Berichte aus der Informationstechnik

Klaus-Dieter Thies

Elementare Einführung in lokale Netzwerke und verteilte Systeme

Architektur, Mechanismen und Protokolle der Layer Physical (1), Medium Access Control (2a), Logical Link Control (2b) und Transport (4) mit User Datagram Protocol (UDP)

Shaker Verlag
Aachen 2014

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.

Copyright Shaker Verlag 2014
Alle Rechte, auch das des auszugsweisen Nachdruckes, der auszugsweisen oder vollständigen Wiedergabe, der Speicherung in Datenverarbeitungsanlagen und der Übersetzung, vorbehalten.

Printed in Germany.

ISBN 978-3-8440-3005-1
ISSN 1610-9406

Shaker Verlag GmbH • Postfach 101818 • 52018 Aachen
Telefon: 02407 / 95 96 - 0 • Telefax: 02407 / 95 96 - 9
Internet: www.shaker.de • E-Mail: info@shaker.de

Vorwort

Das vorliegende Buch nutzt die Methoden der parallelen Programmierung und realisiert auf Kernel-Ebene Algorithmen und Datenstrukturen, die das Zusammenwirken der vielen Teile eines echtzeitfähigen Kommunikationsnetzwerkes transparent machen.
Das Netzwerk wird als Echtzeit-Multitaskingsystem modelliert, das als „Fundamente“ die typischen Kernfunktionen eines Echtzeitbetriebssystems und die Multitasking-Fähigkeit der Intel-Prozessorarchitektur (Pentium) nutzt.
Die Arbeit zeigt, wie die typischen Layer eines lokalen Computernetzwerks (Phys Layer 1, MAC Layer 2a, LLC Layer 2b und Transport Layer 4 mit UDP) als Tasks abgebildet und in das Multitaskingsystem integriert werden können. Die Tasks führen die Protokolle ihrer Layer als parallele Programme aus, die sich als Kernelerweiterungen präsentieren und dafür sorgen, dass der rege vertikale und horizontale Botschaftenaustausch innerhalb der Rechner und zwischen den Rechnern verklemmungsfrei und sicher abläuft.
In Konsequenz führt die Nutzung dieser Methode zu einer Implementierung, die die vielen Rechner im Kommunikationsnetz wie einen einzigen Parallelrechner mit Echtzeitverhalten aussehen lässt.

Das Buch besteht aus fünf Kapiteln.

Das 1. Kapitel behandelt den MAC Layer 2a (Medium Access Control).

Als Mediumzugangskontrolle wird das priorisierte tokenpassing-Verfahren implementiert und verifiziert. Die wesentlichen Strukturmerkmale des MAC-Layers sind die Receive Machine (RxM), die Transmit Machine (TxM) und als Herzstück die Access Control Machine (ACM). Die Receive Machine korrespondiert mit dem Physical Layer und bekommt von dort sogenannte MAC Protocol Data Units (PDUs). Sie überprüft die Frame Check Sequence (FCS) der individuellen PDUs, stellt Fehlermeldungen des UART fest und gibt die Ergebnisse weiter an den LLC Layer und die Access Control Machine. Die Transmit Machine bekommt Daten Frames von der Access Control Machine, berechnet von jedem die Frame Check Sequence, hängt sie an die Daten Frames an und schickt sie als MAC PDUs an den Physical Layer.
Die Access Control Machine stellt sich als eine endliche Zustandsmaschine dar (finite state machine) und bestimmt, wann ein Frame in das Netz geschickt wird. Zu diesem Zweck kooperiert sie mit den ACMs aller anderen Rechner.

Das 2. Kapitel behandelt den LLC Layer 2b (Logical Link Control).

Sie ist die logische Schnittstelle zwischen dem MAC Layer und dem Transport Layer. Ihre Aufgabe ist es, für eine effiziente und zuverlässige Kommunikation zwischen zwei Rechnern zu sorgen. Das dort implementierte Sicherungsprotokoll realisiert einen sogenannten bestätigten verbindungsunabhängigen Dienst, der für tokenbus-basierte Kommunikationsnetze typisch ist. Es zeichnet sich durch seine Robustheit aus, das mit allen Attacken auf den Bitstrom im Netz fertig wird. Durch die Verteilung von Exemplaren dieses Protokolls auf „maßgeschneiderte“ LLC-Tasks, werden sichere Datenübertragungen parallel zwischen den Anwender-Tasks beliebig vieler Rechnern möglich.

Das 3. Kapitel behandelt das User Datagram Protocol (UDP) im Transport Layer 4.

