



Offenlegungsschrift

(51) Int Cl.: **G06F 11/14** (2006.01)

G06F 12/16 (2006.01)

G06F 17/30 (2006.01)

Stolze, Knut, c/o IBM Corp., 71032 Böblingen, DE; Kunioka-Weis, Laura Michiko, c/o IBM Corp., San Jose, Calif., US; Broll, Bjoern, c/o IBM Corp., 71032 Böblingen, DE; Kempfert, Claus, c/o IBM Corp., 71032 Böblingen, DE; Koeth, Oliver, c/o IBM Corp., 71032 Böblingen, DE

**Richardt Patentanwälte PartG mbB, 65185
Wiesbaden, DE**

The flowchart illustrates a data backup method with the following steps and corresponding data states:

- 211** Erstellen einer ersten Sicherungskopie des Teils der Tabelle (Create a first backup copy of the part of the table). This step corresponds to data state **201**, where a table part (111) is backed up to a cylinder (135) and a tape (127) labeled 203A.
- 213** Zuweisen eines Zugriffszustands-Flag zu dem Teil der Tabelle (Assign an access state flag to the part of the table). This step corresponds to data state **203B**, where the table part (111) is marked with a flag (203B) and stored on the cylinder (135).
- 215** Setzen des Zugriffszustands-Flag (Setting the access state flag). This step corresponds to data state **205**, where the flag (203A) is set on the tape (127).
- Verwenden des Zugriffszustands-Flag zum Verhindern einer Abänderung von bestehenden Sicherungskopien des Teils der Tabelle und zum Verhindern weiterer Sicherungsoperationen an dem Teil der Tabelle** (Using the access state flag to prevent modification of existing backup copies of the part of the table and to prevent further backup operations on the part of the table). This step corresponds to data state **203C**, where the table part (111) is marked with a flag (203C) and stored on the cylinder (135).
- 217** Archivieren des Teils der Tabelle (Archiving the part of the table). This step corresponds to data state **205**, where the table part (111) is archived on the cylinder (135) and the tape (127) labeled 203A.
- 219** Löschen des Teils der Tabelle (Deleting the part of the table). This step corresponds to data state **201**, where the table part (111) is deleted from the cylinder (135) and the tape (127) labeled 203A.

Beschreibung**GEBIET DER ERFINDUNG**

[0001] Die Erfindung betrifft Datenverarbeitungssysteme und insbesondere ein Verfahren für ein sicheres Archivieren von Daten.

Hintergrund

[0002] Die Menge an durch Datenbanksysteme gespeicherten und verarbeiteten Daten wächst mit zunehmender Geschwindigkeit. Das Speichern von Milliarden von Datenbanksätzen hat eine Auswirkung auf Plattenplatz, der zum Speichern dieser Daten benötigt wird. Dieser Plattenplatz kann für qualitativ hochwertigen Speicher wie die z/OS-Speicher-Server teuer sein. Außerdem kann sich das Speichern von Milliarden von Datensätzen in einem Datenbanksystem auf die Leistungsfähigkeit des Datenbanksystems auswirken.

[0003] Deshalb könnten weniger häufig geänderte Daten und/oder weniger häufig genutzte Daten davon betroffen sein, auf preisgünstigere Speicher oder Systeme verlagert („archiviert“) zu werden.

[0004] Der Prozess des Verlagerns von Daten von einem Platz zu einem anderen (z. B. auf ein anderes Datenbanksystem) wird in dieser Erfindung „Archivieren“ genannt.

[0005] Nahezu sämtlichen bekannten Lösungen zum „Archivieren“ ist gemein, dass die Daten in ein Archivsystem verlagert werden und die Daten danach von dem Quellsystem gelöscht werden. Diese Datentrennung führt jedoch zu Datenpflegeproblemen auf dem Quellsystem.

Kurzdarstellung der Erfindung

[0006] Es ist ein Ziel von Ausführungsformen der Erfindung, ein verbessertes Verfahren, ein Computersystem und ein Computerprogrammprodukt zum sicheren Archivieren von Daten bereitzustellen. Das Ziel wird durch den Gegenstand der unabhängigen Ansprüche erreicht. Vorteilhafte Ausführungsformen werden in den abhängigen Ansprüchen beschrieben.

[0007] In einem Aspekt betrifft die Erfindung ein auf einem Computer ausgeführtes Verfahren zum Archivieren eines vorher festgelegten Teils einer sich in einem Datenbanksystem befindlichen Tabelle. Das Verfahren umfasst Folgendes:

- a. Erstellen einer ersten Sicherungskopie des Teils der Tabelle;
- b. Zuweisen eines Zugriffszustand-Flag zu dem Teil der Tabelle;
- c. Setzen des Zugriffszustand-Flag;

d. Verwenden des Zugriffszustand-Flag zum Verhindern einer Abänderung von bestehenden Sicherungskopien des Teils der Tabelle und zum Verhindern weiterer Sicherungsoperationen an dem Teil der Tabelle;

e. Archivieren des Teils der Tabelle auf einem separaten System;

f. Löschen des Inhalts des Teils der Tabelle aus dem Datenbanksystem und Beibehalten des Zugriffszustand-Flag in dem Datenbanksystem.

[0008] Diese Merkmale können ein sicheres Verfahren zum Archivieren von Daten durch Verhindern von Datenverlust bereitstellen, indem geschützte Sicherungskopien verwendet werden, die vor dem Löschen der Daten in dem Quellsystem erstellt wurden. Das heißt, die Daten können in dem Datenbanksystem wiederhergestellt werden, falls das Archivsystem ausfällt oder Datenabänderungen an den archivierten Daten benötigt werden. Dies kann außerdem Ressourcen (z. B. Verarbeitungsleistung) einsparen, die andernfalls beim Wiederherstellen der Daten von dem Archivsystem benötigt würden.

[0009] Die Abänderung des Teils der Tabelle umfasst das Einfügen von neuen Daten in den Teil der Tabelle oder das Aktualisieren (Überschreiben) von bestehenden Daten in dem Teil der Tabelle oder ein Löschen des Inhalts des Teils der Tabelle.

[0010] Das vorliegende Verfahren kann das Erstellen mehrerer Archive unter Verwendung einer gleichen einzelnen Datenquelle ermöglichen (d. h. von der ersten Sicherungskopie). Dies steht im Gegensatz zu den herkömmlichen Verfahren, die mehrere Archive beruhend auf mehreren Datenquellen bereitstellen. Dies liegt daran, dass auf das Archivieren von Daten ein Löschen der Datenquelle folgt und somit eine weitere Archivierungsoperation mit dem herkömmlichen Verfahren möglicherweise zunächst das Wiederherstellen der gelöschten Datenquelle erfordern kann. Folglich spart das vorliegende Verfahren möglicherweise Ressourcen ein und verhindert eine Dateninkonsistenz zwischen mehreren Archiven, die beim Bearbeiten mehrerer Quellen derselben Daten auftreten kann.

[0011] Die bestehenden Sicherungskopien umfassen zum Beispiel mindestens die letzte Sicherungskopie des Teils der Tabelle.

[0012] Gemäß einer Ausführungsform weist das Verfahren ferner das Wiederholen des Archivierungsschritts e) mindestens zwei Mal unter Verwendung der ersten Sicherungskopie für den Teil der Tabelle auf. Hierdurch werden möglicherweise mehrere archivierte Kopien desselben Teils der Tabelle unter Verwendung der ersten Sicherungskopie bereitgestellt.

[0013] Gemäß einer Ausführungsform weist das Verfahren ferner das Empfangen einer Abfrage des Teils der Tabelle auf, die Daten aus dem Teil der Tabelle des Quellsystems zurücksenden sollte (d. h. keine archivierten Daten); wobei das Zugriffszustand-Flag zum Zurücksenden eines leeren Ergebnisses als Reaktion auf die empfangene Abfrage verwendet wird. Die Abfrage kann zum Beispiel eine SQL-(Structured Query Language) oder MDX-(Multidimensional Expressions) Abfrage aufweisen, oder sie kann alternativ jedem beliebigen anderen Format oder jeder beliebigen anderen Sprache entsprechen.

[0014] Da der Teil der Tabelle aus dem Datenbanksystem gestrichen wurde, reicht das bloße Vorhandensein des Zugriffszustand-Flag möglicherweise als Indikator dafür aus, dass keine Daten bestehen und nichts durchsucht werden muss. Somit profitiert möglicherweise die Verarbeitung der Abfrage (z. B. durch weniger benötigte Ressourcen), da die Datei für den Teil des Tabellenraums möglicherweise nicht geöffnet werden muss und keine internen Metadaten (Free Space Map, Indizes, Seiteninformationen usw.) geladen und interpretiert werden müssen.

[0015] Gemäß einer Ausführungsform weist die erste Sicherungskopie eine erste Kennung auf, die unter Verwendung eines ersten Identifizierungsalgorithmus festgelegt wurde. Das Verfahren weist ferner auf: nach dem Erstellungsschritt a) und nach dem Setzschrift c), Ermitteln, ob eine zweite Sicherungskopie des Teils der Tabelle mit einer unter Verwendung eines zweiten Identifizierungsalgorithmus festgelegten zweiten Kennung erstellt wurde, wobei, falls die zweite Sicherungskopie erstellt wurde, eine dritte Sicherungskopie des Teils der Tabelle mit einer dritten Kennung erstellt wird, die unter Verwendung des ersten Identifizierungsalgorithmus festgelegt wird.

[0016] Die dritte Sicherungskopie wird die letzte Sicherungskopie des Teils der Tabelle, und somit gilt das unten beschriebene Verfahren für die erste Sicherungskopie (d. h. als letzte Sicherungskopie) auch für die dritte Sicherungskopie.

[0017] Diese Ausführungsform kann vorteilhaft sein, da sie den Sicherheitsaspekt des vorliegenden Verfahrens stärkt, indem die Verwendung verschiedener Identifizierungsalgorithmen für Sicherungskopien archivierter Daten verhindert wird.

[0018] Gemäß einer Ausführungsform weist das Verfahren ferner das Verwenden des Zugriffszustand-Flag zum Verhindern einer Abänderung des Teils der Tabelle auf. Dies kann vorteilhaft sein, da in den Teil der Tabelle keine neuen Datensätze eingefügt werden können, die in den archivierten Daten nicht widergespiegelt würden.

[0019] Gemäß einer Ausführungsform weist das Verfahren ferner die Schritte des Aufhebens des Zugriffszustand-Flag vor dem Löschschritt f); und das Zurücksetzen des Zugriffszustand-Flag nach dem Löschschritt f) auf.

