leitung, es gilt also $r = u = U = 1$. Hier vereinfacht sich Gleichung 7.1 zu der bereits aus Abschnitt 1.3.4.2.3 sowie aus [57] bekannten Beziehung

$$g_2 \geq i_1 + k_3 - 1.$$

7.1.1.2 Begrenzte Maximalbitrate

Es soll nun angenommen werden, daß die Bandbreite einer Verbindung nur nach oben hin durch den (normierten) Grenzwert $U < 1$ begrenzt ist. Mit $u = 0$ ergibt sich aus Gleichung 7.1

$$g_2 > \left\lfloor \frac{i_1 - U}{1 - U} \right\rfloor + \left\lfloor \frac{k_3 - U}{1 - U} \right\rfloor . \tag{7.2}$$

Daraus folgt jedoch [59], daß ein verbindungsorientiertes ATM-Koppelnetz ohne interne Geschwindigkeitserhöhung und ohne Einschränkungen bezüglich der Bandbreite der Verbindungen nicht blockierungsfrei realisiert werden kann ($\lim_{U \to 1} g_2 \to \infty$). Diese Tatsache hängt damit zusammen, daß Verbindungen mit hoher Bandbreite bereits durch eine Verbindung mit niedriger Bitrate blockiert werden können. Das Problem liegt dabei weniger im Vorhandensein von Verbindungen mit hoher Bandbreite, da das Koppelnetz allein für diese Verbindungen durchaus blockierungsfrei gemacht werden kann. Schwierigkeiten bereitet vielmehr die Mischung von Verbindungen mit stark unterschiedlicher Bitrate in einem Koppelnetz. Aus vermittlungstechnischer Sicht besitzt das Konzept der Diensteintegration in einem ATM-Netz also nicht nur Vorteile. In bestimmten Fällen kann es sogar vorteilhaft

sein, bestimmte Verkehrsklassen mit unterschiedlichen Bandbreiteanforderungen getrennt zu vermitteln (siehe 7.1.1.3).

Für symmetrische Koppelnetze ($i_1 = k_3$) kann aus Gleichung 7.2 eine direkte Beziehung zwischen dem internen Aufweitungsfaktor $n = k_1/i_1$ und der maximal zulässigen Bandbreite U einer Verbindung hergeleitet werden:

$$U < \frac{n-2}{n-2/i_1}. \tag{7.3}$$

Als Faustformel kann auch der Grenzwert für $i_1 \to \infty$ verwendet werden, es ergibt sich hier $U < 1 - 2/n$. Bei einer internen Aufweitung um den Faktor $n = 4$ muß die Bandbreite der Verbindungen also auf die Hälfte der Kapazität einer Leitung begrenzt werden.

In bestimmten Fällen unterliegt die maximale Bandbreite der Verbindungen auch gewissen systembedingten Einschränkungen. Ein wichtiges Beispiel sind Koppelnetze mit einer Bitrate von 622 Mb/s in einem ATM-Netz, welches im Anschlußbereich nur mit 155.5 Mb/s arbeitet. Hier ist die maximale Bandbreite einer Verbindung auf ein Viertel der Leitungsbandbreite begrenzt ($U = 0.25$), man erhält daher

$$g_2 > \left\lfloor \frac{4i_1 - 1}{3} \right\rfloor + \left\lfloor \frac{4k_3 - 1}{3} \right\rfloor. \tag{7.4}$$

Unter diesen Voraussetzungen ist ein symmetrisches Koppelnetz beispielsweise mit $i_1 = 8$ und $k_1 = 21$ intern blockierungsfrei.

7.1.1.3 Begrenzte Mindestbitrate

Falls alle Verbindungen mindestens eine (normierte) Bandbreite von $u > 0$ aufweisen, während die maximale Bandbreite nur durch die Leitungskapazität begrenzt ist ($U = 1$), vereinfacht sich Gleichung 7.1 zu

$$g_2 > \left\lfloor \frac{i_1 - 1}{u} \right\rfloor + \left\lfloor \frac{k_3 - 1}{u} \right\rfloor + 2. \tag{7.5}$$

In einem ATM-Netz könnte beispielsweise eine Mindestbitrate von 64 kb/s eingeführt werden, so daß sich bei einer Linkbitrate von 155 Mb/s für u ein Wert von etwa 1/2000 ergibt. Ein symmetrisches Koppelnetz müsste daher rund um den Faktor 4000 aufgeweitet werden, um unter diesen Bedingungen blockierungsfrei zu sein. Selbst bei einer Mindestbitrate von 2 Mb/s wäre noch ein Aufweitungsfaktor von mehr als 100 erforderlich, die Anhebung der Mindestbitrate allein eignet sich daher nicht für die praktische Realisierung einer blockierungsfreien Anordnung.

Durch eine Trennung des Verkehrs auf Grund der benötigten Bandbreite kann ein dreistufiges Clos-Netzwerk jedoch auch ohne Einschränkungen der minimalen und maximalen Bandbreite

blockierungsfrei sein [59]. Zu diesem Zweck werden die Koppelelemente der mittleren Stufe in zwei Gruppen aufgeteilt. Über die erste Gruppe werden nur Verbindungen mit $r \leq 0.5$ geführt, die zweite Gruppe erhält alle Verbindungen mit $r > 0.5$. Für beide Gruppen werden nach den Gleichungen 7.2 und 7.5 jeweils $2(i_1 + k_3) - 1$ Koppelelemente benötigt, bei einer symmetrischen Anordnung genügt also eine Aufweitung um den Faktor 8. Ein derartiges Koppelnetz ist jedoch nur im weiteren Sinne blockierungsfrei, da das Auftreten von Blockierungen nur durch die verkehrsabhängige Auswahl der internen Wege verhindert wird.

7.1.2 Netzwerke mit interner Geschwindigkeitserhöhung

Die interne Geschwindigkeitserhöhung ist ein sehr effektives Mittel zur Reduzierung des Aufweitungsfaktors. Dabei kann die Umsetzung der Geschwindigkeit entweder in der ersten Stufe oder bereits am Eingang des Koppelnetzes erfolgen. Wichtig ist nur, daß die maximale Belastung einer Eingangsleitung um den Faktor m, $m > 1$ geringer ist als die Verkehrskapazität einer internen Leitung.

7.1.2.1 Koppelnetze ohne interne Aufweitung

In einem Koppelnetz ohne interne Aufweitung gilt $g_2 = i_1 = k_1 = i_3 = k_3$. Unter Verwendung von Gleichung 7.1 kann daraus eine Bedingung für die interne Geschwindigkeitserhöhung abgeleitet werden, die für ein blockierungsfreies Koppelnetz mindestens erforderlich ist. Es ergibt sich hier

$$m > 3 - \frac{2}{i_1}, \tag{7.6}$$

wobei keinerlei Einschränkungen bezüglich der Bitrate einer Verbindung notwendig sind. Ein dreistufiges Clos-Netzwerk, welches intern mit der dreifachen Geschwindigkeit betrieben wird, ist also ohne jegliche Aufweitung streng blockierungsfrei.

7.1.2.2 Koppelnetze mit interner Aufweitung

Für den allgemeinen Fall eines Koppelnetzes mit Geschwindigkeitserhöhung und interner Aufweitung folgt aus Gleichung 7.1 für die Anzahl der Koppelelemente in der mittleren Stufe

$$g_2 > \left\lfloor \frac{i_1 - 1}{m - 1} \right\rfloor + \left\lfloor \frac{k_3 - 1}{m - 1} \right\rfloor. \tag{7.7}$$

Ein interessanter Sonderfall ergibt sich, wenn das Koppelnetz mit der doppelten Geschwindigkeit der externen Leitungen betrieben wird. Hier reduziert sich obige Gleichung zu der klassischen Bedingung für die Blockierungsfreiheit eines durchschaltevermittelnden Clos-Netzwerks:

$$g_2 \geq i_1 + k_3 - 1, \quad \text{für } m = 2.$$

7.2 Wegesuche in mehrstufigen ATM-Koppelnetzen

In einem streng blockierungsfreien Koppelnetz hat die Auswahl der internen Wege keinen Einfluß auf die Leistungsfähigkeit, da bei vollständiger Abprüfung aller Möglichkeiten stets ein freier Weg gefunden werden kann. Im Gegensatz dazu hängt die Blockierwahrscheinlichkeit in einem nicht blockierungsfreien Koppelnetz sehr stark von der Strategie zur Wegesuche ab. Insbesondere in einem ATM-Netz, das durch die Integration von Diensten mit den unterschiedlichsten Bandbreiteanforderungen gekennzeichnet ist, kann die Blockierwahrscheinlichkeit von Verbindungen mit hoher Bandbreite durch eine geschickte Auswahl der Wege günstig beeinflußt werden.

7.2.1 Strategien zur Wegesuche

Die Wegesuche ist ein Teil der Koppelnetzsteuerung und wird während des Verbindungsaufbaus durchgeführt. Die Funktionalität der Wegesuche umfaßt das Absuchen des Koppelnetzes nach freien Wegen (hunting) und die Auswahl eines Weges aus den bestehenden Möglichkeiten (path selection). Dabei kann prinzipiell zwischen einer stufenweisen und einer weitspannenden Wegesuche unterschieden werden. Bei einer stufenweisen Wegesuche ist das Absuchen auf die weiterführenden Ausgangsleitungen des gerade betrachteten Koppelelements beschränkt, während bei einer weitspannenden Wegesuche zunächst alle durchgehend freien Wege durch das Koppelnetz gesucht werden. In einer größeren Koppelanordnung können jedoch sehr viele verschiedene Wege existieren, das Abprüfen aller Möglichkeiten stellt daher den Hauptaufwand bei der weitspannenden Wegesuche dar.

Neben einer möglichst geringen Blockierwahrscheinlichkeit werden auch an die Realisierung der Wegesuche verschiedene Anforderungen gestellt. Beispielsweise steht für den gesamten Verbindungsaufbau nur eine begrenzte Zeitdauer zur Verfügung (maximal 4.5 s für eine internationale Verbindung [41]), so daß die Wegesuche in den Vermittlungsknoten gewissen Realzeitanforderungen unterliegt. Insbesondere bei einer weitspannenden Wegesuche erfordert die Einhaltung dieser Grenzwerte daher eine überaus leistungsfähige Koppelnetzsteuerung und eine effiziente Implementierung der Wegesuche.

7.2.1.1 Stufenweise Wegesuchverfahren

Die stufenweise Wegesuche ist durch ein schrittweises Vorgehen bei der Auswahl der internen Zwischenleitungen gekennzeichnet. In jeder Stufe wird die Zielinformation des Verbindungswunsches ausgewertet und daraus die möglichen Wege zur folgenden Stufe bestimmt. Falls mehrere Zwischenleitungen benutzt werden können, muß eine der Leitungen nach bestimmten Kriterien (z.B. zufällig, zyklisch oder zustandsabhängig) ausgewählt werden. Da die

Auswahl des Weges unabhängig vom Belegungszustand des restlichen Koppelnetzes erfolgt, kann es durchaus vorkommen, daß in einer Stufe alle weiterführenden Wege belegt sind. In diesem Fall kehrt die Wegesuchstrategie in die vorhergehende Stufe zurück und setzt die Suche dort fort (backtracking). Die oben beschriebenen Realzeitanforderungen können zwar eine Begrenzung der Anzahl dieser Umkehrvorgänge notwendig machen, in diesem Kapitel wird jedoch davon ausgegangen, daß beliebig viele Umkehrvorgänge möglich sind, so daß alle vorhandenen Wege abgeprüft werden.

Da eine stufenweise Wegesuche nur den lokalen Belegungszustand des gerade betrachteten Koppelelements benötigt, ist sowohl eine verteilte als auch eine zentral gesteuerte Realisierung möglich (Zwischenlösungen wie eine gemeinsame Steuerung für jede Koppelgruppe sind ebenfalls denkbar). Wegesuchverfahren mit verteilter Steuerung haben jedoch verschiedene Vorteile. Zunächst besitzen sie die allgemeinen Vorzüge verteilter Systeme, wie zum Beispiel eine erhöhte Ausfallsicherheit und Fehlertoleranz. Darüber hinaus sind die Anforderungen an Laufzeit und Speicherplatzbedarf der einzelnen Steuereinheiten wesentlich geringer als bei einer zentralen Steuerung, da weniger Zustandsinformationen gespeichert und verarbeitet werden müssen. Auf der anderen Seite erfordert die verteilte Steuerung eine Kommunikation zwischen benachbarten Steuereinheiten, um die für den Verbindungsaufbau benötigten Steuerinformationen auszutauschen.

In dieser Arbeit werden verschiedene Varianten der stufenweisen Wegesuche betrachtet. Jede Variante wird durch eine Abkürzung mit drei Buchstaben bezeichnet, wobei der erste Buchstabe die Art des Verfahrens angibt und die restlichen zwei Buchstaben das Verfahren selbst kennzeichnen.

1. *Zufällige Auswahl* (stage-by-stage/random, SRA)
   Es wird zufällig einer der möglichen Ausgänge gewählt.

2. *Zustandsabhängige Strategien*
   Die Auswahl erfolgt abhängig von der Belastung der einzelnen Ausgänge. Falls der Ausgang mit der niedrigsten Auslastung gewählt wird (stage-by-stage/lowest load, SLL) [77], führt dies zu einer gleichmäßigen Verteilung der Last auf alle Ausgänge. Im Gegensatz dazu kann durch Bevorzugen der am höchsten belasteten Leitung (stage-by-stage/highest load, SHL) die freie Bandbreite zusammenhängend gehalten werden, was für Verbindungen mit hoher Bitrate vorteilhaft ist.

3. *Sequentielles Absuchen*
   Sequentielle Absuchverfahren werden bei durchschaltevermittelten Koppelnetzen sehr häufig verwendet [60, 77]. Hier werden die Ausgänge eines Koppelelementes in zyklischer Reihenfolge abgesucht, wobei entweder bei einer festen Nullstellung (stage-by-stage/sequential with homing, SSH) oder bei einer variablen Nullstellung (stage-by-stage/sequential variable, SSV) begonnen wird.

7.2.1.2 Weitspannende Wegesuchverfahren

Bei der weitspannenden Wegesuche werden alle durchgängig freien Wege zum gewünschten Zielausgang zur Auswahl herangezogen. Um den günstigsten Weg für die betrachtete Verbindung zu bestimmen, muß die bereits reservierte Bandbreite auf diesen Wegen berücksichtigt werden. Im folgenden soll als Entscheidungskriterium die Summe der Belastungen aller Zwischenleitungen entlang eines Weges verwendet werden. Ähnlich wie bei der stufenweisen Wegesuche wird entweder der Weg mit der größten (centralized/highest load, CHL) oder mit der kleinsten Summe (centralized/lowest load, CLL) gewählt. Letztere Strategie wird in [5] auch als 'Least-Cost' Methode bezeichnet. Neben der Summe der Belastungen wurde auch die mit Hilfe einer Geo/D/1-Approximation bestimmte mittlere Durchlaufzeit entlang des Weges als Kriterium getestet. Dabei konnten jedoch keine Unterschiede zu der oben beschriebenen Methode festgestellt werden.

Weitspannende Wegesuchstrategien benötigen den Belegungszustand des gesamten Koppelnetzes und können daher nur mit einer zentralen Steuereinheit realisiert werden. Durch eine geeignete Berücksichtigung des gesamten Belegungszustands sollte die Wegesuchstrategie somit in der Lage sein, den bestmöglichen Weg für eine Verbindung zu wählen. Ein Vergleich der oben beschriebenen Verfahren mit einfacheren Strategien zur stufenweisen Wegesuche zeigt jedoch, daß dies nicht immer der Fall ist (siehe 7.2.3.2).

7.2.2 Modellierung

7.2.2.1 Systemmodell

Die Leistungsfähigkeit der oben beschriebenen Wegesuchverfahren soll anhand des in Bild 7.3 dargestellten Modells untersucht werden. Unabhängig von der gewählten Topologie des Koppelnetzes (siehe Abschnitt 7.2.3) sind dessen Eingänge und Ausgänge in mehrere Leitungsbündel aufgeteilt. Diese Leitungsbündel können beispielsweise als Richtungsbündel aufgefaßt werden, welche die Vermittlungsknoten des ATM-Netzes miteinander verbinden. Innerhalb eines Leitungsbündels kann daher eine beliebige Leitung gewählt werden, die genügend freie Bandbreite für die Annahme der betrachteten Verbindung aufweist. Die Vermittlungsknoten verwalten jedoch nur die Leitungen der abgehenden Bündel, die Auswahl einer bestimmten Leitung innerhalb eines ankommenden Bündels erfolgt bereits in der vorhergehenden Vermittlungsstelle und kann nicht mehr verändert werden.

Ein Verbindungswunsch kann an verschiedenen Stellen des Vermittlungsknotens blockiert werden. Falls die Anforderung bereits am Eingang oder am Ausgang des Koppelnetzes abgewiesen wird, spricht man von einer externer Blockierung. Im Gegensatz dazu bezeichnet man eine erfolglose Wegesuche als interne Blockierung, da diese vom Koppelnetz oder seiner

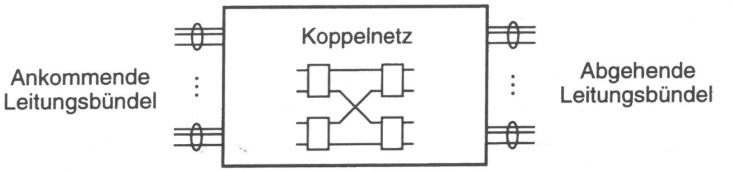


Bild 7.3: Systemmodell für den Vergleich der Wegesuchverfahren

Steuerung hervorgerufen wird. Die interne Blockierwahrscheinlichkeit hängt unter anderem von der Topologie des Koppelnetzes und dem verwendeten Wegesuchverfahren ab.

Durch entsprechende Dimensionierung der Leitungsbündel kann die externe Blockierwahrscheinlichkeit beeinflußt werden. Für den Vergleich der Wegesuchverfahren ist eine geringe externe Blockierwahrscheinlichkeit von großer Bedeutung, da die Unterschiede der einzelnen Strategien sonst nicht zu erkennen sind. Die Dimensionierung der einzelnen Bündel erfolgt normalerweise so, daß bei vorgegebenem Verkehrsangebot eine bestimmte Blockierwahrscheinlichkeit für Verbindungen nicht überschritten wird. Der Einfachheit halber wird im folgenden jedoch davon ausgegangen, daß alle Bündel gleich viele Leitungen enthalten.

7.2.2.2 Modellierung der Verkehrsquellen

Die Modellierung der Verkehrsquellen auf Verbindungsebene erfordert eine Charakterisierung der Verbindungsdauer und des Ankunftsprozesses der Verbindungswünsche (siehe Kapitel 3). Bei der Untersuchung von Koppelnetzen muß zusätzlich auch der Koppelnetzeingang sowie der Zielausgang einer Verbindung festgelegt werden. Für die Leistungsuntersuchung der Wegesuchstrategien werden in diesem Kapitel die folgenden Annahmen getroffen:

- Das betrachtete Koppelnetz sei eine Transit-Vermittlungsstelle innerhalb eines größeren ATM-Netzes, so daß die Verbindungswünsche von sehr vielen Teilnehmern stammen. Der Ankunftsprozeß der Verbindungswünsche ist daher die Überlagerung der Ankunftsprozesse von sehr vielen Einzelquellen mit geringer Ankunftsrate und kann somit als Poisson-Prozeß modelliert werden.
- Die Verbindungsdauer soll eine negativ-exponentielle Verteilung besitzen.
- Jeder Verbindungswunsch wird zufällig und gleichwahrscheinlich einem beliebigen Eingangs- bzw. Ausgangsbündel zugeordnet. Alle Eingänge und Ausgänge des Koppelnetzes sind daher im Mittel gleich belastet. Falls die freie Kapazität der gewählten Eingangsleitung des Koppelnetzes für die Aufnahme der Verbindung nicht ausreicht, wird eine andere Leitung innerhalb der Leitungsgruppe zufällig bestimmt. Analog dazu kann ein anderer Ausgang innerhalb des selben Bündels benutzt werden, falls der zunächst gewählte Ausgang blockiert ist oder die Wegesuche erfolglos war.

7.2.3 Leistungsvergleich

Eine analytische Untersuchung der Blockierwahrscheinlichkeit in mehrstufigen Koppelnetzen ist selbst bei einfachen Wegesuchverfahren bereits sehr komplex. Ein Vergleich verschiedener Wegesuchstrategien ist daher nur mit Hilfe einer ereignisgesteuerten Simulation möglich. In diesem Abschnitt soll die Leistungsfähigkeit der unter 7.2 beschriebenen Wegesuchverfahren am Beispiel einer ausgewählten Koppelnetztopologie verglichen werden.

Für die Bestimmung einer geeigneten Topologie wurden zunächst verschiedene Koppelnetze mit einem einfachen Wegesuchverfahren untersucht. Bei den betrachteten Topologien handelt es sich um ein Banyan-Netzwerk mit zwei Ebenen (siehe 2.3.4.1), ein Beneš-Netzwerk (siehe 2.3.4.2.1) sowie eine Umkehrgruppierung, die ebenfalls auf einem Banyan-Netzwerk basiert [149]. Bild 7.4 zeigt die Topologie dieser Umkehrgruppierung, die aus mehreren Teilnetzen mit der Topologie eines Banyan-Netzwerks besteht. Diese Teilnetze sind über je eine Zwischenleitung mit jedem Koppelelement der Umkehrstufe verbunden, so daß die in Bild 7.4 dargestellte Anordnung mindestens vier verschiedene Wege zwischen den Ein- und Ausgängen aufweist. Wie bei Umkehrgruppierungen üblich können in bestimmten Fällen auch Kurzwege benutzt werden, die bereits vor der eigentlichen Umkehrstufe innerhalb der Teilnetze umkehren (siehe 2.3.4.2.3).

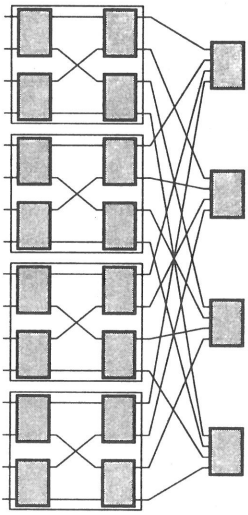


Bild 7.4: Topologie der betrachteten Umkehrgruppierung mit 16 Eingängen

7.2.3.1 Homogenes Verkehrsgemisch

Als Ausgangspunkt für die folgenden Untersuchungen soll zunächst der Sonderfall des homogenen Verkehrs betrachtet werden, alle Verbindungen weisen also einen identischen Bandbreitebedarf auf. Der Einfluß der Koppelnetztopologie wurde am Beispiel eines 256×256 Netzwerks untersucht, welches aus 4×4 Koppelelementen (KE) besteht. Dazu werden benötigt: 448 KE für ein Beneš Netzwerk, 512 KE für ein Banyan Netzwerk mit zwei Ebenen und 832 KE für die beschriebene Umkehrgruppierung. Es wird angenommen, daß jede Verbindung

10% der Kapazität einer Leitung belegt, und daß die Ein- und Ausgänge in Leitungsgruppen zu je 8 Leitungen aufgeteilt sind.

Bild 7.5 zeigt die gesamte Blockierwahrscheinlichkeit für Verbindungswünsche über dem Verkehrsangebot für das Beneš-Netzwerk, die Umkehrgruppierung und das Banyan-Netzwerk mit zwei Ebenen. Ein Banyan-Netzwerk mit nur einer Ebene weist bereits bei einem Angebot von 70% eine Blockierwahrscheinlichkeit von 14% auf und wurde daher nicht berücksichtigt.

Der Vergleich zeigt, daß eine Umkehrgruppierung der angegebenen Topologie und Größe sogar eine höhere Blockierwahrscheinlichkeit aufweist als das entsprechende Beneš-Netzwerk und als das Banyan-Netzwerk mit zwei Ebenen, obwohl diese Netzwerke weniger Koppelelemente benötigen. Die beste Verkehrsleistung erreicht das Banyan-Netzwerk mit zwei Ebenen, bei dem die interne Blockierwahrscheinlichkeit gegenüber der externen Blockierwahrscheinlichkeit vernachlässigbar ist. Bei einem Vergleich der Netzwerktopologien darf jedoch auch die Sensitivität gegenüber ungünstigen Verkehrsbeziehungen nicht außer Acht gelassen werden. Die Leistungsfähigkeit eines Banyan-Netzwerks kann durch bestimmte externe Verkehrsbeziehungen sehr stark beeinträchtigt werden [119, 176, 217], und auch das Hinzufügen einer zusätzlichen Ebene kann dieses Verhalten nicht wesentlich verbessern. Aus diesem Grund wurde das Beneš-Netzwerk als Basis für die weiteren Vergleiche herangezogen.