In tokenbus-basierten Kommunikationsnetzen wird bevorzugt UDP eingesetzt. Die typischen Merkmale für dieses Protokoll sind folgende: Es garantiert, dass Botschaften eines Anwenderprogramms abgeschickt werden, aber es garantiert

nicht, dass diese Botschaften beim Anwenderprogramm des Empfängers auch ankommen. Bei UDP werden als ultimative Ziele nicht Tasks „genannt“, sondern abstrakte Zielpunkte, sogenannte protocol ports. Auch ordnet das Protokoll nicht durcheinandergekommene Botschaften. Hier Ordnung zu schaffen, liegt optional beim Anwender. Um dies zu ermöglichen, sind die erforderlichen Datenstrukturen und Parallel-Algorithmen als Kernelerweiterungen implementiert.

Das 4. Kapitel behandelt die verbindungsorientierte LLC-Kommunikation.

Es analysiert die verbindungslose Kommunikation und zeigt, dass sich der Nutzungsgrad des LANs mit der Anzahl der Rechner dramatisch verringert. In einem solchen Netzwerk sind alle Rechner über „ihre“ LLC-Tasks für immer miteinander in Kontakt. Falls keine Anwender-Daten bereit stehen, tauschen die Rechner ununterbrochen nutzlose Quittungs-PDUs aus. Ein Ausweg aus diesem Dilemma ist ein Netzwerk, das verbindungsorientierte Kommunikation unterstützt. In einem solchen System sind die Rechner lediglich während des Datenaustausch miteinander verbunden. Der Nutzungrad des LANs ist jetzt nicht mehr von der Anzahl der Rechner abhängig und kann den Maximalwert erreichen. Die Funktionen zum Aufbau und Abbau von LLC-Verbindungen werden als Kernel-Erweiterungen entwickelt und stehen zur Unterstützung des User Datagram Protocols zusammen mit dem checksum-Algorithmus als Service Access Points (SAPs) zur Verfügung. Von einem erfolgreichen Verbindungsaufbau/-abbau muss zusätzlich gefordert werden, dass sich die beteiligten Anwender-Tasks synchronisiert verbinden oder trennen. Eine Analyse der Rendezvous-Bedingungen, gefolgt von Lösungen zur Verhinderung von deadlocks, schließt das Kapitel ab.

Das 5. Kapitel enthält eine Einführung in verteilte Systeme.

Die Existenz leistungsstarker Mikroprozessoren auf der einen Seite und lokaler Netzwerke auf der anderen Seite, führt zu der Idee, beide Technologien zu vereinigen und ein Rechner-System zu konstruieren, das aus einer Vielzahl von Prozessoren besteht, die über ein Kommunikationsnetzwerk miteinander verbunden sind. Diese Idee ist in den vorangegangenen Kapiteln bereits konsequent umgesetzt. Ziel eines solchen Systems ist es, dem Benutzer die Illusion zu vermitteln, dass er es mit einem einzigen Rechner zu tun hat und nicht mit einer Ansammlung von mehreren unterschiedlichen Rechnern. Dabei spielt die Task-Kommunikation eine wesentliche Rolle. Ein solches System verhält sich wie ein virtuelles Einprozessorsystem mit dem Makel, dass es dort keinen gemeinsamen Speicher und keine gemeinsamen Betriebsmittel gibt. So müssen ausgesuchte Rechner ihre lokale Einheiten wie Speicher und IO als virtuelle globale Datenstrukturen für den gemeinsamen Zugriff zur Verfügung stellen. Dies macht sie zu kritischen Regionen, die nur unter gegenseitigem Ausschluss betreten werden dürfen. Die Simulation der dafür notwendigen P/V-Operationen in einem zentralen Algorithmus und die Strategie zur Vermeidung von deadlocks bei vielfachem Betriebsmittelbedarf sind die Kernaspekte dieses Kapitels. Damit sich die Kommunikation nicht ausschließlich auf explizite Datenübertragungen beschränkt, greift es die Idee von Nelson und Birrell auf und schließt mit dem Mechanismus des entfernten Prozeduraufrufs (RPC) einschließlich den Input-/Output-RPCs ab.

