

## Informatik-Fachberichte 165

---

Herausgegeben von W. Brauer  
im Auftrag der Gesellschaft für Informatik (GI)

Hans-Ulrich Hei

# berlast in Rechensystemen

Modellierung und Verhinderung



Springer-Verlag  
Berlin Heidelberg New York  
London Paris Tokyo

**Autor**

Hans-Ulrich Heiß  
Institut für Betriebs- und Dialogsysteme, Universität Karlsruhe  
Kaiserstraße 12, 7500 Karlsruhe 1

CR Subject Classifications (1987): C.4, I.6

ISBN-13: 978-3-540-18944-2 e-ISBN-13: 978-3-642-73439-7  
DOI: 10.1007/ 978-3-642-73439-7

CIP-Titelaufnahme der Deutschen Bibliothek

**Heiß, Hans-Ulrich:**

Überlast in Rechensystemen: Modellierung u. Verhinderung / Hans-Ulrich Heiß. –  
Berlin ; Heidelberg ; New York ; London ; Paris ; Tokyo : Springer, 1988  
(Informatik-Fachberichte; 165)

Zugl.: Karlsruhe, Univ., Diss., 1987 u. d. T.: Heiß, Hans-Ulrich: Verhinderung von  
Überlast in Rechensystemen

NE: GT

Dieses Werk ist urheberrechtlich geschützt. Die dadurch begründeten Rechte, insbesondere die der Übersetzung, des Nachdrucks, des Vortrags, der Entnahme von Abbildungen und Tabellen, der Funksendung, der Mikroverfilmung oder der Vervielfältigung auf anderen Wegen und der Speicherung in Datenverarbeitungsanlagen, bleiben, auch bei nur auszugsweiser Verwertung, vorbehalten. Eine Vervielfältigung dieses Werkes oder von Teilen dieses Werkes ist auch im Einzelfall nur in den Grenzen der gesetzlichen Bestimmungen des Urheberrechtsgesetzes der Bundesrepublik Deutschland vom 9. September 1965 in der Fassung vom 24. Juni 1985 zulässig. Sie ist grundsätzlich vergütungspflichtig. Zuwiderhandlungen unterliegen den Strafbestimmungen des Urheberrechtsgesetzes.

© by Springer-Verlag Berlin Heidelberg 1988

2145/3140 – 543210

# Vorwort

Das vorliegende Buch ist eine geringfügig gekürzte und überarbeitete Fassung meiner Dissertation, die unter dem Titel '*Verhinderung von Überlast in Rechensystemen*' von der Fakultät für Informatik der Universität Karlsruhe angenommen wurde.

Es befaßt sich mit einem Phänomen, das jedem Leser aus dem Straßenverkehr vertraut sein wird: daß ein technisches System (*Straßennetz*) einen Leistungskollaps erleiden kann, wenn es zu sehr belastet wird (*zu viele Fahrzeuge*). Während zunächst ein *Mehr* an Fahrzeugen auch zu einem *Mehr* an Durchsatz führt, wird irgendwann ein kritischer Punkt erreicht, an dem das Verhalten umkippt, der Durchsatz absinkt und in der Folge eventuell vollständig zum Erliegen kommt.

Vergleichbare Stau- oder Überlasteffekte finden sich auch in Rechensystemen, so etwa im virtuellen Speicher, in Datennetzen oder in Datenbanksystemen. Die vorliegende Arbeit widmet sich diesen Effekten in zweierlei Weise: in einem mehr analytischen Teil, in dem derartige Verhalten mit den Mitteln der Warteschlangentheorie beschrieben wird, und in einem mehr konstruktiven Teil, in dem Mechanismen zur Verhinderung solcher Leistungseinbrüche entwickelt werden.

