

Institut für Nachrichtenvermittlung und Datenverarbeitung
der Universität Stuttgart
Prof. Dr.-Ing. P. Kühn

Institute of Communications Switching and Data Technics
University of Stuttgart
Prof. Dr.-Ing. P. Kühn

42. Bericht über verkehrstheoretische Arbeiten

42nd Report on Studies in Congestion Theory

Untersuchungen zu Vermittlungssystemen
mit verteilter Steuerung

On the Analysis and Design of Switching Systems
with Distributed Control

von
Wolfgang Denzel

by
Wolfgang Denzel

© 1987 Institut für Nachrichtenvermittlung und Datenverarbeitung Universität Stuttgart

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

ISBN 3-922403-52-2

ABSTRACT

Since the year 1965, telephony switching systems have been built which are controlled by a stored program in a computer. This offers economic advantages, more flexibility, and the possibility for new functions. But in the beginning, a centralized processor was the only economic control structure.

Since that time, the progress in microelectronics has allowed the full digitalization of the switches and so the real integration of voice and data. This opened the possibility for numerous new functions and services and required more and more processing power from the central control. With these higher requirements and the availability of cheap and powerful microprocessors, distributed control structures now become economically feasible and even offer numerous new advantages. But on the other side, new problems and questions arise:

First, one has to find the optimal system structure and an appropriate function and load distribution. Thereby, especially the consequences for the inter-processor communication have to be obeyed.

Because of the complexity of distributed systems, there is also the question how to investigate their performance, and which methods and rules should be applied to design and configure these systems.

The present report focuses on these problems. Various solutions are suggested which allow a detailed modelling and analysis of switching system control structures suitable to support the design of new systems.

CHAPTER 1: Introduction

The first chapter introduces the arguments for distributing the control of SPC switching systems. Furthermore, the essential

problems on the structure and the performance analysis of SPC systems with distributed control are shown.

CHAPTER 2: Conception of Switching Controls

This chapter deals in general with the conception of SPC switching systems. First, the most significant control functions of switching systems are reviewed. Then, the fundamental control structures of centralized, decentralized, and partially centralized systems are compared and their advantages and disadvantages are discussed. The possibilities for sharing the functions as well as the load are considered and the resultant problems are shown.

CHAPTER 3: Integrated Services Switching System "PILOT"

Here, a concrete example for a distributed switching control is presented by means of the implemented laboratory model of the integrated services local switching system "PILOT". The development steps of this system as well as the final partially centralized structure and the functions of the various units are described from the control point of view.

Since the inter-processor communication in a distributed system is a very important subject, the function of the main bus system is shown in very detail. Therefore, a protocol for the scheduling of tasks waiting in transmit queues for being sent to the receive queues of other processing units has been developed. It allows a flexible and dynamic load and function allocation to the receiving processing units.

CHAPTER 4: Analysis of Switching Controls

This chapter deals in general with the modeling of switching controls and subsystems as well as their system environment. Three analysis techniques which are especially useful for the

2.3	Probleme der Funktions- und Lastteilung	35
2.3.1	Funktionsteilung	35
2.3.2	Lastteilung	36
2.3.3	Datenspeicherung	38
2.3.4	Interprozessor-Kommunikation	39
3	DIENSTE-INTEGRIERTES VERMITTLUNGSSYSTEM "PILOT" .	41
3.1	Gesamtkonzept	41
3.1.1	Ausgangssituation: Zentralgesteuerte Fern- sprechvermittlungsstelle	41
3.1.2	Übergangslösung: Kombinierte Fernsprech/Daten- vermittlungsstelle	43
3.1.3	Dienste-integriertes Vermittlungssystem mit verteilter Steuerung	45
3.2	Aufgabenverteilung	49
3.2.1	Ausgelagerter Vermittlungsmodul	49
3.2.1.1	Globale Aufgaben	49
3.2.1.2	Dezentralisierte Funktionen	50
3.2.2	Vermittlungsstelle	50
3.2.2.1	Globale Aufgaben	50
3.2.2.2	Dezentralisierte Funktionen	51
3.3	Interprozessor-Kommunikation	53
3.3.1	Kommunikation innerhalb des Vermittlungsmoduls	54
3.3.2	Signalisier-Kommunikation zwischen Vermitt- lungsmodul und Vermittlungsstelle	55
3.3.3	Bus-Kommunikation innerhalb der Vermittlungs- stelle	55
4	ANALYSE VON VERMITTLUNGSSTEUERUNGEN	62
4.1	Allgemeines über Analysetechniken und -ziele . .	62
4.2	Modellierung von Vermittlungssteuerungen	65
4.2.1	Verkehrsmodelle	65
4.2.1.1	Ankunftsprozesse	65
4.2.1.2	Bedienungsprozesse	68

4.2.2	Systemmodelle	69
4.2.2.1	Gesamtmodell eines Vermittlungssystems . . .	70
4.2.2.2	Teilmodelle von Vermittlungssystemen	72
4.3	Signalflußanalyse	75
4.3.1	Vorgaben für die Signalflußanalyse	75
4.3.1.1	Steuerungsstruktur	76
4.3.1.2	Belegungsarten und Belegungsspektrum	77
4.3.1.3	Meldungsszenarien	78
4.3.2	Prinzip der Signalflußanalyse	80
4.4	Simulation	83
4.4.1	Grundprinzip der zeitreuen Simulation	83
4.4.2	Zeitreue Simulation SDL-spezifizierter Pro- zesse	84
4.5	Warteschlangenmodellanalyse	88
4.5.1	Zustandsprozesse	88
4.5.2	Markoffsche Prozesse	89
4.5.3	Analyse Markoffscher Systeme	90
4.5.3.1	Zustandswahrscheinlichkeiten	90
4.5.3.2	Lösungsschritte	92
5	EINZELUNTERSUCHUNGEN	96
5.1	Strukturuntersuchungen mittels Signalflußanalyse	96
5.1.1	Vorgaben	96
5.1.2	Verallgemeinerte Ergebnisse	97
5.1.3	Anwendungsbeispiel	99
5.1.4	Schlußfolgerungen	101
5.2	Simulation des Vermittlungssystems PILOT	103
5.2.1	Simulationsmodell	103
5.2.1.1	Verkehrsgenerator	104
5.2.1.2	Grundmodul	105
5.2.1.3	Busmodell	107
5.2.2	Simulationsergebnisse	107
5.3	Warteschlangenanalyse von Teilmodellen	113

5.3.1	Analyse von Zuteilungsstrategien	113
5.3.1.1	Zuteilungsstrategien	113
5.3.1.2	Grundmodell	114
5.3.1.3	Analyse	115
5.3.1.4	Implementierungsaspekte	122
5.3.1.5	Ergebnisse	124
5.3.2	Analyse von Abfertigungsstrategien	131
5.3.2.1	Abfertigungsstrategien	131
5.3.2.2	Grundmodell	132
5.3.2.3	Analyse	133
5.3.2.4	Ergebnisse	134
6	ZUSAMMENFASSUNG	141

SCHRIFTTUM

- / 1/ Keister,W. : The Evolution of Telephone Switching. Bell Labs. Rec., 43 (1965) 6, 197-203.
- / 2/ Elektronisches Wählsystem EWS, Systemübersicht. Leistungsmerkmale. Siemens AG, München (1980), 22-24.
- / 3/ Keister,W. et al. : No. 1 ESS. The Bell System Technical Journal, 43 (1964) 5, Part 1, 1831-3075.
- / 4/ System IV - Ein Fernsprech-Vermittlungssystem mit gespeichertem Steuerprogramm. Siemens AG, München (1970).
- / 5/ Elektronisches Wählsystem EWS1. Ortsamtstechnik. Siemens AG, München (1972).
- / 6/ Lotze,A., Rothmaier,K., Scheller,R. : TDM Versus SDM Switching Arrays - A Comparison. 9. ITC, Torremolinos (1979), Paper 345.
- / 7/ Scheller,R. : Verkehrsgerechte Dimensionierung, Analyse und Kostenoptimierung von PCM-Koppelanordnungen. Dissertationsschrift, Universität Stuttgart (1981).
- / 8/ Schwertfeger,H.-J. : Kommunikationseinrichtungen in verteilten Strukturen zur Steuerung von Vermittlungsprozessen. Nachrichtentechn. Elektronik 10, Berlin-Ost (1984), 368-371.
- / 9/ EWSD. Digitales Elektronisches Wählsystem. Systembeschreibung. Siemens AG, München, A30808-X2589-X-51-18.
- /10/ Asmussen,H., Frensch,K.-J., Botsch,D. et al. : Digital-Vermittlungssystem EWSD für öffentliche Fernsprechnetze. (14 Fachartikel), Telcom-Report 4 (1981), Beiheft.
- /11/ System 12. Das Digital-Vermittlungssystem. Technische Beschreibung. Standard Elektrik Lorenz AG, Stuttgart, 1754.3802.5.
- /12/ Langenbach-Belz,M. et al. : System 12. (17 Fachartikel), Elektrisches Nachrichtenwesen 56 (1981), Nr.2/3.

- /13/ Davis,J.H., Janik,J., Royers,R.D., Yokelson,B.J. : No. 5 ESS - System Architecture. ISS, Montreal, Canada (1981), Paper 31 A 2.
- /14/ Bosco,H.L., Eisenhart,R.K., Saal,F.A., Scheerer,W.G. : No. 5 ESS - Hardware Design. ISS, Montreal, Canada (1981), Paper 31 A 3.
- /15/ Kaderali,F., Giebler,H.J. : Characteristics of a Digital Exchange for ISDN. ISS, Florence, Italy (1984), Paper 22 B 4.
- /16/ Becker,D., Schmidt,P. : Zur Integration von Fernmeldediensten in digitalen Netzen. NTZ, Band 34, Heft 5/6 (1981), 288-292/366-369.
- /17/ Rosenbrock,K.-H. : Mögliche Integration von Fernmeldediensten im digitalen Fernsprechnetz der Deutschen Bundespost - ISDN. Zeitschrift für das Post- und Fernmeldewesen, Heft 9 (1982), 24-31.
- /18/ Weiß,W. : Über Struktur, Betriebsweise und Verkehrsleistung eines diensteintegrierten Vermittlungssystems für Fernsprech- und Datenverkehr. Dissertationsschrift, Universität Stuttgart (1984).
- /19/ Eberding,H. : Die Software im System EWSD. Telcom-Report 4 (1981), Beiheft, 13-18.
- /20/ Katzschner,L., van den Brande,F. : System 12. Konzeption und Realisierung der Software. Elektrisches Nachrichtenwesen, 56 (1981), Nr.2/3, 49-54.
- /21/ Duncan,T., Huen,Wing H. : Software Structure of No. 5 ESS - A Distributed Telephone Switching System. IEEE Transactions on Communications, Vol. COM-30, 6 (1982), 1379-1385.
- /22/ Joel,Jr.A.E. : The Classification and Unification of Switching System Functions. ISS, MIT Cambridge, USA (1972), 446-453.
- /23/ Wizgall,M. : Über Architektur, Betriebsweise und Verkehrs-

- verhalten der Steuerung einer rechnergesteuerten Vermittlungsstelle. Dissertationsschrift, Univ. Stuttgart (1980).
- /24/ Kühn,P.J. : Manuskript zur Vorlesung "Nachrichtenvermittlung I", Universität Stuttgart.
- /25/ Cotton,J.M., Giesken,K., Lawrence,A., Upp,D.C. : System 12 Digital-Koppelnetz. Elektrisches Nachrichtenwesen, 56 (1981), Nr.2/3, 148-160.
- /26/ Stallings,W. : Local Networks. Computing Surveys, Vol.16, No.1 (1984), 3-41.
- /27/ Göldner,E.-H., Kühn,P.J. : Integration Of Voice And Data In The Local Area. First International Conference on Data Communications in the ISDN Era, Tel-Aviv, 4-5. March, 1985.
- /28/ CCITT Yellow Book, Volume VIII Fascicle VIII.2 : Data communication networks; services and facilities, terminal equipment and interfaces. Recommendations X.1-X.29. Geneva (1981).
- /29/ CCITT Yellow Book, Volume VIII Fascicle VIII.3 : Data communication networks; transmission, signalling and switching, network aspects, maintenance, administrative arrangements. Recommendations X.40-X.180. Geneva (1981).
- /30/ CCITT Yellow Book, Volume VI Fascicle VI.6 : Specifications of signalling system No. 7. Recommendations Q.701-Q.741. Geneva (1981).
- /31/ Aschrafi,B., Meschkat,P., Szechenyi,K. : Field Trial Results of a Comparison of Time Separation, Echo Compensation and Four-Wire Transmission on Digital Subscriber Loops. ISSLS, Toronto, Canada (1982).
- /32/ CCITT Yellow Book, Volume VI Fascicle VI.7 : Functional Specification and Description Language (SDL). Recommendations Z.101-Z.104. Geneva (1981).
- /33/ Kühn,P.J. : Analyse zufallsabhängiger Prozesse in Systemen zur Nachrichtenvermittlung und Nachrichtenverarbeitung.

Habilitationsschrift, Universität Stuttgart (1981).

- /34/ Lemppenau,W., Tran-Gia,P. : A universal environment simulator for SPC switching system testing. 11. ITC, Kyoto, Japan (1985), Paper 5.1B-3.
- /35/ Huber,M., Wagner,W. : Simulation von Nachrichtenvermittlungssystemen. In: Nichtnumerische Informationsverarbeitung, Hsg. R. Gunzenhäuser, Springer-Verlag, Wien/New York (1968).
- /36/ Kampe,G., Kühn,P., Langenbach-Belz,M. : Simulation in der Nachrichtenverkehrstheorie: Problemstellungen und Programmiersprachen. GI-Workshop über Simulationstechniken, Kernforschungszentrum Karlsruhe, 1973. KFK-Ber. 1845 (1973), 240-263.
- /37/ Kleinrock,L. : Queueing Systems. Vol. I: Theory, Vol. II: Computer Applications. John Wiley & Sons, New York/London/Sydney/Toronto (1975/1976).
- /38/ Tran-Gia,P. : Überlastprobleme in rechnergesteuerten Fernsprechvermittlungssystemen - Modellbildung und Analyse. Dissertationsschrift, Univ.-Gesamthochschule Siegen (1982).
- /39/ Takagi,H., Kleinrock,L. : Analysis of Polling Systems. Japan Science Institute Research Report (1985).
- /40/ Tran-Gia,P., Raith,T. : Multiqueue Systems with Finite Capacity and Nonexhaustive Cyclic Service. International Seminar on Computer Networking and Performance Evaluation, Tokyo, Japan (1985).
- /41/ Kümmerle,K. : Ein Vorschlag zur Berechnung der Vertrauensintervalle bei Verkehrstests. AEU 23 (1969), Heft 10, 507-511.
- /42/ Ide,H.D., Sägebarth,J. : On Properties of Random Numbers and their Influence on Traffic Simulation. 10. ITC, Montreal, Canada (1983), Paper 2.4-6.
- /43/ Sauer,Kl. : Simulationsverfahren für verteilte Systeme ba-

- sierend auf einer formalen Spezifikation in SDL. Diplomarbeit Nr.724, Institut für Nachrichtenvermittlung und Datenverarbeitung, Universität Stuttgart (1984).
- /44/ Syski,R. : Markovian Queues. Proc. of the Symposium on Congestion Theory, University of North Carolina, 1964. The University of North Carolina Press (Ed. Smith,W.L., Wilkinson,W.E.), Chapel Hill (1964), 170-227.
- /45/ Cox,D.R. : A Use of Complex Probabilities in the Theory of Stochastic Processes. Proc. Cambr. Phil. Soc. 51 (1955), 313-319.
- /46/ Cooper,R.B. : Introduction to Queueing Theory. Macmillan Co., New York (1972).
- /47/ Dietrich,G., Salade,R. : Teilbelegungstreue Simulation der Steuerung von Vermittlungssystemen. El. Nachrichtenwesen 52 (1977), Nr. 1, 61-68.
- /48/ Stahl,U. : Einsatz eines Mikroprozessors als Vermittlungsrechner im Labormodell einer kombinierten Vermittlungsstelle. Diplomarbeit Nr. 681, Institut für Nachrichtenvermittlung und Datenverarbeitung, Universität Stuttgart (1982).
- /49/ Eisele,M. : Realisierung einer Vermittlungsmodul-Steuerung und deren Software für Internverkehr. Diplomarbeit Nr.682, Institut für Nachrichtenvermittlung und Datenverarbeitung, Universität Stuttgart (1983).
- /50/ Mönch,J. : Erstellung der Vermittlungssoftware für Externverbindungen im ausgelagerten Vermittlungsmodul. Diplomarbeit Nr.725, Institut für Nachrichtenvermittlung und Datenverarbeitung, Universität Stuttgart (1984).
- /51/ Schassberger,R. : Ein Wartesystem mit zwei parallelen Warteschlangen. Computing (1968) 1, 110-124.
- /52/ Császár,G.Y., Konkoly,R. : Optimal Allocation of Calls to Private Automatic Branch Exchange Operators. Budavox Telecommunication Review (1982) 4, 1-6.

ABKÜRZUNGEN

A	Verkehrs-Angebot
A/D	Analog/Digital-Wandler (Codec)
ADR	Adresse, Adressteil
A _{ges}	Gesamtes Verkehrs-Angebot eines Vermittlungssystems
A-Tln	Rufender Teilnehmer
AVM	Ausgelagerter Vermittlungsmodul
AWS	Ausgangswarteschlange
B , B _i	Blockierungswahrscheinlichkeit, Blockierungswahrscheinlichkeit des Teilsystems i
BE	Bedienungseinheit
BAi	Belegungsart i
BST	Bussteuerung
BST-POLL	Abfrageprozeß (Polling-Prozeß) der BST
BST-ZUT	Zuteilungsprozeß der BST
B-Tln	Gerufener Teilnehmer
BVi	Belegungsversuch i
c , c(i)	Belegungsart, Belegungsart der Meldung i
CCITT	<u>C</u> omité <u>C</u> onsultatif <u>I</u> nternational <u>T</u> élégraphique et <u>T</u> éléphonique
Codec	<u>C</u> oder/ <u>D</u> ecoder (A/D-D/A-Wandler)
D	Bezeichnung für Deterministischen Prozeß
DEE	<u>D</u> aten <u>e</u> nde <u>i</u> nrichtung
DTSIG	Datenteilnehmer-Signalisiereinheit
DÜE	<u>D</u> aten <u>ü</u> bertragung <u>e</u> inrichtung
DVP	Datenvermittlungsprozessor
E[...]	Erwartungswert von ...
E _k	Bezeichnung für Erlang-Prozeß der Ordnung k
Erl	Erlang (dimensionslose Einheit)
ε , ε _i	Bedienungsenderate, Bedienungsenderate der Bedienungseinheit i
EWS	Eingangswarteschlange
F _A (t)	Verteilungsfunktion der Ankunftsabstände T _A

Fe-KN	Fernsprech-Koppelnetz
F _H (t)	Verteilungsfunktion der Bedienungsdauern T _H
FIFO	<u>F</u> irst <u>I</u> n - <u>F</u> irst <u>O</u> ut (Speicherprinzip)
g	Anzahl von Warteschlangen
G	Gewichtungsfaktor für Signalflußanalyse (siehe 4.3.2)
G	Bezeichnung für allgemeinen Bedienungsprozeß
GI	Bezeichnung für allgemeinen unabhängigen Ankunftsprozeß
GVE	Gruppen-Verarbeitungseinrichtung
h , h(i)	Mittlere Bedienungsdauer, mittlere Bedienungsdauer für die Meldung i
HDLC	<u>H</u> igh <u>L</u> evel <u>D</u> ata <u>L</u> ink <u>C</u> ontrol
H _k	Bezeichnung für <u>H</u> yperexponential-Prozeß der Ordnung k
h _k	Mittlere Bedienungsdauer des Teilsystems des Typs k
h _ü	Mittlere Übertragungszeit
h _v	Mittlere Verarbeitungszeit
IDN	<u>I</u> ntegriertes <u>D</u> aten <u>n</u> etz
INF	Information, Informationsteil
ISDN	<u>I</u> ntegrated <u>S</u> ervices <u>D</u> igital <u>N</u> etwork
K	Kennzeichnung, Kennzeichnungsbit
KAS	Konzentrator-Anschlußschaltung
K-KN	Konzentrator-Koppelnetz
KN	Koppelnetz
KNST-D	Koppelnetzsteuerung für das Daten(UK)-KN
KNST-Fe	Koppelnetzsteuerung für das Fe-KN
KST	Konzentrator-Steuerung
λ , λ _i	Ankunftsrate, Ankunftsrate des Teilsystems i
λ _r	Rückkopplungsrate (siehe 5.3.1.4)
LAN	<u>L</u> ocal <u>A</u> rea <u>N</u> etwork
M	Bezeichnung für <u>M</u> arkoff'schen Prozeß
MB-KN	Multi-Bitraten-Koppelnetz
MODUL-E	Empfangsprozeß eines Moduls im System PILOT
MODUL-S	Sendeprozess eines Moduls im System PILOT
n	Anzahl von Bedienungseinheiten

NEE	<u>Netz</u> <u>ende</u> <u>in</u> <u>rich</u> <u>tung</u>
$n_{m/k}$	Anzahl der am Teilsystem k angeschlossenen Anschlußeinheiten des Typs m
n_k	Anzahl paralleler, gleichartiger Teilsysteme des Typs k
Ω	Mittlere Warteschlangenlänge
$P\{\dots\}$	Wahrscheinlichkeit für ...
$p_c, p_c(i)$	Auftrittsw ^{ir} keit der Belegungsart c, Auftrittsw ^{ir} keit der Meldung i der Belegungsart c
PCM	<u>P</u> ulse <u>C</u> ode <u>M</u> odulation
p_D	Wahrscheinlichkeit für eine Datenverbindung
PE , PEi	Periphere Einrichtung, Periphere Einrichtung für die Funktion i
p_{Fe}	Wahrscheinlichkeit für eine Fernsprechverbindung
p_{ext}	Wahrscheinlichkeit für eine Externverbindung
p_{ij}	Zustandsübergangswahrscheinlichkeit von Zustand i nach j
$p_{i,j}$	Verzweigungswahrscheinlichkeit für Übergang von Zustand i nach j im Teilruf-Ankunftsprozeß
PILOT	<u>P</u> artielle <u>I</u> ntegration im <u>l</u> okalen <u>T</u> elefonnetz
p_{int}	Wahrscheinlichkeit für eine Internverbindung
$p_i(\underline{X})$	Bedingte Zustandsübergangswahrscheinlichkeit für einen Übergang vom Zustand <u>X</u> in den Nachbarzustand mit x_{i+1} (siehe 5.3.1.3)
p_j	Zustandswahrscheinlichkeit des Zustands j
$p_k(i)$	Wahrscheinlichkeit, mit der die Meldung i in das Teilsystem k eintritt
PÜE	Periphere Übertragungseinrichtung
PVE	Periphere Verarbeitungseinrichtung
$p(\underline{X})$	Zustandswahrscheinlichkeit des Systemzustands <u>X</u>
$p(x_i)$	Zustandswahrscheinlichkeit des Systemzustands des Teilsystems i
q_{ij}	Zustandsübergangsrate von Zustand i nach j
q_j	Zustandsübergangsrate für das Verlassen des Zustands j
q_z	Wahrscheinlichkeit, daß eine Anforderung eine Warteschlange mit dem Füllgrad z vorfindet

s, s_i	Anzahl der Warteplätze, Anzahl der Warteplätze der Warteschlange i
SDL	(CCITT) <u>S</u> pecification and <u>D</u> escription <u>L</u> anguage
SPC	<u>S</u> tored <u>P</u> rogram <u>C</u> ontrolled
ST	Status, Statusinformationsteil
t	Zeit (absolut)
t	Definierter Schwellwert des Warteschlangen-Füllgrades
T_A	Ankunftsabstand
$t_{A,i}$	Ankunftszeitpunkt der Anforderung i
t_b	Mittlere Dauer einer Belegung
T_D	Zwischeneignisdauer eines Ausgangsprozesses
$t_{E,i}$	Bedienungsendezeitpunkt der Anforderung i
TG	Tongenerator
T_H	Bedienungsdauer
t_i	Zeitpunkt i
$T_{i,j}$	Verweildauer zwischen Ankunftsprozeß-Zustand i und j
TLN	Teilnehmer
TSIG	Teilnehmersignalisiereinheit
T_t	Bedienungsdauer für Anforderung des Typs t
t_w, t_{wi}	Mittlere Wartezeit bzgl. der wartenden Anforderungen, mittlere Wartezeit t_w in der Warteschlange i
T_W	Wartezeit
u	Zeitintervall
ÜE	Übertragungseinrichtung
UEG	<u>Ü</u> bertragungseinrichtung, <u>g</u> ehend
UEK	<u>Ü</u> bertragungseinrichtung, <u>k</u> ommend
UK-KN	Unterkanal-Koppelnetz
UK-MUX	Unterkanal-Multiplexer
VE , VEi	Verarbeitungseinrichtung, Verarbeitungseinrichtung für die Funktion i
VMST	Vermittlungsmodul-Steuerung
w	Mittlere Wartezeit bzgl. aller Anforderungen
W	Wartewahrscheinlichkeit
WZE	Wählzeichenempfänger

\underline{X}	Systemzustandsvektor (mehrdimensionaler Systemzustand)
x_{BE}	Zustand der Bedienungseinheit (frei oder belegt)
x_i	Systemzustand des Teilsystems i
x_j	Systemzustand zum Zeitpunkt t_j
$X(t)$	Systemzustandsfunktion
Y, Y_i	Mittlere Belastung (Verkehrswert), mittlere Belastung der Bedienungseinheit i
Y_k	Mittlere Belastung des Teilsystems k
$Y_{k,j}$	Mittlere Teilbelastung des Teilsystems k, verursacht durch die Belegungsart j
ZST	Zentrale Steuerung
ZUE	Zentrale Übertragungseinrichtung
ZVE	Zentrale Verarbeitungseinrichtung
ZVP	Zentraler Vermittlungsprozessor
ZZK	<u>Zentraler Zeichenkanal</u> (nach CCITT No. 7)
ZZK	Steuerung für zentralen Zeichenkanal

1 EINLEITUNG

1.1 Rechnergesteuerte Vermittlungssysteme

Die Ära der rechnergesteuerten Vermittlungstechnik begann im Jahre 1965 mit der Einschaltung der ersten serienmäßig gefertigten, rechnergesteuerten Fernsprech-Ortsvermittlungsstelle des Typs No. 1 ESS in New Jersey, USA /1/. Rechnergesteuerte Vermittlungssysteme mit gespeichertem Programm, sogenannte SPC-Systeme (engl. stored program controlled), konnten gegenüber herkömmlichen Systemen sowohl dem Benutzer als auch dem Betreiber gewisse Vorteile bieten /2/:

Für den Benutzer wurden neue Leistungsmerkmale bzw. Dienste durch SPC-Systeme wirtschaftlich realisierbar, wie z.B.:

- Kurzwahl
- Anrufumleitung
- Ruhe vor dem Telefon
- Anklopfen
- Warten auf Freiwerden des gerufenen Teilnehmers
- Gebührenübernahme durch den gerufenen Teilnehmer
- Sperre für abgehende Ferngespräche
- Fangen.

Einige Vorteile, die sich dem Betreiber eines SPC-Systems zeigen, sind z.B.:

- Platzeinsparung
- Rationalisierung und Automatisierung von Betrieb und Wartung
- Automatische Gebührenlegung
- Automatische Verkehrsmessung
- Flexibilität bezüglich Erweiterungen und Änderungen.

Bezüglich der Steuerung war die erste Generation von SPC-Vermittlungssystemen durch analoge Durchschaltung im Sprechweg und eine zentralisierte Steuerung gekennzeichnet, wie schon im System No. 1 ESS /3/ oder im ersten in Deutschland im Jahre 1976 in Be-

trieb genommenen SPC-Ortsvermittlungssystem EWSA /4,5/. Der damalige Stand der Halbleiter- und Rechnertechnologie erforderte aus Wirtschaftlichkeitsgründen elektromechanische Kontakte für die Durchschaltung und eine Zentralisierung aller nicht während einer ganzen Verbindung benötigten, intelligenten Funktionen in einen relativ großen, nur sicherheitshalber duplizierten, zentralen Vermittlungsrechner. Dazu war wegen des intensiven Steuerdatenaustausches mit der umfangreichen Peripherie ein aufwendiges, leistungsfähiges Ein/Ausgabewerk notwendig /3,5/.

Inzwischen hat sich die Konzeption von SPC-Vermittlungssystemen durch technologische Fortschritte in einigen wesentlichen Punkten stark verändert, nämlich durch:

- volle Digitalisierung
- Dezentralisierung bzw. Verteilung der Steuerung
- Integration mehrerer Dienste.

1.1.1 Digitalisierung

Durch die schnelle technologische Entwicklung der Mikroelektronik stehen heute hochintegrierte, intelligente Schaltkreise preiswert zur Verfügung. Dadurch wird die volle Digitalisierung, d.h. der Übergang von der analogen Durchschaltung zur störunempfindlichen digitalen Durchschaltung in PCM-Technik (engl. pulse code modulation) wirtschaftlich möglich /6,7/. Gleichzeitig wird damit die Voraussetzung für die Integration von Datenverkehr, d.h. die Durchschaltung von Datenbitströmen, geschaffen. Ebenso wird eine Vereinheitlichung mit der Übertragungstechnik erreicht, wo sich die PCM-Technik schon seit längerem bewährt.

1.1.2 Verteilte Steuerung

Im Bereich der Steuerung wirkt sich die Bereitstellung preiswerter, leistungsfähiger Mikroprozessoren so aus, daß heute eine Verlagerung zunächst zentralisierter Steuerfunktionen wieder ganz oder teilweise in den peripheren Bereich wirtschaftlich möglich

wird. Einzelne Steuerungen können im zentralen wie auch im peripheren Bereich durch mehrere kleinere Teilsteuerungen ersetzt werden, die nach dem Lastteilungsprinzip (engl. load sharing) oder auch nach dem Funktionsteilungsprinzip (engl. function sharing) arbeiten. Dies führt dann zu SPC-Systemen mit verteilter Steuerung, d.h. einer Vielzahl von Mikroprozessoren und einem umfangreichen, leistungsfähigen Kommunikationssystem für den Steuerdatenaustausch zwischen den Prozessoren /8/. Diese Entwicklung wird durch die neuesten SPC-Systeme, wie z.B. EWSD /9,10/, System 12 /11,12/, No. 5 ESS /13,14/ oder auch SPC-Nebenstellenanlagen, wie z.B. T&N 8030 /15/ dokumentiert.

Solche verteilten SPC-Systeme erlauben einen besseren modularen Aufbau der Hard- und Software. Durch definierte Funktionssteuereinheiten erreicht man eine höhere Flexibilität gegenüber technologischen Fortschritten als bei den zentral gesteuerten Systemen. Ebenso wird die Zuverlässigkeit erhöht, da Teilausfälle eine geringere Wirkungsbreite haben und somit eine teure Duplizierung großer, zentraler Baugruppen meist entfallen kann. Ein Systemausbau kann einfach durch Hinzufügen weiterer Moduln erfolgen, und kleinste wie große Systeme können im wesentlichen mit denselben Moduln realisiert werden.

1.1.3 Dienste-Integration

Parallel mit dem technologischen Fortschritt wuchs bei den Benutzern das Bedürfnis nach neuen Kommunikationsdiensten, wie z.B. Text- und Datenübertragung, Faksimile, Fernwirken usw., die durch die Digitalisierung und die Rechnersteuerung der Vermittlungssysteme wirtschaftlich integrierbar werden. Die Dienste-Integration im digitalen Fernsprechnet, dem zukünftigen ISDN (engl. integrated services digital network) /16,17/, könnte weitere dienstspezifische Netze mit eigenen speziellen Vermittlungsstellen, Übertragungseinrichtungen und Leitungsbündeln ersparen und somit wirtschaftliche Vorteile mit sich bringen. Die Steuerungen der zukünftigen ISDN-Vermittlungsstellen werden jedoch durch neue Funktionen, aufwendige Protokolle und zusätzliche Last beeinflusst.

1.2 Grundsätzliche Aufgabenstellungen

Bei der Planung eines neuen SPC-Vermittlungssystems mit verteilter bzw. modularer Steuerung ergeben sich bezüglich der Steuerung eine Reihe neuer Fragestellungen, die bei den zentral gesteuerten Systemen nicht relevant waren.

So ist die Frage nach der optimalen Steuerungsstruktur von großer Bedeutung. Sie hängt im wesentlichen davon ab, wie die Funktionen im System auf separate Steuereinheiten verteilt werden und welche Steuereinheiten nach dem Prinzip der Lastteilung in mehrere Teilsteuern aufgeteilt werden. Je nach Lösung dieser Aufgabenstellung ergeben sich dann kaum mehr zentrale, sondern vielmehr teilzentrale oder dezentrale Steuerungsstrukturen. Für den Steuerdatenaustausch zwischen den einzelnen Teilsteuern ist dann ein leistungsfähiges Verbindungssystem vorzusehen, dessen Realisierung großen Einfluß auf die Leistungsfähigkeit des Gesamtsystems hat.

Letztendlich muß die resultierende Lösung für das Gesamtkonzept durch die Leistungsfähigkeit, die Verfügbarkeit und durch Wirtschaftlichkeit begründet sein. Aus diesem Grund sind Methoden für die Leistungsbeurteilung von solch umfangreichen verteilten Systemen und von Teilsystemen unbedingt notwendig. Dabei müssen die Umwelt des Systems, d.h. die Teilnehmer und deren stochastische Verhaltensweise und auch die eigentliche Aufgabe des Systems, d.h. die Abwicklung des Vermittlungsvorgangs, mit berücksichtigt werden. Denn darin liegt die teilweise unterschiedliche Struktur im Vergleich zu anderen verteilten Rechensystemen begründet. Ferner sind vergleichende analytische Untersuchungen von Teilproblemen für die Wahl der optimalen Realisierung und die Dimensionierung von Parametern nötig.

1.3 Übersicht über die Arbeit

Im Anschluß an die Einleitung wird im zweiten Kapitel die Konzeption verteilter SPC-Vermittlungssteuerungen allgemein behandelt und nach einer Übersicht über die wichtigsten Steuerungsfunktionen auf prinzipielle Lösungen für die Steuerungsstruktur, die Probleme der Funktions- bzw. Lastverteilung und die Auswirkungen auf das Verbindungssystem zwischen den Steuereinheiten eingegangen.

Das dritte Kapitel beschreibt dann eine Realisierungsmöglichkeit am konkreten Beispiel der Laborimplementierung PILOT. Es handelt sich dabei um ein Dienste-integriertes SPC-Vermittlungssystem mit verteilter Steuerung. Das Konzept der Dienste-Integration wurde bereits in /18/ behandelt, während die vorliegende Arbeit den Schwerpunkt auf das verteilte Steuerungskonzept und die Interprocessor-Kommunikation dieses Systems legt.

Im vierten Kapitel werden einige für Vermittlungssteuerungen in Frage kommende Analysetechniken und deren Ziele prinzipiell aufbereitet und im Hinblick auf die im fünften Kapitel zu behandelnden Probleme erweitert. Dies sind Signalflußanalysen, zeitreue Simulationen und stationäre Warteschlangenanalysen Markoffscher Modelle und Teilmodelle.

Das fünfte Kapitel befaßt sich dann schließlich mit konkreten Einzeluntersuchungen. Diese sollen einerseits als Anwendungsbeispiele für die jeweilige Analysemethode dienen, andererseits können aus den Ergebnissen allgemeingültige Aussagen für die Konzeption von Vermittlungssteuerungen und von Teilproblemen gefolgert werden. Es handelt sich um Strukturuntersuchungen mittels der Signalflußanalyse, um eine zeitreue Simulationsstudie am Gesamtmodell des Systems PILOT und um analytische Untersuchungen von Lastzuteilungs- und Abfertigungsstrategien an Teilmodellen.

2 KONZEPTION VON VERMITTLUNGSSTEUERUNGEN

2.1 Grundsätzliche Steuerungsfunktionen

Die Grundaufgabe eines Vermittlungssystems ist die Vermittlung von Verbindungen, d.h. der Aufbau, das Halten und der Abbau von zeitweiligen Verbindungen zwischen einem rufenden Teilnehmer (A-Tln) und einem gerufenen Teilnehmer (B-Tln), wobei letzterer durch eine vom A-Tln dem Vermittlungssystem mitgeteilte Zielinformation bestimmt ist. Diese Vermittlungsaufgabe muß simultan für sehr viele gleichzeitig bestehende, bzw. im Auf- oder Abbau befindliche, unabhängige Verbindungen durchgeführt werden. Man hat es deshalb mit Multiprogramming-Betrieb mit einem sehr hohen Multiprogramming-Grad zu tun.