Klaus-Dieter Thies

Inhalt

Vorwort

Kapitel 1: Die MAC(Medium Access Control)-Teilschicht 1

1.1 Architektur eines bus-basierten Multicomputersystems 1
1.2 Kollisionsfreie Übertragung 6
1.2.1 Tokenbus-Operationen 6
1.3 Die interne Struktur der MAC-Teilschicht 7
1.3.1 Die Receive Machine 8
1.3.1.1 Hardware-basierte Fehlererkennung 16
1.3.1.2 Die Komponenten der MAC PDU 19
1.3.1.2.1 Kommunikation zwischen den Layern 21
1.3.1.3 Cyclic Redundancy Code (CRC) 25
1.3.1.3.1 Die modulo m Rechnung 26
1.3.1.3.2 Die Arithmetik im Galois-Feld GF(2) 27
1.3.1.3.3 Fehlererkennung nach der Polynomcodemethode 28
1.3.1.3.4 Zyklische Eigenschaften 30
1.3.1.3.5 Modulo 2 Division und Hardware 31
1.3.1.3.6 Modulo 2 Division und Software 38
1.3.1.3.7 fcs-Berechnung für 1 Oktet 39
1.3.1.3.8 fcs-Berechnung für n Oktets 42
1.3.1.4 Fehlerauswertung und Botschaften 44
1.3.1.4.1 Botschaften an die Interface Machine 47
1.3.1.4.2 Botschaften an die Access Control Machine 59
1.3.2 Die Access Control Machine 67
1.3.2.1 Access Control Machine (ACM)-Zustände und Transitionen 69
1.3.2.1.1 Der Zustand Idle 74
1.3.2.1.2 Botschaften von der Interface Machine 88
1.3.2.1.3 Der Zustand Use_token 98
1.3.2.1.4 Der Zustand Pass_token 109
1.3.2.1.5 Der Zustand Check_pass_token 110
1.3.2.1.6 Der Zustand Await_ifm_response 125
1.3.2.1.7 Der Zustand Claim_token 128
1.3.3 Die Transmit Machine (TxM) 143

Kapitel 2: Die LLC (Logical Link Control)-Teilschicht 146

2.1 Allgemeines 146
2.1.1 Überblick und Interaktionen 147
2.2 Sicherungsprotokoll 1 150
2.2.1 Ablauf ohne Störung 157
2.2.2 Ablauf mit Störung: LLC PDU kommt nicht an 158
2.2.3 Ablauf mit Störung: Quittungs-PDU kommt nicht an 159
2.2.4 Ablauf mit Störung: LLC PDU und Quittungs-PDU kommen nicht an 160
2.3 Sicherungsprotokoll 2 161
2.4 Sicherungsprotokolle für n Rechner 166
2.4.1 Schreiben in die Interface Machine-Sendeeinheit 177
2.4.2 Lesen von der Interface Machine-Empfangseinheit 183
2.5 Die LLCTasks 190

Kapitel 3: Der Transport Layer und das User Datagram Protocol (UDP) 197

3.1 Einführung 197
3.2 Identifizierung des ultimativen Ziels 197
3.3 Das User Data Protocol 197
3.4 Die UDP-Sendeeinheit 199
3.4.1 Lesen der UDP-Sendeeinheit 203
3.4.2 Schreiben in die UDP-Sendeeinheit 207
3.5 Die UDP Ports (UDP-Empfangseinheit) 213
3.5.1 Demultiplexen 216
3.5.2 Schreiben in die UDP Ports 217
3.5.3 Lesen der UDP Ports 223
3.6 Anwendungsbeispiel 231

Kapitel 4: Verbesserung der LAN-Effizienz durch verbindungsorientierte LLC-Kommunikation 237

4.1 Analyse der verbindungslosen LLC-Kommunikation 237
4.2 Verbindungsorientierte LLC-Kommunikation 238
4.2.2 Aufbau einer LLC-Verbindung 249
4.2.3 Abbau einer LLC-Verbindung 274
4.2.4 Modifizierte Sicherungsprotokolle 292
4.2.5 Programmierbeispiel: Verbindungsaufbau, Verbindungsabbau und checksum-Algorithmus 301
4.2.6 Deadlock-Erkennung und –Verhinderung 308
4.2.6.1 Die Sicht der Anwender-Tasks 309
4.2.6.2 Die Sicht der LLC-Tasks 319

Kapitel 5: Einführung in verteilte System 326

5.1 Kommunikation in verteilten Systemen 326
5.2 Gegenseitiger Ausschluss in verteilten Systemen 334
5.3 Implementierung der P/V-Simulationen 346
5.3.1 Die Betriebssystem-Funktion coordinate(.) 348
5.4 Nutzung der P/V-Simulationen im zentralen Algorithmus 354
5.5 Deadlocks in verteilten Systemen 360
5.5.1 Gemeinsame Nutzung mehrerer Betriebsmittel 360
5.5.2 Echte und scheinbare Deadlocks in verteilten Systemen 363
5.5.3 Verteilte Deadlock-Vermeidung 369
5.5.3.1 Geordnete Betriebsmittelstrategie 369
5.6 Entfernter Prozeduraufruf (Remote Procedure Call; RPC) 383
5.6.1 Die HLL (High Level Language)-Methode der Parameterübergabe 383
5.6.2 Rückgewinnung der Parameter vom Stack 391
5.6.3 Der Mechanismus des entfernten Prozeduraufrufs 398
5.6.3.1 Der Output-RPC 406
5.6.3.2 Der Input-RPC 413

Anhang 420
Quickreferenz der zugänglichen Betriebssystem-Funktionen 420
Task-Zustände 422
Assoziierte Literatur 423