

TEUBNER-TEXTE zur Informatik Band 33

U. Störl

Backup und Recovery in Datenbanksystemen

TEUBNER-TEXTE zur Informatik

Herausgegeben von

Prof. Dr. Johannes Buchmann, Darmstadt

Prof. Dr. Udo Lipeck, Hannover

Prof. Dr. Franz J. Rammig, Paderborn

Prof. Dr. Gerd Wechsung, Jena

Als relativ junge Wissenschaft lebt die Informatik ganz wesentlich von aktuellen Beiträgen. Viele Ideen und Konzepte werden in Originalarbeiten, Vorlesungsskripten und Konferenzberichten behandelt und sind damit nur einem eingeschränkten Leserkreis zugänglich. Lehrbücher stehen zwar zur Verfügung, können aber wegen der schnellen Entwicklung der Wissenschaft oft nicht den neuesten Stand wiedergeben.

Die Reihe „TEUBNER-TEXTE zur Informatik“ soll ein Forum für Einzel- und Sammelbeiträge zu aktuellen Themen aus dem gesamten Bereich der Informatik sein. Gedacht ist dabei insbesondere an herausragende Dissertationen und Habilitationsschriften, spezielle Vorlesungsskripten sowie wissenschaftlich aufbereitete Abschlußberichte bedeutender Forschungsprojekte. Auf eine verständliche Darstellung der theoretischen Fundierung und der Perspektiven für Anwendungen wird besonderer Wert gelegt. Das Programm der Reihe reicht von klassischen Themen aus neuen Blickwinkeln bis hin zur Beschreibung neuartiger, noch nicht etablierter Verfahrensansätze. Dabei werden bewußt eine gewisse Vorläufigkeit und Unvollständigkeit der Stoffauswahl und Darstellung in Kauf genommen, weil so die Lebendigkeit und Originalität von Vorlesungen und Forschungsseminaren beibehalten und weitergehende Studien angeregt und erleichtert werden können.

TEUBNER-TEXTE erscheinen in deutscher oder englischer Sprache.

Backup und Recovery in Datenbanksystemen

Verfahren, Klassifikation, Implementierung
und Bewertung

Von Dr. Uta Störl
Dresdner Bank AG, Frankfurt am Main



Springer Fachmedien Wiesbaden GmbH



Dr. Uta Störl

Geboren 1970 in Jena. Von 1989 bis 1994 Studium der Mathematik an der Friedrich-Schiller-Universität Jena; Diplomarbeit am Wissenschaftlichen Zentrum der IBM in Heidelberg. Anschließend von 1995 bis 1999 wissenschaftliche Mitarbeiterin am Lehrstuhl für Datenbanken und Informationssysteme der Universität Jena.

Forschungsarbeiten im Bereich Sicherungs- und Wiederherstellungstechniken für Datenbanken; Promotion in diesem Themenbereich 1999. Seit November 1999 Research Professional bei der Dresdner Bank AG, Frankfurt am Main. Arbeitsschwerpunkte: Analyse und Bewertung neuer Technologien

Die Deutsche Bibliothek – CIP-Einheitsaufnahme

1. Auflage April 2001

Alle Rechte vorbehalten

© Springer Fachmedien Wiesbaden 2001

Ursprünglich erschienen bei B. G. Teubner GmbH, Stuttgart/Leipzig/Wiesbaden, 2001

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlages unzulässig und strafbar. Das gilt besonders für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

www.teubner.de

Umschlaggestaltung: Peter Pfitz, Stuttgart

ISBN 978-3-519-00416-5 ISBN 978-3-663-09332-9 (eBook)
DOI 10.1007/978-3-663-09332-9

Meinen Eltern

Vorwort

Verfügbarkeit von Daten und Schutz vor Datenverlust sind Hauptanforderungen von Datenbankanwendern. Große betriebliche integrierte Informationssysteme und neue Datenbankanwendungsgebiete haben in den letzten Jahren dazu geführt, daß die Menge der in Datenbanksystemen verwalteten Daten rapide zunimmt. Durch die wachsende Datenmenge und die gleichzeitig steigenden Verfügbarkeitsanforderungen treten bei den heute verwendeten Sicherungs- und Wiederherstellungsverfahren oftmals erhebliche Performance-Probleme auf. Daraus ergibt sich die Notwendigkeit effizienter Techniken für die Datensicherung und -wiederherstellung, deren Darstellung, Klassifikation, Vergleich und Weiterentwicklung Inhalt des vorliegenden Buches ist.