Der einzelne Vermittlungsvorgang kann als Prozeß im Sinne eines endlichen Automaten verstanden werden. Jedes zufällig ankommende Teilnehmerereignis stößt den Ablauf einer vom momentanen Zustand des betreffenden Vermittlungsprozesses abhängigen Vermittlungsfunktion an, die nach Ausführung bestimmter Aufgaben den Vermittlungsprozeß in einen neuen Zustand versetzt. Dies entspricht einem Zustandsübergang des endlichen Automaten. Da die Ausführung der Vermittlungsfunktionen eine gewisse Zeit erfordert und dafür bestimmte Zeitgrenzen eingehalten werden müssen, muß man auch von Realzeit-Betrieb sprechen. Im neuen Zustand wird dann wieder auf weitere Teilnehmerereignisse für den betreffenden Vermittlungsprozeß gewartet, d.h. inzwischen werden eventuell die quasi parallel laufenden Vermittlungsprozesse anderer Verbindungen bearbeitet, aber auch Sicherungsfunktionen oder betriebs- und verwaltungstechnische Funktionen. Alle diese Aufgaben werden schließlich durch Betriebssystem-Funktionen koordiniert und unterstützt. /19,20,21/

2.1.1 Vermittlungsfunktionen

Der Vermittlungsprozeß wird durch die Ausführungsfolge ("Szenario") einzelner Vermittlungs-Teilfunktionen beschrieben. Dabei muß man vier Belegungsarten, d.h. Typen von Verbindungen unterscheiden:

- Internverbindung zwischen zwei Teilnehmern in der betrachteten Vermittlungsstelle
- Abgehende Externverbindung zwischen einem A-Tln in der betrachteten und einem B-Tln in einer anderen Vermittlungsstelle
- Ankommende Externverbindung zwischen einem A-Tln in einer anderen und einem B-Tln in der betrachteten Vermittlungsstelle
- Transitverbindung zwischen zwei Teilnehmern in einer anderen Vermittlungsstelle über die betrachtete Vermittlungsstelle.

In jedem Falle müssen noch der erfolgreiche und die verschiedenen Möglichkeiten des erfolglosen Belegungsversuches und eventuell die Möglichkeiten für die Inanspruchnahme spezieller Leistungsmerkmale unterschieden werden. Darüberhinaus muß noch eine Fülle neuer Merkmale im Zusammenhang mit der Dienste-Integration betrachtet werden. Die Funktionsszenarien für alle diese Fälle bilden insgesamt sehr umfangreiche Listen. Nachstehend ist deshalb nur ein Fall am Beispiel des Funktionsszenarios einer erfolgreichen Internverbindung aufgeführt; er repräsentiert gleichzeitig die klassischen Vermittlungsfunktionen, die auch in der Mehrzahl aller Verbindungen benötigt werden:

- Erkennen des Verbindungswunsches des A-Tln's (A-Tln hebt ab) durch Teilnehmerzustandsüberwachung
- Identifizierung des A-Tln's (Anschlußlage bestimmen, Berechtigung prüfen)
- Bereitstellung einer Einrichtung zum Empfang und zur Verarbeitung von Wählinformationen für den A-Tln (z.B. Wahlaufnahmesatz anschalten)
- Wahlaufforderung an den A-Tln (z.B. Wählton anlegen)

- Empfang und Verarbeitung von Wählinformation (Wahlzeichen empfangen)
- Nach Wahlbeginn: Wahlaufforderung zurücknehmen (z.B. Wählton abschalten)
- Nach Wählende: Freigabe der Einrichtung zum Empfang von Wahlinformationen (z.B. Wahlaufnahmesatz abschalten)
- Wegesuche und Durchschaltung einer Verbindung vom A-Tln zum B-Tln im Koppelnetz der Vermittlungsstelle
- Freimeldung an den A-Tln (z.B. Freiton anlegen)
- Rufanzeige zum B-Tln (z.B. Rufstrom anlegen)
- Erkennen, daß sich der B-Tln meldet (B-Tln hebt ab) durch Teilnehmerzustandsüberwachung
- Rufanzeige beim B-Tln zurücknehmen (z.B. Rufstrom abschalten)
- Meldeanzeige zum A-Tln (z.B. Freiton abschalten)
- Gebührenerfassung starten
- Überwachung der Verbindung durch Teilnehmerzustandsüberwachung beider Teilnehmer
- Erkennen der Auslösung eines Teilnehmers (A- bzw. B-Tln legt auf)
- Gebührenerfassung stoppen
- Auslösen der Verbindung.

Für Extern- oder Transitverbindungen werden zusätzliche Vermittlungsfunktionen benötigt, vor allem Signalisierungsfunktionen für den Steuerdatenaustausch mit anderen Vermittlungsstellen. Für abgehende Externverbindungen werden z.B. gebraucht:

- Belegen einer Leitung zur anderen Vermittlungsstelle
- Signalisierung eines Verbindungswunsches zur anderen Vermittlungsstelle
- Weiterleitung von in der Ursprungs-Vermittlungsstelle nicht verwendeter Wählinformation als Zielinformation für eine andere Vermittlungsstelle.

Für neue Leistungsmerkmale, z.B. Kurzrufnummern, oder neue Dienste, z.B. Datendienste, werden gegebenenfalls zusätzliche Vermittlungsfunktionen erforderlich.

2.1.2 Sicherungsfunktionen

Zur Aufrechterhaltung der Betriebsbereitschaft sind neben den Vermittlungsfunktionen eine Reihe sicherungstechnischer Funktionen vorzusehen. Diese sind von größter Wichtigkeit, da für Vermittlungssysteme eine Verfügbarkeit von 99,992 % oder 2 Stunden Totalausfall in 30 Jahren gefordert wird. Zum Teil sind sie in die Vermittlungsprozesse integriert, wie z.B. die Zeitüberwachung aller Zustände in den Vermittlungsprozessen, damit Betriebsmittel nicht durch falsches Teilnehmerverhalten oder sonstige Fehler im System unnötigerweise lange blockiert werden. Andere Sicherungsfunktionen laufen unabhängig von den Vermittlungsprozessen entweder nach Bedarf oder in regelmäßigen Zeitabständen oder nur, wenn keine Vermittlungsfunktionen anstehen. Dies sind z.B.:

- Routineprüfungen und Messungen für Teilsysteme
- Diagnoseprüfungen zur Lokalisierung von Fehlern im System
- Funktionen zur Beseitigung von Fehlern (z.B. durch Sperrung oder Ersatzschaltung fehlerhafter Teilsysteme)
- Funktionen für den Wiederanlauf abgeschalteter Teilsysteme.

2.1.3 Betriebs- und verwaltungstechnische Funktionen

Zur Unterstützung des Betriebs eines Vermittlungssystems müssen dem Betreiber weitere Verwaltungsfunktionen bereitgestellt werden. Dies sind z.B. Funktionen der Mensch-Maschine-Kommunikation für:

- Verwaltung von Teilnehmern (z.B. Einrichtung, Berechtigungen)
- Verwaltung von Leitungsbündeln (z.B. Einrichtung, Änderung der Verkehrslenkung)
- Verwaltung von vermittlungstechnischen Baugruppen (z.B. Einrichtung, Parameteränderung)
- Verwaltung der Gebührenerfassung (z.B. automatische Rechnungserstellung, Änderung der Tarife)
- Messung und Protokollierung von Verkehr und Dienstgüte
- Entwicklungs-Unterstützung (z.B. Planung, Dimensionierung).

2.1.4 Betriebssystem-Funktionen

Das Betriebssystem eines SPC-Vermittlungssystems steuert den Ablauf aller vermittlungs-, sicherungs- und betriebstechnischen Prozesse. Ferner stellt es den Prozessen einige grundlegende Funktionen zur Verfügung, wie z.B.:

- Zeitverwaltungsfunktionen
- Speicherverwaltungsfunktionen (z.B. in Form einer Datenbank)
- Ein/Ausgabe-Funktionen für die vermittlungs- und betriebstechnische Peripherie
- Funktionen für den Meldungs austausch zwischen verteilten Steuereinheiten (Einreihung in Warteschlangen usw.)
- Funktionen für die Überlastabwehr
- Lade- und Startfunktionen für Programme.

2.2 Grundsätzliche Steuerungsstrukturen

2.2.1 Baugruppen von Vermittlungssteuerungen

Neben dem Koppelnetz (KN) als Grundelement lassen sich die Steuerungseinrichtungen eines Vermittlungssystems in drei Gruppen einteilen /22/ :

- Einrichtungen zum Senden/Empfangen von Steuerinformationen
- Einrichtungen zur Übertragung von Steuerinformationen
- Einrichtungen zur Verarbeitung von Steuerinformationen.

2.2.1.1 Einrichtungen zum Senden/Empfangen von Steuerinformationen

Dieser Gruppe von Einrichtungen können periphere Einrichtungen (PE) zur Signalisierung zugeordnet werden, die sich um das Koppelnetz gruppieren. Typische Beispiele sind:

- Teilnehmersätze als Schnittstelle zu Teilnehmerleitungen
- Identifizierer zur Ermittlung der Anschlußlage von Teilnehmern
- Wahlaufnahmesätze zum Empfang von Wählinformationen
- Verbindungssätze zur Überwachung des Verbindungszustandes
- Tongeneratoren zur Erzeugung von Hör tönen
- Leitungssätze als Schnittstelle zu Verbindungsleitungen
- Wahl nachsendesätze für abgehende Verbindungen
- Einstellsätze zur Einstellung von Verbindungswegen in Koppelnetzen.

In modularen SPC-Vermittlungssystemen werden oft mehrere derartige Funktionen, z.B. zusammengefasst für einige Teilnehmer- oder Verbindungsleitungen, durch periphere Anschlußmoduln oder Signalisierprozessoren realisiert, die dann allerdings zum Teil auch Aufgaben der Verarbeitung bzw. Vorverarbeitung von Steuerinformationen übernehmen.

2.2.1.2 Einrichtungen zur Übertragung von Steuerinformationen

Zu dieser Gruppe gehören Leitungssysteme, Zugriffssysteme und Ein/Ausgabewerke für den Austausch von Steuerinformationen zwischen Einrichtungen zum Senden/Empfangen und Einrichtungen zur Verarbeitung von Steuerinformationen, aber auch für die Kommunikation in verteilten Systemen zwischen den Verarbeitungseinrichtungen /8/. Typische Realisierungen für solche Übertragungseinrichtungen (ÜE) zur Intermodul-Kommunikation in SPC-Vermittlungssystemen sind:

- individuelle Steuerkanäle (voll oder teilweise vermascht)
- parallele oder serielle Bus- oder Ringsysteme
- Speicherkopplung über Mehrfachzugriff-Speicher
- Steuerkanäle über Koppelnetz (separates oder das ohnehin vorhandene Koppelnetz) durch festgeschaltete oder vermittelte Wege.

2.2.1.3 Einrichtungen zur Verarbeitung von Steuerinformationen

Die Verarbeitung von Steuerinformationen geschieht in SPC-Vermittlungssystemen fast ausschließlich in Prozessoren. Dabei soll unter Prozessor ein komplettes Rechnersystem, inklusive Programm- und Datenspeicher, verstanden werden. Der Vermittlungsablauf wird in jedem Fall von Einfach- oder Multiprozessoren koordiniert. Für sonstige Steuerungsfunktionen, insbesondere periphere Steuerungsfunktionen, werden heute zunehmend statt Spezialsteuerwerken Mikroprozessoren eingesetzt. Durch Funktionsteilung bzw. Dezentralisierung von Steuerungsfunktionen lassen sich ferner für gewisse Aufgaben spezialisierte und damit optimierte Rechner vorteilhaft verwenden. Typische Beispiele für Verarbeitungseinheiten (VE) in SPC-Vermittlungssystemen sind:

- Koordinierende Vermittlungsrechner
- Peripherieprozessoren für Anschlußmoduln oder betriebstechnische Geräte
- Signalisierprozessoren für komplizierte Signalisierprotokolle

(z.B. Zentralkanalsignalisierung nach CCITT No. 7 oder ISDN-Teilnehmersignalisierung nach D-Kanal-Protokoll)

- Wegesuchprozessor für autonome Wegesuche im Koppelnetz
- Datenbankrechner für die Verwaltung der Zustandsdaten-Speicherung
- Prozessorgesteuerte Buszugriffssteuerung.

2.2.2 Struktureller Aufbau von Vermittlungssteuerungen

Die Steuerungsstruktur eines SPC-Vermittlungssystems wird hauptsächlich durch den Zentralisierungsgrad der Steuerung und den Grad der Funktions- bzw. Lastteilung bei den Verarbeitungseinrichtungen bestimmt. Die folgenden prinzipiellen Strukturen sind nach dem Zentralisierungsgrad gegliedert und basieren auf den im vorigen Abschnitt vorgestellten Grundelementen:

- | | |
|---------------------------------------|-----|
| - Koppelnetz bzw. Koppelnetzmodul | KN |
| - Periphere Einrichtung, allgemein | PE |
| für eine Funktion i | PEi |
| - Übertragungseinrichtung | ÜE |
| - Verarbeitungseinrichtung, allgemein | VE |
| für eine Funktion i | VEi |

Die jeweilige Realisierungsform dieser Grundelemente ist dabei ohne Belang, und als Verbindungswege werden ausschließlich Steuerdatenwege betrachtet.

2.2.2.1 Zentrale Steuerung

Bild 2.1 zeigt die prinzipielle Steuerungsstruktur eines zentral-gesteuerten SPC-Vermittlungssystems.

Die zahlreichen vermittlungs- und betriebstechnischen peripheren Einrichtungen PE1 ... PE_n und das meist einzige große Koppelnetz KN sind über eine Übertragungstechnische Einrichtung ÜE, z.B. ein Ein/Ausgabewerk, mit der meist konzentrierten, zentralen Verar-

beitungseinrichtung VE, dem Vermittlungsrechner, gekoppelt. Alle Verarbeitungsfunktionen sind in der VE zentralisiert, und deshalb erfolgt auch die Speicherung aller Programme und aller Zustandsdaten der Teilnehmer, Leitungen und Betriebsmittel an zentraler Stelle, im Speicher der VE.

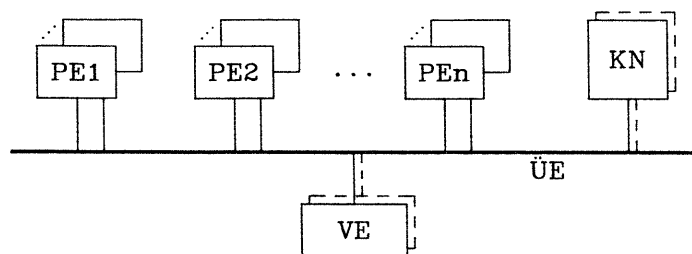


Bild 2.1: Zentrale Steuerungsstruktur

Ein Nachteil der zentralgesteuerten Struktur ist, daß zumindest die VE leistungsmäßig für den möglichen Endausbau konzipiert sein muß oder für verschiedene Ausbaustufen unterschiedliche Versionen der VE bereitstehen müssen. Ferner sollten bei konzentrierter Zentralsteuerung zur Sicherstellung einer hohen Verfügbarkeit die VE, die ÜE und das KN dupliziert sein und geeignete Ersatzschaltungsstrategien /23/ vorgesehen werden. Dabei ist jedoch typisch, daß die aus Sicherheitsgründen duplizierten Einrichtungen nie im Funktions- und nur selten im Lastteilungsbetrieb arbeiten, sondern die zweite Einrichtung dient lediglich als reine Reserve. Die Duplizierung wirkt somit nicht leistungssteigernd, sondern nur sicherheitssteigernd.

Die zentrale Strukturform fand Anwendung in der ersten Generation von SPC-Vermittlungssystemen und wird in den zukünftigen Systemen kaum mehr eine Rolle spielen. Lediglich für kleine Systeme, z.B. kleine Nebenstellenanlagen, könnte die rein zentrale Steuerung noch einen sinnvollen Einsatz finden.

2.2.2.2 Dezentrale Steuerung

Zwei prinzipielle Strukturen mit rein dezentraler Steuerung sind in den Bildern 2.2 und 2.3 dargestellt. Sie unterscheiden sich im Grad der Funktionsteilung. Es sind aber auch Mischformen aus diesen beiden Strukturen denkbar.

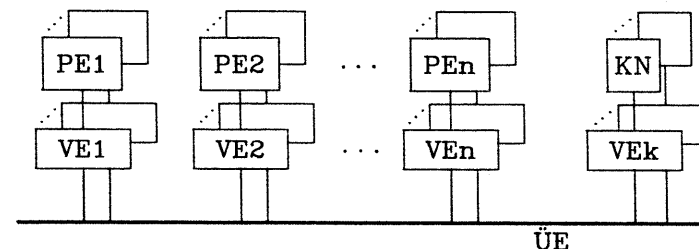


Bild 2.2: Dezentrale Steuerungsstruktur mit Funktionsteilung

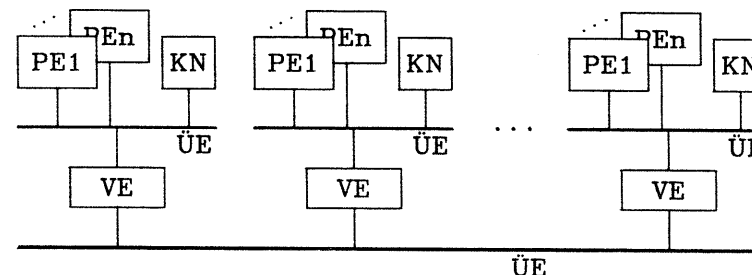


Bild 2.3: Dezentrale Steuerungsstruktur ohne Funktionsteilung

In der dezentralsten Form (siehe Bild 2.2) besitzen alle peripheren Einrichtungen PE_i und eventuell alle Koppelnetzmoduln KN des meist modular aufgebauten Koppelnetzes ihre eigene Steuerung VE_i. Diese übernimmt die autonome Steuerung und Zustandsspeicherung der zugehörigen PE bzw. des KN-Moduls und der damit in Verbindung stehenden Teilnehmer, Leitungen oder Betriebsmittel. Man hat es

deshalb neben der Lastteilung mit ausgeprägter Funktionsteilung im Steuerungsbereich zu tun. Die KN-Moduln könnten jedoch auch jeweils einer PE zugeordnet sein mit einer gemeinsamen VE als Steuerung.

Ferner wäre es möglich (siehe Bild 2.3), daß jeweils eine der wichtigsten für eine Verbindung notwendigen PE, z.B. je ein Modul PE1, PE2, ..., PEn und eventuell ein KN-Modul, über ein kleines peripheres Zugriffssystem ÜE von einer gemeinsamen VE gesteuert werden. In diesem Fall findet somit das Funktionsteilungsprinzip keine Anwendung.

In beiden Fällen kommunizieren die dezentralen VE untereinander über eine Übertragungseinrichtung ÜE, z.B. einen Bus oder das Koppelnetz selbst.

Bei der dezentralen Strukturform erhält man ohne Duplizierung von Einrichtungen eine hohe Zuverlässigkeit, da ein Teilausfall nur eine geringe Wirkungsbreite hat. Lediglich die ÜE müßte eventuell, z.B. bei einem Bus, aus Sicherheitsgründen dupliziert werden. Dies wäre jedoch bei einem Steuerdatenaustausch über ein modulares Koppelnetz nicht notwendig. Schließlich ist durch die hohe Modularisierung größte Flexibilität bezüglich der Erweiterbarkeit und der Anpassbarkeit an neue Technologien gegeben.

Die dezentrale Art der Steuerung ist prinzipiell bereits aus der konventionellen, direktgesteuerten Vermittlungstechnik /24/ bekannt. Dort besaß jeder Wähler bzw. jeder Verbindungssatz seine eigene Steuerung und der Steuerdatenaustausch, die Signalisierung, erfolgte im Sprechkreis selbst, also über das Koppelnetz. Trotzdem sind unter den modernen SPC-Vermittlungssystemen zur Zeit noch keine rein dezentrale Steuerungsstrukturen bekannt, da deren Realisierung prinzipielle Schwierigkeiten bereitet. Die dezentralisierten Teilsteuerwerke benötigen nämlich eine sehr hohe Intelligenz und Speicherfähigkeit. Dieses Problem könnte jedoch in Zukunft durch preiswerte höchstintegrierte Schaltungen behoben werden. Einen teilweisen Schritt in diese Richtung stellt zum ei-

nen z.B. die modulare, dezentrale Koppelnetzsteuerung des Systems 12 /25/ dar, die der konventionellen, direktgesteuerten Technik sehr ähnlich ist. Im Bereich der Datenkommunikation findet man zum andern auch bei einigen lokalen Netzwerken, kurz LANs (engl. local area network /26/), eine dezentrale Steuerung. Dort sind die Probleme jedoch durch die relativ geringe Teilnehmerzahl besser beherrschbar. Im Bereich der LANs wird zur Zeit aber auch die Integration von Sprache diskutiert /27/, sowohl auf synchrone Art als auch in paketierter Form. Auf diese Weise könnten neue Vermittlungssysteme auf LAN-Basis mit dezentraler Steuerung entstehen, zumindest kleinere Nebenstellenanlagen.

2.2.2.3 Teilzentrale Steuerung

Die beiden vorangehenden Abschnitte zeigten, daß die beiden Grenzfälle des Zentralisierungsgrades gravierende Nachteile aufweisen. Eine hybride, teilzentrale Steuerungsstruktur vereint dagegen die meisten Vorteile beider Grenzfälle, ohne deren Nachteile zu besitzen, und sie stellt aus heutiger Sicht die optimale Struktur dar. Deshalb haben fast alle modernen SPC-Systeme diese Struktur. Die Bilder 2.4 und 2.5 zeigen zwei Möglichkeiten für teilzentrale Steuerungen, die sich im dezentralen, peripheren Bereich analog zu den vorgestellten rein dezentralen Strukturen im Grad der Funktionsteilung unterscheiden. Auch hier sind natürlich Mischformen möglich.

Im zentralen Bereich kann sowohl das Last- als auch das Funktionsteilungsprinzip mehr oder weniger Anwendung finden. Wie in beiden Bildern angedeutet, können für beispielsweise m zentrale Funktionen z1 ... zm entsprechende m Steuerungen VEz1 ... VEzm existieren, die ihrerseits wieder nach dem Lastteilungsprinzip in mehrere gleichartige Teilsteuerungen aufgeteilt sein können. Der dezentrale, periphere Bereich kann sich so gestalten wie bei den rein dezentralen Strukturen erläutert.

Die verschiedenen VE kommunizieren untereinander über eine ÜE, in der Regel ein Bus oder das Koppelnetz selbst.

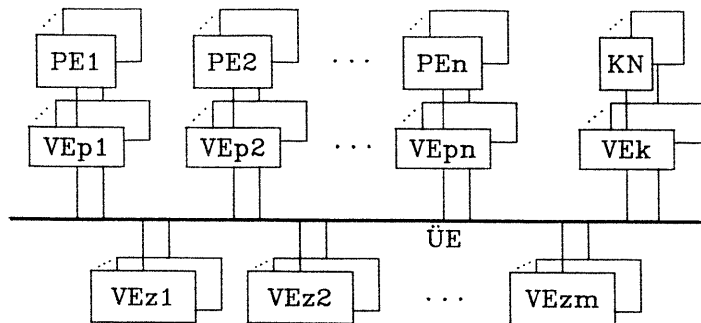


Bild 2.4: Teilzentrale Steuerungsstruktur mit Funktionsteilung im peripheren Bereich

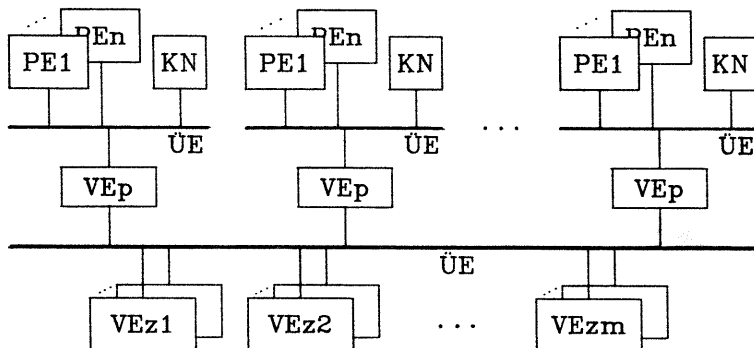


Bild 2.5: Teilzentrale Steuerungsstruktur ohne Funktionsteilung im peripheren Bereich

Bedingt durch das Prinzip der teilzentralen Steuerung hat man es grundsätzlich mit hierarchischen Systemen zu tun. Die in den Bildern 2.4 und 2.5 gezeigten Varianten besitzen eine zweistufige Struktur. Verleiht man jedoch den PE in Bild 2.5 durch eine eigene VE mehr Intelligenz, so erreicht man eine dreistufige Hierar-

chie wie in Bild 2.6 gezeigt. Mehr als dreistufige Systeme sind zwar theoretisch denkbar, finden jedoch kaum praktischen Einsatz, da die Komplexität mit der Anzahl von Stufen zunimmt.

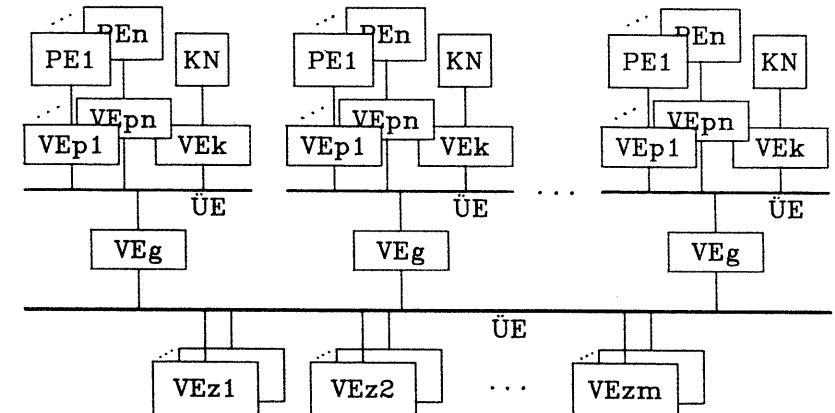


Bild 2.6: Teilzentrale hierarchische Steuerungsstruktur mit drei Stufen

Der zwei- und dreistufigen Strukturform entsprechen die meisten modernen SPC-Vermittlungssysteme. So lassen sich z.B. die Systeme EWSD, System 12, No. 5 ESS, das im dritten Kapitel vorgestellte System PILOT und auch die Nebenstellenanlage T&N 8030 in diese Kategorie einordnen. Beim System 12 ist die Last- und Funktions- teilung im zentralen Bereich am ausgeprägtesten, repräsentiert durch die sogenannten Funktions- und Systemsteuereinheiten. EWSD und No. 5 ESS besitzen in ihrer Grundform einen duplizierten Zentralprozessor. Für sehr große Systeme sind jedoch auch hier verteiltere Strukturen für den zentralen Bereich in Entwicklung.

2.3 Probleme der Funktions- und Lastteilung

In den vorigen Abschnitten wurden Steuerungsstrukturen prinzipiell betrachtet, ohne Rücksicht auf die speziellen Funktionen der einzelnen Verarbeitungseinheiten und deren Anzahl zu nehmen. Im folgenden sollen einige Fragestellungen erläutert werden, die sich aus der Verteilung der Steuerung nach Funktions- und Lastgesichtspunkten ergeben:

- Wie sind die Funktionen auf separate Funktionssteuereinheiten zu verteilen?
- Für welche Funktionen ist zusätzliche Lastteilung vorzusehen? Wieviele Teilsteuerungen braucht man gegebenenfalls? Nach welcher Strategie erfolgt dann die Lastzuteilung?
- Wo erfolgt die Datenspeicherung?
- Wie groß ist das Ausmaß der Interprozessor-Kommunikation? Wie haben demzufolge die Übertragungseinrichtungen auszusehen?

2.3.1 Funktionsteilung

Das Prinzip der Funktionsteilung ist eine Methode, mit der einerseits Modularität erreicht werden kann, die andererseits zentrale oder teilzentrale Steuerungseinheiten entlastet und damit leistungssteigernd wirkt. Sie kann sowohl durch Auslagerung von Funktionen in Richtung Peripherie erfolgen, d.h. durch Dezentralisierung, als auch durch die funktionsmäßige Aufteilung innerhalb der gleichen Steuerungsebene.

Auslagerbar sind z.B. Funktionen zur Vorverarbeitung von Steuerinformationen, Funktionen, für die Spezialprozessoren zur Verfügung stehen oder prinzipiell alle einfach zu realisierenden Funktionen, deren Zentralisierung sich nicht lohnen würde. Sie sollten jedoch möglichst wenige zentralisierte Betriebsmittel bzw. keinen Gesamtüberblick über das ganze Vermittlungssystem benötigen, da dieses einen hohen Aufwand an Interprozessor-Kommunikation zur Folge hätte. Die Auslagerung von Funktionen ist in der Regel zusätzlich mit einer Aufteilung der Last verbunden, und zwar

in dem Sinne, daß die dezentralen Funktionseinheiten nur einer Gruppe von Teilnehmern, peripheren Einrichtungen oder Leitungen zugeordnet sind und somit nur einen Bruchteil der Gesamtlast für die jeweilige Funktion übernehmen müssen. Dies wirkt sich ferner positiv auf die Verfügbarkeit des Gesamtsystems aus, da ein Ausfall einer solchen Funktionseinheit nur einen begrenzten Wirkungsbereich hätte.

Die Funktionsteilung innerhalb der gleichen Steuerungsebene hat gegenüber der Lastteilung den Vorteil, daß nicht jede Verarbeitungseinheit für alle Funktionen ausgelegt sein muß, was sich in geringerem Speicheraufwand bemerkbar macht. Wird sie jedoch in der zentralen Ebene nicht mit Lastteilung kombiniert, so ist pro Funktion aus Sicherheitsgründen eine Duplizierung notwendig. Ferner impliziert die Funktionsteilung in der Regel einen hohen Aufwand an Interprozessor-Kommunikation.

Prinzipiell kann die Aufteilung der Funktionen nach den folgenden Gesichtspunkten erfolgen:

- Aufteilung nach vermittlungstechnischen Grundfunktionen (z.B. Signalisierungsfunktionen, Wegesuche im Koppelnetz usw.)
- Aufteilung nach Verkehrsarten (z.B. getrennt für Intern- und Externverkehr)
- Aufteilung nach Diensten (z.B. getrennt für Fernsprech- und Datendienste)
- Aufteilung nach der Zuordnung zu peripheren Einrichtungen.

Ein Beispiel für eine denkbare Funktionsaufteilung wird im dritten Kapitel anhand des Laborvermittlungssystems PILOT gegeben.

2.3.2 Lastteilung

Das Prinzip der Lastteilung unterstützt wirkungsvoll und relativ einfach die Modularisierung, es wirkt leistungs- und sicherheitssteigernd. Jede definierte Funktionssteuerung kann demnach noch nach Lastgesichtspunkten unterteilt werden, um für den betreffen-

den Satz von Funktionen Flexibilität bezüglich der Erweiterbarkeit zu erreichen. Das Lastteilungsprinzip muß ferner angewendet werden, wenn die Prozessorleistung des gewünschten Prozessortyps für die betreffende Funktion nicht ausreicht, wenn das Ein/Ausgabewerk für einen Prozessor unwirtschaftlich umfangreich wird, und grundsätzlich dort, wo eine hohe Verfügbarkeit notwendig ist.

Aus Gründen der Verfügbarkeit müssen deshalb zumindest in der zentralen Steuerungsebene wenigstens so viele Teilsteuerungen vorhanden sein, daß bei Ausfall einer Teilsteuerung die Leistungsfähigkeit des Gesamtsystems noch tolerierbar reduziert wird. In der Regel wird jedoch die Anzahl der Teilsteuerungen durch die Ausbaugröße des Systems bestimmt bzw. durch den zu verarbeitenden Gesamtverkehr.

Durch die steuerungstechnisch gruppenweise Zusammenfassung von Teilnehmern, Leitungen oder peripheren Einrichtungen realisiert man, wie bereits im Abschnitt zuvor erläutert, ebenfalls eine Aufteilung der Last, allerdings nicht im eigentlichen Sinne des Lastteilungsprinzips, da hier eine Steuerungseinheit nicht die Aufgaben einer anderen übernehmen kann. In diesem Fall spielt oft nicht die Leistungsfähigkeit der Prozessoren eine Rolle, sondern der konstruktive Aufbau, d.h. wieviele Teilnehmerschaltungen z.B. inklusive der zugehörigen Steuerung auf einer Baugruppe Platz finden. Aus diesem Grund ist die Auslastung solcher Steuereinheiten oft verschwindend klein. Dies könnte sich jedoch mit zunehmender Integrationsdichte und wegen komplizierterer Signalisierungsprotokolle in Zukunft noch ändern.

Im letzten Fall sind für die Lastaufteilung keine besonderen Maßnahmen zu treffen, da die Last von der Umwelt statistisch auf die Teilsteuerungen verteilt wird. Bei der eigentlichen Lastteilung in der zentralen Steuerungsebene müssen jedoch geeignete Lastzuteilungsmechanismen angewendet werden. Neben einfachen Zuteilungsverfahren der Aufträge und damit der Last, wie z.B. der statistisch zufälligen Zuteilung oder der gewöhnlichen zyklischen Zuteilung, sind aber auch intelligentere Strategien denkbar, die

sich z.B. am Zustand der Auftragswarteschlangen der einzelnen Teilsteuerungen orientieren. In diesem Fall benötigt man eine Zuteilungsinstanz, z.B. lokalisiert in der Steuerung des Interprozessor-Kommunikationssystems oder verteilt in den einzelnen Aufträge-produzierenden Einheiten. Entsprechende Zuteilungsstrategien werden im fünften Kapitel im Hinblick auf ihre Wirksamkeit und ihren Steuerungsaufwand analytisch untersucht.

2.3.3 Datenspeicherung

Betrachtet man die Speicherung in SPC-Vermittlungssystemen, so muß zwischen Programm- und Datenspeicherung unterschieden werden. Während die Speicherung von Programmen normalerweise auch in verteilten Steuerungen lokal in den Arbeitsspeichern der Teilsteuerungen erfolgt und deshalb unkritisch ist, bereitet die Datenspeicherung in Systemen mit Funktions- und Lastteilung eher Probleme. Man muß in der Regel, unabhängig von der Steuerungsstruktur, dezentrale, teilzentrale und zentrale, d.h. globale Datenspeicherung unterscheiden. Betroffen davon sind feste Daten, wie z.B.

- Teilnehmer-Anschlußlagen
- Teilnehmer-Berechtigungen
- Geräte-Daten
- Leitweg-Tabellen,

Verbindungs-individuelle variable Daten, wie z.B.

- Teilnehmer-Zustände
- Verbindungs-Zustände
- Gebührenstände

und Betriebsmittel-individuelle variable Daten, wie z.B.

- Geräte-Zustände.

Insbesondere bedingt durch die Lastteilung gibt es physikalisch getrennte Steuerungseinheiten, die auf die gleichen Daten bzw. Tabellen obiger Art zugreifen müssen. Es liegt deshalb nahe und

ist am einfachsten, die Speicherung dieser Daten zu zentralisieren und zentral zu verwalten. Der Datenzugriff zu solchen zentralisierten, globalen Daten erfolgt dann z.B. über Meldungen in Form von Aufträgen an die zentralisierte Speicher- bzw. Datenverwaltungseinheit bzw. Rückmeldungen von dieser. Dies hätte jedoch eine zusätzliche Last auf der für den Meldungs austausch vorgesehenen Übertragungseinrichtung zur Folge. Es könnte deshalb auch eine separate Übertragungseinrichtung für globale Speicherzugriffe von Vorteil sein.

Eine dezentrale Speicherung hat zur Folge, daß die betreffenden Daten an verschiedenen Stellen mehrfach gespeichert und verwaltet werden müssen. Dies hat neben einem vielfachen Speicherplatzbedarf und Verwaltungsaufwand den Nachteil, daß wegen der zeitbehafteten Aktualisierung kurzzeitige oder wegen Fehlern dauernde Differenzen zwischen lokalen Datenkopien entstehen können. Diesem Problem kann jedoch durch entsprechende Sicherungsfunktionen begegnet werden, deren Aufwand jedoch nur bei Daten vertretbar ist, die relativ selten aktualisiert werden müssen und die hauptsächlich nur lokal benötigt werden, wie z.B. Teilnehmer-Berechtigungen. Für oft zu aktualisierende Daten, wie z.B. Verbindungszustände, oder besonders zu sichernde Daten, wie z.B. Gebührenstände, scheint die Zentralisierung nach wie vor sinnvoll. Somit ist auch bezüglich der Datenspeicherung die teilzentrale Lösung die zur Zeit optimale und die meist verbreitete Art.

2.3.4 Interprozessor-Kommunikation

Eine ausgeprägte Verteilung der Steuerung stellt hohe Anforderungen an die Übertragungseinrichtungen zur Interprozessor-Kommunikation. Dabei hängt der Kommunikations-Verkehr stark von der gewählten Funktions- und Lastverteilung ab. Es bedarf deshalb genauer Durchsatz-Untersuchungen, um Aussagen über die Anforderungen machen zu können, bzw. Konsequenzen für die Realisierung der Übertragungseinrichtungen ziehen zu können. Geeignete Untersuchungsmethoden werden im vierten Kapitel vorgestellt und exemplarische Ergebnisse im fünften Kapitel gezeigt. Folgerungen können

dann, z.B. wegen hohen Verkehrs auf bestimmten Pfaden, die Notwendigkeit von individuellen Verbindungen zwischen den betreffenden Einrichtungen sein, z.B. in Form von festgeschalteten Steuerkanälen im PCM-Koppelnetz. Andere Lösungen könnten auf einem Bussystem basieren, das in irgend einer Form in fast allen Systemen zu finden ist. Das Funktionsteilungsprinzip bedingt dann, daß die an den Bus gekoppelten Einrichtungen in der Regel unterschiedlich starken Meldungsverkehr liefern. Aus diesem Grund bedarf es dann effektiver Buszugriffs- bzw. Abfertigungsmechanismen. Intelligente Abfertigungsstrategien, die sich am Füllgrad der betreffenden Warteschlangen orientieren, werden im fünften Kapitel hinsichtlich ihrer Wirksamkeit simulativ und analytisch untersucht.