[0020] Gemäß einer Ausführungsform weist das Verfahren ferner auf: nach dem Löschschritt f):

- Feststellen, dass der Teil der Tabelle durch eine Operation mit Ausnahme des Löschschritts f) nach dem Erstellungsschritt a) nicht abgeändert wurde;
- Feststellen, dass die erste Sicherungskopie die letzte Sicherungskopie des Teils der Tabelle ist, und erfolgreiches Beenden des Verfahrens.

[0021] Es kann zum Beispiel andere Operationen als die Operation des Löschschritts f) der vorliegenden Offenbarung geben, die den Teil der Tabelle möglicherweise abändern. Der Schritt des Ermitteln, ob der Teil der Tabelle nicht abgeändert wurde, kann durchgeführt werden, indem geprüft wird, ob der Teil der Tabelle durch derartige andere Operationen abgeändert wurde.

[0022] Gemäß einer Ausführungsform weist das Verfahren ferner auf: nach dem Zurücksetzen des Zugriffszustand-Flag:

- Feststellen, dass der Teil der Tabelle durch eine Operation mit Ausnahme des Löschschritts f) nach dem Erstellungsschritt a) nicht abgeändert wurde;
- Feststellen, dass es sich bei der ersten Sicherungskopie um die letzte Sicherungskopie des Teils der Tabelle handelt, und erfolgreiches Beenden des Verfahrens.

[0023] Hierdurch wird möglicherweise ein Verfahren zum Prüfen der Konsistenz bereitgestellt, ohne dass die Sicherungsdaten mit den archivierten Daten verglichen werden müssen.

[0024] Gemäß einer Ausführungsform weist das Verfahren ferner auf: nach dem Zurücksetzen des Zugriffszustand-Flag:

- Feststellen, dass der Teil der Tabelle durch eine Operation mit Ausnahme des Löschschritts f) nach dem Erstellungsschritt a) nicht abgeändert wurde
- als Reaktion auf eine Feststellung, dass eine zweite spätere Sicherungskopie des Teils der Tabelle vorhanden ist, Feststellen, dass die erste Sicherungskopie ungültig ist;
- Aufheben des Zugriffszustand-Flag;
- Wiederherstellen des Teils der Tabelle unter Verwendung der ersten Sicherungskopie; und
- erneutes Ausführen der Schritte a) bis f), wodurch eine neue Sicherungskopie erstellt wird.

[0025] Da das Zugriffszustand-Flag zum Schützen der zuletzt erstellten, in dem Datenbanksystem gespeicherten Sicherungskopie verwendet werden kann, kann diese Ausführungsform verhindern, dass andere Sicherungskopien (z. B. die zweite Sicherungskopie unter Verwendung anderer Identifizierungsalgorithmen) als die gemäß der vorliegenden Offenbarung (z. B. unter Verwendung des ersten Identifizierungsalgorithmus) erstellten durch das Zugriffszustand-Flag geschützt werden. Das Schützen lediglich der letzten Sicherungskopie kann vorteilhaft sein, da es das Löschen von alten Sicherungskopien und somit das Einsparen von Speicherressourcen ermöglichen kann.

[0026] Gemäß einer Ausführungsform weist das Verfahren ferner auf: nach dem Zurücksetzen des Zugriffszustand-Flag:

- als Reaktion auf eine Feststellung, dass der Teil der Tabelle durch eine Operation mit Ausnahme des Löschschritts f) nach dem Erstellungsschritt a) abgeändert wurde, Ermitteln einer Abänderung des Teils der Tabelle unter Verwendung von dem Teil der Tabelle zugehörigen statistischen Daten,
- Aufheben des Zugriffszustand-Flag;
- Wiederherstellen des Teils der Tabelle unter Verwendung der ersten Sicherungskopie,
- Anwenden der ermittelten Abänderung des Teils der Tabelle und erneutes Ausführen der Schritte a) bis f) für den abgeänderten Teil der Tabelle.

[0027] Wenn zum Beispiel der Löschschritt f nicht möglich ist, wenn das Zugriffszustand-Flag gesetzt ist. Dann wird das Aufheben dessen benötigt, bevor Schritt f) durchgeführt werden kann. In der Zwischenzeit kann eine Abänderung von Daten geschehen. Dies würde im Rahmen der obigen Ausführungsform erkannt und verarbeitet werden.

[0028] Gemäß einer Ausführungsform weist das Verfahren ferner auf: nach dem Löschschritt f)

- Feststellen, dass der Teil der Tabelle durch eine Operation mit Ausnahme des Löschschritts f) nach dem Erstellungsschritt a) nicht abgeändert wurde
- als Reaktion auf eine Feststellung, dass eine zweite spätere Sicherungskopie des Teils der Tabelle vorhanden ist, Feststellen, dass die erste Sicherungskopie ungültig ist;
- Wiederherstellen des Teils der Tabelle unter Verwendung der ersten Sicherungskopie; und
- erneutes Ausführen der Schritte a) bis f), wodurch eine neue Sicherungskopie erstellt wird.

[0029] Gemäß einer Ausführungsform weist das Verfahren ferner auf: nach dem Löschschritt f)

- als Reaktion auf eine Feststellung, dass der Teil der Tabelle durch eine Operation mit Ausnahme des Löschschritts f) nach dem Erstellungsschritt a) abgeändert wurde, Ermitteln einer Abänderung

des Teils der Tabelle unter Verwendung von dem Teil der Tabelle zugehörigen statistischen Daten,

- Wiederherstellen des Teils der Tabelle unter Verwendung der ersten Sicherungskopie,
- Anwenden der ermittelten Abänderung des Teils der Tabelle und erneutes Ausführen der Schritte a) bis f) für den abgeänderten Teil der Tabelle.

[0030] Gemäß einer Ausführungsform wird das Löschen durch ein Löschmodul des Datenbanksystems durchgeführt. Das Verfahren weist ferner das Steuern des Löschmoduls auf, um den Teil der Tabelle unabhängig davon zu löschen, ob das Zugriffszustand-Flag gesetzt ist.

[0031] Diese Ausführungsform kann nahtlos in die bestehenden Systeme integriert werden, da sie nur wenige oder gar keine technischen Einschränkungen für das Ignorieren des Zugriffszustand-Flag aufweist. Dies liegt daran, dass es keine Notwendigkeit dafür gibt, das Zugriffszustand-Flag aufzuheben, und somit keine Notwendigkeit zum Prüfen, ob der Teil der Tabelle abgeändert oder die erste Sicherungskopie ungültig gemacht wurde.

[0032] Gemäß einer Ausführungsform wird das Erstellen einer ersten Sicherungskopie durch ein Sicherungsmodul des Datenbanksystems durchgeführt, wobei das Verfahren ferner das Steuern des Sicherungsmoduls aufweist, damit es die Schritte a, b und c als atomare Operation beinhaltet. Dies kann vorteilhaft sein, da es Abänderungen des Teils der Tabelle verhindern kann sowie verhindern kann, dass neue Sicherungskopien gemacht werden, was andernfalls möglicherweise stattfinden könnte, bevor das Zugriffszustand-Flag gesetzt wurde.

[0033] Gemäß einer Ausführungsform weist das Verfahren ferner das Empfangen einer Anforderung über das Wiederherstellen des Teils der Tabelle, das Aufheben des Zugriffszustand-Flag, und das Wiederherstellen des Teils der Tabelle unter Verwendung der ersten Sicherungskopie auf.

[0034] Gemäß einer Ausführungsform weist das Verfahren ferner das Empfangen von Daten auf, die auf einen unbefugten Zugriffsversuch auf den Teil der Tabelle hinweisen; wobei das Zugriffszustand-Flag zum Verhindern von Zugriff auf die erste Sicherungskopie und die archivierten Daten verwendet wird. Bei dem unbefugten Zugreifen auf den Teil der Tabelle kann es sich zum Beispiel um ein unberechtigtes Erlangen von Zugriff auf Root-Ebene und den sich daraus ergebenden Diebstahl von Informationen handeln. Das untersagte Zugreifen auf Teile der Tabelle kann nur den Eindringling betreffen, der den unbefugten Zugriffsversuch unternimmt (andere berechnete Benutzer können z. B. auf den Teil der Tabelle zugreifen, während das Zugreifen durch den Eindringling blockiert wird). Das Verfahren kann ferner das

Senden einer Zugriffsfehlermeldung an den Eindringling aufweisen. Diese Ausführungsform kann vorteilhaft sein, da sie den Sicherheitsaspekt des vorliegenden Verfahrens stärkt, indem ein benutzerabhängiges Zugriffszustand-Flag bereitgestellt wird.

[0035] In einem anderen Aspekt betrifft die Erfindung ein Computerprogrammprodukt, das auf einem Computer ausführbare Anweisungen zum Durchführen der Verfahrensschritte des Verfahrens nach einer beliebigen der vorhergehenden Ausführungsformen aufweist.

[0036] In einem anderen Aspekt betrifft die Erfindung ein Datenbanksystem zum Archivieren eines vorher festgelegten Teils einer sich in dem Datenbanksystem befindlichen Tabelle, das Datenbanksystem aufweisend:

- ein Sicherungsmodul, das zu Folgendem in der Lage ist:
Erstellen einer ersten Sicherungskopie des Teils der Tabelle;
Zuweisen eines Zugriffszustand-Flag zu dem Teil der Tabelle;
Setzen des Zugriffszustand-Flag;
Verwenden des Zugriffszustand-Flag zum Verhindern einer Abänderung von bestehenden Sicherungskopien des Teils der Tabelle und zum Verhindern weiterer Sicherungsoperationen an dem Teil der Tabelle;
- ein Archivierungsmodul, das in der Lage ist, den Teil der Tabelle auf einem separaten System zu archivieren;
- ein Löschmodul, das in der Lage ist, den Inhalt des Teils der Tabelle aus dem Datenbanksystem zu löschen und das Zugriffszustand-Flag in dem Datenbanksystem beizubehalten.