Der Einfluß verschiedener Wegesuchstrategien auf die interne Blockierwahrscheinlichkeit eines 16×16-Beneš-Netzwerks ist in Bild 7.6 dargestellt. Die Ergebnisse zeigen, daß es am günstigsten ist, gering belastete Leitungen zu benutzen, falls der angebotene Verkehr nur aus Verbin-

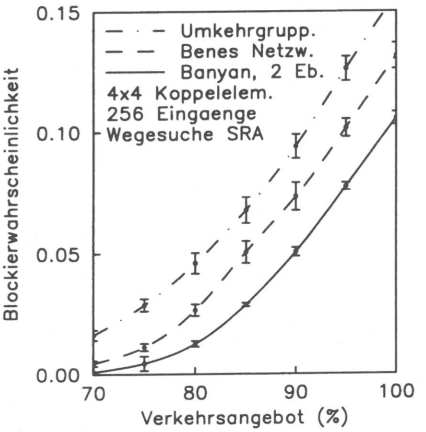


Bild 7.5: Gesamte Blockierwahrscheinlichkeit der Verbindungswünsche bei zufälliger Wegeauswahl (SRA)

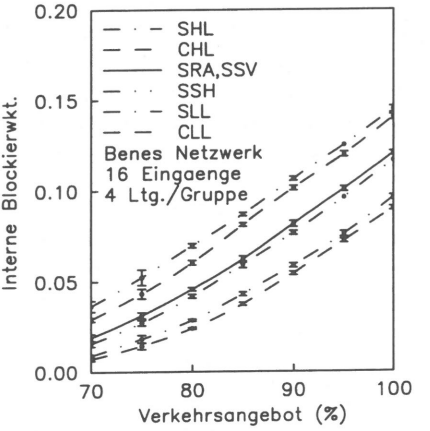


Bild 7.6: Interne Blockierwahrscheinlichkeit bei verschiedenen Wegesuchstrategien (Reihenfolge der Kurven siehe Legende)

dungen derselben Bitrate besteht. Dadurch werden alle internen Leitungen gleichmäßig belastet (siehe auch [77]), so daß möglichst viele verschiedene Wegemöglichkeiten offen bleiben. Im Gegensatz dazu wird die Anzahl der alternativen Wege durch die Auswahl bereits stark belasteter Leitungen schnell reduziert, wodurch sich die Blockierwahrscheinlichkeit erhöht.

Ein Vergleich der lastabhängigen Verfahren mit stufenweiser und weitspannender Wegesuche zeigt, daß Strategien mit stufenweisem Vorgehen eine etwas höhere Blockierwahrscheinlichkeit aufweisen als die entsprechenden weitspannenden Verfahren. Die Unterschiede sind jedoch sehr gering im Vergleich zu dem enormen Aufwand, den die weitspannende Wegesuche für das Absuchen und die Bewertung aller möglichen Wege benötigt.

Innerhalb der Gruppe der stufenweisen Wegesuchverfahren zeigen die sequentiellen Absuchstrategien (SSH und SSV) ein ähnliches Verhalten wie die zufällige Auswahl (SRA). Für den Fall des sequentiellen Absuchens mit variabler Nullstellung stimmen die Ergebnisse sogar exakt überein.

7.2.3.2 Inhomogenes Verkehrsgemisch

Völlig andere Ergebnisse erhält man jedoch für den Fall eines Verkehrsgemischs aus unterschiedlichen Verkehrsklassen. Bild 7.7 zeigt die interne Blockierwahrscheinlichkeit für ein Gemisch von Verbindungen mit 2 Mb/s (Datenkommunikation) und 34 Mb/s (Bewegtbildkommunikation). Unter der Annahme einer maximalen Belastung von 80% auf Zellebene ergibt sich für eine ATM-Leitung mit 155 Mb/s eine nutzbare Kapazität von 112 Mb/s, so daß maximal drei Videoverbindungen über eine Leitung geführt werden können. Da beide Verkehrsklassen denselben Anteil (ausgedrückt als Produkt von Angebot und Bandbreite) zum Gesamtverkehr beitragen, führt dies zu einer relativ großen Blockierwahrscheinlichkeit für Videoverbindungen, wogegen Verbindungen mit 2 Mb/s nur sehr selten abgewiesen werden müssen.

Für dieses Verkehrsgemisch erhält man das zunächst überraschende Ergebnis, daß die sequentielle Absuche mit fester Nullstellung die geringste interne Blockierwahrscheinlichkeit aufweist. Die zufällige Auswahl führt immer zu deutlich schlechteren Werten, wogegen das Abschneiden der lastabhängigen Verfahren stark vom jeweiligen Verkehrsangebot abhängig ist. Bei hohem Angebot verhalten sich alle lastabhängigen Strategien (außer SLL) ähnlich wie die zufällige Auswahl, das SHL Verfahren ist sogar deutlich ungünstiger. Im Gegensatz dazu nähern sich diejenigen Verfahren, die hoch belastete Leitungen bevorzugen, bei niedriger Last (kleiner als 40%) dem Verhalten der Strategie SSH an, während sich die Ergebnisse der Verfahren SLL, CLL und der zufälligen Auswahl kaum unterscheiden.

Zusammenfassend kann festgestellt werden, daß die optimale Strategie sehr stark von der Art des Verkehrsgemischs abhängt. Falls nur eine einzige Verkehrsklasse vorhanden ist, sollte die

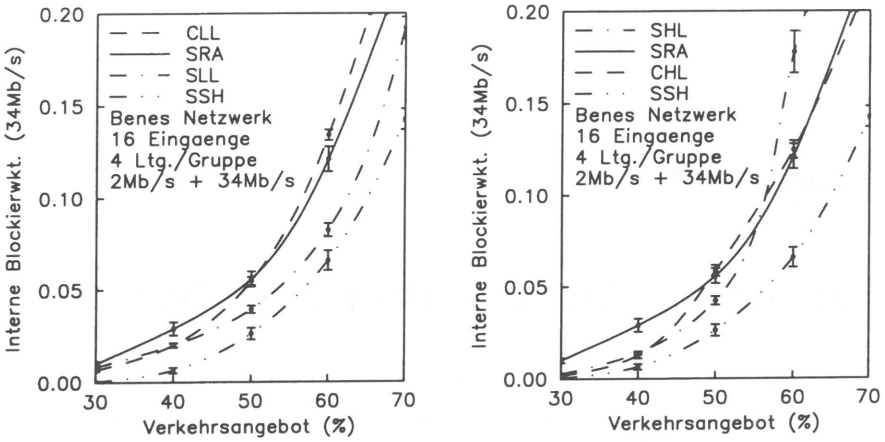


Bild 7.7: Interne Blockierwahrscheinlichkeit der Verbindungen mit 34 Mb/s

Belastung möglichst gleichmäßig auf alle internen Leitungen verteilt werden. Obwohl dies am besten durch eine Strategie erreicht wird, welche die Summe aller Auslastungen entlang eines Weges minimiert (CLL), sollte aus praktischen Erwägungen heraus eine stufenweise Wegesuche bevorzugt werden, bei der die Leitung mit der geringsten Belastung gewählt wird (SLL).

Bei einem Gemisch von Verkehrsklassen mit unterschiedlichen Bandbreiteanforderungen ergibt sich durch sequentielles Absuchen mit fester Nullstellung die geringste Blockierwahrscheinlichkeit aller betrachteten Verfahren. Diese Strategie sorgt dafür, daß die zur Verfügung stehende freie Bandbreite möglichst zusammenhängend bleibt, wodurch die Blockierwahrscheinlichkeit für hochbitratige Verbindungen gesenkt werden kann.

Die Ergebnisse von Abschnitt 7.1.1.3 lassen vermuten, daß durch eine Trennung der beiden Verkehrsklassen eine weitere Leistungssteigerung möglich ist. Diese Trennung könnte beispielsweise durch eine klassenabhängige, sequentielle Absuchstrategie erreicht werden, wobei für die beiden Klassen eine entgegengesetzte Absuchrichtung verwendet wird. Diese einfache Unterscheidung ist jedoch nur bei zwei Verkehrsklassen möglich, bei einem Gemisch aus Verbindungen mit beliebiger Bandbreite muß dagegen eine willkürliche Unterteilung der Verbindungen in zwei Klassen festgelegt werden. Entsprechend Abschnitt 7.1.1.3 müssten alle Verbindungen, deren Bandbreite kleiner bzw. größer als die Hälfte der Kapazität einer Leitung ist, jeweils zu einer Klasse zusammengefaßt werden. Im vorliegenden Fall würden Verbindungen mit 2 Mb/s und mit 34 Mb/s jedoch der gleichen Klasse angehören, so daß kein Unterschied zu einer sequentiellen Absuche mit fester Nullstellung festgestellt werden könnte.

Kapitel 8

Zusammenfassung und Ausblick

In der vorliegenden Arbeit wurden verschiedene Aspekte der Architektur sowie der Verkehrsleistung von ATM-Koppelnetzen behandelt. Ausgehend vom Konzept des zukünftigen Breitband-ISDN und dem dort verwendeten Übermittlungsverfahren, dem *Asynchronus Transfer Mode* (ATM), wurde zunächst die prinzipielle Funktionsweise einer ATM-Koppelanordnung sowie deren Anforderungsprofil beschrieben. Es folgte eine systematische Darstellung der Architektur von ATM-Koppelanordnungen, wobei bereits verschiedene Gesichtspunkte der Verkehrsleistung angesprochen wurden.

Die Leistungsbewertung von Kommunikationssystemen erfordert eine verkehrstheoretische Modellierung des zu untersuchenden Systems sowie des Verkehrs, welcher das System belastet. Neben der Thematik der Quellmodellierung wurden in Kapitel 3 die Grundlagen der zur Leistungsbewertung verwendeten simulativen und analytischen Methoden erläutert. Im Zusammenhang mit den durchgeführten Studien wurden zahlreiche Analyseprogramme und Simulationswerkzeuge erstellt, die eine detaillierte Untersuchung von einstufigen und mehrstufigen Anordnungen erlauben.

ATM-Koppelnetze sind im allgemeinen mehrstufige Netzwerke, die aus mehreren einstufigen Anordnungen, den sogenannten Koppelelementen, aufgebaut sind. Da sich die Verkehrsleistung einer Koppelanordnung aus der Summe der Eigenschaften ihrer einzelnen Komponenten ergibt, müssen sowohl die Eigenschaften der Koppelelemente als auch die des Koppelnetzes optimiert werden, um eine hohe Leistungsfähigkeit zu erhalten.

Beim Entwurf eines ATM-Koppelelements müssen verschiedene Gesichtspunkte wie die Anordnung und Organisation der Pufferspeicher sowie deren Zugriffsgeschwindigkeit berücksichtigt werden. Darüber hinaus spielen auch technologische Randbedingungen wie die maximale Taktrate, die Integrationsdichte und die zur Verfügung stehende Chipfläche eine große Rolle. Die einzelnen Architekturen haben unterschiedliche Anforderungen in Bezug auf diese Parameter, so daß eine sorgfältige Abwägung zwischen der Leistungsfähigkeit einer Variante und dem erforderlichen Implementierungsaufwand notwendig ist.

Koppelelemente mit Ausgangspuffern besitzen bekanntermaßen die besten Eigenschaften in Bezug auf den erreichbaren Durchsatz und die Durchlaufzeitschwankungen. Die Implementierung eines derartigen Koppelelements stellt jedoch hohe Anforderungen an die Zugriffsgeschwindigkeit der Puffer und weist einen vergleichsweise großen Speicherplatzbedarf auf. Aus diesem Grund wird teilweise auf andere Architekturen zurückgegriffen, die mit einer geringeren Zugriffsgeschwindigkeit oder mit weniger Speicherplatz auskommen:

- Koppelelemente mit einem zentralen Speicher haben zwar den geringsten Speicherplatzbedarf, benötigen aber einen sehr schnellen Pufferzugriff und eine aufwendige Speicherverwaltung.
- Eingangspuffer mit wahlfreiem Zugriff ermöglichen ebenfalls eine Reduktion der Speichergröße, für die Auswahl der Zellen ist jedoch eine sehr komplexe Steuerung notwendig.

Ein weiterer Aspekt beim Entwurf eines Koppelelements ist die Auswahl einer geeigneten Arbitrierungsstrategie. Bei Koppelelementen mit Eingangspuffern und mit verteilten Puffern wird eine Arbitrierung benötigt, mit deren Hilfe im Fall einer Abnehmerblockierung eine der Zellen ausgewählt werden kann. Im Hinblick auf eine Reduzierung des Speicherbedarfs sollte jeweils die Zelle aus der längsten Warteschlange gewählt werden, auch wenn dadurch etwas größere Durchlaufzeitschwankungen als bei einer Auswahl nach dem FIFO-Prinzip auftreten.

Ausgangspunkt für die Analyse von mehrstufigen Koppelnetzen war eine genaue Untersuchung der Eigenschaften des koppelnetzinternen Verkehrs. Dabei konnte gezeigt werden, daß die Zellankünfte der internen Stufen selbst bei einem externen Bernoulli-Ankunftsprozeß nicht mehr unabhängig voneinander sind. Die wesentlichen Eigenschaften des internen Verkehrs können durch einen Markov-Modulierten Bernoulli-Prozeß mit drei Zuständen hinreichend genau nachgebildet werden. Dieses Ersatzmodell ermöglicht eine analytische Behandlung der Ausgangspuffer, mit deren Hilfe sowohl die Zustandswahrscheinlichkeiten als auch die Verteilung der Durchlaufzeit in der betreffenden Stufe bestimmt werden können.

Die in Kapitel 5 vorgestellten Ergebnisse zeigen, daß der Bernoulli-Prozeß für eine Dimensionierung der koppelnetzinternen Puffer nicht geeignet ist. Insbesondere bei Koppelelementen mit einer geringen Anzahl von Eingängen muß die Charakteristik des internen Verkehrs berücksichtigt werden, da sonst die zulässige Zellverlustwahrscheinlichkeit überschritten werden kann. Die Dimensionierung der Puffer kann beispielsweise auf der Basis eines „Worst-Case"-Modells für den internen Verkehr erfolgen, welches von dem Ausgangsprozeß eines M/D/1-Systems abgeleitet ist. Unter Verwendung einer Unabhängigkeitsannahme kann auch die Verteilung der gesamten Durchlaufzeit mit Hilfe einer Faltung bestimmt werden. Diese Unabhängigkeitsannahme führt jedoch zu gewissen Abweichungen, da die Durchlaufzeit in aufeinanderfolgenden Stufen des Koppelnetzes eine positive Korrelation aufweist.

In Kapitel 6 wurde der Einfluß der Quellcharakteristik auf das Verhalten von einstufigen und mehrstufigen Koppelanordnungen untersucht. Mit Hilfe einer verfeinerten Modellierung der

Verkehrsquellen konnte gezeigt werden, daß vor allem die Eigenschaften der Burst-Ebene von entscheidender Bedeutung sind. Die Verlustwahrscheinlichkeit hängt außer von der Spitzenbitrate und der mittleren Bitrate der Quellen auch von der gesamten Verteilung der Burst-Länge ab. Der verbindungsorientierte Betrieb von ATM-Netzen hat außerdem zur Folge, daß die Quellcharakteristik auf Burst-Ebene auch innerhalb des Netzes erhalten bleibt. Eine realistische Dimensionierung der Puffer ist daher nur unter Einbeziehung der verwendeten Strategie zur Verbindungsannahme möglich. Falls die Verbindungsannahme auf einer Begrenzung der Überlastwahrscheinlichkeit auf Burst-Ebene basiert, kann die Pufferdimensionierung unter Verwendung eines externen Bernoulli-Ankunftsprozesses erfolgen.

Der Vergleich mit weiteren Architekturen für einstufige Anordnungen hat gezeigt, daß Koppelelemente mit Eingangspuffern sowie mit Zentralspeicher in bestimmten Fällen eine deutlich höhere Verlustwahrscheinlichkeit aufweisen als Koppelelemente mit Ausgangspuffern. Dieser Effekt ist darauf zurückzuführen, daß sich Verbindungen, die für unterschiedliche Ausgänge bestimmt sind, wegen der gemeinsamen Speichernutzung gegenseitig beeinflussen. Die Auswirkungen einer kurzzeitigen Überlast müssen deshalb durch eine geeignete Speicherverwaltung auf den betroffenen Ausgang beschränkt werden, dies kann beispielsweise durch eine Begrenzung der logischen Ausgangswarteschlangen erfolgen.

Die Verkehrsleistung von Koppelnetzen auf Verbindungsebene hängt sowohl von deren Topologie als auch von internen Steuerungsverfahren wie Wegesuche und Verbindungsannahme ab. Am Beispiel eines dreistufigen Clos-Netzwerks wurde in Kapitel 7 die Blockierungsfreiheit von ATM-Koppelanordnungen diskutiert. Ohne eine interne Geschwindigkeitserhöhung kann ein blockierungsfreies Koppelnetz jedoch nur dann realisiert werden, wenn die maximale oder die minimale Bitrate der Verbindungen begrenzt ist. Im Gegensatz dazu ist die betrachtete Anordnung stets blockierungsfrei, wenn intern die dreifache Geschwindigkeit der externen Leitungen verwendet wird.

Bei nicht blockierungsfreien Koppelnetzen hat auch die interne Wegesuche einen wesentlichen Einfluß auf die Blockierwahrscheinlichkeit und den maximalen Durchsatz. Bei der Auswahl eines Wegesuchverfahrens sollte vor allem die Zusammensetzung des Verkehrs berücksichtigt werden. Wenn alle Verbindungen dieselbe Bitrate besitzen (homogenes Verkehrsgemisch), liefert eine möglichst gleichmäßige Verteilung des Verkehrs die besten Ergebnisse. In einem ATM-Netz sollen jedoch beliebige Verkehrsarten mit den unterschiedlichsten Anforderungen an die Bitrate integriert werden. In diesem Fall muß die freie Bandbreite möglichst auf wenige Leitungen konzentriert werden, um die Blockierwahrscheinlichkeit für Verbindungen mit hoher Bitrate zu senken. Es zeigt sich, daß dieses Ziel am besten durch sequentielles Absuchen mit fester Nullstellung erreicht werden kann. Eine mögliche Erweiterung dieses Verfahrens ist die Trennung von Verbindungen mit niedriger und hoher Bandbreite mit Hilfe einer unterschiedlichen Absuchrichtung, um die gegenseitige Beeinflussung zu reduzieren.

Zum gegenwärtigen Zeitpunkt werden von der Industrie Vermittlungssysteme entwickelt, die eine Bitrate von 155 Mb/s bzw. 622 Mb/s unterstützen. Bei der Implementierung sind unter anderem verschiedene Probleme in Bezug auf die Verbindung der einzelnen Baugruppen sowie deren Synchronisation zu lösen [223], [108]. Auch die Kühlung der hochintegrierten Schaltungen erfordert spezielle Maßnahmen, um die bei hohen Arbeitsgeschwindigkeiten entstehende Verlustwärme abzuführen.

In der Einführungsphase des Breitband-ISDN werden voraussichtlich nur Punkt-zu-Punkt-Verbindungen möglich sein, da die entsprechenden Vermittlungssysteme meist keine Multicast-Fähigkeiten besitzen. In der Literatur sind jedoch verschiedene Realisierungsvorschläge für Koppelanordnungen mit Multicast-Betrieb beschrieben [136], [142], [189], [212], [221], die bei der Weiterentwicklung der Vermittlungssysteme berücksichtigt werden könnten.

Aus heutiger Sicht ist bereits abzusehen, daß der Bandbreitebedarf von schnellen Datendiensten sowie der Bewegtbildkommunikation mit hoher Auflösung (zum Beispiel HDTV) wesentlich höhere Geschwindigkeiten innerhalb des Netzes erfordert, um eine effiziente Nutzung der Betriebsmittel zu erzielen. Aus diesem Grund wird in verschiedenen Forschungseinrichtungen auch an Ansätzen der optischen Vermittlungstechnik gearbeitet, die eine wesentliche Steigerung der Bandbreite ermöglichen soll [56], [84], [160], [201]. Das Hauptproblem besteht dabei in der optischen Signalverarbeitung, die in der asynchronen Zeitmultiplextechnik für die Auswertung der Zielinformationen und die Speicherung der Zellen benötigt wird. Die meisten Ansätze sind daher entweder rein theoretischer Natur oder befinden sich noch in einem experimentellen Stadium.

Literaturverzeichnis

[1] AHMADI, H.; DENZEL, W.E.: A Survey of Modern High-Performance Switching Techniques. *IEEE Journal on Selected Areas in Communications*, Vol. 7, No. 7, September 1989, pp. 1091-1103.

[2] AHMADI, H.; DENZEL, W.E.; MURPHY, C.A.; PORT, E.: A High-Performance Switch Fabric for Integrated Circuit and Packet Switching. *International Journal of Digital and Analog Cabled Systems*, Vol. 2, No. 4, October 1989, pp. 277-287.

[3] ALI, M.M.; NGUYEN, H.T.: A Neural Network Controller for a High-Speed Packet Switch. *IEEE International Telecommunications Symposium*, Rio de Janeiro, September 1990, pp. 493-497.

[4] ANICK, D.; MITRA, D.; SONDHI, M.M.: Stochastic Theory of a Data-Handling System with Multiple Sources. *Bell System Technical Journal*, Vol. 61, 1982, pp. 1871-1894.

[5] ANIDO, G.J.; SEETO, A.W.: Multipath Interconnection: A Technique for Reducing Congestion within Fast Packet Switching Fabrics. *IEEE Journal on Selected Areas in Communications*, Vol. 6, No. 9, December 1988, pp. 1480-1488.

[6] ARAKAWA, N.; NOIRI, A.; INOUE, H.: ATM Switch for Multi-Media Switching System. *ISS'90*, Vol. 5, Stockholm, May 1990, pp. 9-14.

[7] ARTHURS, E.; GOODMAN, M.S.; VECCHI, M.P.; KOBRINSKI, H.: HYPASS: An Optoelectronic Hybrid Packet Switching System. *IEEE Transactions on Communications*, Vol. 37, No. 6, June 1989, pp. 645-648.

[8] AUTHENRIETH, F.: *Untersuchung verschiedener Realisierungsmöglichkeiten für Verlustprioritäten in ATM-Netzen.* Semesterarbeit Nr. 1078 am Institut für Nachrichtenvermittlung und Datenverarbeitung, Universität Stuttgart, 1991.

[9] AUTOLITANO, A.; BERNABEI, F.; CIAMPI, M.G.; LISTANTI, M.: Application of Generalized Parallel Delta Networks to a Hybrid Broadband Switch. *ICC'89*, Boston, June 1989, pp. 123-127.

[10] BACHHUBER, W.; FÜRST, W.; KNORR, R.; WILD, T.: Architektur, Technologie und Realisierung eines ATM-Koppelnetzes. *Verbindungsnetzwerke für Parallelrechner und Breitband-Übermittlungssysteme*, ITG/GI Workshop, Stuttgart, 1991, S. 48-52.

[11] BADRAN, H.F.; MOUFTAH, H.T.: Head of Line Arbitration in ATM Switches with Input-Output Buffering and Backpressure Control. *IEEE GLOBECOM'91*, Vol. 1, Phoenix, December 1991, pp. 347-351.

[12] BAE, J.J.; SUDA, T.: Survey of Traffic Control Protocols in ATM Networks. *IEEE GLOBECOM'90*, Vol. 1, San Diego, December 1990, pp. 1-6.

[13] BAIOCCHI, A.; ET AL: Loss Performance Analysis of an ATM Multiplexer Loaded with High-Speed ON-OFF Sources. *IEEE Journal on Selected Areas in Communications*, Vol. 9, No. 3, April 1991, pp. 388-393.

[14] BANNIZA, T.R.; EILENBERGER, G.; PAUWELS, B.; THERASSE, Y.: Design and Technology Aspects of VLSI's for ATM Switches. *IEEE Journal on Selected Areas in Communications*, Vol. 9, No. 8, October 1991, pp. 1255-1264.