Es werden Klassifikationskriterien für Sicherungs- und Wiederherstellungstechniken von Datenbanken eingeführt. Existierende Ansätze für die Behandlung von Externspeicherfehlern in wissenschaftlichen Arbeiten werden präsentiert und eine Einordnung und Abgrenzung zu Arbeiten aus dem Bereich der Transaktions- und Systemfehlerbehandlung vorgenommen. Darüber hinaus wird ein ausführlicher Überblick über den Stand in kommerziellen Datenbanksystemen gegeben.

Neben der Diskussion der Funktionalität von Sicherungs- und Wiederherstellungstechniken bilden quantitative Leistungsbetrachtungen einen Schwerpunkt des Buches. Um quantitative Vergleiche, Bewertungen und Performance-Abschätzungen vornehmen zu können, werden analytische Kostenmodelle für verschiedene Backup- und Recovery-Techniken vorgestellt. Zur Validierung der Modelle wurden verschiedene Verfahren in einem DBMS-Prototyp implementiert und entsprechende Leistungsuntersuchungen durchgeführt. Um die erhaltenen Ergebnisse mit denen in real existierenden Datenbanksystemen vergleichen zu können, wurde ein datenbanksystemunabhängiger Backup- und Recovery-Benchmark entworfen und realisiert.

Ausgehend von den bei der Modellierung und den Messungen erzielten Ergebnissen und dabei aufgedeckten Schwachstellen existierender Ansätze, werden Verbesserungsmöglichkeiten und neue Ansätze für die Behandlung von Externspeicherfehlern vorgeschlagen. Zur Steigerung der Performance des Anwendens von Log-Information während der Recovery wurde ein Log-Clustering-Verfahren zur Nachbehandlung von Log-Information entwickelt. Idee dieses Verfahrens ist eine Neusortierung und physische Neuordnung der Log-Einträge dergestalt, daß beim Anwenden der Log-Einträge die Lokalität der Änderungsoperationen erhöht wird. Dadurch wird die Buffer Hit Ratio während der Wiederherstellung signifikant verbessert und damit die Wiederherstellungszeit verringert.

Für verschiedene Backup- und Restore-Techniken werden Klassifikationen, Implementierungsvarianten und Leistungsuntersuchungen vorgestellt. Ein Schwerpunkt sind dabei Ver-

fahren zur inkrementellen Sicherung von Datenbanken. Durch analytische Modelle und Meßergebnisse wird verdeutlicht, welche entscheidende Rolle die effiziente Implementierung des Lesens der veränderten Seiten für die Leistungsfähigkeit der inkrementellen Sicherungsverfahren darstellt. Es wird ein neues Verfahren vorgestellt, mit dem signifikante Performance-Verbesserungen erreicht werden können. Für Verfahren zum parallelen Backup und Restore wird eine Klassifikation gegeben, und die Eigenschaften der verschiedenen Verfahren werden erläutert. Ihre Leistungsfähigkeit wird anhand analytischer Modelle und prototypischer Implementierungen diskutiert.

Das vorliegende Buch basiert im wesentlichen auf der während meiner Tätigkeit als wissenschaftliche Mitarbeiterin am Lehrstuhl für Datenbanken und Informationssysteme in Jena entstandenen Dissertation [Stö99a, Stö00]. Meinem Doktorvater Herrn Prof. Dr. Klaus Küspert möchte ich für die Anregung zur Beschäftigung mit dem Thema Backup und Recovery in Datenbanksystemen und vor allem für die permanente Unterstützung, die kritischen Diskussionen und die hilfreichen Kommentare während der gesamten Zeit danken.

Den Herren Prof. Dr. Albrecht Blaser, Universität Heidelberg, und Prof. Dr. Peter Dadam, Universität Ulm, danke ich für die Bereitschaft zur Übernahme der weiteren Dissertationsgutachten. Bei Herrn Prof. Blaser möchte ich mich darüber hinaus für seine interessanten und engagierten Vorlesungen, mit denen er mich für die Welt der Datenbanken begeistert hat, und seine unterstützende Begleitung in den vergangenen Jahren bedanken.

