

(19) 世界知的所有権機関
国際事務局



(43) 国際公開日
2001年10月18日 (18.10.2001)

PCT

(10) 国際公開番号
WO 01/77835 A1

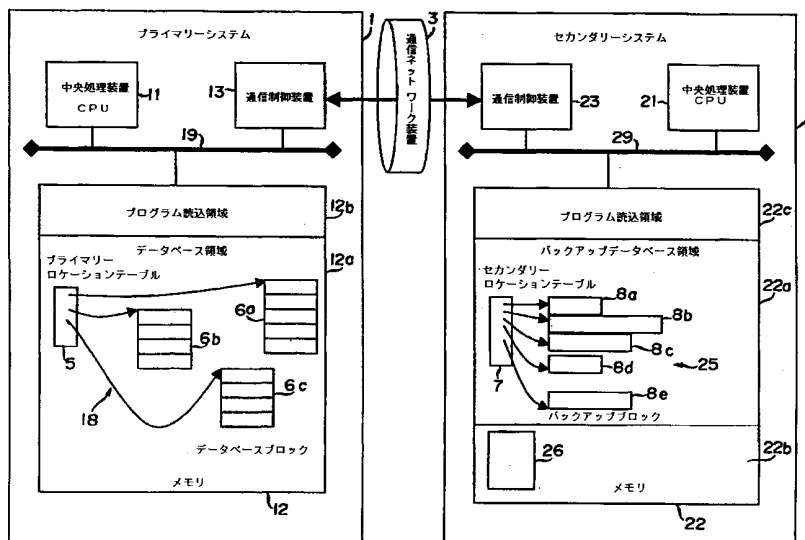
- (51) 国際特許分類: G06F 12/00
- (21) 国際出願番号: PCT/JP01/03126
- (22) 国際出願日: 2001年4月11日 (11.04.2001)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願2000-111299 2000年4月12日 (12.04.2000) JP
特願2001-094678 2001年3月29日 (29.03.2001) JP
- (71) 出願人 および
(72) 発明者: 玉津雅晴 (TAMATSU, Masaharu) [JP/JP]; 〒206-0023 東京都多摩市馬引沢2丁目14番14号 サンセットヒルズⅡ403 Tokyo (JP).
- (74) 代理人: 弁理士 市之瀬宮夫 (ICHINOSE, Miyao); 〒102-0083 東京都千代田区麹町3丁目1番8号 メイゾン麹町604号 Tokyo (JP).
- (81) 指定国 (国内): US.
- (84) 指定国 (広域): ヨーロッパ特許 (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR).
- (71) 出願人 (米国を除く全ての指定国について): アネックスシステムズ株式会社 (ANNEX SYSTEMS INCORPORATED) [JP/JP]; 〒107-0061 東京都港区北青山3丁目7番1号 Tokyo (JP).

添付公開書類:
 — 国際調査報告書
 — 請求の範囲の補正の期限前の公開であり、補正書受領の際には再公開される。

[続葉有]

(54) Title: DATA BACKUP/RECOVERY SYSTEM

(54) 発明の名称: データバックアップ・リカバリー方式



- 1...PRIMARY SYSTEM
- 11...CENTRAL PROCESSING UNIT CPU
- 13...COMMUNICATION CONTROL DEVICE
- 12b...PROGRAM READ AREA
- 12a...DATABASE AREA
- 5...PRIMARY LOCATION TABLE
- 6c...DATABASE BLOCK
- 12...MEMORY

- 3...COMMUNICATION NETWORK DEVICE
- 2...SECONDARY SYSTEM
- 23...COMMUNICATION CONTROL DEVICE
- 21...CENTRAL PROCESSING UNIT CPU
- 22c...PROGRAM READ AREA
- 22a...BACKUP DATABASE AREA
- 7...SECONDARY LOCATION TABLE
- 8e...BACKUP BLOCK
- 22b...MEMORY

(57) Abstract: A data backup/recovery system comprises a primary system for data update and a secondary system for data backup. The secondary system is installed near the primary system or at a place remote therefrom. The secondary system has the same logical structure as that of the primary one and acquires the latest data in real time to backup the data. The state of the primary system can be returned to the one at a past arbitrary time point

[続葉有]



WO 01/77835 A1



2文字コード及び他の略語については、定期発行される各PCTガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

with reference to the update history that the secondary system has. By using the secondary system as a reference system, load distribution is realized; and by stopping the backup operation of the secondary system and then separating it from the on-line processing system, it can be used as a reference batch processing system.

(57) 要約:

本方式はデータ更新を行なうプライマリーシステムとそのバックアップを行うセカンダリーシステムとから構成され、セカンダリーシステムはプライマリーシステムの近傍や遠隔地等に設置される。セカンダリーシステムはプライマリーシステムと論理的に同じ構成であり、リアルタイムに最新データを取得してバックアップを行う。セカンダリーシステムが保持している更新履歴をもとにプライマリーシステムの状態を過去の任意の時点に復元も可能である。セカンダリーシステムを参照系として利用することにより負荷分散を実現したり、セカンダリーシステムのバックアップ処理を停止し、オンライン処理系から切り離すことによつて参照系バッチ処理システムとして利用することもできる。

明 細 書

データバックアップ・リカバリー方式

5 技術分野

本発明はコンピュータにおけるデータバックアップ・リカバリー方式に係り、特に、コンピュータのデータバックアップを、セカンダリーシステムに対してプライマリーシステムのデータの変更がある都度、同じ内容を記憶させることにより、バックアップの時間と手間を大幅に削減するとともに、当該バックアップデータを用いることにより、リカバリーを行なうに要する時間と手間を大幅に削減したデータバックアップ・リカバリー方式に関する。

背景技術

一般に、コンピュータを使用してデータベースを構築し管理する場合には、ハードウェアの故障、ソフトウェアの障害あるいは災害などによるデータ消失、破損を避けるため、データをバックアップしておくことが必須のことである。

このようなコンピュータの利用分野において、データをバックアップする方式は、その対処する方策に応じて、従来、種々のものが提案されている。ここで、従来の各データバックアップ方式の代表的なものを挙げる。まず、従来の第1のデータバックアップ方式は、定期的にファイル全体のコピーを取得する方式である。

従来の第2のデータバックアップ方式は、定期的にファイル全体のコピーをとり、コピーを取らない期間にファイルの更新があったときには、その更新データをログファイルに記憶させるという方式である。

従来の第3のデータバックアップ方式は、プログラムの異常終了やトランザクション・キャンセルに備える方式である。

従来の第4のデータバックアップ方式は、プログラムエラーにより更新結果が誤

っている場合に備える方式である。

従来の第5のデータバックアップ方式は、災害に対処するためにバックアップを行う方式である。

従来の第6のデータバックアップ方式は、ファイル媒体の破壊に備えてバックアップを行なう方式である。

従来の第7のデータバックアップ方式は、正データが更新された時にバックアップを取得する方式である。なお、ここで、正データとは、いわゆる本番データ、すなわち、処理を直接的に行うデータのことをいう。

以下、これら従来の各データバックアップ方式を説明する。

10 従来の第1のデータバックアップ方式は、定期的にファイル全体のコピーを取得する方式である。この第1のデータバックアップ方式では、正データをコピーした後に行われた正データの更新があったときには、この更新データがバックアップファイルに反映されないものであった。したがって、この第1のデータバックアップ方式では、バックアップ間隔によって異なるが、大量の更新データが無くなってしま
15 まう危険性があった。

従来の第2のデータバックアップ方式は、主としてオンライン処理でバックアップ処理を行なうものであり、定期的にファイル全体のコピーを磁気テープなどに取得するほか、コピーを行わない期間に、ファイルの更新が行われるときには、ログ
20 ファイルを磁気ディスク装置や磁気テープ装置によって取得するという方法でバックアップしていた。もう少し詳細に述べると、具体的なバックアップの方式に若干の差はあるものの以下のような方式となる。

すなわち、データの格納してあるファイルの破壊が発生したときに備えて、ファイル全体のコピーを取得する。これは、運用によって周期を決定するが、毎日とか
1週間に1回といった頻度で行われる。その際に、正データの格納されたファイル
25 全体を一体とせず分割してコピーを取得する方法もある。

上述した従来の第2のデータバックアップ方式では、コピーを取得するときには、

データの更新が行われていると整合性が取れなくなるため、更新処理を中断する。

また、この方式の場合、コピー取得後に発生する正データの更新時には、そのログデータを取得しておく。ログデータには、更新データそのものであるトランザクションログ（以下「Tログ」という）、更新対象データの更新前イメージログ（以下「Bログ」という）、更新対象データの更新後ログ（以下「Aログ」という）の3種からなる。

これらのログのことを簡単に説明する。例えば、銀行に残高が「10万円」あり、「1万円」を引き出すことにより、残高が「9万円」になるというケースの場合には、この「10万円」の残高がBログであり、「1万円」の引き出しがTログ、結果として「9万円」の残高がAログということになる。

ここで、何らかの原因でコンピュータが故障して正データが破壊されたときには、最も最近取得したコピー全体または破壊分に相当するコピー全体の一部を使用して、正データをコピー時の状態に復元している。その後、コピーを取得したとき以降のログファイルを使用してファイル破壊直前の状態に復元していた。このような従来の第2のデータバックアップ方式では、正データの容量が増大するに伴って、コピーの取得や、ファイル破壊時の復元作業に相当な時間を要する不都合があった。また、コピーを取得するときに、データ更新を止めなくてはならず、24時間運転をすることは困難であるという不都合があった。

上述した第1のデータバックアップ方式及び第2のデータバックアップ方式は、ファイルや装置の障害に対処したものである。

次に、従来の第3のデータバックアップ方式は、プログラムの異常終了やトランザクション・キャンセルに備える方式である。この第3のデータバックアップ方式では、例えば、プログラムが異常終了した場合や、トランザクション・キャンセルが発生した場合に備えて、1つのトランザクション（1まとまりの処理）がスタートしてから終了するまでの間、更新するデータの更新前の内容（Bログと内容的には同一である）を保存しておく。プログラムの異常終了やトランザクション・キャ

ンセルが発生すると、そのトランザクションで更新されたデータを

更新前の状態に復元するために、更新前の内容を使用して復元を行う。

デッドロックが発生した場合もトランザクション・キャンセルと同様の処理が必要となる。

- 5 この従来の第3のデータバックアップ方式によれば、たまにしか発生しない異常終了やトランザクション・キャンセルに備えて常に正データのコピーデータを保存しなければならないという非効率的な面をもっていた。

- 次に、従来の第4のデータバックアップ方式について説明する。この第4のデータバックアップ方式はプログラムエラーにより更新結果が誤っている場合に備える方式である。この第4のデータバックアップ方式は、プログラムが正しく無い場合に問題になる。例えば、「10万円」の残高の預金から「1万円」を引き出したら、引出し後の残高が「11万円」になっているといった場合を想定したものである。このような場合でも、第4のデータバックアップ方式は、誤ったプログラムが適用される直前の状態に正データを復元し、その後のTログを元に正しいプログラムを稼働させることによりデータの修復を行う。
- 10
- 15

- 次に、災害対策を主目的にした第5のデータバックアップ方式について説明する。この第5のデータバックアップ方式では、バックアップ取得の目的は災害に対処するものである。ここで、「災害」とは、火災や水害、地震などのことである。このような災害によってファイルが消失することを防ぐ目的で、従来の第5のデータバックアップ方式では、バックアップファイルやログファイルのコピーを取り、耐火金庫に入れる。
- 20

さらに、厳重に行なう方法としては、バックアップファイルやログファイルのコピーを取り、それを遠隔地に輸送したり伝送するなどして、本番用ファイルが万一消失してもよいように備えていた。

- 25 しかしながら、遠隔地へは、一旦取得したファイルをコピーして送るので、本番用のファイルのために取得しているコピーやログと全く同じものを保管することが

できず、本番用ファイルが消失した場合には、一定期間分の更新データが無効になるという不都合があった。

また、従来の第6のデータバックアップ方式について説明する。この第6のデータバックアップ方式は、ファイル媒体の破壊に備えてバックアップを行なう方式である。この第6のデータバックアップ方式では、ファイル媒体の破壊に備える方法として、レイド (RAID ; Redundant Array of Inexpensive Disk) と称するバックアップ技術である。

この第6のデータバックアップ方式は、全く同一のファイルを2重に持つ方法や、ファイルの内容を複数の記憶装置に分散して書き込む方法や、あるいは、パリティを創生してデータを分割して記憶装置に書き込む方法などを採用したものである。

このデータバックアップ方式は、CPUやソフトからみると1つのディスク装置に書き込んでいるように見えるものであり、本番用ファイルとバックアップ用ファイルとが同一装置内に収められている。このため、この方式では、災害の対応が全くできないという不都合があった。

また、第6のデータバックアップ方式では、オンラインでの異常終了に対応するバックアウトには対応できておらず、プログラム・エラーによる、過去の時点へデータを戻すことに対応はできていないという不都合があった。

また、このデータバックアップ方式では、書き込み処理が通常の場合に比較して時間が掛かるという欠点もあった上に、データ破壊に対する復元がディスクボリューム単位なため、復元のための時間が長時間になるといった欠点があった。また、レイド (RAID) は、全く同一な装置でなければ構成できないものであった。

これを更に進化させた方式も実現している。これはディスクをミラー化しているが、バックアップ装置を遠隔地にも設置できるようになっている。これは、本番用のディスク装置が更新されると、更新されたデータが格納されている番地と更新された内容を、バックアップ用の装置に送信するものである。更に必要に応じて、バックアップ装置の更新を一定時間止めておき、更新可能になった時点で、バックア

ップ装置に溜まっている更新データにより復元を行い、本番用の装置と同一内容になるまで実行する機能を備えたものもある。

この方式では、リアルタイムでバックアップが行われるといった利点があるが、以下のような欠点があった。

- 5 すなわち、本番用とバックアップ用の装置間のミラーリングは、ディスク装置のハードウェア番地を使用しているため、本番用とバックアップ用のディスク装置は、全く同じ性能・機能を持ったもので構成する必要があった。

また、ハードウェア番地を使用しているため、ファイル毎にミラー化する、しないの設定を行えなかった。

- 10 更に、過去に誤りがあったときに、その過去の時点まで戻り、その時点から正しくデータを更新することができないという欠点があった。

- 従来の第7のデータバックアップ方式は、正データが更新された時にバックアップを取得する方式である。この第7のデータバックアップ方式は、最初に正データの全コピーを取得し、それ以降は、基本的にはAログを取得する方式である。この
15 第7のデータバックアップ方式でも、BログやTログを取得する方式もある。正データへの更新が次々と行われる場合に、Aログを単純に取得して保存すると、データ更新処理が進むに従ってAログ量が増加するため、正データが破壊された場合に
20 元に戻す作業が、非常に長時間掛かることになる。これを避ける為に、定期的に、最初に取得したコピーとAログをマージし、その時点で正データの全コピーを取得したのと同様の効果を持たせている。

しかしながら、原理的には、定期的に全コピーを取得する方式と同一であるため、正データが破壊した場合には、最新の全コピーによってデータを戻した後で、Aログで最新の状態に戻す必要があったため、時間がかかるという欠点があった。

- また、上記の第6を除く、第1～第7の方式に共通するのは、インデックスのバックアップが困難であったり、時間がかかったりすることである。従来のファイル
25 方式ではデータベースと呼ばれる方式がオンラインでは用いられてきた。このイン

デックスは、例えば、数レベルからなりたっていて、多数のインデックスが更新される可能性があるという複雑な形式を持っているため、バックアップ対象としていないファイル方式が多い。例外的にバックアップ対象としているものでは、インデックスへの更新をAログに書き出してリカバリーが可能なようにするか、完全なミラーとするか、であった。

従来の各データバックアップ方式は、ファイル全体のコピー取得に時間がかかり、一旦ファイルが破壊されて、それを修復するためにはコピーに要したと同等以上の時間を要するという欠点があった。これは、従来のバックアップ方式が、基本的に定期的に正データの全コピーを取得すると共に、トランザクション毎にログを書き出し、正データの修復を行う際に、バックアップコピーからの復元によってコピー時の状態に戻した後、ログをファイルに書き込んで最新性を確保する必要がある、という方式を採用していたからである。

また、従来の各データバックアップ方式では、バックアップコピーを記憶する媒体としては磁気テープ装置を使用しているため、装置自体のスピードが遅いことに加えて、コピーされたファイルが順編成でなければならず、必要のないデータも読み込みが必要があるという欠点もあった。これはノンストップ運転が当然視される分野ではもちろんのこと、通常のシステムにおいても大きな問題であった。

また、上記従来のデータバックアップ方式では、正データのコピー取得に要する時間がデータの容量に応じて長くなり、バックアップ取得に必要なコストが大きなものにならざるを得ないという欠点もあった。なお、ここでいうコストには、コピーを取得するための人件費や、バックアップ装置の他に、コピーを格納する記憶媒体の費用や、バックアップデータを記憶した記憶媒体の保管場所に関わる費用のことをいっている。

本発明は、上述した欠点に鑑みなされたものであり、バックアップ及びリカバリーを、短時間でかつ低コストに行なうことができるデータバックアップ・リカバリー方式を提供することを目的としている。

発明の開示

本発明の上記目的は、以下に記載する発明によって達成される。

1. コンピュータにおけるデータバックアップ・リカバリー方式において、一つのユニークなキーとゼロ個または1個以上のノンユニークなキーを持つレコードを
5 順次格納するブロックと、

このブロックの位置管理を、当該ブロックとランダムアクセスメモリの物理アドレスとを対応させたロケーションテーブルを用いて行い、ランダムアクセスメモリに格納されているデータベースを管理するプライマリーシステムと、

- 10 前記プライマリーシステムの正データが格納されているブロックに対応したバックアップブロックを用意し、このブロックの位置管理を、当該ブロックとランダムアクセスメモリの物理アドレスとを対応させたロケーションテーブルを用いて行い、ランダムアクセスメモリに格納されているバックアップデータベースを管理するセカンダリーシステムと