[15] BARRI, P.; GOUBERT, J.A.O.: Implementation of a 16 to 16 Switching Element for ATM Exchanges. *IEEE Journal on Selected Areas in Communications*, Vol. 9, No. 5, June 1991, pp. 751-757.

[16] BASSI, S.; DECINA, M.; PATTAVINA, A.: Performance Analysis of the ATM Shuffle-out Switching Architecture under Non-Uniform Traffic Patterns. *IEEE INFOCOM'92*, Florence, May 1992, pp. 735-742.

[17] BATCHER, K.E.: Sorting Networks and their Applications. *Proceedings of AFIPS Spring Joint Computer Conference*, 1968, pp. 307-314.

[18] BENEŠ, V.E.: *Mathematical Theory of Connecting Networks and Telephone Traffic.* Academic Press, New York, 1965.

[19] BERNABEI, F.; FERRETTI, R.; LISTANTI, M.; ZINGRILLO, G.: A Methodology for Buffer Design in ATM Switches. *European Transaction on Communications*, Vol. 2, No. 4, July 1991, pp. 367-379.

[20] BERNATH, E.; SCHWEDERSKI, T.; SIEGEL, M.; WENZEL, J.: Effiziente VLSI-Strukturen für Mehrstufige Verbindungsnetzwerke. *Verbindungsnetzwerke für Parallelrechner und Breitband-Übermittlungssysteme*, ITG/GI Workshop, Stuttgart, September 1991, pp. 53-58.

[21] BHANDARKAR, D.P.: Analysis of Memory Interference in Multiprocessors. *IEEE Transactions on Computers*, Vol. 24, No. 9, September 1975, pp. 897-908.

[22] BINGHAM, B.; BUSSEY, H.: Reservation-Based Contention Resolution Mechanism for Batcher-Banyan Packet Switches. *Electronic Letters*, Vol. 24, No. 13, June 1988.

[23] BLONDIA, C., THEIMER, T.H.: A Discrete-Time Model for ATM Traffic. *Document PRLB-123-0018-CD-CC/UST-123-0022-CD-CC*, RACE 1022, October 1989.

[24] BLONDIA, C.: A Discrete-Time Batch Markovian Arrival Process. *Document PRLB-123-0028-CD-CC*, RACE 1022, December 1990.

[25] BLONDIA, C.: Performance Evaluation of an M/1-Stage in an ATM Switching Element. *Performance Evaluation*, Vol. 15, No. 1, January 1992, pp. 1-20.

[26] BLONDIA, C.; CASALS, O.: Performance Analysis of Statistical Multiplexing of VBR Sources. *IEEE INFOCOM'92*, Florence, May 1992, pp. 828-838.

[27] BONOMI, F.; MONTAGNA, S.; PAGLINO, R.: Busy Period Analysis for an ATM Switching Element Output Line. *IEEE INFOCOM'92*, Florence, May 1992, pp. 544-551.

[28] BOYER, P.; BOYER, J.; LOUVION, J.-R.; ROMOEUF, L.: Modelling the ATD Transfer Technique. *5th ITC Seminar: Traffic Engineering for ISDN Design and Planning*, Lake Como, May 1987, pp. 26.

[29] BOYER, P.: Evaluation of the Multiplex Activity in an ATM-Based Network. *Document CNET-123-042-CD-CC*, RACE 1022, January 1989.

[30] BRIEM, U.: *Analyse von Warteschlangenmodellen mit korreliertem Eingangsprozeß.* Diplomarbeit Nr. 997 am Institut für Nachrichtenvermittlung und Datenverarbeitung, Universität Stuttgart, 1990.

[31] BRUZZI, G.; PATTAVINA, A.: Performance Evaluation of an Input-Queued ATM Switch with Internal Speed-Up and Finite Output Queues. *IEEE GLOBECOM'90*, San Diego, California, December 1990, pp. 1455-1459.

[32] BUBENIK, R.G.; TURNER, J.S.: Performance of a Broadcast Packet Switch. *IEEE Transactions on Communications*, Vol. 37, No. 1, January 1989, pp. 60-69.

[33] BÖTTLE, D.; HENRION, M.A.: Das ATM-Koppelfeld von Alcatel und seine Eigenschaften. *Elektrisches Nachrichtenwesen*, Band 64, Nr. 2, 1990, S. 156-165.

[34] BÖTTLE, D; EILENBERGER, G.J.; SCHRODI, K.: Flexibles ATM-Koppelfeldkonzept für die Breitbandkommunikation. *NTZ*, Band 43, Nr. 11, 1990, S. 810-815.

[35] CCITT, DRAFT RECOMMENDATION I.35B: *B-ISDN ATM Layer Cell Transfer Performance.* Geneva, 1992.

[36] CCITT, RECOMMENDATION I.113: *Vocabulary of Terms for Broadband Aspects of ISDN.* Geneva, 1991.

[37] CCITT, RECOMMENDATION I.121: *Broadband Aspects of ISDN.* Blue Book, Fascicle III.7, Geneva, 1989.

[38] CCITT, RECOMMENDATION I.150: *B-ISDN ATM Functional Characteristics.* Geneva, 1992.

[39] CCITT, RECOMMENDATION I.311: *B-ISDN General Network Aspects.* Geneva, 1992.

[40] CCITT, RECOMMENDATION I.321: *B-ISDN Protocol Reference Model and its Application.* Geneva, 1991.

[41] CCITT, RECOMMENDATION I.352: *Network Performance Objectives for Connection Processing Delays in an ISDN.* Blue Book, Fascicle III.8, Melbourne, 1988.

[42] CCITT, RECOMMENDATION I.361: *B-ISDN ATM Layer Specification.* Geneva, 1992.

[43] CCITT, RECOMMENDATION I.362: *B-ISDN ATM Adaptation Layer (AAL) Functional Description.* Geneva, 1992.

[44] CCITT, RECOMMENDATION I.363: *B-ISDN ATM Adaptation Layer (AAL) Specification.* Geneva, 1992.

[45] CCITT, DRAFT RECOMMENDATION I.371: *Traffic Control and Resource Management in B-ISDN.* Geneva, 1992.

[46] CCITT, RECOMMENDATION I.432: *B-ISDN User Network Interface – Physical Layer Specification.* Geneva, 1991.

[47] CHEN, J.S.-C.; STERN, T.E.: Performance of a Generic Non-Blocking Packet Switch. *4th International Conference on Data Communication Systems and their Performance,* Barcelona, June 1990, pp. 99-115.

[48] CHEN, J.S.-C.; STERN, T.E.: Optimal Buffer Allocation for Packet Switches with Input and Output Queueing. *IEEE GLOBECOM'90,* San Diego, California, December 1990, pp. 1936-1941.

[49] CHEN, D.X.; MARK, J.W.: Performance Analysis of Output Buffered Fast Packet Switches with Bursty Traffic Loading. *IEEE GLOBECOM'91,* Vol. 1, Phoenix, December 1991, pp. 455-459.

[50] CHEN, X.: A Survey of Multistage Interconnection Networks in Fast Packet Switching. *International Journal of Digital and Analog Communication Systems,* Vol. 4, No. 1, January 1991, pp. 33-59.

[51] CHEN, X.; HAYES, J.F.: A Shared Buffer Memory Switch with Maximum Queue and Minimum Allocation. *Canadian Conference on Electrical and Computer Engineering,* Quebec, September 1991, pp. 7.1.1-7.1.4.

[52] CHEN, D.X.; MARK, J.W.: A Buffer Management Scheme for the SCOQ Switch under Nonuniform Traffic. *IEEE INFOCOM'92,* Vol. 1, Florence, May 1992, pp. 132-140.

[53] CHEN, J.S.-C.; GUÉRIN, R.; STERN, T.E.: Markov-Modulated Flow Model for the Output Queues of a Packet Switch. *IEEE Transactions on Communications,* Vol. 40, No. 6, June 1992, pp. 1098-1110.

[54] CINLAR, E.: Markov Renewal Theory. *Advances in Applied Probability*, Vol. 1, 1969, pp. 123-187.

[55] CINLAR, E.: *Introduction to Stochastic Processes.* Prentice Hall, Englewood Cliffs, New Jersey, 1975.

[56] CISNEROS, A.; BRACKETT, C.A.: A Large ATM Switch Based on Memory Switches and Optical Star Couplers. *IEEE Journal on Selected Areas in Communications*, Vol. 9, No. 8, October 1991, pp. 1348-1360.

[57] CLOS, C.: A Study of Non-Blocking Switching Networks. *Bell System Technical Journal*, Vol. 32, 1953, pp. 406-424.

[58] COOPER, R.B.: *Introduction to Queueing Theory.* Macmillan, New York, London, 1972.

[59] COPPO, P.; D'AMBROSIO, M.; MELEN, R.: Optimal Cost/Performance Design of ATM Switches. *IEEE INFOCOM'92*, Vol. 1, Florence, May 1992, pp. 446-458.

[60] COTTON, J.M.; GIESKEN, K.; LAWRENCE, A.; UPP, D.C.: ITT 1240 Digital Exchange Switching Network. *Electrical Communication*, Vol. 56, No. 2, 1981, pp. 148-160.

[61] DAIGLE, J.N.; LANGFORD, J.D.: Models for Analysis of Packet Voice Communications Systems. *IEEE Journal on Selected Areas in Communications*, Vol. 4, No. 6, September 1986, pp. 847-855.

[62] DAY, C.; GIACOPELLI, J.; HICKEY, J.: Applications of Self-Routing Switches to LATA Fiber Optic Networks. *ISS'87*, Phoenix, March 1987, pp. 519-523.

[63] DECINA, M.; GIACOMAZZI, P.; PATTAVINA, A.: Shuffleout Switching for ATM Interconnection Networks. *7th ITC Seminar: Broadband Technologies*, Morristown, October 1990, pp. 8.

[64] DE PRYCKER, M.; DE SOMER, M.: Performance of a Service Independent Switching Network with Distributed Control. *IEEE Journal on Selected Areas in Communications*, Vol. 5, No. 8, October 1987, pp. 1293-1301.

[65] DE PRYCKER, M.; BAUWENS, J.: A Broadband Central Office for ATM. *International Journal of Digital and Analog Cabled Systems*, Vol. 1, No. 2, 1988, pp. 87-94.

[66] DE PRYCKER, M.; ET AL: An ATM Switching Architecture with Intrinsic Multicast Capabilities for the Belgian Broadband Experiment. *ISS'90*, Vol. 5, Stockholm, May 1990, pp. 111-118.

[67] DE PRYCKER, M.; PESCHL, R.; VAN LANDEGEM, T.: B-ISDN and the OSI Protocol Reference Model. *Third IFIP WG 6.4 Conference on High Speed Networking, March 1991*, Berlin, March 1991, pp. 39-57.

[68] De Vries, R.J.F.: Gauss: A Simple High Performance Switch Architecture for ATM. *SIGCOMM'90 Symposium, Communications Architectures and Protocols*, Philadelphia, September 1990, pp. 126-134.

[69] De Vries, R.J.F.: ATM Multicast Connections using the Gauss Switch. *IEEE GLOBECOM'90*, San Diego, California, December 1990, pp. 211-217.

[70] Desmet, E.; Petit, G.H.: Performance Analysis of the Discrete Time Multiserver Queueing System Geo(N)/D/c/K. *RACE Workshop on Traffic and Performance Aspects in IBCN*, Munich, July 1990, pp. 20.

[71] Desmet, E.: Performance Assessment of an ATM Self-Routing Switching Network using Parallel Programming Techniques and a Network of Transputers. *International Computer and Communication Conference*, Beijing, October 1991.

[72] Dörken, H.; Habermann, R.; Sommer, M.: TMS - Transputer Management System. *3. Transputer Anwendertreffen, 1991*, in: Informatik Fachberichte, Springer Verlag, Hrsg. W. Brauer, Aachen, 1992, S. 338.

[73] Eckberg, A.E.; Hou, T.-C.: Effects of Output Buffer Sharing on Buffer Requirements in an ATDM Packet Switch. *IEEE INFOCOM'88*, New Orleans, March 1988, pp. 459-465.

[74] Endo, N.; Ohuchi, T.; Kozaki, T.; Kuwahara, H.; Mori, M.: Traffic Characteristics of a Shared Buffer ATM Switch. *IEEE GLOBECOM'90*, San Diego, California, December 1990, pp. 1913-1918.

[75] Endres, R.: *Wegesuche in ATM Koppelnetzen.* Semesterarbeit Nr. 1019 am Institut für Nachrichtenvermittlung und Datenverarbeitung, Universität Stuttgart, 1990.

[76] Filipiak, J.: Structured Systems Analysis Methodology for Design of an ATM Network Architecture. *IEEE Journal on Selected Areas in Communications*, Vol. 7, No. 8, October 1989, pp. 1263-1273.

[77] Fischer, W.; Huber, M.N.; Schrodi, K.J.: Performance Analysis of a Modular Videophone Switching Network. *ISS'87*, Vol. 3, Phoenix, March 1987, pp. 766-771.

[78] Fischer, W.; Fundneider, O.; Goeldner, E.H.; Lutz, K.A.: A Scalable ATM Switching System Architecture. *IEEE Journal on Selected Areas in Communications*, Vol. 9, No. 8, October 1991, pp. 1299-1307.

[79] Fried, J.: A VLSI Chip Set for Burst and Fast ATM Switching. *ICC'89*, Boston, June 1989, pp. 128-135.

[80] Gail, H.R.; Grover, G.; Guerin, R.; Hantler, S.L.: Buffer Size Requirements under Longest Queue First. *Intern. Conference on the Performance of Distributed Systems and Integrated Communication Networks*, Kyoto, September 1991, pp. 383-394.

[81] GARD, I.; ROOTH, J.: An ATM Switch Implementation - Technique and Technology. *ISS'90*, Vol. 4, Stockholm, May 1990, pp. 23-27.

[82] GARETTI, E.; ET AL: An Experimental ATM Switching Architecture for the Evolving B-ISDN Scenario. *ISS'90*, Vol. 4, Stockholm, May 1990, pp. 15-22.

[83] GIACOPELLI, J.N.; ET AL: Sunshine: A High-Performance Self-Routing Broadband Packet Switch Architecture. *IEEE Journal on Selected Areas in Communications*, Vol. 9, No. 8, October 1991, pp. 1289-1298.

[84] GIDRON, R.: TeraNet: A Multi-Gigabits per Second ATM Network. *Computer Communications*, Vol. 15, No. 3, April 1992, pp. 143-152.

[85] GIORCELLI, S.; DEMICHELIS, C.; GIANDONATO, G.; MELEN, R.: Experimenting with Fast Packet Switching Techniques in First Generation ISDN Environment. *ISS'87*, Phoenix, March 1987, pp. 388-394.

[86] GOKE, L.R.; LIPOVSKI, G.J.: Banyan Networks for Partitioning Multiprocessor Systems. *First Annual Symposium on Computer Architecture*, 1973, pp. 21-28.

[87] GOLI, P.; KUMAR, V.: Performance of a Crosspoint Buffered ATM Switch Fabric. *IEEE INFOCOM'92*, Vol. 1, Florence, May 1992, pp. 426-435.

[88] GRAVEY, A.; LOUVION, J.-R.; BOYER, P.: On the Geo/D/1 and Geo/D/1/n Queues. *Performance Evaluation*, Vol. 11, No. 2, 1990, pp. 117-125.

[89] GRAVEY, A.; HÉBUTERNE, G.: Simultaneity in Discrete-Time Single Server Queues with Bernoulli Inputs. *Performance Evaluation*, Vol. 14, No. 2, January 1992, pp. 123-131.

[90] GROSS, D.; HARRIS, C.M.: *Fundamentals of Queueing Theory.* John Wiley and Sons, New York, 1981.

[91] GUPTA, A.K.; BARBOSA, L.O.; GEORGANAS, N.D.: 16x16 Limited Intermediate Buffer Switch Module for ATM Networks. *IEEE GLOBECOM'91*, Vol. 2, Phoenix, December 1991, pp. 939-943.

[92] HABERMANN, R.; HEUER, P.: Verkehrstheoretische Untersuchung eines NxN-Koppelvielfachs in ATD-Koppelnetzen. *ntz-Archiv*, Band 10, Nr. 12, 1988, S. 325-330.

[93] HELD, A.: *Untersuchung von Prioritätsmechanismen für ATM Netze.* Diplomarbeit Nr. 964 am Institut für Nachrichtenvermittlung und Datenverarbeitung, Universität Stuttgart, 1989.

[94] HENRION, M.A.; ET AL: Switching Network Architecture for ATM-Based Broadband Communications. *ISS'90*, Vol. 5, Stockholm, May 1990, pp. 1-8.

[95] HEUER, P.: Konzept der virtuellen Wege in ATM-Netzen. *ITG-Fachbericht 115: Kommunikationsnetze der Zukunft*, Essen, Oktober 1990, S. 63-69.

[96] HLUCHYJ, M.G.; KAROL, M.J.: Queueing in High-Performance Packet Switching. *IEEE Journal on Selected Areas in Communications*, Vol. 6, No. 9, December 1988, pp. 1587-1597.

[97] HÖGLINGER, F.: *Simulative Untersuchung von ATM-Koppelelementen.* Semesterarbeit Nr. 1034 am Institut für Nachrichtenvermittlung und Datenverarbeitung, Universität Stuttgart, 1990.

[98] HUANG, A.; KNAUER, S.: STARLITE: A Wideband Digital Switch. *IEEE GLOBECOM'84*, November 1984, pp. 121-125.

[99] HUI, J.Y.; ARTHURS, E.: A Broadband Packet Switch for Integrated Transport. *IEEE Journal on Selected Areas in Communications*, Vol. 5, No. 8, October 1987, pp. 1264-1273.

[100] HUI, J.Y.: *Switching and Traffic Theory for Integrated Broadband Networks.* Kluwer Academic Publishers, Boston/Dordrecht/London, 1990.

[101] HUI, J.Y.: Switching Integrated Broadband Services by Sort-Banyan Networks. *Technical Report 90-04, Department of Electrical and Computer Engineering, Rutgers University*, Piscataway, 1990, pp. 23.

[102] HÜBNER, F.; TRAN-GIA, P.: Quasi-Stationary Analysis of a Finite Capacity Asynchronous Multiplexer with Modulated Deterministic Input. *ITC 13, Proceedings in 'North Holland Studies in Telecommunication'*, Vol. 14, Copenhagen, June 1991, pp. 723-729.

[103] ILIADIS, I.; DENZEL, W.E.: Performance of Packet Switches with Input and Output Queueing. *ICC'90*, Atlanta, April 1990, pp. 747-753.

[104] ILIADIS, I.: Performance of a Packet Switch with Shared Buffer and Input Queueing. *ITC 13, Proceedings in 'North Holland Studies in Telecommunication'*, Vol. 14, Copenhagen, June 1991, pp. 911-916.

[105] ILIADIS, I.: Head of the Line Arbitration of Packet Switches with Combined Input and Output Queueing. *International Journal of Digital and Analog Communication Systems*, Vol. 4, No. 3, July 1991, pp. 181-190.

[106] ILIADIS, I.: Performance of a Packet Switch with Input and Output Queueing under Unbalanced Traffic. *IEEE INFOCOM'92*, Florence, May 1992, pp. 743-752.

[107] ISO 7498: *Information Processing Systems – Open Systems Interconnection – Basic Reference Model.* International Standard, 1984.

[108] ITOH, A.; ET AL: Practical Implementation and Packaging Technologies for a Large-Scale ATM Switching System. *IEEE Journal on Selected Areas in Communications*, Vol. 9, No. 8, October 1991, pp. 1280-1288.

[109] KALTENMORGEN, B.; KLEINEWILLINGHÖFER-KOPP, R.: Verkehrskontrolle in ATM-Netzen. *Der Fernmelde-Ingenieur*, Band 46, Nr. 1, Januar 1992, S. 1-33.

[110] KAMOUN, F.; KLEINROCK, L.: Analysis of Shared Finite Storage in a Computer Network Node Environment under General Traffic Conditions. *IEEE Transactions on Communications*, Vol. 28, No. 7, July 1980, pp. 992-1003.

[111] KAROL, M.J.; HLUCHYJ, M.G.; MORGAN, S.P.: Input Versus Output Queueing on a Space-Division Packet Switch. *IEEE Transactions on Communications*, Vol. 35, No. 12, December 1987, pp. 1347-1356.

[112] KAROL, M.J.; ENG, K.Y.; OBARA, H.: Improving the Performance of Input-Queued ATM Packet Switches. *IEEE INFOCOM'92*, Vol. 1, Florence, May 1992, pp. 110-115.

[113] KATEVENIS, M.; SIDIROPOULOS, S.; COURCOUBETIS, C.: Weighted Round-Robin Cell Multiplexing in a General-Purpose ATM Switch Chip. *IEEE Journal on Selected Areas in Communications*, Vol. 9, No. 8, October 1991, pp. 1265-1279.

[114] KATO, Y.; SHIMOE, T.; HAJIKANO, K.; MURAKAMI, K.: Experimental Broadband ATM Switching System. *IEEE GLOBECOM'88*, Hollywood, December 1988, pp. 1288-1292.

[115] KATO, Y.; HAYAMI, H.; KAMOI, J.; MURAKAMI, K.: A Bi-CMOS ATM Switching LSI. *International Zürich Seminar*, Zürich, March 1990, pp. 195-210.

[116] KERMANI, P.; KLEINROCK, L.: Virtual Cut-Through: A New Computer Communication Switching Technique. *Computer Networks*, Vol. 3, No. 4, 1979, pp. 267-286.

[117] KILLAT, U.; ET AL: A Versatile ATM Switch Concept. *ISS'90*, Vol. 4, Stockholm, May 1990, pp. 127-134.

[118] KIM, H.S.; LEON-GARCIA, A.: A Self-Routing Multistage Switching Network for Broadband ISDN. *IEEE Journal on Selected Areas in Communications*, Vol. 8, No. 3, April 1990, pp. 459-466.

[119] KIM, H.S.; LEON-GARCIA, A.: Performance of Buffered Banyan Networks under Nonuniform Traffic Patterns. *IEEE Transactions on Communications*, Vol. 38, No. 5, May 1990, pp. 648-658.

[120] KIM, H.S.; LEON-GARCIA, A.: Nonblocking Property of Reverse Banyan Networks. *IEEE Transactions on Communications*, Vol. 40, No. 3, March 1992, pp. 472-476.

[121] KLEINROCK, L.: *Queueing Systems.* John Wiley and Sons, New York, 1975.

[122] KOBAYASHI, H.: *Geometric Bounds on the Waiting Time Distribution of the Discrete-Time GI/G/1-System.* in: Probability Theory and Computer Science, Part II, edited by G. Louchard & G. Latouche, Academic Press, London, 1983.

[123] KOWALK, W.: Numerically Stable Method for the Evaluation of Markov Chains. *Document PFH-123-0012-CD-CC*, RACE 1022, August 1989.

[124] KOZAKI, T.; ET AL: 32x32 Shared Buffer Type ATM Switch VLSI for B-ISDN's. *IEEE Journal on Selected Areas in Communications*, Vol. 9, No. 8, October 1991, pp. 1239-1247.

[125] KRUSKAL, C.P.; SNIR, M.; WEISS, A.: The Distribution of Waiting Times in Clocked Multistage Interconnection Networks. *IEEE Transactions on Computers*, Vol. 37, No. 11, November 1988, pp. 1337-1352.

[126] KRÖNER, H.; THEIMER, T.; BRIEM, U.: Queueing Models for ATM Systems - A Comparison. *7th ITC Seminar: Broadband Technologies*, Morristown, October 1990, pp. 8.

[127] KRÖNER, H.: Statistical Multiplexing of Sporadic Sources - Exact and Approximate Performance Analysis. *ITC 13, Proceedings in 'North Holland Studies in Telecommunication'*, Vol. 14, Copenhagen, June 1991, pp. 787-793.

[128] KRÖNER, H.; HÉBUTERNE, G.; BOYER, P.; GRAVEY, A.: Priority Management in ATM Switching Nodes. *IEEE Journal on Selected Areas in Communications*, Vol. 9, No. 3, April 1991, pp. 418-427.

[129] KRÖNER, H.; EBERSPÄCHER, M.; THEIMER, T.; KÜHN, P.J.; BRIEM, U.: Approximate Analysis of the End-to-End Delay in ATM Networks. *IEEE INFOCOM'92*, Florence, May 1992, pp. 978-986.

[130] KÜHN, P.J.: Input Process Model GMDP. *Document UST-123-0003-CD-CC*, RACE 1022, June 1988.