3 DIENTE-INTEGRIERTES VERMITTLUNGSSYSTEM "PILOT"

3.1 Gesamtkonzept

Zur Dienste-Integration entstand anfangs der 80er Jahre am Institut für Nachrichtenvermittlung und Datenverarbeitung der Universität Stuttgart das Konzept "PILOT" /18/. Die Abkürzung "PILOT" steht dabei für "Partielle Integration im lokalen Telefonnetz", d.h., die Dienste-Integration erfolgt teilweise durch neue Vermittlungsstellen im Ortsnetz, die den Anschluß von Teilnehmern aller Dienste erlauben und den Zugang zum bestehenden Fernsprechnetz wie auch zu bestehenden oder zukünftigen dienstspezifischen Teilnetzen ermöglichen.

Für dieses Konzept wurde ein Labormodell einer Dienste-integrierten SPC-Ortsvermittlungsstelle mit verteilter Steuerung entwickelt und aufgebaut. Die Ausgangssituation war dabei das bestehende, in den 70er Jahren realisierte Labormodell einer zentralgesteuerten, digitalen Fernsprechvermittlungsstelle /23/. Daraus entstand durch die Einführung der Dienste-Integration /18/ und durch die Verteilung der Steuerung in einem ersten Entwicklungsschritt eine Übergangslösung und in einem zweiten Schritt die endgültige Version, deren Steuerungskonzept in diesem Kapitel nach einem kurzen Abriß der vorausgegangenen Entwicklungsstufen näher erläutert werden soll.

3.1.1 Ausgangssituation: Zentralgesteuerte Fernsprechvermittlungsstelle

Als Grundlage für die nachfolgenden Entwicklungen diente eine zentralgesteuerte digitale Fernsprechvermittlungsstelle mit Konzentratoren im Vorfeld zum Anschluß der Teilnehmer. Die Steuerungsstruktur ist in Bild 3.1 grob dargestellt.

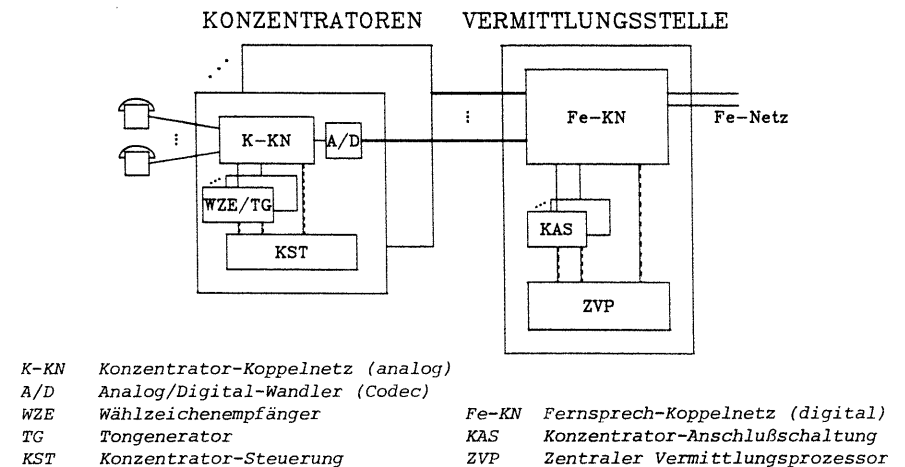


Bild 3.1: Zentralgesteuerte Fernsprechvermittlungsstelle mit Konzentratoren im Vorfeld

Die Vermittlungsstelle selbst besteht im wesentlichen aus dem zentralen Vermittlungsprozessor ZVP zur Koordinierung der gesamten Vermittlungsabläufe, dem PCM-Fernsprech-Koppelnetz Fe-KN für die Vermittlung von 64 Kbit/s -Kanälen und pro angeschlossenem Konzentrator einer Konzentrator-Anschlußschaltung KAS zur Fernsteuerung der Konzentratoren.

Das Fe-KN ermöglicht zum einen mittels entsprechender Übertragungstechnischer Anpassungseinrichtungen den Zugang zum Fernsprechnetz als auch über PCM-30/32 -Übertragungssysteme den Anschluß von Konzentratoren.

Diese Konzentratoren besitzen steuerungstechnisch keine Intelligenz, d.h., deren Konzentrator-Steuerung KST ist lediglich ein von der Vermittlungsstelle ferngesteuertes, ausführendes Organ zur Durchschaltung von Wegen im analogen Konzentrator-Koppelnetz K-KN und zum Austausch von Teilnehmersignalisierinformationen

zwischen den Signalisiereinrichtungen im Konzentrator und der dafür reservierten 16. Zeitlage des PCM-Übertragungssystems zur Vermittlungsstelle. Zur Teilnehmersignalisierung stehen neben den üblichen Teilnehmerschaltungen eine Reihe von Wählzeichenempfängern WZE und Tongeneratoren TG zur Verfügung, die jedoch, wie auch die Wege im K-KN, zentral im ZVP der Vermittlungsstelle verwaltet werden.

Die Analog/Digital-Wandlung A/D erfolgt durch 30-Kanal-Codecs im Konzentrator, welche zwischen dem analogen K-KN und den digitalen PCM-Übertragungssystemen zur Vermittlungsstelle angeordnet sind.

3.1.2 Übergangslösung: Kombinierte Fernsprech/Daten-Vermittlungsstelle

Der erste Schritt zur partiellen Integration von Datendiensten bestand in der Ergänzung des bestehenden Ortsvermittlungssystems um Einrichtungen für den Datenverkehr. Der Konzentrator wurde deshalb um einen Unterkanal-Multiplexer UK-MUX und die Vermittlungsstelle selbst durch einen Datenmodul erweitert. Die dadurch entstandene Steuerungsstruktur einer kombinierten Vermittlungsstelle für Sprache und Daten mit Konzentratoren im Vorfeld ist in Bild 3.2 grob dargestellt.

Die Datenendeinrichtungen DEE der Datenteilnehmer sind über Datenübertragungseinrichtungen DÜE am UK-MUX des Konzentrators angeschlossen. Der UK-MUX hat die Aufgabe, die verschiedenen Bitströme der Datenteilnehmer zu einer oder mehreren für den Datenverkehr reservierten Zeitlagen des PCM-Übertragungssystems zur Vermittlungsstelle zusammenzufassen. Diese Datenzeitlagen sind in 80 Unterkanäle zu je 600 bit/s unterteilt, so daß die Bitraten der in der CCITT-Empfehlung X.1 /28/ festgelegten Benutzerklassen jeweils einen oder mehrere Unterkanäle belegen. Die Belegung einer Datenzeitlage mit Unterkanälen erfolgt dabei gemäß der CCITT-Vorschrift X.51 /29/.

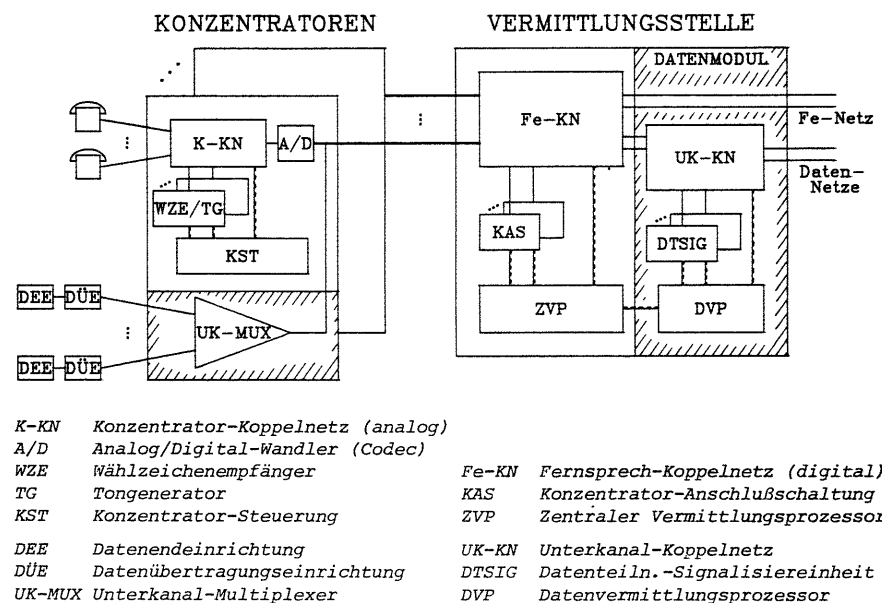


Bild 3.2: Kombinierte Fernsprech/Daten-Vermittlungsstelle mit Konzentratoren im Vorfeld

In der Vermittlungsstelle können nun die Datenzeitlagen durch das Fe-KN in das spezielle Unterkanal-Koppelnetz UK-KN im Datenmodul festgeschaltet vermittelt werden. Dieses UK-KN erlaubt dann die Vermittlung auf der Ebene der Unterkanäle und über entsprechende Übertragungstechnische Anpassungseinrichtungen den Zugang zu Datennetzen, z.B. zum Integrierten Datennetz IDN.

Die Auskopplung der Teilnehmersignalisierung aus den Bitströmen der Datenteilnehmer, bzw. deren Einkopplung sowie die Signalisierungsvorverarbeitung erfolgen in den Datenteilnehmer-Signalisierungseinheiten DTSIG. Die komplette Steuerung der Datenverbindungen übernimmt ein separater und bezüglich Datenverbindungen zentralisierter Datenvermittlungsprozessor DVP.

Durch die Einführung des eigenständigen Datenmoduls erfolgte somit steuerungstechnisch der erste Schritt in Richtung Verteilung der Steuerung. Die Aufteilung war jedoch rein funktionsbezogen in der zentralen Steuerungsebene, getrennt nach Fernsprech- und Datenvermittlungsfunktionen. Gleichzeitig mit der Implementierung der ersten Integrationsstufe erfolgten auch erste Bestrebungen in Richtung Dezentralisierung der Steuerung, nämlich indem die bislang festverdrahtete Konzentratoren-Steuerung KST und die Konzentratoren-Anschlußschaltungen KAS durch mikroprozessorgesteuerte Versionen ersetzt wurden. Dadurch bot sich die Gelegenheit, einige einfache, bisher zentralisierte Funktionen in die KAS bzw. in die KST aus dem ZVP auszulagern.

3.1.3 Dienste-integriertes Vermittlungssystem mit verteilter Steuerung

Für die Zukunft stellte die im vorangehenden Abschnitt beschriebene Übergangslösung einer kombinierten Vermittlungsstelle mit Konzentratoren noch kein befriedigendes Konzept dar.

Für eine vollständige Dienste-Integration sollte die Teilnehmer-schnittstelle für alle betroffenen Dienste und Nutzbitraten vereinheitlicht werden. Diese Forderung konnte jedoch nur durch neue Vorfeld-einheiten anstatt der bisherigen Konzentratoren erfüllt werden. Somit entstanden intelligente ausgelagerte Vermittlungsmoduln AVM mit einheitlicher Teilnehmerschnittstelle, die nicht mehr nur zur Verkehrskonzentration bzw. -Expansion dienen, sondern die u.a. Internverkehr innerhalb des Moduls eigenständig abwickeln können. Zur signalisieretechnischen Ankopplung der AVM an die Vermittlungsstelle fanden erstmals zentrale Zeichenkanäle entsprechend dem vom CCITT vorgeschlagenen Signalisiersystem No.7 /30/ Anwendung.

Innerhalb der Vermittlungsstelle konnte das Prinzip der separaten Vermittlung von Fernsprech-Zeitlagen und Daten-Unterkanälen beibehalten werden. Die Steuerungsstruktur wurde jedoch im Zuge der

Verteilung der Steuerung in eine Multi-Mikroprozessor-Steuerung geändert, wobei die Teilsteuerungen über einen gemeinsamen Bus kommunizieren können.

In Bild 3.3 ist die grobe Steuerungsstruktur des Dienste-integrierten Vermittlungssystems mit verteilter Steuerung gezeigt.

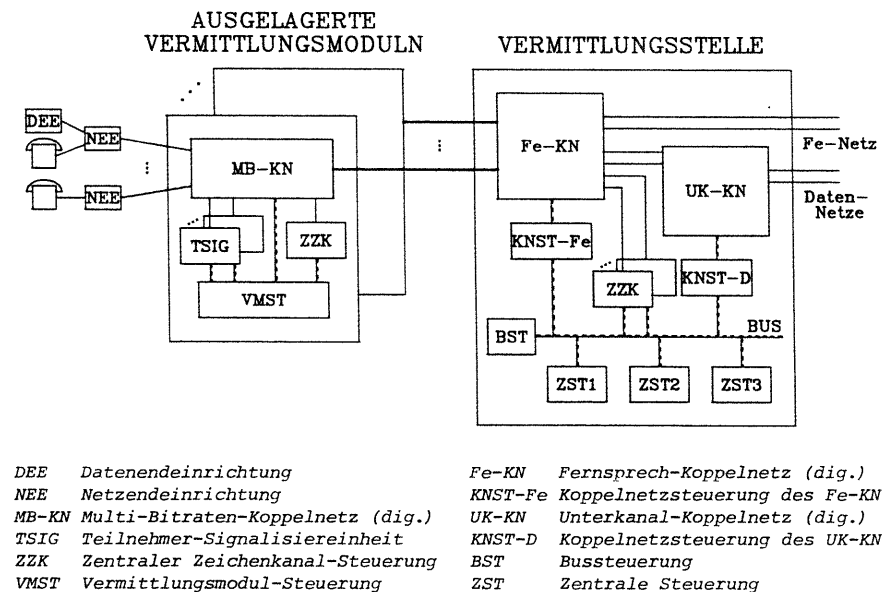


Bild 3.3: Dienste-integriertes Vermittlungssystem mit verteilter Steuerung

Alle Teilnehmer sind über Netzendeinrichtungen NEE mit einheitlicher Teilnehmerschnittstelle und einheitlichem Signalisierprotokoll /18/ an den ausgelagerten Vermittlungsmoduln AVM angeschlossen. Es stehen verschiedene NEE für Fernsprechteilnehmer, für Datenteilnehmer und für kombinierte Fernsprech/Datenteilnehmer zur Verfügung. Die Teilnehmerleitungen werden unabhängig von der

Nutzbitrate mit einer Geschwindigkeit von 80 Kbit/s pro Richtung betrieben. Übertragungstechnisch kann dabei auf bestehenden 2-Draht-Teilnehmerleitungen das Zeitgetrenntlage- oder das Echo-Kompensationsverfahren Anwendung finden /31/.

Die tatsächlichen Nutzbitraten von 600 bit/s für langsame Datenteilnehmer bis zu 64 Kbit/s für Fernsprechteilnehmer können im Multi-Bitraten-Koppelnetz MB-KN des AVM /18/ beliebig vermittelt werden, wobei auf den PCM-Leitungen zur Vermittlungsstelle für die Belegung der Datenzeitlegen mit Unterkanälen wie bisher die CCITT-Empfehlung X.51 Anwendung findet. Das MB-KN ist im Vollausbau für 512 Teilnehmer und bis zu 8 PCM-30/32 -Systemen konzipiert und besitzt eine modulare Struktur. Prinzipiell besteht das MB-KN aus einer Kombination eines PCM- und eines Unterkanalkoppelnetzes, welche sich wegen der geringen Größe trotz aufwendiger Technik im AVM lohnt.

Die Ein- bzw. Auskopplung und Vorverarbeitung der Teilnehmersignalisierung nach der einheitlichen Signalisierungsprozedur erfolgt in speziellen Teilnehmersignalisierungseinheiten TSIG. Aus technologischen und strukturellen Gründen wird pro 64 Teilnehmer je eine TSIG benötigt.

Die Signalisierung mit der Vermittlungsstelle erfolgt gemäß dem Signalisierungsprotokoll CCITT No.7. Sie wird physikalisch als 64 Kbit/s -Bitrate in der 16. Zeitlege eines PCM-Systems übertragen. Die Ebenen 1 und 2 des No.7 -Protokolls werden von einem separaten Signalisierungsprozessor ZZK bearbeitet. Dieser beinhaltet neben einem Mikroprozessor einen handelsüblichen HDLC-Steuerbaustein zur Unterstützung des No.7 -Protokolls, da dieses teilweise dem HDLC-Protokoll sehr ähnlich ist.

Die Vermittlungsmodul-Steuerung VMST koordiniert weitgehend autonom Internverbindungen innerhalb des AVM und bearbeitet Externverbindungen von oder nach außerhalb des AVM, also die Ebenen 3 und 4 des No.7 -Protokolls. Somit bildet der AVM eine vollständige kleine Dienste-integrierte Vermittlungsstelle und könnte

deshalb auch ohne große Änderungen z.B. als kleine Nebenstellenanlage eingesetzt werden.

Innerhalb der Vermittlungsstelle wurden die beiden Koppelnetze Fe-KN und UK-KN der kombinierten Vermittlungsstelle beibehalten. Ein separater modularer Ausbau im Fernsprech- und im Datenteil wird dadurch ermöglicht. Die Steuerung der Koppelnetze erfolgt jedoch nicht mehr im Bereich der Zentralsteuerung, sondern in speziellen peripheren Koppelnetzsteuerungen KNST-Fe für das Fe-KN und KNST-D für das UK-KN.

Zu den peripheren Einrichtungen ist auch eine Reihe von Signalisierungsprozessoren ZZK für zentrale Zeichenkanäle nach CCITT No.7 zu rechnen. Sie sind identisch aufgebaut wie ihre Partneereinrichtungen im AVM. Neben je einem ZZK pro AVM sind weitere für die Signalisierung mit anderen Vermittlungsstellen einsetzbar.

Weitere periphere Ein/Ausgabeprozessoren zum Anschluß von Massenspeichern, Druckern und Bildschirmgeräten für die Mensch/Maschine-Kommunikation sind in Bild 3.3 nicht eingezeichnet, da sie für den Vermittlungsvorgang nicht von direkter Bedeutung sind.

Der zentrale Steuerungsbereich der Vermittlungsstelle besteht aus mehreren zentralen Steuerungen ZST. Im Labormodell sind zunächst drei ZST vorgesehen, vorerst wurde jedoch nur eine implementiert. Bei Bedarf können aber leicht weitere ZST modular ergänzt werden.

Die Interprozessor-Kommunikation aller Teilsteuerungen erfolgt über einen gemeinsamen Bus durch Meldungsaustausch. Die Bussteuerung BST vergibt die Buszugriffsrechte und übernimmt in einer speziellen Betriebsart teilweise die Meldungslenkung, z.B. zur Lastaufteilung auf mehrere ZST. Die Buskopplung im zentralen Bereich hat gegenüber einer Kopplung über Steuerkanäle des Koppelnetzes Vorteile bezüglich der Geschwindigkeit und ermöglicht einfachere Anpassungsschaltungen. Trotzdem könnte die letztere Möglichkeit durch entsprechende Ein- und Auskoppelschaltungen ebenfalls realisiert werden.

3.2 Aufgabenverteilung

Die globale Aufgabenverteilung im Gesamtsystem ist hauptsächlich bestimmt durch das teilzentrale Strukturkonzept mit ausgelagerten Vermittlungsmoduln mit eigener Steuerung für eine Gruppe von Teilnehmern. Im AVM wird trotz einer einfachen Steuerungsstruktur mit einer zum größten Teil konzentrierten Steuerung und nur zwei ausgelagerten Funktionen der Signalisierung eine möglichst autonome Verbindungssteuerung angestrebt. Die Vermittlungsstelle selbst soll nur als Bindeglied für AVM-externe und Vermittlungsstellen-externe Verbindungen, für besonders zu sichernde Funktionen und für selten in Anspruch genommene Betriebsmittel benötigt werden. Ihre Steuerung ist teilzentral und modular im zentralen Bereich mit Funktions- und Lastteilung.

3.2.1 Ausgelagerter Vermittlungsmodul

3.2.1.1 Globale Aufgaben

Die globalen Aufgaben des AVM sind:

- Teilnehmersignalisierung nach einheitlichem Protokoll für alle Dienste
- Vollständige Verbindungssteuerung für AVM-interne Verbindungen
- Teilweise Verbindungssteuerung für gehende bzw. kommende AVM-externe Verbindungen
- Signalisierung zur Vermittlungsstelle nach CCITT No.7 -Protokoll für AVM-externe Verbindungen
- Teilweise Überlastabwehr
- Routine- und bedarfsmäßige Prüfung von AVM-Baugruppen
- Routine- und bedarfsmäßige Erfassung von Verkehrsmeßdaten und ihre Übermittlung zur Vermittlungsstelle.

Mit Ausnahme der beiden Signalisierungsfunktionen werden diese Aufgaben von der Vermittlungsmodul-Steuerung VMST koordiniert und abgewickelt.

3.2.1.2 Dezentralisierte Funktionen

Eine Teilnehmersignalisiereinheit TSIG koppelt die Signalisierungsinformationen aus den Bitströmen von jeweils 64 Teilnehmern aus und führt eine Vorverarbeitung entsprechend dem mittels eines endlichen Automaten definierten einheitlichen Signalisierprotokoll durch. Nur die für die Weiterverarbeitung notwendigen Informationen werden an die VMST weitergegeben. Zum Teilnehmer zu sendende Signalisierungsinformationen, die sich aus der Vorverarbeitung bereits in der TSIG ergeben oder von der VMST übergeben werden, fügt die TSIG in die Bitströme zu den Teilnehmern wieder ein. Die Behandlung der Teilnehmersignalisierung muß aus Geschwindigkeitsgründen ausgelagert werden, da bei 512 Teilnehmern ein Signalisierungswort bereits nach 240 ns auf das vorangehende folgen kann. Selbst die Gruppenbildung in 8 TSIG für je 64 Teilnehmer erfordert noch eine Hardware-Lösung in Form eines mikroprogrammierten sequentiellen Netzwerkes.

Die Zentralkanal-Signalisiereinheit ZZK übernimmt die Ebenen 1 und 2 des CCITT No.7 -Protokolls zur Signalisierung mit der Vermittlungsstelle. Diese Funktionen wurden einerseits aus der VMST ausgelagert, da der Steuerungsaufwand für die Ebene 2 des No.7 -Protokolls bei gleichzeitiger Bearbeitung des Sendens und Empfangens mit 64 Kbit/s -Bitrate sehr hoch ist und einen normalen 8 bit -Mikroprozessor voll auslastet. Der zweite Grund besteht darin, daß für die Ebene 1 vorteilhaft HDLC-Steuerbausteine verwendet werden können.

3.2.2 Vermittlungsstelle

3.2.2.1 Globale Aufgaben

Die globalen Aufgaben der Vermittlungsstelle sind:

- Verbindungssteuerung für Internverbindungen zwischen zwei verschiedenen AVM
- Verbindungssteuerung für abgehende und ankommende Externver-

bindungen mit Leitweglenkung

- Signalisierung mit dem AVM und mit anderen Vermittlungsstellen nach CCITT No.7 -Protokoll
- Gesamte Gebührenerfassung und -verwaltung
- Routine- und bedarfsmäßige Prüfung von Baugruppen der Vermittlungsstelle
- Routine- und bedarfsmäßige Erfassung und Auswertung von Verkehrsmeßdaten
- Verwaltung von Teilnehmer- und Gerätedaten.

Die Koordinierung und Bearbeitung dieser Funktionen erfolgt, abgesehen von den dezentralisierten Funktionen der Koppelnetzsteuerung und der Zentralkanalsignalisierung, im zentralen Steuerungskomplex, bestehend aus mehreren ZST mit Lastteilung und eventuell auch Funktionsteilung. Der Lastteilungsbetrieb im zentralen Bereich sollte aus Sicherheitsgründen vorgesehen werden. Die Lastzuteilung und die eventuelle Funktionszuteilung auf die Teilsteuerwerke ZST erfolgt teilweise dynamisch beim Meldungs austausch durch die Bussteuerung BST. Dadurch erfüllt die BST gleichzeitig Aufgaben für eine begrenzte Überlastabwehr und eine Ersatzschaltsteuerung. Diese Funktionen der BST sind im Abschnitt 3.3 näher erläutert.

3.2.2.2 Dezentralisierte Funktionen

Die Zentralkanal-Signalisierung ist auch in der Vermittlungsstelle aus den gleichen Gründen wie beim AVM ausgelagert. Die entsprechenden Moduln im AVM und der Vermittlungsstelle sind identisch.

Die speziellen ausgelagerten Funktionen der KNST-Fe sind:

- Wegesuche im Fe-KN
- Wededurchschaltung bzw. -auslösung
- Feste Durchschaltung der Datenzeitlagen zum UK-KN
- Wegespeicherung
- Routine- und bedarfsmäßige Wegeprüfung.

Die Funktionen der KNST-D sind:

- Wegesuche im UK-KN in Unterkanaltechnik
- Wededurchschaltung bzw. -auslösung
- Wegespeicherung
- Routine- und bedarfsmäßige Wegeprüfung.

Die Dezentralisierung dieser Funktionen beider Koppelnetze hat zwei Gründe. Die Schnittstelle zu PCM-Koppelnetzen besteht aus sogenannten Steuerspeichern, und ein Durchschalte- bzw. Auslösebefehl erfolgt durch einen Zugriff auf diesen Steuerspeicher. Dazu wird auf jeden Fall eine Speicherzugriffseinheit benötigt, wofür sich ein Mikroprozessor besonders gut eignet. Mit dem Einsatz von Mikroprozessoren an dieser Stelle schafft man aber gleichzeitig weiter zu nutzende Rechnerkapazität für intelligente Funktionen. Für eine Auslagerung von Funktionen aus dem zentralen Steuerungsbereich sind dann die Wegesuche und Prüffunktionen für das Koppelnetz naheliegend. Gleichzeitig erreicht man damit eine wesentliche Entlastung der Zentralsteuerung. Ferner bietet eine separate Koppelnetzsteuerung die Möglichkeit, neben dem Koppelnetz selbst auch die zugehörige Steuerung modular zu gestalten, und Modularität bedeutet Vorteile bezüglich der Ausbaubarkeit und der Verfügbarkeit. Im Labormodell wurde jedoch wegen der relativ kleinen Ausbaustufe der Koppelnetze nur je eine Koppelnetzsteuerung implementiert.

3.3 Interprozessor-Kommunikation

Sämtliche Aktivitäten im System werden vom Betriebssystem der jeweiligen Teilsteuerungen im Sinne von Prozessen koordiniert. Dies sind die eigentlichen Vermittlungsprozesse in den koordinierenden Steuerungen, die Signalisierungsprozesse in den ausgelagerten Signalisierungsprozessoren, die Wegesuchprozesse in den Koppelnetzsteuerungen und eine Reihe anderer Hilfsprozesse.

Die Kommunikation zwischen Prozessen erfolgt ausnahmslos durch Meldungs austausch über FIFO-Warteschlangen. Das Senden von Meldungen, d.h. der Eintrag in eine Ausgangswarteschlange AWS, und das Empfangen, d.h. der Austrag aus einer Eingangswarteschlange EWS, wird von entsprechenden Betriebssystemfunktionen übernommen. Sofern zwei kommunizierende Prozesse in derselben Steuereinheit lokalisiert sind, erfolgt der Meldungs austausch über eine interne Software-Warteschlange im Arbeitsspeicher der jeweiligen Steuereinheit. Handelt es sich jedoch um Prozesse in unterschiedlichen Steuereinheiten, so muß man im Gesamtsystem die drei in Bild 3.4

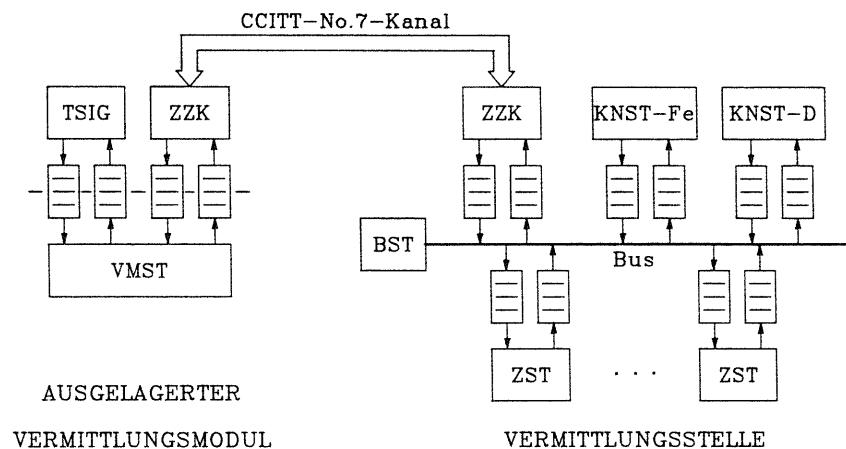


Bild 3.4: Prozessor-Kopplung im System "PILOT"

schematisch dargestellten Fälle unterscheiden, nämlich:

- Direkte Warteschlangen-gekoppelte Interprozessor-Kommunikation innerhalb des AVM über individuelle parallele Ein/Ausgabekanäle
- Indirekte Warteschlangen-gekoppelte Interprozessor-Kommunikation zwischen AVM und Vermittlungsstelle über eine Ebene 2 - Verbindung eines CCITT No.7 -Signalisierungskanals, gesteuert von zwei korrespondierenden ZZK-Steuerungen
- Indirekte Warteschlangen-gekoppelte Interprozessor-Kommunikation in der Vermittlungsstelle über den gemeinsamen Bus, gesteuert von der Bussteuerung BST.

3.3.1 Kommunikation innerhalb des Vermittlungsmoduls

Im AVM sind die Warteschlangen physikalisch je zur Hälfte in den betreffenden Partnereinheiten untergebracht, so daß sich die parallelen Hardware-Schnittstellen in Bild 3.4 in der Mitte der gezeichneten Warteschlangen befinden. Die Warteschlangenhälften sind in Software realisiert, mit Ausnahme in der TSIG, wo sich wegen des Hardware-Aufbaus FIFO-Speicherbausteine besser eignen.

Den physikalischen Datenaustausch zwischen den Warteschlangenhälften steuern Ein/Ausgabe-Routinen der betreffenden Betriebssysteme mittels eines einfachen Handshake-Protokolls. Dieses wird von den eingesetzten Ein/Ausgabe-Bausteinen mit unterstützt und läuft wie folgt ab: Das vorderste Wort in der Sendewarteschlange wird in den Ausgabebaustein übergeben, der es an den Schnittstellenleitungen anlegt und darauf ein Bereit-Signal sendet. Dieses löst bei der empfangenden Steuereinheit eine Unterbrechung aus. Wenn das zugehörige Unterbrechungsprogramm zum Zuge kommt, was vom Betriebssystem prioritätsabhängig gesteuert wird, übernimmt es das Wort in seine Empfangswarteschlange, falls dort noch ein Wartepplatz frei ist. Gleichzeitig wird ein Quittungs-Signal zurückgeschickt, welches von der sendenden Einheit erkannt wird und zur Freigabe für das Senden eines eventuellen nächsten Wortes dient.

3.3.2 Signalisier-Kommunikation zwischen Vermittlungsmodul und Vermittlungsstelle

Die Schnittstelle für den Meldungsaustausch zwischen dem AVM und der Vermittlungsstelle bilden aus Sicht der VMST bzw. des ZST-Komplexes die Warteschlangen zu und von den ZZK-Steuerungen. Die ZZK bieten logisch eine gesicherte Ebene 2 -Verbindung über den CCITT No.7 -Kanal an.

Die Arbeitsweise und Eigenschaften von No.7 -Signalisierkanälen sowie eine Leistungsuntersuchung sind in /18/ ausführlich beschrieben.

3.3.3 Bus-Kommunikation innerhalb der Vermittlungsstelle

Die Teilsteuerungen in der Vermittlungsstelle besitzen alle identische Buszugriffseinheiten mit je einer Eingangswarteschlange EWS und einer Ausgangswarteschlange AWS. Unter Kontrolle der Bussteuerung BST hat der Bus die Aufgabe, die Meldungen aus den AWS der Ursprungseinheiten möglichst schnell zu den richtigen EWS der Zieleinheiten zu übertragen.

Dabei ergibt sich ein erstes Problem aus der Tatsache, daß die einzelnen Steuerungen statisch und dynamisch verschieden starken Meldungsverkehr erzeugen. Eine starre, z.B. rein zyklische Busvergabe ist deshalb nicht die sinnvollste Lösung, da sie die einzelnen Teilsteuerungen nicht priorisiert. Besser ist bereits die meist eingesetzte statische Priorisierung bestimmter Steuerungen, z.B. durch mehrfache Busvergabe innerhalb eines Zyklus. Im betrachteten System sollte jedoch eine noch intelligentere Möglichkeit, nämlich eine dynamische Priorisierung, eingesetzt werden. Sie wird ebenfalls durch mehrfache Busvergabe an zu priorisierende Einheiten pro Zyklus erreicht, wobei sich aber die Anzahl der Busvergaben an jede Einheit dynamisch am Zustand der AWS orientiert und damit am momentanen Meldungsverkehrsaufkommen der einzelnen Einheiten. Zur Steuerung dieser Aufgabe wurde u.a. eine zentrale Bussteuerung BST eingeführt. Diese fragt die einzelnen

AWS entsprechend einer momentanen Zyklusfolge ab (engl. Polling). Dabei erfährt sie, ob zu übertragende Meldungen anstehen und wieviele Meldungen sich gegebenenfalls in der AWS befinden. Mit dieser Information ist die BST in der Lage, die Abfrage-Zyklusfolge eventuell neu zu optimieren und gegebenenfalls eine oder mehrere Meldungen zu übertragen. Dieser Algorithmus, im folgenden "Dynamisches Polling" genannt, wurde im Rahmen einer detaillierten Simulationsstudie untersucht, deren Ergebnisse im fünften Kapitel beschrieben sind.

Ein zweites Problem bei der Übertragung von Ursprungs- zu Zieleinheiten ist durch das Lastteilungsprinzip bedingt. Im betrachteten System sind aus Sicherheitsgründen im zentralen Steuerungsbereich mehrere gleichberechtigte Lastteilungsprozessoren ZST vorgesehen, so daß bei Ausfall einer dieser ZST kein Totalausfall erfolgt, sondern nur die Leistungsfähigkeit reduziert wird. Dabei stellt sich jedoch die Frage, wie Anforderungen an den zentralen Steuerungskomplex auf die einzelnen ZST-Einheiten verteilt werden. Die einfachste Möglichkeit wäre, wenn die Meldungen zum ZST-Bereich von der sendenden Ursprungseinheit zufällig oder z.B. zyklisch den verschiedenen ZST-Einheiten zugeteilt würden, d.h. wenn in diesem Falle die Meldungs-Zieladressen abgewechselt würden. Im fünften Kapitel beschriebene analytische Untersuchungen zeigen jedoch, daß Zuteilungsstrategien, die sich am Zustand der EWS der Lastteilungsprozessoren orientieren, ähnlich wie beim dynamischen Polling, eine bessere Verkehrsleistung aufweisen. Zur Implementierung einer solchen flexiblen, dynamischen Zuteilungsstrategie bietet sich ebenfalls die zentrale BST an. Wenn sie über die Zustände der EWS der ZST-Einheiten informiert ist, ist sie auch in der Lage, den Meldungstransfer durch Adressierung der als optimal ermittelten Zieleinheit zu lenken. Mittels änderbarer Meldungslenkungstabellen hat sie dabei auch gleichzeitig die Möglichkeit, für ausgefallene Steuereinheiten bestimmte Meldungen entsprechend umzulenken, was einer Ersatzschaltung gleichkommt.

Prinzipiell könnte nun eine Busübertragung mit intelligenter Meldungslenkung so ablaufen, daß Ursprungseinheiten I ihre Meldungen

grundsätzlich an die BST senden, die sie dann in einem zweiten Schritt an die in der BST ermittelten Zieleinheiten J schickt, wie in Bild 3.5 schematisch dargestellt.

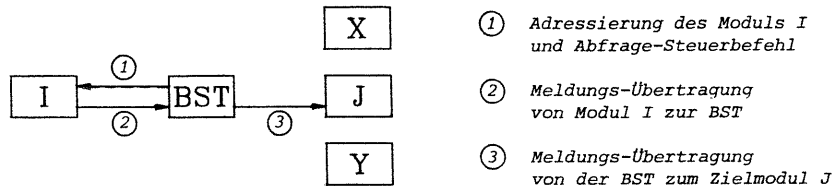


Bild 3.5: Einfacher Bustransfer mit Meldungslenkung in der Bussteuerung BST

Bei dieser Lösung wäre die BST jedoch ein zentrales Element, dessen Ausfall den Totalausfall des Gesamtsystems zur Folge hätte. Statt einer naheliegenden Duplizierung der BST wurde für den Bustransfer ein Protokoll entwickelt, das bei Ausfall der leistungssteigernden dynamischen Meldungslenkung nur den Verlust dieser Leistungssteigerung zur Folge hat. Lediglich die grundlegende Bussteuerlogik für den Polling-Betrieb muß in der BST dupliziert sein, wie auch der Bus selbst. In Bild 3.6 ist das Prinzip des vorgeschlagenen Busprotokolls der einfachen Lösung von Bild 3.5 gegenübergestellt.