In einem einleitenden Kapitel wird das Phänomen anhand einiger Beispiele näher beschrieben, bevor im zweiten Kapitel die Einzelstation mit Überlastverhalten im Hinblick auf Stabilität und elementare Mechanismen zur Lastkontrolle untersucht wird. Dieser Ansatz wird dann im dritten Kapitel auf Warteschlangennetze übertragen, wobei mehrere Freiheitsgrade zu berücksichtigen sind. Mit dem vierten Kapitel beginnt der Teil, der dem '*online*'-Einsatz von Regelungsmechanismen zur Lastkontrolle gewidmet ist. Dabei werden zunächst die Probleme bei der Beherrschung dynamischen Verhaltens aufgezeigt, um darauf aufbauend im fünften Kapitel konkrete Verfahren vorzuschlagen, die im Kapitel 6 durch Simulation vergleichend untersucht und bewertet werden. Das siebte und letzte Kapitel zeigt dann, wie die erforderlichen Meß- und Stellglieder in die betroffenen Programmstücke einer Betriebsmittelverwaltung einzubauen sind.

Damit wird versucht, nicht nur den Theoretiker anzusprechen, sondern auch dem Praktiker einerseits zu einem verbesserten grundsätzlichen Verständnis des so häufig vorzufindenden Überlastphänomens zu verhelfen, andererseits ihm Anregungen dafür zu geben, wie er Meß- und Regelverfahren in modularer und effizienter Weise realisieren kann, zumal derartige Mechanismen im Zuge der Verbreitung verteilter Systeme immer größere Bedeutung gewinnen.

Ich möchte nicht versäumen, an dieser Stelle einigen Personen für ihre Mithilfe und Unterstützung zu danken:

Mein besonderer Dank gilt meinem Betreuer, Herrn Professor Dr.-Ing. Horst Wettstein, der mir die Anfertigung dieser Arbeit ermöglichte und dessen Anregungen zu ihrer Ab-  
rundung beigetragen haben. Ebenso danke ich Herrn Professor Dr. Adolf Schreiner für die Übernahme des Korreferats und seine nützlichen Hinweise zur schriftlichen Aus-  
gestaltung. Auch dem mir unbekanntem Gutachter, der die Arbeit im Hinblick auf die  
Veröffentlichung in den Informatik-Fachberichten des Springer-Verlags geprüft hat,  
verdanke ich hilfreiche Anmerkungen.

Unter den Kollegen fand ich im Frühstadium der Arbeit in Dr. Günter Totzauer einen  
kompetenten Gesprächspartner auf dem Felde der Warteschlangentheorie, in Alfred  
Schätter einen kritischen Prüfer des Beweises des Optimalitätstheorems. Hilfreiche  
Unterstützung leistete auch Herr cand.inform. Wolfgang Gerteis bei der Implementie-  
rung des Simulationsmodells. Ihnen sei an dieser Stelle herzlich gedankt.

Mein Dank gilt auch der IBM Deutschland GmbH, die im Rahmen eines Koopera-  
tionsprojektes Hard- und Software zur Verfügung stellte, ohne die manches in dieser  
Arbeit mühsamer hätte ablaufen müssen.