[0037] In der Verwendung hierin umfasst ein ‚durch einen Computer lesbares Speichermedium‘ jedes beliebige physische Speichermedium, das Anweisungen speichern kann, die von einem Prozessor einer Datenverarbeitungseinheit ausgeführt werden können. Das durch einen Computer lesbare Medium kann als ein durch einen Computer lesbares nichtflüchtiges Speichermedium bezeichnet werden. Das durch einen Computer lesbare Speichermedium kann auch als ein durch einen Computer lesbares physisches Medium bezeichnet werden. In einigen Ausführungsformen kann ein durch einen Computer lesbares Speichermedium auch in der Lage sein, Daten zu speichern, auf die durch den Prozessor der Datenverarbeitungseinheit zugegriffen werden kann. Zu Beispielen von durch einen Computer lesbaren Speichermedien gehören: eine Diskette, ein magnetisches Festplattenlaufwerk, eine Solid-State-Festplatte, Flash-Speicher, ein USB-Thumb-Drive, Direktzugriffsspeicher (RAM, Random Access Memory), Nur-Lese-Speicher (ROM, Read Only Memory), eine optische Platte, eine magnetooptische Platte und die

Registerdatei des Prozessors, aber nicht darauf beschränkt. Zu Beispielen von optischen Platten gehören Compact Disks (CDs) und Digital Versatile Disks (DVDs) wie zum Beispiel CD-ROM, CD-RW, CD-R, DVD-ROM, DVD-RW oder DVD-R Platten. Der Begriff „durch einen Computer lesbares Speichermedium“ bezieht sich auch auf verschiedene Typen von Aufzeichnungsmedien, auf die durch die Computereinheit über ein Netzwerk oder eine Datenübertragungsverbindung zugegriffen werden kann. Daten können zum Beispiel über einen Modem, über das Internet oder über ein lokales Netzwerk abgerufen werden. Auf einem durch einen Computer lesbaren Medium enthaltener auf einem Computer ausführbarer Code kann unter Verwendung eines beliebigen geeigneten Mediums übertragen werden, darunter drahtlos, drahtgebunden, Lichtwellenleiter-Kabel, HF usw. oder jede geeignete Kombination daraus, aber nicht darauf beschränkt.

[0038] Ein durch einen Computer lesbares Signalmedium kann ein weitergeleitetes Datensignal mit darin enthaltenem auf einem Computer ausführbarem Code beinhalten, zum Beispiel im Basisband oder als Teil einer Trägerwelle. Ein derartiges weitergeleitetes Signal kann eine beliebige Form aus einer Vielfalt an Formen annehmen, darunter elektromagnetische, optische bzw. jede geeignete Kombination daraus, jedoch nicht darauf beschränkt. Bei einem durch einen Computer lesbaren Signalmedium kann es sich um ein beliebiges durch einen Computer lesbares Medium handeln, das kein durch einen Computer lesbares Speichermedium ist und das ein Programm zur Verwendung durch oder in Verbindung mit einem System, einer Vorrichtung oder Einheit zum Ausführen von Anweisungen übertragen, weiterleiten bzw. transportieren kann.

[0039] ‚Computer-Hauptspeicher‘ oder ‚Hauptspeicher‘ ist ein Beispiel für ein durch einen Computer lesbares Speichermedium. Bei Computer-Hauptspeicher handelt es sich um jeden beliebigen Speicher, der direkt für einen Prozessor zugänglich ist. ‚Computer-Massenspeicher‘ oder ‚Massenspeicher‘ ist ein weiteres Beispiel für ein durch einen Computer lesbares Speichermedium. Bei Computer-Massenspeicher handelt es sich um jedes beliebige nichtflüchtige durch einen Computer lesbare Speichermedium. In einigen Ausführungsformen kann Computer-Massenspeicher auch Computer-Hauptspeicher sein oder umgekehrt.

[0040] In der Verwendung hierin umfasst ein ‚Prozessor‘ eine elektronische Komponente, die in der Lage ist, ein Programm oder eine durch eine Maschine ausführbare Anweisung oder auf einem Computer ausführbaren Code auszuführen. Verweise auf die Datenverarbeitungseinheit, die „einen Prozessor“ aufweist, sollten so ausgelegt werden, dass sie möglicherweise mehr als einen Prozessor oder Verar-

beitungskern enthält. Bei dem Prozessor kann es sich zum Beispiel um einen Mehrkernprozessor handeln. Ein Prozessor kann auch auf eine Sammlung von Prozessoren innerhalb eines einzelnen Computersystems oder verteilt über mehrere Computersysteme verweisen. Der Begriff „Datenverarbeitungseinheit“ sollte so ausgelegt werden, dass er möglicherweise auf eine Sammlung oder ein Netzwerk von Datenverarbeitungseinheiten verweist, die jeweils einen Prozessor oder Prozessoren aufweisen. Der auf einem Computer ausführbare Code kann durch mehrere Prozessoren ausgeführt werden, die sich innerhalb derselben Datenverarbeitungseinheit befinden können oder die sogar über mehrere Datenverarbeitungseinheiten verteilt sein können.

[0041] Auf einem Computer ausführbarer Code kann durch eine Maschine ausführbare Anweisungen oder ein Programm aufweisen, das einen Prozessor dazu veranlasst, einen Aspekt der vorliegenden Erfindung durchzuführen. Auf einem Computer ausführbarer Code für das Ausführen von Arbeitsschritten für Aspekte der vorliegenden Erfindung kann in einer beliebigen Kombination aus einer oder mehreren Programmiersprachen geschrieben werden, darunter objektorientierte Programmiersprachen wie Java, Smalltalk, C++ o. ä. sowie herkömmliche prozedurale Programmiersprachen wie die Programmiersprache „C“ oder ähnliche Programmiersprachen und kann in durch eine Maschine ausführbare Anweisungen kompiliert werden. In einigen Fällen kann der auf einem Computer ausführbare Code in Form einer hohen Sprache oder einer vorkompilierten Form vorliegen und in Verbindung mit einem Interpreter verwendet werden, der die durch eine Maschine ausführbaren Anweisungen während der Verarbeitung erzeugt.

[0042] Der auf einem Computer ausführbare Code kann vollständig auf dem Computer des Benutzers, teilweise auf dem Computer des Benutzers, als eigenständiges Software-Paket, teilweise auf dem Computer des Benutzers und teilweise auf einem fernen Computer oder vollständig auf dem fernen Computer oder Server ausgeführt werden. In letzterem Fall kann der entfernt angeordnete Computer mit dem Computer des Benutzers durch eine beliebige Art Netzwerk verbunden sein, darunter ein lokales Netzwerk (LAN) oder ein Weitverkehrsnetz (WAN), oder die Verbindung kann mit einem externen Computer hergestellt werden (zum Beispiel über das Internet unter Verwendung eines Internet-Dienstanbieters).

[0043] Aspekte der vorliegenden Erfindung werden unter Bezugnahme auf Ablaufpläne und/oder Blockschaltbilder bzw. Schaubilder von Verfahren, Vorrichtungen (Systemen) und Computerprogrammprodukten gemäß Ausführungsformen der Erfindung beschrieben. Es wird darauf hingewiesen, dass jeder Block oder ein Teil der Blöcke der Ablaufpläne, Veranschaulichungen und/oder der Blockschalt-

bilder bzw. Schaubilder durch Computerprogramm-anweisungen, falls zutreffend, in Form von auf einem Computer ausführbaren Code ausgeführt werden kann. Es wird ferner darauf hingewiesen, dass Kombinationen von Blöcken in verschiedenen Ablaufplänen, Veranschaulichungen und/oder Blockschaltbildern bzw. Schaubildern kombiniert werden können, sofern sie sich nicht gegenseitig ausschließen. Diese Computerprogrammanweisungen können einem Prozessor eines Universalcomputers, eines Spezialcomputers oder einer anderen programmierbaren Datenverarbeitungsvorrichtung bereitgestellt werden, um eine Maschine zu erzeugen, so dass die über den Prozessor des Computers bzw. der anderen programmierbaren Datenverarbeitungsvorrichtung ausgeführten Anweisungen ein Mittel zur Umsetzung der in dem Block bzw. den Blöcken der Ablaufpläne und/oder der Blockschaltbilder bzw. Schaubilder festgelegten Funktionen/Schritte erzeugen.

[0044] Diese Computerprogrammanweisungen können auch auf einem durch einen Computer lesbaren Medium gespeichert sein, das einen Computer oder eine andere programmierbare Datenverarbeitungsvorrichtung bzw. andere Einheiten so steuern kann, dass sie auf eine bestimmte Art funktionieren, so dass die auf dem durch einen Computer lesbaren Medium gespeicherten Anweisungen ein Herstellungsprodukt herstellen, darunter Anweisungen, welche die/den in dem Block bzw. den Blöcken des Ablaufplans und/oder der Blockschaltbilder bzw. Schaubilder angegebene(n) Funktion/Schritt umsetzen.

[0045] Die Computerprogrammanweisungen können auch auf einen Computer oder eine andere programmierbare Datenverarbeitungsvorrichtung bzw. andere Einheiten geladen werden, um das Ausführen einer Folge von Prozessschritten auf dem Computer, der anderen programmierbaren Vorrichtung bzw. den anderen Einheiten zu veranlassen, um einen auf einem Computer ausgeführten Prozess zu erzeugen, so dass die auf dem Computer oder einer anderen programmierbaren Vorrichtung ausgeführten Anweisungen Verfahren zur Umsetzung der in dem Block bzw. den Blöcken der Ablaufpläne und/oder der Blockschaltbilder bzw. Schaubilder angegebenen Funktionen/Schritte erzeugen.

[0046] Der Fachmann wird verstehen, dass Aspekte der vorliegenden Erfindung als Vorrichtung, Verfahren oder Computerprogrammprodukt ausgeführt werden können. Entsprechend können Aspekte der vorliegenden Erfindung die Form einer vollständigen Hardware-Ausführungsform, einer vollständigen Software-Ausführungsform (darunter Firmware, im Speicher befindliche Software, Mikrocode, usw.) oder einer Software- und Hardware-Aspekte kombinierenden Ausführungsform annehmen, die hierin alle allgemein als „Schaltkreis“, „Modul“ oder „System“ be-

zeichnet sein können. Des Weiteren können Aspekte der vorliegenden Erfindung die Form eines auf einem oder mehreren durch einen Computer lesbaren Medien enthaltenen Computerprogrammprodukts annehmen, die auf einem Computer ausführbaren Code enthalten.

[0047] Es sei darauf hingewiesen, dass ein oder mehrere der vorher angeführten Ausführungsformen kombiniert werden können, solange die kombinierten Ausführungsformen sich nicht gegenseitig ausschließen.