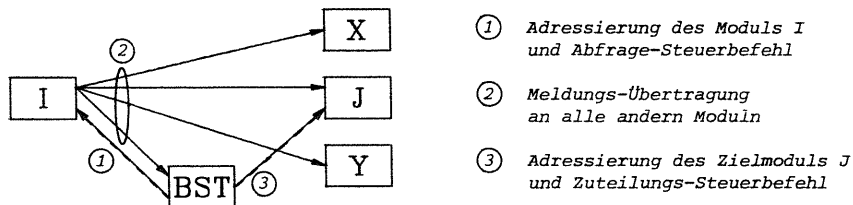


Bild 3.6: Verbesserter Bustransfer mit Meldungslenkung durch die Bussteuerung BST

Eine Übertragung von Einheit I mit Meldungslenkung nach Einheit J läuft vereinfacht wie folgt ab:

Der Modul I erhält das Buszugriffsrecht, wenn er von der Polling-Logik der BST mittels eines Abfrage-Steuerbefehls adressiert wird. Falls keine Meldung in der AWS ansteht, wird eine entsprechende Steuermeldung an die BST zurückgeschickt. Andernfalls wird die anstehende Meldung gesendet. Obwohl alle beteiligten Einheiten diese Meldung empfangen und zwischenspeichern, ist die Zieladresse dieser Meldung in jedem Falle fest definiert, hier z.B. die Adresse von Modul X. Dadurch erreichen die Meldungen auch bei Ausfall der Meldungslenkungsfunktion der BST ihr Ziel, wenn auch nicht immer das optimale unter den Lastteilungsprozessoren. Im Normalfall hört die BST jedoch mit und kann sich bei Bedarf in die Übertragung einschalten. Will sie die Meldung beispielsweise zu Modul J umlenken, so sendet sie einen Zuteilungs-Steuerbefehl mit der Adresse des Moduls J. Darauf übernimmt dieser die empfangene und zwischengespeicherte Meldung in seine EWS. Für die anderen Einheiten bedeutet der Zuteilungs-Steuerbefehl, daß sie ihren Zwischenspeicher löschen können. Wäre der Zuteilungs-Befehl ausgeblieben, so hätte im Beispiel der Modul X die Meldung übernommen und die anderen Einheiten hätten ihre Zwischenspeicher gelöscht.

Das Busprotokoll wird durch vier Prozesse bestimmt, nämlich den Sende-Prozeß MODUL-S und den Empfangs-Prozeß MODUL-E in jeder Teilsteuerung sowie den Polling-Prozess BST-POLL und den Zuteilungs-Prozess BST-ZUT in der BST. Die vom CCITT für derartige Prozesse vorgeschlagene Spezifikations- und Beschreibungssprache SDL /32/ ist in den Bildern 3.8 und 3.9 zur genaueren Spezifikation dieser vier Prozesse benutzt. Bild 3.7 zeigt dazu alle möglichen über den Bus ausgetauschten Botschaften und von welchen Prozessen diese empfangen werden. Jede Botschaft setzt sich dabei aus den vier Teilen ADR, K, ST und INF zusammen, die wie folgt definiert sind:

ADR, K, ST, INF

Informationsteil; er enthält:

- Inhalt der Meldung, falls K=I
- Art der Steuermeldung, falls K=S

Eventuelle Statusinformation; gibt Aufschluß
über den Zustand der EWS bzw. AWS der absen-
denden Einheit

Kennzeichnung für den Informationsteil:

- K=S für Steuermeldungen
- K=I für Meldungen mit Nutzinformation

Adressteil für Adresse der Zieleinheit oder Zielgruppe. Er dient der gezielten Adressierung bei der Abfrage und Zuteilung bzw. enthält die vorbestimmte Zieladresse für den Fall, daß die Meldungslenkungsfunktion, d.h. der Zuteilungs-Prozeß BST-ZUT, ausfällt.

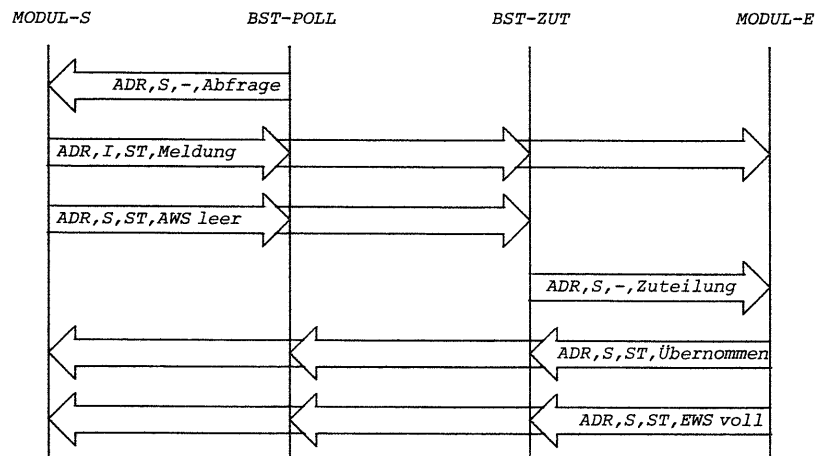


Bild 3.7: Mögliche Botschaften zwischen den Prozessen für den Bustransfer

In den SDL-Diagrammen sind die Wartezustände der Prozesse, soweit notwendig, durch Zeitüberwachungen abgesichert. Dafür stehen jeweils Prozeß-interne Zeitgeber (Timer) zur Verfügung. Das Signal "Restart" im Prozeß BST-POLL dient zum erstmaligen Anlauf der

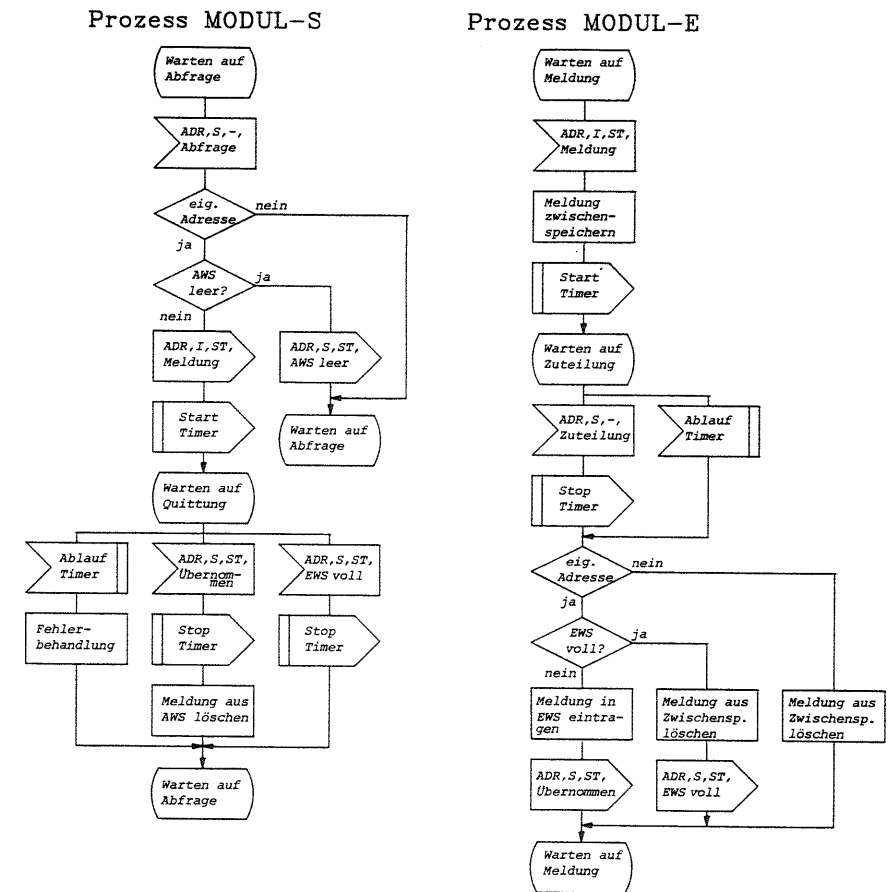


Bild 3.8: SDL-Diagramme der Sende- und Empfangsprozesse der Teilsteuerungen

Polling-Logik. Das Protokoll ist so ausgelegt, daß auch im Falle eines eventuellen Ausfalls des Prozesses BST-ZUT noch ein Bus-transfer ohne Meldungslenkungsfunktion möglich ist.

4 ANALYSE VON VERMITTLUNGSSTEUERUNGEN

4.1 Allgemeines über Analysetechniken und -ziele

Komplizierte Vorgänge in realen Systemen können prinzipiell durch Messungen und Beobachtungen am realen System quantitativ und qualitativ erfasst werden. Für die Untersuchung der Leistungsfähigkeit und des Verhaltens von Vermittlungssystemen verbietet sich diese Methode jedoch meist, da es sich entweder um Konzepte handelt, die sich erst im Planungs- oder Entwicklungsstadium befinden, oder um Prototypen, für die im Labor die realen Umweltbedingungen fehlen. Selbst bei Systemen in der realen Umwelt sind manche Vorgänge durch die stochastische Verhaltensweise der Teilnehmer schwer zu erfassen, weil sie z.B. sehr selten auftreten.

Aus diesen Gründen muß die Untersuchung von Vermittlungssystemen und deren Teilsysteme durch analytische und/oder simulative Methoden erfolgen /33/. Dafür werden verkehrstheoretische Modelle bzw. Teilmodelle benötigt, die die logischen und zeitlichen Zusammenhänge der Realität in einfacher Weise beschreiben und leicht durchschaubar sind. Oft approximieren allerdings solche Modelle die Realität nur und sind deshalb teilweise fehlerbehaftet.

Die Systemumwelt muß dabei in jedem Falle in Verkehrsmodelle abgebildet werden, die den zufallsbedingten Ankunftsprozeß des Testverkehrs beschreiben, der dann einem analytischen Systemmodell, einem Simulationsmodell oder, in Hardware nachgebildet, einem realen System zugeführt wird. In letzterem Falle hat man es dann mit der sogenannten Umweltsimulation /34/ zu tun.

Als Systemmodelle eignen sich Warteschlangenmodelle mit den gängigen Strukturelementen wie beispielsweise Bedienungseinheiten, Warteschlangen, oder Schalter. Neben der Simulation /35,36/, d.h. der Nachbildung des Ablaufgeschehens im Warteschlangenmodell auf einem Rechner, stellt die Bedienungs- oder Warteschlangentheorie

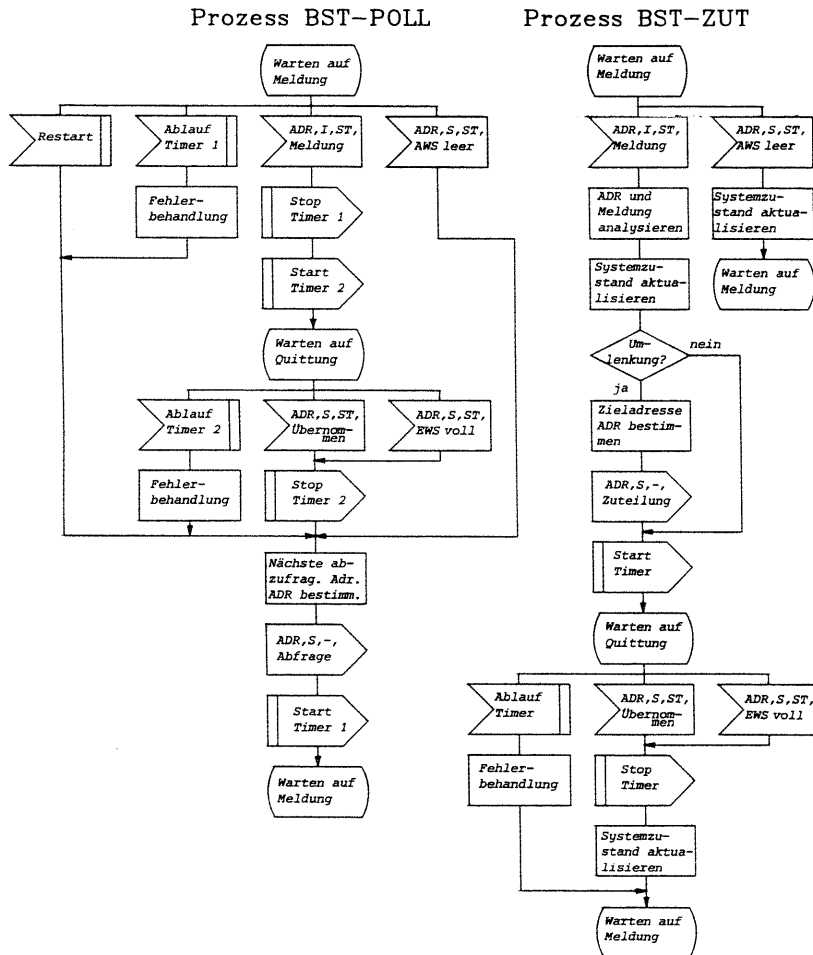


Bild 3.9: SDL-Diagramme des Abfrage- und des Zuteilungsprozesses der Bussteuerung

/37/ eine Reihe analytischer Methoden zur Modellanalyse bereit, die jedoch oftmals mit vertretbarem Aufwand nur auf stark vereinfachte Modelle, Teilmodelle oder Teilaspekte anwendbar sind.

Nach einer Beschreibung der wichtigsten Punkte der Modellbildung werden im Verlauf dieses Kapitels schwerpunktmäßig nur die Techniken der Modellanalyse aufbereitet und teilweise erweitert, die sich für die Untersuchung von verteilten Vermittlungssteuerungen besonders eignen und die zur Lösung der im fünften Kapitel behandelten Probleme benötigt werden:

- Signalflußanalyse :

Das Ziel von Signalflußanalysen ist, möglichst schnell, früh und einfach einen groben Überblick über die Belastungen der Teilsysteme des vereinfachten Gesamtmodells von verteilten Strukturen zu erhalten, abhängig von der vorgesehenen Aufgaben- und Verkehrsverteilung. Somit können wenig sinnvolle Strukturen, Systemkonfigurationen und Aufgabenverteilungen erkannt und geeignete Konzepte ausgewählt werden.

- Zeittreue Simulation :

Die Simulation des detaillierten Gesamtmodells hat das Ziel, theoretisch beliebig genaue Ergebnisse über die Leistungsfähigkeit, das Verhalten und über Engpässe des Systems zu liefern. Dafür ist jedoch ein teilweise erheblicher Zeit- und Arbeitsaufwand erforderlich, auch schon bei der etwas einfacheren stationären, zeittreuen Simulation, die aber i.a. den meisten Anforderungen genügt. Lediglich bei der Untersuchung des dynamischen Überlastverhaltens ist eine instationäre Simulationsmethode notwendig /38/. Mit Hilfe der Simulation lassen sich einerseits charakteristische Leistungsgrößen ermitteln, die sich auf das gesamte Vermittlungssystem beziehen, wie z.B. :

- Wähltonverzug
- Rufverzug

- Maximaler Durchsatz (akzeptierbare Anrufrate)
- Blockierung aus steuerungstechnischen Gründen.

Es können aber auch andererseits charakteristische Größen von Teilsystemen erhalten werden wie z.B. :

- Wartewahrscheinlichkeiten
- Blockierungswahrscheinlichkeiten
- Mittelwert und Verteilungsfunktion von Warte- oder Durchlaufzeiten
- Mittelwert und Verteilungsfunktion von Warteschlangenlängen
- Mittelwert und Verteilungsfunktion von Belastungen der Bedienungseinheiten
- Verteilungsfunktionen von Steuerdatenverkehrsströmen.

- Warteschlangenmodellanalyse :

Die analytische Untersuchung von Teilmodellen bzw. Teilaspekten mittels bedienungstheoretischer Methoden ist notwendig, da sie i.a., wo möglich, schneller Ergebnisse liefert als die Simulation. Dadurch werden dann erst umfangreiche Parameterstudien über die Leistungsfähigkeit von Betriebsweisen, Strategien oder Teilsystemen ermöglicht. Dies gilt insbesondere auch für die Untersuchung seltener Ereignisse. Ferner führen analytische Lösungen eher zu allgemein handhabbaren Dimensionierungsrichtlinien, z.B. in Gestalt mathematischer Formeln. Wie bei der Simulation genügt auch bei der Warteschlangenmodellanalyse meist die einfachere stationäre Analyse den gestellten Anforderungen, sofern nicht das dynamische Überlastverhalten von Interesse ist.

4.2 Modellierung von Vermittlungssteuerungen

4.2.1 Verkehrsmodelle

Die Abläufe innerhalb von Vermittlungssystemen sind im wesentlichen durch zufallsabhängig eintreffende Anforderungen der Teilnehmer an Betriebsmittel und durch die zufallsabhängige oder deterministische Bedienung dieser Anforderungen durch die Betriebsmittel bestimmt.

4.2.1.1 Ankunftsprozesse

Die Anforderungen der Teilnehmer (Rufe) werden i.a. durch einen zufallsabhängigen Ankunftsprozeß aus Verkehrsquellen modelliert, der die Ankunftsabstände T_A der Rufe beschreibt. Eine durch Messungen bestätigte, gute Näherung des stationären Ankunftsprozesses einer großen Anzahl von Teilnehmern ist der Poisson-Prozeß mit negativ exponentieller Verteilungsfunktion der Ankunftsabstände

$$F_A(t) = P\{T_A \leq t\} = 1 - e^{-\lambda \cdot t}, \quad (4.1)$$

mit der Ankunftsrate, d.h. dem Kehrwert des Erwartungswertes des Ankunftsabstandes

$$\lambda = 1 / E[T_A]. \quad (4.2)$$

Für andere Ankunftsprozesse kann auch der Modellansatz einer konstanten, einer Erlang-k- oder hyperexponentiellen Verteilungsfunktion sinnvoll sein, und zur Modellierung von instationären Rufankunftsprozessen wird der verallgemeinerte Poisson-Prozeß mit zeitabhängiger Ankunftsrate $\lambda(t)$ angesetzt /38/.

Bei genauerer Betrachtung besteht jedoch ein Ruf aus mehreren Teilnehmerereignissen (Teilrufe), wie z.B. dem Verbindungswunsch (Abheben), den Wählfziffern und dem Verbindungsende (Auflegen). Diese Teilrufe treten je nach Art des Rufes in einer bestimmten

zeitlichen Reihenfolge auf, wobei die Zwischenankunftsabstände $T_{i,j}$ entweder zufallsabhängig oder abhängig von Rückwirkungen des Systems (Wählton, Besetztton usw.) sind. Diese sogenannten Ereignisketten /23/ können durch einen Zustandsprozeß teilnehmerindividuell modelliert werden, der die Teilrufanforderungen an das System generiert.

Eine Ereigniskette wird durch eine Anforderung eines Rufankunftsprozesses (i.a. Poisson-Prozeß) gestartet, d.h. aus ihrem Ruhezustand gebracht. Mit gewissen Verzweigungswahrscheinlichkeiten $p_{i,j}$ erfolgen dann Zustandsübergänge mit zufallsabhängigen oder durch Systemrückwirkungen bestimmten Verweildauern $T_{i,j}$. Bild 4.1 zeigt dazu exemplarisch einen Ausschnitt eines entsprechenden Zustandsdiagramms in SDL-Darstellung. Die zufallsabhängigen Verweildauern $T_{i,j}$ werden dabei durch Prozeß-interne Zeitgeber (Timer) realisiert, deren Ablaufzeiten einer geeigneten Verteilungsfunktion genügen, z.B. näherungsweise wieder einer Exponential-Verteilung. Für die Wahrscheinlichkeiten $p_{i,j}$ und die mittleren Verweildauern $E[T_{i,j}]$ können z.B. typische Meßwerte verwendet werden.

Ein Verkehrsmodell mit derartigen Ereignisketten kann z.B. in sehr detaillierten Simulationen Anwendung finden, wie z.B. bei der belegungsstreu Simulation (siehe Abschnitt 4.4). Falls teilnehmerbezogene Ereignisketten jedoch zu aufwendig sind und trotzdem eine Modellierung auf Teilrufebene notwendig ist, kann aufgrund der hohen Teilnehmeranzahl im System auch auf der Teilrufebene ein Poisson-Prozeß aus Teilrufquellen zugrunde gelegt werden, d.h. man betrachtet ein stochastisches Gemisch von negativ exponentiell verteilt eintreffenden Teilrufen, deren Typ quantitativ durch eine Auftrittswahrscheinlichkeit bestimmt ist. Ein derartiges Verkehrsmodell findet z.B. bei der sogenannten belegungsphasentreuen Simulation (siehe Abschnitt 4.4) Anwendung. Es geht dabei jedoch die Abhängigkeit zwischen den einzelnen Teilrufen eines bestimmten Rufes verloren, so daß keine rufbezogenen Ergebnisse mehr produzierbar sind, wie z.B. Verteilungsfunktionen charakteristischer Systemreaktionszeiten.

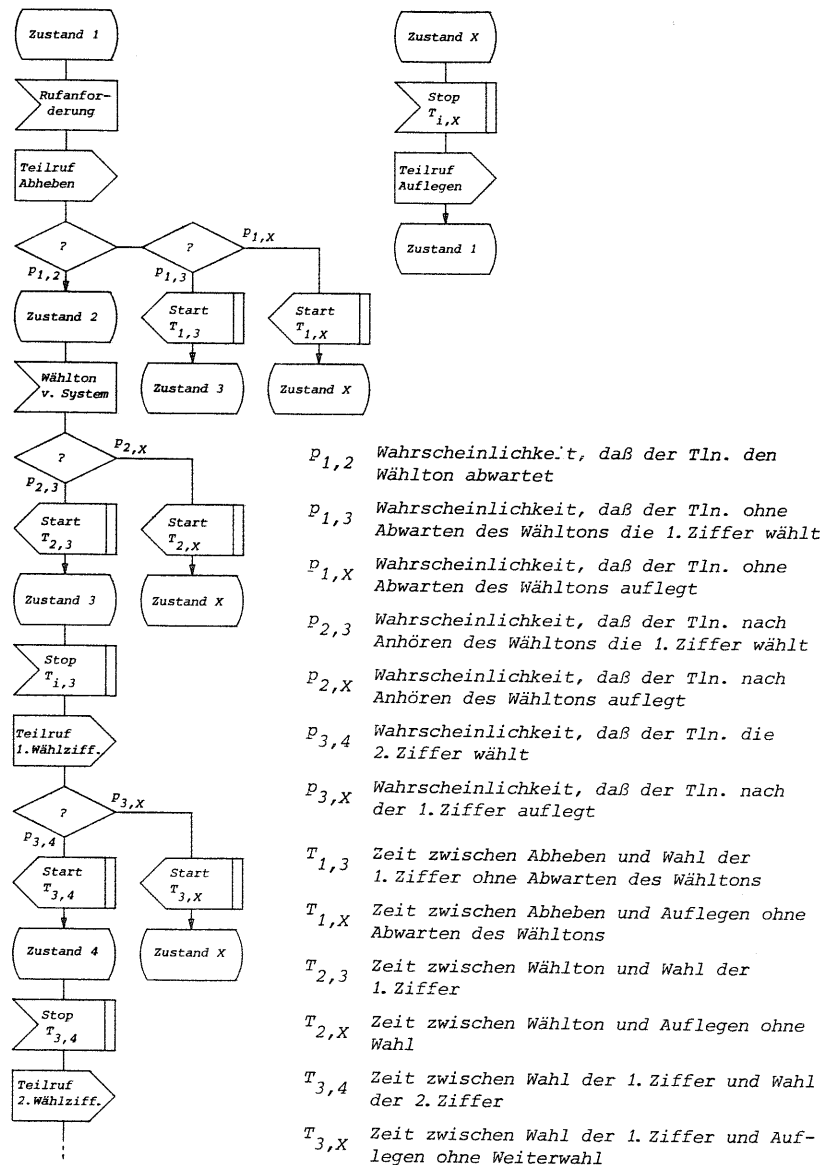


Bild 4.1: Ausschnitt des SDL-Diagramms eines Teilrufprozesses

4.2.1.2 Bedienungsprozesse

Analog zu den zufallsabhängigen Ankunftsprozessen kann die Bedienung von Rufen bzw. Teilrufen in Bedienungseinheiten des Systems durch zufallsabhängige Bedienungsprozesse modelliert werden, die die Bedienungsdauern T_H beschreiben. Bedeutend für die Praxis sind neben der Erlang-k- und der Hyperexponential-Verteilung insbesondere die Exponential-Verteilung

$$F_H(t) = P\{T_H \leq t\} = 1 - e^{-t/h} \quad (4.3)$$

und die konstante Verteilung

$$F_H(t) = P\{T_H \leq t\} = \begin{cases} 0, & t < h \\ 1, & t \geq h \end{cases}, \quad (4.4)$$

mit der mittleren bzw. konstanten Bedienungsdauer als Erwartungswert der Bedienungsdauer

$$h = E[T_H]. \quad (4.5)$$

Bei größerer Modellierungstiefe, z.B. in detaillierten Simulationen, ist es auch möglich, eine vom Typ der Anforderung, z.B. des Teilrufes, abhängige konstante oder zufallsabhängige Bedienungsdauer T_t zu definieren. Dies erfordert jedoch genaue Kenntnisse der individuellen Bedienungsdauern, z.B. ermittelt durch Messungen am realen System oder durch manuelle Laufzeitanalysen der Programme. Dieses Modell läßt sich, ähnlich wie die Teilnehmerereignisketten, durch einen Zustandsprozeß beschreiben, der für jeden möglichen Anforderungstyp t einen Zustandsübergang in der Bedienungseinheit mit der Bedienungsdauer T_t bewirkt. Im einfachsten Fall existiert dabei nur ein Zustand, aber es sind bei noch größerer Modellierungstiefe durchaus mehrere Zustände pro Bedienungseinheit denkbar, denn dann ist der Zustandsprozeß ein direktes Abbild des Realzeitprozesses, der auf der betrachteten Bedienungseinheit läuft. Da Realzeitprozesse in der Vermittlungssoftware meist in SDL spezifiziert sind, bietet sich SDL auch für

das Modell des Zustandsprozesses an. Ein exemplarisches Beispiel dazu ist in Bild 4.2 gezeigt.

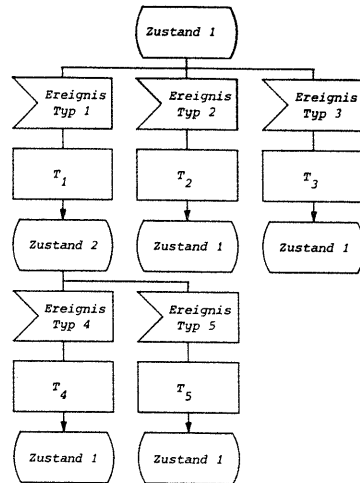


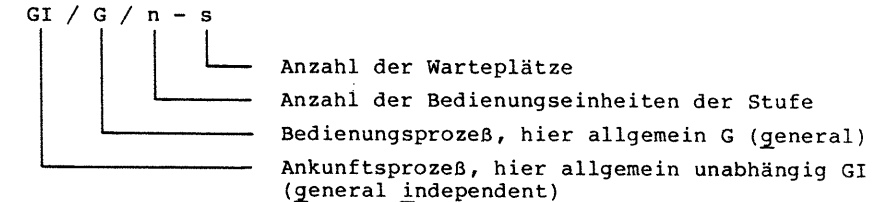
Bild 4.2: Beispiel für einen Bedienungsprozeß in SDL-Darstellung

4.2.2 Systemmodelle

Die wichtigsten strukturbezogenen Elemente von Warteschlangenmodellen sind:

- Verkehrsquellen zur Erzeugung von Anforderungen mit einem Ankunftsprozeß gemäß Abschnitt 4.2.1.1
- Warteschlangen zur Zwischenspeicherung wartender Anforderungen mit unendlicher oder endlicher Anzahl s von Warteplätzen und definierter Abfertigungsstrategie (z.B. FIFO)
- Bedienungseinheiten zur Ausführung der angeforderten Funktionen mit einem Bedienungsprozeß gemäß Abschnitt 4.2.1.2
- Schalter zur zeitweisen Schaltung von Verbindungswegen gemäß definierten Schaltstrategien
- Verzweigungen zur richtungsmäßigen Aufteilung von Anforderungen gemäß definierten Verzweigungsstrategien.

Die bekannteste Grundform eines Modells und die Basis der meisten Gesamt- und Teilmodelle ist das allgemeine, einstufige Warteschlangenmodell, das nach einer erweiterten 1954 von Kendall eingeführten Notation bezeichnet wird:



Falls die Art des Ankunfts- oder des Bedienungsprozesses genauer bekannt ist, können statt der Kennzeichnungen GI bzw. G für den allgemeinen Fall u.a. auch folgende Bezeichnungen verwendet werden:

- M Markoffscher Prozeß (Exponentialverteilung)
- D Determinierter Prozeß (Konstante Verteilung)
- E_k Erlang-Prozeß der Ordnung k
- H_k Hyperexponentieller Prozeß der Ordnung k

4.2.2.1 Gesamtmodell eines Vermittlungssystems

Bild 4.3 zeigt ein verallgemeinertes, strukturelles Gesamtmodell eines Vermittlungssystems mit verteilter Steuerung.

Jede Steuerungseinheit, im folgenden Prozessor genannt, ist durch eine Bedienungseinheit mit einer Eingangswarteschlange und einer Ausgangswarteschlange modelliert. Für jede Funktion ist im allgemeinsten Fall entsprechend dem Funktionsteilungsprinzip eine Gruppe von Prozessoren vorgesehen, nämlich z.B. im peripheren Bereich eine Gruppe peripherer Prozessoren und im zentralen Bereich je eine Gruppe von Prozessoren für allgemein n zentrale Funktionen. Die peripheren Prozessoren empfangen u.a. die ankommenden Rufe bzw. Teilrufe, die durch einen geeigneten Ankunftsprozeß modelliert werden. Jede funktionelle Gruppe von Prozessoren kann

weiterhin noch entsprechend dem Lastteilungsprinzip in mehrere Einzelprozessoren unterteilt sein.

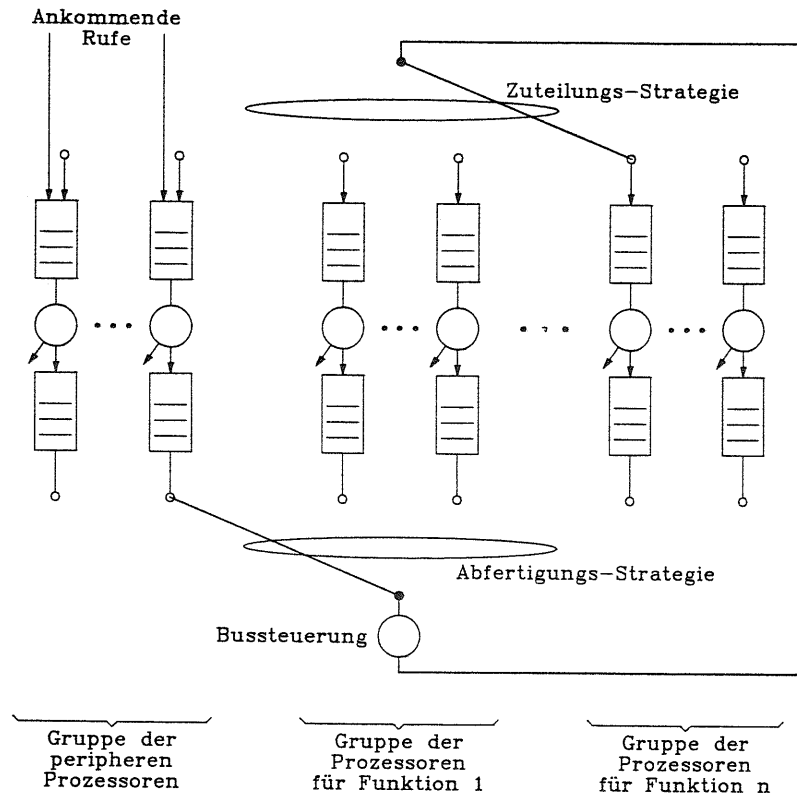


Bild 4.3: Verallgemeinertes Gesamtmodell eines Vermittlungssystems mit verteilter Steuerung

Jede Bedienungseinheit bearbeitet die Anforderungen in ihrer Eingangswarteschlange. Falls während der Bedienungsphase neue Anforderungen an andere Prozessoren gesendet werden sollen, so werden diese in die Ausgangswarteschlange eingetragen, wo sie auf

die Übertragung zu den Zielprozessoren warten. Falls während der Bedienungsphase keine weiteren Anforderungen an andere Prozessoren entstehen, endet der entsprechende Anforderungsstrom, was im Modell durch die von der Bedienungseinheit ausgehende Verzweigung gekennzeichnet ist. Diese Verzweigung kann je nach Modellierungstiefe zufallsabhängig oder determiniert, abhängig vom Typ der Anforderung, angesetzt werden.

Der Transfer der Anforderungen zwischen den Prozessoren erfolgt über ein Kommunikationssystem, z.B. einen Bus, wie im Modell angenommen. Die durch eine einfache Bedienungseinheit modellierte Bussteuerung koordiniert die Übertragung von Anforderungen von den Ausgangswarteschlangen der Ursprungsprozessoren zu den Eingangswarteschlangen der Zielprozessoren. Dieser Transfer lässt sich logisch in zwei Teilaufgaben gliedern, nämlich die Abfertigung der Ausgangswarteschlangen und die Zuteilung der Anforderungen zu den Eingangswarteschlangen, jeweils gemäß einer definierten Strategie. Im Modell ist ersteres durch den unteren Schalter und letzteres durch den oberen Schalter repräsentiert. Die für diese beiden Aufgaben und für die Übertragung selbst benötigte Zeit wird durch die Bedienungsdauer der Bedienungseinheit modelliert, die die Bussteuerung repräsentiert. Sie kann, wie auch die Bedienungsdauern der Prozessoren, je nach Modellierungstiefe probabilistisch oder detaillierter, abhängig vom Typ der Anforderungen oder von bestimmten Zuständen, modelliert werden.

4.2.2.2 Teilmodelle von Vermittlungssystemen

Selbst unter vereinfachten Annahmen für die Ankunfts- und Bedienungsprozesse ist die Analyse von Gesamtmodellen aufwendig und schwierig. Durch eine Dekomposition in kleinere und überschaubare Teilmodelle lassen sich jedoch Teilaspekte des Systems einfacher untersuchen, z.B. analytisch lösen und optimieren. Für die Ankunftsprozesse eines Teilmodells können dann entweder allgemeine Annahmen gemacht werden, oder es kann ein Ansatz Anwendung finden, der die Zwischenankunfts-Verteilungsfunktion annähert, die durch Messung in einer Simulation des Gesamtsystems an der be-

treffenden Stelle ermittelt wurde oder die zuvor als Ausgangsprozess des vorgeschalteten Teilmodells ermittelt wurde.

An dieser Stelle soll nur exemplarisch ein typisches Teilmodell aus dem im vorigen Abschnitt beschriebenen Gesamtmodell aufgezeigt werden. Weitere Teilmodelle sind in den Analysen im fünften Kapitel zu finden. Bild 4.4 zeigt ein Teilmodell, das den Vorgang der Abfertigung der Ausgangswarteschlangen durch die Bussteuerung (Polling) beschreibt.

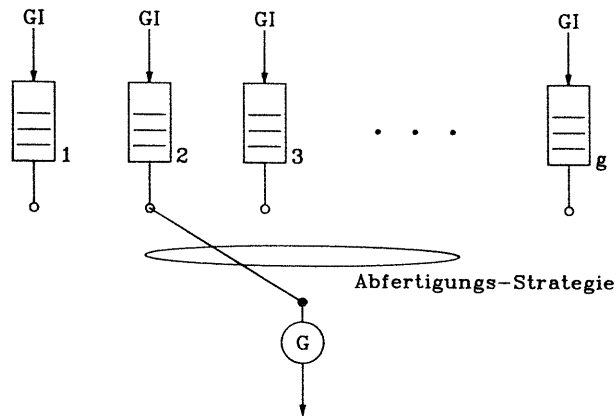


Bild 4.4: Teilmodell für das Polling der Ausgangswarteschlangen

Das Modell besteht aus g Warteschlangen endlicher Kapazität mit allgemeinen Ankunftsprozessen GI und i.a. unterschiedlichen Intensitäten der g Ankunftsströme. Ein gemäß einer definierten Abfertigungs-Strategie umlaufender Schalter ermöglicht jeweils den Transfer von Anforderungen, wobei die Transferzeit durch einen allgemeinen Bedienungsprozeß G modelliert sei. In praktischen Anwendungsfällen in Vermittlungssteuerungen können die allgemeinen Prozesse oft durch eine exponentielle und im Falle konstanter Meldungslänge der Bedienungsprozeß auch durch eine konstante Verteilungsfunktion modelliert werden. Messungen und Ergebnisse sehr

detaillierter Simulationen belegen, daß der Variationskoeffizient der Ankunftsabstände nahe bei eins liegt, d.h. daß der approximative Ansatz einer Exponentialverteilung für den Ankunftsprozeß der Teilmodelle von Vermittlungssystemen i.a. gerechtfertigt ist. In realen Polling-Systemen benötigt auch der Umschaltvorgang zu einer anderen Warteschlange eine gewisse Umschaltzeit, deren Berücksichtigung im Modell bei entsprechender Größe im Verhältnis zur Transferzeit notwendig ist. Analysen derartiger Polling-Systeme für zyklische Abfertigung sind z.B. in /39/ und /40/ beschrieben, während eine Studie über zustandsabhängige Abfertigungsstrategien im fünften Kapitel dieser Arbeit behandelt wird.

