

Föderierte Datenbanksysteme

Springer-Verlag Berlin Heidelberg GmbH

Stefan Conrad

Föderierte Datenbanksysteme

Konzepte der Datenintegration

Mit 83 Abbildungen



Springer

Dr. habil. Stefan Conrad
Otto-von-Guericke-Universität Magdeburg
Institut für Technische Informationssysteme
Postfach 4120, D-39016 Magdeburg

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Conrad, Stefan:

Föderierte Datenbanksysteme: Konzepte der Datenintegration/Stefan Conrad.

Berlin; Heidelberg; New York; Barcelona; Budapest; Hongkong; London; Mailand; Paris;
Santa Clara; Singapur; Tokio: Springer, 1997

DBN: 95.093754.1

SG: 28

ISBN 978-3-540-63176-7

ISBN 978-3-642-59028-3 (eBook)

DOI 10.1007/978-3-642-59028-3

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 jeweils geltenden Fassung zulässig. Sie ist grundsätzlich vergütungspflichtig. Zuwiderhandlungen unterliegen den Strafbestimmungen des Urheberrechtsgesetzes.

© Springer-Verlag Berlin Heidelberg 1997

Ursprünglich erschienen bei Springer-Verlag Berlin Heidelberg New York 1997

Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Werk berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, daß solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften.

Umschlaggestaltung: Künkel + Lopka, Heidelberg

Satz: Reproduktionsfertige Autorenvorlage

Gedruckt auf säurefreiem Papier SPIN: 10632558 45/3142 - 5 4 3 2 1 0

Vorwort

Die Integration bestehender und bislang unabhängig voneinander verwalteter Datenbestände wird zu einem immer größeren wirtschaftlichen Problem, dessen Lösung wesentlich zum Erfolg von Unternehmen beitragen kann. Föderierte Datenbanksysteme bieten spezielle Konzepte für diese Integration heterogener Datenbanken. Neben der Schaffung eines einheitlichen und möglichst transparenten Zugriffs auf heterogene Datenbestände, der auch neue, systemübergreifende Anwendungen erlaubt, steht vor allem die Erhaltung der Autonomie der bestehenden Systeme im Vordergrund. Diese gewährleistet, daß existierende Anwendungsprogramme uneingeschränkt weiterverwendet werden können.

Im Rahmen dieser Arbeit stellen wir mehrere Anwendungsgebiete vor, in denen der Einsatz föderierter Datenbanksysteme sinnvoll und gewinnbringend ist. Verschiedene in der Literatur vorkommende Begriffsdefinitionen für föderierte Datenbanksysteme und Multidatenbanksysteme werden zueinander in Bezug gesetzt. Drei Architekturen, die praktisch allen weiteren Arbeiten in diesem Bereich zugrunde liegen und daher als Referenzarchitekturen bezeichnet werden können, werden ausführlich erläutert und verglichen.

Besonderes Gewicht wird den Gebieten der Schemaintegration, der semantischen Integrität und der Transaktionsverwaltung eingeräumt. Weiterhin behandeln wir aber auch Multidatenbanksprachen, die Anfragebearbeitung und -optimierung, Sicherheitsaspekte in föderierten Datenbanksystemen sowie Kopplungsprinzipien für Datenbanksysteme, die als technische Basis für die Realisierung föderierter Systeme einsetzbar sind. Abschließend werden ausgewählte aktuelle Systementwicklungen betrachtet.

Neben der reinen Darstellung der Probleme und der Diskussion verschiedener Lösungsansätze vermitteln wir auch eigene Anregungen und Ideen. Insbesondere Aspekte wie die durchgängige Berücksichtigung von Integritätsbedingungen bei der Schemaintegration, die systematische Analyse der Reaktionsmöglichkeiten auf Verletzungen globaler Integritätsbedingungen sowie die Analyse der Probleme, die bei der Integration konkreter Instanzen bzw. Datenbankausprägungen auftreten, wurden bislang nicht oder nur unvollständig untersucht.