Kurzbeschreibung der Zeichnungen

[0048] Im Folgenden werden bevorzugte Ausführungsformen der Erfindung ausführlicher und lediglich beispielhaft unter Bezugnahme auf die Zeichnungen beschrieben, in denen:

[0049] Fig. 1 ein Datenbanksystem veranschaulicht, das in der Lage ist, das Verfahren zum Archivieren von Daten auszuführen;

[0050] Fig. 2 ein Ablaufplan eines beispielhaften Verfahrens zum Archivieren von Daten gemäß der vorliegenden Offenbarung ist;

[0051] Fig. 3 ein Ablaufplan eines anderen beispielhaften Verfahrens zum Archivieren von Daten gemäß der vorliegenden Offenbarung ist; und

[0052] Fig. 4 ein Ablaufplan eines weiteren beispielhaften Verfahrens zum Archivieren von Daten gemäß der vorliegenden Offenbarung ist.

Ausführliche Beschreibung

[0053] Im Folgenden geben gleich nummerierte Elemente in den Figuren entweder gleichartige Elemente oder eine gleichwertige Funktion durchführende Elemente an.

[0054] Vorher bereits erörterte Elemente werden nicht unbedingt in späteren Figuren erörtert, wenn die Funktion gleichwertig ist.

[0055] Die vorliegende Offenbarung geht die Unzulänglichkeit von Datenarchivierungsverfahren an, bei denen archivierte Daten sowie zugehörige Sicherungskopien möglicherweise nicht z. B. vor Datenverlust geschützt sind. Wenn Daten beruhend auf irgendwelchen Kriterien wie z. B. archivspezifischen Partitionen archiviert werden, setzt das Datenbanksystem möglicherweise nicht durch und kann nicht gewährleisten, dass keine neuen, die Kriterien des Teils der Tabelle (z. B. Partition) erfüllenden Datensätze hinzugefügt werden – oder bestehende Datensätze derart geändert werden, dass sie nun die Kriterien erfüllen. Auch sind die Sicherungsimagen mögli-

cherweise nicht geschützt, da sie zum Beispiel regelmäßig und automatisch gelöscht werden, wenn sie einen bestimmten Schwellwert erreichen (z. B. Anzahl von Tagen).

[0056] Das Einführen eines Tabellenraum/Partitions-Zustands und dessen Integrieren in die betroffenen Datenbankoperationen verhindern möglicherweise, dass neue Daten in archivierte Partitionen einer Tabelle eingefügt werden, verhindern möglicherweise das Erstellen von neuen Sicherungsimagen und verhindern möglicherweise das Löschen der n jüngsten Sicherungsimagen.

[0057] Fig. 1 zeigt ein beispielhaftes Datenbanksystem **100** für einen Datenarchivierungsprozess. Das Datenbanksystem **100** kann zum Beispiel Teil eines zEnterprise-Systems von IBM® sein, das ein IBM DB2-Datenbanksystem aufweist.

[0058] Das Datenbanksystem **100** stellt ein Computersystem **101** bereit. Die Komponenten des Computersystems **101** können eine(n) oder mehrere Prozessoren oder Verarbeitungseinheiten **103**, ein Massenspeichersystem **111**, ein Hauptspeichersystem **105** und einen Bus **107** aufweisen, der verschiedene Systemkomponenten, darunter das Hauptspeichersystem **105**, mit dem Prozessor **103** verbindet, sind aber nicht darauf beschränkt. Das Hauptspeichersystem **105** kann durch ein Computersystem lesbare Medien in Form eines flüchtigen Speichers wie einen Direktzugriffsspeicher (RAM) und/oder einen Cachespeicher enthalten.

[0059] Das Computersystem **101** enthält üblicherweise eine Vielfalt von durch ein Computersystem lesbaren Medien. Bei derartigen Medien kann es sich um jedes beliebige Medium handeln, auf welches das Computersystem **101** zugreifen kann, und es beinhaltet sowohl flüchtige als auch nichtflüchtige Medien, austauschbare und nicht austauschbare Medien.

[0060] Das Computersystem **101** kann auch mit einer oder mehreren externen Einheiten wie einer Tastatur, einer Zeigeeinheit, einer Anzeige **113** usw., einer oder mehreren Einheiten, die es einem Benutzer ermöglichen, mit dem Computersystem **101** Daten auszutauschen, und/oder beliebigen Einheiten (z. B. Netzwerkkarten, Modems usw.) Daten austauschen, die es dem Computersystem **101** ermöglichen, mit einer oder mehreren anderen Datenverarbeitungseinheiten Daten auszutauschen. Ein derartiger Datenaustausch kann über eine oder mehrere E/A-Schnittstellen **115** erfolgen. Außerdem kann das Computersystem **101** über einen Netzwerkadapter **109** mit einem oder mehreren Netzwerken Daten austauschen, wie zum Beispiel einem lokalen Netzwerk (LAN), einem allgemeinen Weitverkehrsnetz (WAN) und/oder einem öffentlichen Netzwerk (z. B. dem Internet). Wie abgebildet ist, tauscht der Netzwerkadapter **109** mit

den anderen Komponenten des Computersystems **101** über den Bus **107** Daten aus.

[0061] Das Computersystem **101** kann als ein mit einer Datenbankinfrastruktur **125** verbundenes Datenbankverwaltungs-Serversystem dienen. Das Computersystem **101** kann ein Datenbankverwaltungssystem beinhalten und umsetzen. Die Datenbankinfrastruktur **125** weist eine Quelldatenbank **127** auf. Die Datenbankinfrastruktur **125** kann zum Beispiel einen Cachespeicher **139** aufweisen. Der Inhalt des Cachespeichers **139** kann einen Einblick darin bereitstellen, welche Datenbankobjekte der Quelldatenbank **127** häufig verwendet werden und vor kurzem verwendet wurden. Dies kann zum Beispiel bei der Entscheidung darüber helfen, welche Daten zu archivieren sind. Das Datenbanksystem **100** weist ferner eine Zieldatenbank **135** auf. Alternativ kann es sich bei der Datenbank **135** um eine externe und/oder eine unabhängige Ressource des Datenbanksystems **100** handeln (in diesem Fall kann das Datenbanksystem **100** mit der Zieldatenbank **135** verbunden sein). Die Zieldatenbank **135** ist zum Speichern von archivierten Daten geeignet. Die Zieldatenbank **135** kann Teil eines Akzelerator-Serversystems sein (nicht gezeigt). Das Akzelerator-Serversystem kann zum Beispiel durch ein IBM Netezza-System umgesetzt sein, das mit IBM DB2 für z/OS integriert ist und als ein IBM DB2 Analytics Accelerator (IDAA) dient. Das Akzelerator-Serversystem kann eine überlegene Abfrageleistungsfähigkeit bereitstellen, wenn auf die archivierten Daten mindestens auf Nur-Lesen-Weise zugegriffen wird. Die in der Zieldatenbank **135** gespeicherten Daten können direkt z. B. über SQL abgefragt werden.

[0062] Bei der Quelldatenbank **127** kann es sich um eine interne Ressource des Computersystems **101** handeln, während es sich bei der Zieldatenbank **135** um eine externe und/oder eine unabhängige Ressource des Computersystems **101** handeln kann.

[0063] Das Hauptspeichersystem **105** ist so konfiguriert, dass es Anwendungen speichert, die auf dem Prozessor **103** ausführbar sind. Das Hauptspeichersystem **105** kann zum Beispiel ein Betriebssystem sowie Anwendungsprogramme aufweisen. Die Anwendungsprogramme weisen zum Beispiel ein Sicherungsmodul **129**, ein Archivierungsmodul **131** und ein Löschmodul **133** auf.

[0064] Das Sicherungsmodul **129** verwendet die Speichereinheit **111** als Sicherungsspeichereinheit. Bei der Speichereinheit **111** kann es sich um eine interne oder externe Ressource des Computersystems **101** oder eine Kombination aus den beiden handeln. Von dem Sicherungsmodul **129** kann jede beliebige Anzahl von lokalen und/oder externen Speichereinheiten zum Speichern von Sicherungsdaten verwendet werden.

[0065] Das Archivierungsmodul **131** kann zum Beispiel den Inhalt des Cachespeichers **139** verwenden, um zu ermitteln, welche in der Quelldatenbank **127** gespeicherten Daten in der Zieldatenbank **135** archiviert werden können. Das Archivierungsmodul **131** kann die Daten in der Zieldatenbank **135** archivieren.

[0066] Das Löschmodul **133** kann veraltete und/oder weniger häufig verwendete Daten aus der Quelldatenbank **127** löschen.

[0067] Das Sicherungs-, das Archiv- und das Löschmodul können Teil des umgesetzten Datenbankverwaltungssystems sein.

[0068] Der Betrieb des Datenbanksystems **100** wird unter Bezugnahme auf die Fig. 2 bis Fig. 4 ausführlich beschrieben.

[0069] Fig. 2 ist ein Ablaufplan eines beispielhaften Verfahrens zum Archivieren von Daten gemäß der vorliegenden Offenbarung. Fig. 2 zeigt die Quelldatenbank **127** als eine Datentabelle **201** enthaltend. Die Datentabelle **201** kann zum Beispiel Betriebsdaten aufweisen, die möglicherweise für nachfolgende Operationen auf dem Datenbanksystem **100** verwendet werden. Die Datentabelle **201** weist einen Teil **203A** auf.

[0070] Der Einfachheit der Beschreibung halber werden lediglich eine einfache Tabelle und ein einzelner Teil dieser Tabelle beschrieben; der Fachmann wird jedoch leicht verstehen, dass das vorliegende Verfahren auf mehr als eine Tabelle und mehr als einen Teil pro Tabelle verallgemeinert werden kann.

[0071] Der Teil der Tabelle **203A** weist möglicherweise veraltete Daten und/oder weniger häufig verwendete Daten auf. Der Teil der Tabelle **203A** kann von einem Benutzer des Datenbanksystems **100** festgelegt werden, der den Teil der Tabelle **203A** in der Zieldatenbank **135** zu archivieren wünscht. Dafür kann das Computersystem **101** von dem Benutzer zum Beispiel über die E/A-Schnittstelle **115** eine Auswahl des Teils der Tabelle **203A** empfangen. Alternativ kann der Teil der Tabelle **203A** automatisch durch das Computersystem **101** ausgewählt werden, indem zum Beispiel die Zugriffshäufigkeit (durch einen einzelnen oder mehrere Benutzer) auf den Teil der Tabelle **203A** und/oder das letzte Abänderungsdatum des Teils der Tabelle **203A** verwendet werden. Wenn die Zugriffshäufigkeit auf den Teil der Tabelle **203A** zum Beispiel unter einem vorher festgelegten Schwellwert liegt, kann das Computersystem **101** entscheiden, den Teil der Tabelle **203A** zu archivieren.