4.3 Signalflußanalyse

Die mittleren Belastungen der Teilsysteme von verteilten Vermittlungssteuerungen, d.h. der Verarbeitungs- und Übertragungseinrichtungen, sowie die Beeinflussung der Belastungen durch verschiedene Belegungsarten und Dienste lassen sich einfach und bereits in einem frühen Entwicklungsstadium mit einer Signalflußanalyse ermitteln.

Prinzipiell wird aus den Meldungsszenarien der einzelnen Belegungsarten und deren Auftretswahrscheinlichkeiten gemäß dem Belegungsspektrum für das jeweils betrachtete Teilsystem die gesamte für einen Ruf aufzuwendende Bearbeitungs- bzw. Übertragungszeit ermittelt, woraus dann mit der globalen Rufankunftsrate die mittlere Belastung des Teilsystems berechnet werden kann. Diese Belastung stellt sich in Realität jedoch nur ein, sofern im System keine Verluste auftreten und keine Überlastabwehr den Rufankunftsstrom drosselt. Bei Hochlast kann sie deshalb nur ein Schätzwert sein, bei niedrigem und normalem Verkehrsangebot ist sie aber hinreichend genau.

Signalflußanalysen bieten sich an für Studien über neue Strukturen und Systemkonfigurationen sowie auch für die Untersuchung des Einflusses neuer Aufgabenverteilungen, neuer Dienste und der verschiedener Fälle des Teilnehmerverhaltens. Aufgrund des Umfangs der notwendigen Meldungsszenarien sollte trotz der Einfachheit des Verfahrens Rechnerunterstützung in Anspruch genommen werden.

4.3.1 Vorgaben für die Signalflußanalyse

Die notwendigen Vorgaben für eine Signalflußanalyse sind:

- das Modell der Steuerungsstruktur des betrachteten Systems, bestehend aus Verarbeitungs- und Übertragungseinrichtungen
- die Belegungsarten und deren Auftretswahrscheinlichkeiten im Belegungsspektrum für alle relevanten Fälle

- die Meldungsszenarien aller betrachteten Belegungsarten, mit geschätzten oder gemessenen, individuellen oder pauschalen Meldungsbedienungs- bzw. Übertragungszeiten.

4.3.1.1 Steuerungsstruktur

Prinzipiell bestehen bei der Signalflußanalyse bezüglich der Struktur des Systems keine Einschränkungen. Es ist jedoch zweckmäßig, sich bei der Implementierung eines universellen Software-Werkzeuges für Signalflußanalysen auf maximal dreistufige Strukturen (siehe Abschnitt 2.2.2.3) zu beschränken. Dann besteht das allgemeinste und gleichzeitig auch typische Strukturmodell, gemäß Bild 4.5, aus zentralen Verarbeitungseinrichtungen ZVE, über eine zentrale Übertragungseinrichtung ZÜE gekoppelte Gruppen-Verarbeitungseinrichtungen GVE und über periphere Übertragungseinrichtungen PÜE an die GVE gekoppelte periphere Verarbeitungseinrichtungen PVE.

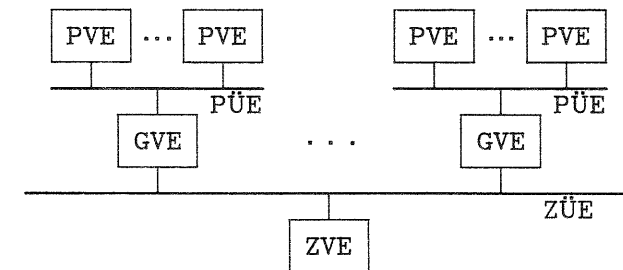


Bild 4.5: Allgemeine teilzentrale dreistufige Steuerungsstruktur

Eine Struktur gemäß Bild 4.5 könnte für eine Signalflußanalyse z.B. wie in Tabelle 4.1 Rechner-gerecht spezifiziert werden. Das Beispiel beschreibt eine Struktur mit einer ZVE und einer Koppelnetzsteuerung ZVE-KN in der zentralen Ebene und vier Gruppen-Verarbeitungseinrichtungen GVE in der Gruppen-Ebene. Die Peripherie der ersten drei Gruppen besteht aus je 50 PVE für die Steuerung von Teilnehmeranschlüssen (-TLN), und

die vierte Gruppe ist mit je 10 PVE zur Steuerung gehender (-UEG) bzw. kommender (-UEK) Leitungsübertragungen beschaltet.

Hierarchieebene	UE	Anzahl	VE	Anzahl
ZENTRAL	ZÜE	1	ZVE ZVE-KN	1 1
GRUPPE(1..3)	PÜE	1	GVE	1
GRUPPE(4)	PÜE	1	GVE	1
PERIPHERIE(1..3)	-		PVE-TLN	50
PERIPHERIE(4)	-		PVE-UEG	10
			PVE-UEK	10

Tabelle 4.1: Beispiel für eine Rechner-gerechte Spezifikation einer dreistufigen Steuerungsstruktur

4.3.1.2 Belegungsarten und Belegungsspektrum

Für eine Signalflußanalyse müssen alle relevanten Belegungsarten betrachtet werden, d.h. alle Arten vollständiger und unvollständiger Belegungsversuche.

Die wichtigsten Belegungsarten der einzelnen Dienste sind z.B.:

- Internrufe
- Abgehende Ortsrufe
- Abgehende Fernrufe
- Ankommende Ortsrufe
- Abgehende Fernrufe.

Ein typisches Belegungsspektrum besteht z.B. aus den Fällen:

- Erfolgreiche Rufe
- Rufe mit abgebrochener Wahl

- Nach Wahlende gassenbesetzte Rufe
- Rufe, bei denen der B-Tln besetzt ist
- Rufe, bei denen sich der B-Tln nicht meldet.

Für die Signalflußanalyse wird dann die Auftrittswahrscheinlichkeit p_c jeder Belegungsart c des Belegungsgemisches benötigt.

4.3.1.3 Meldungsszenarien

Bild 4.6 zeigt mittels eines Signalflußgraphen exemplarisch einen Ausschnitt aus einem Meldungsszenario, dem eine dreistufige Steuerungsstruktur gemäß Bild 4.5 zugrunde liegt.

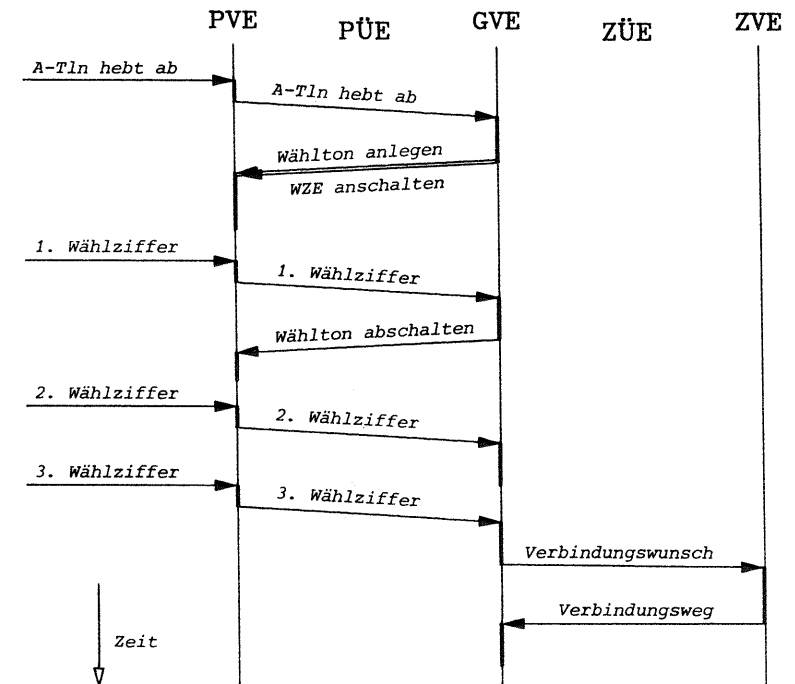


Bild 4.6: Ausschnitt eines Signalflußgraphen für ein Meldungsszenario

Jedes Teilnehmerereignis bzw. jede Meldung hat eine Vermittlungsfunktion zur Folge, die in der jeweiligen Verarbeitungseinrichtung eine gewisse Verarbeitungszeit h_v und zur Übertragung auf der betreffenden Übertragungseinrichtung eine gewisse Übertragungszeit h_u benötigt.

Derartige Meldungsszenarien müssen für alle zu betrachtenden Belegungsarten spezifiziert sein. Für eine rechnerunterstützte Signalflußanalyse wäre zu diesem Zweck prinzipiell eine textuelle Version von SDL geeignet, mit der üblicherweise auch die Software von SPC-Systemen spezifiziert wird. Einfacher zu behandeln ist jedoch pro Belegungsart eine tabellarische Darstellung, z.B. wie in Tabelle 4.2 für den Verbindungsaufbau eines erfolgreichen Internrufs exemplarisch gezeigt. Diese Tabellen sind zwar insgesamt sehr umfangreich, aber gewisse Teile wiederholen sich bei den einzelnen Belegungsversuchen immer wieder, so daß sie mittels eines Text-Editors doch relativ schnell erstellt werden können. Falls die Meldungsbedienungszeiten h_v bzw. Übertragungszeiten h_u zunächst pauschale Werte haben sollen, können sie z.B. auch durch Variablennamen ersetzt werden, denen dann später Werte zugewiesen werden.

Aus dem Beispiel, dem eine Steuerungsstruktur gemäß Tabelle 4.1 zugrunde liegt, wird deutlich, daß in der mittleren Hierarchieebene die eigentlich gleichartigen GVE und PÜE unterschieden werden müssen, da die Gruppen in der peripheren Ebene unterschiedlich beschaltet sind. Bei dem Internruf ist die vierte Gruppe der Beispielstruktur unbeteiligt, da sie nur PVE-UEG bzw. PVE-UEK zur Steuerung von Leitungsübertragungen enthält. Die Kennzeichnung der jeweils betroffenen GVE bzw. PÜE erfolgt im Meldungsszenario am zweckmäßigsten wie bei den PVE durch Bezeichnungszusätze wie z.B. -TLN, -UEG oder -UEK.

von VE	zu VE	über UE	h_u /ms	h_v /ms	Kommentar
-	PVE-TLN	-	-	1.5	Teilnehmerereignis
PVE-TLN	GVE-TLN	PÜE-TLN	1	3	A-Tln hebt ab
GVE-TLN	PVE-TLN	PÜE-TLN	1	2	Wählton anlegen
GVE-TLN	PVE-TLN	PÜE-TLN	1	2	WZE anschalten
-	PVE-TLN	-	-	1.5	Teilnehmerereignis
PVE-TLN	GVE-TLN	PÜE-TLN	1	3	1. Wählziffer
GVE-TLN	PVE-TLN	PÜE-TLN	1	2	Wählton abschalten
-	PVE-TLN	-	-	1.5	Teilnehmerereignis
PVE-TLN	GVE-TLN	PÜE-TLN	1	3	2. Wählziffer
-	PVE-TLN	-	-	1.5	Teilnehmerereignis
PVE-TLN	GVE-TLN	PÜE-TLN	1	3	3. Wählziffer
GVE-TLN	ZVE	ZÜE	0.1	4	Verbindungswunsch
ZVE	GVE-TLN	ZÜE	0.1	3	Verbindungsweg
ZVE	ZVE-KN	ZÜE	0.1	7.5	KN-Weg einstellen
GVE-TLN	GVE-TLN	ZÜE	0.1	3	Belegungswunsch
GVE-TLN	PVE-TLN	PÜE-TLN	1	2	Rufstrom anlegen
GVE-TLN	GVE-TLN	ZÜE	0.1	3	B-Tln ist frei
GVE-TLN	PVE-TLN	PÜE-TLN	1	2	Freiton anlegen
GVE-TLN	PVE-TLN	PÜE-TLN	1	2	WZE abschalten
-	PVE-TLN	-	-	1.5	Teilnehmerereignis
PVE-TLN	GVE-TLN	PÜE-TLN	1	3	B-Tln hebt ab
.
.

Tabelle 4.2: Ausschnitt einer Rechner-gerechten Spezifikation des Meldungsszenarios eines erfolgreichen Internrufs

4.3.2 Prinzip der Signalflußanalyse

Der Signalflußanalyse liegen folgende Überlegungen zugrunde. Betrachtet werde eine beliebige Meldung i aus der Gesamtheit der Meldungsszenarien aller Belegungsarten. $c(i)$ sei die Belegungsart, zu der diese Meldung i gehört. Die Meldung i tritt dann pro

Ruf mit der Auftrittswahrscheinlichkeit $p_c(i)$ ihrer Belegungsart $c(i)$ auf. Ferner betrifft sie aber auch nur mit einer bestimmten Wahrscheinlichkeit $p_k(i)$ ein betrachtetes Teilsystem k , d.h. sie gehört mit dieser Wahrscheinlichkeit zur Untermenge der Meldungen, die in das Teilsystem k eintreten. Somit erfährt die Meldung i mit der Wahrscheinlichkeit $p_c(i) \cdot p_k(i)$ im Teilsystem k eine Bedienung mit konstanter oder mittlerer Bedienungszeit $h(i)$. Initiiert wird die Meldung i jedoch durch Rufe mit der Ankunftsrate λ . Die mittlere Belastung Y_k eines jeden Teilsystems k ergibt sich dann zu

$$Y_k = \lambda \cdot \sum_i p_c(i) \cdot p_k(i) \cdot h(i) \quad (4.6)$$

Die durch die Belegungsart j verursachte mittlere Teilbelastung $Y_{k,j}$ eines jeden Teilsystems k ist

$$Y_{k,j} = \lambda \cdot \sum_{i|c(i)=j} p_c(i) \cdot p_k(i) \cdot h(i) \quad (4.7)$$

Damit ergibt sich der durch eine Rufart j verursachte Belastungsanteil der Belastung des Teilsystems k zu $Y_{k,j} / Y_k$.

Die Wahrscheinlichkeit $p_k(i)$ ist dabei wie folgt zu bestimmen:

$$p_k(i) = \begin{cases} 0 & , \text{ falls das Teilsystem } k \text{ nicht} \\ & \text{Ziel-VE bzw. UE der Meldung } i \text{ ist.} \\ 1/n_k & , \text{ falls das Teilsystem } k \text{ die Ziel-} \\ & \text{ZVE oder Ziel-PVE bzw. die ZUE} \\ & \text{der Meldung } i \text{ ist.} \\ G/n_k & , \text{ falls das Teilsystem } k \text{ die Ziel-} \\ & \text{GVE bzw. die PUE der Meldung } i \text{ ist.} \end{cases} \quad (4.8)$$

n_k ist dabei die Anzahl gleichartiger, eventuell paralleler Teilsysteme vom Typ des Teilsystems k im Gesamtsystem.

G ist ein Gewichtungsfaktor, der die Bewertung der Meldung i für

Teilsysteme der mittleren Hierarchieebene angibt, die aufgrund der Beschaltung in der peripheren Ebene erforderlich ist. Im betrachteten Fall betrifft die Meldung i im Teilsystem k eine Funktion, die nur durch die Beschaltung mit der speziellen Art PVE bedingt ist, die den Funktionsbezug herstellt. G ist das Verhältnis der Anzahl all der betreffenden gleichartigen, am Teilsystem k angeschlossenen PVE, zur Anzahl aller derartigen PVE im Gesamtsystem.

Die Bestimmung des Faktors G sei an folgendem Beispiel verdeutlicht. Es sei u.a. eine GVE(1) mit 50 PVE-TLN und eine GVE(2) mit 40 PVE-TLN für die Steuerung von Teilnehmeranschlüssen beschaltet. Die Gesamtanzahl der PVE-TLN im System sei 200. Ist nun die Ziel-Verarbeitungseinrichtung einer Meldung i gemäß oben vorgeschlagener Nomenklatur eine GVE-TLN, so ergibt sich der Faktor G für die Bestimmung der Belastung der GVE(1) zu $50/200 = 0,25$ und für die GVE(2) zu $40/200 = 0,2$.

4.4 Simulation

Die Nachbildung dynamischer Vorgänge in Vermittlungssteuerungen auf einem Rechner erfolgt i.a. mittels einer zeittreuen, ereignisgesteuerten Simulationstechnik (engl. time true, event-by-event simulation), da als Ergebnisse nicht nur Belastungen, sondern auch Zeitgrößen von Interesse sind. Je nach Modellierungstiefe und Anzahl der gewünschten Ergebnisparameter stehen drei unterschiedlich aufwendige, zeittreue Verfahren zur Auswahl:

- die belastungstreue Simulation (engl. load-type simulation)
- die belegungsphasentreue Simulation (engl. subcall-type simulation)
- die belegungstreue Simulation (engl. call-type simulation).

Die Unterschiede der Verfahren sind in /18/ näher erläutert. Das Verfahren der belegungstreuen Simulation ist das genaueste und aufwendigste. Für große Simulationsstudien, wie z.B. die im fünften Kapitel beschriebene Simulation des Systems PILOT, muß i.a. aus Aufwandsgründen das zweitgenannte Verfahren Anwendung finden. In den folgenden Abschnitten soll jedoch lediglich nach einer kurzen Beschreibung des Grundprinzips der zeittreuen Simulation eine Erweiterung für die beiden letztgenannten Verfahren auf der Basis SDL-spezifizierter Prozesse erläutert werden.

4.4.1 Grundprinzip der zeittreuen Simulation

Grundsätzlich wird bei der zeittreuen Simulation das Ablaufgeschehen durch die Abfolge der Ereignisse charakterisiert, die die Veränderungen im System bewirken. Derartige Ereignisse sind im wesentlichen die Ankünfte von Anforderungen an das System und die Beendigung von Bedienungsphasen dieser Anforderungen im System. Das Simulationsprogramm "springt" dann jeweils zeitlich von einem Ereignis zum nächsten (daher engl. event-by-event), indem es die notwendigen Veränderungen des Systems für das momentan bearbeitete Ereignis ausführt, den Zeitpunkt des nächsten Ereignisses derselben und eventuell auch einer anderen Art bestimmt, diese Zeitpunkte in einen sogenannten "Kalender" einträgt und dann zum

zeitlich nächsten Ereignis springt, d.h. im Kalender den Zeitpunkt des zeitlich nächsten Ereignisses sucht, die momentane Zeit auf diesen Zeitpunkt aktualisiert und das neue Ereignis wieder wie bekannt ausführt.

Die Bestimmung des Zeitpunktes eines nächsten Zufallsereignisses, z.B. des nächsten Ankunftszeitpunktes, besteht in der Ermittlung eines zufallsbehafteten Ankunftsabstandes T_A , dessen Verteilungsfunktion $F_A(t)$ vorgegeben ist. Praktisch erfolgt dies durch die Transformation einer gleichverteilten Pseudo-Zufallszahl aus dem Intervall zwischen 0 und 1 durch die Verteilungsfunktion $F_A(t)$. Die Erzeugung von Pseudo-Zufallszahlen erfolgt mittels eines Software-Zufallszahlengenerators, wie z.B. in /42/ vorgeschlagen.

In einem Simulationsprogramm werden gleichzeitig neben dem zeitlich gerauschten Ereignis-Ablauf bei jeder Veränderung des Systems bzw. eines Teilsystems alle interessierenden Meßdaten erfaßt und für eine spätere statistische Auswertung akkumuliert abgespeichert. Ein Simulationslauf wird üblicherweise nach einem Vorlauf zur Erreichung eines eingeschwungenen Zustandes in einige Meßintervalle oder Teiltests mit z.B. einigen Tausend Ankunftsereignissen unterteilt, nach denen jeweils zur Gewinnung statistischer Stichprobenwerte die akkumulierten Meßdaten ausgewertet werden. Am Ende des Simulationslaufes können dann aus den Stichprobenwerten der Teiltests die interessierenden charakteristischen Größen des Systems (Belastungen, Blockierungswahrscheinlichkeiten, Wartezeiten usw.) und ihre statistische Aussagesicherheit z.B. in Form von Vertrauensintervallen nach der in /41/ beschriebenen Methode errechnet werden.

4.4.2 Zeittreue Simulation SDL-spezifizierter Prozesse

Bisher erfolgte die Beschreibung der Handlungsweise von Simulationsprogrammen bei konventionellen belegungsphasen- oder belegungstreuen Simulationen zum Teil fest programmiert und mittels umfangreicher Ereignistabellen für viele unterschiedliche Betriebsfälle. Die Bedienungsprozesse im Systemmodell sind jedoch

bei größtmöglicher Modellierungstiefe identisch mit den Vermittlungsprozessen, und für die Spezifikation von Vermittlungsprozessen hat sich seit einiger Zeit die Sprache SDL weit etabliert. Ferner wurde im Abschnitt 4.2.1.1 gezeigt, daß auch die Modelle für die Ankunftsprozesse bei größerer Modellierungstiefe vorteilhaft in SDL beschreibbar sind. Somit liegt es nahe, auch das Ablaufgeschehen in einem Simulationsprogramm in SDL zu spezifizieren, d.h. die Handlungsweise von Ereignisgeneratoren und Bedienungseinheiten im Simulationsmodell durch übersichtliche, SDL-gemäße Datensätze flexibel vorzugeben, die ohnehin teilweise schon in Form der in SDL spezifizierten Vermittlungsprozesse vorliegen.

Zur Realisierung eines derartigen Simulationsprogrammes /43/ ist für jeden vorkommenden Typ von Ankunfts- oder Bedienungsprozessen ein Speicherabbild seines SDL-Graphen vorzusehen, für dessen Eingabe sich z.B. eine textuelle Darstellung von SDL eignet. Falls ein Prozeßtyp dann in mehrfacher Verkörperung existiert, genügt jeweils ein Zustandsspeicher, der im SDL-Graphen des Prozeßtyps die momentan bearbeitete Stelle angibt, z.B. mittels eines auf das entsprechende SDL-Symbol verweisenden Zustandszeigers. Das Simulationsprogramm bearbeitet dann alle Prozesse laut der SDL-Vorschrift schrittweise ab, indem es jedes SDL-Symbol definitionsgemäß bearbeitet und danach den Zustandszeiger gemäß dem SDL-Graphen auf das SDL-Folgesymbol aktualisiert. Für die wesentlichen SDL-Symbole können dann die folgend beschriebenen Bedeutungen und die zugehörigen Bearbeitungsregeln definiert werden:

- TASK :

Die TASK-Anweisungen in SDL-Programmen repräsentieren die Phasen der Ankunftsabstände bzw. Bedienungsdauern und sind mit Ausnahme des Wartens bei der STATE-Anweisung für die Simulation die einzigen zeitbehafteten SDL-Elemente. Das Ende einer TASK entspricht folglich einem Simulationereignis und muß im Kalender als solches mit seinem Ereignis-Zeitpunkt und der zugehörigen Prozeß-Kennung abgelegt sein. Zusätzlich muß im Kalender vermerkt sein, um welche TASK innerhalb des Prozesses

es sich handelt, z.B. mittels eines Zeigers auf das Folgesymbol des entsprechenden TASK-Symbols im Speicherabbild des zugehörigen SDL-Graphen. Tritt nun im Simulationsverlauf ein derartiges Ereignis ein, indem es im Kalender als zeitlich nächstes gefunden wird, so können mit Hilfe dieses Zeigers sofort die nächsten auszuführenden SDL-Symbole ermittelt werden. Bei der Bearbeitung einer TASK ist eine neue Belegung der Bedienungseinheit erforderlich. Dabei wird die als Attribut der TASK-Anweisung festgelegte Zeitdauer (eventuell gemäß einer gegebenen Verteilungsfunktion) ermittelt und deren Ablaufzeitpunkt, deren Prozeß-Kennung und ein Zeiger auf das Folgesymbol im SDL-Graphen in den Kalender eingetragen. Die Fortsetzung der Simulation geschieht durch Suchen des nächsten Ereignisses im Kalender.

- STATE :

Bei einem STATE-Symbol muß die Eingangswarteschlange für den betreffenden Prozeß auf alle im SDL-Graphen nach dem STATE-Symbol gegebenen Alternativen von Anforderungen (INPUT-Symbole) abgefragt werden, sofern die Bedienungseinheit nicht durch Anforderungen höherer Priorität belegt ist. Kann keine Belegung der Bedienungseinheit erfolgen, so wird der Prozeß in einen Wartezustand versetzt und bleibt auf dem STATE-Symbol stehen. Die Fortsetzung der Simulation geschieht dann durch Suchen des nächsten Ereignisses im Kalender. Andernfalls wird der SDL-Graph, ausgehend vom STATE-Symbol, bis zur nächsten zeitbehafteten TASK abgearbeitet und dort dann die Bedienungseinheit erneut belegt.

- INPUT :

Ein INPUT-Symbol repräsentiert eine Anforderung eines bestimmten Typs. Die INPUT-Anweisung enthält als Attribut den Anforderungstyp, z.B. den Teilruftyp oder eine Klassenkennung. Die Behandlung des INPUT-Symbols erfolgt in gleicher Weise wie beim STATE-Symbol beschrieben.

- OUTPUT :

Die OUTPUT-Anweisung dient zum Senden von Anforderungen an andere Prozesse. Dabei wird die Anforderung mit ihrem Anforderungstyp als Attribut in eine entsprechende Warteschlange des Zielprozesses eingetragen. Falls sich der Zielprozeß in einem Wartezustand, d.h. auf einem STATE-Symbol, befindet, muß unter Umständen direkt eine Belegung der Zielbedienungseinheit erfolgen, indem deren SDL-Graph, ausgehend von diesem STATE-Symbol, bis zur nächsten zeitbehafteten TASK abgearbeitet wird. Ansonsten wird sofort die nächste Anweisung gemäß SDL-Graph abgearbeitet.

- CONDITION :

Aufgrund der im Attribut der Anweisung festgelegten Bedingung erfolgt eine Verzweigung im SDL-Graphen. Anschließend wird sofort die nächste Anweisung gemäß SDL-Graph abgearbeitet. Bei der Simulation von Protokollen kann die Einführung von Prozeßvariablen von Bedeutung sein. Die Bedingung der CONDITION-Anweisung kann z.B. durch den Vergleich zweier Prozeßvariablen definiert werden. Es sollte dann bei der Bearbeitung von TASKS auch die Zuweisung von arithmetischen Ausdrücken aus Variablen zu Variablen ermöglicht werden, wie auch das Senden von Variablen zu anderen Prozessen als Attribut von INPUT-Anweisungen.

4.5 Warteschlangenmodellanalyse

Die analytische Behandlung von Warteschlangenmodellen basiert i.a. auf der Betrachtung der Ankunfts- und Bedienungsprozesse als gegebene zufallsabhängige (stochastische) Prozesse. Eine besondere Stellung nimmt dabei die Klasse der Markoffschen Prozesse ein, da diese oft eine gute Näherung für die Realität darstellen, die Analyse Markoffscher Systeme /44,37/ exakt und meist relativ einfach möglich ist und auch die Grundlage für die Analyse Nichtmarkoffscher Systeme bildet, z.B. nach der Methode von Cox /45,37/. Ferner finden in der Praxis auch zahlreiche approximative Analysemethoden /33/ Anwendung. Das Ziel all dieser Techniken ist jedoch, die das Ablaufgeschehen im System charakterisierenden stochastischen Prozesse zu ermitteln, nämlich i.a. den Systemzustandsprozeß und auch z.B. den Warteprozeß oder den Ausgangsprozeß.

Im folgenden wird zunächst exemplarisch der Begriff des Zustandsprozesses eingeführt. Nach einer Definition der Markoff-Eigenschaft folgt dann ein Abschnitt, der sich auf eine Beschreibung der Analyse Markoffscher Systeme beschränkt, insbesondere für den stationären Fall und im Hinblick auf die Anwendung auf Teilmodelle von Vermittlungssteuerungen im fünften Kapitel dieser Arbeit.

4.5.1 Zustandsprozesse

Betrachtet werde ein allgemeines Warteverlustsystem vom Typ GI/G/1-s mit einer Bedienungseinheit und einer Warteschlange der Kapazität $s = 2$. Die Zufallsvariable $X(t) = x$ beschreibe den Systemzustand, d.h. den zeitlichen Verlauf der Anzahl x der Anforderungen im System. In Bild 4.7 ist prinzipiell ein Verlauf von $X(t)$ gezeigt, wobei die durch den Ankunftsprozeß festgelegten Ankunftszeitpunkte $t_{A,i}$ und die den Ausgangsprozeß kennzeichnenden Bedienungsendezeitpunkte $t_{E,j}$ jeweils durch Pfeile markiert sind. Zu jedem Ankunftszeitpunkt $t_{A,i}$ erhöht sich $X(t)$ um eins, sofern das System nicht schon mit $x = s + 1 = 3$ Anforderungen voll belegt ist (alle s Wartepplätze und die Bedienungseinheit belegt).

Ist letzteres aber der Fall, so wird eine eintreffende Anforderung blockiert und abgewiesen. Zu jedem Bedienungsendezeitpunkt $t_{E,j}$ erniedrigt sich $X(t)$ wieder um eins, und eine eventuell wartende Anforderung mit der bisherigen Wartezeit T_W rückt aus der Warteschlange in die Bedienungseinheit vor, um dort die Bedienungsdauer T_H zu erfahren.

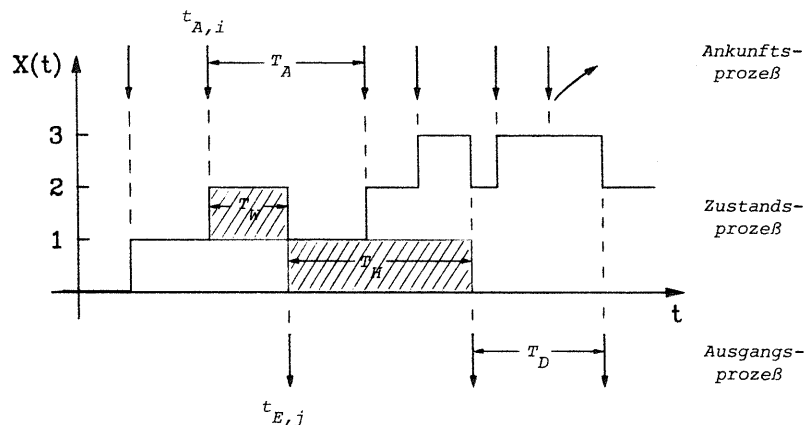


Bild 4.7: Zustandsprozeß eines Warteverlustsystems GI/G/1-s, s=2

4.5.2 Markoffsche Prozesse

Einen stochastischen Prozeß $X(t)$ nennt man einen Markoffschen Prozeß, wenn seine zukünftige Entwicklung nicht vom vergangenen Prozeßverlauf, sondern nur vom gegenwärtigen Prozeßzustand abhängt. Diese Eigenschaft der "Gedächtnislosigkeit" wird für alle Zeitpunkte t_i ($t_0 < t_1 < t_2 < \dots < t_n < t_{n+1}$) durch folgende Gleichung definiert:

$$\begin{aligned} P\{X(t_{n+1})=x_{n+1} \mid X(t_n)=x_n, \dots, X(t_0)=x_0\} \\ = P\{X(t_{n+1})=x_{n+1} \mid X(t_n)=x_n\}. \end{aligned} \quad (4.9)$$

Dabei stellt x_i den Zustand des Prozesses $X(t)$ zur Zeit t_i dar.

Bezieht man die Markoff-Eigenschaft z.B. auf einen Ankunftsprozeß, der zum Zeitpunkt t_1 betrachtet werde, so gilt:

$$P\{T_A > t_2 \mid T_A > t_1\} = P\{T_A > t_2 - t_1\}, \quad (4.10)$$

wobei die letzte Anforderung zur Zeit $t = 0$ eingetroffen sei und t_2 der Eintreffzeitpunkt der nächsten Anforderung ist. Dies bedeutet, daß zu jedem Betrachtungszeitpunkt die Zeit bis zum Eintreffen der nächsten Anforderung unabhängig davon ist, wann die letzte eingetroffen ist.

Die einzige Verteilungsfunktion der Ankunftsabstände T_A , die der Markoff-Eigenschaft genügt, ist die negativ exponentielle Verteilungsfunktion (siehe Gleichung 4.1). Analoges gilt auch für Bedienungsprozesse bezüglich der Verteilungsfunktion der Bedienungsdauer T_H (siehe Gleichung 4.3).

4.5.3 Analyse Markoffscher Systeme

4.5.3.1 Zustandswahrscheinlichkeiten

Für einen durch $X(t)$ beschriebenen Systemzustandsprozeß mit der Markoffschen Eigenschaft betrachte man zunächst die Wahrscheinlichkeit für einen Zustandsübergang von einem Zustand i zu einem um die Zeitdauer u späteren Zustand j

$$p_{ij}(u) = P\{X(t+u) = j \mid X(t) = i\}. \quad (4.11)$$

Der Prozeß sei homogen, d.h. p_{ij} sei unabhängig von der absoluten Zeit t . Die Wahrscheinlichkeit für einen Übergang von einem Zustand i über alle beliebigen Zwischenzustände k zum Zustand j in zwei aufeinander folgenden Zeitspannen t und u wird durch die Chapman-Kolmogoroff-Gleichung /46/ beschrieben

$$p_{ij}(t+u) = \sum_k p_{ik}(t) \cdot p_{kj}(u). \quad (4.12)$$

Für den Grenzübergang $u \rightarrow 0$ erhält man daraus dann die Kolmogoroffsche Vorwärtsgleichung

$$\frac{d}{dt} p_{ij}(t) = -q_j \cdot p_{ij}(t) + \sum_{k \neq j} q_{kj} \cdot p_{ik}(t) \quad (4.13a)$$

mit den sogenannten Übergangsraten

$$q_j = \lim_{u \rightarrow 0} \frac{1 - p_{jj}(u)}{u} \quad \text{und} \quad q_{kj} = \lim_{u \rightarrow 0} \frac{p_{kj}(u)}{u}, \quad (4.13b)$$

die nur vom gegenwärtigen Zustand abhängen, worin sich die Markoffsche Eigenschaft äußert.

Als nächstes betrachte man die Zustandswahrscheinlichkeiten

$$p_j(t) = P\{X(t) = j\}, \quad j = 0, 1, 2, \dots \quad (4.14)$$

Sie lassen sich aus den Anfangswahrscheinlichkeiten $p_j(0)$ zum Zeitpunkt $t = 0$ über das Gesetz der totalen Wahrscheinlichkeit berechnen

$$p_j(t) = \sum_i p_i(0) \cdot p_{ij}(t) \quad (4.15)$$

Mit Gleichung 4.13a ergibt sich dann die Kolmogoroffsche Vorwärtsgleichung der Zustandswahrscheinlichkeiten

$$\frac{d}{dt} p_j(t) = -q_j \cdot p_j(t) + \sum_{k \neq j} q_{kj} \cdot p_k(t) \quad (4.16)$$

mit den Übergangsraten gemäß Gleichung 4.13b. Mit Hilfe der Normierungsbedingung

$$\sum_j p_j(t) = 1 \quad (4.17)$$

lassen sich dann aus den Differentialgleichungen 4.16 alle zeit-

abhängigen Zustandswahrscheinlichkeiten berechnen.

Wenn jedoch, wie in vielen Anwendungsfällen, nur der stationäre Fall (eingeschwungener Systemzustand) von Interesse ist, kann in Gleichung 4.16 der Grenzübergang $t \rightarrow \infty$ durchgeführt werden. Somit entstehen aus den zeitabhängigen Zustandswahrscheinlichkeiten $p_j(t)$ die stationären Zustandswahrscheinlichkeiten p_j und aus dem Kolmogoroffschen Differentialgleichungssystem wird ein lineares Kolmogoroffsches Gleichungssystem für den stationären Fall

$$q_j \cdot p_j = \sum_{k \neq j} q_{kj} \cdot p_k, \quad j = 0, 1, 2, \dots \quad (4.18)$$

Diese Gleichung beschreibt das "statistische Gleichgewicht", welches besagt, daß in der infinitesimal kleinen Zeitspanne dt die Wahrscheinlichkeit für das Verlassen des Zustands j (Rate q_j) gleich der Summe der Wahrscheinlichkeiten für das Erreichen dieses Zustands aus allen anderen Zuständen k (Rate q_{kj}) ist.

Zusammen mit der Normierungsbedingung der stationären Zustandswahrscheinlichkeiten

$$\sum_j p_j = 1 \quad (4.19)$$

lassen sich dann durch Lösen des linearen Gleichungssystems 4.18 die stationären Zustandswahrscheinlichkeiten p_j ermitteln.

4.5.3.2 Lösungsschritte

Die zur stationären Untersuchung eines Markoffschen Systems notwendigen fünf Lösungsschritte sind im folgenden prinzipiell und exemplarisch anhand des Warteverlustsystems M/M/n-s beschrieben. Als weitere Anwendungsbeispiele und zur Verdeutlichung der Abhängigkeit der Details der Lösungsschritte vom jeweiligen Problem können die entsprechenden Untersuchungen im fünften Kapitel herangezogen werden.