Dank gebührt auch den Studenten, die durch ihre Studien- und Diplomarbeiten bei der Produktevaluierung bzw. der Implementierung des DBMS-Prototyps einen großen Beitrag zum Gelingen der Arbeit geleistet haben. Hier seien besonders Robert Baumgarten, Christoph Gollmick, Gert Großmann, Marco Maxeiner und Arndt Schweinitz genannt.

Den Firmen IBM, Software AG und SQL Datenbanksysteme danke ich für die Bereitstellung ihrer Datenbankprodukte. Bei Herrn Dr. Harald Schöning von der Software AG Darmstadt möchte ich mich für die vielen interessanten Diskussionen in den vergangenen Jahren bedanken.

Meinem ehemaligen Kollegen Dr. Ralf Schaarschmidt danke ich für das sorgfältige und kritische Korrekturlesen vieler Veröffentlichungen und der Dissertation sowie für die interessanten Diskussionen und Anregungen. Für die große technische und persönliche Unterstützung in den vergangenen Jahren möchte ich mich bei Jutta Sieron bedanken. Jan Nowitzky und Jens Lufter danke ich für ihren Beitrag zur Schaffung eines angenehmen Arbeitsumfeldes am Lehrstuhl.

Dem Verlag B. G. Teubner und Herrn Jürgen Weiß in Leipzig sowie dem Mitherausgeber Prof. Dr. Udo Lipeck, Universität Hannover, danke ich für die Möglichkeit, dieses Buch im Teubner-Verlag zu veröffentlichen. Bei der Firma IBM, insbesondere bei den Herren Manfred Päckler, Düsseldorf, und Dr. Hans-Joachim Renger, Böblingen, möchte ich mich für die Unterstützung der Publikation bedanken.

Besonderer Dank gebührt meinem Mann und meinen Freunden für die motivierende Unterstützung in den vergangenen Jahren und das Verständnis für den leider oft notwendigen Verzicht auf gemeinsame Zeit.

Jena und Frankfurt am Main, im Januar 2001

Uta Störl
Uta.Stoerl@gmx.de

Inhaltsverzeichnis

1	Einleitung	15
1.1	Motivation	15
1.2	Ziele des Buches	16
1.3	Abgrenzung zu anderen Arbeiten	17
1.4	Aufbau des Buches	18
2	Grundlagen	21
2.1	Architektur eines Datenbanksystems	21
2.2	Fehlerbehandlung in Datenbanksystemen	22
2.3	Backup	24
2.3.1	Partielles Backup	25
2.3.2	Inkrementelles Backup	26
2.3.3	Online-Backup	27
2.3.4	Paralleles Backup	27
2.4	Recovery	28
2.4.1	Partielle Recovery	28
2.4.2	Point-In-Time-Recovery	29
2.4.3	Online-Recovery	29
2.4.4	Paralleles Restore	30
2.5	Stand der Forschung	30
2.6	Stand der Technik	33
2.7	Weitere Ansätze zum Schutz vor Datenverlust	36
2.7.1	Spiegelung	37
2.7.2	RAID	37
2.7.3	Remote-Backup	39
2.8	Zusammenspiel mit anderen Sicherungssystemen	41

2.8.1	Betriebssystemdienste	41
2.8.2	Speichermanagementsysteme	43
2.8.3	Koordination der Sicherungsstrategien	43
3	Konzepte und Werkzeuge für quantitative Untersuchungen	45
3.1	Überblick	45
3.2	DBMS-Prototyp	47
3.2.1	Architektur	47
3.2.2	Buffer Manager	49
3.2.3	Log Manager	50
3.2.4	Transaction Manager	51
3.2.5	Backup Manager	51
3.2.6	Restore Manager	52
3.3	Analyse und Konvertierung von DB2-Logs	52
3.3.1	Architektur	52
3.3.2	Lesen von Log-Einträgen	52
3.3.3	Statistische Auswertungen	53
3.3.4	Konvertierung	56
3.4	Backup- und Recovery-Benchmark	56
3.4.1	Architektur	57
3.4.2	Funktionalität	58
3.4.3	Schnittstelle	59
3.5	Untersuchungsszenarien	63
3.6	Meßumgebung	64
3.6.1	Speicherverwaltung	65
3.6.2	Verwendete Meßumgebung	66
3.6.3	Verwaltung und Auswertung der Meßdaten	67
4	Logging und Reapply	69
4.1	Log-Protokollierung in Datenbanksystemen	69
4.1.1	Grundlagen	69
4.1.2	Archivierung von Log-Information	73
4.1.3	Physisches Logging	75
4.1.4	Logisches Logging	76
4.1.5	Physiological Logging	77