[131] KUMAR, V.P.; KNEUER, J.G.; PAL, D.; BRUNNER, B.: PHOENIX: A Building Block for Fault Tolerant Broadband Packet Switches. *IEEE GLOBECOM'91*, Vol. 1, Phoenix, December 1991, pp. 228-233.

[132] LAMPE, D.: Transfer Delay Deviation of Packets in ATD Switching Matrices and its Effect on Dimensioning a Depacketizer Buffer. *ICCC'88*, Tel Aviv, October 1988, pp. 55-60.

[133] LANGER, K.-D.: Konzept eines Koppelelements für die Hochratige ATM-Vermittlung. *Verbindungsnetzwerke für Parallelrechner und Breitband-Übermittlungssysteme*, ITG/GI Workshop, Stuttgart, September 1991, S. 42-47.

[134] LAVENBERG, S.S. (ED.): *Computer Performance Modeling Handbook*. Academic Press, New York, 1983.

[135] LAWRIE, D.H.: Access and Alignment of Data in an Array Processor. *IEEE Transactions on Computers*, Vol. C-24, 12, 1975, pp. 1145-1155.

[136] LEE, T.T.; BOORSTYN, R.; ARTHURS, E.: The Architecture of a Multicast Broadband Packet Switch. *IEEE INFOCOM'88*, New Orleans, March 1988, pp. 1-8.

[137] LEE, H.; KOOK, K.H.; RIM, C.S.; JUN, K.P.; LIM, S.K.: A Limited Shared Output Buffer Switch for ATM. *4th International Conference on Data Communication Systems and their Performance*, Barcelona, June 1990, pp. 163-179.

[138] LEE, M.J.; LI, S.-Q.: Performance Trade-Offs in Input/Output Buffer Design for a Non-Blocking Space-Division Packet Switch. *International Journal of Digital and Analog Communication Systems*, Vol. 4, No. 1, January 1991, pp. 21-31.

[139] LEHNERT, R.: On Performance Aspects of Broadband-Switches Using the Asynchronous Transfer Mode. *Proc. Nachrichtentechnisches Kolloquium, Universität Braunschweig, Teubner Verlag*, Braunschweig, 1988, pp. 69-93.

[140] LIAO, K.-Q.; MASON, L.G.: A Heuristic Approach for Performance Analysis of ATM Systems. *IEEE GLOBECOM'90*, San Diego, December 1990.

[141] LIEW, S.C.; LU, K.W.: A Three-Stage Architecture for Very Large Packet Switches. *International Journal of Digital and Analog Cabled Systems*, Vol. 2, No. 4, October 1989, pp. 303-316.

[142] LIEW, S.C.: Multicast Routing Algorithms for 3-Stage Clos ATM Switching Networks. *IEEE GLOBECOM'91*, Vol. 3, Phoenix, December 1991, pp. 1619-1625.

[143] LIN, A.Y.M.; SILVESTER, J.A.: Priority Queueing Strategies and Buffer Allocation Protocols for Traffic Control at an ATM Integrated Broadband Switching System. *IEEE Journal on Selected Areas in Communications*, Vol. 9, No. 9, December 1991, pp. 1524-1536.

[144] LIN, T.; KLEINROCK, L.: Performance Analysis of Finite-Buffered Multistage Interconnection Networks with a General Traffic Pattern. *1991 ACM SIGMETRICS*, San Diego, May 1991, pp. 68-78.

[145] LINDBERGER, K.: Analytical Methods for the Traffical Problems with Statistical Multiplexing in ATM-Networks. *ITC 13, Proceedings in 'North Holland Studies in Telecommunication'*, Vol. 14, Copenhagen, June 1991, pp. 807-813.

[146] LUCANTONI, D.M.; MEIER-HELLSTERN, K.S.; NEUTS, M.F.: A Single-Server Queue with Server Vacations and a Class of Non-Renewal Arrival Processes. *Advances in Applied Probability*, Vol. 22, No. 1, 1989, pp. 676-705.

[147] LUCANTONI, D.M.: New Results on the Single Server Queue with a Batch Markovian Arrival Process. *Communications in Statistics - Stochastic Models*, Vol. 7, No. 1, 1991, pp. 1-46.

[148] LUTZ, K.A.: ATM ermöglicht unterschiedliche Bitraten im einheitlichen Breitbandnetz. *Siemens Telcom Report*, Band 11, Nr. 6, 1988, S. 210-213.

[149] LUTZ, K.A.: Considerations on ATM Switching Techniques. *International Journal of Digital and Analog Cabled Systems*, Vol. 1, No. 4, 1988, pp. 237-243.

[150] MAGLARIS, B.; ET AL: Performance Models of Statistical Multiplexing in Packet Video Communications. *IEEE Transactions on Communications*, Vol. 36, No. 7, July 1988, pp. 834-844.

[151] MEDOW, R.: *Untersuchung und Entwurf eines Ein-Chip ATM Koppelelementes.* Diplomarbeit Nr. 1014 am Institut für Nachrichtenvermittlung und Datenverarbeitung, Universität Stuttgart, 1990.

[152] MEISLING, T.: Discrete Time Queueing Theory. *Operations Research*, Vol. 6, 1958, pp. 96-105.

[153] MELEN, R.; TURNER, J.S.: Nonblocking Networks for Fast Packet Switching. *IEEE INFOCOM'89*, Ottawa, Ontario, April 1989, pp. 548-557.

[154] MELEN, R.; TURNER, J.S.: Nonblocking Multirate Distribution Networks. *IEEE INFOCOM'90*, San Francisco, CA, June 1990, pp. 1234-1241.

[155] MELEN, R.: Current Architectures for ATM Implementation. *European Transactions on Telecommunications*, Vol. 3, No. 2, March 1992, pp. 53-63.

[156] MORRIS, T.D.; PERROS, H.G.: Performance Analysis of a Multi-Buffered Banyan ATM Switch Under Bursty Traffic. *Technical Report, Department of Computer Science, North Carolina State University*, Raleigh, July 1991, pp. 30.

[157] NEUTS, M.F.: *Structured Stochastic Matrices of M/G/1 Type and their Applications.* Marcel Dekker, New York and Basel, 1989.

[158] NEWMAN, P.: A Fast Packet Switch for the Integrated Services Backbone Network. *IEEE Journal on Selected Areas in Communications*, Vol. 6, No. 9, December 1988, pp. 1468-1479.

[159] NIESTEGGE, G.: Nonblocking Multirate Switching Networks. *5th ITC Seminar: Traffic Engineering for ISDN Design and Planning*, Lake Como, May 1987, pp. 10.

[160] NISHIKIDO, J.; OKUNO, M.; HIMENO, A.: Optical Routing Control using Coherent Pattern-Matching Circuit for Photonic Self-Routing Switch. *IEEE INFOCOM'92*, Vol. 1, Florence, May 1992, pp. 459-467.

[161] NOJIMA, S.; TSUTSUI, E.; FUKUDA, H.; HASHIMOTO, M.: Integrated Services Packet Network Using Bus Matrix Switch. *IEEE Journal on Selected Areas in Communications*, Vol. 5, No. 8, October 1987, pp. 1284-1291.

[162] OBARA, H.; YASUSHI, T.: An Efficient Contention Resolution Algorithm for Input Queueing ATM Cross-Connect Switches. *International Journal of Digital and Analog Cabled Systems*, Vol. 2, No. 4, October 1989, pp. 261-267.

[163] OIE, Y.; MURATA, M.; KUBOTA, K.; MIYAHARA, H.: Effect of Speedup in Nonblocking Packet Switch. *ICC'89*, Boston, June 1989, pp. 410-414.

[164] OIE, Y.; SUDA, T.; MURATA, M.; MIYAHARA, H.: Survey of the Performance of Non-Blocking Switches with FIFO Input Buffers. *ICC'90*, Atlanta, GA, April 1990, pp. 737-741.

[165] OIE, Y.; MURATA, M.; KUBOTA, K.; MIYAHARA, H.: Performance Analysis of Nonblocking Packet Switch with Input and Output Buffers. *IEEE Transactions on Communications*, Vol. 40, No. 8, August 1992, pp. 1294-1297.

[166] PALMER, R.: An Experimental ATM Switch for BISDN Studies. *International Journal of Digital and Analog Communication Systems*, Vol. 3, No. 4, 1990, pp. 341-349.

[167] PATTAVINA, A.: Multichannel Bandwidth Allocation in a Broadband Packet Switch. *IEEE Journal on Selected Areas in Communications*, Vol. 6, No. 9, December 1988, pp. 1489-1499.

[168] PATTAVINA, A.: Performance Evaluation of ATM Switches with Input and Output Queueing. *International Journal of Digital and Analog Communication Systems*, Vol. 3, No. 3, 1990, pp. 277-286.

[169] PATTAVINA, A.: Broadband Switching Systems: First Generation. *European Transactions on Telecommunications*, Vol. 2, No. 1, January 1991, pp. 75-87.

[170] PATTAVINA, A.: An ATM Switch Architecture for Provision of Integrated Broadband Services. *IEEE Journal on Selected Areas in Communications*, Vol. 9, No. 9, December 1991, pp. 1537-1548.

[171] PHILLIPS, C.I.; CUTHBERT, L.G.: Concurrent Discrete Event-Driven Simulation Tools. *IEEE Journal on Selected Areas in Communications*, Vol. 9, No. 3, April 1991, pp. 477-485.

[172] PROCTOR, R.J.; MADDERN, T.S.: Synchronous ATM Switching Fabrics. *ISS'90*, Vol. 4, Stockholm, May 1990, pp. 109-114.

[173] RASMUSSEN, C.; SORENSEN, J.; KVOLS, K.S.; JACOBSEN, S.B.: Source-Independent Call Acceptance Procedures in ATM Networks. *IEEE Journal on Selected Areas in Communications*, Vol. 9, No. 3, April 1991, pp. 351-358.

[174] RATHGEB, E.P.; THEIMER, T.H.; HUBER, M.N.: ATM Switches - Basic Architectures and their Performance. *International Journal of Digital and Analog Cabled Systems*, Vol. 2, No. 4, October 1989, pp. 227-236.

[175] RATHGEB, E.P.: *Verkehrsflüsse in ATM-Netzen – Modellierung und Analyse von Verkehrsquellen und Quellflusskontrollverfahren.* Dissertation am Institut für Nachrichtenvermittlung und Datenverarbeitung, Universität Stuttgart, 1991.

[176] REITER, A.: *Untersuchung von Koppelnetzstrukturen für ATM.* Diplomarbeit Nr. 949 am Institut für Nachrichtenvermittlung und Datenverarbeitung, Universität Stuttgart, 1989.

[177] ROSS, F.E.: An Overview of FDDI: The Fiber Distributed Data Interface. *IEEE Journal on Selected Areas in Communications*, Vol. 7, No. 7, September 1989.

[178] ROTHERMEL, K.; SEEGER, D.: Traffic Studies of Switching Networks for Asynchronous Transfer Mode. *12th International Teletraffic Congress*, Torino, June 1988, pp. 1.3A5.

[179] ROTHERMEL, K.: Priority Mechanisms in ATM Networks. *IEEE GLOBECOM'90*, San Diego, December 1990, pp. 847-851.

[180] SAITO, H.; KAWASHIMA, K.; SATO, K.: Traffic Control Technologies in ATM Networks. *IEICE Transactions*, Vol. E74, No. 4, April 1991, pp. 761-771.

[181] SAKURAI, Y.; IDO, N.; GOHARA, S.; ENDO, N.: Large Scale ATM Multi-Stage Switching Network with Shared Buffer Memory Switches. *ISS'90*, Vol. 4, Stockholm, May 1990, pp. 121-126.

[182] SALTER, J.; EVANS, P.: The IEEE 802.6 Standard for MANs, its Scope and Purpose. *EFOC/LAN 90*, Paper 4.6.4, Munich, June 1990.

[183] SCHOUTE, F.C.: Simple Decision Rules for Acceptance of Mixed Traffic Streams. *12th International Teletraffic Congress*, Torino, June 1988, pp. 4.2A5.

[184] SCHWARTZ, D.B.; SAMALAM, V.K.: VLSI for ATM Network Management. *IEEE GLOBECOM'90*, San Diego, California, December 1990, pp. 1387-1392.

[185] SCHWARZ, A.; THEIMER, T.H.: Simulative Untersuchungen an einem ATM-Koppelelement mit Zentralspeicher. *Verbindungsnetzwerke für Parallelrechner und Breitband-Übermittlungssysteme*, ITG/GI Workshop, Stuttgart, September 1991, pp. 95-100.

[186] SCHWARZ, A.: *Untersuchung und VLSI-Entwurf einer Kontrolleinheit für ATM-Schaltelemente mit Zentralspeicher.* Diplomarbeit Nr. 1115 am Institut für Nachrichtenvermittlung und Datenverarbeitung, Universität Stuttgart, 1991.

[187] SEN, P.; MAGLARIS, B.; RIKLI, N.-E.; ANASTASSIOU, D.: Models for Packet Switching of Variable-Bit-Rate Video Sources. *IEEE Journal on Selected Areas in Communications*, Vol. 7, No. 5, September 1989, pp. 865-869.

[188] SENNEMA, R.T.: Switching Based on Crosspoints. *Document APT-311&321-0020-CD-CC*, RACE 1022, December 1988.

[189] SEZAKI, K.; TANAKA, Y.; AKIYAMA, M.: The Cascade Clos Broadcast Switching Network - A New ATM Switching Network which is Multiconnection Non-Blocking. *ISS'90*, Vol. 4, Stockholm, May 1990, pp. 143-147.

[190] SEZAKI, K.; TANAKA, Y.; AKIYAMA, M.: A New ATM Switching Network which is Robust for Multicast. *IEICE Transactions*, Vol. E74, No. 9, September 1991, pp. 2779-2784.

[191] SHIOMOTO, K.; MURATA, M.; OIE, Y.; MIYAHARA, H.: Performance Evaluation of Cell Bypass Queueing Discipline for Buffered Banyan Type ATM Switches. *IEEE INFOCOM'90*, Vol. 2, San Francisco, June 1990, pp. 677-685.

[192] SHOBATAKE, Y.; ET AL: A One-Chip Scalable 8x8 ATM Switch LSI Employing Shared Buffer Architecture. *IEEE Journal on Selected Areas in Communications*, Vol. 9, No. 8, October 1991, pp. 1248-1254.

[193] SIEGEL, H.J.: Interconnection Networks for Large-Scale Parallel Processing. *Lexington*, 1985.

[194] SIEGEL, M.; MEDOW, R.; THEIMER, T.; SCHWEDERSKI, T.: Design of a Single-Chip ATM Switching Element. *ICCC'90*, New Dehli, November 1990, pp. 804-811.

[195] SUMITA, S.; OZAWA, T.: Achievability of Performance Objectives in ATM Switching Nodes. *International Seminar on Performance of Distributed and Parallel Systems*, Kyoto, December 1988, pp. 45-56.

[196] SUZUKI, H.; NAGANO, H.; SUZUKI, T.; TAKEUCHI, T.; IWASAKI, S.: Output-Buffer Switch Architecture for Asynchronous Transfer Mode. *International Journal of Digital and Analog Cabled Systems*, October 1989, pp. 269-276.

[197] SVINNSET, I.: Nonblocking ATM Switching Networks. *ITC 13, Proceedings in 'North Holland Studies in Telecommunication'*, Vol. 14, Copenhagen, June 1991, pp. 1011-1016.

[198] TAKACS, L.: Delay Distributions for One Line with Poisson Input, General Holding Times, and Various Orders of Service. *Bell System Technical Journal*, Vol. 42, March 1963, pp. 487-503.

[199] TAKACS, L.: *Combinatorial Methods in the Theory of Stochastic Processes.* John Wiley & Sons, New York, 1967.

[200] TAKAHASHI, A.; NISHINO, T.; MURAKAMI, K.; DUNNING, S.: A Broadband Switching System for Public Network. *ISS'90*, Vol. 5, Stockholm, May 1990, pp. 103-109.

[201] TAKAHASHI, Y.; ANDO, K.; INOUE, H.; AMADA, E.: An Experimental Photonic ATM Switching System Using an InP Semiconductor Optical Switch. *IEEE GLOBECOM'90*, San Diego, California, December 1990, pp. 1586-1589.

[202] TAKEUCHI, T.; SUZUKI, H.; ARAMAKI, T.: Switch Architectures and Technologies for Asynchronous Transfer Mode. *IEICE Transactions*, Vol. 74, No. 4, April 1991, pp. 752-760.

[203] TAMIR, Y.; FRAZIER, G.: High-Performance Multiqueue Buffers for VLSI Communication Switches. *15th Annual International Symposium on Computer Architecture*, Honolulu, May 1988, pp. 343-354.

[204] THEIMER, T.H.: How to Compute the Moments of a GMDP. *Document UST-123-0023-CD-CC*, RACE 1022, November 1989.

[205] THEIMER, T.H.: Performance Comparison of Routing Strategies in ATM Switch Fabrics. *ITC 13, Proceedings in 'North Holland Studies in Telecommunication'*, Vol. 14, Copenhagen, June 1991, pp. 923-928.

[206] THOMAS, A.; COUDREUSE, J.P.; SERVEL, M.: Asynchronous Time-Division Techniques: An Experimental Packet Network Integrating Videocommunication. *ISS'84*, Florence, May 1984, pp. 32 C 2.

[207] TOBAGI, F.A.: Fast Packet Switch Architectures for Broadband Integrated Services Digital Networks. *Proceedings of the IEEE*, Vol. 78, No. 1, January 1990, pp. 133-167.

[208] TOBAGI, F.A.; KWOK, T.; CHIUSSI, F.M.: Architecture, Performance, and Implementation of the Tandem Banyan Fast Packet Switch. *IEEE Journal on Selected Areas in Communications*, Vol. 9, No. 8, October 1991, pp. 1173-1193.

[209] TRAN-GIA, P.; AHMADI, H.: Analysis of a Discrete-Time G[X]/D/1-S Queueing System with Applications in Packet-Switching Systems. *INFOCOM'88*, New Orleans, March 1988, pp. 9A.1.

[210] TRAN-GIA, P.; GOLD, H.: Anwendung Neuronaler Netze in Kommunikationssystemen. *Workshop "Neuronale Netze", Deutsche Informatik Akademie der GI*, Bonn, May 1990, pp. 8.

[211] TROUDET, T.P.; WALTERS, S.M.: Neural Network Architecture for Crossbar Switch Control. *IEEE Transactions on Circuits and Systems*, Vol. 38, No. 1, January 1991, pp. 42-56.

[212] TU, S.-C.; LEUNG, W.-H.: Multicast Connection-Oriented Packet Switching Networks. *ICC'90*, Atlanta, GA, April 1990, pp. 495-501.

[213] TURNER, J.S.: Design of an Integrated Services Packet Network. *9th Data Communications Symposium, ACM SIGCOMM Computer Communication Review*, Vol. 15, No. 4, September 1985, pp. 124-133.

[214] VILLEN-ALTAMIRANO, M.; VILLEN-ALTAMIRANO, J.: RESTART: A Method for Accelerating Rare Event Simulations. *13th International Teletraffic Congress, Workshop on Queueing, Performance and Control in ATM*, Copenhagen, June 1991, pp. 71-76.

[215] WEI, S.X.; COYLE, E.J.; HSIAO, M.-T.T.: An Optimal Buffer Management Policy for High-Performance Packet Switching. *IEEE GLOBECOM'91*, Vol. 2, Phoenix, December 1991, pp. 924-928.

[216] WU, C.; FENG, T.-Y.: On a Class of Multistage Interconnection Networks. *IEEE Transactions on Computers*, Vol. 29, No. 8, August 1980, pp. 694-702.

[217] WU, L.T.: Mixing Traffic in a Buffered Banyan Network. *Ninth Data Communications Symposium, ACM SIGCOMM, Computer Communication Review*, Vol. 15, No. 4, 1985, pp. 134-139.

[218] WULLEMAN, R.; VAN LANDEGEM, T.: Comparison of ATM Switching Architectures. *International Journal of Digital and Analog Cabled Systems*, Vol. 2, No. 4, October 1989, pp. 211-225.

[219] YAMADA, H.; YAMADA, S.; KAI, H.; TAKAHASHI, T.: A Multi-Purpose Memory Switch LSI for ATM-Based Systems. *IEEE GLOBECOM'90*, San Diego, December 1990, pp. 1602-1608.

[220] YAMASHITA, H.; PERROS, H.G.; HONG, S.-W.: Performance Modelling of a Shared Buffer ATM Switch Architecture. *ITC 13, Proceedings in 'North Holland Studies in Telecommunication'*, Vol. 14, Copenhagen, June 1991, pp. 993-998.

[221] YANG, Y.; MASSON, G.M.: Nonblocking Broadcast Switching Networks. *IEEE Transactions on Computers*, Vol. 40, No. 9, September 1991, pp. 1005-1015.

[222] YEH, Y.S.; HLUCHYJ, M.G.; ACAMPORA, A.S.: The Knockout Switch: A Simple, Modular Architecture for High Performance Packet Switching. *IEEE Journal on Selected Areas in Communications*, Vol. 5, No. 8, October 1987, pp. 1274-1283.

[223] YOUNG, K.C.; BANWELL, T.C.; CHENG, S.S.; ESTES, R.C. ET AL.: Physical Design Issues for Very Large ATM Switching Systems. *IEEE GLOBECOM'90*, San Diego, California, December 1990, pp. 1590-1593.

Anhang A

Analyse des zeitdiskreten $D^{[X]}/GI/1$-Systems

A.1 Modellbeschreibung

Das $D^{[X]}/GI/1$-System ist ein zeitdiskretes Warteschlangenmodell mit Gruppenankünften. Der Ankunftsprozeß ist die zeitdiskrete Version eines Poisson-Prozesses mit Gruppenankünften (siehe z.B. [122]). In jedem Zeitschlitz kommt eine Gruppe von $\Xi \geq 0$ Anforderungen an, wobei die in aufeinanderfolgenden Zeitschlitzen eintreffenden Gruppen statistisch voneinander unabhängig sind. Der Ankunftsprozeß ist also gedächtnisfrei. Die Gruppengröße besitzt die Verteilung $\xi_k = \mathrm{P}\{\Xi = k\}$, mit der Erzeugenden Funktion $\xi(z) = \sum_{k=0}^{\infty} \xi_k \, z^k$. Die Ankunftsrate ist $\lambda = \mathrm{E}[\Xi] = \xi'(1)$.

Die wartenden Anforderungen werden von der Bedieneinheit in der Reihenfolge ihres Eintreffens (Abfertigungsdisziplin FIFO) bearbeitet. Bei gleichem Ankunftszeitpunkt erfolgt die Bearbeitung der betreffenden Anforderungen in zufälliger Reihenfolge. Die Bedienzeit H der Anforderungen ist unabhängig vom Zustand des Systems und besitzt die Wahrscheinlichkeitsverteilung $h_k = \mathrm{P}\{H = k\}, k > 0$, mit der Erzeugenden Funktion $h(z) = \sum_{k=1}^{\infty} h_k \, z^k$. Der Fall $H = 0$ entspricht der Bedienung von mehreren Anforderungen in einem Zeitschlitz und wird hier absichtlich ausgeschlossen. Aus der mittleren Bedienzeit $\mathrm{E}[H] = h'(1)$ und der Ankunftsrate λ ergibt sich das Angebot zu $A = \lambda \mathrm{E}[H] = \xi'(1)h'(1)$.