Die Gesamtdarstellung des Forschungs- und Entwicklungsbereiches föderierter Datenbanksysteme wird abgerundet durch umfangreiche Literaturan-

gaben und Hinweise auf zukünftige Entwicklungen und Einsatzmöglichkeiten föderierter Systeme.

Weitere Informationen

Im Laufe der nächsten Zeit sollen weitere Informationen als Ergänzung zu diesem Buch über das WWW verfügbar gemacht werden. Unter der URL

<http://wwiti.cs.uni-magdeburg.de/~conrad/FDBS/Buch.html>

sind zunächst Folien als PostScript-Dateien zugänglich, die als Grundlage für eine Vorlesung „Föderierte Datenbanken“ verwendet werden können. Weitere Vorlesungsunterlagen und sonstige Begleitmaterialien werden Schritt für Schritt hinzugefügt.

Danksagung

Jede wissenschaftliche Arbeit wird von dem Umfeld beeinflusst, in dem sie entsteht. In diesem Fall bestand das „Umfeld“ vor allem aus den Kollegen des Institutes für Technische Informationssysteme sowie der Infrastruktur, die mir während der Anfertigung der Arbeit zur Verfügung stand. Da gerade für eine Arbeit wie diese ein möglichst umfassender Zugriff auf die wissenschaftliche Literatur notwendig ist, möchte ich hier besonders die Bibliothek des Internationalen Begegnungs- und Forschungszentrums für Informatik (IBFI) in Schloß Dagstuhl hervorheben. Dort habe ich einen wesentlichen Teil der relevanten Literatur gefunden.

Viel zu verdanken habe ich Prof. Dr. Gunter Saake. Er hat mir die notwendigen wissenschaftlichen Freiheiten gegeben und mich auch sonst in jeder denkbaren Weise gefördert. Darüber hinaus gilt mein Dank meinen Kollegen Andreas Christiansen, Eyk Hildebrandt, Michael Höding, Ingo Schmitt, Kerstin Schwarz und Can Türker für Anregungen, Diskussionen und das Korrekturlesen von Vorversionen einzelner Abschnitte. Nicht vergessen möchte ich meine ehemaligen Kollegen Dr. Thorsten Hartmann, Dr. Jan-Hardy Kusch und Sven Janssen, die ebenfalls in der Arbeitsgruppe von Prof. Saake mit mir zusammengearbeitet haben.

Ferner möchte ich auch all jenen danken, mit denen ich über einzelne Fragen diskutiert habe oder die mir wertvolle Anregungen für meine eigene Arbeit gegeben haben. Erwähnt seien hier insbesondere Prof. Dr. Andreas Heuer (Universität Rostock), Prof. Dr. Michael Schrefl (Universität Linz) und Dr. Myra Spiliopoulou (Humboldt-Universität Berlin) sowie alle Teilnehmer des Workshops über föderierte Datenbanken, der am 22./23. April 1996 in Magdeburg stattgefunden hat.

Ganz besonderer Dank gilt meiner Frau Sibylle. Sie hat meine monatelange Arbeit an diesem Buch in bewundernswerter Weise „ertragen“ und mich dabei in jeder Hinsicht unterstützt, so daß das Buch zügig entstehen konnte.