- Definition der Zustände :

Sämtliche existierenden Zustände des Warteschlangenmodells sind aufzustellen, so z.B. für das System M/M/n-s die Zustände

$$x_j, \quad j = 0, 1, \dots, s+n.$$

Bei Modellen mit mehreren Warteschlangen werden die Zustände durch entsprechend mehrdimensionale Vektoren dargestellt, z.B. für ein Warteschlangennetz, das zwei Systeme M/M/n-s enthält

$$(x_j, x_k), \quad j, k = 0, 1, \dots, s+n.$$

- Bestimmung der Übergangsraten :

Für sämtliche möglichen Zustandsübergänge von Zustand i nach Zustand j im definierten Zustandsraum des Modells sind die Übergangsraten q_{ij} zu bestimmen. In vielen Anwendungen gibt es aber nur Übergänge zwischen Nachbarzuständen ($j = i \pm 1$). Man spricht dann von Geburts- und Sterbeprozessen. Die Raten sind im wesentlichen durch die Ankunftsrate λ oder die Bedienungsrate ϵ (Kehrwert der mittleren Bedienungsdauern h) bestimmt. Für das System M/M/n-s ergeben sich z.B. die Raten

$$q_{ij} = \begin{cases} \lambda & , \quad j = i+1, \quad 0 \leq j \leq s+n \\ i \cdot \epsilon & , \quad j = i-1, \quad 0 \leq j < n \\ n \cdot \epsilon & , \quad j = i-1, \quad n \leq j < s+n \end{cases} \quad (4.20)$$

- Aufstellung des Zustandsgleichungssystems :

Für die definierten Zustände ist mittels der Übergangsraten das Kolmogoroffsche lineare, homogene Gleichungssystem der Zustandswahrscheinlichkeiten gemäß Gleichung 4.18 aufzustellen. Die Übergangsrate q_j für das Verlassen eines Zustands j ist dann die Summe der Raten q_{jk} aller Übergänge, die den Zustand j verlassen. Für das Beispiel des Systems M/M/n-s ergibt sich das Gleichungssystem

$$\begin{aligned} p_0 \cdot \lambda &= p_1 \cdot 1 \cdot \epsilon \\ p_1 \cdot (1 \cdot \epsilon + \lambda) &= p_0 \cdot \lambda + p_2 \cdot 2 \cdot \epsilon \end{aligned}$$

$$\begin{aligned} p_2 \cdot (2 \cdot \epsilon + \lambda) &= p_1 \cdot \lambda + p_3 \cdot 3 \cdot \epsilon \\ \vdots & \quad \quad \quad \vdots \\ p_n \cdot (n \cdot \epsilon + \lambda) &= p_{n-1} \cdot \lambda + p_{n+1} \cdot n \cdot \epsilon \\ p_{n+1} \cdot (n \cdot \epsilon + \lambda) &= p_n \cdot \lambda + p_{n+2} \cdot n \cdot \epsilon \\ \vdots & \quad \quad \quad \vdots \\ p_{s+n-1} \cdot (n \cdot \epsilon + \lambda) &= p_{s+n-2} \cdot \lambda + p_{s+n} \cdot n \cdot \epsilon \\ p_{s+n} \cdot n \cdot \epsilon &= p_{s+n-1} \cdot \lambda \end{aligned} \quad \dots (4.21)$$

- Lösung des Zustandsgleichungssystems :

Die stationären Zustandswahrscheinlichkeiten können durch die Lösung des aufgestellten homogenen Gleichungssystems zusammen mit der Normierungsbedingung nach Gleichung 4.19 eindeutig ermittelt werden. Je nach Problemstellung kann dies explizit, rekursiv oder iterativ, z.B. mit dem Gauß-Seidel-Algorithmus erfolgen. In vielen praktischen Anwendungsfällen, besonders im Falle mehrdimensionaler Zustände, ist das Gleichungssystem von hoher Ordnung und oft nur iterativ lösbar. Ferner zwingt sich, auch bei relativ unregelmäßigen Zustandsräumen, eine automatische, algorithmische Erstellung des Gleichungssystems auf einem Rechner auf. Das Beispiel des Warteverlustsystems M/M/n-s stellt dagegen einen Fall mit expliziter Angabe der Lösung dar:

$$\begin{aligned} p_j &= \begin{cases} p_0 \cdot A^j / j! & , \quad 1 \leq j < n \\ p_0 \cdot A^n / n! \cdot \rho^{j-n} & , \quad n \leq j \leq s+n \end{cases} \\ p_0 &= \begin{cases} \left[\sum_{j=0}^{n-1} A^j / j! + A^n / n! \cdot (1 - \rho^{s+1}) / (1 - \rho) \right]^{-1} & , \quad \rho \neq 1 \\ \left[\sum_{j=0}^{n-1} n^j / j! + n^n / n! \cdot (s+1) \right]^{-1} & , \quad \rho = 1 \end{cases} \end{aligned} \quad \dots (4.22)$$

mit dem Angebot

$$A = \lambda / \epsilon$$

in der Pseudo-Einheit "Erlang"

und

$$\rho = A / n.$$

- Ermittlung charakteristischer Größen :

Aus den stationären Zustandswahrscheinlichkeiten können dann durch einfache Summationen oder Beziehungen, wie z.B. Little's Theorem *), charakteristische Systemgrößen gewonnen werden.

Für das Warteverlustsystem M/M/n-s ergibt sich z.B.:

- Blockierungswahrscheinlichkeit $B = p_{n+s}$ (4.23)

- Wartewahrscheinlichkeit $W = \sum_{j=n}^{s+n-1} p_j$ (4.24)

- mittlere Warteschlangenlänge $\Omega = \sum_{j=0}^s j \cdot p_{j+n}$ (4.25)

- mittlere Wartezeit w bezüglich aller Anforderungen (gemäß Little's Theorem) $w = \Omega / \lambda$ (4.26)

- mittlere Wartezeit t_w bezüglich der wartenden Anforderungen $t_w = w / W$ (4.27)

- Verkehrswert (Belastung) $Y = A \cdot (1 - B)$ (4.28)

*) Little's Theorem besagt, daß in einem beliebigen System bei einem beliebigen Ankunftsprozeß die mittlere Anzahl der Anforderungen im System gleich dem Produkt aus der Ankunftsrate und der mittleren Aufenthaltsdauer im System ist.

5 EINZELUNTERSUCHUNGEN

5.1 Strukturuntersuchungen mittels Signalflußanalyse

Dieser Abschnitt beschreibt die Ergebnisse und die sich daraus ergebenden Folgerungen einer Anwendung der im vierten Kapitel eingeführten Signalflußanalyse auf eine verallgemeinerte, aber realistische Steuerungsstruktur mit verallgemeinerten Meldungsszenarien.

5.1.1 Vorgaben

Als typischer Vertreter verteilter Steuerungsstrukturen von Vermittlungssystemen wurde die allgemeine dreistufige Struktur gemäß Bild 4.5 einer Signalflußanalyse unterzogen. Dabei wurden 20 ausführliche, typische Meldungsszenarien mit insgesamt ca. 1250 Meldungen zugrunde gelegt, die die fünf relevanten Typen von Belegungsversuchen BV1 ... BV5 der fünf wichtigsten Belegungsarten BA1 ... BA5 für Fernsprechkverkehr abdecken (siehe Tabelle 5.1).

<u>Belegungsarten</u>	Internrufe	BA1
	Abgehende Ortsrufe	BA2
	Abgehende Fernrufe	BA3
	Ankommende Ortsrufe	BA4
	Ankommende Fernrufe	BA5
<u>Belegungsversuche</u>	Erfolgreiche Rufe	BV1
	Rufe mit abgebrochener Wahl	BV2
	Nach Wahlende gassenbesetzte Rufe	BV3
	Rufe, bei denen der B-Tln besetzt ist	BV4
	Rufe, bei denen sich der B-Tln nicht meldet	BV5

Tabelle 5.1: Belegungsarten (BAi) und Belegungsversuche (BV_i)

Zwei Belegungsspektren, d.h. zwei Fälle der Gewichtung der Auftrittswahrscheinlichkeiten, wurden betrachtet. Sie basieren auf realen Verkehrsmessungen der Fernmeldeindustrie. Der erste Fall ist durch den relativ geringen Gesamtanteil von 40% Internverkehr gekennzeichnet (siehe Tabelle 5.2), und der zweite Fall durch den höheren Anteil von 70% Internverkehr (siehe Tabelle 5.3).

	Gesamtanteil	BV1	BV2	BV3	BV4	BV5
BA1	0,4	0,22	0,032	-	0,088	0,06
BA2	0,2	0,11	0,018	0,016	0,036	0,02
BA3	0,1	0,05	0,009	0,013	0,018	0,01
BA4	0,2	0,126	-	-	0,044	0,03
BA5	0,1	0,063	-	-	0,022	0,015

Tabelle 5.2: Belegungsspektrum für 40% Internverkehr

	Gesamtanteil	BV1	BV2	BV3	BV4	BV5
BA1	0,7	0,385	0,056	-	0,154	0,105
BA2	0,1	0,055	0,009	0,008	0,018	0,01
BA3	0,05	0,025	0,0045	0,0065	0,009	0,005
BA4	0,1	0,063	-	-	0,022	0,015
BA5	0,05	0,0315	-	-	0,011	0,0075

Tabelle 5.3: Belegungsspektrum für 70% Internverkehr

5.1.2 Verallgemeinerte Ergebnisse

Zur Verallgemeinerung der Ergebnisse der Signalflußanalyse des definierten Systems sollen die mittleren Belastungen der Teilsysteme in normierter Form gegeben werden, nämlich bezogen auf die mittlere Gesamtanrufrate, die mittleren Bedienungsdauern und die

Anzahl der an den peripheren Einheiten angeschlossenen Teilnehmer und Übertragungsleitungen. Dadurch bleiben die Ergebnisse unabhängig von der Systemgröße, der Systemkonfiguration und der Leistungsfähigkeit der Prozessoren. Folgende Vereinbarungen gelten:

- Mittlere Anrufrate λ
- Mittlere Meldungsbearbeitungsdauer der PVE h_{PVE}
- Mittlere Meldungsbearbeitungsdauer der GVE h_{GVE}
- Mittlere Meldungsbearbeitungsdauer der ZVE h_{ZVE}
- Mittlere Meldungsübertragungszeit der PÜE $h_{PÜE}$
- Mittlere Meldungsübertragungszeit der ZÜE $h_{ZÜE}$
- Anzahl der Teilnehmer pro PVE $n_{TLN/PVE}$
- Anzahl gehender Übertragungsleitungen pro PVE $n_{UEG/PVE}$
- Anzahl kommender Übertragungsleitungen pro PVE $n_{UEK/PVE}$
- Anzahl der Teilnehmer pro GVE $n_{TLN/GVE}$
- Anzahl gehender Übertragungsleitungen pro GVE $n_{UEG/GVE}$
- Anzahl kommender Übertragungsleitungen pro GVE $n_{UEK/GVE}$

Mit diesen Vereinbarungen und den erhaltenen Ergebnissen der Signalflußanalyse lassen sich die mittleren Belastungen der Teilsysteme durch die folgenden Beziehungen darstellen.

Mittlere Belastung der PVE für 40% Internverkehr:

$$Y_{PVE} = (1,61 \cdot 10^{-3} \cdot n_{TLN/PVE} + 7,45 \cdot 10^{-3} \cdot n_{UEG/PVE} + 3,81 \cdot 10^{-3} \cdot n_{UEK/PVE}) \cdot \lambda \cdot h_{PVE} \quad (5.1a)$$

Mittlere Belastung der PVE für 70% Internverkehr:

$$Y_{PVE} = (1,76 \cdot 10^{-3} \cdot n_{TLN/PVE} + 3,72 \cdot 10^{-3} \cdot n_{UEG/PVE} + 1,91 \cdot 10^{-3} \cdot n_{UEK/PVE}) \cdot \lambda \cdot h_{PVE} \quad (5.1b)$$

Mittlere Belastung der GVE für 40% Internverkehr:

$$Y_{GVE} = (7,54 \cdot 10^{-4} \cdot n_{TLN/GVE} + 5,10 \cdot 10^{-3} \cdot n_{UEG/GVE} + 2,18 \cdot 10^{-3} \cdot n_{UEK/GVE}) \cdot \lambda \cdot h_{GVE} \quad (5.2a)$$

Mittlere Belastung der GVE für 70% Internverkehr:

$$Y_{GVE} = (8,30 \cdot 10^{-4} \cdot n_{TLN/GVE} + 2,55 \cdot 10^{-3} \cdot n_{UEG/GVE} + 1,09 \cdot 10^{-3} \cdot n_{UEK/GVE}) \cdot \lambda \cdot h_{GVE} \quad (5.2b)$$

Mittlere Belastung der ZVE für 40% Internverkehr:

$$Y_{ZVE} = 2,696 \cdot \lambda \cdot h_{ZVE} \quad (5.3a)$$

Mittlere Belastung der ZVE für 70% Internverkehr:

$$Y_{ZVE} = 2,268 \cdot \lambda \cdot h_{ZVE} \quad (5.3b)$$

Mittlere Belastung der PÜE für 40% Internverkehr:

$$Y_{PÜE} = (1,67 \cdot 10^{-3} \cdot n_{TLN/GVE} + 7,95 \cdot 10^{-3} \cdot n_{UEG/GVE} + 4,29 \cdot 10^{-3} \cdot n_{UEK/GVE}) \cdot \lambda \cdot h_{PÜE} \quad (5.4a)$$

Mittlere Belastung der PÜE für 70% Internverkehr:

$$Y_{PÜE} = (1,81 \cdot 10^{-3} \cdot n_{TLN/GVE} + 3,97 \cdot 10^{-3} \cdot n_{UEG/GVE} + 2,15 \cdot 10^{-3} \cdot n_{UEK/GVE}) \cdot \lambda \cdot h_{PÜE} \quad (5.4b)$$

Mittlere Belastung der ZÜE für 40% Internverkehr:

$$Y_{ZÜE} = 12,06 \cdot \lambda \cdot h_{ZÜE} \quad (5.5a)$$

Mittlere Belastung der ZÜE für 70% Internverkehr:

$$Y_{ZÜE} = 10,14 \cdot \lambda \cdot h_{ZÜE} \quad (5.5b)$$

5.1.3 Anwendungsbeispiel

Es werde ein realistisches Systembeispiel betrachtet.

Das Vermittlungssystem sei für 10.000 Teilnehmer konzipiert. Jeder Teilnehmer erzeuge ein mittleres Verkehrsangebot von 0,15 Erl (Erlang) in der Hauptverkehrsstunde. Somit ist das von den Teil-

nehmern erzeugte Gesamtangebot $A_{ges} = 1.500$ Erl. Mit der typischen mittleren Gesprächsdauer von $t_b = 60$ s folgt dann eine Gesamtankunftsrate $\lambda = A_{ges} / t_b = 25$ Rufe/s.

Ferner seien für die i.a. mit 8-bit-Mikroprozessoren realisierte periphere Ebene typische mittlere Meldungsbearbeitungszeiten von 5 ms in den PVE angenommen, in den GVE z.B. 10 ms und in der ZVE 8 ms, was für die i.a. mit 16-bit-Mikroprozessoren bestückten zentraleren Ebenen ebenfalls typisch ist. Die mittlere Meldungsübertragungszeit der PÜE sei 0,25 ms und die der ZÜE 0,15 ms. Letzteren Werten liegt die Annahme zugrunde, daß es sich bei der PÜE um einen Bus mit 256 Kbit/s Übertragungsrate mit im Mittel 8 byte langen Meldungen handelt, und im Fall der ZÜE um einen Bus mit 1 Mbit/s und z.B. 20 byte langen Meldungen.

Zur Konfigurierung eines konkreten Vermittlungssystems wird weiterhin angenommen, daß eine PVE wahlweise mit 64 Teilnehmern oder 8 gehenden oder 8 kommenden Übertragungsleitungen beschaltet werden kann. An eine GVE können ihrerseits maximal 16 beliebig beschaltete PVE angeschlossen sein.

Aus den Gleichungen 5.1 bis 5.5 ergeben sich dann konkrete mittlere Belastungen für die Teilsysteme des Beispiels, wie in Tabelle 5.4 gezeigt. Bei den PVE, GVE und PÜE sind jeweils die Werte für eine periphere Beschaltung ausschließlich mit Teilnehmerleitungen (TLN), nur mit gehenden (UEG) oder nur mit kommenden (UEK) Übertragungsleitungen gegeben. Den Werten für eine gemischte Beschaltung ist ein gängiges Verhältnis der Anzahl der Teilnehmerleitungen zu gehenden und zu kommenden Übertragungsleitungen von 10 : 1 : 1 zugrunde gelegt.

Die Anwendung der Ergebnisse ist auch auf nur zweistufige Systemstrukturen möglich. In diesem Fall werden die Funktionen der PVE von den GVE mit übernommen und es sind lediglich die einzelnen Belastungen der durch eine GVE mit abgedeckten PVE zur mittleren Belastung der GVE zu addieren.

Teilsystem	periphere Beschaltung	Mittlere Belastung bei	
		40% Internverkehr	70% Internverkehr
PVE	nur TLN	0.0129	0.0141
	nur UEG	0.0075	0.0037
	nur UEK	0.0038	0.0019
GVE	nur TLN	0.1930	0.2125
	nur UEG	0.1632	0.0816
	nur UEK	0.0698	0.0349
	gemischt	0.1460	0.1176
ZVE	gemischt	0.5392	0.4536
PÜE	gemischt	0.0107	0.0116
	nur UEG	0.0064	0.0032
	nur UEK	0.0034	0.0017
	gemischt	0.0071	0.0060
ZÜE	gemischt	0.0458	0.0385

Tabelle 5.4: Mittlere Belastungen für das Beispielsystem

5.1.4 Schlußfolgerungen

Aus diesen Ergebnissen können nun folgende Schlüsse für den Entwurf von Systemstrukturen oder Konfigurationen gezogen werden:

- PVE und PÜE :

Die PVE und PÜE weisen unter den im Beispiel angenommenen Voraussetzungen extrem geringe mittlere Belastungen auf. Da die Voraussetzungen jedoch aus konstruktiven Gründen realistisch sind (siehe auch Abschnitt 2.3.2), sind im peripheren Bereich keine Engpässe zu erwarten.

- GVE :

Die Annahme von maximal $16 \cdot 64 = 1024$ Teilnehmern oder $16 \cdot 8 = 64$ Übertragungsleitungen pro GVE stellt bei der vorausgesetzten, aber großzügig abgeschätzten Prozessorgeschwindigkeit eine vernünftige Größenordnung dar. Eine wesentliche Vergrößerung der peripheren Beschaltung erscheint aus Gründen der Modularität nicht sinnvoll. Bei eventuell weniger leistungsfähigen Prozessoren wäre eine Lastreduzierung aber leicht durch eine geringere periphere Beschaltung erzielbar.

- ZVE :

Bei insgesamt 10.000 Teilnehmern pro Vermittlungsstelle und der definierten Prozessorgeschwindigkeit erhält man bereits eine relativ hohe mittlere Belastung von ca. 50% in der ZVE. Eine Erhöhung der Teilnehmerzahl oder eine Verringerung der Prozessorgeschwindigkeit sollte aus Sicherheitsgründen nicht mehr erfolgen. Für größere Systeme ist deshalb unbedingt ein leistungsfähigerer Prozessor notwendig, oder besser und sicherer der Einsatz mehrerer Prozessoren im Lastteilungsbetrieb.

- ZÜE :

Die ZÜE weist im Beispielsystem eine relativ geringe Belastung auf. Sie besitzt somit noch genügend Leistungsreserven für eine merkliche Zunahme der gesamten Teilnehmerzahl. Ferner wäre darüber hinaus eine höhere Übertragungsrate als 1 Mbit/s technisch noch leicht realisierbar.

5.2 Simulation des Vermittlungssystems PILOT

In diesem Abschnitt werden das Simulationsmodell und einige typische Ergebnisse einer zeitreuen Simulationsstudie des im dritten Kapitel vorgestellten Vermittlungssystems PILOT beschrieben.

5.2.1 Simulationsmodell

Als Grundlage für ein Simulationsmodell des Systems PILOT kann das im Abschnitt 4.2.2.1 eingeführte, verallgemeinerte Systemmodell dienen. Es ergibt sich somit das in Bild 5.1 gezeigte spezifische Simulationsmodell.

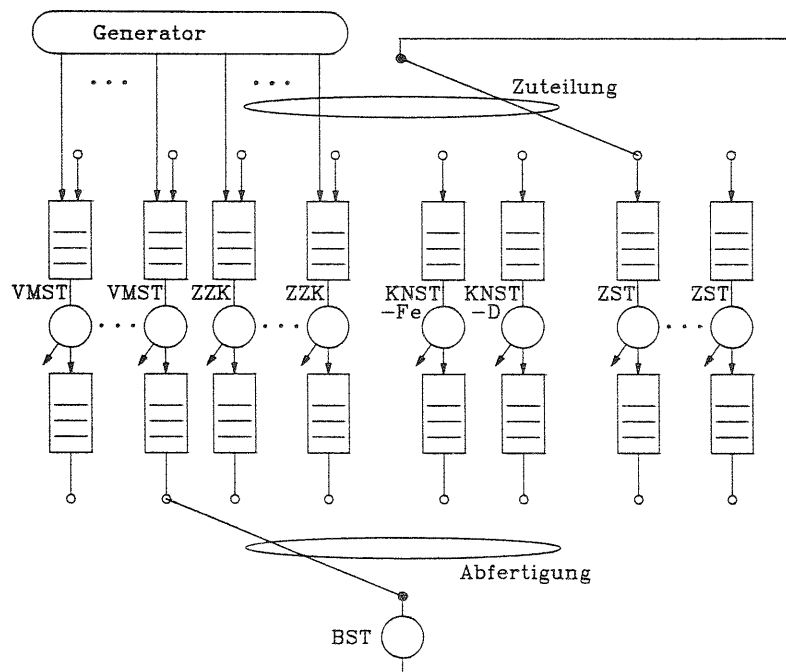


Bild 5.1: Simulationsmodell des Vermittlungssystems PILOT

Eine vorgebbare, von der jeweiligen Systemkonfiguration abhängige Anzahl von Vermittlungsmodul-Steuerungen VMST, ZZK-Steuerungen für Signalisierkanäle und zentralen Steuerungen ZST sowie die beiden Koppelnetzsteuerungen KNST-Fe und KNST-D sind durch je einen Grundmodul repräsentiert, der seinerseits eine Eingangswarteschlange (EWS), eine Bedienungseinheit (BE) und eine Ausgangswarteschlange (AWS) enthält. Der vom Verkehrsgenerator erzeugte Internverkehr, der real den VMST von den Teilnehmer-Signalisierereinheiten TSIG zugeführt wird, wird gleichwahrscheinlich in die EWS der VMST eingespeist. Der ebenfalls vom Generator erzeugte ankommende Externverkehr, der real steuerungstechnisch über zentrale Zeichenkanäle abgewickelt wird, führt gleichwahrscheinlich in die EWS der durch Grundmoduln repräsentierten ZZK-Steuerungen.

Die Interprozessor-Kommunikation zwischen allen Grundmoduln erfolgt über den Bus, dessen Bussteuerung BST durch eine spezielle Bedienungseinheit dargestellt wird. Dabei ist zu bemerken, daß die in Wirklichkeit über das Protokoll der CCITT No.7 -Kanäle abgewinkelte Kopplung der VMST an den zentralen Bus im Simulationsmodell vereinfachend nur durch eine entsprechende Übertragungszeit der Meldungen simuliert wird, weshalb die dafür vorgesehenen ZZK-Steuerungen im Modell nicht enthalten sind. Die Mechanismen der Abfertigung der AWS und der Zuteilung der Meldungen zu den EWS sind durch die entsprechenden Schalter abgebildet.

5.2.1.1 Verkehrsgenerator

Die Modellierungstiefe des Verkehrsgenerators ist im wesentlichen durch die hier verwendete teilbelegungstreue Simulationemethode /47/ bestimmt. Diese wurde gewählt, da sie einerseits bei der großen Anzahl gleichzeitig im System befindlicher Anforderungen (Teilrufe) die Realität genügend genau annähert und andererseits bezüglich des Rechenzeit- und Speicheraufwands erheblich günstiger als eine belegungstreue Simulation ist. Der Generator hat bei der teilbelegungstreuen Simulation die Aufgabe, ein Gemisch von verschiedenen Teilrufen zu erzeugen, das im vorliegenden Fall folgende drei Kriterien erfüllen muß:

- Die Ankunftsabstände T_A der Teilrufe gehorchen einer Exponential-Verteilung (siehe auch Abschnitt 4.2.1.1)
- Die zielbezogene Aufteilung der Teilrufe zu einer VMST oder zu einer ZZK-Steuerung entspricht der Auftrittswahrscheinlichkeit P_{int} des Internverkehrs bzw. P_{ext} des Externverkehrs, wobei die Aufteilung unter den VMST einer Gleichverteilung genügt.
- Die Häufigkeit jedes Teilruftyps entspricht seiner realen Auftrittswahrscheinlichkeit, wobei sich die Teilrufe für Fernsprech- und Datenverbindungen entsprechend der Wahrscheinlichkeit p_{Fe} für Fernsprech- und p_D für Datenverkehr aufteilen.

Somit ist jeder Teilruf durch seinen Ankunftszeitpunkt, sein Ziel und den Teilruftyp gekennzeichnet. Zur Generierung dieser drei Merkmale werden drei gleichverteilte Pseudozufallszahlen im Intervall $[0,1)$ herangezogen, die jeweils über eine entsprechende Verteilungsfunktion in den gewünschten Funktionswert transformiert werden. Im Falle des Ankunftsabstandes handelt es sich um eine kontinuierliche Zufallsvariable mit analytisch invertierbarer Verteilungsfunktion, so daß die Ermittlung exponentiell verteilter Ankunftsabstände durch Logarithmierung der gleichverteilten Zufallszahl möglich ist. Das Ziel und der Typ der Teilrufe sind dagegen diskrete Zufallsvariablen. Die Verteilungsfunktion ist dann treppenförmig, wobei die Stufenhöhen den Auftrittswahrscheinlichkeiten der diskreten Zufallsvariablen entsprechen. In einer programmierten Abfragekette wird dann ermittelt, in welche Stufe die Zufallszahl fällt.

5.2.1.2 Grundmodul

Aus Gründen der Modularität wird im Simulationsprogramm jede Verarbeitungseinheit durch einen gleichartigen Grundmodul gemäß der Darstellung in Bild 5.2 gebildet. Die Speichergröße der Warteschlangen ist vorgebbar. Die Speicherplätze nehmen Teilrufe bzw. Meldungen mit den Attributen Zeit, Ziel und Typ auf. Wie in Bild 5.2 schematisch angedeutet, wird die Handlungsweise der BE durch einen SDL-Graphen genau spezifiziert. Diese SDL-Graphen geben im

wesentlichen durch TASKs die Bedienungsdauern und durch OUTPUTs die eventuell resultierenden Folgemeldungen für alle möglichen durch INPUTs spezifizierte Eingangsmeldungen wieder. Sie bilden somit vereinfacht die im realen System PILOT implementierte Software /48,49,50/ ab. Als Bedienungsdauern der einzelnen Teilruf- bzw. Meldungstypen werden teilweise im realen System gemessene Programmlaufzeiten oder anhand der Programme geschätzte Werte verwendet. Die simulative Abhandlung der SDL-Graphen erfolgt analog der Beschreibung in Abschnitt 4.4.2. Schließlich muß noch angemerkt werden, daß der Einfluß der Koppelnetze des Systems und der abgehenden Leitungsbündel vereinfacht durch eine zufallsabhängige Verzweigung in den jeweiligen SDL-Prozessen entsprechend einer vorgebbaren Blockierwahrscheinlichkeit modelliert wird.

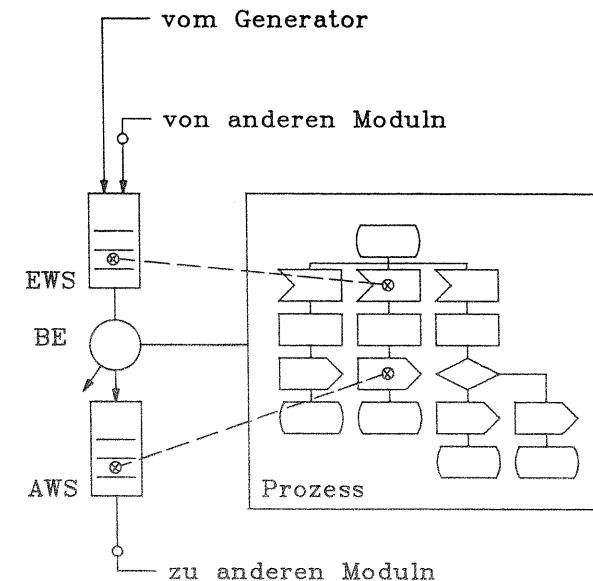


Bild 5.2: Grundmodul des Simulationsmodells

5.2.1.3 Busmodell

Die Abfertigung der AWS durch die BST erfolgt gemäß einem vorgebbaren Zyklus. Durch mehrfache Nennung einer bestimmten AWS in diesem Zyklus kann der entsprechenden Verarbeitungseinheit bezüglich der Abfertigung eine höhere Priorität verliehen werden. Im System PILOT ist jedoch auch eine dynamische Priorisierung vorgesehen, die sich an den Warteschlangenlängen der AWS orientiert. Im Simulationsprogramm wird dies dadurch realisiert, daß die AWS im Abfragezyklus automatisch zweifach eingetragen werden, deren Füllgrad einen vorgebbaren oberen Schwellwert übersteigt. Dies wird bei Unterschreitung eines unteren Schwellwertes wieder rückgängig gemacht.

Die Bedienungsdauer der BST modelliert die Verzögerung, die eine Meldung bei einem Bustransfer erfährt. Sie besteht aus der Übertragungsphase, einer Verwaltungsphase für die Ermittlung des Zieles der Meldung und einer weiteren Übertragungsphase für die Zuteilungs-Steuerungsmeldung (siehe auch Abschnitt 3.3.3). Bevor dann eine neue Übertragung erfolgen kann, ist grundsätzlich eine separate Verwaltungsphase zur Bestimmung der als nächstes abzufragenden Einheit, deren Adressierung und Abfrage notwendig, auch wenn dann keine Übertragung zustande kommt.

Im Bereich der zentralen Steuerungen wird in der vorgestellten Simulationsstudie ein Fall der Funktionsteilung betrachtet. Deshalb ist an dieser Stelle und für alle anderen Moduln eine feste Zuteilung der Meldungen notwendig, gesteuert durch das Ziel-Attribut der Meldungen. Lediglich, wenn das Ziel einer Meldung eine VMST oder eine ZZK-Steuerung ist, erfolgt eine zufallsgesteuerte, gleichwahrscheinliche Zuteilung zu einer VMST bzw. ZZK-Steuerung.

5.2.2 Simulationsergebnisse

Den in diesem Abschnitt gezeigten Simulationsergebnissen liegen die folgenden wichtigsten Parameter zugrunde:

- Systemkonfiguration: 8 AVM, 2 ZZK,
1 KNST-Fe, 1 KNST-D,
3 ZST
- Funktionsteilung: ZST1: Behandlung von Internverkehr
ZST2: Behandlung von Externverkehr
ZST3: Behandlung von Datenverbindungen
- Verkehrsaufteilung: $P_{int} : P_{ext} = 30\% : 70\%$
 $P_{Fe} : P_D = 80\% : 20\%$
- Warteschlangen: EWS, AWS: je 50 Warteschlangenplätze
- Bus:
Bedienungsdauer der BST: 1 ms
Abfrage-Verwaltungszeit der BST: 0,1 ms
Abfertigungsstrategie:
a) zyklisches Polling ohne Prioritäten
b) dynamisches Polling mit zweifacher Abfrage
Füllgradschwellen der AWS: 25 bzw. 10
Zuteilungsstrategie:
VMST, ZZK: je zufällig, gleichverteilt
sonst: entsprechend dem Meldungsziel
- Statistik: 10 Teilttests mit je 5000 Teilrufen

Das Diagramm Bild 5.3 zeigt die mittleren Belastungen der einzelnen Bedienungseinheiten des Systems in Abhängigkeit der vom Generator erzeugten Teilrufrate. Zur Verdeutlichung der Größenordnung der Teilrufrate sei bemerkt, daß z.B. der sehr hohe Wert von 100.000 Belegungsversuchen pro Stunde einer Teilrufrate von ca. 400 Teilrufen/s entspricht, wenn man die mittlere Anzahl von 14 Teilrufen pro Ruf bei der vorgegebenen Verkehrsaufteilung zugrunde legt. Im gezeigten Lastbereich steigen alle Belastungen noch linear, da im System bei Anwendung der dynamischen Bus-Abfertigungsstrategie bis zu 750 Teilrufen/s keine Blockierungen bei den Warteschlangen auftreten. Hätte allerdings die rein zyklische Abfertigung Anwendung gefunden, wären ab ca. 600 Teilrufen/s Verluste in den AWS aufgetreten. Sämtliche Belastungswerte liegen

also auch im Hochlastbereich in einer akzeptablen Größenordnung. Insbesondere die sehr niedrigen Werte der peripheren Einheiten bestätigen auch die bei den allgemeinen Signalflußanalyse-Studien gemachten Feststellungen.

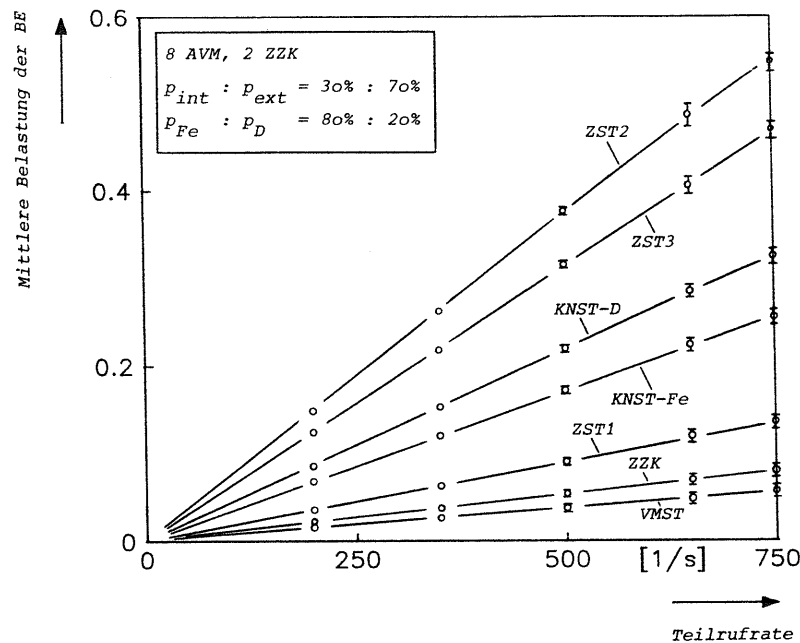


Bild 5.3: Mittlere Belastung der Bedienungseinheiten

Eines der für die Beurteilung der Leistungsfähigkeit eines Vermittlungssystems wichtigsten Kriterien, die mittlere Rufverzugszeit, zeigt das zweite Ergebnisbeispiel in Bild 5.4 für Internverbindungen in Abhängigkeit der Teilrufrate. Die teilbelegungs-treue Simulation erlaubt zwar nicht, eine derartige Durchlaufzeit explizit zu messen, aber der Mittelwert kann durch Aufsummierung aller betroffenen, gemessenen mittleren Wartezeiten und vorgegebenen Bedienungsdauern gemäß dem Meldungsszenario der betreffen-

den Belegungsart berechnet werden. Das Diagramm zeigt auf diese Weise gewonnene Ergebnisse für das Szenario von Fernsprech- und von Datenverbindungen. In beiden Fällen steigt die anfangs sehr niedrige Rufverzugszeit erst im Höchstlastbereich stark an, aber immer noch ohne Verluste bei den Warteschlangen. Dies gilt allerdings wieder nur bei Anwendung der dynamischen Abfertigungsstrategie der BST.

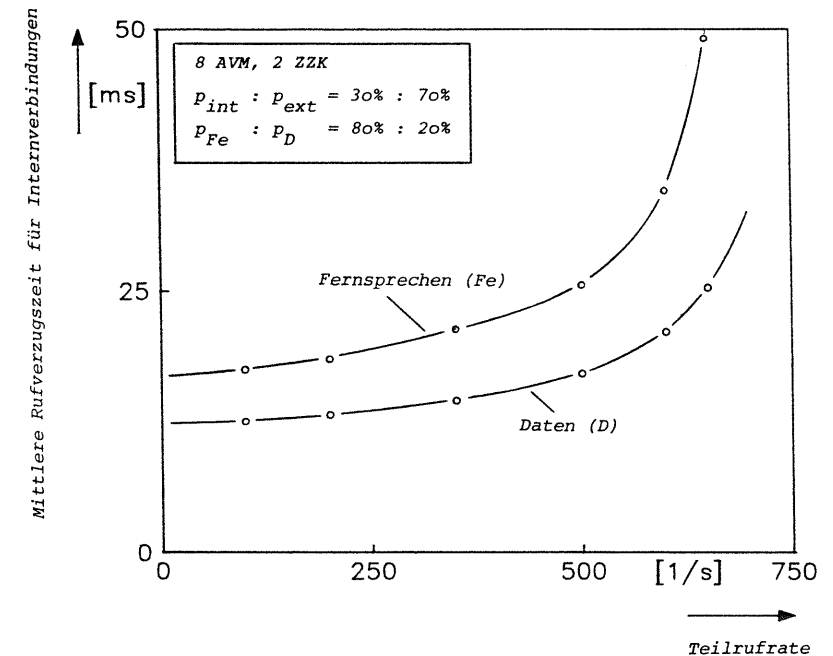


Bild 5.4: Mittlere Rufverzugszeit für Fernsprech- und Daten-Internverbindungen

In den folgenden beiden Diagrammen soll der Einfluß der beschriebenen dynamischen Abfertigungsstrategie der BST im Vergleich zur rein zyklischen Strategie gezeigt werden. Bild 5.5 zeigt dazu für die dynamische Strategie die Simulationsergebnisse des mittleren

Ankunftsabstandes der Meldungen am Eingang der einzelnen AWS. Zusätzlich ist die gemessene mittlere Buszykluszeit eingezeichnet. Falls nun in diesem Fall die AWS der ZST2 grundsätzlich nur einfach pro Zyklus abgefragt würde, so würde sie ab ca. 600 Teilrufen/s instabil, da sie dann schneller gefüllt als geleert wird. Aus diesem Grunde erklärt sich die Tatsache, daß bei rein zyklischer, prioritätsloser Abfrage oberhalb von dieser Grenzrate Verluste hauptsächlich bei der AWS der ZST2 auftreten. Im Falle der dynamischen Strategie wird jedoch oberhalb dieser Teilrufrate die AWS der ZST2 vorwiegend zweifach abgefragt, was aus ihrer Sicht einer scheinbaren Halbierung der Buszykluszeit entspricht und somit Verluste erst später ab ca. 750 Teilrufen/s eintreten.