Karlsruhe, im Januar 1988

Hans-Ulrich Heiß

# Inhalt

<b>1 Überlast in Rechensystemen</b>	<b>1</b>
1.1 Das Überlastphänomen	1
1.2 Ursachen der Überlast	3
1.3 Das Überlastproblem in der Literatur	8
1.4 Motivation, Zielsetzung und Vorgehensweise	12
<b>2 Die Einzelstation mit Überlast</b>	<b>15</b>
2.1 Das unbegrenzte System	15
2.2 Die Einzelstation mit Abweisung	24
2.2.1 Eigenschaften	24
2.2.2 Durchsatzbegrenzende Faktoren	29
2.2.3 Ansätze zur Regelung	31
2.3 Einzelstation mit Aufstauen	33
<b>3 Netze von Überlaststationen</b>	<b>35</b>
3.1 Allgemeine Eigenschaften	35
3.1.1 Typisierung und Beschreibung	35
3.1.2 Lokale und globale Lastabhängigkeit	36
3.1.3 Reduzierbarkeit von Netzen durch Norton's Theorem	39
3.2 Offene Netze	41
3.2.1 Offene Netze ohne Lastkontrolle	41
3.2.2 Überlastkontrolle durch Abweisung	44
3.2.2.1 Lokale Lastabhängigkeit	44
3.2.2.2 Globale Lastabhängigkeit	45
3.2.3 Überlastkontrolle durch Aufstauen	46
3.2.3.1 Lokale Lastabhängigkeit	46
3.2.3.2 Globale Lastabhängigkeit	47
3.3 Geschlossene Netze	48
3.3.1 Das geschlossene Netz ohne Lastkontrolle	48
3.3.2 Überlastkontrolle durch Aufstauen	55
3.3.2.1 Lokale Lastabhängigkeit	55
3.3.2.2 Globale Lastabhängigkeit	56
3.4 Optimale Lastkontrolle	56
3.4.1 Globale optimale Überlastkontrolle	56
3.4.2 Lokale optimale Überlastkontrolle	57

<b>4 Das Überlastproblem unter dynamischen Aspekten</b>	<b>61</b>
4.1 Dynamisches Verhalten von Rechensystemen	61
4.1.1 Notwendigkeit dynamischer Sichtweise	61
4.1.2 Modellierung des dynamischen Verhaltens von Rechensystemen	62
4.2 Das dynamische Überlastproblem	64
4.3 Das dynamische Überlastproblem aus regelungstheoretischer Sicht	67
4.4 Messung von Leistungsgrößen in Rechensystemen	70
<b>5 Konkrete Verfahren</b>	<b>80</b>
5.1 Direkte Optimum-Suchverfahren	80
5.1.1 Inkrementelle Testschritte	80
5.1.2 Tripel-Verfahren	84
5.2 Verfahren mit expliziter Modellbildung	87
5.2.1 Einsatz von Warteschlangenmodellen in Regelmechanismen	87
5.2.2 Polynomansätze	88
5.2.2.1 Parameterschätzung auf der Basis der 'kleinsten Quadrate'	89
5.2.2.2 Linearer Ansatz	94
5.2.2.3 Quadratischer Ansatz	95
5.3 Adaptionfähigkeit zweiter Stufe	97
<b>6 Simulationsergebnisse</b>	<b>100</b>
6.1 Das Simulationsmodell	101
6.2 Dynamische Lage des Optimums	106
6.2.1 Driftartige Änderungen	106
6.2.2 Sprungartige Änderungen	112
6.3 Konstante Lage des Optimums	118
6.3.1 Stationäres Verhalten	118
6.3.2 Niveauänderungen des Optimums	118
6.4 Nachbemerkungen	119
<b>7 Realisierungsaspekte</b>	<b>120</b>
7.1 Überlastkontrolle als Aufgabe der Betriebsmittelverwaltung	120
7.2 Einbettung von Meßoperationen	124
7.2.1 Elementargrößen	127
7.2.1.1 Synchrone Auswertung	128
7.2.1.2 Asynchrone Auswertung	129
7.2.2 Zusammengesetzte Größen	135
7.2.2.1 Dezentrale Auswertung	138
7.2.2.2 Zentrale Auswertung	139
7.3 Einbettung von Regelungsoperationen	140
7.3.1 Synchrone Einbettung	141
7.3.2 Asynchrone Einbettung	142
7.3.3 Wirkungsweise des Reglereingriffs	143
7.4 Nachbemerkungen	145

<b>8 Zusammenfassung und Ausblick</b> .....	<b>147</b>
<b>Literatur</b> .....	<b>151</b>
<b>Anhang A: Abschätzung der Meßintervallängen</b> .....	<b>165</b>
Durchsatz .....	166
Auslastung .....	169
<b>Anhang B: Verzeichnis der Symbole</b> .....	<b>175</b>