を備えたことを特徴とするデータバックアップ・リカバリー方式。

15

2. 前記プライマリーシステムは、アプリケーション処理を行うプライマリ処理装置のメインメモリを前記ランダムアクセスメモリとして使用し、前記ランダムアクセスメモリ内のデータベースの内容を変更するデータベース制御機構と、前記データベース制御機構が前記データベースの内容を変更したときに当該変更内容のデータを送出するプライマリバックアップリカバリ制御機構とを備え、
20

- 前記セカンダリーシステムは、セカンダリ処理装置のメインメモリをランダムアクセスメモリとして使用し、前記プライマリバックアップリカバリ制御機構から送られてくる当該データで前記ランダムアクセスメモリ内のバックアップデータベースを変更するセカンダリバックアップリカバリ制御機構を備えたことを特徴とする前
25 記1記載のデータバックアップ・リカバリー方式。

3. 前記プライマリーシステムは、アプリケーション処理を行うプライマリ処理装置と、このプライマリ処理装置のメインメモリとは別にデータベースを格納するランダムアクセスメモリからなるプライマリ記憶装置とを設け、前記セカンダリシステムは、各種の処理を実行するセカンダリ処理装置と、このセカンダリ処理装置のメインメモリとは別にデータベースを格納するランダムアクセスメモリからなるセカンダリ記憶装置とを設けたことを特徴とする前記1記載のデータバックアップ・リカバリー方式。

4. 前記プライマリーシステムは、アプリケーション処理を行うプライマリ処理装置と、このプライマリ処理装置のメインメモリとは別にデータベースを格納するランダムアクセスメモリからなるプライマリ記憶装置とを設け、

前記セカンダリシステムは、データベースを格納するランダムアクセスメモリからなるセカンダリ記憶装置のみを設け、

前記プライマリ記憶装置は、バックアップデータの通信を行う手段と、前記データベースの内容を変更するデータベース制御機構と、前記データベース制御機構が前記データベースの内容を変更したときに当該変更内容のデータを前記通信手段を介して送出するプライマリバックアップリカバリ制御機構とを備え、

前記セカンダリ記憶装置は、バックアップデータの通信を行う手段と、前記プライマリバックアップリカバリ制御機構から前記通信手段を介して送られてくる当該データで前記バックアップデータベースを変更するセカンダリバックアップリカバリ制御機構とを備え、

たことを特徴とする前記1記載のデータバックアップ・リカバリー方式。

5. 前記プライマリーシステムは、トランザクション処理が開始されたときにトランザクション開始情報を送信し、更新後のデータの内容と更新内容、データが格納されているブロックを特定する情報を前記セカンダリシステムに送信し、

前記セカンダリシステムは、更新後データ情報を受信するたびに、当該トランザクションの更新後データ情報に基づき該当データの更新を行ない、

前記プライマリーシステムは、トランザクションのデータ更新が終了したときに、更新終了情報を前記セカンダリシステムに送信するようにした同期密結合方式を採用したものであることを特徴とする前記1記載のデータバックアップ・リカバリー方式。

6. 前記プライマリーシステムは、更新後のデータの内容と更新内容、データが格納されているブロックを特定する情報をセカンダリシステムに送信し、

前記セカンダリシステムは、前記プライマリーシステムからトランザクション開始の情報を受信後、トランザクション内のログデータを受信して当該データの更新をおこない、前記プライマリーシステムからトランザクション終了情報を受け取った後、当該バックアップ更新処理が終了するまでバックアップ終了情報を前記プライマリーシステムに送信しないようにした非同期疎結合方式を採用したものであることを特徴とする前記1記載のデータバックアップ・リカバリー方式。

7. 前記プライマリーシステムは、トランザクション処理が開始されたときにトランザクション開始情報を送信し、更新後のデータの内容と更新内容、データが格納されているブロックを特定する情報を前記セカンダリシステムに送信し、前記セカンダリシステムは、更新後データ情報を受信するたびに、当該トランザクションの更新後データ情報に基づき該当データの更新を行ない、前記プライマリーシステムは、トランザクションのデータ更新が終了したときに、更新終了情報を前記セカンダリシステムに送信するようにした同期密結合方式を採用したものであることを特徴とする前記2記載のデータバックアップ・リカバリー方式。

8. 前記プライマリーシステムは、更新後のデータの内容と更新内容、データが

格納されているブロックを特定する情報をセカンダリシステムに送信し、

前記セカンダリシステムは、前記プライマリシステムからトランザクション開始の情報を受信後、トランザクション内のログデータを受信して当該データの更新をおこない、前記プライマリシステムからトランザクション終了情報を受け取った
5 後、当該バックアップ更新処理が終了するまでバックアップ終了情報を前記プライマリシステムに送信しないようにした非同期疎結合方式を採用したものであることを特徴とする前記 2 記載のデータバックアップ・リカバリー方式。

9. 前記プライマリ処理装置は、バックアップデータの通信を行う通信手段と、
10 前記データベースの内容を変更するデータベース制御機構と、前記データベース制御機構が前記データベースの内容を変更したときに当該変更内容のデータを前記通信手段を介して送出するプライマリバックアップリカバリ制御機構とを備え、

前記セカンダリ処理装置は、バックアップデータの通信を行う通信手段と、前記
プライマリバックアップリカバリ制御機構から前記通信手段を介して送られてくる
15 当該データで前記バックアップデータベースを変更するセカンダリバックアップリカバリ制御機構と、
を備えたことを特徴とする前記 3 記載のデータバックアップ・リカバリー方式

10. 前記プライマリ処理装置及びセカンダリ処理装置は、両者の間でバックアップデータの通信を行う通信手段のみを設けてあり、かつ、
20 前記プライマリ記憶装置は、前記データベースの内容を変更するデータベース制御機構と、前記データベース制御機構が前記データベースの内容を変更したときに当該変更内容のデータを前記通信手段を介して送出するプライマリバックアップリカバリ制御機構とを備え、

25 前記セカンダリ記憶装置は、前記プライマリバックアップリカバリ制御機構から前記通信手段を介して送られてくる当該データで前記バックアップデータベースを

変更するセカンダリバックアップリカバリ制御機構を備え、
たことを特徴とする前記3記載のデータバックアップ・リカバリー方式。

1 1. 前記プライマリーシステムは、トランザクション処理が開始されたときに
5 トランザクション開始情報を送信し、更新後のデータの内容と更新内容、データが
格納されているブロックを特定する情報を前記セカンダリシステムに送信し、

前記セカンダリシステムは、更新後データ情報を受信するたびに、当該トランザ
クションの更新後データ情報に基づき該当データの更新を行ない、

前記プライマリーシステムは、トランザクションのデータ更新が終了したときに、
10 更新終了情報を前記セカンダリシステムに送信するようにした同期密結合方式を採
用したものであることを特徴とする前記3記載のデータバックアップ・リカバリー
方式。

1 2. 前記プライマリーシステムは、更新後のデータの内容と更新内容、データが
15 格納されているブロックを特定する情報をセカンダリシステムに送信し、

前記セカンダリシステムは、前記プライマリーシステムからトランザクション開
始の情報を受信後、トランザクション内のログデータを受信して当該データの更新
をおこない、前記プライマリーシステムからトランザクション終了情報を受け取った
後、当該バックアップ更新処理が終了するまでバックアップ終了情報を前記プライ
20 マリーシステムに送信しないようにした非同期疎結合方式を採用したものであるこ
とを特徴とする前記3記載のデータバックアップ・リカバリー方式。

1 3. 前記プライマリーシステムは、トランザクション処理が開始されたときにト
ランザクション開始情報を送信し、更新後のデータの内容と更新内容、データが格
25 納されているブロックを特定する情報を前記セカンダリシステムに送信し、

前記セカンダリシステムは、更新後データ情報を受信するたびに、当該トランザ

クシヨンの更新後データ情報に基づき該当データの更新を行ない、

前記プライマリーシステムは、トランザクシヨンのデータ更新が終了したときに、更新終了情報を前記セカンダリシステムに送信するようにした同期密結合方式を採用したものであることを特徴とする前記4記載のデータバックアップ・リカバリー方式。

14. 前記プライマリーシステムは、更新後のデータの内容と更新内容、データが格納されているブロックを特定する情報をセカンダリシステムに送信し、

前記セカンダリシステムは、前記プライマリーシステムからトランザクシヨン開始の情報を受信後、トランザクシヨン内のログデータを受信して当該データの更新をおこない、前記プライマリーシステムからトランザクシヨン終了情報を受け取った後、当該バックアップ更新処理が終了するまでバックアップ終了情報を前記プライマリーシステムに送信しないようにした非同期疎結合方式を採用したものであることを特徴とする前記4記載のデータバックアップ・リカバリー方式。

15

15. コンピュータにおけるデータバックアップ・リカバリー方式において、

一つのユニークなキーとゼロ個または1個以上のノンユニークなキーを持つレコードを順次格納するブロックと、このブロックの位置管理を、当該ブロックとランダムアクセスメモリの物理アドレスとを対応させたロケーションテーブルを用いて行い、ランダムアクセスメモリに格納されているデータベースを管理するプライマリーシステムを備えたことを特徴とするデータバックアップ・リカバリー方式。

16. 前記プライマリーシステムは、アプリケーション処理を行うプライマリ処理装置のメインメモリを前記ランダムアクセスメモリとして使用し、前記ランダムアクセスメモリ内のデータベースの内容を変更するデータベース制御機構と、前記データベース制御機構が前記データベースの内容を変更したときに当該変更内容の

データを送出するプライマリバックアップリカバリ制御機構とを備えたことを特徴とする前記 1 5 記載のデータバックアップ・リカバリー方式。

1 7. 前記プライマリーシステムは、アプリケーション処理を行うプライマリ処理装置と、このプライマリ処理装置のメインメモリとは別にデータベースを格納するランダムアクセスメモリからなるプライマリ記憶装置とを設けたことを特徴とする前記 1 5 記載のデータバックアップ・リカバリー方式。

1 8. 前記プライマリーシステムは、前記同期密結合方式または前記非同期疎結合方式を採用したものであることを特徴とする前記 1 5 記載のデータバックアップ・リカバリー方式。

1 9. 前記プライマリーシステムは、前記同期密結合方式または前記非同期疎結合方式を採用したものであることを特徴とする前記 1 6 記載のデータバックアップ・リカバリー方式。

2 0. 前記プライマリ処理装置は、バックアップデータの通信を行う通信手段と、前記データベースの内容を変更するデータベース制御機構と、前記データベース制御機構が前記データベースの内容を変更したときに当該変更内容のデータを前記通信手段を介して送出的るプライマリバックアップリカバリ制御機構とを備えことを特徴とする前記 1 7 記載のデータバックアップ・リカバリー方式。

2 1. 前記プライマリ処理装置は、バックアップデータの通信を行う通信手段のみを設け、前記プライマリ記憶装置は、前記データベースの内容を変更するデータベース制御機構と、前記データベース制御機構が前記データベースの内容を変更したときに当該変更内容のデータを前記通信手段を介して送出的るプライマリバック

アップリカバリ制御機構とを備えたことを特徴とする前記 17 記載のデータバックアップ・リカバリー方式。

22. 前記プライマリ記憶装置は、バックアップデータの通信を行う通信手段と、
5 前記データベースの内容を変更するデータベース制御機構と、前記データベース制御機構が前記データベースの内容を変更したときに当該変更内容のデータを前記通信手段を介して送出するプライマリバックアップリカバリ制御機構とを備えたことを特徴とする前記 17 記載のデータバックアップ・リカバリー方式。

10 23. 前記プライマリシステムは、前記同期密結合方式または前記非同期疎結合方式を採用したものであることを特徴とする前記 17 記載のデータバックアップ・リカバリー方式。

24. コンピュータにおけるデータバックアップ・リカバリー方式において、
15 バックアップしようとするプライマリシステムの正データが格納されているブロックに対応したバックアップブロックを用意し、このブロックの位置管理を、当該ブロックとランダムアクセスメモリの物理アドレスとを対応させたロケーションテーブルを用いて行い、ランダムアクセスメモリに格納されているバックアップデータベースを管理するセカンダリシステムを備えたことを特徴とするデータバックアップ・リカバリー方式。
20

25. 前記セカンダリシステムは、アプリケーション処理を行うセカンダリ処理装置のメインメモリをランダムアクセスメモリとして使用し、前記バックアップしようとするプライマリシステムから送られてくる当該データで前記ランダムアクセスメモリ内のバックアップデータベースを変更するセカンダリバックアップリカバリ制御機構を備えたことを特徴とする前記 24 記載のデータバックアップ・リカバ

リー方式。

26. 前記セカンダリシステムは、アプリケーション処理を行うセカンダリ処理装置と、このセカンダリ処理装置のメインメモリとは別にデータベースを格納するランダムアクセスメモリからなるセカンダリ記憶装置とを設けたことを特徴とする前記24記載のデータバックアップ・リカバリー方式。

27. 前記セカンダリシステムは、データベースを格納するランダムアクセスメモリからなるセカンダリ記憶装置のみを設け、

- 10 前記セカンダリ記憶装置は、バックアップデータの通信を行う手段と、バックアップしようとするプライマリシステムから前記通信手段を介して送られてくる当該データで前記バックアップデータベースを変更するセカンダリバックアップリカバリ制御機構を備え、

たことを特徴とする前記24記載のデータバックアップ・リカバリー方式。

15

28. 前記セカンダリシステムは、前記同期密結合方式または前記非同期疎結合方式を採用したものであることを特徴とする前記24記載のデータバックアップ・リカバリー方式。

- 20 29. 前記セカンダリシステムは、前記同期密結合方式または前記非同期疎結合方式を採用したものであることを特徴とする前記25記載のデータバックアップ・リカバリー方式。

- 25 30. 前記セカンダリ処理装置は、バックアップデータの通信を行う通信手段と、前記バックアップしようとするプライマリシステムから前記通信手段を介して送られてくる当該データで前記バックアップデータベースを変更するセカンダリバック

アプリケーション制御機構と備えたことを特徴とする前記 2 6 記載のデータバックアップ・リカバリー方式。

5 3 1. 前記セカンダリ処理装置は、バックアップデータの通信を行う通信手段のみを設け、

前記セカンダリ記憶装置は、バックアップしようとするプライマリシステムから前記通信手段を介して送られてくる当該データで前記バックアップデータベースを変更するセカンダリバックアップアプリケーション制御機構とを備えたことを特徴とする前記 2 6 記載のデータバックアップ・リカバリー方式。

10

3 2. 前記セカンダリーシステムは、前記同期密結合方式または前記非同期疎結合方式を採用したものであることを特徴とする前記 2 6 記載のデータバックアップ・リカバリー方式。

15 3 3. 前記セカンダリーシステムは、前記同期密結合方式または前記非同期疎結合方式を採用したものであることを特徴とする前記 2 7 記載のデータバックアップ・リカバリー方式。

20 前記 1 の発明では、セカンダリーシステムを少なくとも 1 個用意し、プライマリシステムの正データと論理的に同じ形式として、プライマリシステムとセカンダリーシステムのブロックを対応させている。また、セカンダリーシステムは、正データが更新された時に更新を行うため、常に最新の正データのコピーデータがバックアップファイルに保持される。セカンダリーシステムは通常一個あれば良いが、必要に応じて複数個用意することも可能である。

25 また、前記 1 の発明において、バックアップ対象とするファイルの形式は、レコードをブロックに格納し、ブロックの位置等の管理をロケーションテーブルと呼ぶ

テーブルで管理するものである。

さらに、前記 1 の発明において、プライマリーシステムでは、ファイルの格納方式として、ロケーションテーブルとブロックを使用しており、セカンダリーシステムでも、バックアップファイルのバックアップブロックをプライマリーシステムで
5 使用しているブロックと対応させている。

図面の簡単な説明

第 1 図は、本発明の第 1 の実施の形態であるデータバックアップ・リカバリー方式を実現するプライマリーシステムとセカンダリーシステムの構成を示したブロッ
10 ク図である。

第 2 図は、本発明の第 1 の実施の形態であるデータバックアップ・リカバリー方式を実現するプライマリーシステムとセカンダリーシステムとをさらに詳細に説明したブロック図である。

第 3 図は、本発明の第 1 の実施の形態であるデータバックアップ・リカバリー方式を実現するプライマリーシステムとセカンダリーシステムで使用されるロケーシ
15 ョンテーブルとデータベースブロックとを示す説明図である。

第 4 図は本発明の第 1 の実施の形態であるデータバックアップ・リカバリー方式を実現するプライマリーシステムで使用されるブロックの構成を説明するための図
である。

第 5 図は、本発明の第 1 の実施の形態であるデータバックアップ・リカバリー方式を実現するデータバックアップ・リカバリー方式において、一つのトランザクシ
20 ョンが開始されて終了されるまでの一連の動作を説明するためのチャートである。

第 6 図は、本発明の第 1 の実施の形態であるデータバックアップ・リカバリー方式を実現するデータバックアップ・リカバリー方式において、同期密結合方式の動
25 作を説明するためのチャートである。

第 7 図は、本発明の第 1 の実施の形態であるデータバックアップ・リカバリー方

式を実現するデータバックアップ・リカバリー方式において、非同期疎結合方式の動作を説明するためのチャートである。

第8図は、本発明の第2の実施の形態に係るデータバックアップ・リカバリー方式を実現するプライマリーシステムとセカンダリーシステムの他の構成を示したブロック図である。

第9図は、本発明の第3の実施の形態に係るデータバックアップ・リカバリー方式を実現するプライマリーシステムとセカンダリーシステムの他の構成を示したブロック図である。

10 発明を実施するための最良の形態

以下、本発明の実施の形態に係るデータバックアップ・リカバリー方式について図面を参照して説明する。なお、本発明らは、データ格納方式（特開平11-31096号公報）を提案しており、このデータ格納方式をデータバックアップ・リカバリーする方式として以下説明してゆくので、説明の都合上、このデータ格納方式も盛り込んで説明することにする。

本明細書において、「テーブル」、「ファイル」、「データベース」という用語が用いられているが、これらは以下のように定義される。

「テーブル」と「ファイル」とは、それぞれ同義である。両者の関係は論理的に見た場合には、「テーブル」であり、物理的に見た場合には「ファイル」である。「データベース」とはファイルの集合体である。レコードを格納するブロックの集合体が一つのファイルであるが、この他にロケーションテーブル（ファイル）と代替キーテーブル（ファイル）を組み合わせたものが、一つのデータベースを構成する。