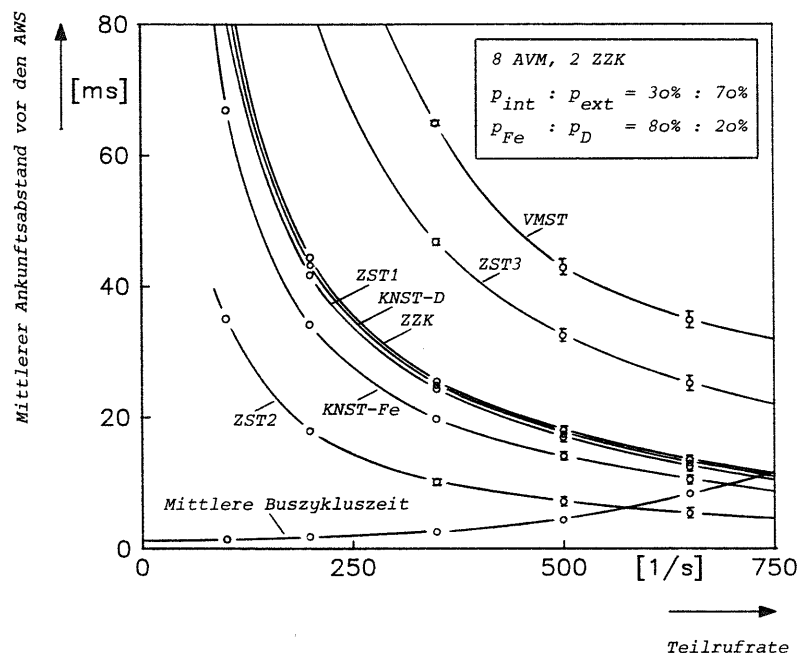


Bild 5.5: Mittlerer Ankunftsabstand vor den AWS und mittlere Buszykluszeit

Diese Betrachtung zur Bestimmung der Grenzrate ist jedoch insofern nur eine Näherung, daß bei mehrfacher Abfrage einer AWS der Buszyklus auf Kosten der anderen Einheiten länger wird. Simulationsergebnisse für den rein zyklischen Fall zeigen jedoch, daß der Unterschied unwesentlich ist, was durch die große Anzahl wenig belasteter abzufragender Einheiten im System begründbar ist. So zeigen z.B. die mittleren Warteschlangenlängen der AWS mit Ausnahme der hochbelasteten AWS der ZST2 nur wenig Veränderung. Dies ist in Bild 5.6 für die hohe Teilrufrate von 650 Teilrufen/s veranschaulicht. Durch den Regelmechanismus der dynamischen Abfertigung pendelt sich die Warteschlangenlänge der hochbelasteten AWS im Mittel nur knapp über der unteren Grenze des Hysteresebereichs des Füllgrades stabil ein, ohne Verluste aufzuweisen. Bei den anderen AWS wird dagegen die Warteschlangenlänge nur wenig größer und pendelt selbständig stabil um Werte unterhalb von 1.

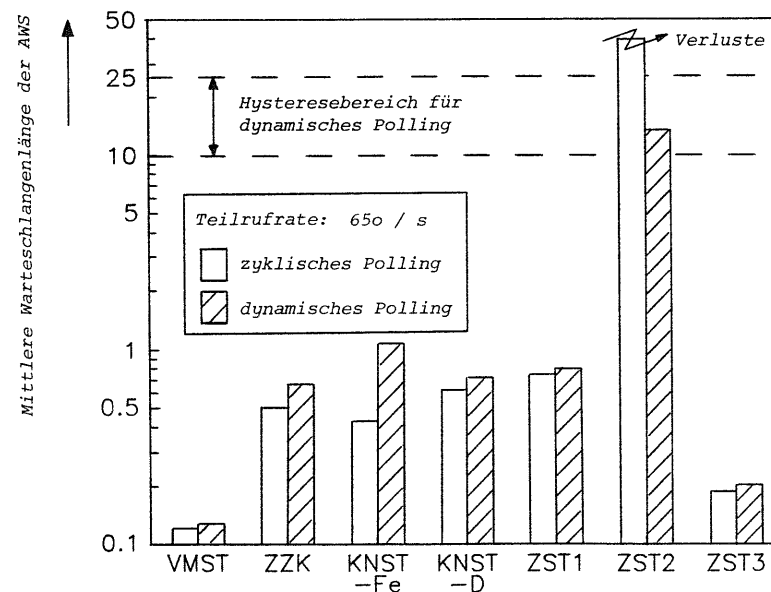


Bild 5.6: Mittlere Warteschlangenlängen der AWS

5.3 Warteschlangenanalyse von Teilmodellen

Im folgenden werden die beiden Teilaspekte der Zuteilung von Anforderungen zu Eingangswarteschlangen von Lastteilungsprozessoren und der Abfertigung von Ausgangswarteschlangen verschiedener Prozessoren analytisch untersucht. Mit dem Ziel, für den jeweiligen Anwendungsfall die optimale Strategie zu finden, werden dabei neben einfachen Mechanismen insbesondere intelligente Strategien berücksichtigt, d.h. Strategien, die sich am Systemzustand orientieren. Dazu werden für beide Aspekte grundsätzliche Teilmodelle definiert und für verschiedene, zustandsabhängige Betriebsweisen charakteristische Leistungsmerkmale mittels der stationären Markoffschen Zustandsanalysetechnik berechnet. Ein Vergleich der Ergebnisse ermöglicht dann die Wahl der optimalen Strategie.

5.3.1 Analyse von Zuteilungsstrategien

5.3.1.1 Zuteilungsstrategien

Es werden fünf prinzipiell unterschiedliche Zuteilungsstrategien gemäß folgenden Definitionen betrachtet:

- ZUFÄLLIG :

Die Zuteilung der Anforderungen erfolgt zufallsabhängig und gleichverteilt zu den einzelnen Warteschlangen.

- ZYKLISCH :

Die Zuteilung der Anforderungen erfolgt rein zyklisch, d.h. der Reihe nach abwechselnd zu den einzelnen Warteschlangen.

- DYNAMISCH :

Die Anforderungen werden dynamisch zugeteilt, d.h. zur jeweils kürzesten Warteschlange. Falls mehrere Warteschlangen den gleichen niedrigsten Füllgrad aufweisen, erfolgt die Zuteilung unter diesen gemäß der Strategie "ZUFÄLLIG".

- TEILDYNAMISCH :

Die Zuteilung der Anforderungen erfolgt teilweise dynamisch, d.h. solange noch Warteschlangen einen definierten Füllgrad nicht übersteigen, wird unter diesen gemäß der Strategie "ZUFÄLLIG" zugeteilt. Ansonsten erfolgt die Zuteilung gemäß der Strategie "DYNAMISCH".

- 2 VERSUCHE :

Die Zuteilung der Anforderungen geschieht, falls notwendig, in zwei Versuchen, wobei der erste Versuch entsprechend der Strategie "ZUFÄLLIG" erfolgt, und falls dieser scheitert, wird dann ein zweiter Versuch durchgeführt, bei dem gemäß der Strategie "ZUFÄLLIG" zugeteilt wird, und zwar

- a) unter allen anderen Warteschlangen,
- b) unter allen nicht vollen Warteschlangen.

Die beiden ersten Strategien werden als nicht intelligente Zuteilungsverfahren den anderen drei intelligenten, zustandsabhängigen Verfahren gegenübergestellt. Die Strategie "ZUFÄLLIG" wird nur zum Vergleich wegen ihrer einfachen Analyse betrachtet. In der Praxis ist sie schwieriger implementierbar als die meist verwendete Strategie "ZYKLISCH". Die beiden Strategien "TEILDYNAMISCH" und "2 VERSUCHE" stellen ferner einen Versuch dar, den hohen Verwaltungs- und Implementierungsaufwand der Strategie "DYNAMISCH" unter Beibehaltung einer hohen Leistungsfähigkeit zu reduzieren.

5.3.1.2 Grundmodell

Bild 5.7 zeigt das Grundmodell für die Warteschlangenanalyse von Zuteilungsstrategien. Mit der Ankunftsrate λ eintreffende Anforderungen eines Markoffschen Ankunftsprozesses werden abhängig von der jeweiligen Zuteilungsstrategie zu einem von g Warteverlustsystemen zugeteilt, falls sie dort nicht mit der Wahrscheinlichkeit B blockiert werden. Jedes Warteverlustsystem i besitzt eine Warteschlange der Kapazität s_i und eine Bedienungseinheit mit einem

Markoffschem Bedienungsprozeß mit der Bedienungsenderate ϵ_i . Die Zustandsvariable x_i repräsentiert die Anzahl von Anforderungen im Teilsystem i , und der gesamte Systemzustand wird durch den Zustandsvektor $\underline{X} = (x_1, x_2, \dots, x_g)$ beschrieben.

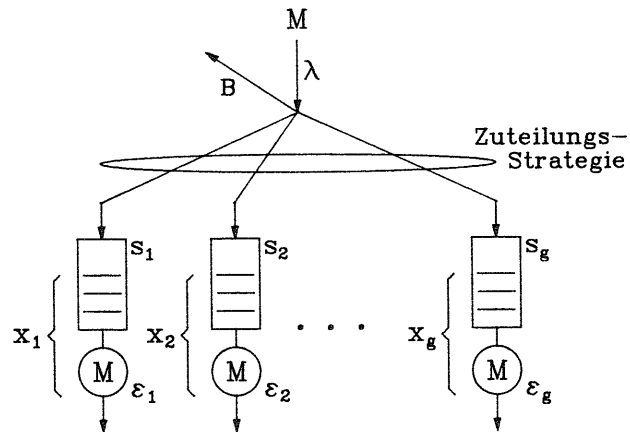


Bild 5.7: Grundmodell für die Analyse von Zuteilungsstrategien

5.3.1.3 Analyse

In der Literatur sind nur wenige analytische Untersuchungen von Zuteilungsverfahren bekannt, die meist Sonderfälle der Strategie "DYNAMISCH" behandeln oder auf einem spezielleren Modell basieren. So wird z.B. in /51/ der Sonderfall eines aus nur zwei Teilsystemen bestehenden symmetrischen Systems für die Strategie "DYNAMISCH" gelöst. In /52/ wird dagegen zur Untersuchung einer optimalen Zuteilung von Anrufen zu Vermittlungsplätzen ein Modell mit einer vorgeschalteten gemeinsamen Warteschlange im Zuteiler behandelt. Da die Ergebnisse solcher spezieller Untersuchungen jedoch für den Vergleich der genannten Strategien nicht geeignet sind, wird im folgenden die Lösungsmethode für jede zu untersuchende Strategie aufgezeigt.

Bei der Strategie "ZUFÄLLIG" wird der Markoffsche Ankunftsprozeß mit der Rate λ durch die zufällige, gleichverteilte Zuteilung in g Markoffsche Prozesse mit jeweils der Rate λ/g aufgeteilt (siehe z.B. /33/). Somit erhält man g unabhängige Warteverlustsysteme des Typs M/M/1-s. Für jedes dieser Teilsysteme sind dann die Zustandswahrscheinlichkeiten nach Gleichung 4.22 und die charakteristischen Größen nach den Gleichungen 4.23 bis 4.28 mit $n=1$ einfach ermittelbar.

Bei der Zuteilung nach der Strategie "ZYKLISCH" wird einem Teilsystem jede g -te Ankunft aus dem Markoffschen Ankunftsstrom zugeteilt. Dies bedeutet, daß sich der Abstand zwischen zwei aufeinander folgenden Ankünften bei einem Teilsystem aus g seriellen und jeweils exponentiell mit gleichem Mittelwert verteilten Teilphasen zusammensetzen läßt. Der resultierende Ankunftsprozeß eines einzelnen Teilsystems entspricht deshalb einem Erlang-Prozeß der Ordnung g mit einer Ankunftsrate von λ/g . Die Analyse der Strategie "ZYKLISCH" besteht somit darin, das Warteverlustsystem vom Typ $E_k/M/1-s$ zu lösen, wobei $k=g$ und $s=s_i$ zu setzen sind. Dazu splittet man z.B. jeden Zustand eines Teilsystems i in g Unterzustände auf, indem man die zweidimensionale Zustandsvariable $\underline{X}_i = (x_i, j)$ einführt, wobei x_i der Anzahl der Anforderungen im Teilsystem i und j einer Zählvariablen entspricht, die bei jeder Ankunft einer neuen Anforderung ins Gesamtsystem modulo g inkrementiert wird. Dabei darf x_i mit jedem Durchlauf von j nur einmal (z.B. bei $j=g$) inkrementiert werden, solange x_i kleiner als s_i+1 ist. Es ergibt sich somit ein Zustandsraum mit Übergangsraaten gemäß Bild 5.8, für den das lineare Gleichungssystem der Zustandswahrscheinlichkeiten $p(\underline{X}_i)$ für das statistische Gleichgewicht aufzustellen ist. Dieses kann zusammen mit der Normierungsbedingung rekursiv gelöst werden. Die Summe der jeweils g Unterzustandswahrscheinlichkeiten $p(\underline{X}_i)$ über alle j ergibt dann die eigentlichen Zustandswahrscheinlichkeiten $p(x_i)$ eines Teilsystems i , die dann die Ermittlung der charakteristischen Größen wie beim System M/M/1-s gemäß der Gleichungen 4.23 bis 4.28 mit $n=1$ erlauben.

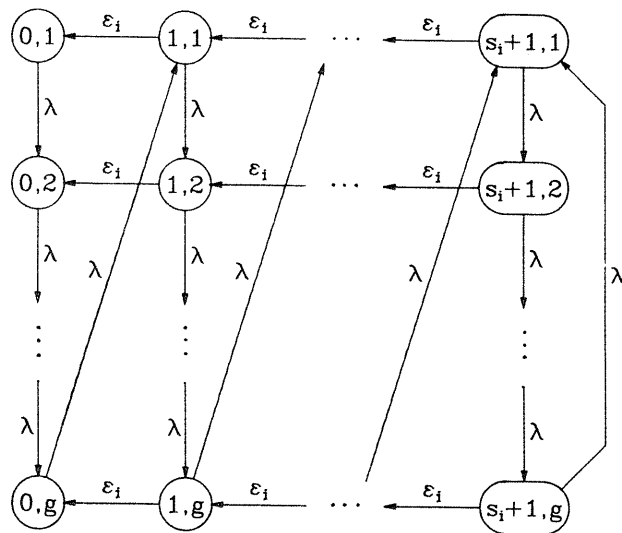


Bild 5.8: Zustandsraum für die Zuteilungsstrategie "ZYKLISCH"

Der Lösungsweg für die zustandsabhängigen Strategien "DYNAMISCH", "TEILDYNAMISCH" und "2 VERSUCHE" ist prinzipiell identisch. In jedem Falle sind g -dimensionale Zustandsräume für die Zustandsvariable $\underline{x} = (x_1, x_2, \dots, x_g)$ mit den von der Strategie abhängigen Zustandsübergangsraten aufzustellen. Das Prinzip soll anhand zweier einfacher Beispiele in den Bildern 5.9 und 5.10 gezeigt werden.

In Bild 5.9 handelt es sich um ein System aus zwei Warteschlangen mit je drei Warteschlangenplätzen und der Zuteilungsstrategie "DYNAMISCH". Die Zuteilungswahrscheinlichkeiten p_1 und $p_2 = 1-p_1$ für den Fall, daß beide Warteschlangen den gleichen Füllgrad aufweisen, sind hierbei noch allgemein gehalten. Bei der gleichverteilten Zuteilung gilt $p_1 = p_2 = 0,5$.

Bei dem in Bild 5.10 gezeigten Beispiel handelt es sich um ein gleich großes System mit der Zuteilungsstrategie "TEILDYNAMISCH". Die Füllgradgrenze t , unterhalb der eine zufällige Zuteilung erfolgt und ab der die dynamische Zuteilung einsetzt, ist dabei auf $x_i=3$ festgelegt. Die Zuteilungswahrscheinlichkeiten sind unterhalb von t grundsätzlich und ab t im Falle des Gleichstandes beider Warteschlangen wieder allgemein als p_1 bzw. p_2 angegeben.

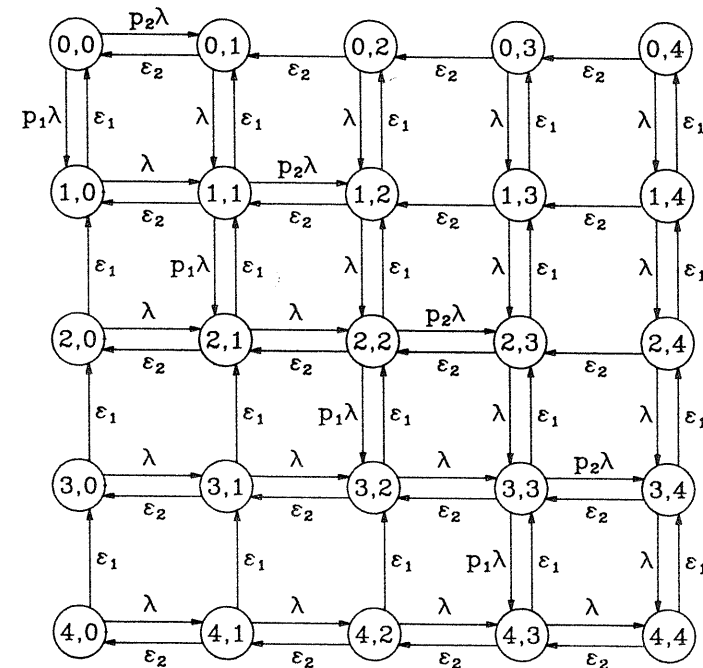


Bild 5.9: Beispiel eines Zustandsraums für die Zuteilungsstrategie "DYNAMISCH" ($g = 2, s_i = 3$)

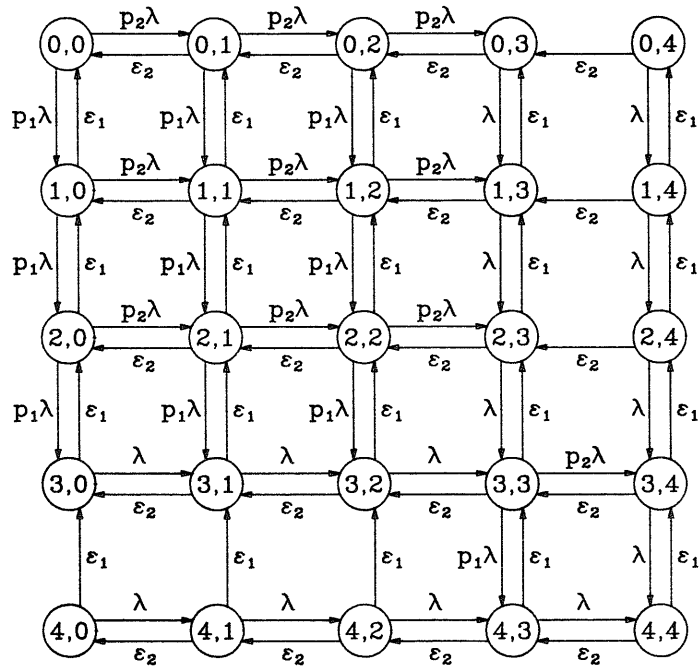


Bild 5.10: Beispiel eines Zustandsraums für die Zuteilungsstrategie "TEILDYNAMISCH" ($g = 2$, $s_i = 3$, $t = 3$)

Die Strategie "2 VERSUCHE" stellt bezüglich ihrer Analyse einen Sonderfall der Strategie "TEILDYNAMISCH" dar. Setzt man nämlich bei dieser die Füllgradgrenze t auf den Maximalwert s_i+1 , so wird prinzipiell zunächst zufällig zugeteilt, solange keine volle Warteschlange angetroffen wird. Erst in letzterem Falle setzt eine intelligentere Strategie ein, die entweder nur die Kenntnis der als voll angetroffenen Warteschlange (Fall a) oder die Kenntnis aller vollen Warteschlangen (Fall b) benützt. Praktisch kann dies durch zwei Versuche realisiert werden, wobei bei Mißerfolg des ersten Versuches ein zweiter folgt, der die Erfahrung des ersten Versuches oder auch mehr Information zur Entscheidung benützt.

Die Fälle a und b unterscheiden sich bei nur zwei Teilsystemen nicht. Bei größeren Systemen ist der Zustandsraum für den Fall b identisch mit der Strategie "TEILDYNAMISCH" für $t = s_i+1$, während der Fall a an den Rändern des Zustandsraumes leicht modifiziert ist.

Bei den drei beschriebenen, zustandsabhängigen Zuteilungsstrategien ergeben sich für realistische Systeme i.a. mehrdimensionale und größere Zustandsräume als die gezeigten Beispiele. Es ist deshalb notwendig, deren Speicherabbilder im Rechner oder direkt die Koeffizienten der Zustandsgleichungen automatisch zu erstellen. Dazu ist eine Beschreibung erforderlich, die für jeden beliebigen Zustand des Zustandsraumes algorithmisch sämtliche davon ausgehenden Zustandsübergänge angibt. Darin spiegelt sich jedoch hier im wesentlichen der Algorithmus der Zuteilungsstrategie wieder. Die Lösung der Gleichungssysteme, also die Ermittlung der Zustandswahrscheinlichkeiten $p(\underline{x})$, kann in allen beschriebenen Fällen nur iterativ erfolgen, wofür sich z.B. der Gauß-Seidel-Algorithmus besonders eignet. Es lassen sich dann mit vertretbarem Rechenzeit- und Speicheraufwand noch Systeme mit z.B. sechs Teilsystemen und je vier Warteplätzen oder vier Teilsystemen mit je zehn Warteplätzen lösen.

Aus den Zustandswahrscheinlichkeiten $p(\underline{x})$ können nun die charakteristischen Leistungsgrößen der Systeme mit den zustandsabhängigen Strategien "DYNAMISCH", "TEILDYNAMISCH" und "2 VERSUCHE" allgemein durch die einfachen Beziehungen gemäß den Gleichungen 5.6 bis 5.11 ermittelt werden. Dabei ist $p_i(\underline{x})$ die für jeden Zustand durch die jeweilige Strategie vorgegebene, bedingte Zustandsübergangswahrscheinlichkeit für einen Übergang vom Zustand \underline{x} in den Nachbarzustand $(x_1, x_2, \dots, x_{i+1}, \dots, x_g)$, also z.B. die Zuteilungswahrscheinlichkeit p_1 bzw. p_2 der Beispiele in Bild 5.9 und 5.10 in den betreffenden Zuständen.

- Die globale Blockierungswahrscheinlichkeit B , mit der eine ankommende Anforderung abgewiesen wird, ergibt sich zu

$$B = \sum_{i=1}^g \sum_{\underline{X} | x_i = s_i + 1} p(\underline{X}) \cdot p_i(\underline{X}) \quad (5.6)$$

- Die globale Wartewahrscheinlichkeit W , mit der eine ankommende Anforderung warten muß, ist

$$W = \sum_{i=1}^g \sum_{\underline{X} | 0 < x_i < s_i + 1} p(\underline{X}) \cdot p_i(\underline{X}) \quad (5.7)$$

- Die mittlere Anzahl Ω von Anforderungen im Gesamtsystem, d.h. die Summe aller Warteschlangenlängen, berechnet sich zu

$$\Omega = \sum_{i=1}^g \sum_{\underline{X} | x_i > 0} (x_i - 1) \cdot p(\underline{X}) \quad (5.8)$$

- Somit folgt nach Littles Theorem als mittlere Wartezeit w bezüglich aller ankommender Anforderungen

$$w = \Omega / \lambda \quad (5.9)$$

- Die mittlere Wartezeit t_w bezüglich der wartenden Anforderungen ergibt sich daraus zu

$$t_w = w / W \quad (5.10)$$

- Für den Verkehrswert, d.h. die Belastung Y_i einer Bedienungseinheit i gilt ferner

$$Y_i = 1 - \sum_{\underline{X} | x_i = 0} p(\underline{X}) \quad (5.11)$$

5.3.1.4 Implementierungsaspekte

Zur Ausführung zustandsabhängiger Strategien benötigt der Zuteiler zu bestimmten Zeiten Informationen über den Systemzustand, und auch bei einfachen Strategien ist i.a. eine Information bei Blockierung notwendig, um Verluste von Meldungen durch Rückstau zu vermeiden. Im Falle eines zentralen Zuteilers, wie z.B. die BST im System PILOT, müssen diese Informationen von den Buszugriffseinheiten der Teilsysteme über den Bus zurückgemeldet werden, was eine zusätzliche Belastung des Busses und einen Steuerungsaufwand im Zuteiler zur Folge hat. In beiden Fällen ist der Einfluß direkt der Rate λ_r der Rückmeldungen proportional. Deshalb soll diese Rückkopplungsrate λ_r als Maß für den Steuerungsaufwand der Strategien benützt werden. Da deren Bestimmung jedoch von der Art der Implementierung abhängig ist, wird im folgenden versucht, für jede Strategie möglichst die Implementierung zugrunde zu legen, die eine minimale Rate λ_r zur Folge hat.

- Bei den Strategien "ZUFÄLLIG" und "ZYKLISCH" ist lediglich bei jeder Blockierung eine Rückmeldung notwendig. Deshalb ist λ_r proportional der Rate der blockierten Anforderungen, also

$$\lambda_r = \lambda \cdot B \quad (5.12)$$

- Im Falle der Strategie "DYNAMISCH" benötigt der Zuteiler zu jeder Zeit den gesamten Systemzustand. Da er jedoch aufgrund seiner Zuteilungen alle positiven Änderungen des Systemzustandes kennt, genügt die Rückmeldung aller negativen Zustandsänderungen, um den Systemzustand rekonstruieren zu können. Somit ist bei jeder Beendigung einer Bedienungsphase im System eine Rückmeldung notwendig, und λ_r ergibt sich zu

$$\lambda_r = \sum_{i=1}^g Y_i \cdot \epsilon_i = \lambda \cdot (1 - B) \quad (5.13)$$

- Bei der Strategie "TEILDYNAMISCH" ist jeweils eine Rückmeldung erforderlich, wenn durch die Ankunft einer Anforderung der de-

finierte Schwellwert t des Füllgrades der betreffenden Warteschlange überschritten wird, oder wenn durch die Beendigung einer Bedienungsphase dieser Schwellwert unterschritten wird. Somit folgt

$$\lambda_r = \lambda \cdot q_{t-1} + \lambda \cdot (1 - B) \cdot q_t, \quad (5.14a)$$

wobei q_z die Wahrscheinlichkeit ist, daß das Teilsystem, dem eine ankommende Anforderung zugewiesen wird, mit dem Zustand z vorgefunden wird. Sie ermittelt sich im Falle gleich großer Teilsysteme zu

$$q_z = \sum_{\underline{x} | x_i = z} p(\underline{x}). \quad (5.14b)$$

- Bei der Strategie "2 VERSUCHE" ist eine Möglichkeit der Implementierung bereits durch die Definition gegeben. Im Fall a ist nach dem ersten Versuch bei jeder Ankunft, die eine volle Warteschlange vorfindet, eine Rückmeldung notwendig und eine zweite, wenn ein Folgeversuch endgültig scheitert. Ersteres geschieht wiederum im Fall gleich großer Teilsysteme mit der Wahrscheinlichkeit q_{s_i+1} und letzteres mit Wahrscheinlichkeit B . Somit ist

$$\lambda_r = \lambda \cdot (q_{s_i+1} + B). \quad (5.15a)$$

Der Fall b könnte prinzipiell realisiert werden, indem sich nach einem gescheiterten ersten Versuch zusätzlich alle nicht vollen Teilsysteme zurückmelden. Dann wäre

$$\lambda_r = \lambda \cdot q_{s_i+1} \cdot (1 + g \cdot (1 - q_{s_i+1})). \quad (5.15b)$$

Da der Fall b jedoch ein Sonderfall der Strategie "TEILDYNAMISCH" ist, sollte bei entsprechender Implementierung besser die Gleichung 5.14a mit $t = s_i+1$ Anwendung finden, was Ergebnisvergleiche im folgenden Abschnitt zeigen werden. Dann gilt

$$\lambda_r = \lambda \cdot q_{s_i} + \lambda \cdot (1 - B) \cdot q_{s_i+1}. \quad (5.15c)$$

5.3.1.5 Ergebnisse

Die folgenden Diagramme zeigen einige exemplarische, vergleichende Ergebnisse der Untersuchungen von Zuteilungsstrategien. Die charakteristischen Größen sind jeweils in Abhängigkeit des mittleren Angebots A je Bedienungseinheit dargestellt, welches analog zum Wartesystem GI/G/n definiert sei als

$$A = \lambda / \sum_{i=1}^g \epsilon_i. \quad (5.16)$$

Zunächst sollen am Beispiel eines Systems mit vier Teilsystemen und je sechs Warteschlangenplätzen die prinzipiellen Unterschiede

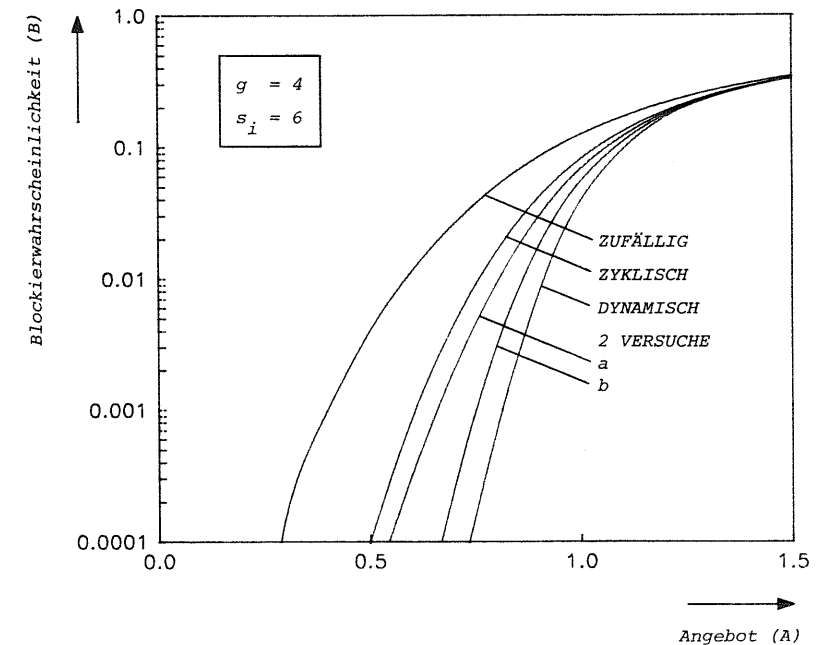


Bild 5.11: Blockierungswahrscheinlichkeit B für verschiedene Zuteilungsstrategien

der Strategien "ZUFÄLLIG", "ZYKLISCH", "DYNAMISCH" und "2 VERSUCHE" (Fälle a und b) aufgezeigt werden.

Bild 5.11 stellt die Blockierungswahrscheinlichkeiten B für die Vergleichsstrategien dar. Im normalen Lastbereich ist die beste Strategie "DYNAMISCH" diesbezüglich mehr als 100 mal besser als die Strategie "ZUFÄLLIG" und mehr als zehn mal besser als die Strategie "ZYKLISCH". Mit der Strategie "2 VERSUCHE" lassen sich ebenfalls Verbesserungen gegenüber der Strategie "ZYKLISCH" erzielen, wobei der Fall b annähernd die Qualität der Strategie "DYNAMISCH" erreicht. Somit sind die Strategien "DYNAMISCH" und der Fall b von "2 VERSUCHE" empfehlenswert, wenn eine sehr kleine Blockierungswahrscheinlichkeit gefordert wird.

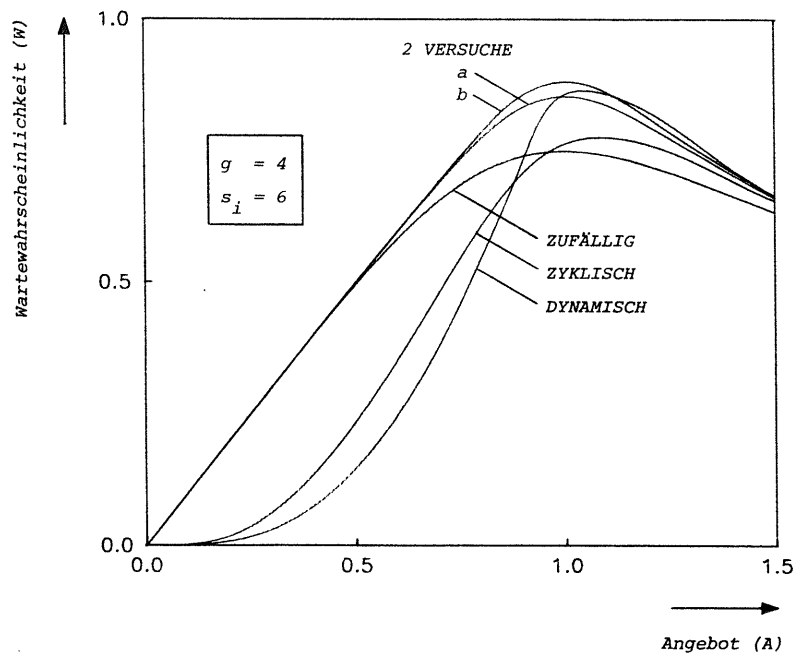


Bild 5.12: Wartewahrscheinlichkeit W für verschiedene Zuteilungsstrategien

Die Wartewahrscheinlichkeiten W sind in Bild 5.12 gezeigt. Sie weisen einen für Warteverlustsysteme typischen Verlauf auf. Der Abfall im Überlastbereich ist durch die vielen Verluste dort bedingt. Im normalen Lastbereich warten die Anforderungen bei der Strategie "DYNAMISCH" wesentlich seltener und bei der Strategie "ZYKLISCH" nur unwesentlich häufiger. Bei der Strategie "2 VERSUCHE" erkennt man sowohl im Fall a als auch b deutlich das Verhalten der Strategie "ZUFÄLLIG" im normalen Lastbereich und der Strategie "DYNAMISCH" im Überlastbereich.

Diese Tendenzen spiegeln sich in gleicher Form auch bei den mittleren Wartezeiten t_w wieder, dargestellt in Bild 5.13. Die größere Wartezeit der intelligenten Strategien im Überlastfall ist

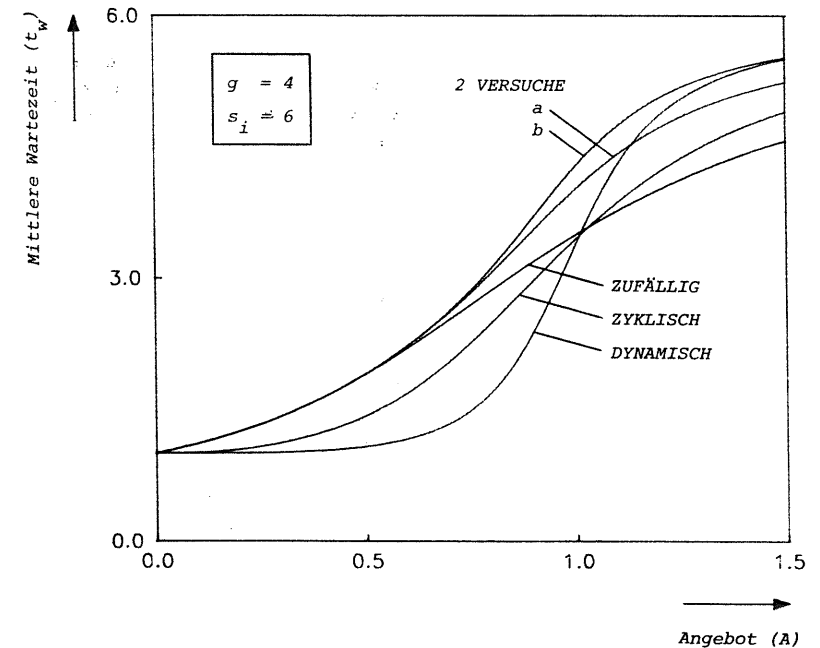


Bild 5.13: Mittlere Wartezeit t_w für verschiedene Zuteilungsstrategien

auf den höheren Durchsatz dort zurückzuführen. Bezüglich des Kriteriums Wartezeit ist auf jeden Fall wieder die Strategie "DYNAMISCH" vorzuziehen. Die Strategie "2 VERSUCHE" bringt hier im Gegensatz zur Blockierungswahrscheinlichkeit keine Vorteile.