4.1.6	Logging im DBMS-Prototyp	78
4.1.6.1	Physisches Logging	79
4.1.6.2	Physiological Logging	80
4.1.7	Logging in DB2	81
4.1.7.1	Log-Protokollierungstechnik	81
4.1.7.2	Archivierung von Log-Dateien	82
4.2	Reapply-Algorithmen	83
4.2.1	Reapply-Algorithmen mit Analysephase	83
4.2.2	Reapply-Algorithmen mit verkürzter Analysephase	85
4.2.3	Reapply-Algorithmen ohne Analysephase	86
4.3	Recovery-Verfahren	88
4.3.1	Point-In-Time-Recovery	88
4.3.2	Partielle Recovery	91
4.3.3	Online-Recovery	93
4.4	Leistungsuntersuchungen	94
4.4.1	Analytische Modelle	94
4.4.1.1	Größe des Log	95
4.4.1.2	Reapply	97
4.4.1.3	Beispiele	101
4.4.2	Messungen	104
4.5	Log-Clustering-Verfahren LogSplit	105
4.5.1	LogSplit-Algorithmus	106
4.5.2	Reapply	109
4.5.3	Leistungsuntersuchungen	113
4.5.3.1	Größe des Log	113
4.5.3.2	Reapply	114
4.5.4	Anwendung von LogSplit für verschiedene Recovery-Verfahren	116
4.5.4.1	Partielle Recovery	116
4.5.4.2	Online-Recovery	117
4.5.4.3	Paralleles Reapply	118

5 Backup und Restore	121
5.1 Komplett-Backup und -Restore	121
5.1.1 Implementierung	121
5.1.2 Leistungsuntersuchungen	122
5.1.2.1 Analytische Modelle	122
5.1.2.2 Messungen	124
5.2 Online-Backup	128
5.2.1 Klassifikation	128
5.2.2 Implementierung	131
5.2.2.1 Eingeschränktes Online-Backup	131
5.2.2.2 Vollständiges Online-Backup	133
5.2.3 Bewertung	135
5.3 Inkrementelles Backup	135
5.3.1 Grundlagen	136
5.3.1.1 Einfaches inkrementelles Backup	136
5.3.1.2 Inkrementelles Multilevel-Backup	137
5.3.2 Klassifikation	139
5.3.2.1 Klassifikationskriterien	139
5.3.2.2 Einfaches inkrementelles Backup	139
5.3.2.3 Inkrementelles Multilevel-Backup	142
5.3.3 Implementierung	143
5.3.3.1 FullRead	144
5.3.3.2 SelectiveRead	146
5.3.4 Leistungsuntersuchungen	146
5.3.4.1 Analytische Modelle	147
5.3.4.2 Messungen	149
5.3.5 SelectiveRead _{Gap}	152
5.3.5.1 Algorithmus	152
5.3.5.2 Leistungsuntersuchungen	155
5.4 Paralleles Backup und Restore	157
5.4.1 Grundlagen	157
5.4.1.1 I/O-Parallelität	158
5.4.1.2 Paralleles Backup und Restore	158
5.4.1.3 Physische Speichereinheiten der Datenbank	159

- 5.4.1.4 Datenverteilung beim Backup 160
- 5.4.2 Klassifikation 163
 - 5.4.2.1 Klassifikationskriterien 163
 - 5.4.2.2 Bewertungskriterien 166
- 5.4.3 Implementierung 170
 - 5.4.3.1 Dynamische Verfahren 170
 - 5.4.3.2 Statische Verfahren 172
- 5.4.4 Leistungsuntersuchungen 176
 - 5.4.4.1 Analytische Modelle 176
 - 5.4.4.2 Messungen 180
- 6 Zusammenfassung und Ausblick 183**
 - 6.1 Ergebnisse 183
 - 6.2 Weiterführende Arbeiten 185
- Anhang 191**
- Literaturverzeichnis 195**
- Stichwortverzeichnis 203**