第1図は、本発明の実施の形態であるデータバックアップ・リカバリー方式を実現するプライマリーシステムとセカンダリーシステムの構成を示したブロック図である。

本発明は、通常使用されているコンピュータのメインメモリの所定の領域をバッ

クアップ用記憶装置として使用することができる他に、当該バックアップ用記憶装置のみを従来のハードディスク装置の代替として使用することもできる。第1図は通常使用されているコンピュータのメインメモリの所定の領域（一部）をバックアップ用記憶装置として使用する場合を示している。

- 5 この第1図において、符号1はコンピュータで構成したプライマリーシステムであり、符号2は同コンピュータで構成したセカンダリーシステムである。ここで、プライマリーシステム1とはデータの更新を最初に行う記憶装置（メモリ）の組のことである。セカンダリーシステム2とはバックアップを取得するシステムのことである。プライマリーシステム1もセカンダリーシステム2も、記憶装置（メモリ）
- 10 の障害の被害を最小限に止めるために、単一の記憶装置ではなく、複数の記憶装置（メモリ）に分散して格納することが好ましい。また、このようにデータを格納することで障害対応のために用意する記憶装置（メモリ）の容量を少なくすることが可能となる。

- 上記プライマリーシステム1とセカンダリーシステム2とは通信ネットワーク3
- 15 を介して接続されており、両者の間でデータのやりとりができるようになっている。

- プライマリーシステム1は、CPU11と、メインメモリとを使用するとともに当該メインメモリの所定の領域をバックアップ用記憶装置としても使用するランダムアクセスメモリ12と、通信制御装置13と、プライマリバックアップリカバリ
- 20 制御機構14と、データベース制御機構15と、入出力端末通信制御機構16と、アプリケーションプログラム17と、その他の手段（図示せず）とからなる。また、ランダムアクセスメモリ12には、データベース18を記憶するデータベース領域12aが設けられている。ここで、ランダムアクセスメモリ12は、例えば半導体記憶装置、その他ランダムにアクセスできるメモリであればどのようなものであ
- 25 てもよい。

また、プライマリーシステム1には、入出力端末通信制御機構16を介して入出

力端末4, …が接続されている。

セカンダリーシステム2は、CPU21と、ランダムアクセスメモリ22と、通信制御装置23と、バックアップリカバリ制御機構24と、その他の手段(図示せず)とからなる。ランダムアクセスメモリ22は、バックアップデータベース25
5 を記憶するバックアップデータベース領域22aと、ログ履歴データ26を記憶するログ履歴記憶領域22bと、他の記憶領域(図示せず)とからなる。ここで、ランダムアクセスメモリ22は、例えば半導体記憶装置、その他ランダムにアクセスできるメモリであればどのようなものであってもよい。

次に、上記プライマリーシステム1とセカンダリーシステム2との信号の流れについて簡単に説明する。入出力端末4に入力されたデータは、入出力端末通信制御機構16に送られる(S1)。入出力端末通信制御機構16では、このデータをアプリケーションプログラム17が受け取る(S2)。すると、アプリケーションプログラム17は、一連のデータベース操作指示及びその受信したデータ(Tログ)をデータベース制御機構15に送信する(S3)。同時に、入出力端末通信制御機構16からプライマリバックアップリカバリ制御機構14に対してもデータを送信する(S3)。
15

ここで、プライマリーシステム1において、ファイルの数は必要に応じて作成することができ制限はない。主にオンライン処理を念頭においた説明を行うが、バッチ処理であってもトランザクションを切って行う処理であれば、同一のロジックで
20 実行できる。データベース制御機構15は、対象データベース18を更新処理する(S4)。また、プライマリバックアップリカバリ制御機構14は、これらの指示に基づいて、通信制御装置13、通信ネットワーク3を介してセカンダリーシステム2へログデータを送信する(S5)。

セカンダリーシステム2では、バックアップリカバリ制御機構24により、受信した各種のログデータをログ履歴データ26として記憶領域に記憶させる(S6)。
25 また、バックアップリカバリ制御機構24により、ログデータを用いてバックアッ

データベース 25 を更新処理する (S 7)。また、セカンダリーシステム 2 では、セカンダリバックアップリカバリ制御機構 24 により、バックアップ処理完了情報を通信制御装置 23、通信ネットワーク 3 を介してプライマリーシステム 1 に通知する (S 8)。これにより、プライマリーシステム 1 では、プライマリバックアップリカバリ制御機構 14 によって当該アプリケーションプログラム 17 の排他レコードを排他解除する (S 9)。上述したようにプライマリーシステム 1 とセカンダリーシステム 2 とは動作をし、プライマリーシステム 1 のデータベース 18 が更新された時にセカンダリーシステム 2 のバックアップデータベース 25 が更新されることになる。

10 上述したデータバックアップ・リカバリー方式を実現するための装置は、本番用に 1 組の記憶装置 (メモリ) を、バックアップ用に 1 組以上の記憶装置 (メモリ) を用意し、それらの装置のアクセススピードは本番用の装置と同等で独立したものとしている。記憶装置 (メモリ) は各々が処理装置に接続されており、処理装置の指令で読み出し、書き込み、更新、削除を行う。処理装置は記憶装置毎に独立して
15 いることが好ましい。1 組の記憶装置 (メモリ) とは以下のような概念である。ファイル全体の容量が大きくて 1 台の記憶装置 (メモリ) に収まらず複数台の記憶装置 (メモリ) に収める場合、または、パフォーマンス上の問題や、故障などの被害を局所化したい場合に、2 台以上の記憶装置 (メモリ) に対してデータを分散させる場合など、複数の記憶装置 (メモリ) で必要とするファイルの記憶を行う場合は、
20 その複数の記憶装置 (メモリ) を 1 組とする。1 組の記憶装置 (メモリ) をプライマリーシステム 1 とし、残りの記憶装置 (メモリ) を 1 組毎にセカンダリーシステム 2 とし、以下第 2 図に示すようなシステムを構築する。

第 2 図は、プライマリーシステムとセカンダリーシステムとをさらに詳細に説明したブロック図である。この図では、プライマリーシステム 1 は、CPU 11 と、
25 ランダムアクセスメモリ 12 と、通信制御装置 13 とがバスライン 19 で接続された構成で示されている。また、ランダムアクセスメモリ 12 は、データベース領域

1 2 a と、プログラム読込領域 1 2 b と、から構成されている。このデータベース領域 1 2 a には、データベース 1 8 が記憶されている。データベース 1 8 には、プライマリロケーションテーブル 5 と、プライマリーブロック 6 a, 6 b, 6 c, …とが記憶されている。

- 5 セカンダリシステム 2 は、CPU 2 1 と、ランダムアクセスメモリ 2 2 と、通信制御装置 2 3 とがバスライン 2 9 で接続された構成で示されている。また、ランダムアクセスメモリ 2 2 は、バックアップデータベース領域 2 2 a と、ログ履歴記憶領域 2 2 b と、プログラム読込領域 2 2 c とから構成されている。このバックアップデータベース領域 2 2 a には、セカンダリロケーションテーブル 7 と、バックアップブロック 8 a, 8 b, 8 c, …とが記憶されている。また、ログ履歴記憶領域 2 2 b にはログ履歴データ 2 6 が記憶されている。このプライマリシステム 1 におけるプライマリロケーションテーブル 5 とプライマリーブロック 6 a, 6 b, 6 c, …と、セカンダリシステム 2 におけるセカンダリロケーションテーブル 7 とバックアップブロック 8 a, 8 b, 8 c, …とは対応させている。

- 15 なお、図では、このプライマリシステム 1 におけるプライマリロケーションテーブル 5 とプライマリーブロック 6 a, 6 b, 6 c, …からなる 1 種類のデータベースしか示していないが、通常は、多数のデータベースが存在する。例えば、人事管理データベース、給与データベース、在庫管理データベース、あるいは、顧客管理データベース等である。

- 20 第 3 図は、プライマリシステム 1 とセカンダリシステム 2 で使用されるロケーションテーブルとデータベースブロックとを示す説明図である。プライマリシステム 1 で使用するプライマリロケーションテーブル 5 は、上からブロック "0", ブロック "1", ブロック "2", ブロック "3", …というようにブロック番号 5 1 a, 5 1 b, 5 1 c, …が割り振られており、このブロック番号 5 1 a, 5 1 b, 5 1 c, …に対応してそれぞれランダムアクセスメモリ 1 2 の物理アドレス 5 2 a, 5 2 b, 5 2 c, …が割り振られている。

また、このプライマリーロケーションテーブル5に記載されたブロック番号5 1 a, 5 1 b, 5 1 c, …に相当するプライマリーブロック6 a, 6 b, 6 c, …は、そのテーブル5に記述されている物理アドレス5 2 a, 5 2 b, 5 2 c, …に応じてランダムアクセスメモリ1 2のデータベース記憶領域1 2 a内に配置され記憶されることになる。

なお、図では、プライマリーブロック6 c, 6 eにオーバーフローブロック9 c, 9 eが添付された状態を示している。また、プライマリーロケーションテーブル5とセカンダリーロケーションテーブル7は、物理的なものがすべて同じ内容ではなく、論理的に同じ内容になっていればよい。

10 セカンダリーシステム2で使用するセカンダリーロケーションテーブル7も、上からブロック"0", ブロック"1", ブロック"2", ブロック"3", …というようにブロック番号7 1 a, 7 1 b, 7 1 c, …が割り振られており、このブロック番号7 1 a, 7 1 b, 7 1 c, …対応してそれぞれランダムアクセスメモリ2 2の物理アドレス7 2 a, 7 2 b, 7 2 c, …が割り振られている。

15 また、このセカンダリーロケーションテーブル7に記載されたブロック番号7 1 a, 7 1 b, 7 1 c, …に相当するバックアップブロック8 a, 8 b, 8 c, …は、そのテーブル7に記述されている物理アドレス7 2 a, 7 2 b, 7 2 c, …に応じてランダムアクセスメモリ2 2内に配置され記憶されることになる。

次に、プライマリーシステムのデータベースにデータを格納する動作について第20 4図を参照して説明する。ここで、第4図は、プライマリーシステムで使用されるブロックの構成を説明するための図である。

この第4図において、ブロックは、プライマリブロック6と、オーバーフローブロック9とに分類できる。プライマリブロック6は、ブロック番号6 1と、プライマリキー値6 2と、オーバーフローキー値6 3と、レコード6 4 a, 6 4 b…と、オーバーフローブロックアドレス6 5とを備えている。また、プライマリキー値6 2にはFROM、TOが設けられており、FROMとTOとは当該ブロック6の中のキ

一値の最小値と最大値とを示している。

同様に、オーバーフローキー値 6 3 には FROM、TO が設けられており、FROM と TO とはオーバーフローブロックの中のキー値の最小値と最大値とを示している。なお、キー値は FROM と TO の両方または何れか一方を持ってもよい。さらに、プライマリキー値 6 2 とオーバーフローキー値 6 3 に FROM と TO を持たしたが、双方合わせて 1 組にすることも可能である。

上記オーバーフローブロック 9 は、プライマリブロック 6 にレコード 6 4 が格納できないときに使用されるものである。このオーバーフローブロック 9 は、プライマリブロック 6 の従属ブロックとして管理され、プライマリブロック 6 からポインティングされるのみで、ロケーションテーブル 5 では管理されないようになっている。

オーバーフローブロック 9 は 1 つで足りない場合は、さらに 1 つずつ追加を行う。このような状態で最初のオーバーフローブロック 9 から 2 番目のオーバーフローブロック 9 の位置を指し示すためにオーバーフローブロック・アドレス 9 1 を使用する。オーバーフローブロック 9 は、プライマリブロック 6 と同様に、その内部にレコード 9 0 を保持する。

また、プライマリーシステム 1 で使用するレコード 6 4 は、第 4 図に示すように、データレコード中に一つのユニークなキー（異なるレコードのキー値が重複しないもので、以降、「主キー」と呼ぶ）6 4 1 と、ゼロ個もしくは 1 個以上のノンユニークなキー（異なるレコードのキー値が重複してもかまわないもので、以降、「代替キー」と呼ぶ）6 4 2 と、データ 6 4 3 とを持つ構造のものを使用する。

また、レコード 6 4 がそれぞれ異なることを示すために、符号 a, b, c, … を使用している。ここで、データが入力されると、プライマリーシステム 1 のデータベース制御機構 1 5 は、データレコード中に一つのユニークな主キー 6 4 1 と、代替キー 6 4 2 と、データ 6 4 3 とを持つレコード 6 4 を、固定長のブロック 6 の中に主キー 6 4 1 の順番に並ぶように 1 個以上格納する。そして、レコード 6 4 はデータベース制御機構 1 5 により、まずプライマリブロック 6 に格納される。レコ

ード64 a, 64 b, …の挿入によりプライマリーブロック6中に格納できなくなった場合に、そのプライマリーブロック6に対してオーバーフローブロック9を割り当て、1つのオーバーフローブロック9では格納できない場合はさらに1つオーバーフローブロック9'を割り当て、各々のブロック6, 9, 9'を連携してブロックとしてレコード64 a, 64 b, …を格納する。

レコード64 a, 64 b, …, 64 n, …の追加に対して最終プライマリーブロック6に格納できない場合は、データベース制御機構15は、新たなプライマリーブロック6を割り当ててレコード64 n, …を格納する。ブロック6 a, 6 b, 6 c, …の位置管理は、ロケーションテーブル5を用いることにより、各々のブロック6 a, 6 b, 6 c, …の物理的な位置に関しては何らの制限を受けず配置でき、また、各々のブロック6 a, 6 b, 6 c, …は予め作成しておく必要が無く、必要に応じて作成すればよく、かつ、物理的なデータ格納エリアが満杯になるまで作成できる。

また、複数の特定のプライマリーキーの後に、レコード挿入が多数行われる型のファイルに対しては、挿入が行われる位置で複数のサブレンジに分割し、挿入ではなくレコード追加とし、オーバーフローレコードの発生を防いでいる。

次に、データバックアップ・リカバリー方式の動作を、第1図ないし第4図を基に、第5図を参照して説明する。

ここで、第5図は、一つのトランザクションが開始されて終了されるまでの一連の動作を説明するためのチャートである。

ここで、「データの更新」とは、データの新規追加、削除も含み、ファイルに変更を与える総ての動作のことをいう。

また、「トランザクション」とは、コンピュータシステム上で関連する一連の動作をいう。例えば、一人の顧客が預金を引き出す際の、一連のプロセスがそれに該当する。この方式を適用する場合は、最初に、一度だけ、正データの全コピーを、正データを格納しているプライマリーシステム1からバックアップデータベース2

5を格納しているセカンダリーシステム2に対して行う。

プライマリーシステム1ではデータはブロック6に格納されており、各々のブロック6は順番にブロック番号61が付いていたが、セカンダリーシステム2でもプライマリーシステム1と1対1に対応するバックアップブロック8a, 8b, 8c, 5
…を用意し、その中にデータを格納する。バックアップブロック8a, 8b, 8c, …を管理する為にセカンダリーロケーションテーブル7を用意する。

最初の準備が完了したら、プライマリーシステム1及び入出力端末4等からなるデータ処理システムの稼動を行う。すると、プライマリーシステム1のデータベース制御機構15はトランザクション開始通知をし(S101)、データT1をデータ転送する(S102)。データ処理システムによってデータ処理がなされて正データに対する更新処理が行われると(S103)、プライマリーシステム1のバックアップリカバリ制御機構14は、更新後のデータ(Aログ(A1))と必要に応じてBログ(B1)、Tログ(図示しない)をセカンダリーシステム2に対して送信する(S104)。また、それらのログをプライマリーシステム上にも保存しておくことは、必要となって使用する際に有用である。15

セカンダリーシステム2のバックアップリカバリ制御機構24は、データT1を受信すると(S201)、ログ履歴データ26内に記憶させる(S202)。ついで、セカンダリーシステム2のバックアップリカバリ制御機構24は、Aログ(A1)を受信すると(S203)、これを格納すべきバックアップデータベース25の内部のバックアップブロック8a, 8b, 8c, …を探す。これは、プライマリーシステム1のブロック6のブロック番号61と同じ番号のバックアップブロック8a, 8b, 8c, …を探せばよい。20

バックアップリカバリ制御機構24は、そのバックアップブロック8a, 8b, 8c, …内の該当するデータを探して、Aログ(A1)で書き換える(S204)。また、バックアップリカバリ制御機構24は、Bログ(B1)をログ履歴データ26に格納する(S205)。これにより、常にプライマリーシステム1とセカンダ25

リーシステム 2 のデータが一致した状態に保たれ、プライマリー・システムの障害に対処する事が可能となる。

再び、データ処理システムによってデータ処理がなされて正データに対する追加の発生処理が行われると (S 1 0 5)、プライマリーシステム 1 のバックアップリ
5 カバリ制御機構 1 4 は、追加のデータ (A ログ (A2)) をセカンダリーシステム 2 に対して送信する (S 1 0 6)。

セカンダリーシステム 2 のバックアップリカバリ制御機構 2 4 は、A ログ (A2) を受信すると (S 2 0 6)、プライマリーシステム 1 のブロック 6 のブロック番号
6 1 と同じ番号のバックアップブロック 8 a, 8 b, 8 c, …を探して、当該バック
10 アップブロック 8 に格納する (S 2 0 7)。

また、バックアップリカバリ制御機構 2 4 は、A ログ (A2) をログ履歴データ
2 6 に格納する (S 2 0 8)。これにより、常にプライマリーシステム 1 とセカン
ダリーシステム 2 のデータが一致した状態に保たれ、プライマリー・システムの障
害に対処する事が可能となる。

15 さらに、データ処理システムによってデータ処理がなされて正データに対する削除の処理が行われると (S 1 0 7)、プライマリーシステム 1 のバックアップリカ
バリ制御機構 1 4 は、削除データ (A ログ (A3)、B ログ (B3)) をセカンダリ
ーシステム 2 に対して送信する (S 1 0 8)。