[0072] In Schritt **211** kann das Sicherungsmodul **129** eine erste Sicherungskopie **203B** des Teils der Tabelle **203A** erstellen. Die erste Sicherungskopie **203B**

kann in der Speichereinheit **111** gespeichert werden. Die Speichereinheit **111** kann mehrere Speicherschichten mit verschiedenen Leistungsfähigkeitseigenschaften aufweisen.

[0073] Bei den Leistungsfähigkeitseigenschaften kann es sich zum Beispiel E/A-Antwortzeit und/oder E/A-Durchsatz und/oder eine Kombination daraus handeln. Die Speicherschichten können beruhend auf deren Leistungsfähigkeitseigenschaften eingeordnet oder eingestuft werden. Die am höchsten eingestufte Speicherschicht kann zum Beispiel die geringste E/A-Antwortzeit haben. Die erste Sicherungskopie **203B** kann während einer ersten Zeitspanne in der am höchsten eingestuften Speicherschicht gespeichert werden (z. B. eine Woche nach der Erstellung der ersten Sicherungskopie **203B**). Die erste Sicherungskopie **203B** kann während einer zweiten Zeitspanne, die auf die erste Zeitspanne folgt, in der am zweithöchsten eingestuften Speicherschicht gespeichert werden. Dies kann vorteilhaft sein, da die erste Sicherungskopie **203B** während der ersten Zeitspanne möglicherweise häufiger verwendet wird als in der zweiten Zeitspanne und somit während der ersten Zeitspanne möglicherweise eine Speicherschicht mit besserer Leistungsfähigkeit benötigt wird. Alternativ kann die Speichereinheit **111** mit langsamerem Festplattenspeicher umgesetzt werden. Dies kann vorteilhaft sein, da hierdurch eine preisgünstigere Speichereinheit bereitgestellt werden kann, die insbesondere dann verwendet werden kann, wenn die Zugriffshäufigkeit auf die erste Sicherungskopie **203B** gering ist. Das Sicherungsmodul **129** kann einen Identifizierungsalgorithmus aufweisen, wobei der Identifizierungsalgorithmus der ersten Sicherungskopie **203B** eine erste Kennung zuweisen kann. Die erste Kennung kann einen eindeutigen Namen der ersten Sicherungskopie **203B** und/oder einen Zeitstempel der Erstellung der ersten Sicherungskopie **203B** aufweisen. Die erste Sicherungskopie **203B** kann mit einem Zeitstempel versehen werden, um die zeitliche Reihenfolge anzugeben, in der sie in der Speichereinheit **111** gespeichert wurde.

[0074] In Schritt **213** kann dem Teil **203A** der Tabelle ein Zugriffszustand-Flag zugewiesen werden. Das Zugriffszustand-Flag kann in der Quelldatenbank **127** gespeichert werden. Dies kann zum Beispiel durch das Hinzufügen von Metadaten zu der die Zugriffszustand-Flags gemäß der vorliegenden Offenbarung speichernden Tabelle **205** durchgeführt werden. Die Metadaten können den Teil der Tabelle **203A** und den Zustand des zugehörigen Zugriffszustand-Flag angeben. Bei dem Zugriffszustand-Flag kann es sich zum Beispiel um ein binäres Flag wie z. B. ein boolesches Flag handeln, und der Zustand des Zugriffszustand-Flag kann ein wahrer oder falscher Zustand sein. Alternativ kann das Zugriffszustand-Flag eine Vielzahl von Bits aufweisen. Jedes Bit kann einem Zustand des Zugriffszustand-Flag zugewiesen wer-

den. Das Setzen eines ersten Bits (z. B. auf den Wert 1) aus der Vielzahl von Bits kann eine Abänderung von bestehenden Sicherungskopien des Teils der Tabelle **203A** verhindern und zum Verhindern weiterer Sicherungsoperationen an dem Teil der Tabelle **203A**. Das Setzen eines zweiten Bits (z. B. auf den Wert 1) aus der Vielzahl von Bits kann eine Abänderung des Teils der Tabelle **203A** verhindern.

[0075] Zum Steuern des Zugriffs auf den Teil der Tabelle können (parallel zu dem Zugriffszustand-Flag) andere Datenbank-Flags verwendet werden. Bei diesen anderen Flags kann es sich um die Flags „Nur Dienstprogrammoperationen sind zulässig“, „Nur-Lesen“, „Dienstprogrammoperationen erforderlich“, „Dienstprogrammoperationen empfohlen“ handeln. In diesem Fall kann die erste Sicherungskopie **203B** durch das Zugriffszustand-Flag geschützt werden, während der Teil der Tabelle **203A** unter Verwendung der anderen Datenbank-Flags und des Zugriffszustand-Flag geschützt werden kann.

[0076] Das Zugriffszustand-Flag kann in betroffene, den Teil der Tabelle **203A** verwendende Datenbankoperationen integriert werden. Falls das Computersystem **101** zum Beispiel ein SQL-Datenbankverwaltungssystem umsetzt, können die Aussagen INSERT, UPDATE oder MERGE SQL so konfiguriert werden, dass sie beim Zugreifen auf den Teil der Tabelle **203A** das Zugriffszustand-Flag berücksichtigen.

[0077] Falls das Computersystem **101** in einem anderen Beispiel ein DB2-Datenbankverwaltungssystem umsetzt, können verschiedene mit der physischen Darstellung des Teils der Tabelle **203A** in der Quelldatenbank **127** arbeitende Dienstprogramme wie REORG, COPY, RUNSTATS und RESTORE so konfiguriert werden, dass sie das Zugriffszustand-Flag berücksichtigen. RUNSTATS kann zum Beispiel mit dem Teil der Tabelle **203A** als Nur-Lesen-Daten arbeiten und Statistiken darüber sammeln, wann das Zugriffszustand-Flag gesetzt ist. REORG schlägt möglicherweise beim Zugreifen auf den Teil der Tabelle **203A** fehl, da möglicherweise keine Änderungen an dem Teil der Tabelle **203A** gestattet sind, wenn das Zugriffszustand-Flag gesetzt ist und wenn das Zugriffszustand-Flag zum Verhindern einer Abänderung des Teils der Tabelle **203A** verwendet wird.

[0078] Gleichermaßen schlägt möglicherweise das Laden neuer Daten in den Teil der Tabelle **203A** fehl.

[0079] In Schritt **215** kann das Zugriffszustand-Flag zum Beispiel durch Setzen des booleschen Zustands in der Tabelle **205** auf „wahr“ (oder durch Setzen des ersten Bits auf 1) gesetzt werden. Wenn er auf „wahr“ gesetzt ist, kann das Sicherungsmodul **129** den Zustand des Zugriffszustand-Flag zum Verhindern einer Abänderung von bestehenden Sicherungskopien des Teils der Tabelle **203A** und zum Verhindern weiterer

Sicherungsoperationen an dem Teil der Tabelle **203A** verwenden.

[0080] Das Sicherungsmodul **129** kann zum Beispiel so konfiguriert sein, dass es die Schritte **211** bis **215** als atomare Operation ausführt.

[0081] Nach Schritt **215** kann das Sicherungsmodul **129** bestimmen (oder überprüfen), ob eine zweite Sicherungskopie des Teils der Tabelle **203A** mit einer zweiten Kennung (d. h. mit einem jüngsten Zeitstempel) vor dem Setzen des Zugriffszustand-Flag erstellt wurde. Die zweite Kennung wird unter Verwendung eines zweiten Identifizierungsalgorithmus festgelegt. Und falls die zweite Sicherungskopie vorhanden ist, wird eine dritte Sicherungskopie des Teils der Tabelle erstellt, und der erstellten dritten Sicherungskopie wird eine dritte Kennung zugewiesen. Die dritte Kennung kann unter Verwendung des ersten Identifizierungsalgorithmus festgelegt werden. Bei dem ersten Identifizierungsalgorithmus kann es sich um einen GUID-Algorithmus (Globally Unique Identifier, global eindeutige ID) handeln. Der erste Identifizierungsalgorithmus kann ein vorher festgelegtes Namensgebungsschema zum Kennzeichnen der ersten Sicherungskopie **203B** verwenden. Der erste Identifizierungsalgorithmus kann zum Beispiel die erste Sicherungskopie **203B** unter Verwendung eines Namens kennzeichnen, der einen Zeitstempel der Erstellung der ersten Sicherungskopie **203B** und einen eindeutigen Präfix enthält, der nur für archivierte Daten verwendet wird. Das vorher festgelegte Namensgebungsschema kann zum Kennzeichnen der dritten Sicherungskopie verwendet werden. Hierdurch kann die unbeabsichtigte Verwendung einer zweiten Sicherungskopie verhindert werden, die sich möglicherweise von der ersten Sicherungskopie unterscheidet (z. B. durch Verwendung eines anderen Identifizierungsalgorithmus).

[0082] In Schritt **217** kann das Archivierungsmodul **131** den Teil der Tabelle **203A** in der Zieldatenbank **135** archivieren, woraus sich eine archivierte Kopie **203C** des Teils der Tabelle **203A** ergibt.

[0083] Jeder Versuch des Einfügens von Datensätzen in die archivierte Kopie **203C** schlägt möglicherweise mit einer entsprechenden Fehlermeldung fehl. Das Computersystem **101** kann jedoch noch immer die archivierte Kopie **203C** abfragen, so dass Durchsuchungen von Tabellen zum Beispiel nicht fehlschlagen.

[0084] In Schritt **219** kann das Löschmodul **133** den Inhalt des Teils der Tabelle **203A** löschen, während es das Zugriffszustand-Flag beibehält. Falls das Zugriffszustand-Flag auch zum Verhindern einer Abänderung des Teils der Tabelle **203A** verwendet wird (d. h., wenn das Zugriffszustand-Flag gesetzt ist, ist die Abänderung des Teils der Tabelle **203A** verboten),

kann das Löschmodul **133** den Teil der Tabelle **203A** unabhängig davon, ob das Zugriffszustand-Flag gesetzt ist, löschen (d. h., das Löschmodul kann so konfiguriert sein, dass es den Zustand des Zugriffszustand-Flag während des Archivierungsprozesses ignoriert).

