

## Informatik-Fachberichte 242

---

Herausgeber: W. Brauer  
im Auftrag der Gesellschaft für Informatik (GI)

Dieter Wybranietz

# Multicast-Kommunikation in verteilten Systemen



Springer-Verlag  
Berlin Heidelberg New York  
London Paris Tokyo Hong Kong

**Autor**

Dieter Wybranietz

TELENORMA - Bosch Telecom, Zentrale Entwicklung  
Mainzer Landstraße 128-146, 6000 Frankfurt/Main 1

CR Subject Classifications (1987): C.2.4, D.3.3, D.4.4

ISBN-13: 978-3-540-52551-6 e-ISBN-13: 978-3-642-75689-4

DOI: 10.1007/978-3-642-75689-4

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, bei auch 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.

© Springer-Verlag Berlin Heidelberg 1990

2145/3140-543210 – Gedruckt auf säurefreiem Papier

# Vorwort

Der Datenaustausch zwischen Prozessen und damit die Abwicklung der Interprozeßkommunikation spielt in verteilten Systemen eine zentrale Rolle. Als Ergänzung zu den bekannten Formen der 1:1-Kommunikation, bei denen aus der Sicht eines Senders stets nur ein einziger Empfänger für eine Nachricht vorhanden ist, werden in dieser Arbeit 1:n-Kommunikationen (Multicasts) untersucht, bei denen mit einer Botschaft mehrere Prozesse erreicht werden können. Der Multicast kann als die allgemeine Kommunikationsform aufgefaßt werden, von der Broadcast und Unicast (1:1-Kommunikation) Spezialfälle darstellen: Beim Unicast enthält die Multicast-Gruppe nur einen einzigen Empfänger, beim Broadcast alle potentiellen Empfänger einer zugrundegelegten Menge.

Bei der Festlegung der Semantik von Multicasts treten im Vergleich zu 1:1-Kommunikationen zusätzliche Schwierigkeiten auf, die ihre Ursache in der Existenz mehrerer Empfänger für eine abgesandte Nachricht haben. Es empfiehlt sich daher, bei Multicast-Operationen vom Zuverlässigkeitsgrad anstatt einfach von der Zuverlässigkeit zu sprechen. Der Begriff "Zuverlässigkeit von Multicast-Operationen" wird durch eine formale Definition verschiedener Zuverlässigkeitsgrade präzisiert, so daß eine Basis für die Klassifikation von Multicast-Protokollen geschaffen wird. Die Protokollklassen reichen von 0-zuverlässigen über atomar-zuverlässige bis hin zu ordnungserhaltenden Multicasts.

Mit Hilfe eines Simulators wurden ausgewählte Vertreter der Protokollklassen unter verschiedenen Arbeitslasten bezüglich einer Reihe von Kriterien miteinander verglichen, wie etwa Nachrichten- und Zeitkomplexität, Belastung des Übertragungsmediums, Verlustaten von Nachrichten, Rechenzeitverbrauch und Einfluß des Zuverlässigkeitsgrads auf Ausführungszeiten.

In die verteilte Systemimplementierungssprache LADY wurde eine k-zuverlässige Multicast-Kommunikationsmöglichkeit integriert und auf einem experimentellen Mehrrechner-system implementiert. Aufgrund der durch Messungen im Experimentalsystem und durch Simulationen gewonnenen Resultate wurde ein besseres Verständnis der Problematik von Multicast-Kommunikationen erreicht, das zu neuen Vorschlägen für leistungsfähige Protokolle führte.

Die vorliegende Arbeit entstand während meiner Tätigkeit als Mitarbeiter im Sonderforschungsbereich 124 "VLSI-Entwurf und Parallelität" im Fachbereich Informatik der Universität Kaiserslautern.

An dieser Stelle möchte ich meinen Kollegen für viele anregende Gespräche und Diskussionen danken - insbesondere Peter Buhler, Dr. Dieter Haban, Dr. Friedemann Mattern und Peter Sturm. Mein besonderer Dank gilt Herrn Prof. Dr. J. Nehmer für die langjährige Unterstützung und Förderung meiner Arbeit sowie seine Anregungen und fortwährende Diskussionsbereitschaft wie auch Herrn Prof. Dr. Th. Härder für die gründliche Durchsicht des Manuskripts.