セカンダリーシステム 2 のバックアップリカバリ制御機構 2 4 は、A ログ (A3、
20 B ログ (B3)) を受信すると (S 2 0 9)、プライマリーシステム 1 のブロック 6
のブロック番号 6 1 と同じ番号のバックアップブロック 8 a, 8 b, 8 c, …を探
して、当該バックアップブロック 8 内のデータを削除する (S 2 1 0)。

また、バックアップリカバリ制御機構 2 4 は、削除データ (A ログ (A3)、B
ログ (B3)) をログ履歴データ 2 6 に格納する (S 2 1 1)。これにより、常にプ
25 ライマリーシステム 1 とセカンダリーシステム 2 のデータが一致した状態に保たれ、
プライマリー・システム 1 の障害に対処する事が可能となる。このように動作し、

一連の動作が終了すると、プライマリーシステム1のバックアップリカバリ制御機構14は、トランザクション終了通知をセカンダリーシステム2に送出する（S109）。

5 セカンダリーシステム2はトランザクション終了通知を受信すると（S212）、バックアップリカバリ制御機構24は、当該トランザクションに係わるデータの更新を全て完了し、その直後にバックアップ終了情報をプライマリーシステム1に送出する（S213）。

すると、プライマリーシステム1のバックアップリカバリ制御機構14では、排他解除処理を実行する（S110）。

10 なお、Bログ、Tログについて説明する。まず、Bログを上述したように時系列的にログ履歴データ26に収集しておくことにより、プライマリーシステム1を一定時間分、溯った状態に戻すことが可能となる。これは、Bログを最新のものから、順次、時系列と逆に掛けて行くことにより、任意の時点のファイル状態に戻すことになるからである。これにより、プログラム異常時にしか使用しなかった更新前情報
15 を有効に使用できることとなる。

また、Tログを時系列的にログ履歴データ26に収集しておくことにより、プログラムエラーで更新誤りが発生した場合に、Bログで過去の時点に戻し、正しいプログラムをTログに基づいて実行することにより、データ内容を正しく復元することができることになる。また、Bログ、Tログの保持期間に関しては、必要に応じて個別に決めればよい。
20

この方法では常にファイルのバックアップがセカンダリーシステム2に保持されており、別の媒体にコピーする必要性はないが、ある時点でのファイルのコピーを取得しておくことは、本発明の趣旨に反するものではない。インデックスに関しては、プライマリーシステム1において採用している格納方式の構造がシンプルであるためと、データから簡単に短時間で再生が行えるため、バックアップを取ってお
25 かななくてもよいが、バックアップを取得する場合でも、データと同様の形式で簡単

に行える。

セカンダリーシステム2におけるバックアップデータベース25にバックアップデータを格納方法の詳細について以下に説明する。バックアップデータを非圧縮で格納させる方法の場合にはそのまま書き込めば良いので、以下の説明では圧縮する

5 場合を基本にして述べる。

プライマリーシステム1では、データは圧縮されず固定長のプライマリブロック6に格納され、必要に応じてオーバーフローブロック9に格納されることになる(第3図、第4図(c)参照)。また、さらに必要ならオーバーフローブロック9'が作成される(第3図、第4図(c)参照)。

10 セカンダリーシステム2では、ロケーションテーブル7を、プライマリーシステム1で用意したエントリーの分だけ用意する。セカンダリーシステム2のレコードは圧縮して格納することにより必要な記憶装置容量を小さくすることが可能である。第2図及び第3図に示すように、バックアップブロック8a, 8b, 8c, …のブロック長が変化する。このため、セカンダリーシステム2では、バックアップブロッ
15 ック8a, 8b, 8c, …のブロック長は可変長とし、プライマリーシステム1のプライマリブロック6にオーバーフロー・ブロック9がある場合には、オーバーフローブロック9を含めて一つのバックアップブロック8c, 8e, …と見なす。

しかしながら、セカンダリーシステム2で圧縮を行わない場合には、当該セカンダリーシステム2でも、プライマリーシステム1と同様に、プライマリブロック
20 6とオーバーフローブロック9という形式を採用してもよい。

バックアップブロック8は、セカンダリーロケーションテーブル7により管理する。特定のバックアップブロック8a, 8b, 8c, …を探すには、セカンダリーシステム2のバックアップリカバリ制御機構24からバックアップブロック8a, 8b, 8c, …のアドレスを探してアクセスを行う。

25 プライマリーシステム1での代替キー・テーブルは、復元が容易であるため、バックアップの必要は無いが、リカバリーを高速に行いたい場合には、セカンダリー

システム 2 に保持する。

代替キー・テーブルのバックアップ方法を以下に述べる。代替キー・テーブルは、データ格納方式（特開平 1 1 - 3 1 0 9 6 号公報）で説明してあるが、第 1 0 図の形式で格納されている。セカンダリーシステムにプライマリーシステムと全く同一のサイズ・形式でセカンダリー・代替キー・テーブルを用意する。これは、ブロックの集合からなるファイルであり、各ブロックにはブロック番号が付してある。プライマリー・代替キー・テーブルの内容が更新された場合には、プライマリー・ブロックが更新された場合の処理と同様に、まず、更新前の情報である B ログを取得し、その後更新内容である A ログを、セカンダリー・システムに送信する。セカンダリー・システムでは A ログの情報に基づいて、セカンダリー・代替キー・テーブルの更新を行う。代替キー・テーブルにはブロック番号が付してあるので、A ログにはブロック番号を含むことが好ましい。セカンダリー・システムでは、ブロック番号を元に、対象セカンダリー・代替キー・テーブルを検出し、該当ブロックに対して更新を行う。

15 また、プレ代替キー・テーブルを使用する場合は、セカンダリー・システム上にも、プライマリー・システムと同一形式でプレ代替キー・テーブルを作成する。

データの更新を行う場合を述べる。オンラインによるデータ処理は常にプライマリーシステム 1 で行い、データの更新はプライマリーシステム 1 上でダイレクトに実行される。プライマリーシステム 1 とセカンダリーシステム 2 との間のデータの通信のシーケンスの方式には、「同期密結合方式」と「非同期疎結合方式」とがある。

「同期密結合方式」は、セカンダリーシステム 2 によるバックアップを、プライマリーシステム 1 での更新と同期させて行うもので、セカンダリーシステム 2 がプライマリーシステム 1 の近くに高速な伝送手段によって結合されている場合を想定した方式である。

「非同期疎結合方式」は、災害対処を主たる目的として、セカンダリーシステム

2をプライマリーシステム1から離れた場所に設置し、通信回線を用いて接続する方式を想定している。

(同期密結合方式の説明)

- 5 まず、同期密結合方式について、第1図ないし第4図を基に、第5図及び第6図を参照して説明する。ここで、第6図は、同期密結合方式の動作を説明するためのチャートである。

10 トランザクション処理がプライマリーシステム1で開始されるときに（第5図及び第6図のS101）、プライマリーシステム1からセカンダリーシステム2に対して、トランザクション開始情報を送信する（第5図及び第6図のS102）。この送信する情報は、トランザクションを特定する情報を含んだものとする。この情報は、トランザクション番号を送信するものとする。

15 トランザクションが開始された後、データに対して更新処理（更新、追加、削除を含む）が行われる場合は、データの更新はプライマリーシステム1のランダムアクセスメモリ12の対象データベース18に対して直接的に行われる。更新後のデータの内容（Aログ）と更新内容（更新、追加、削除の区別）、ファイル識別の他、データが格納されているブロック番号、ブロック内レコード先頭アドレスを付加することが好ましい。このデータを送信することにより、セカンダリーシステム2に書き込む時の位置検出が早くなる。

- 20 ここで、「ファイル識別」とは、1つのシステム内に複数のファイルが存在する場合、どのファイルが更新されたかが分からないと、処理が不可能になってしまうが、これを識別するためのものである。

25 プライマリーシステム1のバックアップリカバリ制御機構14は、これらの情報にトランザクション番号と送信時刻等のデータの順序が峻別できる情報を付けて更新後データ情報として、セカンダリーシステム2に送信する（第5図及び第6図のS104、S106、S108）。

また、トランザクション・キャンセルに備えて、プライマリーシステム1では、正データの更新前に、更新前のデータ内容（Bログ）を取得しておく。Bログは、必要に応じて、セカンダリーシステム2に送信しても良いし、しなくても良い。セカンダリーシステム2に送信するメリットは、セカンダリーシステム2でもBログ
5 を保持するために、ログ消失の危険性が少なくなることと、トランザクション・キャンセルのあった場合には、セカンダリーシステム2での作業が早くなることが挙げられる。

セカンダリーシステム2では、更新後データ情報（Aログ）が到着するたびに、直ちに、当該トランザクションの更新後データ情報に基づき該当データの更新を行
10 う（第5図のS204、S207、S210）。当該「データの更新」とは、当該データに対して更新後データ情報の内容に置き換えて、データの書き換えを行うことをいう。内容は圧縮された状態で格納してもよい。

当該トランザクションのデータの更新が総て終了したら（第5図のS210）、更新終了情報をプライマリーシステム1から、セカンダリーシステム2に送信する
15 （第5図及び第6図のS213）。セカンダリーシステム2が複数ある場合は、すべてのセカンダリーシステム2に対して同様の処理を行う。

次に、トランザクション2の説明をする。再び、トランザクション処理がプライマリーシステム1で開始されるときに（第6図のS121）、プライマリーシステム1からセカンダリーシステム2に対して、トランザクション開始情報を送信する
20 （第6図のS122）。トランザクションが開始された後、データに対して更新処理（更新、追加、削除を含む）が行われる場合は、データの更新はプライマリーシステム1のランダムアクセスメモリ12の対象データベース18に対して直接的に行われる。プライマリーシステム1のバックアップリカバリ制御機構14は、これらの情報にトランザクション番号と送信時刻等のデータの順序が峻別できる情報を
25 付けて更新後データ情報として、セカンダリーシステム2に送信する（第6図のS123、S124、S125、S126）。セカンダリーシステム2では、更新後

データ情報が到着するたびに、直ちに、当該トランザクションの更新後データ情報に基づき該当データの更新を行う（第6図のS 2 2 3、S 2 2 4、S 2 2 5、S 2 2 6）。プライマリーシステム1では、データを更新する場合、該当データに対して排他をかけて、2重更新が発生しないようにしているが、セカンダリーシステム
5 2からのバックアップ終了情報が上がってきた時点で排他を解除する（第6図のS 1 2 8）。

したがって、プライマリーシステム1がトランザクションを終了して（第6図のS 1 2 7）、セカンダリーシステム2からバックアップ終了情報が上がってくる（第6図のS 1 2 8）までの間に、次のトランザクション処理を待つ必要は無く、トランザクションを開始する（第6図のS 1 3 1）。データへの更新が必要になった場合、そのデータに対して排他をかける時点で排他待ちとなるので、セカンダリーシステム2でのバックアップデータ更新が終了する前にプライマリーシステム1で2重更新が行われる可能性は無い。

次に複数のセカンダリーシステム2が存在する場合の排他解除の処理について考
15 える。複数のセカンダリーシステム2が存在するときに、安全性を重視した処理の場合には、全てのセカンダリーシステム2からバックアップ終了情報がプライマリーシステム1に上がってくるのを待って排他解除を行うが、高速性を重視する場合には、複数のセカンダリーシステム2のうち少なくとも1つのセカンダリーシステム2からのバックアップ終了情報をプライマリーシステム1が受信したことをもっ
20 て、排他を解除するようにすることも可能である。

さらに、トランザクション3の説明をする。再び、トランザクション処理がプライマリーシステム1で開始されるときに（第6図のS 1 2 3）、プライマリーシステム1からセカンダリーシステム2に対して、トランザクション開始情報を送信する（第6図のS 1 3 2）。トランザクションが開始された後、データに対して更新
25 処理（更新、追加、削除を含む）が行われる場合は、データの更新はプライマリーシステム1のランダムアクセスメモリ12の対象データベース18に対して直接的

に行われる。プライマリーシステム1のバックアップリカバリ制御機構14は、これらの情報にトランザクション番号と送信時刻等のデータの順序が峻別できる情報を付けて更新後データ情報として、セカンダリーシステム2に送信する（第6図のS133、S134、S135、S136）。セカンダリーシステム2では、更新後データ情報が到着するたびに、直ちに、当該トランザクションの更新後データ情報に基づき該当データの更新を行う（第6図のS233、S234、S235、S236）。

セカンダリシステム2では、データを更新する場合、該当データに対して排他をかけて、2重更新が発生しないようにしているが、セカンダリーシステム2からの更新終了情報が上がってきた時点で排他を解除する（第6図のS138）。

したがって、プライマリーシステム1がトランザクションを終了して（第6図のS127）、セカンダリーシステム2からバックアップ終了情報が上がってくる（第6図のS128）までの間に、次のトランザクション処理を待つ必要は無く、トランザクションを開始する（第6図のS131）。データへの更新が必要になった場合、そのデータに対して排他をかける時点で排他待ちとなるので、セカンダリーシステム2でのバックアップデータ更新が終了する前にプライマリーシステム1で2重更新が行われる可能性は無い。このように次々と処理をしてトランザクション、データ更新等の処理をしてゆくことにより、データバックアップ・リカバリーが可能になる。

上記の場合、ログの取得方法は以下のようなになる。Bログは、更新前情報をそのまま使用する。Bログの必要性は、正データを任意の時間分だけ戻す際に使用する。Tログは入ってきた更新データをそのまま使用する。Tログの必要性は、プログラム・エラーにより処理のやり直しが必要になった場合に使用するものである。Bログ、Tログは、プライマリーシステム1上で保持すれば良いものであるが、これらに対してバックアップを行う場合には、セカンダリーシステム2に送信して保持する。上記では、データ用のブロックのバックアップに対して述べてあるが、ロケ

ーションテーブルと代替キーテーブルのバックアップに対して説明する。

プライマリーシステム1のプライマリーロケーションテーブル5は、データの追加に伴ってプライマリーブロック6が追加されるに従って、変更される。セカンダリーシステム2でデータの追加が行われ、それに伴ってバックアップブロック8の追加が行われると、セカンダリーシステム2でも同様にバックアップブロック8の追加が行われる。その際に、セカンダリーシステム2のセカンダリーロケーションテーブル7も自動的に生成されるので、直接的なバックアップは不要である。代替キーテーブルのバックアップは必須ではないが、高速なリカバリーを実現するためにはバックアップ対象にするのが望ましい。

10 代替キーテーブルは、代替キーブロック中のデータの変更と代替キー・オーバーフロー・ブロックの追加が行われる。この場合、代替キー・テーブルの格納ブロックをプライマリーシステム1からセカンダリーシステム2に対して送信する。セカンダリーシステム2では、該当代替キーブロックをそのまま更新する。該当代替キーブロックの検出は、代替キーテーブルの先頭からの変位で行なうか、または、代替
15 キーブロックに番号を付けて、それを持って検出するのが好ましいが、代替キーを用いて検出することも可能である。

(非同期疎結合方式の説明)

次に非同期疎結合方式について、第1図ないし第3図を基に、第7図を参照して
20 説明する。ここで、第7図は、非同期疎結合方式の動作を説明するためのチャートである。トランザクション処理がプライマリーシステム1で開始される時に、プライマリーシステム1からセカンダリーシステム2に対して、同期密結合方式と同様に、トランザクション開始情報を送信する（第7図のS301）。

これは、トランザクションを特定する情報を含んだものとする。ここでは、トランザクション番号としておく。トランザクションが開始された後、データに対して
25 更新処理（更新、追加、削除を含む）が行われる場合は、データの更新はプライマ

リーシステム1のランダムアクセスメモリ12のデータベース18に対して直接的に行われる。更新後のデータの内容(Aログ)と更新内容(更新、追加、削除の区別)、ファイル識別の他、データが格納されているブロック番号、ブロック内レコード先頭アドレスを付加することが好ましい。これにより、セカンダリーシステム
5 2に書き込む時の位置検出が早くなる。また、「ファイル識別」については既に説明しているが、ここでも説明する。この「ファイル識別」とは、1つのシステム内に複数のファイルが存在する場合、どのファイルが更新されたかが分からないと、処理が不可能になってしまうが、これを識別するためのものである。

プライマリーシステム1は、更新後のデータの内容(Aログ)と更新内容(更新、追加、削除の区別)、ファイル識別の他、データが格納されているブロック番号、
10 ブロック内レコード先頭アドレスを付加した情報をセカンダリーシステム2に送信する(S302~S304)。

プライマリーシステム1は、これらの情報にトランザクション番号と送信時刻等のデータの順序が峻別できる情報を付けて更新後データ情報として、セカンダリー
15 システム2に送信する。また、Tログ、Bログに関しては、同期密結合方式と同様である。データの更新後の内容は、レコード全体を送信することも可能であるが、送信量が大きくなるため、変更のあった部分だけを抜き出して、オフセット値と長さを付けて送信することも可能である。

セカンダリーシステム2は、トランザクション開始の情報を受信し(第7図のS
20 401)、以後、トランザクション内のログデータを受信しセカンダリーシステム2の該当データの更新を行なう(第7図のS402~S404)。プライマリーシステム1からトランザクション終了情報を受け取った(S405)後、セカンダリーシステムでのバックアップ更新処理がすべて終了したらバックアップ終了情報をプライマリー・システムに送信する(第7図のS406)。プライマリーシステム
25 1では、その情報が来たトランザクションのAログを破棄するが(第7図のS306)、必要に応じて保管することも可能である。Aログの情報は、セカンダリーシ

システムのバックアップブロックに直接反映される為、通常は保管する必要はないが、更新差分を保持したい場合に有効である。

本方式の場合、セカンダリーシステム2でのバックアップ処理とプライマリーシステム1での更新処理は非同期に実行されるため、データへの更新内容の反映が、

5 セカンダリーシステム2上で逆転してしまい、最新性の確保が行えなくなる危険性はあるが、これに対しては、以下のような方法で避けることができる。

プライマリーシステム1からセカンダリーシステム2にトランザクション開始や終了、ログなどを送信する場合に、処理の順番が分かるように実行時刻を含んで通知するとともに、通信障害などで途中のログなどが消失してしまわないように、連番

10 を付して送付することにより、一般の通信で行われているように、データの抜けを検出して、抜けている場合には再送信を要求し、連番の順に処理を行なうことで整合性を保つことができる。

