

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4390618号  
(P4390618)

(45) 発行日 平成21年12月24日(2009.12.24)

(24) 登録日 平成21年10月16日(2009.10.16)

(51) Int.Cl. F I  
G O 6 F 12/00 (2006.01) G O 6 F 12/00 5 O 1 S

請求項の数 5 (全 21 頁)

(21) 出願番号	特願2004-132350 (P2004-132350)	(73) 特許権者	000005223 富士通株式会社 神奈川県川崎市中原区上小田中4丁目1番1号
(22) 出願日	平成16年4月28日(2004.4.28)	(74) 代理人	100092152 弁理士 服部 毅巖
(65) 公開番号	特開2005-316624 (P2005-316624A)	(72) 発明者	西野 晃英 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内
(43) 公開日	平成17年11月10日(2005.11.10)	(72) 発明者	圓佛 久幸 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内
審査請求日	平成18年9月25日(2006.9.25)	審査官	工藤 嘉晃

最終頁に続く

(54) 【発明の名称】 データベース再編成プログラム、データベース再編成方法、及びデータベース再編成装置

(57) 【特許請求の範囲】

【請求項1】

コンピュータに、複数のページそれぞれに対して格納すべきレコードが決められているデータベースを再編成処理を行わせるデータベース再編成プログラムにおいて、

前記コンピュータを、

前記複写元データベースの再編成要求に応じて、複写元データベース内のページを順次選択し、選択した前記ページ内のレコードを複写先データベースに複写すると共に、前記複写元データベース内のページ毎の複写の進捗状況をページ状態管理テーブルに設定するレコード複写手段、

更新レコードの書き込み要求に応じて前記ページ状態管理テーブルを参照し、前記複写元データベース内の書き込み対象ページが未複写の状態であれば前記複写元データベース内の該当ページに対して前記更新レコードを書き込み、前記書き込み対象ページが複写済の状態であれば前記複写元データベースと前記複写先データベースとに前記更新レコードを書き込むレコード書き込み手段、

として機能させることを特徴とするデータベース再編成プログラム。

【請求項2】

前記コンピュータを前記レコード書き込み手段として機能させる際、前記書き込み対象ページが複写中の状態であれば複写済に遷移するのを待ち、複写中から複写済の状態に遷移したときに、前記複写元データベースと前記複写先データベースとに前記更新レコードを書き込むことを特徴とする請求項1記載のデータベース再編成プログラム。

10

20

## 【請求項 3】

前記複写元データベースは、格納すべきレコードの番号が予め規定されている複数のプライムページと、前記プライムページに格納できなかったレコードが格納される複数のオーバフローページとを有しており、

前記コンピュータを前記レコード複写手段として機能させる際、前記プライムページを順次選択し、選択された前記プライムページと、選択された前記プライムページに格納できなかったレコードが格納された前記オーバフローページとの複写処理を行うことを特徴とする請求項 1 記載のデータベース再編成プログラム。

## 【請求項 4】

複数のページそれぞれに対して格納すべきレコードが決められているデータベースを再編成するデータベース再編成方法において、

レコード複写手段が、前記複写元データベースの再編成要求に応じて、複写元データベース内のページを順次選択し、選択した前記ページ内のレコードを複写先データベースに複写すると共に、前記複写元データベース内のページ毎の複写の進捗状況をページ状態管理テーブルに設定し、

レコード書き込み手段が、更新レコードの書き込み要求に応じて前記ページ状態管理テーブルを参照し、前記複写元データベース内の書き込み対象ページが未複写の状態であれば前記複写元データベース内の該当ページに対して前記更新レコードを書き込み、前記書き込み対象ページが複写済の状態であれば前記複写元データベースと前記複写先データベースとに前記更新レコードを書き込む、

ことを特徴とするデータベース再編成方法。

## 【請求項 5】

複数のページそれぞれに対して格納すべきレコードが決められているデータベースを再編成するデータベース再編成装置において、

前記複写元データベースの再編成要求に応じて、複写元データベース内のページを順次選択し、選択した前記ページ内のレコードを複写先データベースに複写すると共に、前記複写元データベース内のページ毎の複写の進捗状況をページ状態管理テーブルに設定するレコード複写手段と、

更新レコードの書き込み要求に応じて前記ページ状態管理テーブルを参照し、前記複写元データベース内の書き込み対象ページが未複写の状態であれば前記複写元データベース内の該当ページに対して前記更新レコードを書き込み、前記書き込み対象ページが複写済の状態であれば前記複写元データベースと前記複写先データベースとに前記更新レコードを書き込むレコード書き込み手段と、