Bei zeitdiskreten Verkehrsmodellen können Ankunfts- und Bedienendeereignisse gleichzeitig, das heißt im selben Zeitschlitz, auftreten. Die Reihenfolge, in der diese Ereignisse bearbeitet werden, muß daher ebenfalls festgelegt werden, da das Systemverhalten insbesondere bei endlicher Warteschlangenlänge von der Bearbeitungsreihenfolge abhängig ist. Aus der Literatur sind die beiden folgenden Varianten bekannt [89, 93]:

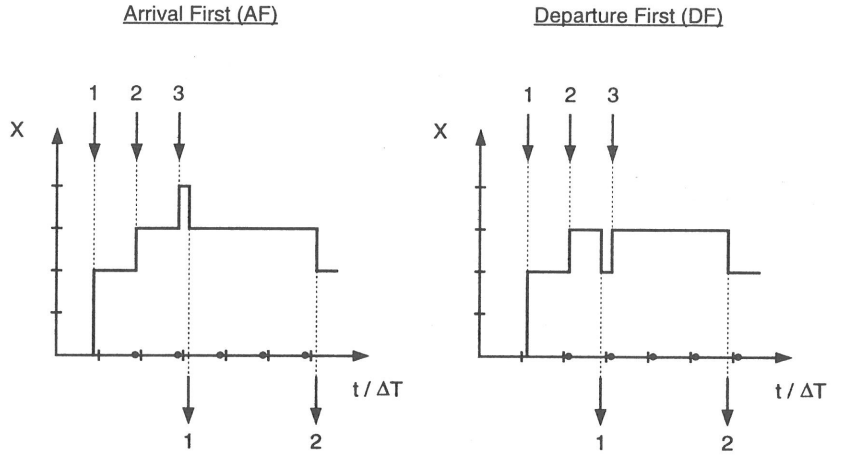


Bild A.1: Bearbeitungsreihenfolge in zeitdiskreten Warteschlangenmodellen; das Bild zeigt die Anzahl der Anforderungen im System (X) über der Zeit

AF: Das Ankunftsereignis wird vor dem Bedienende bearbeitet (Arrival First, AF).
DF: Das Bedienende wird vor dem Ankunftsereignis bearbeitet. (Departure First, DF).

In Bild A.1 sind gleichzeitig auftretende Ereignisse absichtlich nacheinander eingezeichnet, um die beiden Varianten AF und DF zu veranschaulichen. Mögliche Ankunftszeitpunkte sind außerdem durch Punkte auf der Zeitachse markiert. Man erkennt, daß sich die Varianten während der Bedienzeit der ersten Anforderung im Anschluß an ein leeres System unterscheiden. Bei der Bearbeitungsreihenfolge AF finden die Ankünfte 2 und 3 noch innerhalb der Bedienzeit dieser Anforderung statt, wogegen im DF-System die Ankunft 3 erst nach dem Bedienende der ersten Anforderung bearbeitet wird. Die Bearbeitungsreihenfolge muß folglich bei der Berechnung der Zustandswahrscheinlichkeiten des Systems beachtet werden.

Das oben beschriebene Warteschlangenmodell ist ein Sonderfall des in [24] untersuchten D-BMAP/GI/1/K-Systems. Für dieses sehr allgemeine Modell, dessen Ankunftsprozeß ein zeitdiskreter MAP [146, 147] mit Gruppenankünften ist, wurden in [24] die Zustandswahrscheinlichkeiten nach einem Bedienende sowie die zeitgemittelten Zustandswahrscheinlichkeiten für den Fall der endlichen Warteschlangenlänge bestimmt. Die Wartezeitverteilung des $D^{[X]}$/GI/1-Systems mit unbegrenzter Warteschlangenlänge ist in [125] hergeleitet. Ein Sonderfall des hier untersuchten Modells ist das Geo/GI/1-System, für das in [122, 152] die Zustandswahrscheinlichkeiten berechnet wurden. In [88] werden die Zustandswahrscheinlichkeiten sowie die Wartezeitverteilung für das Geo/D/1-System mit begrenzter Warteschlangenlänge bestimmt. Für den Fall der konstanten Bedienzeit sind in [175] und [209] zwei numerische Verfahren zur Analyse des endlichen $GI^{[X]}$/D/1-Systems angegeben.

A.2 Analyse des Systems mit unbegrenztem Wartespeicher

Da der Ankunftsprozeß des oben beschriebenen Systems gedächtnisfrei ist, kann die Analyse mit Hilfe der Methode der eingebetteten Markov-Kette durchgeführt werden. Dazu wird das System zu bestimmten Zeitpunkten betrachtet, den sogenannten Regenerationszeitpunkten, an denen seine zukünftige Entwicklung nur vom gegenwärtigen Zustand abhängt.

Für das betrachtete Modell bieten sich hier die Zeitpunkte unmittelbar nach einem Bedienende an, da zu diesem Zeitpunkt die gedächtnisbehaftete Bedienphase abgeschlossen ist. Zu diesen Zeitpunkten erfüllt das System also die Voraussetzungen für die Einbettung einer Markov-Kette.

A.2.1 Zustandswahrscheinlichkeiten nach einem Bedienende

Der Zustand des Systems zu den Regenerationszeitpunkten wird durch die Anzahl X der Anforderungen im System (Warteschlange und Bedieneinheit) unmittelbar nach einem Bedienende beschrieben. Die Verteilung von X sei $x_k = \mathrm{P}\{X = k\}$, und deren Erzeugende Funktion sei $x(z) = \sum_{k=0}^{\infty} x_k \, z^k$. Zur Bestimmung der Verteilung von X werden nun alle möglichen Übergänge von einem Bedienende zum nächsten betrachtet.

Zwischen zwei Bedienendezeitpunkten kann X im allgemeinen um beliebig viele Anforderungen zunehmen, aber um höchstens eine Anforderung abnehmen. Es gilt also

$$X_{n+1} = \max(0, X_n + A_n - 1), \tag{A.1}$$

wobei X_n den Systemzustand nach dem n-ten Bedienende bezeichnet, und A_n die Anzahl der zwischen dem n-ten und dem $(n+1)$-ten Bedienende eingetroffenen Anforderungen angibt.

Eine graphische Darstellung der möglichen Übergänge zwischen zwei Regenerationszeitpunkten mit den zugehörigen Übergangswahrscheinlichkeiten zeigt Bild A.2. Bei der Bestimmung der Übergangswahrscheinlichkeiten ist zu beachten, daß der Ausgangszustand 0 im allgemeinen getrennt behandelt werden muß.

Die Übergangswahrscheinlichkeiten v_k in Bild A.2 geben die Wahrscheinlichkeit an, daß während der Bedienzeit einer Anforderung k neue Anforderungen eintreffen. Unter der Bedingung, daß die Bedienzeit genau ν Zeitschlitze beträgt, ergibt sich die Verteilung der Anzahl der neuen Ankünfte aus der $(\nu - 1)$-fachen Faltung von ξ_k mit sich selbst. Die Verteilung dieser mit V bezeichneten Zufallsvariable läßt sich also wie folgt berechnen:

$$v_k = \mathrm{P}\{V = k\} = \sum_{\nu=1}^{\infty} [\xi^{(\nu)}]_k \cdot h_\nu. \tag{A.2}$$

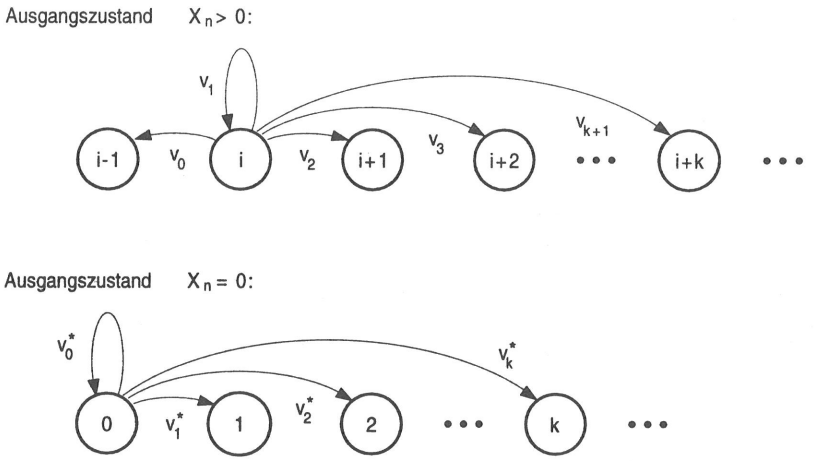


Bild A.2: Schematische Darstellung der möglichen Übergänge zwischen X_n und X_{n+1}

Dabei bezeichnet $[\xi^{(\nu)}]_k$ das k-te Element der $(\nu-1)$-fachen Faltung der Verteilung von Ξ mit sich selbst. Die Erzeugende Funktion $v(z)$ ist dann

$$\begin{aligned} v(z) &= \sum_{k=0}^{\infty}(\sum_{\nu=1}^{\infty}[\xi^{(\nu)}]_k h_\nu)\cdot z^k \\ &= \sum_{\nu=1}^{\infty}[\xi(z)]^\nu h_\nu \\ &= h(\xi(z)) \end{aligned} \tag{A.3}$$

mit Mittelwert $\mathrm{E}[V] = h'(1)\xi'(1) = \lambda \mathrm{E}[H] = A$.

Die Zufallsvariable V^* beschreibt die Anzahl der Anforderungen im System unmittelbar nach dem ersten Bedienende einer Betriebsperiode (*Busy Period*). Die Verteilung $v_k^* = \mathrm{P}\{V^* = k\}$ und deren Erzeugende Funktion $v^*(z) = \sum_{k=0}^{\infty} v_k^*\, z^k$ sind von der Bearbeitungsreihenfolge abhängig und werden deshalb in den Abschnitten A.2.1.1 und A.2.1.2 behandelt.

Zur Vereinfachung der Schreibweise können die Zustandswahrscheinlichkeiten x_k auch zu dem Zeilenvektor $\mathbf{x} = (x_0, x_1, x_2 \ldots)$ zusammengefaßt werden. Die Matrix $\mathbf{P} = [p_{ij}]$ soll die Übergangswahrscheinlichkeiten von Zustand i in den Zustand j enthalten, so daß $p_{ij} = \mathrm{P}\{X_{n+1} = j \mid X_n = i\}$ ist. Die Struktur der Matrix $\mathbf{P}$ ist somit

$$\mathbf{P} = \begin{pmatrix} v_0^* & v_1^* & v_2^* & \ldots \\ v_0 & v_1 & v_2 & \ldots \\ 0 & v_0 & v_1 & \ldots \\ 0 & 0 & v_0 & \ldots \\ \vdots & \vdots & \ddots & \ddots \end{pmatrix}.$$

Die stationären Zustandswahrscheinlichkeiten erhält man als Lösung des Gleichungssystems

$$\mathbf{x} \cdot \mathbf{P} = \mathbf{x}, \quad \mathbf{x} \cdot \mathbf{e} = 1 \qquad \text{mit } \mathbf{e} = (1, 1, \ldots)^T.$$

Aus diesem Gleichungssystem kann die allgemeine Form der Übergangsgleichungen abgeleitet werden:

$$x_i = v_i^* x_0 + \sum_{j=1}^{i+1} v_{i-j+1} x_j. \tag{A.4}$$

Die Erzeugende Funktion von x_k ergibt sich daraus zu

$$\begin{aligned} x(z) &= \sum_{i=0}^{\infty} x_0 v_i^* z^i + \sum_{i=0}^{\infty} \sum_{j=1}^{i+1} v_{i-j+1} x_j z^i \\ &= x_0 v^*(z) + z^{-1} v(z)[x(z) - x_0] , \end{aligned}$$

und nach $x(z)$ aufgelöst erhält man

$$x(z) = x_0 \frac{z v^*(z) - v(z)}{z - v(z)}. \tag{A.5}$$

Da x_k eine Verteilung ist, muß $x(z)|_{z=1} = 1$ gelten. Aus Gleichung A.5 folgt durch Grenzwertbildung und nach einmaliger Anwendung des Satzes von L'Hospital

$$x_0 = \frac{1 - A}{1 - A + \mathrm{E}[V^*]}.$$

Die allgemeine Form der Erzeugenden Funktion $x(z)$ ist somit

$$x(z) = \frac{1 - A}{1 - A + \mathrm{E}[V^*]} \frac{z v^*(z) - v(z)}{z - v(z)}. \tag{A.6}$$

Es soll nun $x(z)$ für die Varianten AF und DF bestimmt werden. Aus Sicht der Analyse unterscheiden sich die beiden Varianten nur durch ihr Verhalten während der ersten Bedienzeit einer Betriebsperiode. Die Anzahl V^* der Anforderungen im System unmittelbar nach dem ersten Bedienende einer Betriebsperiode ist die Summe aus zwei voneinander unabhängigen Zufallsvariablen:

- Eine Betriebsperiode wird von einer Gruppenankunft ausgelöst, die mindestens eine Anforderung enthält. Außer der ersten Anforderung, die sofort bedient wird, können gleichzeitig weitere Anforderungen eintreffen, die nicht sofort bedient werden können. Ξ^* bezeichnet die Anzahl derjenigen Anforderungen, die nicht sofort bedient werden.

  Es sei außerdem γ die Gruppengröße der ersten Ankunft zu Beginn einer Betriebsperiode. Da diese Gruppe auf jeden Fall mindestens eine Anforderung enthält, ist

$$\gamma_k = \mathrm{P}\{\gamma = k\} = \begin{cases} 0, & k = 0, \\ \dfrac{1}{1 - \xi_0} \xi_k, & k \geq 1, \end{cases} \tag{A.7}$$

mit der Erzeugenden Funktion

$$\gamma(z) = \sum_{k=1}^{\infty} \gamma_k z^k = \frac{\xi(z) - \xi_0}{1 - \xi_0}. \tag{A.8}$$

Falls die erste Gruppenankunft insgesamt $\gamma = k$ Anforderungen enthält, müssen davon $\Xi^* = k - 1$ Anforderungen warten. Die Verteilung von Ξ^* ist also

$$\xi_k^* = \mathrm{P}\{\Xi^* = k\} = \gamma_{k+1} = \frac{\xi_{k+1}}{1 - \xi_0} \tag{A.9}$$

mit der Erzeugenden Funktion

$$\xi^*(z) = \sum_{k=0}^{\infty} \xi_k^* \, z^k = z^{-1} \, \gamma(z) = z^{-1} \, \frac{\xi(z) - \xi_0}{1 - \xi_0} \tag{A.10}$$

und dem Mittelwert

$$\mathrm{E}[\Xi^*] = \frac{\mathrm{E}[\Xi]}{1 - \xi_0} - 1.$$

- Während der Bedienung der ersten Anforderung können weitere Gruppenankünfte auftreten. Die Anzahl der möglichen Gruppenankünfte hängt von der Bearbeitungsreihenfolge ab und muß für beide Varianten getrennt bestimmt werden.

Zur besseren Unterscheidung der Varianten werden im folgenden die Indices AF und DF verwendet.

A.2.1.1 Bearbeitungsreihenfolge AF

Es soll zunächst angenommen werden, daß die Dauer der ersten Bedienzeit einer Betriebsperiode ν Zeiteinheiten beträgt. Wie aus Bild A.1 hervorgeht, sind während dieser Bedienzeit insgesamt ν neue Gruppenankünfte möglich. Die Anzahl der Anforderungen, die während dieser Bedienzeit eintreffen, wird also durch die Zufallsvariable V beschrieben. Daraus folgt für die Erzeugende Funktion $v_{AF}^*(z)$:

$$\begin{aligned} v_{AF}^*(z) &= v(z) \, \xi^*(z) \\ &= v(z) \, z^{-1} \, \frac{\xi(z) - \xi_0}{1 - \xi_0}. \end{aligned} \tag{A.11}$$

Die numerische Berechnung von v_k^* für die Variante AF erfordert demnach eine Faltung von v_k mit ξ_k^*.

Aus Gleichung A.11 erhält man für den Mittelwert von V_{AF}^*

$$\mathrm{E}[V_{AF}^*] = \frac{\mathrm{E}[\Xi]}{1 - \xi_0} - (1 - A),$$

so daß sich die Erzeugende Funktion $x_{AF}(z)$ wie folgt angeben läßt:

$$x_{AF}(z) = \frac{1 - A}{\lambda} \, v(z) \, \frac{\xi(z) - 1}{z - v(z)}. \tag{A.12}$$

A.2.1.2 Bearbeitungsreihenfolge DF

Bei der Bearbeitungsreihenfolge DF wird, falls Bedienende und Ankunftsereignis gleichzeitig auftreten, das Bedienende vor dem Ankunftsereignis bearbeitet. Dies gilt auch für das erste Bedienende einer Betriebsperiode, so daß eventuell gleichzeitig mit diesem Bedienende eintreffende Anforderungen noch nicht gespeichert sind. Falls die Dauer dieser ersten Bedienzeit ν Zeiteinheiten beträgt, sind deshalb nur $\nu - 1$ weitere Gruppenankünfte möglich (vgl. Bild A.1).

Es gilt daher

$$\begin{aligned} v^*_{DF}(z) &= \sum_{\nu=1}^{\infty} \xi^*(z)\,[\xi(z)]^{\nu-1} h_\nu \\ &= \xi^*(z) \frac{v(z)}{\xi(z)}, \end{aligned} \tag{A.13}$$

und

$$\mathrm{E}[V^*_{DF}] = \frac{\xi_0\ \mathrm{E}[\Xi]}{1-\xi_0} - (1-A).$$

Man erhält also für die Bearbeitungsreihenfolge DF

$$x_{DF}(z) = \frac{1-A}{\lambda} \frac{v(z)}{\xi(z)} \frac{\xi(z)-1}{z-v(z)}. \tag{A.14}$$

Aus den Gleichungen A.12 und A.14 läßt sich folgende einfache Beziehung zwischen $x_{AF}(z)$ und $x_{DF}(z)$ ableiten:

$$x_{DF}(z) = \frac{x_{AF}(z)}{\xi(z)} \tag{A.15}$$

Für den Sonderfall des Geo/GI/1-Systems wurde in [89] eine zu Gleichung A.15 analoge Beziehung hergeleitet.

A.2.2 Zeitgemittelte Zustandswahrscheinlichkeiten

Die zeitgemittelten Zustandswahrscheinlichkeiten geben den relativen Zeitanteil an, mit dem sich das System in einem bestimmten Zustand befindet. Dies entspricht den Zustandswahrscheinlichkeiten, die von einem außenstehenden Beobachter zu einem zufälligen Zeitpunkt gesehen werden. Es sei $y_k = \mathrm{P}\{Y = k\}$ die Wahrscheinlichkeit, daß zu einem zufälligen Zeitpunkt $Y = k$ Anforderungen im System sind.

Zur Berechnung der zeitgemittelten Zustandswahrscheinlichkeiten wird das sogenannte *Key Renewal Theorem* [54, 55] benötigt, welches eine Beziehung zwischen den Zustandswahrscheinlichkeiten zu den Regenerationszeitpunkten und den zeitgemittelten Zustandswahrscheinlichkeiten herstellt. Es sei

S_{t_n}	der Systemzustand zum letzten Regenerationszeitpunkt (Bedienende) vor dem Beobachtungszeitpunkt,
S_{t_n+u}	der Systemzustand zum Beobachtungszeitpunkt $t_n + u$,
$T_{n+1} = t_{n+1} - t_n$	der Abstand zwischen dem n-ten und dem $n+1$-ten Bedienende,
$\mathrm{E}[D]$	der mittlere Abstand zwischen zwei Regenerationszeitpunkten.

Nach dem *Key Renewal Theorem* berechnen sich die zeitgemittelten Zustandswahrscheinlichkeiten wie folgt:

$$y_j = \sum_{i=0}^{j} \frac{x_i}{\mathrm{E}[D]} \sum_{u=1}^{\infty} \mathrm{P}\{S_{t_n+u} = j, T_{n+1} > u \mid S_{t_n} = i\} \tag{A.16}$$

Die Bestimmung von $\mathrm{P}\{S_{t_n+u} = j, T_{n+1} > u \mid S_{t_n} = i\}$, also der Wahrscheinlichkeit, daß sich das System zu einem beliebigen Zeitpunkt $t_n + u$ zwischen den Regenerationszeitpunkten t_n und t_{n+1} im Zustand j befindet, unter der Bedingung daß zum letzten Regenerationszeitpunkt i Anforderungen im System waren, stellt die Hauptschwierigkeit bei der Berechnung der zeitgemittelten Zustandswahrscheinlichkeiten dar.

Im folgenden soll die Berechnung für die Bearbeitungsreihenfolge AF skizziert werden. Der Lösungsweg für das DF-System ist völlig analog und führt zum selben Ergebnis, da die zeitgemittelten Zustandswahrscheinlichkeiten im unendlichen System nicht von der Bearbeitungsreihenfolge abhängig sind.

Entsprechend der in [30] beschriebenen Vorgehensweise können drei verschiedene Fälle unterschieden werden:

1. Der Beobachtungszeitpunkt fällt in eine Freiperiode

Dies ist nur möglich, wenn das System zum letzten Regenerationszeitpunkt leer war, d.h. $S_{t_n} = 0$ gilt. Um zum Zeitpunkt $t_n + u$ ein leeres System vorzufinden, dürfen während der vergangenen $u - 1$ Zeiteinheiten keine Anforderungen ankommen. Daraus folgt

$$\mathrm{P}\{S_{t_n+u} = 0 \mid S_{t_n} = 0\} = \xi_0^{u-1},$$

und

$$y_0 = \frac{x_0}{\mathrm{E}[D]} \sum_{u=1}^{\infty} \mathrm{P}\{S_{t_n+u} = 0 \mid S_{t_n} = 0\} = \frac{x_0}{\mathrm{E}[D]} \frac{1}{1 - \xi_0}. \tag{A.17}$$

2. Der Beobachtungszeitpunkt fällt in die erste Bedienzeit nach einer Freiperiode

Auch hier war das System zum letzten Bedienende leer, d.h. $S_{t_n} = 0$. Ein Beobachter findet außer der Gruppenankunft, die die Freiperiode beendet hat, auch die vom Bedienungsbeginn bis zum Beobachtungszeitpunkt eingetroffenen Anforderungen im System vor. Unter der Bedingung, daß die Dauer der vorhergehenden Freiperiode l Zeiteinheiten beträgt, verbleiben noch $u - l - 1$ Zeitschlitze vom Bedienungsbeginn der ersten Anforderung bis zum Beobachtungszeitpunkt $t_n + u$.

Es sei $F^*(j,k)$ die Wahrscheinlichkeit, daß am Ende der Freiperiode und in den folgenden k Zeitschlitzen insgesamt j Anforderungen ankommen. $F^*(j,k)$ berechnet sich zu $F^*(j,k) = \left[\gamma_i * \xi_i^{(k)}\right]_j$, wobei das Symbol $*$ die Faltungsoperation bezeichnet, $[\cdot]_j$ das j-te Element der innerhalb der Klammern stehenden Verteilung ist und γ_i durch Gleichung A.7 gegeben ist.

Damit läßt sich $\mathrm{P}\{S_{t_n+u} = j, T_{n+1} > u \mid S_{t_n} = 0\}$ wie folgt angeben:

$$\mathrm{P}\{S_{t_n+u} = j, T_{n+1} > u \mid S_{t_n} = 0\} = \sum_{l=1}^{u-1} \xi_0^{l-1}(1-\xi_0)\, F^*(j, u-l-1) \sum_{\nu=u-l}^{\infty} h_\nu \qquad \text{(A.18)}$$

3. Der Beobachtungszeitpunkt fällt in eine spätere Bedienzeit

In diesem Fall waren nach dem letzten Bedienende noch $i > 0$ Anforderungen im System. Damit zum Beobachtungszeitpunkt $t_n + u$ insgesamt j Anforderungen enthalten sind, müssen in den verbleibenden $u-1$ Zeiteinheiten insgesamt $j-i$ Anforderungen eingetroffen sein. Unter Verwendung der Wahrscheinlichkeit $F(j,k) = [\xi^{(k)}]_j$, daß in k Zeitschlitzen j Anforderungen eintreffen, ergibt sich

$$\mathrm{P}\{S_{t_n+u} = j, T_{n+1} > u \mid S_{t_n} = i\} = F(j-i, u-1) \sum_{\nu=u}^{\infty} h_\nu. \qquad \text{(A.19)}$$

Damit sind alle zur Bestimmung der zeitgemittelten Zustandswahrscheinlichkeiten benötigten Größen bekannt. Unter Verwendung der Beziehungen A.18 und A.19 erhält man für y_j:

$$\begin{aligned} y_j \;=\; & \frac{1}{\mathrm{E}[D]} \left[x_0 \sum_{u=2}^{\infty} \sum_{l=1}^{u-1} \xi_0^{l-1}(1-\xi_0)\, F^*(j, u-l-1) \sum_{\nu=u-l}^{\infty} h_\nu \right. \\ & \left. + \sum_{i=1}^{j} x_i \sum_{u=1}^{\infty} F(j-i, u-1) \sum_{\nu=u}^{\infty} h_\nu \right], \quad j > 0. \end{aligned} \qquad \text{(A.20)}$$

Das weitere Vorgehen ist völlig analog zu der Herleitung in [30]. Zunächst wird die Erzeugende Funktion $y(z) = \sum_{j=0}^{\infty} y_j z^j$ gebildet. Eine etwas langwierige Umformung liefert dann

$$y(z) = y_0 + \frac{1}{\mathrm{E}[D]} \left[x_0\, \gamma(z) + x(z) - x_0\right] \frac{1 - v(z)}{1 - \xi(z)}.$$

Mit Hilfe der Gleichungen A.8 und A.17 folgt daraus

$$y(z)\mathrm{E}[D]\,(1-\xi(z)) = x(z)\,(1-v(z)) - x_0\,(z\, v^*(z) - v(z))\,.$$

Nach Gleichung A.5 ist jedoch

$$x_0\,(z\, v^*(z) - v(z)) = x(z)\,(z - v(z))\,,$$

und somit

$$y(z) = \frac{1-z}{\mathrm{E}[D]\;(1-\xi(z))}\; x(z). \qquad \text{(A.21)}$$

Da in einem reinen Wartesystem keine Anforderungen verloren gehen, muß die Abgangsrate gleich der Ankunftsrate sein, d.h. $\mathrm{E}[D] = 1/\lambda$. Nach Einsetzen von Gleichung A.12 ergibt sich schließlich

$$y(z) = (1 - A)\ (z - 1)\ \frac{v(z)}{z - v(z)}. \tag{A.22}$$

Diese Beziehung ist, allerdings ohne Beweis, auch in [122] angegeben.