密結合同期方式の場合には、通信障害の可能性は少ないが、安全性を高めるために、上述の方式を採用することは意味のあることである。

15 セカンダリーシステム2では、更新後データ情報が到着するたびに、直ちに、当該トランザクションの更新後のデータ情報に基づき、該当データの更新を行う（第7図のS402～S404）。当該データの更新とは、当該データに対して更新後データ情報の内容に置き換えて、データの書き換えを行うことをいう。内容は圧縮された状態で格納される。当該トランザクションのデータの更新が総て終了したら、

20 次のトランザクションの処理を行う。以後、上記処理を繰り返すことにより、データバックアップ・リカバリーを行なうことができる。

次に、トランザクションが異常終了した場合やキャンセルされた場合について説明する。デッドロック発生によるトランザクション・キャンセルも同様であるので、一緒に説明する。この場合には該当トランザクションで更新されたデータを元の状態に復元した上で、次のトランザクションの処理を行う必要がある。トランザクシ

25 ョンの異常終了やキャンセルが発生した場合は、プライマリーシステム1からセカ

ンダリーシステム2にトランザクション・キャンセル情報を送信する。この情報には、「どのトランザクションが異常終了したのか」または「キャンセルされたのか」という情報の他に、当該トランザクションで更新されたすべてのデータの更新前情報（Bログ）を含む。トランザクション・キャンセル情報を受信したセカンダリー

5 システム2は、更新前情報に基づき、当該データすべてを更新前の状態に修復する。これは、プライマリーシステム1で実行される内容と同じである。すべてのセカンダリーシステム2は、修復の終了をプライマリーシステム1に送信する。

ファイル障害への対応は、従来方式では、装置毎とかファイル毎の修復が必要であったが、本方式ではブロック番号による管理を行っているため、多段階での対応

10 が可能である。障害が局所的であれば、ブロックをいくつか修復する方式が採用できる。また、従来と同様に、ファイル毎とか装置毎、または、プライマリーシステム1全体などの多段階の修復が可能である。また、従来方式では、順編成ファイルに格納されているため、目的の部分だけを抜き出すためには、最初から磁気テープを読まなくてはならず、修復のための時間がかかる要因になっていたが、本方式で

15 は、セカンダリーロケーションテーブル7によって直接的に目的のバックアップブロックから読み取ることが可能である。また、プライマリーシステム1のブロック6とセカンダリーシステム2のバックアップブロック8は対応しており、簡単に目的のブロックが探し出せる。

次に、セカンダリーシステム2の障害の場合の対処について説明する。プライマ

20 リーシステム1は複数の記憶装置（メモリ）に分散する事が好ましい。この場合、障害は、プライマリーシステム1の一部分に発生することが殆どである。この場合の復旧方法は以下の通りである。

装置が故障した場合に備えて、予め予備の装置を用意しておき、これを使用する。予備の装置は、プライマリーシステム1で、分割されている記憶装置（メモリ）と

25 同等以上の容量・性能を持つものとする。障害が発生したプライマリーシステム1の記憶装置（メモリ）を切り離し、予備の装置をその記憶装置（メモリ）の代わり

に充当する。障害が発生した装置に格納されていたデータに該当するデータをセカンダリーシステム2からコピーをする。この作業を行っている時間は、他の一切の処理は停止しておく。そして、コピーが終了したら処理を再開する。この方法では、一定の時間、処理の中断が発生するが、それ以降の処理は通常の実行速度で行えるし、

5 バックアップの取得も確実に実行できる。

上記方式は、安全性を優先した方法である。これに対して、切り換え時間を短縮する方式がある。この方式として採用できるものは、ホットスワップである。このホットスワップとは、セカンダリーシステム2のランダムアクセスメモリ22をそのままプライマリーシステム1の記憶装置（メモリ）として使用することである。

10 この場合、セカンダリーシステム2が2組以上あれば、バックアップの問題は発生しないが、セカンダリーシステム2が1組しかない場合には、バックアップシステムが一時的になくなることになるので、運用上で注意が必要である。

障害発生時には、セカンダリーシステム2を最新の状態にした後に、セカンダリーシステム2のランダムアクセスメモリ22をプライマリーシステム1の記憶装置

15 （メモリ）として使用する。プライマリーシステム1で障害のあったランダムアクセスメモリ12に関しては、障害のあった装置を切り離し、障害用の予備装置を充当する。これをセカンダリーシステム2の記憶装置（メモリ）として使用する。プライマリーシステム1では、その後、正常運転を続行する。また、平行して、新たなセカンダリーシステム2の予備装置のデータ復元を実行する。これは、予備装置

20 のデータを順にプライマリーシステム1から読み込んで、セカンダリーシステム2に送り、セカンダリーシステム2で書込を行なう。しかしながら、その作業を行っている間にも、プライマリーシステム1のデータ更新が続行されるため、セカンダリーシステム2に書き込んだデータが古くなってしまいう問題があるが、セカンダリーシステム2の予備装置を充当した記憶装置（メモリ）の先頭からそれまで書き込んだ

25 ンドブロックまでに関して、更新が行われれば、通常の実行と同様に、プライマリーシステム1からセカンダリーシステム2へ送信を行って、ブロック8の

更新を行なえばよいので、問題なく同期させることが可能である。ロケーション・テーブルや代替キー・テーブルに関しては、セカンダリーシステム 2 にバックアップを保持しておき、障害時にコピーして戻す方法の他、データを元に再作成する方法がある。

- 5 非同期疎結合方式の場合は、プライマリーシステム 1 とセカンダリーシステム 2 のデータの内容の同期がとれず、データの内容がずれた状態になる可能性がある。そのような状態で、プライマリーシステム 1 の破壊が発生すると、セカンダリーシステム 2 のデータをプライマリーシステム 1 にコピーすると、データが古い状態になってしまい、用をなさないことになってしまう。セカンダリーシステム 2 で更新
- 10 が終了していないデータが存在する場合、セカンダリーシステム 2 にはトランザクション開始・終了情報と共に、T ログ、A ログが送信されている。プライマリーシステム 1 で障害が発生した場合には、プライマリーシステム 1 では直ちにデータ処理を中断することになる。セカンダリーシステム 2 では、その時点で処理中のトランザクションの処理を終了した後も、残っている A ログの処理を行って、セカンダ
- 15 リーシステム 2 のデータ内容をプライマリーシステム 1 と同じにしてから、セカンダリーシステム 2 に必要な部分のデータが圧縮されている場合には、そのデータを伸長して圧縮前の状態にして送信する。

- 上記では、プライマリーシステム 1 の障害の場合を述べたが、セカンダリーシステム 2 でも同様の障害が発生する可能性がある。この場合も、上記とほぼ同様であるが、セカンダリーシステム 2 の障害に備えるために予備の装置を用意する。これは、
- 20 プライマリーシステム 1 の障害用の予備装置と同一のものを共用してもよい。セカンダリーシステム 2 の障害が発生した装置を切り離し、予備の装置をその記憶装置（メモリ）の代わりに充当する。障害が発生した装置に格納されていたデータに該当するデータをプライマリーシステム 1 からコピーを行う。この作業を行って
- 25 いる時間は、他の一切の処理は停止しておく。そして、コピーが終了したら処理を再開する。

上述したセカンダリーシステム2の障害に対処する方法は、安全性を重視した場合の方法であるが、セカンダリーシステム2が2台以上ある場合や、1台の場合でも、処理を重視した場合には、以下の方法が採用できる。

セカンダリーシステム2に障害が発生した場合に、予備の装置を障害が発生した装置に割り当てては、全く同様である。割当が終了したら、プライマリーシステム1では処理を継続する。この間、セカンダリーシステム2が1台の場合には、予備に切り換えた装置に格納されているデータに関しては、バックアップが行われていない状態になる。プライマリーシステム1での処理と併行して、プライマリーシステム1からセカンダリーシステム2の予備装置を充当した分へのデータの転送を行い、バックアップファイルの修復を行う。この際に、順次データ転送を行っていくため、一旦転送が終了したブロックのデータに更新が発生する場合があります。最新性が確保されなくなる恐れがあるが、通常のバックアップを行っている場合と同様に、更新が行われるとそのデータはセカンダリーシステム2に送信されるので、通常の処理と同様に処理することで、最新性の確保は行われる。また、セカンダリーシステム2で、バックアップが未了のブロック中のデータに対して、プライマリーシステム1で更新が行われた時には、そのデータを適用するブロックが存在しないため、Aログ等のログはセカンダリーシステム2において破棄する。

また、このような場合に備えて、セカンダリーシステム2を複数用意した場合には、障害が発生した装置を切り離して、残った装置だけで運転を継続することが可能である。切り離した装置の内容を最新の状態にするには、プライマリーシステム1の運転を停止した時点で行うか、上記方法を使用して行なう。

非同期疎結合方式の場合は、前述の場合と同様に、プライマリーシステム1とセカンダリーシステム2のデータの内容の同期がとれず、データの内容がずれた状態になる可能性がある。セカンダリーシステム2で障害が発生した場合は、プライマリーシステム1から必要なデータをセカンダリーシステム2にコピーし、完了したらシステムの稼働を開始するが、その時点で、セカンダリーシステム2をプライマ

リーシステム1との差がある場合、まず、プライマリーシステム1からセカンダリーシステム2へ必要なブロックのコピーを行う。その後、セカンダリーシステム2からのトランザクション終了情報を用いて、プライマリーシステム1で終了しており、セカンダリーシステム2で終了していないトランザクションを検出し、それに

5 該当するAログを用いてセカンダリーシステム2のデータ内容を更新する。

本発明の実施の形態に係るデータバックアップ・リカバリー方式を用いた場合の副次的な効果として、以下の点が挙げられる。

本方式では、常に最新のバックアップがセカンダリーシステム2上に保持される。このことは、セカンダリーシステム2の上のデータを参照系のトランザクション処

10 理に使用することが可能であるということになる。一般的には、更新系と参照系の比率は1対10程度と言われている。このことは、参照系のシステムを切り離すことにより、更新系を扱っているシステムの負荷を削減することが可能になることを意味している。セカンダリーシステム2を参照系のデータとして使用することで、リアルタイムのデータを用いた参照系システムを構築することが遥かに容易になり、

15 セカンダリーシステム2を単なるバックアップ用途以外に使用することになり、投資対効果が大きくなる。

セカンダリーシステム2を参照系システムとして使用する場合には、データの圧縮をすると、その処理に時間がかかることになり好ましくない。また、代替キー・テーブルのバックアップは必須となる。

20 このほか、本実施の形態に係るデータバックアップ・リカバリー方式よれば、次のような利点がある。

(1) バックアップに関する時間と手間を大幅に削減できる。

(2) バックアップデータを用いたリカバリーに要する時間と手間を同様に大幅に削減できる。

25 (3) バックアップと、バックアップデータを用いたリカバリーとを確実にこなうことができる。

(4) 記憶資源を少なくでき、簡単な操作でリカバリーができる。

(5) プログラムの誤った更新を行った場合でもデータ修復することも可能である。

[第2の実施の形態]

5 第8図は、本発明の第2の実施の形態に係るデータバックアップ・リカバリー方式を実現するプライマリシステムとセカンダリシステムの他の構成を示したブロック図である。

この第8図に示す第2の実施の形態におけるデータバックアップ・リカバリー方式も、プライマリシステム1aと、セカンダリシステム2aとから構成されている。

10 さらに詳細に説明すると、前記プライマリシステム1aは、アプリケーション処理を行うプライマリ処理装置110に対し、ランダムアクセスメモリを内蔵するプライマリ記憶装置120を別に設けている。また、前記セカンダリシステム2aは、各種の処理を実行するセカンダリ処理装置210に対し、ランダムアクセスメモリを内蔵するセカンダリ記憶装置220を別に設けている。

15 ここで、前記プライマリ処理装置110は、図示しないが、バックアップデータの通信を行う通信手段と、前記プライマリ記憶装置内のデータベースの内容を変更するデータベース制御機構と、前記データベース制御機構が前記データベースの内容を変更したときに当該変更内容のデータを前記通信手段を介して送出するプライマリバックアップリカバリ制御機構とを備えている。ここで、第2の実施の形態に
20 における通信手段は、第1の実施の形態の通信制御装置13と同一あるいは同一の機能を有している。また、第2の実施の形態におけるデータベース制御機構は、第1の実施の形態のデータベース制御機構15と同一または同一機能を有している。さらに、第2の実施の形態におけるプライマリバックアップリカバリ制御機構は、第1の実施の形態のプライマリバックアップリカバリ制御機構14と同一または同
25 一機能を有している。

なお、第2の実施の形態のプライマリ記憶装置120におけるデータベース領域

では、第1の実施の形態と同様に、ロケーションテーブル5と、ブロック6、…によりバックアップ動作が行われている。

前記セカンダリ処理装置210は、バックアップデータの通信を行う通信手段と、前記プライマリバックアップリカバリ制御機構から前記通信手段を介して送られてくる当該データで前記セカンダリ記憶装置内のバックアップデータベースを変更するセカンダリバックアップリカバリ制御機構とを備えている。ここで、第2の実施の形態における通信手段は、第1の実施の形態の通信制御装置23と同一あるいは同一の機能を有している。ここで、第2の実施の形態におけるセカンダリバックアップリカバリ制御機構は、第1の実施の形態のセカンダリバックアップリカバリ制御機構24と同一または同一機能を有している。

なお、第2の実施の形態のセカンダリ記憶装置220におけるデータベース領域では、第1の実施の形態と同様に、ロケーションテーブル7と、ブロック8、…によりバックアップ動作が行われている。

このような第2の実施の形態によれば、上記第1の実施の形態における同期密結合方式や非同期疎結合方式によるバックアップリカバリ動作を実現できる。

また、第2の実施の形態によれば、上記第1の実施の形態の利点に加えて、記憶装置の増設、変更が容易に行なうことができる。

[第2の実施の形態の変形例]

第2の実施の形態の変形例に係るデータバックアップ・リカバリー方式は、第9図と同じ構成図で表すことができるが、各ブロックで処理する内容が第2の実施の形態とは異なる。この第2の実施の形態の変形例に係るデータバックアップ・リカバリー方式も、第2の実施の形態と同様に、プライマリシステムと、セカンダリを備えたものである。

また、前記プライマリシステムが、アプリケーション処理を行うプライマリ処理装置に対し、ランダムアクセスメモリを内蔵するプライマリ記憶装置を別に設け

ている点も、また、前記センカダリシステムが、各種の処理を実行するセカンダリ処理装置に対し、ランダムアクセスメモリを内蔵するセカンダリ記憶装置を別に設けている点も、上記第2の実施の形態と外形的に同じである。

さらに、第2の実施の形態に変形例が、前記プライマリ記憶装置1bに通信手段
5 を設け、かつ、セカンダリ記憶装置に通信手段を設け、両者の間で両通信手段によりバックアップデータの通信を行うようにしている点でも第2の実施の形態と同じである。

これに対して、第2の実施の形態の変形例が第2の実施の形態と異なるところは、
プライマリ記憶装置が当該プライマリ記憶装置内のデータベースの内容を変更する
10 データベース制御機構と、前記データベース制御機構が前記データベースの内容を変更した時に当該変更内容のデータを前記通信手段を介して送出するプライマリバックアップリカバリ制御機構とを備えている点と、前記セカンダリ記憶装置が、前記プライマリバックアップリカバリ制御機構から前記通信手段を介して送られてくる当該データで当該センカダリ記憶装置内のバックアップデータベースを変更する
15 セカンダリリカバリ制御機構を備えている点にある。

なお、第2の実施の形態の変形例のプライマリ記憶装置におけるデータベース領域では、第1の実施の形態と同様に、ロケーションテーブルと、一つ以上のブロックとによりバックアップ動作が行われている。

また、第2の実施の形態の変形例のセカンダリ記憶装置におけるデータベース領域では、第1の実施の形態と同様に、ロケーションテーブルと、一つ以上のブロックとによりバックアップ動作が行われている。

このような第2の実施の形態の変形例によっても、上記第1の実施の形態における同期密結合方式や非同期疎結合方式によるバックアップリカバリ動作を実現できる。

25 また、第2の実施の形態の変形例によれば、上記第1の実施の形態の利点に加えて、記憶装置の増設、変更が容易に行なうことができる他、各処理装置に通信手段

を設ける他にバックアップ機構を持たせることなく、バックアップが可能になる。

- 上記各実施の形態では、基本的に、セカンダリーシステム 2 は 1 組を設置する例で説明したが、バックアップを万全にしたいときには、セカンダリーシステム 2 を複数用意することが望ましい。特に、災害対処を考慮する場合には、セカンダリーシステム 2 を 2 台用意して、1 台のセカンダリーシステム 2 はプライマリーシステム 1 の近くに設置し、もう 1 台のセカンダリーシステム 2 はプライマリーシステム 1 から離れた場所に設置するようにすればよい。

[第 3 の実施の形態]

- 10 第 9 図は、本発明の第 3 の実施の形態に係るデータバックアップ・リカバリー方式を実現するプライマリーシステムとセカンダリーシステムの他の構成を示したブロック図である。

この第 9 図に示す第 3 の実施の形態におけるデータバックアップ・リカバリー方式も、プライマリーシステム 1 b と、セカンダリーシステム 2 b とから構成されている。

- 15 さらに詳細に説明すると、前記プライマリーシステム 1 b は、アプリケーション処理を行うプライマリ処理装置 1 1 0 b に対し、ランダムアクセスメモリを内蔵するプライマリ記憶装置 1 2 0 b を別に設けている。

前記セカンダリーシステム 2 b は、ランダムアクセスメモリを内蔵するセカンダリ記憶装置 2 2 0 b のみ設けている。

- 20 前記プライマリ記憶装置 1 2 0 b に通信手段を設けるとともに、セカンダリ記憶装置 2 2 0 b に通信手段を設け、前記プライマリ記憶装置 1 2 0 b とセカンダリ記憶装置 2 2 0 b との間（両者の間）でバックアップデータの通信を行えるように構成している。