を有することを特徴とするデータベース再編成装置。

## 【発明の詳細な説明】

## 【技術分野】

## 【0001】

本発明はデータベースの再編成をするためのデータベース再編成プログラム、データベース再編成方法、及びデータベース再編成装置に関し、特にデータベースに対するランザクションを継続したままデータベースの再編成を行うデータベース再編成プログラム、データベース再編成方法、及びデータベース再編成装置に関する。

## 【背景技術】

## 【0002】

多くのデータベースシステムでは、効率的なデータ入出力環境を維持するためデータベースの再編成が必要となる。たとえば、以下のような目的で再編成が行われる。

- ・ページ内のレコード間にある空き領域の統合（フラグメンテーション解消）
- ・オーバフローページにあるレコードをプライムページに格納することによるアクセス性能向上
- ・データベース内においてレコードが満杯状態になる前の領域の拡張

従来一般的な再編成処理では、まず、データベース中にある全てのレコードが中間フ

10

20

30

40

50

ファイルに退避される。そこで、データベースの容量拡張等を実施した後、退避したデータが復元される。

【 0 0 0 3 】

図 1 0 は、従来の一般的なデータベース再構築方法を示す図である。図 1 0 には、データベース 9 1 1 のデータを再構築して、データベース 9 2 1 に格納する場合の例が示されている。

【 0 0 0 4 】

データベース 9 1 1 は、複数のプライムページが格納されたプライムページ格納領域 9 1 1 a と複数のオーバフローページが格納されたオーバフローページ格納領域 9 1 1 b とを有している。プライムページは、レコードの基本的な格納領域である。各レコードを格納すべきプライムページは、インデックス 9 1 2 によって予め定義されている。

10

【 0 0 0 5 】

オーバフローページは、プライムページの空き容量が足りないために格納できなかったレコードを格納するための格納領域である。オーバフローページにレコードが格納された場合、そのレコードが本来格納されるべきプライムページから実際に格納されたオーバフローページに対して、ポインタが設定される。

【 0 0 0 6 】

このデータベース 9 1 1 を再構築する際には、まず、データベース 9 1 1 内のレコードが磁気テープ等の記録媒体 9 3 0 に退避される（ステップ S 1 0 1）。退避の順番は、インデックス情報に従ったプライムページの先頭から昇順で行われる。プライムページからポインタで示されたオーバフローページがある場合は、1 つのプライムページのレコードを退避後、対応するオーバフローページのレコードを順に退避していく。この退避処理が最終ページまで繰り返される。

20

【 0 0 0 7 】

次に、インデックス 9 2 2 が再作成され、容量の拡張されたデータベース 9 2 1 が構築される（ステップ S 1 0 2）。作成されたインデックス 9 2 2 では、全てのレコードがプライムページ格納領域 9 2 1 a 内に納めるように、レコードを格納するページが定義される。データベース 9 1 1 はストレージデバイス上で一旦削除され、その後領域拡張したデータベース 9 2 1 が構築される。そして、記録媒体 9 3 0 に退避したレコードをデータベース 9 2 1 に復元する（ステップ S 1 0 3）。

30

【 0 0 0 8 】

このようにしてデータベースを再構築することができる。新たに生成されたインデックス 9 2 2 の指定に沿ってデータベース 9 2 1 へのレコード格納が行われることにより、たとえば、より多くのプライムページが利用可能となり、アクセス性能が向上する。また、その後、レコード格納があった場合においても、ある程度プライムページに空きができるため、オーバフローページを切り出すことはなく、アクセス性能劣化を防止できる。

【 0 0 0 9 】

ただし、このようなデータベース再構築手法では、レコードを退避させてから復元するまでの間、データベース 9 1 1 に対する更新等の処理を停止する必要がある。これは、データベース再編成時にレコードを格納すると、データ保証ができないためである。そのため、24時間連続して稼働させることが要求されるシステムには適用できない。そこで、データベースへのアクセスを伴う処理機能動作中にデータベース再編成を実行するための方式が必要となる。

40

【 0 0 1 0 】

図 1 1 は、アクセスを停止せずにデータベース再編成を実行する方式を示す図である。この方式では、データベースを使用する処理機能からアクセスしているデータベース 9 4 1（複写元）とは別に、新規に再編成を実施するためのデータベース 9 5 1（複写先）を用意する。各データベース 9 4 1, 9 5 1 には、対応するインデックス 9 4 2, 9 5 2 が設けられている。また、データベース 9 4 1 は、複数のプライムページを格納するプライムページ格納領域 9 4 1 a と、複数のオーバフローページを格納するオーバフローページ

50

格納領域 9 4 1 b とを有している。同様に、データベース 9 5 1 は、複数のプライムページを格納するプライムページ格納領域 9 5 1 a と、複数のオーバフローページを格納するオーバフローページ格納領域 9 5 1 b とを有している。なお、この例では、データベース 9 5 1 は、データベース 9 4 1 よりも記憶容量が拡張されているものとする。