Es kann leicht nachgeprüft werden, daß $y_0 = y(z)|_{z=0} = 1 - A$ erfüllt ist. Die mittlere Anzahl der Anforderungen im System ist

$$\mathrm{E}[Y] = \frac{d}{dz} y(z)|_{z=1} = \frac{h''(1)\lambda^2 + \mathrm{E}[H]\xi''(1)}{2(1 - A)} + A \tag{A.23}$$

A.2.3 Zustandswahrscheinlichkeiten vor der Ankunft einer Anforderung

Für Systeme, deren Zustand sich um jeweils eine Anforderung erhöht oder erniedrigt, wurde in [58] bewiesen, daß die Zustandswahrscheinlichkeiten bei Eintritt einer Anforderung identisch sind mit den Zustandswahrscheinlichkeiten unmittelbar nach einem Bedienende. Bezogen auf eine Gruppenankunft als Ganzes ist diese Bedingung im $\mathrm{D}^{[X]}$/GI/1-System zwar nicht erfüllt, wohl aber bezogen auf jede einzelne Anforderung innerhalb einer Gruppe.

Konzeptionell kann die Gruppenankunft auch so interpretiert werden, daß die Anforderungen zwar gleichzeitig ankommen, aber nacheinander in die Warteschlange eingereiht werden. Falls eine Anforderung unmittelbar vor ihrem Eintritt in die Warteschlange bereits k Anforderungen im System vorfindet, löst die Anforderung einen Übergang vom Zustand k in den Zustand $k + 1$ aus. Da jedes Bedienende den Systemzustand um jeweils eine Einheit erniedrigt, muß innerhalb dieser Betriebsperiode zwangsläufig auch ein Übergang vom Zustand $k + 1$ in den Zustand k erfolgen. Die Kopplung dieser beiden Übergänge hat zur Folge, daß ihre relative Häufigkeit innerhalb einer Betriebsperiode identisch ist. Daher sind die Zustandswahrscheinlichkeiten bei Eintritt einer einzelnen Anforderung und unmittelbar nach einem Bedienende ebenfalls identisch.

Diese Identität kann auch rein formal gezeigt werden. Es sei Z der Systemzustand aus der Sicht einer eintreffenden Anforderung, mit der Erzeugenden Funktion $z_{AF}(z)$ bzw. $z_{DF}(z)$. Der Systemzustand, den eine eintreffende Anforderung vorfindet, setzt sich aus zwei Anteilen zusammen:

- der Anzahl Z^* der Anforderungen im System vor Ankunft der Gruppe, zu der die betrachtete Anforderung gehört, und
- der Anzahl Ψ derjenigen Anforderungen, die gleichzeitig eintreffen, aber bereits gespeichert sind.

Z^* ist der Systemzustand vor dem Eintreffen einer beliebigen Gruppe, die auch die Größe 0 haben kann. Da die Gruppengröße ξ jedoch vom Zustand des Systems unabhängig ist, beschreibt Z^* auch den Zustand des Systems vor dem Eintreffen einer Gruppe, die mindestens eine Anforderung enthält.

Zunächst soll die Verteilung ψ_k der Zufallsvariable Ψ berechnet werden, da diese von der Bearbeitungsreihenfolge unabhängig ist. Dazu wird die Wahrscheinlichkeit ψ_i^* benötigt, daß sich die betrachtete Anforderung in einer Gruppe der Größe i befindet. Nach [70, 125] gilt

$$\psi_i^* = \frac{i \cdot \xi_i}{\sum_{\nu=1}^{\infty} \nu \cdot \xi_\nu} = \frac{1}{\lambda}\, i \cdot \xi_i.$$

Es wird nun angenommen, daß die Anforderungen einer Gruppe in zufälliger Reihenfolge gespeichert werden. Unter der Bedingung, daß die Gruppengröße i Anforderungen beträgt, ist die Wahrscheinlichkeit, mit der $k < i$ Anforderungen bereits gespeichert sind, gleich $1/i$.

$$\mathrm{P}\{\Psi = k \mid \Psi^* = i\} = \frac{1}{i}, \quad 0 \le k \le i-1.$$

Nach dem Gesetz der totalen Wahrscheinlichkeit folgt daraus

$$\begin{aligned} \psi_k &= \sum_{i=k+1}^{\infty} \mathrm{P}\{\Psi = k \mid \Psi^* = i\}\ \mathrm{P}\{\Psi^* = i\} \\ &= \frac{1}{\lambda} \sum_{i=k+1}^{\infty} \xi_i \\ &= \frac{1}{\lambda}\,(1 - \mathrm{P}\{\Xi \le k\}). \end{aligned} \tag{A.24}$$

Die Erzeugende Funktion $\psi(z) = \sum_{k=0}^{\infty} \psi_k$ ist

$$\begin{aligned} \psi(z) &= \frac{1}{\lambda}\left(\frac{1}{1-z} - \frac{\xi(z)}{1-z}\right) \\ &= \frac{1-\xi(z)}{\lambda(1-z)}. \end{aligned} \tag{A.25}$$

Die Verteilung $z_k^* = \mathrm{P}\{Z^* = k\}$ kann mit Hilfe der zeitgemittelten Zustandswahrscheinlichkeiten berechnet werden. Die Beziehung zwischen Y und Z^* ist jedoch von der Bearbeitungsreihenfolge abhängig.

A.2.3.1 Bearbeitungsreihenfolge AF

Eine eintreffende Gruppe findet hier den zeitlichen Mittelwert der Systemzustandsverteilung vor, da ein eventuell gleichzeitig auftretendes Bedienende erst nach der Ankunft bearbeitet wird. Es gilt deshalb $z_{AF}^*(z) = y(z)$, und mit Gleichung A.25 ergibt sich

$$\begin{aligned} z_{AF}(z) &= y(z)\ \psi(z) \\ &= \frac{1-A}{\lambda}\, v(z)\, \frac{\xi(z)-1}{z-v(z)}. \end{aligned} \tag{A.26}$$

Mit Gleichung A.12 folgt daraus $z_{AF}(z) = x_{AF}(z)$.

A.2.3.2 Bearbeitungsreihenfolge DF

Hier ist der zeitlich gemittelte Systemzustand der Zustand nach dem Eintreffen einer Gruppe beliebiger Größe. Es gilt daher $Y = Z^*_{DF} + \Xi$, und somit ist $z^*_{DF}(z) = y(z)/\xi(z)$. Für $z_{DF}(z)$ erhält man also

$$\begin{aligned} z_{DF}(z) &= \frac{y(z)}{\xi(z)}\psi(z) \\ &= \frac{1-A}{\lambda}\,\frac{v(z)}{\xi(z)}\,\frac{\xi(z)-1}{z-v(z)}. \end{aligned} \tag{A.27}$$

Ein Vergleich mit Gleichung A.14 zeigt, daß die Beziehung $z_{DF}(z) = x_{DF}(z)$ auch hier gilt.

A.2.4 Wartezeitverteilung

Ein häufig verwendeter Ansatz zur Berechnung der Wartezeitverteilung basiert auf dem Konzept der Restarbeit im System. Diese Restarbeit ist die Summe aus der Restbedienzeit der gerade bedienten Anforderung und der gesamten Bedienzeit aller anderen Anforderungen im System. Die Restarbeit ist also äquivalent zu der Zeit, die benötigt wird, um alle momentan im System befindlichen Anforderungen vollständig zu bearbeiten.

Für Systeme mit FIFO-Abfertigungsstrategie ist daher die Wartezeit einer Anforderung gleich der Restarbeit im System zum Zeitpunkt ihres Eintreffens. Es sei W^*_n die Restarbeit *vor* dem Eintreffen von Anforderungen zum Zeitpunkt $n \cdot \Delta T$. Bis zum Zeitpunkt $(n+1) \cdot \Delta T$ wird die Restarbeit, die einer Zeiteinheit entspricht, erledigt. Gleichzeitig kommt jedoch die mit C_n bezeichnete gesamte Restarbeit aller neu eingetroffenen Anforderungen hinzu. Analog zu Gleichung A.1 gilt somit:

$$W^*_{n+1} = \max(0, W^*_n + C_n - 1). \tag{A.28}$$

Die weitere Vorgehensweise zur Bestimmung der stationären Verteilung von W^* orientiert sich an der Berechnung der Zustandswahrscheinlichkeiten nach einem Bedienende. Die Verteilung $c_k = \mathrm{P}\{C = k\}$ der gesamten Restarbeit, die durch eine Gruppenankunft hinzu kommt, berechnet sich zu [125]

$$c_k = \sum_{\nu=0}^{\infty} [h^{(\nu)}]_k \cdot \xi_\nu \quad , \tag{A.29}$$

mit der Erzeugenden Funktion

$$\begin{aligned} c(z) &= \sum_{k=0}^{\infty} c_k\, z^k = \sum_{k=0}^{\infty} \Big(\sum_{\nu=0}^{\infty} [h^{(\nu)}]_k \xi_\nu\Big) \cdot z^k \\ &= \xi\,(h(z))\,. \end{aligned} \tag{A.30}$$

$c(z)$ ist also in gewisser Weise eine zu $v(z)$ komplementäre Funktion. Dies verdeutlicht auch Bild A.3, in dem die möglichen Übergänge von W^*_n nach W^*_{n+1} dargestellt sind: Ein Vergleich

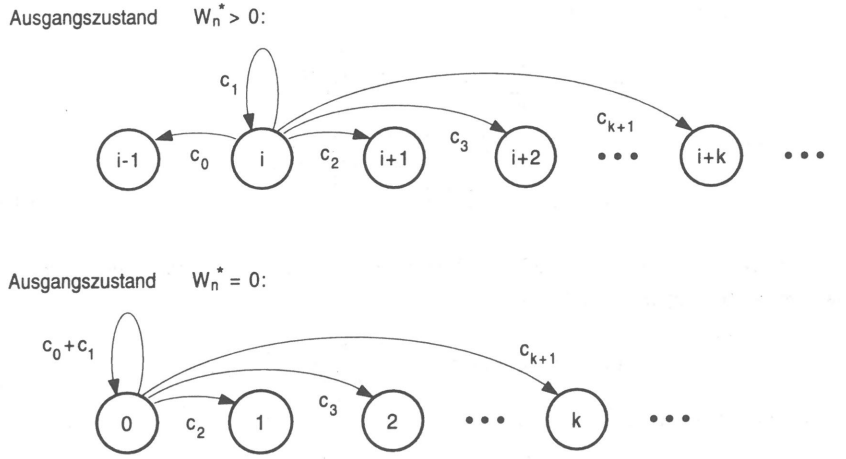


Bild A.3: Schematische Darstellung der möglichen Übergänge zwischen W_n^* und W_{n+1}^*

von Abbildung A.2 mit Bild A.3 zeigt, daß beide Übergangsdiagramme äquivalent sind, falls v_k durch c_k ersetzt wird und

$$v_k^* = \begin{cases} c_0 + c_1, & k = 0, \\ c_{k+1}, & k > 0 \end{cases} \quad \text{gilt.}$$

Die Erzeugende Funktion $w^*(z)$ kann somit sehr einfach aus Gleichung A.6 durch entsprechendes Ersetzen von $v(z)$ und $v^*(z)$ bestimmt werden. Mit Hilfe von

$$\begin{aligned} v(z) &= c(z) \qquad \text{und} \\ v^*(z) &= c_0 + c_1 + \sum_{k=1}^{\infty} c_{k+1}\, z^k \\ &= z^{-1}\ (c(z) + c_0(z-1)) \end{aligned}$$

erhält man aus A.6

$$w^*(z) = (1 - A)\ \frac{z-1}{z - c(z)}\ . \tag{A.31}$$

Zur Bestimmung der Wartezeit W einer beliebigen Anforderung muß noch die gesamte Bedienzeit ω aller Anforderungen berücksichtigt werden, die in der gleichen Gruppe eintreffen und vor der betrachteten Anforderung gespeichert werden:

$$W = W^* + \omega$$

Die Verteilung $\omega_k = \mathrm{P}\{\omega = k\}$ kann unter Verwendung von Gleichung A.25 sofort angegeben werden:

$$\omega_k = \sum_{\nu=1}^{\infty} [h^{(\nu)}]_k \psi_k\ , \tag{A.32}$$

mit der Erzeugenden Funktion

$$\begin{aligned}\omega(z) &= \sum_{k=0}^{\infty} \omega_k\, z^k = \sum_{k=0}^{\infty}\sum_{\nu=1}^{\infty} [h^{(\nu)}]_k \psi_k\, z^k \\ &= \psi\,(h(z)) = \frac{1-\xi(h(z))}{\mathrm{E}[\Xi]\,(1-h(z))}, \\ \omega(z) &= \frac{1}{\lambda}\,\frac{1-c(z)}{1-h(z)}. \end{aligned} \tag{A.33}$$

Da ω und W^* statistisch voneinander unabhängig sind, ist die Erzeugende Funktion $w(z)$ der Wartezeitverteilung $w_k = \mathrm{P}\{W = k\}$

$$w(z) = w^*(z)\,\omega(z) = \frac{1-A}{\lambda}\,\frac{z-1}{1-h(z)}\,\frac{1-c(z)}{z-c(z)}. \tag{A.34}$$

Die mittlere Wartezeit berechnet sich zu

$$\mathrm{E}[W] = \frac{d}{dz} w(z)|_{z=1} = \frac{h''(1)\lambda^2 + \mathrm{E}[H]\xi''(1)}{2\lambda(1-A)}. \tag{A.35}$$

Nach dem Gesetz von Little muß $\lambda(\mathrm{E}[W] + \mathrm{E}[H]) = \mathrm{E}[Y]$ gelten. Dies kann durch Einsetzen der Beziehungen A.23 und A.35 leicht nachgeprüft werden.

A.2.5 Verteilung der Betriebsperiode

Die Betriebsperiode (*Busy Period*) ist ein Zeitintervall, während dessen die Bedieneinheit des Systems ununterbrochen belegt ist. In einem AF-System ist dies die Zeit von der Ankunft der ersten Anforderung, die auf ein leeres System trifft, bis zum nächsten Bedienende, welches das System wieder leer zurückläßt. In einem DF-System kann auch während einer Betriebsperiode der Fall auftreten, daß nach einem Bedienende keine Anforderungen mehr vorhanden sind. Es müssen dann jedoch gleichzeitig mit diesem Bedienende weitere Anforderungen ankommen, so daß die Betriebsperiode fortgesetzt wird.

Die Bedieneinheit wechselt also zwischen Betriebsperioden, in denen die Bedieneinheit ständig belegt ist, und Freiperioden, in denen das System (und damit auch die Bedieneinheit) leer ist. Die Folge der Betriebsperioden und Freiperioden bildet einen rekurrenten Prozeß, da die einzelnen Phasen voneinander unabhängig sind.

Die Verteilung der Dauer I einer Freiperiode kann wegen der Gedächtnisfreiheit des Ankunftsprozesses sofort angegeben werden:

$$\mathrm{P}\{I = k\} = \xi_0^{k-1}(1-\xi_0), \qquad k \geq 1, \tag{A.36}$$

$$\mathrm{E}[I] = \frac{1}{1-\xi_0}. \tag{A.37}$$

Die Analyse der Dauer einer Betriebsperiode ist dagegen etwas aufwendiger. Für den allgemeinen Fall mit Gruppenankünften und beliebiger Verteilung der Bedienzeit kann nur eine implizite Beziehung für die Erzeugende Funktion der Betriebsperiode hergeleitet werden. Auf diese Weise lassen sich zwar deren Momente bestimmen, für die Berechnung der gesamten Verteilung muß jedoch auf ein numerisches Verfahren zurückgegriffen werden.

Es seien

- B die Dauer einer Betriebsperiode,
- b_k die Verteilung von B, d.h. $b_k = \mathrm{P}\{B = k\}$,
- $b(z)$ die Erzeugende Funktion von b_k,
- γ die Gruppengröße der ersten Ankunft der betrachteten Betriebsperiode,
- C die gesamte Bedienzeit aller Anforderungen der ersten Gruppenankunft (abhängig von γ), und
- N die Anzahl der Gruppen mit mindestens einer Anforderung, die während der Bedienung aller Anforderungen der ersten Gruppenankunft eintreffen (abhängig von C).

Unter der Bedingung, daß γ, C und N bekannt sind, läßt sich die „bedingte" Erzeugende Funktion der Verteilung einer Betriebsperiode leicht angeben. Der Gedankengang ist völlig analog zu der in [121] beschriebenen Vorgehensweise für das M/GI/1-System.

Die Betriebsperiode wird durch eine Gruppenankunft ausgelöst, die γ Anforderungen enthält. Während der Bedienung dieser Anforderungen können weitere Gruppen eintreffen, wobei jede Gruppe, die mindestens eine Anforderung enthält, eine weitere Betriebsperiode auslöst. Die Betriebsperiode besteht also aus den Bedienzeiten aller Anforderungen der ersten Gruppenankunft und den dabei initiierten weiteren Betriebsperioden, die sich nahtlos anschließen. Die bedingte Erzeugende Funktion ist daher gegeben durch [121]

$$b(z \mid N = i, C = k, \gamma = l) = z^k \, b(z)^i .$$

Im folgenden werden die einzelnen, voneinander abhängigen Bedingungen eliminiert. Die Wahrscheinlichkeit, daß in k Zeitschlitzen i Gruppen mit mindestens einer Anforderung eintreffen, ist $\binom{k}{i}(1-\xi_0)^i \, \xi_0^{k-i}$. Damit kann die Bedingung $N = i$ entfernt werden:

$$\begin{aligned} b(z \mid C = k, \gamma = l) &= z^k \sum_{i=0}^{k} b(z)^i \binom{k}{i} (1-\xi_0)^i \, \xi_0^{k-i} \\ &= z^k \left[\xi_0 + (1-\xi_0) b(z)\right]^k . \end{aligned}$$

Die Verteilung der gesamten Bedienzeit aller in der ersten Gruppenankunft eintreffenden Anforderungen ist die $(l-1)$-fache Faltung von h_k mit sich selbst, so daß

$$\begin{aligned} b(z \mid \gamma = l) &= \sum_{k=1}^{\infty} z^k \left[\xi_0 + (1-\xi_0) b(z)\right]^k \left[h^{(l)}\right]_k \\ &= h\left(z \left[\xi_0 + (1-\xi_0) b(z)\right]\right)^l . \end{aligned}$$

Die Wahrscheinlichkeit, daß die erste Gruppenankunft γ Anforderungen enthält, ist γ_k (siehe Gleichung A.7). Damit ergibt sich schließlich die Erzeugende Funktion $b(z)$ zu

$$\begin{aligned} b(z) &= \sum_{l=1}^{\infty} h\left(z\left[\xi_0 + (1-\xi_0)b(z)\right]\right)^l \gamma_l \\ &= \gamma\left(h\left(z\left[\xi_0 + (1-\xi_0)b(z)\right]\right)\right). \end{aligned} \tag{A.38}$$

Die Momente von B können nun aus Gleichung A.38 abgeleitet werden:

$$\mathrm{E}[B] = b'(1) = \frac{1}{1-\xi_0}\,\frac{A}{1-A}, \tag{A.39}$$

$$\mathrm{E}[B^2] = b''(1) + b'(1) = \frac{1}{1-\xi_0}\left[\frac{v''(1)}{(1-A)^3} + \frac{A^2}{(1-A)^2} + \frac{A}{1-A}\right]. \tag{A.40}$$

Die Verteilung b_k der Dauer einer Betriebsperiode kann mit Hilfe eines numerischen Verfahrens berechnet werden, welches auf einen Ansatz von Cooper [58] zurückgeht. Es sei M die Anzahl der Anforderungen, die während einer Betriebsperiode bedient werden. Falls es gelingt, die Verteilung $\mathrm{P}\{M = k\}$ zu berechnen, läßt sich die Verteilung der Dauer einer Betriebsperiode daraus sehr einfach ableiten:

$$b_k = \mathrm{P}\{B = k\} = \sum_{i=1}^{\infty} \mathrm{P}\{M = i\}[h^{(i)}]_k. \tag{A.41}$$

Die Berechnung von $\mathrm{P}\{M = k\}$ beruht auf einem Ergebnis von Takács [199], welches sich auf die Anzahl der Anforderungen bezieht, die in einer Betriebsperiode bedient werden. Es sei $\mathrm{P}\{M = k \mid \gamma = j\}$ die Wahrscheinlichkeit, daß in einer Betriebsperiode, die mit einer Gruppenankunft von j Anforderungen beginnt, insgesamt k Anforderungen bedient werden. Nach Takács [199] gilt:

$$\begin{aligned} \mathrm{P}\{M = k \mid \gamma = j\} &= \frac{j}{k}\,\mathrm{P}\{(k-j) \text{ Ankünfte in der Bedienzeit von } k \text{ Anforderungen}\} \\ &= \frac{j}{k}\,[v^{(k)}]_{k-j}. \end{aligned}$$

Daraus folgt

$$\mathrm{P}\{M = k\} = \sum_{j=1}^{\infty} \frac{j}{k}\,[v^{(k)}]_{k-j}\,\gamma_j.$$

Die Verteilung der Dauer einer Betriebsperiode kann nun mit Hilfe von Gleichung A.41 berechnet werden. Dieser Ansatz wurde bereits in [29] für den Sonderfall des Geo/D/1-Systems beschrieben.

A.3 Analyse des Systems mit begrenztem Wartespeicher

Systeme mit unbegrenztem Wartespeicher können mit Hilfe von Erzeugenden Funktionen sehr elegant analysiert werden, wobei sich insbesondere die Momente der interessierenden Größen in geschlossener Form angeben lassen. Bei Systemen mit begrenztem Wartespeicher ist es im allgemeinen nicht mehr möglich, eine geschlossene Lösung zu bestimmen. Es muß meist auf numerische Verfahren zur Berechnung der Verteilungen und deren Momente zurückgegriffen werden.

A.3.1 Zustandswahrscheinlichkeiten nach einem Bedienende

Im folgenden wird das $D^{[X]}/GI/1$-System mit begrenzter Warteschlangenlänge S betrachtet (insgesamt können sich also bis zu $S+1$ Anforderungen im System befinden). Die Analyse basiert auch hier auf einer Markov-Kette, die in die Bedienendezeitpunkte eingebettet ist. Nach einem Bedienende kann das System allerdings höchstens S Anforderungen enthalten, so daß der Zustandsraum $S+1$ Zustände umfaßt. Die Übergangsmatrix $\mathbf{P}$ der eingebetteten Markov-Kette hat die gleiche Struktur wie im reinen Wartesystem:

$$\mathbf{P} = \begin{pmatrix} v_0^* & v_1^* & v_2^* & \dots & v_{S-1}^* & \sum_{i=S}^{\infty} v_i^* \\ v_0 & v_1 & v_2 & \dots & v_{S-1} & \sum_{i=S}^{\infty} v_i \\ 0 & v_0 & v_1 & \dots & v_{S-2} & \sum_{i=S-1}^{\infty} v_i \\ 0 & 0 & v_0 & \dots & v_{S-3} & \sum_{i=S-2}^{\infty} v_i \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & v_0 & \sum_{i=1}^{\infty} v_i \end{pmatrix}.$$

Da die Übergangsmatrix $\mathbf{P}$ hier endlich ist, kann das Gleichungssystem $\mathbf{x} \cdot \mathbf{P} = \mathbf{x}, \quad \mathbf{x} \cdot \mathbf{e} = 1$ numerisch gelöst werden, um den Zustandsvektor $\mathbf{x}$ zu bestimmen. Wegen der besonderen Struktur der Matrix $\mathbf{P}$ können die Zustandswahrscheinlichkeiten jedoch auch auf eine wesentlich einfachere und schnellere Weise berechnet werden. Ein Vergleich der Übergangsmatrizen der Systeme mit unendlicher und mit begrenzter Warteschlangenlänge zeigt, daß die Übergangsgleichungen für die Zustände 0 bis $S-1$ identisch sind. Es gilt also (siehe A.4)

$$x_i = v_i^* x_0 + \sum_{j=1}^{i+1} v_{i-j+1} x_j, \qquad 0 \le i < S.$$

Die letzte Spalte der Übergangsmatrix wird zur Berechnung der x_i nicht benötigt, da zusammen mit der Normierungsbedingung $\mathbf{x} \cdot \mathbf{e} = 1$ insgesamt $S+2$ Gleichungen für $S+1$ Unbekannte zur Verfügung stehen. Aus der Identität der Gleichungssysteme folgt außerdem,

daß sich die Zustandswahrscheinlichkeiten $x_i, 0 \leq i < S$, in den Systemen mit endlicher und unbegrenzter Warteschlangenlänge nur durch einen Normierungsfaktor unterscheiden.