[0085] Wenn das Zugriffszustand-Flag gesetzt ist, können Anforderungen über neue Sicherungen des Teils der Tabelle **203A** Ignoriert werden oder mit einem Fehler fehlschlagen. Dies liegt daran, dass die erste Sicherungskopie **203B** mit den jetzt archivierten Daten (**203C**) noch immer vorhanden und gültig ist, d. h., in ihr sind die Daten zum Zwecke des Wiederherstellens gespeichert. Neue Sicherungen wären leer und somit nutzlos.

[0086] Nach dem Löschschritt **219** wird festgestellt, dass der Teil der Tabelle **203A** durch eine Operation mit Ausnahme des Löschschritts f) nach dem Erstellungsschritt **211** nicht abgeändert wurde, und falls die erste Sicherungskopie **203B** noch gültig ist, endet das Verfahren erfolgreich. Eine gültige Sicherungskopie kann eine nichtbeschädigte Sicherungskopie sein, bei der es sich um die letzte Sicherungskopie handelt, die für den Teil der Tabelle **203A** erstellt wird.

[0087] Genauso wichtig wie das Verhindern des Erstellens neuer Sicherungsbilder ist die Anforderung, das die Archivdaten enthaltende Sicherungsbild zu beibehalten. Das Datenbanksystem kann sämtliche Sicherungsbilder verwalten und löscht alte Sicherungsbilder zum Beispiel implizit oder explizit beruhend auf irgendwelchen Kriterien. Der dafür verantwortliche Algorithmus kann so konfiguriert sein, dass er das Zugriffszustand-Flag berücksichtigt. Wenn dieses Zugriffszustand-Flag gefunden wird, darf das letzte Sicherungsbild (oder die letzten „n“ Sicherungsbilder) nicht gelöscht werden.

[0088] Fig. 3 ist ein Ablaufplan eines beispielhaften Verfahrens zum Archivieren des Teils der Tabelle **203A**. In diesem beispielhaften Verfahren wird das Zugriffszustand-Flag ebenfalls zum Verhindern einer Abänderung des Teils der Tabelle **203A** verwendet.

[0089] In Schritt **301** werden die Schritte **211** bis **217** ausgeführt. In Schritt **303** wird das Zugriffszustand-Flag vor dem in Schritt **305** durchgeführten Löschschritt **219** aufgehoben (z. B. durch Umschalten des Zustands des Zugriffszustand-Flag in der Tabelle **205** auf „falsch“ oder durch Setzen des ersten Bitwerts auf 0). In Schritt **307** wird das Zugriffszustand-Flag zurückgesetzt.

[0090] In Schritt **309** wird ermittelt, ob der Teil **203A** der Tabelle durch eine Operation mit Ausnahme des Löschschritts f) nach dem Schritt **211** abgeändert wurde. Dies kann zum Beispiel unter Verwendung von historischen Daten (z. B. durch Verwendung des

Protokolls) durchgeführt werden, um zu ermitteln, ob in den Teil der Tabelle neue Daten eingefügt wurden oder ob an dem Teil der Tabelle Aktualisierungen des Inhalts vorgenommen wurden, bevor sie gelöscht wird oder eine Löschung des Inhalts des Teils der Tabelle durch eine andere Operation als Löschschritt f) vorgenommen wird, z. B. eine Löschoperation durch eine andere Anwendung, die möglicherweise vor dem Löschschritt f) der vorliegenden Offenbarung geschieht. Die Abänderung darf geschehen, wenn das Zugriffszustand-Flag aufgehoben ist. Falls der Teil der Tabelle durch eine Operation mit Ausnahme des Löschschritts f) nach dem Schritt **211** abgeändert wurde, wird in Schritt **311** eine Abänderung des Teils **203A** der Tabelle festgestellt. Dies kann zum Beispiel unter Verwendung von dem Teil der Tabelle **203A** zugehörigen statistischen Daten durchgeführt werden. Die statistischen Daten können im Fall von DB2 unter Verwendung der Echtzeitstatistiken (RTS, Real Time Statistics) erfasst werden.

[0091] In Schritt **313** wird das Zugriffszustand-Flag aufgehoben, um den Teil der Tabelle in Schritt **315** wiederherzustellen. In Schritt **317** wird die festgestellte Abänderung auf den Teil der Tabelle angewendet. Die Schritte werden für den abgeänderten Teil der Tabelle wiederholt.

[0092] Falls der Teil der Tabelle nach dem Schritt **211** nicht abgeändert wurde, wird in Schritt **319** ermittelt, ob eine zweite spätere Sicherungskopie vorhanden ist. Wenn die zweite Sicherungskopie vorhanden ist, wird das Zugriffszustand-Flag in Schritt **321** aufgehoben, der Teil der Tabelle wird unter Verwendung der ersten Sicherungskopie **203B** wiederhergestellt (Schritt **323**), und die Schritte des Verfahrens werden zum Erstellen einer neuen Sicherungskopie wiederholt. Wenn die zweite Sicherungskopie nicht vorhanden ist, endet das Verfahren erfolgreich.

[0093] Fig. 4 ist ein Ablaufplan eines beispielhaften Verfahrens zum Archivieren des Teils **203A** der Tabelle. In diesem beispielhaften Verfahren wird das Zugriffszustand-Flag nicht zum Verhindern einer Abänderung des Teils der Tabelle **203A** verwendet.

[0094] In Schritt **401** werden die Schritte **211** bis **219** ausgeführt. In Schritt **403** wird ermittelt, ob der Teil **203A** der Tabelle durch eine Operation mit Ausnahme des Löschschritts f) nach dem Schritt **211** abgeändert wurde.

[0095] Falls der Teil der Tabelle durch eine Operation mit Ausnahme des Löschschritts f) nach dem Schritt **211** abgeändert wurde, wird in Schritt **405** eine Abänderung des Teils **203A** der Tabelle festgestellt. Dies kann zum Beispiel unter Verwendung von dem Teil der Tabelle zugehörigen statistischen Daten durchgeführt werden.

[0096] In Schritt **407** wird der Teil der Tabelle unter Verwendung der ersten Sicherungskopie wiederhergestellt. In Schritt **409** wird die festgestellte Abänderung auf den wiederhergestellten Teil der Tabelle angewendet. Die Schritte des Verfahrens werden für den abgeänderten Teil der Tabelle wiederholt.

[0097] Falls der Teil der Tabelle nach dem Schritt **211** nicht abgeändert wurde, wird in Schritt **411** ermittelt, ob eine zweite spätere Sicherungskopie vorhanden ist. Wenn die zweite Sicherungskopie vorhanden ist, wird der Teil der Tabelle unter Verwendung der ersten Sicherungskopie in Schritt **413** wiederhergestellt, und die Schritte des Verfahrens werden zum Erstellen einer neuen Sicherungskopie wiederholt. Wenn die zweite Sicherungskopie nicht vorhanden ist, endet das Verfahren erfolgreich.

Bezugszeichenliste

100	Datenbanksystem
101	Computersystem
103	Prozessor
105	Hauptspeicher
107	Bus
109	Netzwerkadapter
111	Speichersystem
113	Anzeige
115	E/A-Schnittstelle
125	Datenbankinfrastruktur
127	Quelldatenbank
129	Sicherungsmodul
131	Archivierungsmodul
133	Löschmodul
135	Zieldatenbank
139	Cachespeicher
201, 205	Tabelle
203A	Teil der Tabelle
203B	erste Sicherungskopie
203C	archivierte Kopie.

Patentansprüche

1. Auf einem Computer ausgeführtes Verfahren zum Archivieren eines vorher festgelegten Teils (**203A**) einer sich in einem Datenbanksystem (**100**) befindlichen Tabelle (**201**), das Verfahren aufweisend:
 - a. Erstellen einer ersten Sicherungskopie (**203B**) des Teils der Tabelle (**203A**);
 - b. Zuweisen eines Zugriffszustand-Flag zu dem Teil der Tabelle (**203A**);
 - c. Setzen des Zugriffszustand-Flag;
 - d. Verwenden des Zugriffszustand-Flag zum Verhindern einer Abänderung von bestehenden Sicherungskopien des Teils der Tabelle (**203A**) und zum Verhindern weiterer Sicherungsoperationen an dem Teil der Tabelle (**203A**);
 - e. Archivieren des Teils der Tabelle (**203A**) auf einem separaten System (**135**);

f. Löschen des Inhalts des Teils der Tabelle (**203A**) aus dem Datenbanksystem und Beibehalten des Zugriffszustand-Flag in dem Datenbanksystem.

2. Verfahren nach Anspruch 1, wobei die erste Sicherungskopie (**203B**) eine erste Kennung aufweist, die unter Verwendung eines ersten Identifizierungsalgorithmus festgelegt wurde, das Verfahren ferner aufweisend: nach dem Erstellungsschritt a) und nach dem Setzschrift c), Ermitteln, ob eine zweite Sicherungskopie des Teils der Tabelle (**203A**) mit einer unter Verwendung eines zweiten Identifizierungsalgorithmus festgelegten zweiten Kennung erstellt wurde, wobei, falls die zweite Sicherungskopie erstellt wurde, eine dritte Sicherungskopie des Teils der Tabelle (**203A**) mit einer dritten Kennung erstellt wird, die unter Verwendung des ersten Identifizierungsalgorithmus festgelegt wird.

3. Verfahren nach Anspruch 1, ferner aufweisend das Verwenden des Zugriffszustand-Flag zum Verhindern einer Abänderung des Teils der Tabelle (**203A**).

4. Verfahren nach Anspruch 3, ferner aufweisend die folgenden Schritte:

- Aufheben des Zugriffszustand-Flag vor dem Löschschritt f); und
- Zurücksetzen des Zugriffszustand-Flag nach dem Löschschritt f).

5. Verfahren nach Anspruch 1, ferner aufweisend nach dem Löschschritt f):

- Feststellen, dass der Teil der Tabelle (**203A**) durch eine Operation mit Ausnahme des Löschschritts f) nach dem Erstellungsschritt a) nicht abgeändert wurde;
- Feststellen, dass es sich bei der ersten Sicherungskopie (**203B**) um die letzte Sicherungskopie des Teils der Tabelle handelt, und erfolgreiches Beenden des Verfahrens.