Weiterer Dank gebührt Diplomanden und studentischen Mitarbeitern, die viele der Implementierungsarbeiten und Messungen durchgeführt haben, hier vor allem Harry Köhler und Hans-Joachim Walter für die gute und fruchtbare Zusammenarbeit während der Erstellung des Simulators und der Auswertung von Simulationsläufen.

Danken möchte ich zudem für die technische und organisatorische Unterstützung während der Fertigstellung meiner Arbeit: T. Gauweiler, R. Reske und U. Sellentin für die Betreuung und Wartung der Rechnersysteme sowie P. Berners und B. Büssow für das Editieren des Manuskripts.

Frankfurt, im Januar 1990

D. Wybranietz

# Inhaltsverzeichnis

<b>1. Einleitung</b> .....	1
<b>2. Anwendungsbeispiele für Multicasts</b> .....	5
<b>3. Problematik der Multicast-Kommunikation</b> .....	13
3.1 Begriffsfestlegungen .....	14
3.2 Kommunikationsmodelle.....	17
3.3 Zuverlässigkeitsaspekte .....	24
3.4 Festlegung und Verwaltung von Multicast-Gruppen.....	28
<b>4. Multicast-Protokolle</b> .....	34
4.1 Basistechniken für Multicast-Protokolle.....	34
4.1.1 Simulation von Multicast-Nachrichten .....	39
4.1.2 Direkte Multicast-Übertragung mit Quittungen.....	43
4.1.3 Verfahren zur Reduktion von Quittungen.....	43
4.1.4 Negative Quittungen .....	45
4.1.5 Kombination verschiedener Techniken.....	46
4.2 Direkte Protokolle mit steigenden Zuverlässigkeitsgraden .....	48
4.2.1 k-zuverlässige Protokolle .....	48
4.2.2 Vermeidung von Pufferüberläufen bei der Paketisierung.....	56
4.3 Atomar-zuverlässige Multicasts.....	63
4.4 Ordnungserhaltende Multicasts .....	64
4.4.1 Direkte Übertragung: das ISIS-System.....	64
4.4.2 Ring-basierte Verfahren .....	65
4.4.2.1 Eine zentral-gesteuerte Protokollfamilie.....	65
4.4.2.2 Dezentrale Multicast-Protokolle .....	72
4.4.3 Quittungen als Multicast-Nachrichten .....	83
4.5 Byzantinische Protokolle .....	85
4.6 Zusammenfassung.....	87

<b>5. Unterstützung von Multicasts durch Sprachen und Betriebssysteme.....</b>	<b>89</b>
5.1 Die verteilte Systemimplementierungssprache LADY.....	90
5.1.1 Sprachkonzepte von LADY .....	91
5.1.2 Die Multicast-Kommunikation in LADY .....	94
5.1.3 Semantik der Multicast-Kommunikation .....	97
5.1.4 Implementierungsaspekte.....	99
5.1.5 Diskussion der Multicast-Semantik in LADY .....	110
5.2 CSSA .....	112
5.2.1 Multicasts in CSSA .....	115
5.3 StarMod .....	117
5.4 Eine Multicast-Erweiterung für LINDA .....	119
5.5 Multicast-Kommunikation in MEGLOS .....	122
5.6 Eine Multicast-Schnittstelle für UNIX .....	124
5.7 Der V-Kernel und das VMTP-Protokoll.....	127
5.8 Das ISIS-System .....	129
<b>6. Hardware-Unterstützung .....</b>	<b>133</b>
6.1 Existierende Hardware-Unterstützung für Multicasts .....	134
6.2 Verbesserungsmöglichkeiten der Hardware-Unterstützung .....	136
<b>7. Vergleichende Leistungsbewertung ausgewählter Multicast-Protokolle.....</b>	<b>140</b>
7.1 Messungen im INCAS-Experimentalsystem .....	140
7.1.1 Die Meßumgebung.....	141
7.1.2 Meßergebnisse.....	144
7.2 Das Simulationssystem INCSIM .....	148
7.3 Simulationsergebnisse .....	154
7.3.1 Lasterzeugung .....	154
7.3.2 Ergebnisse .....	158
7.3.3 Zusammenfassende Bewertung.....	172
<b>8. Schlußbemerkungen .....</b>	<b>175</b>
<b>Literatur.....</b>	<b>180</b>