Ein möglicher Ansatz zur numerischen Berechnung der Zustandswahrscheinlichkeiten x_i besteht darin, für x_0 zunächst einen beliebigen Wert zu wählen. Mit Hilfe von Gleichung A.4 erhält man die unnormierten Werte für $x_i, 1 \leq i < S$ aus der Rekursion

$$x_i = \frac{1}{v_0}\left[x_{i-1} - \sum_{j=1}^{i-1} v_{i-j}x_j - v_{i-1}^* x_0\right]. \tag{A.42}$$

Anschließend müssen die Zustandswahrscheinlichkeiten noch normiert werden, so daß $\mathbf{x}\cdot\mathbf{e} = 1$ ist. Der Algorithmus nach Gleichung A.42 ist jedoch numerisch nicht sehr stabil, da innerhalb der Rekursion Differenzen gebildet werden müssen. Bei großen Werten von S können daher wegen der endlichen Genauigkeit der Zahlendarstellung numerische Probleme auftreten.

In [123] ist ein numerisch stabiler Ansatz zur Berechnung der Zustandswahrscheinlichkeiten angegeben, der ohne Differenzbildung auskommt. Bei diesem Verfahren werden zunächst die Übergangsgleichungen der Zustände 0 bis k aufsummiert.

$$\begin{aligned}\sum_{i=0}^{k} x_i &= \sum_{i=0}^{k}\left[v_i^* x_0 + \sum_{j=1}^{i+1} v_{i-j+1}x_j\right] \\ &= x_0\sum_{i=0}^{k} v_i^* + v_0\sum_{i=1}^{k} x_i + v_0 x_{k+1} + \sum_{j=1}^{k} x_j \sum_{i=1}^{k-j+1} v_i.\end{aligned}$$

Nach x_{k+1} aufgelöst, erhält man

$$x_{k+1} = \frac{1}{v_0}\left[x_0\bar{v}_k^* + \sum_{j=1}^{k} x_j \bar{v}_{k+1-j}\right], \qquad 0 \leq k < S-1,$$

wobei die abkürzende Schreibweise $\bar{v}_k = \sum_{i=k+1}^{\infty} v_i$ für die komplementäre Verteilungsfunktion benutzt wurde. Schließlich kann $k+1$ noch durch k ersetzt werden, und es ergibt sich

$$x_k = \begin{cases} \dfrac{\bar{v}_0^*}{v_0} x_0, & k = 1, \\ \dfrac{1}{v_0}\left(x_0\bar{v}_{k-1}^* + \displaystyle\sum_{i=1}^{k-1} x_i \bar{v}_{k-i}\right), & 2 \leq k \leq S. \end{cases} \tag{A.43}$$

Anschließend müssen die Zustandswahrscheinlichkeiten x_k noch normiert werden, so daß $\mathbf{x}\mathbf{e} = 1$ gilt. Diese Rekursion ist numerisch wesentlich stabiler als der Algorithmus nach Gleichung A.42. Es wird allerdings vorausgesetzt, daß die komplementären Verteilungsfunktionen $\bar{v}_k$ und $\bar{v}_k^*$ hinreichend genau berechnet werden können. Der Algorithmus eignet sich natürlich auch für die Berechnung der Zustandswahrscheinlichkeiten x_i im reinen Wartesystem, wobei hier x_0 explizit gegeben ist, so daß die Normierung entfallen kann.

Eine einfache Umrechnung der Zustandswahrscheinlichkeiten nach einem Bedienende zwischen Systemen mit Bearbeitungsreihenfolge AF und DF ist im allgemeinen nicht möglich. Nur für den Sonderfall des Geo/GI/1-Systems ohne Gruppenankünfte kann eine einfache Beziehung für die Umrechnung gefunden werden [89].

A.3.2 Zeitgemittelte Zustandswahrscheinlichkeiten

Falls ein Beobachter, der das System zu einem zufälligen Zeitpunkt betrachtet, einen der Zustände 0 bis einschließlich S vorfindet, können zwischen dem Beobachtungszeitpunkt und dem vorhergehenden Bedienende keine Verluste aufgetreten sein. Es gelten also für y_k, $0 \leq k \leq S$, dieselben Beziehungen wie im reinen Wartesystem. Der Zustand $S+1$ muß jedoch getrennt betrachtet werden, da hier zwischen dem Beobachtungszeitpunkt und dem vorhergehenden Bedienende auch Verlustereignisse möglich sind.

Die zeitgemittelten Zustandswahrscheinlichkeiten sind im Warte-Verlustsystem auch von der Bearbeitungsreihenfolge abhängig, da das AF- und das DF-System unterschiedliche Verlustwahrscheinlichkeiten aufweisen.

A.3.2.1 Bearbeitungsreihenfolge AF

Hier können die Gleichungen für y_k, $0 \leq k \leq S$, direkt aus Abschnitt A.2.2 übernommen werden. Es gilt

$$y_0 = \frac{x_0}{\mathrm{E}[D]} \frac{1}{1-\xi_0}, \tag{A.44}$$

$$y_j = \frac{1}{\mathrm{E}[D]} \left[x_0 \sum_{u=2}^{\infty} \sum_{l=1}^{u-1} \xi_0^{l-1}(1-\xi_0)\, F^*(j, u-l-1) \sum_{\nu=u-l}^{\infty} h_\nu + \sum_{i=1}^{j} x_i \sum_{u=1}^{\infty} F(j-i, u-1) \sum_{\nu=u}^{\infty} h_\nu \right], \quad 0 < j \leq S, \tag{A.45}$$

mit $F(j,k) = [\xi^{(k)}]_j$ und $F^*(j,k) = \left[\gamma_i * [\xi^{(k)}]_i\right]_j$.

Bei der Bestimmung von y_{S+1} ist die endliche Warteschlange zu berücksichtigen, so daß zusätzlich über die Anzahl der Ankünfte bis zum Beobachtungszeitpunkt aufsummiert werden muß.

$$y_{S+1} = \frac{1}{\mathrm{E}[D]} \left[x_0 \sum_{u=2}^{\infty} \sum_{l=1}^{u-1} \xi_0^{l-1}(1-\xi_0) \sum_{j=S+1}^{\infty} F^*(j, u-l-1) \sum_{\nu=u-l}^{\infty} h_\nu + \sum_{i=1}^{S} x_i \sum_{u=1}^{\infty} \sum_{j=S+1}^{\infty} F(j-i, u-1) \sum_{\nu=u}^{\infty} h_\nu \right]. \tag{A.46}$$

Mit Hilfe von Gleichung A.21 kann ein rekursiver Algorithmus zur numerischen Berechnung der zeitgemittelten Zustandswahrscheinlichkeiten abgeleitet werden (siehe auch [30]). Dieser bezieht sich zwar zunächst nur auf das reine Wartesystem, kann jedoch unverändert auch auf das Warte-Verlustsystem übernommen werden, da die entsprechenden Gleichungen für x_k und y_k, $0 \leq k \leq S$, in beiden Systemen identisch sind. Aus A.21 folgt

$$\begin{aligned} y(z)\,(1-\xi(z)) &= \frac{1}{\mathrm{E}[D]}\,x(z)\,(1-z), \\ y(z)\,(1-\xi_0) &= y(z)\sum_{i=1}^{\infty}\xi_i z^i + \frac{1}{\mathrm{E}[D]}\,x(z)\,(1-z). \end{aligned}$$

Durch Rücktransformation erhält man daraus folgende rekursive Beziehung:

$$y_k = \frac{1}{1-\xi_0}\left[\sum_{i=1}^{k}\xi_i\,y_{k-i} + \frac{1}{\mathrm{E}[D]}\,(x_k - x_{k-1})\right], \qquad 1 \leq k \leq S. \tag{A.47}$$

Es muß beachtet werden, daß wegen der Verluste im System mit begrenzter Warteschlange die Abgangsrate nicht mehr gleich der Ankunftsrate ist. Der mittlere Abstand zwischen zwei Endeereignissen berechnet sich in einem AF-System zu

$$\begin{aligned} \mathrm{E}[D] &= x_0(\mathrm{E}[H]+\mathrm{E}[I]) + (1-x_0)\mathrm{E}[H] \;=\; \mathrm{E}[H] + x_0\mathrm{E}[I] \\ &= \mathrm{E}[H] + \frac{x_0}{1-\xi_0}. \end{aligned} \tag{A.48}$$

Die Wahrscheinlichkeit y_{S+1} kann nun aus $y_{S+1} = 1 - \sum_{k=0}^{S} y_k$ bestimmt werden. Für sehr kleine Werte von y_{S+1} können dabei allerdings numerische Probleme auftreten, so daß besser auf Gleichung A.46 zurückgegriffen wird. Diese Beziehung läßt sich außerdem in den meisten Sonderfällen stark vereinfachen.

A.3.2.2 Bearbeitungsreihenfolge DF

Bei der Bearbeitungsreihenfolge DF ist zu beachten, daß nach einem Bedienende im selben Zeitschlitz bereits weitere Anforderungen eintreffen können. Die drei Fälle der Herleitung in Abschnitt A.2.2 müssen daher entsprechend angepaßt werden.

1. Der Beobachtungszeitpunkt fällt in eine Freiperiode

Dies ist nur möglich, wenn das System zum letzten Regenerationszeitpunkt leer war, d.h. $S_{t_n} = 0$ gilt. Um zum Zeitpunkt $t_n + u$ ein leeres System vorzufinden, dürfen während der vergangenen u Zeiteinheiten keine Anforderungen ankommen. Daraus folgt

$$\mathrm{P}\{S_{t_n+u} = 0 \mid S_{t_n} = 0\} = \xi_0^u.$$

2. Der Beobachtungszeitpunkt fällt in die erste Bedienzeit nach einer Freiperiode

Auch hier unterscheidet sich das DF-System vom AF-System durch die Wahrscheinlichkeit, daß die Dauer der Freiperiode l Zeiteinheiten beträgt:

$$\mathrm{P}\{S_{t_n+u} = j, T_{n+1} > u \mid S_{t_n} = 0\} = \sum_{l=1}^{u-1} \xi_0^l (1-\xi_0)\, F^*(j, u-l-1) \sum_{\nu=u-l}^{\infty} h_\nu, \quad 1 \le j \le S.$$

3. Der Beobachtungszeitpunkt fällt in eine spätere Bedienzeit

In diesem Fall waren nach dem letzten Bedienende noch $i > 0$ Anforderungen im System. Damit zum Beobachtungszeitpunkt $t_n + u$ insgesamt j Anforderungen enthalten sind, müssen in den verbleibenden u Zeiteinheiten insgesamt $j - i$ Anforderungen eingetroffen sein:

$$\mathrm{P}\{S_{t_n+u} = j, T_{n+1} > u \mid S_{t_n} = i\} = F(j-i, u) \sum_{\nu=u}^{\infty} h_\nu,$$

für $1 \le j \le S$. Daraus folgt

$$y_0 = \frac{x_0}{\mathrm{E}[D]} \sum_{u=1}^{\infty} \mathrm{P}\{S_{t_n+u} = 0 \mid S_{t_n} = 0\} = \frac{x_0}{\mathrm{E}[D]} \frac{\xi_0}{1-\xi_0}, \tag{A.49}$$

$$y_j = \frac{1}{\mathrm{E}[D]} \left[x_0 \sum_{u=2}^{\infty} \sum_{l=1}^{u-1} \xi_0^l (1-\xi_0)\, F^*(j, u-l-1) \sum_{\nu=u-l}^{\infty} h_\nu + \sum_{i=1}^{j} x_i \sum_{u=1}^{\infty} F(j-i, u) \sum_{\nu=u}^{\infty} h_\nu \right], \quad 0 < j \le S, \tag{A.50}$$

$$y_{S+1} = \frac{1}{\mathrm{E}[D]} \left[x_0 \sum_{u=2}^{\infty} \sum_{l=1}^{u-1} \xi_0^l (1-\xi_0) \sum_{j=S+1}^{\infty} F^*(j, u-l-1) \sum_{\nu=u-l}^{\infty} h_\nu + \sum_{i=1}^{S} x_i \sum_{u=1}^{\infty} \sum_{j=S+1}^{\infty} F(j-i, u) \sum_{\nu=u}^{\infty} h_\nu \right]. \tag{A.51}$$

Auch im DF-System ist eine rekursive Berechnung der zeitgemittelten Zustandswahrscheinlichkeiten möglich. Da im reinen Wartesystem $x_{AF}(z) = \xi(z) x_{DF}(z)$ gilt, ergibt sich

$$y(z)\,(1-\xi(z)) = \frac{1}{\mathrm{E}[D]}\, \xi(z) x(z)(1-z),$$

$$y(z)\,(1-\xi_0) = y(z) \sum_{i=1}^{\infty} \xi_i z^i + \frac{1}{\mathrm{E}[D]}\, \xi(z) x(z)(1-z).$$

Die Rücktransformation liefert

$$y_k = \frac{1}{1-\xi_0} \left[\sum_{i=1}^{k} \xi_i\, y_{k-i} + \frac{1}{\mathrm{E}[D]} \sum_{i=0}^{k} \xi_i (x_{k-i} - x_{k-i-1}) \right], \qquad 1 \le k \le S, \tag{A.52}$$

wobei der mittlere Abstand zwischen zwei Endeereignissen im DF-Warte-Verlustsystem gegeben ist durch

$$\mathrm{E}[D] = x_0 \xi_0 (\mathrm{E}[H] + \mathrm{E}[I]) + (1 - x_0 \xi_0) \mathrm{E}[H] = \mathrm{E}[H] + \frac{x_0 \xi_0}{1-\xi_0}. \tag{A.53}$$

A.3.3 Zustandswahrscheinlichkeiten vor der Ankunft einer Anforderung

In diesem Abschnitt sollen die Zustandswahrscheinlichkeiten vor der Ankunft einer beliebigen Anforderung bestimmt werden, unabhängig davon, ob die Anforderung gespeichert werden kann oder verloren geht. Daraus läßt sich die Verlustwahrscheinlichkeit, also die Wahrscheinlichkeit, daß eine ankommende Anforderung ein volles System antrifft, direkt angeben:

$$\mathrm{P}\{\text{Anforderung geht verloren}\} = z_{S+1}.$$

Die Zustandsverteilung vor der Ankunft einer Anforderung hängt, wie im reinen Wartesystem auch, von zwei Komponenten ab:

- der Anzahl Z^* der Anforderungen im System vor Ankunft der Gruppe, zu der die betrachtete Anforderung gehört, und
- der Anzahl Ψ derjenigen Anforderungen, die gleichzeitig eintreffen, aber vor der betrachteten Anforderung gespeichert werden sollen.

Unter Berücksichtigung der begrenzten Warteschlangenlänge ergibt sich folgende Beziehung zwischen den Zustandswahrscheinlichkeiten z_k^* (vor dem Eintreffen der gesamten Gruppenankunft) und z_k (vor dem Eintreffen einer Anforderung):

$$z_k = \begin{cases} \displaystyle\sum_{i=0}^{k} z_i^* \psi_{k-i}, & 0 \le k \le S, \\ \displaystyle\sum_{i=0}^{S+1} z_i^* \sum_{j=S+1-i}^{\infty} \psi_j, & k = S+1. \end{cases} \tag{A.54}$$

A.3.3.1 Bearbeitungsreihenfolge AF

Wie auch im reinen Wartesystem sieht eine ankommende Gruppe den zeitlichen Mittelwert der Systemzustandsverteilung, d.h. $z_k^* = y_k$. Daraus folgt für die Zustandsverteilung aus der Sicht ankommender Anforderungen

$$z_k = \begin{cases} \displaystyle\sum_{i=0}^{k} y_i \psi_{k-i}, & 0 \le k \le S, \\ \displaystyle\sum_{i=0}^{S+1} y_i \sum_{j=S+1-i}^{\infty} \psi_j, & k = S+1. \end{cases} \tag{A.55}$$

A.3.3.2 Bearbeitungsreihenfolge DF

Für das DF-System erhält man, ebenfalls unter Berücksichtigung der begrenzten Warteschlange, folgende Beziehung zwischen y_k und den Zustandswahrscheinlichkeiten z_k^* vor dem Eintreffen der gesamten Gruppenankunft:

$$y_k = \begin{cases} \sum_{i=0}^{k} z_i^* \xi_{k-i}, & 0 \leq k \leq S, \\ \sum_{i=0}^{S} z_i^* \sum_{j=S+1-i}^{\infty} \xi_j + z_{S+1}^*, & k = S+1. \end{cases} \tag{A.56}$$

Daraus kann die Verteilung z_k^* durch „Rückfaltung" berechnet werden:

$$z_k^* = \begin{cases} \frac{y_0}{\xi_0}, & k = 0, \\ \frac{1}{\xi_0}(y_k - \sum_{i=0}^{k-1} z_i^* \xi_{k-i}), & 1 \leq k \leq S, \\ y_{S+1} - \sum_{i=0}^{S} z_i^* \sum_{j=S+1-i}^{\infty} \xi_j, & k = S+1. \end{cases} \tag{A.57}$$

Wegen der Differenzbildung kann die Rückfaltung allerdings zu numerischen Problemen führen. Anschließend muß die Verteilung z_k noch mit Hilfe von Gleichung A.54 berechnet werden.

A.3.4 Verteilung der Betriebsperiode

Auch im System mit endlicher Warteschlangenlänge wird zunächst die Verteilung von M, also der Anzahl der Anforderungen, die während einer Betriebsperiode bedient werden, benötigt. Mit Hilfe von Gleichung A.41 läßt sich daraus dann die Verteilung der Betriebsperiode bestimmen. Der in Abschnitt A.2.5 beschriebene Ansatz kann allerdings nicht auf Systeme mit begrenzter Warteschlange übertragen werden. Hier wird ein anderer Ansatz verwendet, welcher auf der Theorie der zeitdiskreten Phasenverteilungen basiert (siehe [157], Seite 240). Es folgt daher eine Zusammenfassung der wesentlichen Ergebnisse dieser Theorie, anschließend werden die Parameter der Verteilung getrennt für die Systemtypen AF und DF abgeleitet.

A.3.4.1 Zeitdiskrete Phasenverteilungen

Die Grundlage jeder Phasenverteilung ist eine Markov-Kette mit $m+1$ Zuständen, wobei die Zustände $\{1,\ldots,m\}$ transiente Zustände sind. Der Zustand $m+1$ ist ein absorbierender Zustand, das heißt, es sind keine Übergänge aus diesem Zustand heraus möglich. Eine zeitdiskrete Phasenverteilung gibt nun die Anzahl der Übergänge bis zum Erreichen des absorbierenden Zustands in einer zeitdiskreten Markov-Kette an, wobei der Ausgangszustand i ein transienter Zustand ist und σ_i die Wahrscheinlichkeit angibt, daß sich das System zum Startzeitpunkt im Zustand i, $1 \leq i \leq m$ aufhält.

Die Übergangsmatrix der Markov-Kette hat folgende allgemeine Form:

$$\mathbf{Q} = \begin{pmatrix} \mathbf{T} & \mathbf{t}_0 \\ \mathbf{0} & 0 \end{pmatrix}.$$

Die $m \times m$ Matrix $\mathbf{T}$ beschreibt alle Übergänge zwischen den transienten Zuständen, und der Spaltenvektor $\mathbf{t}_0$ enthält die Übergangswahrscheinlichkeiten in den absorbierenden Zustand. $\mathbf{T}$ ist eine sub-stochastische Matrix, das heißt alle Zeilensummen sind kleiner oder gleich 1, und deshalb ist die Matrix $\mathbf{I}-\mathbf{T}$ in jedem Fall nicht-singulär und somit invertierbar. Da $\mathbf{Q}$ eine stochastische Matrix ist, muß $\mathbf{T}\cdot\mathbf{e}+\mathbf{t}_0=\mathbf{e}$ gelten.

Die Anzahl N der Übergänge, die die Markov-Kette bis zum Erreichen des absorbierenden Zustands benötigt, hat nun die folgende zeitdiskrete Phasenverteilung [157]:

$$\mathrm{P}\{N=k\} = \boldsymbol{\sigma}\,\mathbf{T}^{k-1}\mathbf{t}_0, \qquad k \geq 1, \tag{A.58}$$

wobei die Wahrscheinlichkeiten σ_i zu dem „Startvektor" $\boldsymbol{\sigma} = (\sigma_1,\ldots,\sigma_m)$ zusammengefaßt wurden. Die Erzeugende Funktion der Phasenverteilung ist

$$f(z) = \sum_{k=1}^{\infty} \mathrm{P}\{N=k\}z^k = \boldsymbol{\sigma}\, z(\mathbf{I}-z\mathbf{T})^{-1}\mathbf{t}_0.$$

Durch vollständige Induktion läßt sich zeigen, daß die n-te Ableitung von $f(z)$ wie folgt geschrieben werden kann:

$$\frac{d^n}{dz^n} f(z) = n!\;\boldsymbol{\sigma}\;(\mathbf{I}-z\mathbf{T})^{-(n+1)}\;\mathbf{T}^{n-1}\;\mathbf{t}_0.$$

Die ersten beiden gewöhnlichen Momente ergeben sich daraus zu

$$\mathrm{E}[N] = f'(z)|_{z=1} = \boldsymbol{\sigma}\,(\mathbf{I}-\mathbf{T})^{-2}\mathbf{t}_0, \tag{A.59}$$

$$\mathrm{E}[N^2] = [f''(z)+f'(z)]_{z=1} = \boldsymbol{\sigma}\,(\mathbf{I}-\mathbf{T})^{-3}\,(\mathbf{I}+\mathbf{T})\,\mathbf{t}_0. \tag{A.60}$$

Für eine ausführlichere Behandlung zeitdiskreter und zeitkontinuierlicher Phasenverteilungen wird auf [157] verwiesen.

Die Verteilung von M, der Anzahl der Anforderungen, die während einer Betriebsperiode bedient werden, soll nun in Form einer Phasenverteilung beschrieben werden. Da M gleichzeitig die Anzahl der Übergänge der eingebetteten Markov-Kette vom Beginn der Betriebsperiode bis zum Erreichen der folgenden Freiperiode ist, muß diese Darstellung auf jeden Fall möglich sein. Die Übergangsmatrix der Phasenverteilung kann aus der eingebetteten Markov-Kette durch entsprechende Wahl des absorbierenden Zustands abgeleitet werden.

A.3.4.2 Bearbeitungsreihenfolge AF

In einem AF-System endet die Betriebsperiode genau dann, wenn das System nach einem Bedienende leer ist. Der Zustand $X = 0$ entspricht daher dem absorbierenden Zustand der Phasenverteilung, die Zustände $X = k$, $1 \leq k \leq S$ sind transiente Zustände. Es muß außerdem beachtet werden, daß nach der ersten Gruppenankunft zu Beginn der Betriebsperiode bis zu $S + 1$ Anforderungen im System sein können. Der Zustand $S + 1$ kann jedoch nur als Ausgangszustand auftreten, nach einem Bedienende gilt immer $X \leq S$.