6. Verfahren nach Anspruch 4, ferner aufweisend nach dem Zurücksetzen des Zugriffszustand-Flag:

- Feststellen, dass der Teil der Tabelle (**203A**) durch eine Operation mit Ausnahme des Löschschritts f) nach dem Erstellungsschritt a) nicht abgeändert wurde
- als Reaktion auf eine Feststellung, dass eine zweite spätere Sicherungskopie des Teils der Tabelle (**203A**) vorhanden ist, Feststellen, dass die erste Sicherungskopie (**203B**) ungültig ist;
- Aufheben des Zugriffszustand-Flag;
- Wiederherstellen des Teils der Tabelle unter Verwendung der ersten Sicherungskopie; und
- erneutes Ausführen der Schritte a) bis f), wodurch eine neue Sicherungskopie erstellt wird.

7. Verfahren nach Anspruch 4, ferner aufweisend nach dem Zurücksetzen des Zugriffszustand-Flag:

- als Reaktion auf eine Feststellung, dass der Teil der Tabelle (**203A**) durch eine Operation mit Ausnahme des Löschschritts f) nach dem Erstellungsschritt a) abgeändert wurde, Ermitteln einer Abänderung des Teils der Tabelle (**203A**) unter Verwendung von dem Teil der Tabelle zugehörigen statistischen Daten,
- Aufheben des Zugriffszustand-Flag;
- Wiederherstellen des Teils der Tabelle (**203A**) unter Verwendung der ersten Sicherungskopie,
- Anwenden der ermittelten Abänderung des Teils der Tabelle (**203A**) und erneutes Ausführen der Schritte a) bis f) für den abgeänderten Teil der Tabelle.

8. Verfahren nach Anspruch 4, ferner aufweisend nach dem Zurücksetzen des Zugriffszustand-Flag:

- Feststellen, dass der Teil der Tabelle (**203A**) durch eine Operation mit Ausnahme des Löschschritts f) nach dem Erstellungsschritt a) nicht abgeändert wurde;
- Feststellen, dass es sich bei der ersten Sicherungskopie um die letzte Sicherungskopie des Teils der Tabelle handelt, und erfolgreiches Beenden des Verfahrens.

9. Verfahren nach Anspruch 1, ferner aufweisend nach dem Löschschritt f)

- Feststellen, dass der Teil der Tabelle (**203A**) durch eine Operation mit Ausnahme des Löschschritts f) nach dem Erstellungsschritt a) nicht abgeändert wurde
- als Reaktion auf eine Feststellung, dass eine zweite spätere Sicherungskopie des Teils der Tabelle (**203A**) vorhanden ist, Feststellen, dass die erste Sicherungskopie (**203B**) ungültig ist;
- Wiederherstellen des Teils der Tabelle unter Verwendung der ersten Sicherungskopie; und
- erneutes Ausführen der Schritte a) bis f), wodurch eine neue Sicherungskopie erstellt wird,

10. Verfahren nach Anspruch 1, ferner aufweisend nach dem Löschschritt f)

- als Reaktion auf eine Feststellung, dass der Teil der Tabelle (**203A**) durch eine Operation mit Ausnahme des Löschschritts f) nach dem Erstellungsschritt a) abgeändert wurde, Ermitteln einer Abänderung des Teils der Tabelle unter Verwendung von dem Teil der Tabelle zugehörigen statistischen Daten,
- Wiederherstellen des Teils der Tabelle (**203A**) unter Verwendung der ersten Sicherungskopie (**203B**),
- Anwenden der ermittelten Abänderung des Teils der Tabelle (**203A**) und erneutes Ausführen der Schritte a) bis f) für den abgeänderten Teil der Tabelle.

11. Verfahren nach Anspruch 3, wobei das Löschen durch ein Löschmodul (**133**) des Datenbanksystems (**100**) durchgeführt wird, das Verfahren ferner aufweisend das Steuern des Löschmoduls (**133**)

zum Löschen des Teils der Tabelle unabhängig davon, ob das Zugriffszustand-Flag gesetzt ist.

12. Verfahren nach Anspruch 1, wobei das Erstellen einer ersten Sicherungskopie (**203B**) durch ein Sicherungsmodul (**129**) des Datenbanksystems durchgeführt wird, wobei das Verfahren ferner das Steuern des Sicherungsmoduls (**129**) aufweist, damit es die Schritte a, b und c als atomare Operation beinhaltet.

13. Verfahren nach Anspruch 3, ferner aufweisend: Empfangen einer Anforderung über das Wiederherstellen des Teils der Tabelle (**203A**); Aufheben des Zugriffszustand-Flag; Wiederherstellen des Teils der Tabelle (**203A**) unter Verwendung der ersten Sicherungskopie (**203B**).

14. Computerprogrammprodukt, das auf einem Computer ausführbare Anweisungen zum Durchführen der Verfahrensschritte des Verfahrens nach einem der vorhergehenden Ansprüche aufweist.

15. Datenbanksystem (**100**) zum Archivieren eines vorher festgelegten Teils einer Tabelle (**203A**), die sich in dem Datenbanksystem befindet, das Datenbanksystem (**100**) aufweisend:

– ein Sicherungsmodul (**129**), das zu Folgendem in der Lage ist:

Erstellen einer ersten Sicherungskopie (**203B**) des Teils der Tabelle (**203A**);

Zuweisen eines Zugriffszustand-Flag zu dem Teil der Tabelle;

Setzen des Zugriffszustand-Flag;

Verwenden des Zugriffszustand-Flag zum Verhindern einer Abänderung von bestehenden Sicherungskopien des Teils der Tabelle und zum Verhindern weiterer Sicherungsoperationen an dem Teil der Tabelle;

– ein Archivierungsmodul (**131**), das in der Lage ist, den Teil der Tabelle (**203A**) auf einem separaten System (**135**) zu archivieren;

– ein Löschmodul (**133**), das in der Lage ist, den Inhalt des Teils der Tabelle (**203A**) aus dem Datenbanksystem zu löschen und das Zugriffszustand-Flag in dem Datenbanksystem beizubehalten.