- 25 また、前記プライマリ記憶装置 1 2 0 b は、当該プライマリ記憶装置 1 2 0 b 内のデータベースの内容を変更するデータベース制御機構と、前記データベース制御機構が前記データベースの内容を変更した時に当該変更内容のデータを前記通信手

段を介して送出するプライマリバックアップリカバリ制御機構とを備えている。ここで、第3の実施の形態における通信手段は、第1の実施の形態の通信制御装置13と同一あるいは同一の機能を有している。

また、第3の実施の形態におけるデータベース制御機構は、第1の実施の形態のデータベース制御機構15と同一または同一機能を有している。さらに、第3の実施の形態におけるプライマリバックアップリカバリ制御機構は、第1の実施の形態のプライマリバックアップリカバリ制御機構14と同一または同一機能を有している。

前記セカンダリ記憶装置220bは、前記プライマリバックアップリカバリ制御機構から前記通信手段を介して送られてくる当該データで当該セカンダリ記憶装置220b内のバックアップデータベースを変更するセカンダリバックアップリカバリ制御機構を備えている。

ここで、第3の実施の形態におけるセカンダリバックアップリカバリ制御機構は、第1の実施の形態のセカンダリバックアップリカバリ制御機構24と同一または同一機能を有している。なお、第3の実施の形態におけるプライマリ記憶装置120bにおけるデータベース領域では、第1の実施の形態と同様に、ロケーションテーブル5と、ブロック6、…によりバックアップ動作が行われている。

また、第3の実施の形態におけるセカンダリ記憶装置220bにおけるデータベース領域では、第1の実施の形態と同様に、ロケーションテーブル7と、ブロック8、…によりバックアップ動作が行われている。

このような第3の実施の形態によっても、上記第1の実施の形態における同期密結合方式や非同期疎結合方式によるバックアップリカバリ動作を実現できる。

上記第3の実施の形態によれば、上記第1の実施の形態の利点に加えて、記憶装置の増設、変更が容易に行なうことができる他、各処理装置に通信手段を設ける他にバックアップ機構を持たせることなく、バックアップが可能になる。

また、セカンダリシステムには、セカンダリ記憶装置のみ設けるのみでよく、ハー

ドウェア構成が簡素になる。

本発明を利用するとオンライン処理を止めずに、リスト出力等の参照系バッチ処理の実行が可能になる。従来の方式ではオンラインを停止し、データ更新が行なわれない状態にして月次レポート作成などのバッチ処理を行なっている為オンライン

5 の連続運転が困難であった。本方式では、特定の任意の時点でセカンダリーシステムのバックアップ更新処理を止めることにより実現できる。バックアップ一時停止情報をプライマリーシステムからセカンダリーシステムに送信し、セカンダリーシステムでは以降のログは受けとるがバックアップ更新を行なわない。

また、プライマリーシステムでは、同期密結合方式の場合排他解除をセカンダリー

10 ーシステムからのバックアップ終了情報によって行なっているが、これを非同期疎結合方式と同様に行なう。セカンダリーシステムでは、バックアップ一時停止情報受け取り後、参照系バッチ処理を起動する。参照系バッチ処理が終了すると手動又は自動的にセカンダリーシステムはバックアップ更新処理が保留されたいたログを

15 が保留されていたログが無くなるとバックアップ一時停止解除情報をセカンダリーシステムからプライマリーシステムに送信し、通常のバックアップモードに戻る。この方式の場合には、一時的にプライマリーシステムとセカンダリーシステムが同期しない形となる為安全性を重視する場合はセカンダリーシステムを複数用意することが好ましい。

20 なお、上記各々の実施の形態においては、前記プライマリバックアップリカバリ制御機構は、前記データベース制御機構が前記データベースの内容を変更した時に当該変更内容のデータを送出するようにしているが、処理の条件によっては、前述したように同時に当該変更内容のデータを送出せず、他のタイミングにおいて当該変更内容のデータを送出するようにしてもよい。

25 また、本発明の上記各実施の形態に係るデータバックアップ・リカバリー方式では、プライマリーシステム1とセカンダリーシステム2の双方を日本国内に設置し

てバックアップをとる構成で説明したが、プライマリーシステム 1 を日本国内に設置しセカンダリーシステム 2 を外国に設置してバックアップをとる構成や、プライマリーシステム 1 を外国に設置しセカンダリーシステム 2 を日本国内に設置してバックアップをとる構成でもよく、これらは、すべて本発明の範囲に属するものである。

5 また、本発明の上記各実施の形態に係るデータバックアップ・リカバリー方式では、プライマリーシステム 1 とセカンダリーシステム 2 の双方を同一人格が設置してバックアップをとる構成で説明したが、プライマリーシステム 1 を設置する者とセカンダリーシステム 2 を設置する者とが別人格であってもよく、上記プライマリーシステム 1 またはセカンダリーシステム 2 を設置すれば本発明の範囲に属するものである。

産業上の利用可能性

以上説明したように本発明に係るデータバックアップ・リカバリー方式によれば、

15 次のような利点がある。

(1) バックアップに関する時間と手間を大幅に削減できる。

(2) バックアップデータを用いたリカバリーに要する時間と手間を同様に大幅に削減できる。

(3) バックアップと、バックアップデータを用いたリカバリーとを確実にこなうことができる。

(4) インデックスのバックアップが確実に簡単に行なえるためリカバリー所用時間が短縮できる。

(5) プライマリーシステムが故障した場合にセカンダリーシステムに切り替えることにより無停止運転が実現できる。

25 (6) バックアップ用の装置はプライマリーシステムと同一の形式の装置である必要はない。

請求の範囲

1. コンピュータにおけるデータバックアップ・リカバリー方式において、一つのユニークなキーとゼロ個または1個以上のノンユニークなキーを持つレコードを
5 順次格納するブロックと、

このブロックの位置管理を、当該ブロックとランダムアクセスメモリの物理アドレスとを対応させたロケーションテーブルを用いて行い、ランダムアクセスメモリに格納されているデータベースを管理するプライマリーシステムと、

10 前記プライマリーシステムの正データが格納されているブロックに対応したバックアップブロックを用意し、このブロックの位置管理を、当該ブロックとランダムアクセスメモリの物理アドレスとを対応させたロケーションテーブルを用いて行い、ランダムアクセスメモリに格納されているバックアップデータベースを管理するセカンダリーシステムと
を備えたことを特徴とするデータバックアップ・リカバリー方式。

15

2. 前記プライマリーシステムは、アプリケーション処理を行うプライマリ処理装置のメインメモリを前記ランダムアクセスメモリとして使用し、前記ランダムアクセスメモリ内のデータベースの内容を変更するデータベース制御機構と、前記データベース制御機構が前記データベースの内容を変更したときに当該変更内容のデータを送出するプライマリバックアップリカバリ制御機構とを備え、
20

前記セカンダリーシステムは、セカンダリ処理装置のメインメモリをランダムアクセスメモリとして使用し、前記プライマリバックアップリカバリ制御機構から送られてくる当該データで前記ランダムアクセスメモリ内のバックアップデータベースを変更するセカンダリバックアップリカバリ制御機構を備えたことを特徴とする請求項1記載のデータバックアップ・リカバリー方式。
25

3. 前記プライマリーシステムは、アプリケーション処理を行うプライマリ処理装置と、このプライマリ処理装置のメインメモリとは別にデータベースを格納するランダムアクセスメモリからなるプライマリ記憶装置とを設け、前記セカンダリシステムは、各種の処理を実行するセカンダリ処理装置と、このセカンダリ処理装置のメインメモリとは別にデータベースを格納するランダムアクセスメモリからなるセカンダリ記憶装置とを設けたことを特徴とする請求項1記載のデータバックアップ・リカバリー方式。

4. 前記プライマリーシステムは、アプリケーション処理を行うプライマリ処理装置と、このプライマリ処理装置のメインメモリとは別にデータベースを格納するランダムアクセスメモリからなるプライマリ記憶装置とを設け、

前記セカンダリシステムは、データベースを格納するランダムアクセスメモリからなるセカンダリ記憶装置のみを設け、

前記プライマリ記憶装置は、バックアップデータの通信を行う手段と、前記データベースの内容を変更するデータベース制御機構と、前記データベース制御機構が前記データベースの内容を変更したときに当該変更内容のデータを前記通信手段を介して送出するプライマリバックアップリカバリ制御機構とを備え、

前記セカンダリ記憶装置は、バックアップデータの通信を行う手段と、前記プライマリバックアップリカバリ制御機構から前記通信手段を介して送られてくる当該データで前記バックアップデータベースを変更するセカンダリバックアップリカバリ制御機構とを備え、

たことを特徴とする請求項1記載のデータバックアップ・リカバリー方式。

5. 前記プライマリーシステムは、トランザクション処理が開始されたときにトランザクション開始情報を送信し、更新後のデータの内容と更新内容、データが格納されているブロックを特定する情報を前記セカンダリシステムに送信し、

前記セカンダリシステムは、更新後データ情報を受信するたびに、当該トランザクションの更新後データ情報に基づき該当データの更新を行ない、

前記プライマリーシステムは、トランザクションのデータ更新が終了したときに、更新終了情報を前記セカンダリシステムに送信するようにした同期密結合方式を採用したものであることを特徴とする請求項1記載のデータバックアップ・リカバリー方式。

6. 前記プライマリーシステムは、更新後のデータの内容と更新内容、データが格納されているブロックを特定する情報をセカンダリシステムに送信し、

10 前記セカンダリシステムは、前記プライマリーシステムからトランザクション開始の情報を受信後、トランザクション内のログデータを受信して当該データの更新をおこない、前記プライマリーシステムからトランザクション終了情報を受け取った後、当該バックアップ更新処理が終了するまでバックアップ終了情報を前記プライマリーシステムに送信しないようにした非同期疎結合方式を採用したものであること

15 とを特徴とする請求項1記載のデータバックアップ・リカバリー方式。

7. 前記プライマリーシステムは、トランザクション処理が開始されたときにトランザクション開始情報を送信し、更新後のデータの内容と更新内容、データが格納されているブロックを特定する情報を前記セカンダリシステムに送信し、

20 前記セカンダリシステムは、更新後データ情報を受信するたびに、当該トランザクションの更新後データ情報に基づき該当データの更新を行ない、

前記プライマリーシステムは、トランザクションのデータ更新が終了したときに、更新終了情報を前記セカンダリシステムに送信するようにした同期密結合方式を採用したものであることを特徴とする請求項2記載のデータバックアップ・リカバリー方式。

25

8. 前記プライマリーシステムは、更新後のデータの内容と更新内容、データが格納されているブロックを特定する情報をセカンダリシステムに送信し、

前記セカンダリシステムは、前記プライマリーシステムからトランザクション開始の情報を受信後、トランザクション内のログデータを受信して当該データの更新

5. をおこない、前記プライマリーシステムからトランザクション終了情報を受け取った後、当該バックアップ更新処理が終了するまでバックアップ終了情報を前記プライマリーシステムに送信しないようにした非同期疎結合方式を採用したものであることを特徴とする請求項2記載のデータバックアップ・リカバリー方式。

10. 9. 前記プライマリ処理装置は、バックアップデータの通信を行う通信手段と、前記データベースの内容を変更するデータベース制御機構と、前記データベース制御機構が前記データベースの内容を変更したときに当該変更内容のデータを前記通信手段を介して送出するプライマリバックアップリカバリ制御機構とを備え、

15. 前記セカンダリ処理装置は、バックアップデータの通信を行う通信手段と、前記プライマリバックアップリカバリ制御機構から前記通信手段を介して送られてくる当該データで前記バックアップデータベースを変更するセカンダリバックアップリカバリ制御機構と、
を備えたことを特徴とする請求項3記載のデータバックアップ・リカバリー方式

20. 10. 前記プライマリ処理装置及びセカンダリ処理装置は、両者の間でバックアップデータの通信を行う通信手段のみを設けてあり、かつ、
前記プライマリ記憶装置は、前記データベースの内容を変更するデータベース制御機構と、前記データベース制御機構が前記データベースの内容を変更したときに当該変更内容のデータを前記通信手段を介して送出するプライマリバックアップリカ
25. バリ制御機構とを備え、

前記セカンダリ記憶装置は、前記プライマリバックアップリカバリ制御機構から

前記通信手段を介して送られてくる当該データで前記バックアップデータベースを変更するセカンダリバックアップリカバリ制御機構を備え、
たことを特徴とする請求項3記載のデータバックアップ・リカバリー方式。

- 5 1 1. 前記プライマリーシステムは、トランザクション処理が開始されたときにトランザクション開始情報を送信し、更新後のデータの内容と更新内容、データが格納されているブロックを特定する情報を前記セカンダリシステムに送信し、

前記セカンダリシステムは、更新後データ情報を受信するたびに、当該トランザクションの更新後データ情報に基づき該当データの更新を行ない、

- 10 前記プライマリーシステムは、トランザクションのデータ更新が終了したときに、更新終了情報を前記セカンダリシステムに送信するようにした同期密結合方式を採用したものであることを特徴とする請求項3記載のデータバックアップ・リカバリー方式。

- 15 1 2. 前記プライマリーシステムは、更新後のデータの内容と更新内容、データが格納されているブロックを特定する情報をセカンダリシステムに送信し、

前記セカンダリシステムは、前記プライマリーシステムからトランザクション開始の情報を受信後、トランザクション内のログデータを受信して当該データの更新をおこない、前記プライマリーシステムからトランザクション終了情報を受け取った
20 後、当該バックアップ更新処理が終了するまでバックアップ終了情報を前記プライマリーシステムに送信しないようにした非同期疎結合方式を採用したものであることを特徴とする請求項3記載のデータバックアップ・リカバリー方式。

- 1 3. 前記プライマリーシステムは、トランザクション処理が開始されたときにト
25 ランザクション開始情報を送信し、更新後のデータの内容と更新内容、データが格納されているブロックを特定する情報を前記セカンダリシステムに送信し、

前記セカンダリシステムは、更新後データ情報を受信するたびに、当該トランザクションの更新後データ情報に基づき該当データの更新を行ない、

前記プライマリシステムは、トランザクションのデータ更新が終了したときに、更新終了情報を前記セカンダリシステムに送信するようにした同期密結合方式を採用したものであることを特徴とする請求項4記載のデータバックアップ・リカバリー方式。

14. 前記プライマリシステムは、更新後のデータの内容と更新内容、データが格納されているブロックを特定する情報をセカンダリシステムに送信し、

10 前記セカンダリシステムは、前記プライマリシステムからトランザクション開始の情報を受信後、トランザクション内のログデータを受信して当該データの更新をおこない、前記プライマリシステムからトランザクション終了情報を受け取った後、当該バックアップ更新処理が終了するまでバックアップ終了情報を前記プライマリシステムに送信しないようにした非同期疎結合方式を採用したものであること
15 を特徴とする請求項4記載のデータバックアップ・リカバリー方式。

15. コンピュータにおけるデータバックアップ・リカバリー方式において、

一つのユニークなキーとゼロ個または1個以上のノンユニークなキーを持つレコードを順次格納するブロックと、このブロックの位置管理を、当該ブロックとランダムアクセスメモリの物理アドレスとを対応させたロケーションテーブルを用いて
20 行い、ランダムアクセスメモリに格納されているデータベースを管理するプライマリシステムを備えたことを特徴とするデータバックアップ・リカバリー方式。

16. 前記プライマリシステムは、アプリケーション処理を行うプライマリ処理装置のメインメモリを前記ランダムアクセスメモリとして使用し、前記ランダム
25 アクセスメモリ内のデータベースの内容を変更するデータベース制御機構と、前記

データベース制御機構が前記データベースの内容を変更したときに当該変更内容のデータを送出するプライマリバックアップリカバリ制御機構とを備えたことを特徴とする請求項15記載のデータバックアップ・リカバリー方式。

5 17. 前記プライマリーシステムは、アプリケーション処理を行うプライマリ処理装置と、このプライマリ処理装置のメインメモリとは別にデータベースを格納するランダムアクセスメモリからなるプライマリ記憶装置とを設けたことを特徴とする請求項15記載のデータバックアップ・リカバリー方式。

10 18. 前記プライマリーシステムは、前記同期密結合方式または前記非同期疎結合方式を採用したものであることを特徴とする請求項15記載のデータバックアップ・リカバリー方式。

15 19. 前記プライマリーシステムは、前記同期密結合方式または前記非同期疎結合方式を採用したものであることを特徴とする請求項16記載のデータバックアップ・リカバリー方式。

20 20. 前記プライマリ処理装置は、バックアップデータの通信を行う通信手段と、前記データベースの内容を変更するデータベース制御機構と、前記データベース制御機構が前記データベースの内容を変更したときに当該変更内容のデータを前記通信手段を介して送出的るプライマリバックアップリカバリ制御機構とを備えことを特徴とする請求項17記載のデータバックアップ・リカバリー方式。

25 21. 前記プライマリ処理装置は、バックアップデータの通信を行う通信手段のみを設け、

前記プライマリ記憶装置は、前記データベースの内容を変更するデータベース制

御機構と、前記データベース制御機構が前記データベースの内容を変更したときに当該変更内容のデータを前記通信手段を介して送出するプライマリバックアップリカバリ制御機構とを備えたことを特徴とする請求項 17 記載のデータバックアップ・リカバリー方式。

5

22. 前記プライマリ記憶装置は、バックアップデータの通信を行う通信手段と、前記データベースの内容を変更するデータベース制御機構と、前記データベース制御機構が前記データベースの内容を変更したときに当該変更内容のデータを前記通信手段を介して送出するプライマリバックアップリカバリ制御機構とを備えたことを特徴とする請求項 17 記載のデータバックアップ・リカバリー方式。

10

23. 前記プライマリーシステムは、前記同期密結合方式または前記非同期疎結合方式を採用したものであることを特徴とする請求項 17 記載のデータバックアップ・リカバリー方式。

15

24. コンピュータにおけるデータバックアップ・リカバリー方式において、バックアップしようとするプライマリシステムの正データが格納されているブロックに対応したバックアップブロックを用意し、このブロックの位置管理を、当該ブロックとランダムアクセスメモリの物理アドレスとを対応させたロケーションテーブルを用いて行い、ランダムアクセスメモリに格納されているバックアップデータベースを管理するセカンダリーシステムを備えたことを特徴とするデータバックアップ・リカバリー方式。