Die Matrix $\mathbf{T}$ entspricht im wesentlichen der Übergangsmatrix der eingebetteten Markov-Kette ohne die erste Zeile und Spalte. Außerdem muß der Zustand $S + 1$ als möglicher Ausgangszustand hinzugefügt werden.

$$\mathbf{T} = \begin{pmatrix} v_1 & v_2 & \cdots & v_{S-1} & \sum_{i=S}^{\infty} v_i & 0 \\ v_0 & v_1 & \cdots & v_{S-2} & \sum_{i=S-1}^{\infty} v_i & 0 \\ 0 & v_0 & \cdots & v_{S-3} & \sum_{i=S-2}^{\infty} v_i & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & v_0 & \sum_{i=1}^{\infty} v_i & 0 \\ 0 & \cdots & 0 & 0 & 1 & 0 \end{pmatrix}.$$

Ein Übergang in den absorbierenden Zustand ist nur aus dem Zustand $X = 1$ möglich, die entsprechende Übergangswahrscheinlichkeit ist v_0. Der Spaltenvektor $\mathbf{t}_0$ ist daher

$$\mathbf{t}_0 = (v_0, 0, \ldots, 0)^T.$$

Der Ausgangszustand der Phasenverteilung hängt von der Gruppengröße der ersten Ankunft zu Beginn der Betriebsperiode ab:

$$\boldsymbol{\sigma} = (\gamma_1, \gamma_2, \ldots, \gamma_S, \sum_{i=S+1}^{\infty} \gamma_i).$$

Mit Hilfe von Beziehung A.58 kann nun die Verteilung $\mathrm{P}\{M = k\}$ numerisch durch fortgesetzte Matrizenmultiplikation berechnet werden.

A.3.4.3 Bearbeitungsreihenfolge DF

Bei der Ableitung der Verteilung von M im DF-System muß beachtet werden, daß hier auch während einer Betriebsperiode die Möglichkeit besteht, daß das System nach einem Bedienende leer ist. Die Betriebsperiode endet erst, wenn keine neuen Anforderungen gleichzeitig mit diesem Bedienende eintreffen. Der Zustand $X = 0$ ist daher in der Markov-Kette der Phasenverteilung ebenfalls ein transienter Zustand, an den sich mit Wahrscheinlichkeit ξ_0 eine Freiperiode anschließt. Mit Wahrscheinlichkeit $1 - \xi_0$ bleibt das System innerhalb der transienten Zustände, wobei die Übergangswahrscheinlichkeit in den Zustand k durch $(1 - \xi_0)v_k^*$, $0 \leq k < S$, gegeben ist.

Eine weitere Besonderheit des DF-Systems besteht darin, daß die erste Bedienzeit einer Betriebsperiode gesondert behandelt werden muß (siehe auch Abschnitt A.2.1.2). Aus diesem Grund wird der Zustand des Systems nach dem ersten Bedienende einer Betriebsperiode als Ausgangszustand der Phasenverteilung gewählt. Die Anzahl der Ankünfte bis zum ersten Bedienende muß also durch den Startvektor entsprechend berücksichtigt werden. Wegen der Verschiebung des Startzeitpunktes hinter das erste Bedienende kann auch der Zustand $S+1$ in der Markov-Kette der Phasenverteilung nicht mehr auftreten. Die obigen Überlegungen führen somit zu folgender Struktur der Matrix $\mathbf{T}$:

$$\mathbf{T} = \begin{pmatrix} (1-\xi_0)v_0^* & (1-\xi_0)v_1^* & (1-\xi_0)v_2^* & \cdots & (1-\xi_0)v_{S-1}^* & (1-\xi_0)\sum_{i=S}^{\infty} v_i^* \\ v_0 & v_1 & v_2 & \cdots & v_{S-1} & \sum_{i=S}^{\infty} v_i \\ 0 & v_0 & v_1 & \cdots & v_{S-2} & \sum_{i=S-1}^{\infty} v_i \\ 0 & 0 & v_0 & \cdots & v_{S-3} & \sum_{i=S-2}^{\infty} v_i \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & v_0 & \sum_{i=1}^{\infty} v_i \end{pmatrix}.$$

Der absorbierende Zustand wird, wie bereits erwähnt, mit Wahrscheinlichkeit ξ_0 aus dem Zustand 0 erreicht:

$$\mathbf{t}_0 = (\xi_0, 0, \ldots, 0)^T.$$

Die Elemente σ_i des Startvektors geben die Wahrscheinlichkeit an, daß sich das System nach dem ersten Bedienende im Zustand i befindet. Dies entspricht exakt der Definition der Übergangswahrscheinlichkeiten v_i^*, es folgt also

$$\boldsymbol{\sigma} = (v_0^*, v_1^*, \ldots, v_{S-1}^*, \sum_{i=S}^{\infty} v_i^*).$$

Anhang B

Verteilung der Durchlaufzeit im $D^{[X]}/D/1$-System mit zufälliger Abfertigung

Die Abfertigungsdisziplin bestimmt die Reihenfolge, in der wartende Anforderungen aus der Warteschlange entnommen werden. Häufig anzutreffende Abfertigungsdisziplinen sind

- FIFO: Abfertigung in der Reihenfolge des Eintreffens,
- LIFO: zuletzt eingetroffene Anforderungen werden zuerst bedient, und
- SIRO: Abfertigung in zufälliger Reihenfolge.

Da die Anzahl der Anforderungen im System von der Reihenfolge ihrer Bedienung unabhängig ist, beeinflußt die Abfertigungsdisziplin nur die Wartezeitverteilung, nicht jedoch die Zustandsverteilung. Die Berechnung der Wartezeitverteilung des M/G/1-Systems ist in [198] für die oben angegebenen Abfertigungsdisziplinen beschrieben. Diese Vorgehensweise kann prinzipiell auch auf das $D^{[X]}/G/1$-System mit unbegrenzter Warteschlangenlänge übertragen werden, im folgenden soll jedoch nur der Sonderfall der zufälligen Abfertigung (SIRO) bei konstanter Bedienzeit $H = 1$ betrachtet werden.

Die Zufallsvariable Δ sei die Durchlaufzeit einer Anforderung durch das gesamte System, also die Summe aus Wartezeit und Bedienzeit. Unter der Bedingung, daß die Anzahl Y^+ der Anforderungen im System unmittelbar vor dem Beginn einer Bedienung gegeben ist, läßt sich die Verteilung der Durchlaufzeit mit Hilfe einer rekursiven Beziehung angeben [198]. Falls das System zu diesem Zeitpunkt $Y^+ = i$ Anforderungen enthält, wird wegen der zufälligen Abfertigung jede dieser Anforderungen gleichwahrscheinlich bedient. Die Wahrscheinlichkeit dafür, daß die verbleibende Durchlaufzeit der betrachteten Anforderung genau eine Zeiteinheit beträgt, ist somit $1/i$:

$$P\{\Delta = 1 \mid Y^+ = i\} = \frac{1}{i}, \quad i \geq 1.$$

Mit Wahrscheinlichkeit $1-1/i$ muß die betrachtete Anforderung bis zum nächsten Zeitschlitz warten. Die Durchlaufzeit erhöht sich in diesem Fall um eine Zeiteinheit und um die verbleibende Durchlaufzeit nach Ankunft der Anforderungen im nächsten Zeitschlitz. Falls j neue Anforderungen im nächsten Zeitschlitz eintreffen, befinden sich anschließend $\max(0, i-1+j)$ Anforderungen im System. Nach dem Gesetz der totalen Wahrscheinlichkeit gilt daher

$$\mathrm{P}\{\Delta = k \mid Y^+ = i\} = \begin{cases} 0, & i = 1,\ k \geq 2, \\ (1-\frac{1}{i}) \sum_{j=0}^{\infty} \mathrm{P}\{\Delta = k-1 \mid Y^+ = i-1+j\} \cdot \xi_j, & i \geq 2,\ k \geq 2. \end{cases}$$

Beginnend mit $k = 1$ kann die bedingte Verteilung $\mathrm{P}\{\Delta = k \mid Y^+ = i\}$ auf diese Weise numerisch berechnet werden.

Um die Bedingung $Y^+ = i$ zu eliminieren, wird auch die Verteilung von Y^+ benötigt, die man durch eine einfache Betrachtung aus den zeitgemittelten Zustandswahrscheinlichkeiten erhält. Wegen der Gedächtnisfreiheit des Ankunftsprozesses ist die Zustandsverteilung des Systems im Zeitschlitz vor der Ankunft der betrachteten Anforderung durch y_n gegeben. Entsprechend Gleichung A.24 besitzt die Größe Ψ der Gruppe, in der sich die betrachtete Anforderung befindet, die folgende Verteilung

$$\psi_k = \frac{1}{\lambda}\,(1 - \mathrm{P}\{\Xi \leq k\}),$$

wobei Ξ die Größe einer beliebigen Gruppenankunft ist.

Falls das System im vorhergehenden Zeitschlitz leer war und $\Psi = i$ neue Anforderungen eintreffen, befinden sich zu Beginn der Bedienung ebenfalls $Y^+ = i$ Anforderungen im System. Andernfalls muß auch das gleichzeitig mit der Ankunft stattfindende Bedienende berücksichtigt werden. Es ergibt sich somit

$$\mathrm{P}\{Y^+ = i\} = y_0\psi_i + \sum_{\nu=1}^{i} y_\nu \psi_{i-\nu+1}, \quad i \geq 1.$$

Die Durchlaufzeitverteilung erhält man schließlich nach dem Gesetz der totalen Wahrscheinlichkeit:

$$\mathrm{P}\{\Delta = k\} = \sum_{i=1}^{\infty} \mathrm{P}\{\Delta = k \mid Y^+ = i\} \cdot \mathrm{P}\{Y^+ = i\}$$

In Bild B.1 ist ein Vergleich der Durchlaufzeitverteilungen für die Abfertigungsdisziplinen SIRO und FIFO anhand eines Beispiels dargestellt. Es handelt sich hier um das Modell für einen Ausgangspuffer, so daß die Gruppengröße des Ankunftsprozesses durch eine Binomialverteilung beschrieben werden kann (siehe Abschnitt 4.1). Die Ergebnisse zeigen deutlich, daß eine Abfertigung in zufälliger Reihenfolge zu wesentlich größeren Durchlaufzeitschwankungen führt als die Abfertigungsdisziplin FIFO.

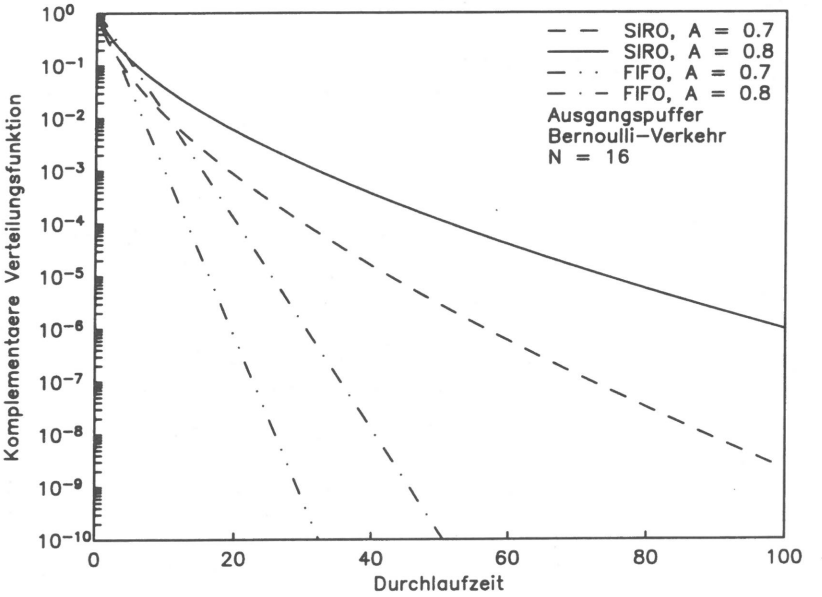


Bild B.1: Durchlaufzeitverteilung für die Abfertigungsdisziplinen SIRO und FIFO

Anhang C

Der Markov-Modulierte Bernoulli-Prozeß

C.1 Definition

Der Markov-Modulierte Bernoulli-Prozeß (MMBP) ist ein doppelt-stochastischer Prozeß, dessen momentane Ankunftsrate durch eine Markov-Kette gesteuert wird. Der Prozeß basiert auf einer zeitdiskreten Markov-Kette mit n Zuständen, deren Verhalten durch die Übergangsmatrix $\mathbf{P} = [p_{ij}]$ bestimmt wird. Dabei ist p_{ij} die Wahrscheinlichkeit für einen Übergang vom Zustand i in den Zustand j. Die mittlere Dauer des Zustands i ist durch $1/(1-p_{ii})$ gegeben.

Wenn sich die Markov-Kette in Zustand i befindet, ist die Wahrscheinlichkeit für ein Ankunftsereignis durch q_i gegeben. Innerhalb eines Zustands werden die Ankünfte also wie bei einem Bernoulli-Prozeß mit Ankunftsrate q_i erzeugt. Der MMBP ist somit eine zeitdiskrete Version des Markov-Modulierten Poisson-Prozesses (MMPP, siehe z.B. [157]). Zur Vereinfachung der Schreibweise werden die Ankunftswahrscheinlichkeiten q_i zu dem Spaltenvektor $\mathbf{q} = (q_1, q_2, \ldots, q_n)^T$ zusammengefaßt. Der Markov-Modulierte Bernoulli-Prozeß ist somit durch die Übergangsmatrix $\mathbf{P}$ und den Vektor $\mathbf{q}$ vollständig definiert.

C.2 Charakteristische Größen

Die stationären Zustandswahrscheinlichkeiten der Markov-Kette (π_i) erhält man als Lösung des folgenden linearen Gleichungssystems:

$$\boldsymbol{\pi}\,\mathbf{P} = \boldsymbol{\pi}, \quad \boldsymbol{\pi}\,\mathbf{e} = 1. \tag{C.1}$$

Der Zeilenvektor $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_n)$ enthält die stationären Zustandswahrscheinlichkeiten, und der Spaltenvektor $\mathbf{e}$ ist $\mathbf{e} = (1, 1, \ldots, 1)^T$. π_i ist die Wahrscheinlichkeit, daß sich die Markov-Kette zu einem beliebigen Zeitpunkt im Zustand i befindet.

Es sei Ξ die Anzahl der Ankünfte (0 oder 1) des stationären MMBP zu einem beliebigen Zeitpunkt. Die mittlere Ankunftsrate ist

$$\lambda = \mathrm{E}[\Xi] = \mathrm{P}\{\Xi = 1\} = \sum_{i=1}^{n} \pi_i q_i = \boldsymbol{\pi}\ \mathbf{q}. \tag{C.2}$$

Der mittlere Ankunftsabstand $\mathrm{E}[T_A]$ kann mit $\mathrm{E}[T_A] = 1/\lambda$ ebenfalls sofort angegeben werden. Für das zweite gewöhnliche Moment von Ξ gilt

$$\mathrm{E}[\Xi^2] = \mathrm{P}\{\Xi = 1\} = \lambda = \boldsymbol{\pi}\ \mathbf{q}. \tag{C.3}$$

Die Varianz von Ξ ist daher

$$\mathrm{VAR}[\Xi] = \mathrm{E}[\Xi^2] - \mathrm{E}[\Xi]^2 = \lambda(1 - \lambda). \tag{C.4}$$

C.3 Analyse des Korrelationsverhaltens

Das Korrelationsverhalten eines Markov-Modulierten Bernoulli-Prozesses ist durch die Korrelation zwischen den Ankünften zu verschiedenen Zeitpunkten bestimmt. Ein Maß für die Korrelation zwischen zwei Zufallsvariablen ist die Kovarianz bzw. der Korrelationskoeffizient.

Es seien Ξ_j und Ξ_{j+k} jeweils die Anzahl der Ankünfte zu den Zeitpunkten j und $j+k$. Falls sich die Markov-Kette zum Zeitpunkt j bereits im stationären Zustand befindet, hängt die Korrelation nur vom Abstand k der beiden Zeitpunkte ab.

Die Kovarianz $\mathrm{COV}[\Xi_j, \Xi_{j+k}]$ ist wie folgt definiert:

$$\mathrm{COV}[\Xi_j, \Xi_{j+k}] = \mathrm{E}[(\Xi_j - \mathrm{E}[\Xi_j])(\Xi_{j+k} - \mathrm{E}[\Xi_{j+k}])].$$

Im stationären Zustand gilt jedoch $\mathrm{E}[\Xi_j] = \mathrm{E}[\Xi_{j+k}] = \lambda$, es ergibt sich daher

$$\mathrm{COV}[\Xi_j, \Xi_{j+k}] = \mathrm{E}[\Xi_j \Xi_{j+k}] - \lambda^2 = \mathrm{P}\{\Xi_j = 1, \Xi_{j+k} = 1\} - \lambda^2. \tag{C.5}$$

Die Wahrscheinlichkeit $\mathrm{P}\{\Xi_j = 1, \Xi_{j+k} = 1\}$ kann mit Hilfe der Übergangsmatrix $\mathbf{P}$ berechnet werden. Unter der Voraussetzung der Stationarität ist $\pi_i q_i$ die Wahrscheinlichkeit dafür, daß sich die Markov-Kette zum Zeitpunkt j im Zustand i befindet und ein Ankunftsereignis auftritt. Der Zeilenvektor mit den Wahrscheinlichkeiten $\pi_i q_i$ ergibt sich aus einer Multiplikation von $\boldsymbol{\pi}$ mit der Diagonalmatrix $\mathbf{Q} = \mathrm{diag}(\mathbf{q})$. Vom Zeitpunkt j bis zum Zeitpunkt $j+k$ erfolgen k Übergänge der Markov-Kette, die entsprechende Übergangsmatrix ist daher

durch $\mathbf{P}^k$ gegeben. Die Wahrscheinlichkeit, daß zum Zeitpunkt $j+k$ ebenfalls ein Ankunftsereignis auftritt, kann durch eine erneute Multiplikation mit der Matrix $\mathbf{Q}$ berücksichtigt werden, und durch eine anschließende Multiplikation mit dem Spaltenvektor $\mathbf{e}$ werden alle Wahrscheinlichkeiten aufsummiert. Es gilt also

$$\mathrm{P}\{\Xi_j = 1, \Xi_{j+k} = 1\} = \boldsymbol{\pi}\ \mathbf{Q}\ \mathbf{P}^k\ \mathbf{Q}\ \mathbf{e}, \quad k \geq 1, \tag{C.6}$$

mit

$$\mathbf{Q} = \mathrm{diag}(\mathbf{q}) = \begin{pmatrix} q_1 & 0 & 0 & \dots & 0 \\ 0 & q_2 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & q_{n-1} & 0 \\ 0 & \dots & 0 & 0 & q_n \end{pmatrix}.$$

Falls die Übergangsmatrix $\mathbf{P}$ diagonalisierbar ist, kann die k-te Potenz $\mathbf{P}^k$ mit Hilfe der Eigenwerte und Eigenvektoren von $\mathbf{P}$ berechnet werden (siehe [55]). Eine $n \times n$-Matrix ist diagonalisierbar, wenn sie n linear unabhängige Eigenvektoren besitzt. Diese Bedingung ist in jedem Fall erfüllt, wenn alle n Eigenwerte verschieden sind.

Es seien μ_i, $1 \leq i \leq n$, die Eigenwerte der Matrix $\mathbf{P}$. Jedem Eigenwert μ_i ist ein linksseitiger Eigenvektor $\boldsymbol{\varphi}_i$ sowie ein rechtsseitiger Eigenvektor $\mathbf{f}_i$ zugeordnet. Die Eigenvektoren sind jeweils die Lösung eines homogenen linearen Gleichungssystems:

$$\boldsymbol{\varphi}_i\ \mathbf{P} = \mu_i\ \boldsymbol{\varphi}_i, \tag{C.7}$$

$$\mathbf{P}\ \mathbf{f}_i = \mu_i\ \mathbf{f}_i. \tag{C.8}$$

Weiterhin wird vorausgesetzt, daß die Eigenvektoren folgende Bedingung erfüllen:

$$\boldsymbol{\varphi}_i\ \mathbf{f}_i = 1. \tag{C.9}$$

Diese Normierungsbedingung bedeutet keine Einschränkung der Allgemeinheit, da die Eigenvektoren ohnehin nur bis auf eine multiplikative Konstante bestimmt sind. Unter diesen Voraussetzungen ist die k-te Potenz der Matrix $\mathbf{P}$ durch folgende Beziehung gegeben [55]:

$$\mathbf{P}^k = \sum_{i=1}^{n} \mu_i^k \mathbf{F}_i. \tag{C.10}$$

Die $n \times n$-Matrizen $\mathbf{F}_i$ erhält man aus einer Multiplikation der Eigenvektoren $\mathbf{f}_i$ und $\boldsymbol{\varphi}_i$, wobei der Spaltenvektor $\mathbf{f}_i$ als eine $n \times 1$-Matrix und der Zeilenvektor $\boldsymbol{\varphi}_i$ als eine $1 \times n$-Matrix aufgefaßt wird.

$$\mathbf{F}_i = \mathbf{f}_i\ \boldsymbol{\varphi}_i. \tag{C.11}$$

Gleichung C.10 bezeichnet man auch als Spektraldarstellung der Matrix $\mathbf{P}^k$. Einsetzen von Gleichung C.10 in Gleichung C.6 liefert

$$\mathrm{P}\{\Xi_j = 1, \Xi_{j+k} = 1\} = \sum_{i=1}^{n} \mu_i^k (\boldsymbol{\pi}\ \mathbf{Q}\ \mathbf{F}_i\ \mathbf{Q}\ \mathbf{e}) = \sum_{i=1}^{n} \mu_i^k a_i, \tag{C.12}$$

mit $a_i = \boldsymbol{\pi}\ \mathbf{Q}\ \mathbf{F}_i\ \mathbf{Q}\ \mathbf{e}$. Es ergibt sich somit

$$\mathrm{COV}[\Xi_j, \Xi_{j+k}] = \sum_{i=1}^{n} \mu_i^k a_i - \lambda^2. \tag{C.13}$$

Da $\mathbf{P}$ eine Übergangsmatrix ist, sind alle Elemente von $\mathbf{P}$ reelle Zahlen zwischen 0 und 1, wobei die Summe der Elemente einer Zeile stets 1 ergibt. Unter der zusätzlichen Bedingung, daß $\mathbf{P}$ irreduzibel ist, kann gezeigt werden, daß $\mathbf{P}$ den einfachen Eigenwert 1 besitzt [55]. Der stationäre Zustandsvektor $\boldsymbol{\pi}$ erfüllt das Gleichungssystem $\boldsymbol{\pi}\ \mathbf{P} = \boldsymbol{\pi}$ und ist daher ein linksseitiger Eigenvektor zum Eigenwert 1. Der rechtsseitige Eigenvektor zum Eigenwert 1 ist durch den Spaltenvektor $\mathbf{e}$ gegeben. Außerdem gilt, daß der Betrag aller anderen Eigenwerte von $\mathbf{P}$ kleiner als 1 ist.

Es soll nun angenommen werden, daß die Eigenwerte nach ihrem Betrag geordnet werden, so daß

$$1 > |\mu_2| \geq |\mu_3| \geq \ldots \geq |\mu_n|.$$

gilt. Für die Kovarianz erhält man folglich

$$\begin{aligned} \mathrm{COV}[\Xi_j, \Xi_{j+k}] &= \boldsymbol{\pi}\ \mathbf{Q}\ \mathbf{F}_1 \mathbf{Q}\ \mathbf{e} + \sum_{i=2}^{n} \mu_i^k a_i - \lambda^2 \\ &= \boldsymbol{\pi}\ \mathbf{Q}\ \mathbf{e}\ \boldsymbol{\pi}\ \mathbf{Q}\ \mathbf{e} + \sum_{i=2}^{n} \mu_i^k a_i - \lambda^2 \\ &= \sum_{i=2}^{n} \mu_i^k a_i. \end{aligned} \tag{C.14}$$

Dabei wurde von der Beziehung $\boldsymbol{\pi}\ \mathbf{Q}\ \mathbf{e} = \boldsymbol{\pi}\ \mathbf{q} = \lambda$ Gebrauch gemacht. Die Abhängigkeit der Kovarianz von k ist also durch eine Summe aus $(n-1)$ exponentiell abnehmenden Termen gegeben. Das asymptotische Verhalten für große k wird dabei durch den Eigenwert mit dem größten Betrag (μ_2) bestimmt. Gleichung C.14 bestätigt außerdem, daß die Kovarianz der Ankünfte mit zunehmendem Abstand gegen 0 geht.