Die Strategie "TEILDYNAMISCH" wurde in den bisherigen Vergleichen nicht berücksichtigt, da die Ergebniskurven stark vom Schwellwert t abhängen. Durch geeignete Wahl von t lassen sich nämlich Abstufungen zwischen den Strategien "DYNAMISCH" und dem Fall b von "2 VERSUCHE" einstellen. Bild 5.14 zeigt dies am Beispiel der mittleren Wartezeiten t_w für das bisher betrachtete System ($g=4$, $s_i=6$), wobei der Schwellwert t als Parameter zwischen 1 und 7 variiert. Dabei entspricht $t=1$ dem Grenzfall der Strategie "DYNAMISCH" und $t=7$ dem Grenzfall "2 VERSUCHE" (Fall b). Eine Erhöhung

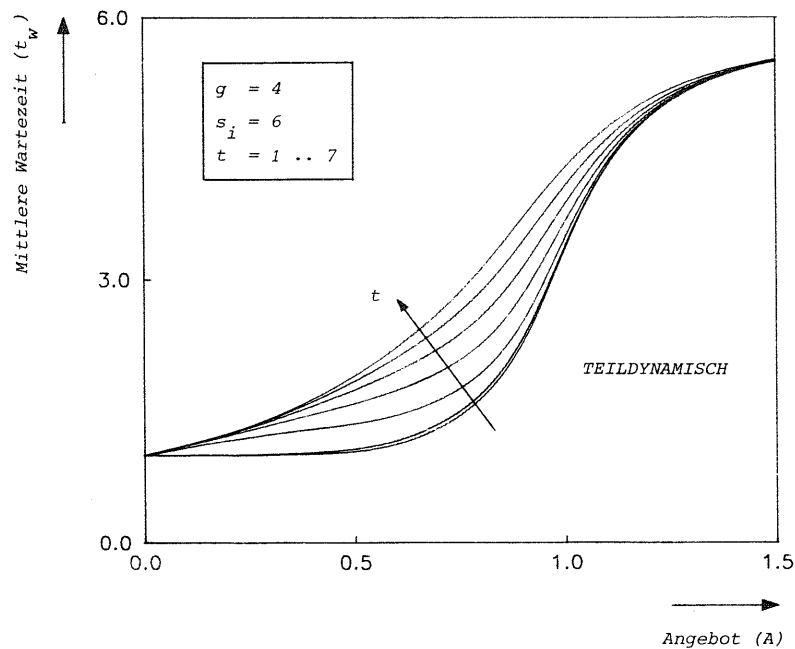


Bild 5.14: Mittlere Wartezeit t_w für Strategie "TEILDYNAMISCH"

des Schwellwertes t soll die Rückkopplungsrate λ_r und somit den Steuerungsaufwand reduzieren. Sie bringt jedoch eine Qualitätseinbuße mit sich, die bei der Blockierungswahrscheinlichkeit allerdings nicht so gravierend ausfällt, wie bei der Wartezeit.

Die gezeigten, exemplarischen Ergebnisse geben über die Abhängigkeit von der Systemgröße, d.h. die Anzahl g der Teilsysteme und s_i der Warteplätze noch keinen Aufschluß. Andere Ergebnisse zeigen jedoch, daß der relative Unterschied zwischen den Strategien im wesentlichen nur von der Anzahl g der Teilsysteme abhängt, und zwar indem die Effektivität der intelligenten Strategien mit der Anzahl g steigt. Dieser Effekt ist dadurch begründbar, daß eine intelligente Strategie umso mehr freie Plätze ausnützen kann, je

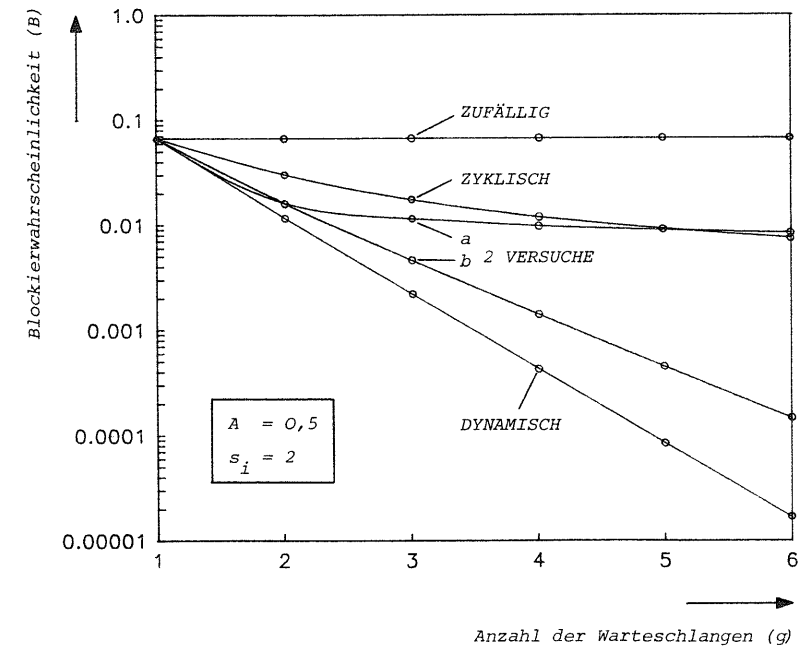


Bild 5.15: Blockierungswahrscheinlichkeit B in Abhängigkeit der Anzahl g von Teilsystemen für verschiedene Strategien

mehr Teilsysteme vorhanden sind. Dies wird am Beispiel der Blockierungswahrscheinlichkeit in Bild 5.15 für Systeme mit bis zu sechs Teilsystemen und konstantem Angebot $A=0,5$ verdeutlicht. Es ist bemerkenswert, daß die Blockierungswahrscheinlichkeit bei der Strategie "DYNAMISCH" nahezu exponentiell mit der Anzahl g fällt, wodurch auch Interpolationen für größere Systeme möglich werden.

Alle bisher betrachteten Leistungsmerkmale sprechen für den Einsatz der Strategie "DYNAMISCH". Vergleicht man jedoch den Steuerungsaufwand der Strategien, der in Bild 5.16 in Form der Rückkopplungsrate λ_r dargestellt ist, so schneidet diese Strategie im normalen Lastbereich am schlechtesten ab, da λ_r anfangs linear mit der Systembelastung ansteigt. Im Gegensatz dazu schneiden die

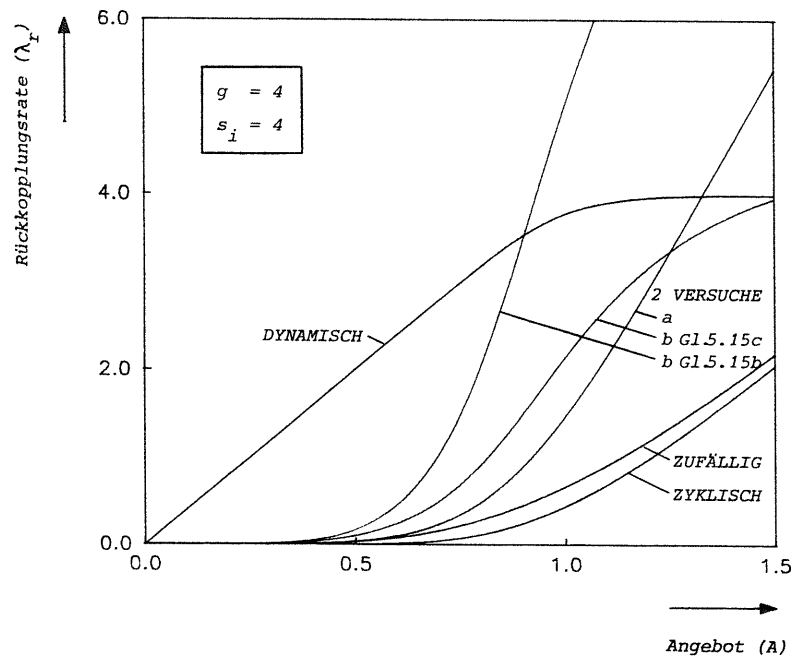


Bild 5.16: Rückkopplungsrate λ_r für verschiedene Zuteilungsstrategien

nicht intelligenten Strategien am besten ab, da λ_r mit der Blockierungswahrscheinlichkeit erst im Hochlastbereich ansteigt. Die Strategie "2 VERSUCHE", die gleichzeitig den diesbezüglich optimalen Grenzfall der Strategie "TEILDYNAMISCH" darstellt, kann mit ihren verschiedenen Implementierungsarten als akzeptabler Kompromiß betrachtet werden. Insbesondere die Implementierung des Falles b nach Gleichung 5.15c ist dabei hervorzuheben, da sie bis in den Überlastbereich günstig bleibt.

Aus den vorangehenden Ergebnissen kann man zusammenfassend folgende Schlußfolgerungen ziehen: Die Anwendung der Strategie "DYNAMISCH" ist grundsätzlich vorteilhaft, besonders bei großen Systemen. Ihr Einsatz ist sogar unbedingt erforderlich, wenn man höchste Anforderungen an die Wartezeit stellt. Kann ihr Einsatz jedoch wegen ihres Steuerungsaufwandes nicht vertreten werden, so bietet sich eine Version der Strategien "TEILDYNAMISCH" oder "2 VERSUCHE" an, die allerdings nur bezüglich der Blockierungswahrscheinlichkeit Vorteile bieten.

Im Abschnitt über die Analyse von Zuteilungsstrategien wurde die exakte Lösung für alle Strategien und Parameterfälle aufgezeigt, die jedoch bei großen Systemen mit intelligenten Strategien sehr aufwendig werden kann. Es liegt deshalb nahe, exakte Ergebnisse auch mit Approximationen zu vergleichen. Da die zur Verfügung stehenden Warteplätze und Bedienungseinheiten von der Strategie "DYNAMISCH" offensichtlich sehr gut ausgenutzt werden, liegt ein Vergleich mit einem Warteverlustsystem des Typs M/M/n-s mit gleicher Anzahl von Bedienungseinheiten und Warteplätzen nahe, d.h. mit $n=g$ und $s=\sum s_i$. Diesbezügliche Ergebnisse zeigen eine sehr genaue Übereinstimmung der Blockierungs- und eine gute Übereinstimmung der Wartewahrscheinlichkeiten. Lediglich die mittleren Wartezeiten weisen natürlich bei geringer Last bei dem System M/M/n-s einen um den Faktor n kleineren Wert auf. Somit stellen die Gleichungen 4.23 und 4.24 für die Blockierungs- bzw. Wartewahrscheinlichkeit des Systems M/M/n-s auch für die Zuteilungsstrategie "DYNAMISCH" eine gute Approximation dar.

5.3.2 Analyse von Abfertigungsstrategien

5.3.2.1 Abfertigungsstrategien

Analog zu einigen der untersuchten Zuteilungsstrategien können auch Abfertigungsstrategien mit prinzipiell entsprechenden Mechanismen definiert werden:

- ZUFÄLLIG :

Die Folge der Abfertigungen der Warteschlangen ist zufallsabhängig und gleichverteilt unter allen Warteschlangen.

- ZYKLISCH :

Die Folge der Abfertigungen der Warteschlangen ist rein zyklisch, d.h. alle Warteschlangen werden der Reihe nach abwechselnd abgefertigt.

- DYNAMISCH :

Die Abfertigung der Warteschlangen erfolgt dynamisch, d.h. die nächste abzufertigende Warteschlange ist die längste Warteschlange. Falls mehrere Warteschlangen den gleichen höchsten Füllgrad aufweisen, erfolgt die Abfertigung unter diesen gemäß der Strategie "ZUFÄLLIG".

- TEILDYNAMISCH :

Die Abfertigung der Warteschlangen erfolgt nur teilweise dynamisch, d.h. solange noch keine Warteschlange einen definierten Füllgrad übersteigt, wird gemäß der Strategie "ZUFÄLLIG" abgefertigt. Ansonsten erfolgt die Abfertigung gemäß der Strategie "DYNAMISCH".

Die beiden ersten Strategien werden wiederum als nicht intelligente Abfertigungsmechanismen den beiden anderen intelligenten, zustandsabhängigen Verfahren gegenübergestellt. Die Strategie "ZYKLISCH" ist in der Praxis am meisten eingesetzt, wobei aller-

dings oft eine Priorisierung einzelner Warteschlangen durch eine fest definierte mehrfache Abfertigung pro Zyklus eingeführt werden muß. Die Strategien "DYNAMISCH" und "TEILDYNAMISCH" haben den Vorteil, daß sie eventuell notwendige und vielleicht auch unbekannte Priorisierungen automatisch und sogar dynamisch einstellen. Aus diesem Grund kann der Einsatz dieser Strategien bereits ohne die Berücksichtigung eventuell besserer Leistungsmerkmale sinnvoll sein, was auch schon in der Simulationsstudie am Vermittlungssystem PILOT (siehe Abschnitt 5.2) gezeigt wurde.

5.3.2.2 Grundmodell

In Bild 5.17 ist das Grundmodell für die Warteschlangenanalyse von Abfertigungsstrategien dargestellt.

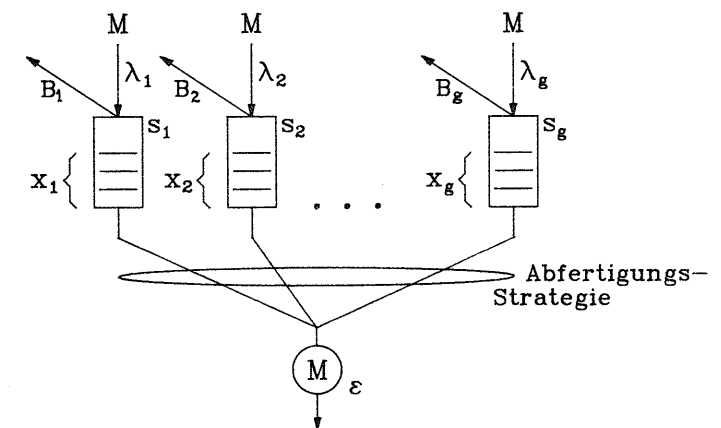


Bild 5.17: Grundmodell für die Analyse von Abfertigungsstrategien

Im Hinblick auf eine Markoffsche Zustandsanalyse wird bei jeder Warteschlange i ein Markoff'scher Ankunftsprozeß mit der Rate λ_i vorausgesetzt. Die Kapazität jeder Warteschlange i ist s_i und

die Abfertigung der g Warteschlangen geschieht durch eine Bedienungseinheit mit einem Markoffschen Bedienungsprozeß und der Bedienungsenderate ϵ . Die Zustandsvariable x_i repräsentiert die Anzahl der Anforderungen in der Warteschlange i . Der gesamte Systemzustand wird durch den Zustandsvektor $\underline{x} = (x_1, x_2, \dots, x_g, x_{BE})$ beschrieben, wobei x_{BE} den Zustand der Bedienungseinheit repräsentiert, der jedoch nur die Zustände 1 (belegt) oder 0 (frei) annehmen kann.

5.3.2.3 Analyse

In der Literatur sind zahlreiche analytische Untersuchungen von Abfertigungs- bzw. Polling-Systemen bekannt, z.B. /39/ und /40/. Sie behandeln meist die zyklische Abfertigung mit oder ohne Prioritäten, und es existieren exakte und approximative Lösungen für Systeme mit i.a. unendlichen, aber zum Teil auch endlichen Warteschlangen. Ferner wird häufig die Umschaltzeit mit modelliert, die die abfertigende Einheit braucht, um von einer Warteschlange zur nächsten zu kommen. Diese Ergebnisse könnten jedoch hier prinzipiell nur für die Strategie "ZYKLISCH" herangezogen werden, während zur Untersuchung der zustandsabhängigen Abfertigungsstrategien, die eventuell nur feine Unterschiede zeigen, die exakte Markoffsche Zustandsanalyse zugrunde gelegt werden muß. Diese soll im folgenden jedoch für alle betrachteten Strategien angewendet werden, da zwecks Reduzierung des Umfangs der Zustandsräume auf eine Berücksichtigung der Umschaltzeit verzichtet wird.

Der Lösungsweg für die hier betrachteten Abfertigungsstrategien ist prinzipiell ähnlich wie die für Zuteilungsstrategien im Abschnitt 5.3.1.3 beschriebene Analysetechnik. Es wird deshalb an dieser Stelle auf eine genaue Ausführung der Analyse verzichtet. Es soll lediglich anhand eines einfachen Beispiels die prinzipielle Struktur der Zustandsräume gezeigt werden. In Bild 5.18 ist der Zustandsraum mit den Zustandsübergängen für die Abfertigungsstrategie "DYNAMISCH" eines Systems mit zwei Warteschlangen zu jeweils drei Warteplätzen dargestellt. Die charakteristischen Leistungsgrößen ermitteln sich dann aus den Zustandswahrschein-

lichkeiten durch ähnliche Beziehungen wie bei den Zuteilungsstrategien.

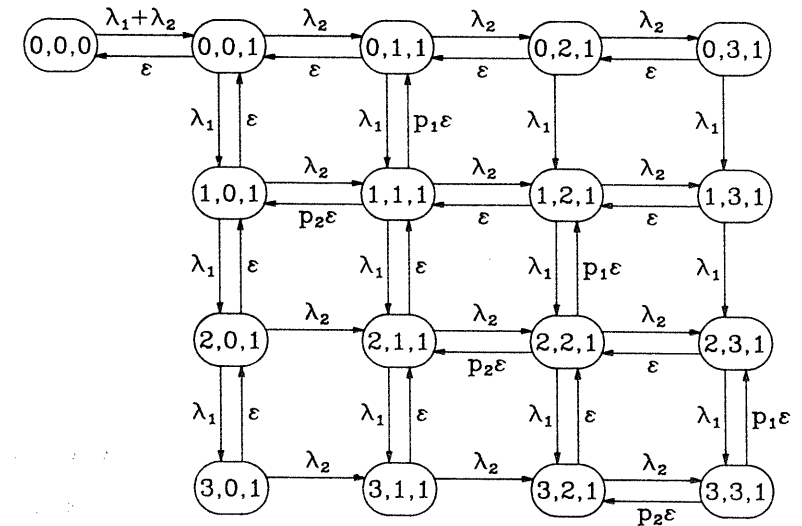


Bild 5.18: Beispiel eines Zustandsraums für die Abfertigungsstrategie "DYNAMISCH" ($g = 2$, $s_i = 3$)

5.3.2.4 Ergebnisse

Im folgenden werden einige exemplarische Ergebnisse zum Vergleich der verschiedenen Abfertigungsstrategien gezeigt, und zwar für ein System mit vier Warteschlangen und jeweils vier Warteplätzen. Dieses System wird sowohl bei symmetrischer Last als auch bei unsymmetrischer Last betrachtet. Die gezeigten charakteristischen Größen sind in Abhängigkeit des Angebots des Gesamtsystems dargestellt, welches definiert ist als

$$A = \sum_{i=1}^g \lambda_i / \epsilon \quad (5.17)$$

In Bild 5.19 ist die bei symmetrischer Last für alle Warteschlangen einheitliche Blockierungswahrscheinlichkeit B_i für die Strategien "ZUFÄLLIG", "ZYKLISCH" und "DYNAMISCH" dargestellt. Die Ergebnisse zeigen, daß die Abfertigungsstrategien prinzipiell wesentlich geringere Unterschiede aufweisen als die Zuteilungsstrategien. Dieser Effekt ist durch die gemeinsame Bedienungseinheit erklärbar, denn den Durchsatz verbessernde und durch die Abfertigungsstrategie bedingte Eingriffe in ein Teilsystem haben auf die anderen Teilsysteme einen negativen Einfluß und heben sich somit gegenseitig teilweise wieder auf. Deshalb erreicht man mit der Strategie "DYNAMISCH" nur geringfügige Verbesserungen der Blockierungswahrscheinlichkeit gegenüber den anderen Strategien, so daß sich ihr Einsatz bei symmetrischer Last selten lohnt.

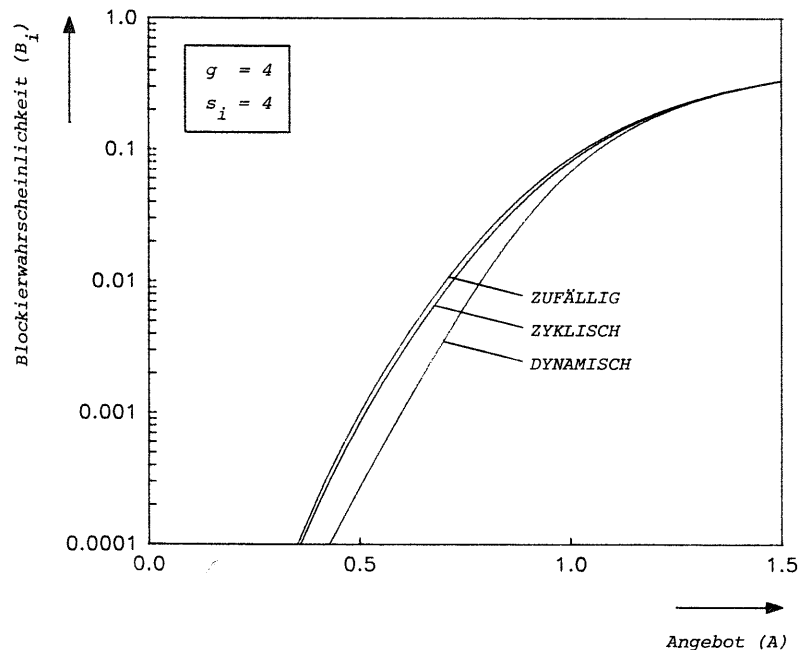


Bild 5.19: Blockierungswahrscheinlichkeit B_i für verschiedene Abfertigungsstrategien

Die mittlere Wartezeit t_{wi} in jeder Warteschlange i ist, bedingt durch den geringfügig höheren Durchsatz bei der Strategie "DYNAMISCH", grundsätzlich etwas höher als bei den nicht intelligenten Strategien, siehe Bild 5.20. Dies wirkt sich allerdings erst im Hochlastfall aus. Prinzipiell lohnt sich also der Einsatz intelligenter Abfertigungsstrategien aus Gründen der Wartezeit nicht.

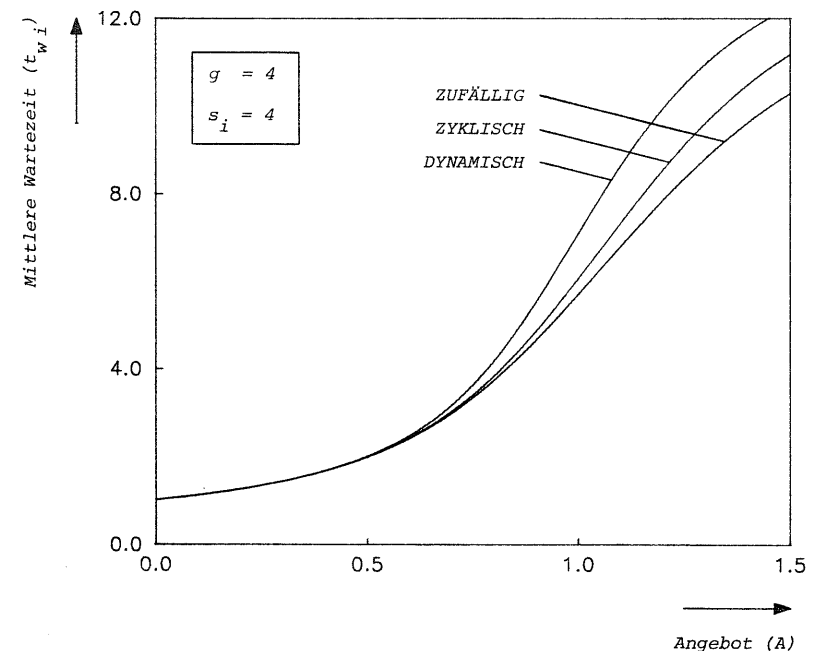


Bild 5.20: Mittlere Wartezeit t_{wi} für verschiedene Abfertigungsstrategien

Auf eine Darstellung der Ergebnisse für die Strategie "TEILDYNAMISCH" wird an dieser Stelle verzichtet, denn diese liegen je nach Wahl des definierten Schwellwertes des Warteschlangenfüllgrades zwischen den Strategien "ZUFÄLLIG" und "DYNAMISCH", die aber ohnehin nur geringe Unterschiede aufweisen.

Im folgenden wird das Beispielsystem mit unsymmetrischer Last betrachtet, z.B. für den Fall einer doppelt so hohen Ankunftsrate λ_1 bei der ersten Warteschlange im Vergleich zu den Ankunftsrate $\lambda_{2..4}$ der anderen drei Warteschlangen. Aufgrund der geringen Unterschiede der Strategien "ZUFÄLLIG" und "ZYKLISCH" sollen nur noch die Fälle "DYNAMISCH" und "ZUFÄLLIG" verglichen werden. In Bild 5.21 sind die Blockierungswahrscheinlichkeiten in Abhängigkeit des Angebots A dargestellt. Bei der höher belasteten Warteschlange ergibt sich durch die dynamische Strategie eine Verbesserung, jedoch nur von der Größenordnung wie bei symmetrischer Last. Bei den weniger belasteten anderen Warteschlangen ergibt sich im normalen Lastbereich ebenfalls noch eine Verbesserung, aber im höheren Lastbereich eine geringfügige Verschlechterung.

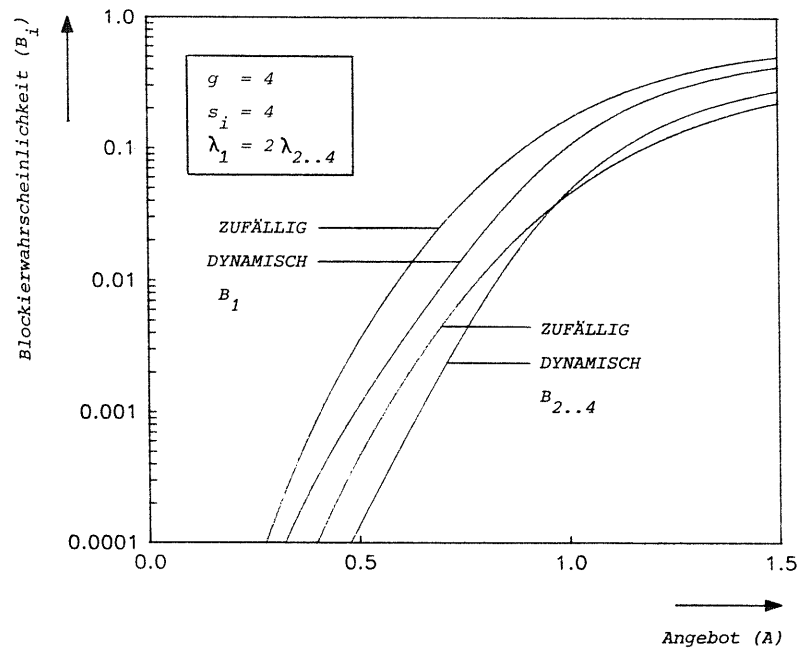


Bild 5.21: Blockierungswahrscheinlichkeiten B_i für die Strategien "ZUFÄLLIG" und "DYNAMISCH" (Last 2:1:1:1)

Betrachtet man in Bild 5.22 eine extremere Unsymmetrie, z.B. eine zehnfache Last der ersten Warteschlange im Vergleich zu den anderen, so reduziert sich die durch die Strategie "DYNAMISCH" bedingte Verbesserung bei der ersten Warteschlange, und die Verschlechterung bei den anderen Warteschlangen wird auch schon im normalen Lastbereich größer. Dieser Effekt ist dadurch begründbar, daß die Bedienungseinheit bei beiden Strategien wegen der vernachlässigten Umschaltzeit im wesentlichen nur noch für die erste Warteschlange arbeitet. Die Verschlechterung bei den weniger belasteten Warteschlangen ist allerdings nicht schwerwiegend, da deren Blockierungswahrscheinlichkeit ohnehin wesentlich kleiner ist.

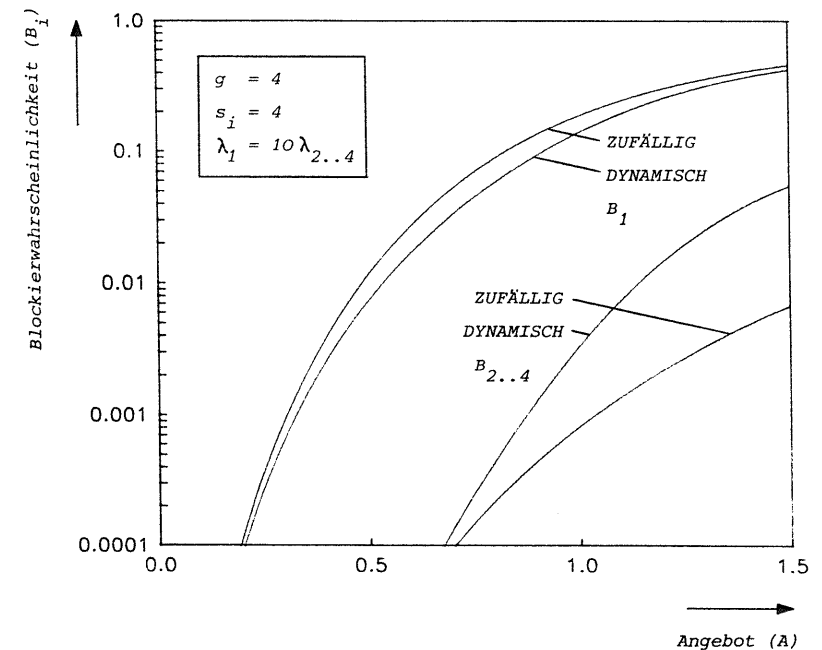


Bild 5.22: Blockierungswahrscheinlichkeiten B_i für die Strategien "ZUFÄLLIG" und "DYNAMISCH" (Last 10:1:1:1)

Für das letztere Lastverhältnis ist in Bild 5.23 schließlich die mittlere Wartezeit der wartenden Anforderungen über dem Angebot des Gesamtsystems dargestellt. Bei der höher belasteten ersten Warteschlange kann durch die Strategie "DYNAMISCH" eine kleine Verbesserung erzielt werden, während dafür bei den anderen Warteschlangen im höheren Lastbereich eine erhebliche Erhöhung der Wartezeit in Kauf genommen werden muß. Diese Tendenz ist prinzipiell auch bei geringerer Unsymmetrie der Last in abgeschwächter Form erkennbar. Somit kann im Gegensatz zur symmetrischen Last der Einsatz einer intelligenten Abfertigungsstrategie auch aus Gründen der Wartezeit eventuell sinnvoll sein, wenn die Wartezeit in hochbelasteten Warteschlangen optimiert werden muß.

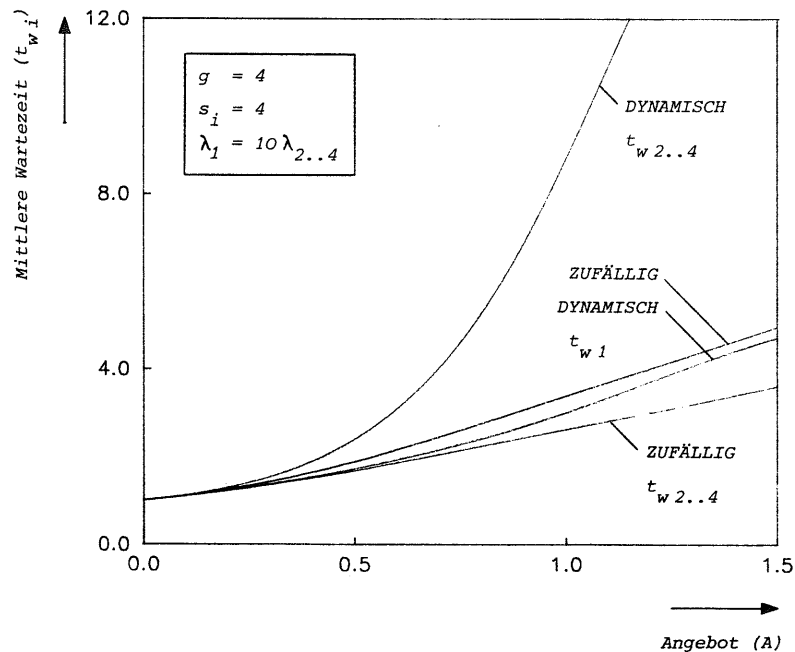


Bild 5.23: Mittlere Wartezeiten t_{wi} für die Strategien "ZUFÄLLIG" und "DYNAMISCH" (Last 10:1:1:1)

Aus den erzielten Ergebnissen können zusammenfassend folgende Schlußfolgerungen gezogen werden, die jedoch nur für Systeme mit einer gegenüber der Bedienungszeit vernachlässigbaren Umschaltzeit gelten: Die Unterschiede der Strategien sind relativ gering. Die Strategie "DYNAMISCH" ist nur bezüglich der Blockierungswahrscheinlichkeit die beste Strategie. Lediglich bei unsymmetrischer Last können auch bezüglich der mittleren Wartezeit bei hochbelasteten Warteschlangen leichte Verbesserungen erzielt werden. Da der Verwaltungsaufwand für die intelligenten Strategien vergleichbar mit dem der entsprechenden Zuteilungsstrategien (siehe Abschnitt 5.3.1.5) ist, erscheint der Einsatz einer intelligenten Abfertigungsstrategie unter der gemachten Voraussetzung bezüglich der Umschaltzeit nur in besonders kritischen Fällen sinnvoll.

Die Simulationsstudie am System PILOT (siehe Abschnitt 5.2) zeigt jedoch, daß bei nicht mehr zu vernachlässigender Umschaltzeit ein Einsatz intelligenter Strategien wieder sinnvoll wird. Falls man den Implementierungsaufwand nicht in Kauf nehmen will, bietet sich zumindest der Einsatz einer intelligenten Strategie in der Simulation an, woraus sich dann durch Messung der Abfertigungshäufigkeiten der einzelnen Warteschlangen die sich automatisch einstellenden Prioritäten ergeben, die dann für eine zyklische Abfertigung mit statischen Prioritäten verwendet werden können.

6 ZUSAMMENFASSUNG

Im ersten Kapitel wurden einleitend die Gründe für die Verteilung der Steuerung von SPC-Vermittlungssystemen genannt und die grundsätzlichen Probleme aufgezeigt, die mit der Strukturierung und Leistungsuntersuchung verteilter Vermittlungssteuerungen verbunden sind.

Die Konzeption von SPC-Vermittlungssystemen wurde im zweiten Kapitel allgemein behandelt. Dazu wurden sowohl die wichtigsten Steuerungsfunktionen von SPC-Vermittlungssteuerungen als auch die prinzipiell sinnvollen zentralen, dezentralen und teilzentralen Steuerungsstrukturen zusammengestellt und deren Vor- und Nachteile abgewägt. Die prinzipiellen Möglichkeiten für eine Aufteilung der Steuerungsfunktionen nach Last- oder Funktions Gesichtspunkten und die daraus resultierenden Probleme wurden dazu diskutiert.

Im dritten Kapitel wurde schließlich anhand des als Labormodell implementierten Dienst-integrierten Ortsvermittlungssystems "PILOT" ein Realisierungsbeispiel für verteilte Vermittlungssteuerungen gegeben. Aus steuerungstechnischer Sicht wurden die Entwicklungsschritte und die endgültige Struktur sowie die Funktionen dieses Systems beschrieben. Ein Schwerpunkt bildete insbesondere die Interprozessor-Kommunikation in diesem System. Ein Protokoll für die Abfertigung von Sendewarteschlangen, welches gleichzeitig die Möglichkeit einer flexiblen und dynamischen Last- bzw. Funktionszuteilung zu den Empfangswarteschlangen erlaubt, wurde vorgestellt.

Das vierte Kapitel befasste sich allgemein mit der Modellierung von Vermittlungssteuerungen und deren Umwelt und behandelte speziell drei Analysetechniken, die sich für die Beurteilung und Leistungsuntersuchung verteilter Vermittlungssteuerungen bzw. deren Strukturen besonders eignen. Es wurden ein rechnerunterstütztes Verfahren für Signalflußanalysen, ein zeittreues Simulationsverfahren für verteilte Vermittlungssteuerungen mit einem in SDL spezifizierten Ablaufgeschehen und die Methode der stationären

Warteschlangenanalyse Markoffscher Teilmodelle behandelt.

Für die drei im vierten Kapitel beschriebenen Analysetechniken wurden im fünften Kapitel konkrete Anwendungen gezeigt und deren Ergebnisse vorgestellt. Für eine allgemeine teilzentrale Systemstruktur und ein verallgemeinertes Funktionsszenario wurde eine Signalflußanalyse durchgeführt, aus der allgemeine Aussagen über die Strukturierung und Konfigurierung von Systemen gefolgert werden konnten. Eine detaillierte, zeittreue Simulationsstudie des im dritten Kapitel vorgestellten Systems PILOT stellte die Leistungsfähigkeit dieses Systems bei Hochlast unter Beweis. Gleichzeitig wurden damit die Auswirkungen eines durch dynamische Prioritäten verbesserten Abfertigungsverfahrens der Sendewarteschlangen durch die Bussteuerung des Systems gezeigt. Mittels Markoffscher Zustandsanalysen wurden schließlich Teilmodelle untersucht, welche intelligente Strategien für die Zuteilung der Anforderungen zu Lastteilungsprozessoren sowie intelligente Abfertigungsmechanismen der Sendewarteschlangen in einem System beschreiben. Die Ergebnisse zeigten die Wirksamkeit verschiedener Strategien und gaben darüber Aufschluß, unter welchen Voraussetzungen der Einsatz bestimmter Strategien lohnenswert ist.