20

25. 前記セカンダリーシステムは、アプリケーション処理を行うセカンダリ処理装置のメインメモリをランダムアクセスメモリとして使用し、前記バックアップしようとするプライマリシステムから送られてくる当該データで前記ランダムアクセ

25

メモリ内のバックアップデータベースを変更するセカンダリバックアップリカバリ制御機構を備えたことを特徴とする請求項 2 4 記載のデータバックアップ・リカバリ方式。

5 2 6. 前記セカンダリシステムは、アプリケーション処理を行うセカンダリ処理装置と、このセカンダリ処理装置のメインメモリとは別にデータベースを格納するランダムアクセスメモリからなるセカンダリ記憶装置とを設けたことを特徴とする請求項 2 4 記載のデータバックアップ・リカバリ方式。

10 2 7. 前記セカンダリシステムは、データベースを格納するランダムアクセスメモリからなるセカンダリ記憶装置のみを設け、

前記セカンダリ記憶装置は、バックアップデータの通信を行う手段と、バックアップしようとするプライマリシステムから前記通信手段を介して送られてくる当該データで前記バックアップデータベースを変更するセカンダリバックアップリカバリ制御機構を備え、

15 たことを特徴とする請求項 2 4 記載のデータバックアップ・リカバリ方式。

2 8. 前記セカンダリシステムは、前記同期密結合方式または前記非同期疎結合方式を採用したものであることを特徴とする請求項 2 4 記載のデータバックアップ・リカバリ方式。

2 9. 前記セカンダリシステムは、前記同期密結合方式または前記非同期疎結合方式を採用したものであることを特徴とする請求項 2 5 記載のデータバックアップ・リカバリ方式。

25

3 0. 前記セカンダリ処理装置は、バックアップデータの通信を行う通信手段

と、前記バックアップしようとするプライマリシステムから前記通信手段を介して送られてくる当該データで前記バックアップデータベースを変更するセカンダリバックアップリカバリ制御機構と備えたことを特徴とする請求項 26 記載のデータバックアップ・リカバリー方式。

5

31. 前記セカンダリ処理装置は、バックアップデータの通信を行う通信手段のみを設け、

前記セカンダリ記憶装置は、バックアップしようとするプライマリシステムから前記通信手段を介して送られてくる当該データで前記バックアップデータベースを変更するセカンダリバックアップリカバリ制御機構とを備えたことを特徴とする請求項 26 記載のデータバックアップ・リカバリー方式。

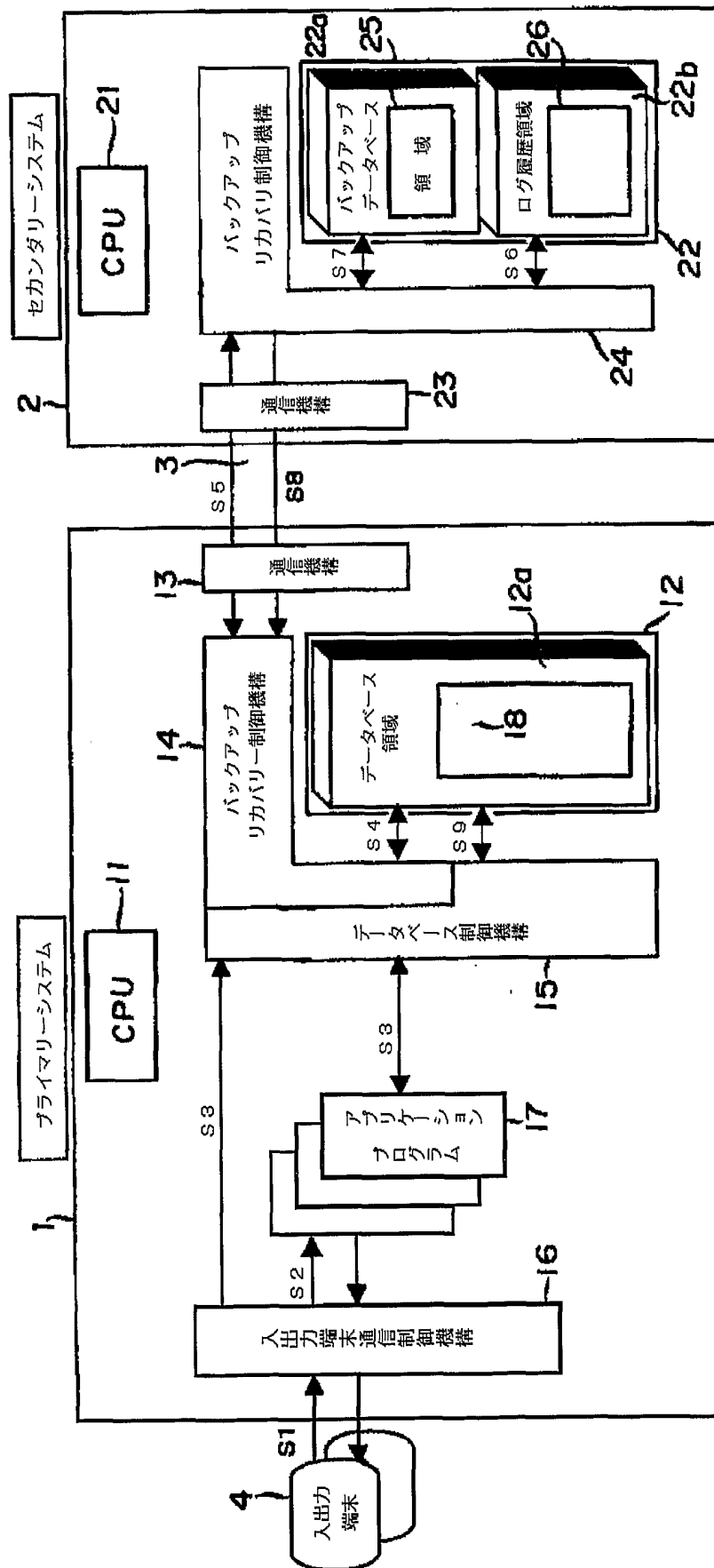
10

32. 前記セカンダリーシステムは、前記同期密結合方式または前記非同期疎結合方式を採用したものであることを特徴とする請求項 26 記載のデータバックアップ・リカバリー方式。

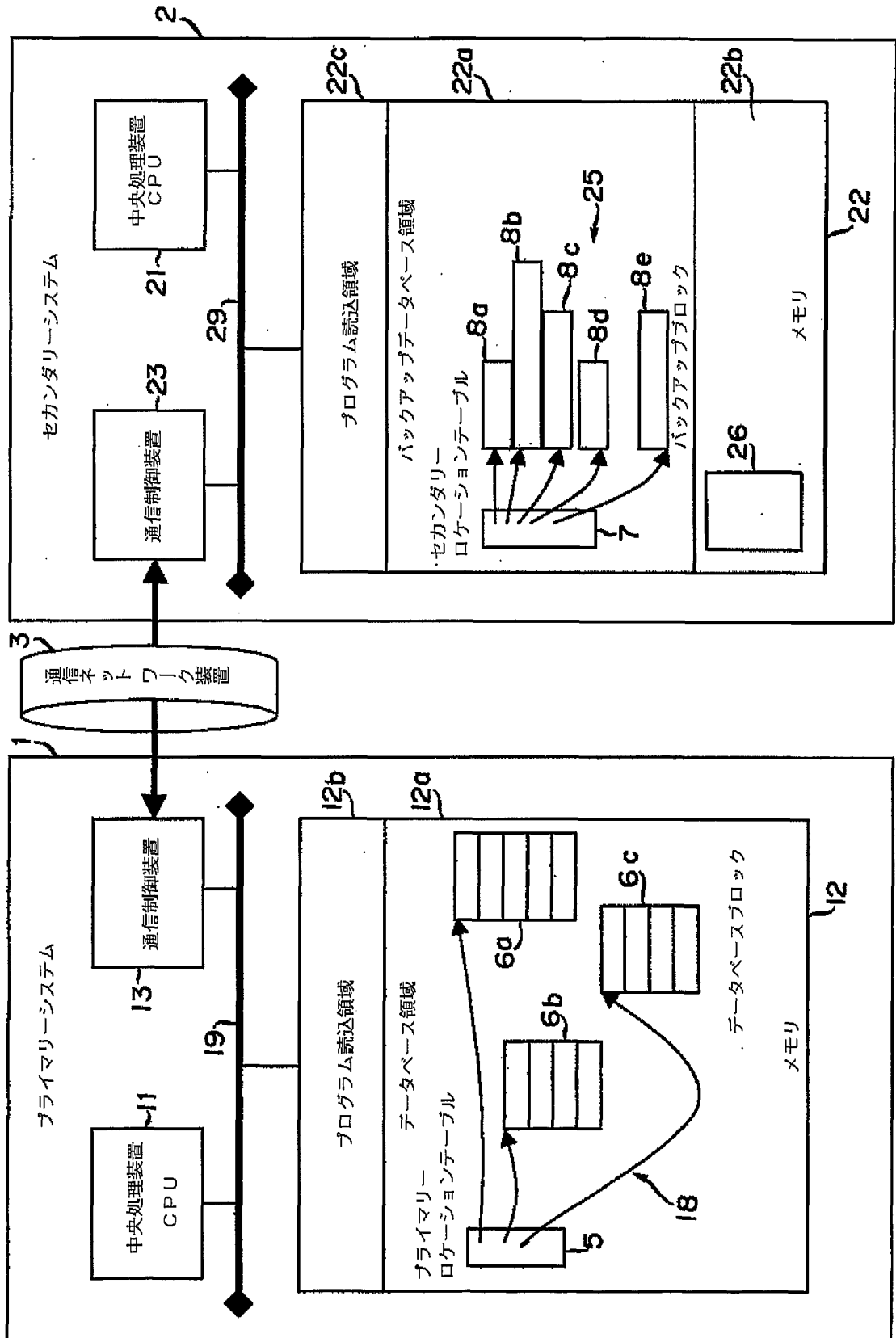
15

33. 前記セカンダリーシステムは、前記同期密結合方式または前記非同期疎結合方式を採用したものであることを特徴とする請求項 27 記載のデータバックアップ・リカバリー方式。

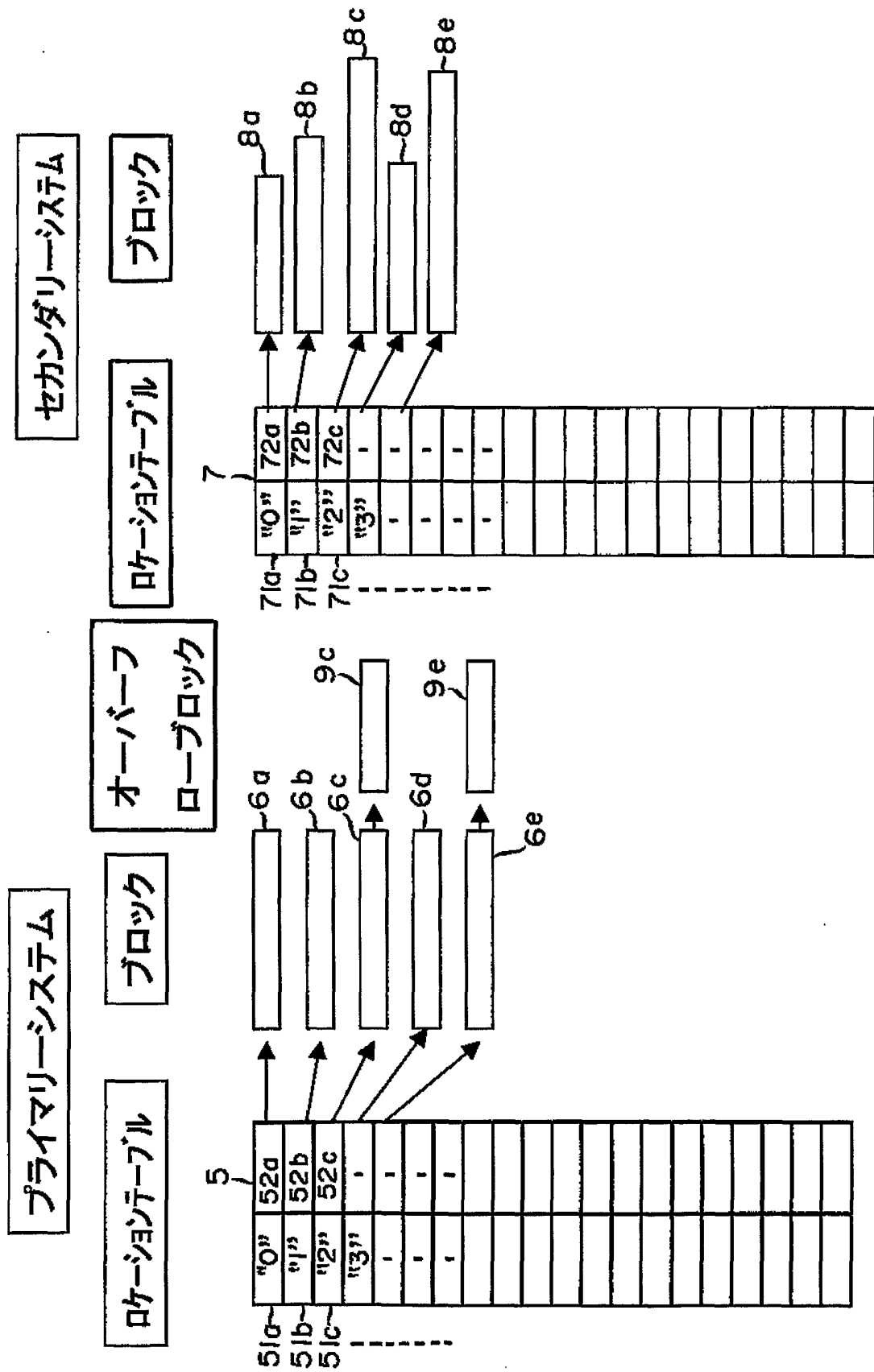
第1図



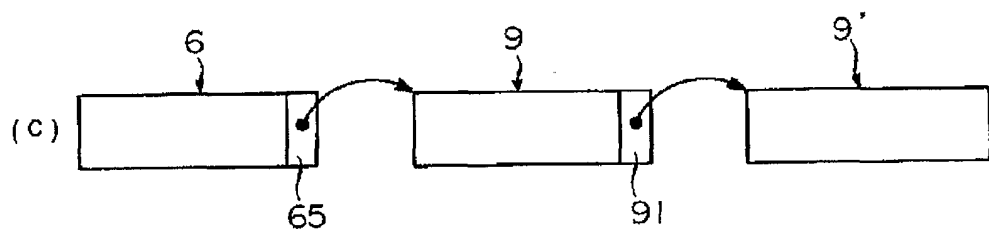
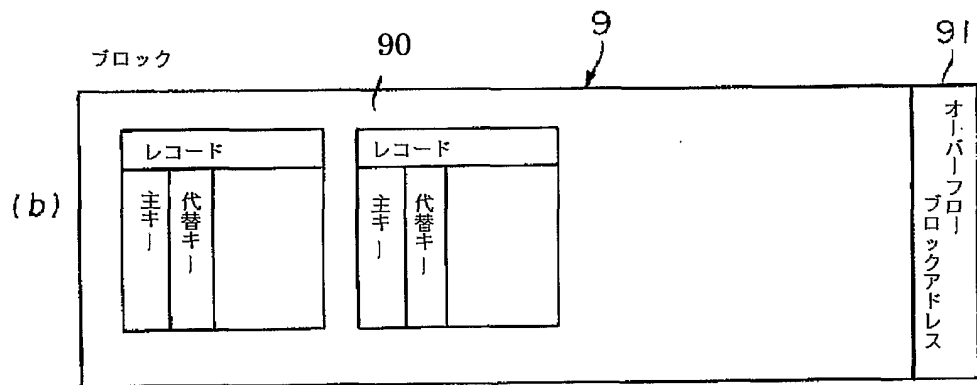
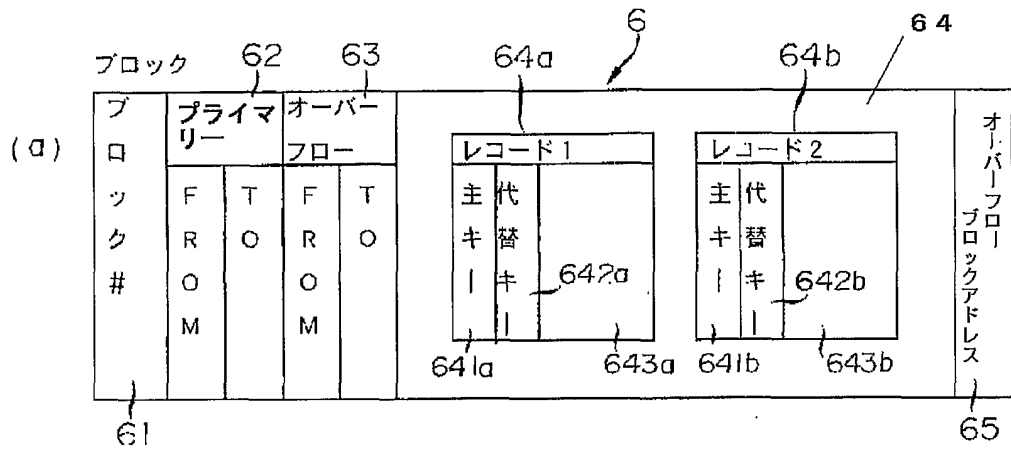
第2図