Magdeburg, im August 1997

Stefan Conrad

Inhaltsverzeichnis

1. Einleitung	1
1.1 Entwicklung des Gebietes	1
1.2 Was ist eine Föderation?	4
1.3 Föderierte Datenbanksysteme: eine grobe Skizze	6
1.4 Struktur dieses Buches	8
2. Einsatzgebiete und Anwendungsszenarien	13
2.1 Integration bereits existierender Datenbestände	14
2.2 Fabrikplanung: Werkzeugintegration und kooperierendes Arbeiten	16
2.3 Migration	21
2.4 Krankenhausinformationssysteme	24
2.5 Zusammenfassung und Literaturhinweise	28
3. Föderierte Datenbanksysteme: Begriffe und Architekturen	31
3.1 Begriffsbestimmung und Einordnung	32
3.1.1 Zentrale Datenbanksysteme	34
3.1.2 Verteilte Datenbanksysteme	38
3.1.3 Multidatenbanksysteme	40
3.2 Verteilung, Heterogenität und Autonomie	44
3.3 Architekturen für föderierte Datenbanksysteme	50
3.3.1 Import-/Export-Schema-Architektur	51
3.3.2 Die Multidatenbank-Architektur	54
3.3.3 5-Ebenen-Schema-Architektur	57
3.3.4 Vergleich der Architekturen	62
3.4 Zusammenfassung und Literaturhinweise	68
4. Föderierter Datenbankentwurf	71
4.1 Allgemeine Problemstellung	73
4.1.1 Integrationsstrategien	74
4.1.2 Integrationsprozeß	77
4.1.3 Kriterien für Integrationsmethoden	78
4.2 Klassifizierung von Integrationskonflikten	79

4.3	Ausgewählte Lösungsansätze	84
4.3.1	Zusicherungsbasierte Integration	85
4.3.2	Integration von Klassenhierarchien: <i>Upward Inheritance</i>	99
4.3.3	Formalisierte objektorientierte Integration	105
4.3.4	Generisches Integrationsmodell (GIM)	116
4.3.5	Vergleich der Integrationsmethoden	131
4.4	Zusammenfassung und Literaturhinweise	140
5.	Semantische Integrität	143
5.1	Schemaintegration und Integritätsbedingungen	147
5.1.1	Bildung globaler Integritätsbedingungen am Beispiel . .	150
5.1.2	Formale Integrationsregeln für Integritätsbedingungen	155
5.2	Mögliche Reaktionen auf Konsistenzverletzungen	158
5.2.1	Einfügen oder Löschen von Objekten	159
5.2.2	Updates	166
5.3	Konsistenzsicherungsmechanismen	169
5.3.1	Protokollbasierte Ansätze	170
5.3.2	Erweiterte Architektur für Integritätsüberwachung und -sicherung	176
5.4	Zusammenfassung und Literaturhinweise	182
6.	Transaktionen in föderierten Datenbanksystemen	185
6.1	Der klassische Transaktionsbegriff	186
6.2	Allgemeine Grundprinzipien erweiterter Transaktionsmodelle	191
6.2.1	Geschachtelte Transaktionen	191
6.2.2	Unterstützung geschachtelter Transaktionen in föderierten Systemen	194
6.3	Probleme mit autonomen Komponentensystemen	197
6.3.1	Globale Serialisierbarkeit	197
6.3.2	Globale Atomarität	200
6.3.3	Globale Deadlock-Erkennung und -Vermeidung	203
6.4	Spezielle Transaktionsmodelle	205
6.4.1	Multi-Level-Transaktionen	206
6.4.2	Das ConTract-Modell	213
6.4.3	X/OpenDTP	221
6.5	Zusammenfassung und Literaturhinweise	223
7.	Weitere Aspekte	227
7.1	Multidatenbanksprachen	228
7.1.1	MSQL	228
7.1.2	SchemaSQL	231
7.1.3	Diskussion	236
7.2	Anfragebearbeitung und -optimierung	238
7.2.1	Anfragebearbeitung	238

7.2.2	Anfrageoptimierung	242
7.3	Integration von Instanzen	245
7.4	Sicherheit in föderierten Systemen	251
7.5	Gateways und andere Kopplungsprinzipien	254
7.5.1	Gateways	255
7.5.2	ODBC — Open Database Connectivity	258
7.5.3	RDA — Remote Database Access	260
7.5.4	CORBA — Common Object Request Broker Architecture	261
7.6	Zusammenfassung und Literaturhinweise	262
8.	Aktuelle Systementwicklungen	265
8.1	IRO-DB	265
8.2	OpenDM/Efendi	271
8.3	PEGASUS	275
8.4	MARIPOSA	277
8.5	Weitere Projekte	280
8.6	Zusammenfassung und Literaturhinweise	283
9.	Zusammenfassung und Ausblick	287
	Literaturverzeichnis	293
	Abbildungsverzeichnis	319
	Index	323