Es folgen 4 Seiten Zeichnungen

Anhängende Zeichnungen

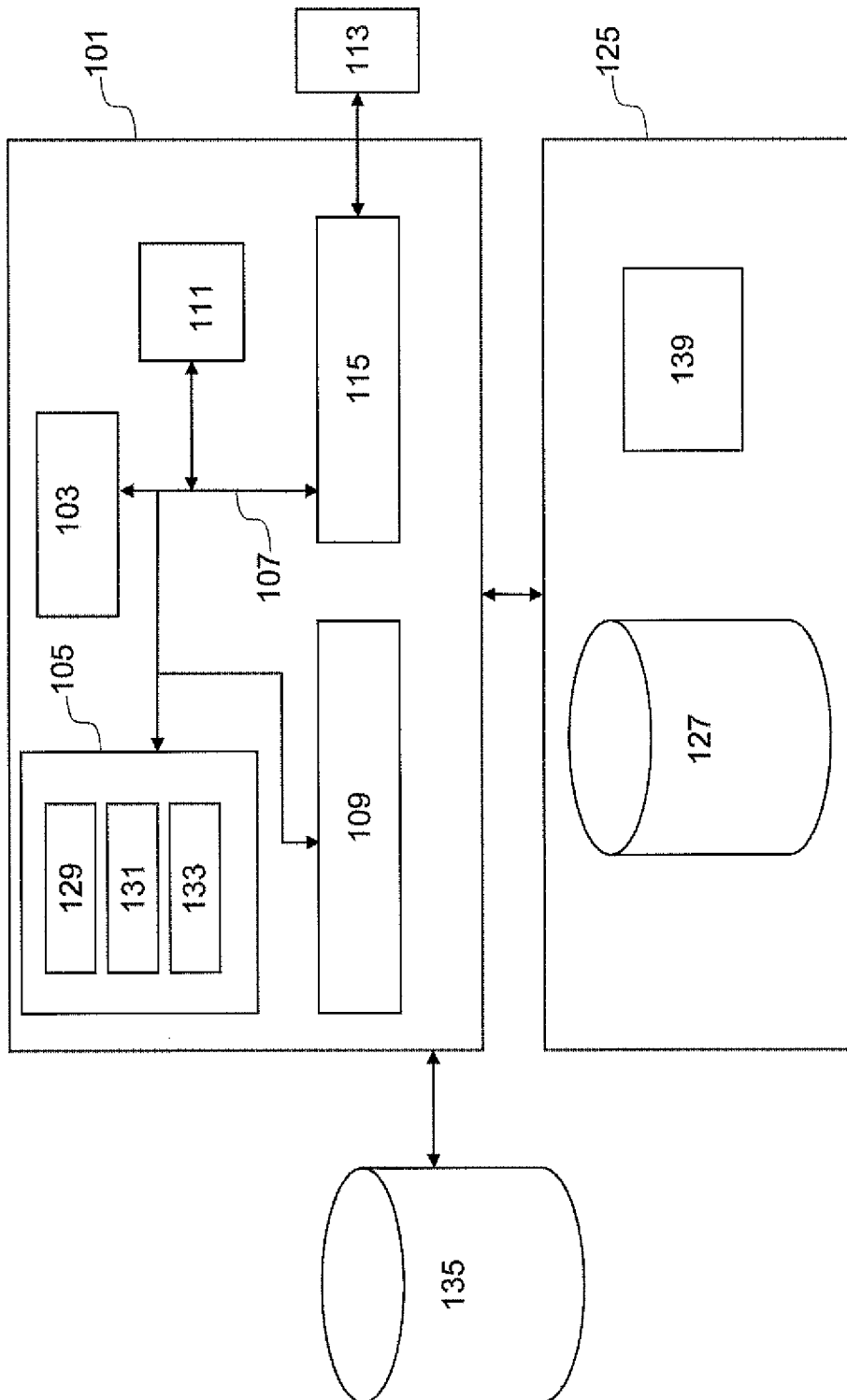


Fig. 1

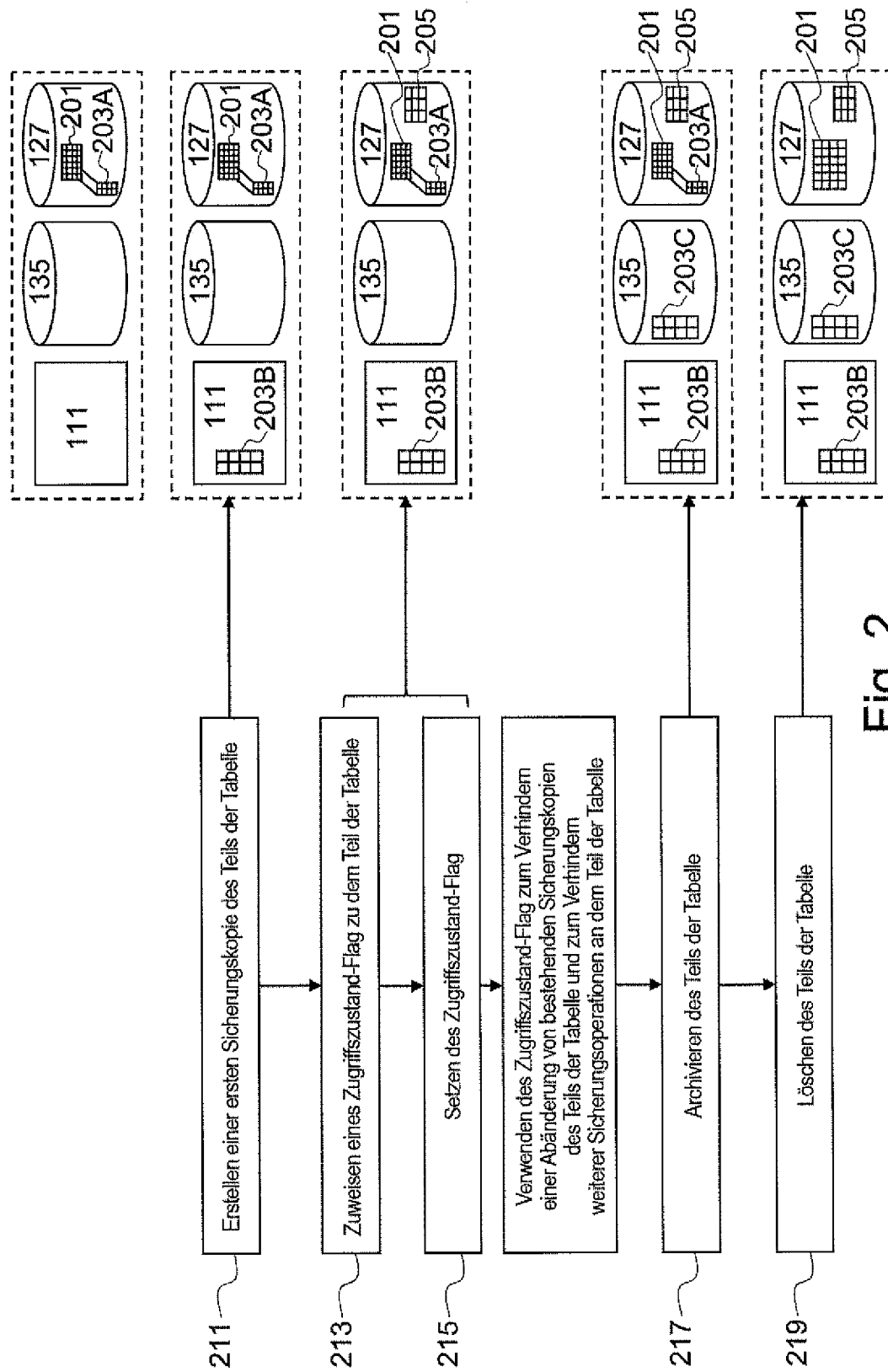


Fig. 2

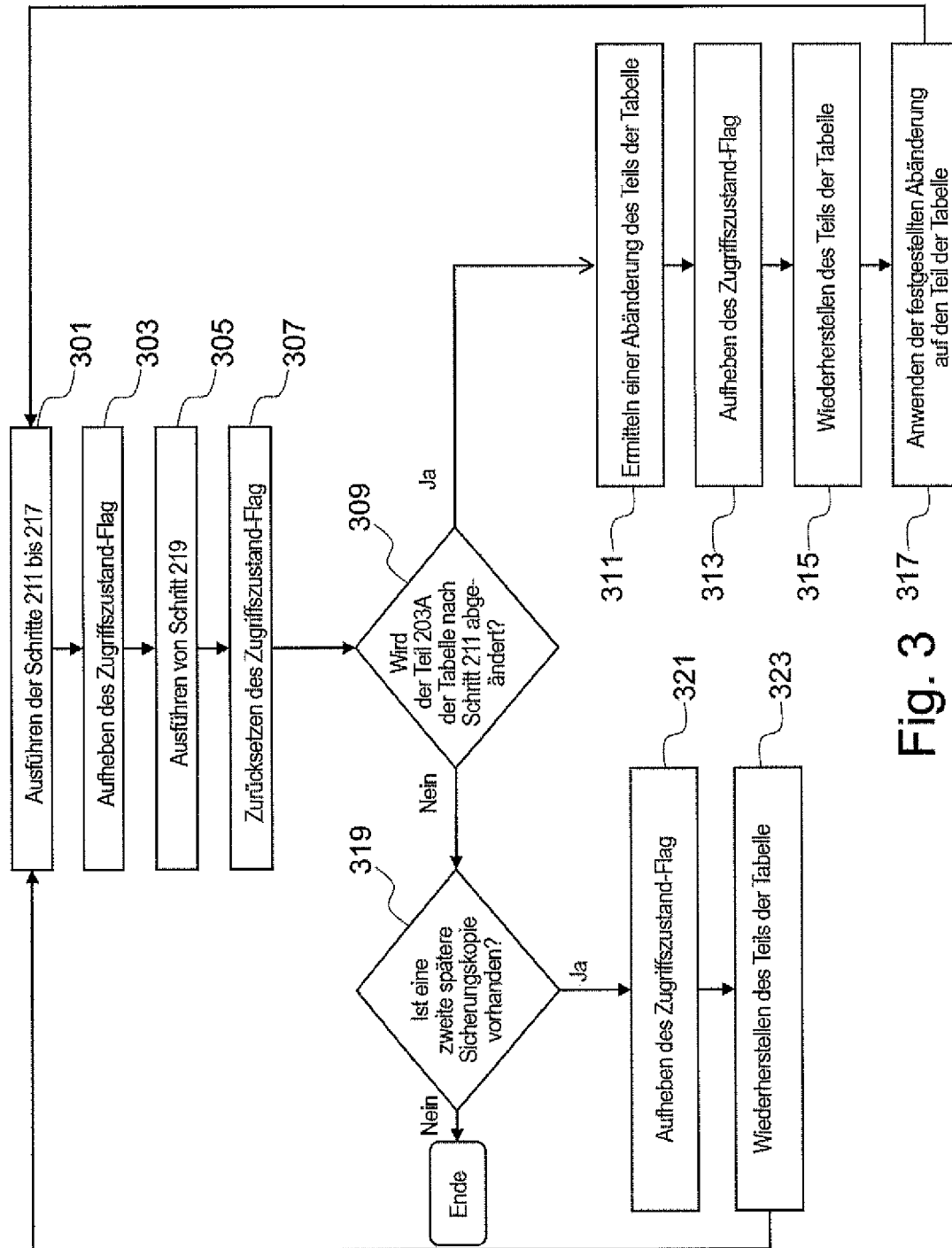


Fig. 3

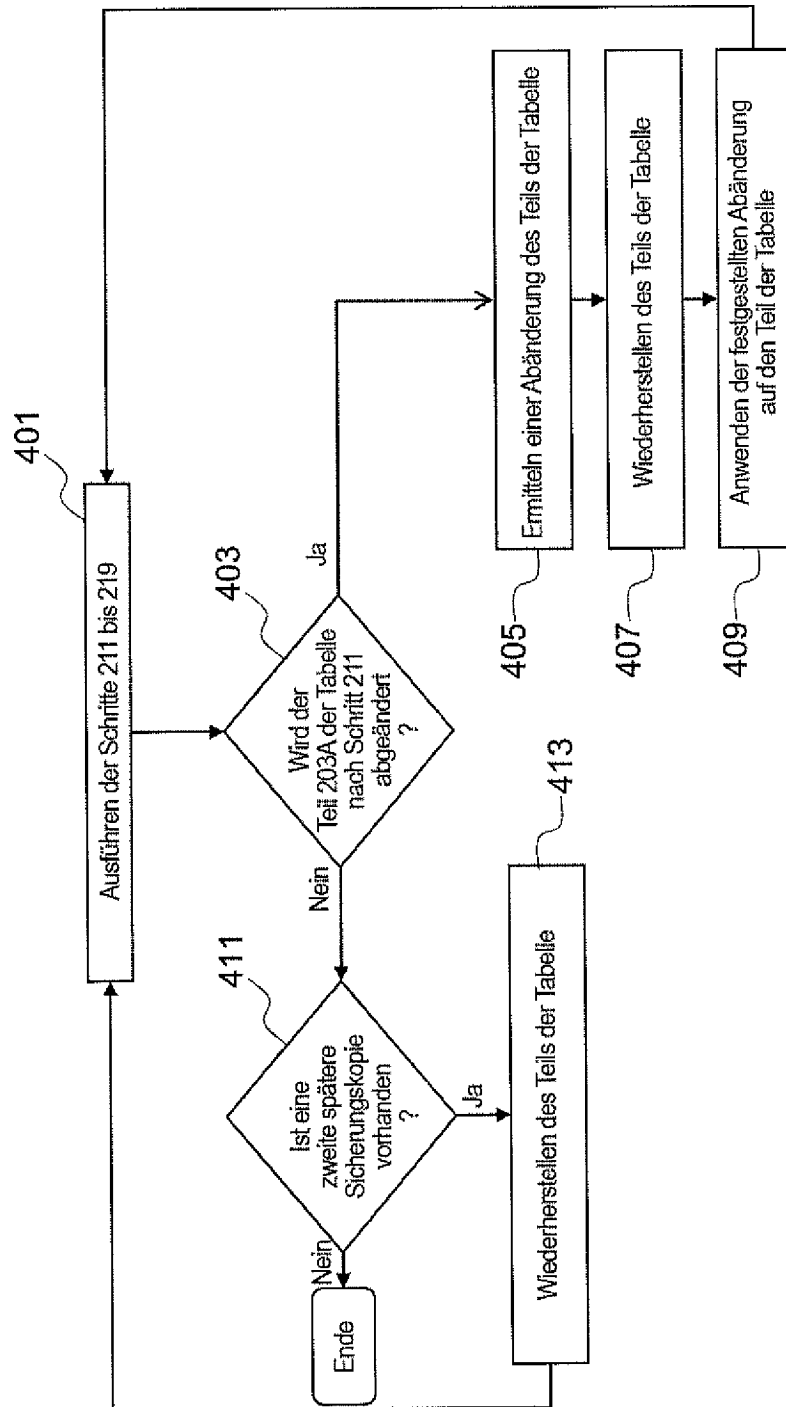


Fig. 4