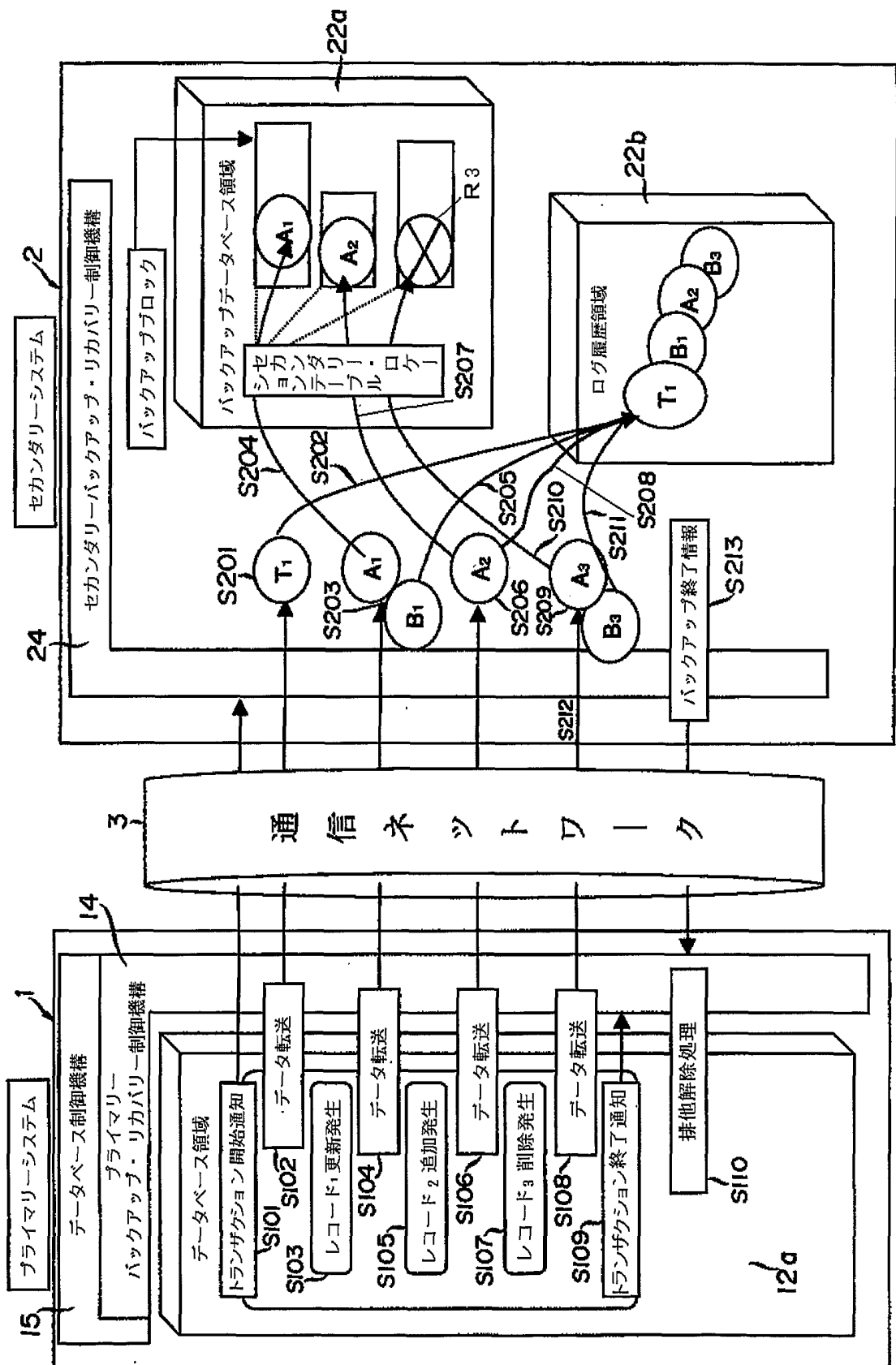
第 3 図



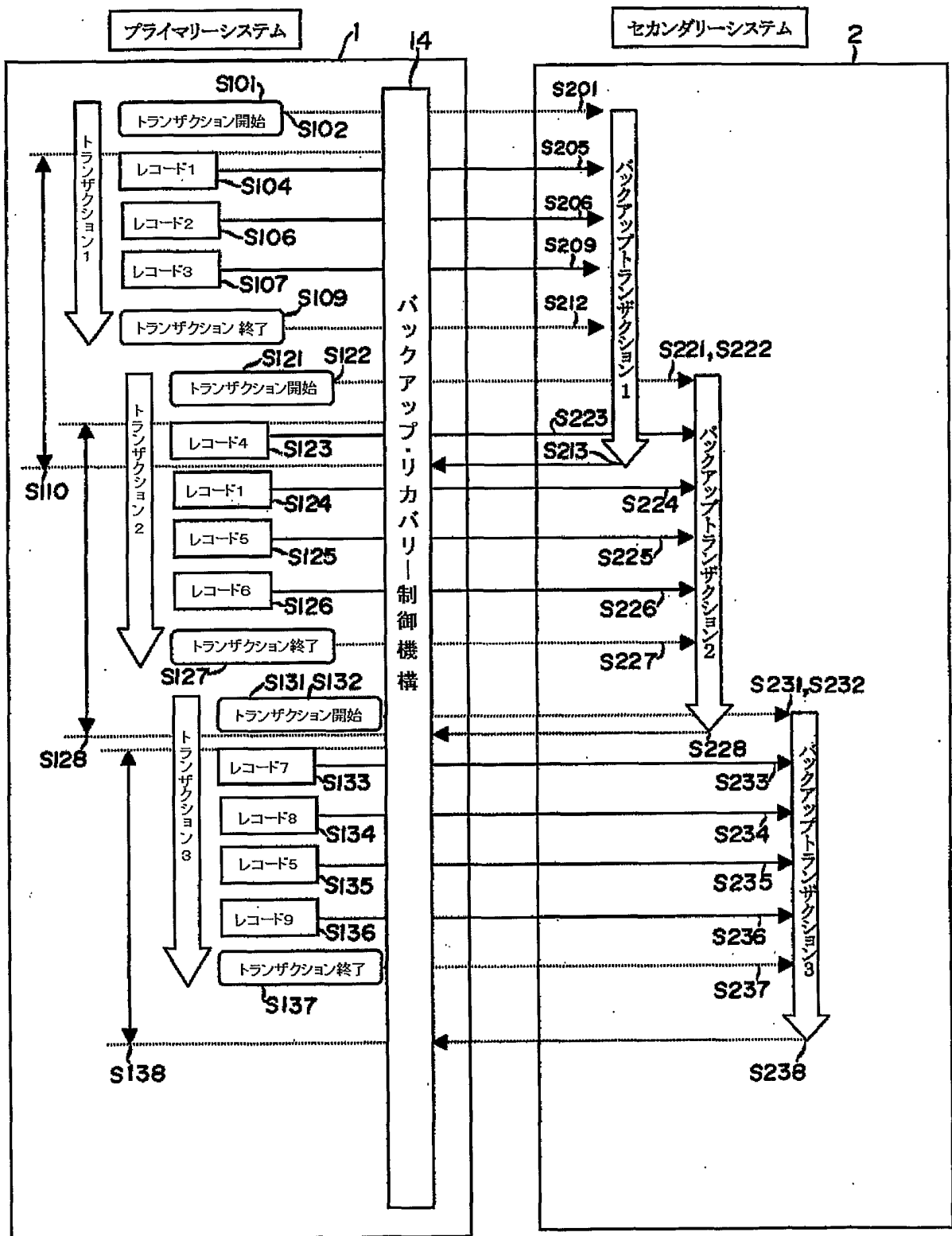
第4図



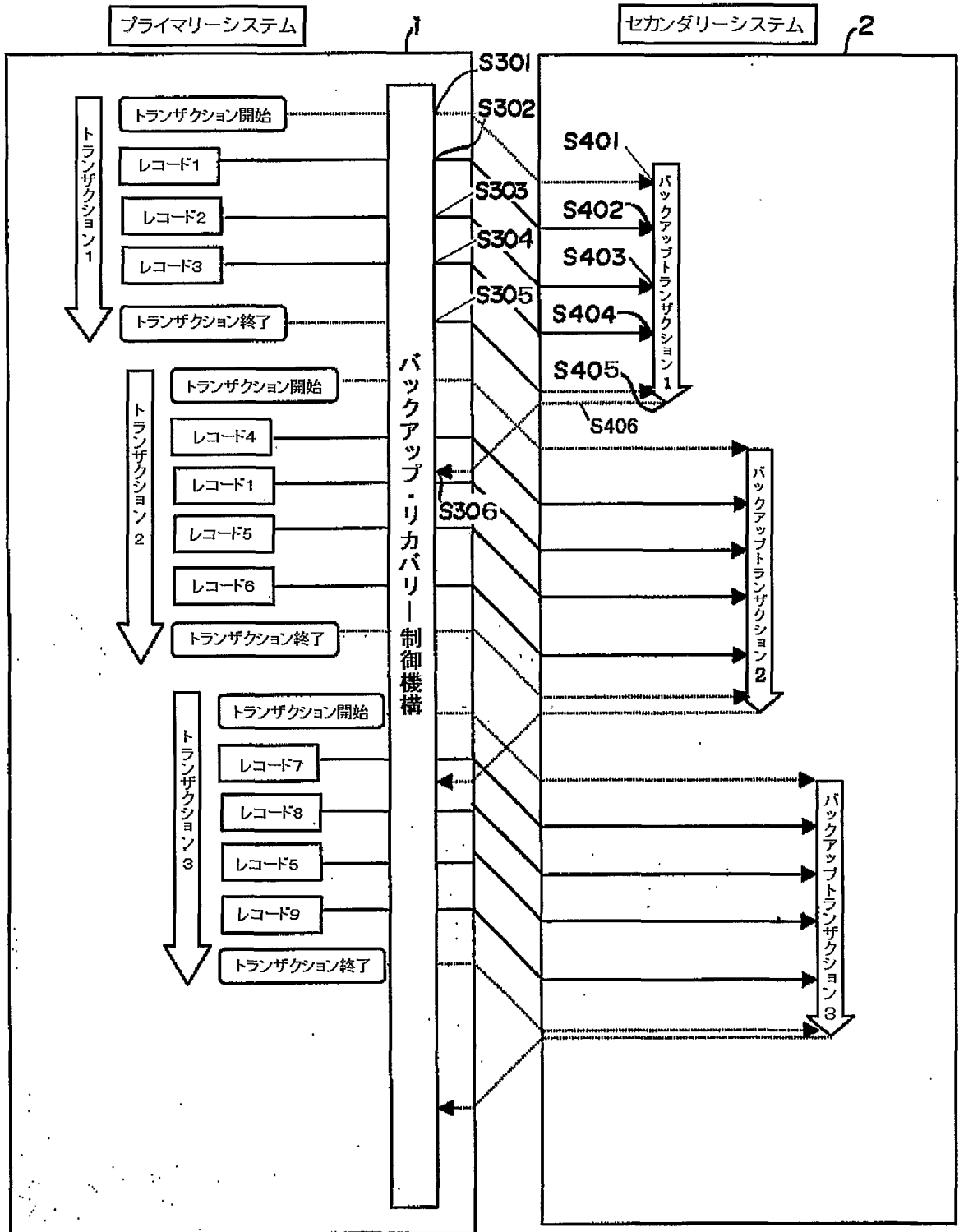
第5図



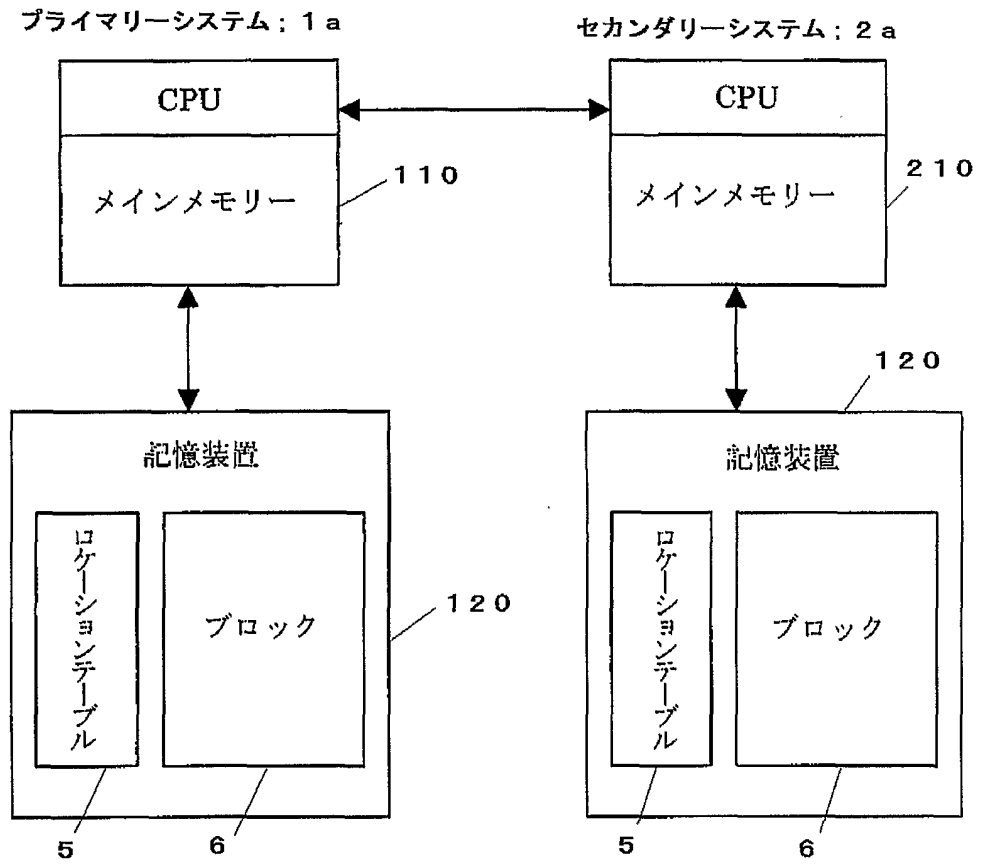
第6図



第7図

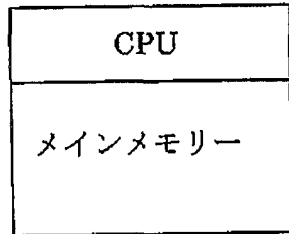


第 8 図

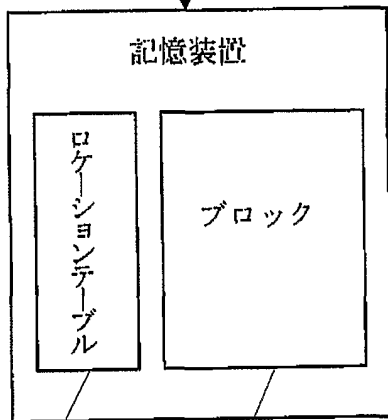


第 9 図

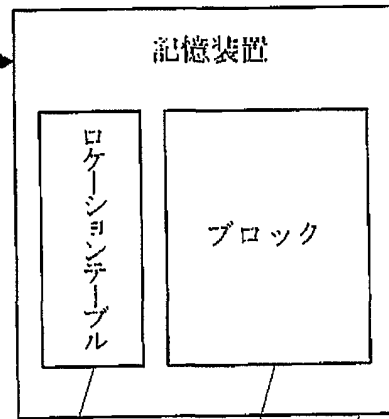
プライマリーシステム：1 b



セカンダリーシステム：2 b



120 b



5

6

5

6

220 b

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP01/03126

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl⁷ G06F12/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl⁷ G06F12/00, G06F11/20

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1926-1996	Toroku Jitsuyo Shinan Koho	1994-2001
Kokai Jitsuyo Shinan Koho	1971-2001	Jitsuyo Shinan Toroku Koho	1996-2001

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	J. Gray et al, translation: Eichi WATANABE, "OLTP System, 1 st edition", McGraw-Hill ed., 25 July, 1991 (25.07.91), pp. 81 - 92	1-33
Y	Masahiro NAKAMURA, "Fukusei File no Hidouki Koushin ga Bunsan Database no Chuushin Gijutsu ni; 2 sou Commit no Daikou Shudan to shite Fujou", Nikkei Electronics, 06 June, 1994 (06.06.94), No.609, pp.101-110	1-33
Y	Mamoru MAEKAWA, et al., "Bunsan Operating System: UNIX no Tsugi ni Kurumono, 1 st edition, Kyouritsu Shuppan, 25 December, 1991 (25.12.91), pp.258-261	1-33

 Further documents are listed in the continuation of Box C. See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art


"&" document member of the same patent family

Date of the actual completion of the international search
10 July, 2001 (10.07.01)Date of mailing of the international search report
07 August, 2001 (07.08.01)Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

A. 発明の属する分野の分類 (国際特許分類 (IPC)) Int. Cl ⁷ G06F12/00		
B. 調査を行った分野 調査を行った最小限資料 (国際特許分類 (IPC)) Int. Cl ⁷ G06F12/00, G06F11/20		
最小限資料以外の資料で調査を行った分野に含まれるもの 日本国実用新案公報 1926-1996 日本国公開実用新案公報 1971-2001 日本国登録実用新案公報 1994-2001 日本国実用新案登録公報 1996-2001		
国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)		
C. 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
Y	J. グレイ、外著、渡辺 榮一訳、OLTPシステム 第1版、マグロウヒル出版、2 5.7月.1991(25.07.91), p. 81 - 92	1 - 33
Y	中村 正弘、複製ファイルの非同期更新が分散データベースの中心技術に 2相コミットの代替手段として浮上、日経エレクトロニクス、6.6月.1994(06.06.94)、 第609号, p. 101 - 110	1 - 33
Y	前川 守、外2名著、分散オペレーティングシステム -UNIXの次にくるもの 第1版、共立出版、25.12月.1991(25.12.91), p. 258 - 261	1 - 33
<input type="checkbox"/> C欄の続きにも文献が列挙されている。		<input type="checkbox"/> パテントファミリーに関する別紙を参照。
* 引用文献のカテゴリー 「A」 特に関連のある文献ではなく、一般的技術水準を示すもの 「E」 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの 「L」 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す) 「O」 口頭による開示、使用、展示等に言及する文献 「P」 国際出願日前で、かつ優先権の主張の基礎となる出願		の日の後に公表された文献 「T」 国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの 「X」 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの 「Y」 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの 「&」 同一パテントファミリー文献
国際調査を完了した日 10.07.01	国際調査報告の発送日 07.08.01	
国際調査機関の名称及びあて先 日本国特許庁 (ISA/JP) 郵便番号100-8915 東京都千代田区霞が関三丁目4番3号	特許庁審査官 (権限のある職員) 原 秀人	5N 9644 
	電話番号 03-3581-1101	内線 3585