【 0 0 1 1 】

そして、複写元のデータベース 9 4 1 の先頭ページから順に、格納されているレコードが複写先のデータベース 9 5 1 に複写され、データベースの再編成が実施される（ステップ S 1 1 1）。

【 0 0 1 2 】

このとき、複写元のデータベース 9 4 1 に対するレコードの更新を意図するトランザクション 9 6 1 を受け付けることができる。たとえば、トランザクション 9 6 1 によってプライムページ内のレコード（レコード番号「R 1」）が変更され、オーバフローページにレコード（レコード番号「R 3 3」）が追加される（ステップ S 1 1 2）。再編成中にレコードの更新処理が行われると、その処理内容が更新後ログ 9 4 3 に記録される（ステップ S 1 1 3）。

10

【 0 0 1 3 】

ここで、この複写されたレコードに対して応用プログラムからの更新があった場合、複写のレコードに対する更新後ログ 9 4 3 を複写先のデータベース 9 5 1 のレコードに付け込む（ステップ S 1 1 4）。付け込みとは、更新されたレコードの内容をデータベース 9 5 1 に反映させる処理である。付け込みによりレコードの上書きを行うことで、整合性が保証される。このようなデータベースの再編成方式は、DB snapper と呼ばれる（たとえば、特許文献 1 参照）。

20

【 0 0 1 4 】

なお、図 1 1 に示す方式では、更新後ログ 9 4 3 を付け込むまでの間、最新性が保証できない。そのため、再編成後のデータベース 9 5 1 へのアクセスを伴うトランザクション 9 6 2 は、データベース 9 5 1 の参照であってもアクセスが拒否される（ステップ S 1 1 5）。図 1 1 の例では、付け込み処理中に、レコード番号「R 1 1」のレコードの参照を伴うトランザクション 9 6 2 が行われている。このとき、複写先のデータベース 9 5 1 に該当レコードに対する複写が完了していないため、即時参照できない（タイムラグあり）。複写先のデータベース 9 5 1 に更新後ログ 9 4 3 の付け込みが完了すると、それ以降、そのレコードの参照が可能となる。

30

【 0 0 1 5 】

なお、データベース内のデータをブロック単位で複製し、ブロック毎に再編成が完了したか否かを管理することもできる。この場合、再編成が完了したブロックに対するアクセスは、複写先のデータベースに対して行われる。また、再編成が未完了のブロックに対するアクセスは、複写元のデータベースに対して行われる。そして、複写中のブロックに対するアクセスは、複写の完了を待ってから、複写先のデータベースに対して行われる（たとえば、特許文献 2 参照）。

【特許文献 1】特開 2 0 0 0 - 3 3 9 2 1 0 号公報

【特許文献 2】特開平 5 - 1 5 1 0 3 7 号公報

40

【発明の開示】

【発明が解決しようとする課題】

【 0 0 1 6 】

しかし、図 1 0 に示した方式では、ログの付け込みを行う間、データベースにアクセスできない。すなわち、レコード更新してからログ反映するまでの間、そのレコードに対して最新性を保証できない。そのため、応用プログラムからの即時利用ができない。

【 0 0 1 7 】

なお、特許文献 2 記載の方式によれば、レコードへのアクセスを停止せずにデータベースの再編成が可能である。ただし、この方式では、システムに障害等が発生し、データベースの内容が壊れたとき、正しく復旧するのが困難である。すなわち、この方式では、複

50

写されたブロックの内容を更新する場合、複写先のデータベースに対してのみレコードが書き込まれる。したがって、複写先のデータベースに対する更新処理のみが更新後ログとして記録される。

【0018】

ここで、再編成の途中で複写元のデータベースに障害が発生すると、複写元のデータベースの内容が過去のバックアップ時の状態に戻される。その状態のデータベースに複写元のデータベースに対する更新後ログの内容を反映させても、複写先のデータベースに対して行われた更新処理の内容は反映されない。そのため、複写先のデータベースに対して書き込まれたレコードの保証が困難となる。したがって、この方式は、銀行の口座管理のような高度な信頼性を要求されるシステムでは利用することができない。

10

【0019】

本発明はこのような点に鑑みてなされたものであり、高度な信頼性を維持しながら、サービスを停止せずにデータベースを再編成可能なデータベース再編成プログラム、データベース再編成方法、及びデータベース再編成装置を提供することを目的とする。

【課題を解決するための手段】

【0020】

本発明では上記課題を解決するために、図1に示すようなデータベース再編成プログラムが提供される。本発明に係るデータベース再編成プログラムは、コンピュータに、複数のページそれぞれに対して格納すべきレコードが決められている複写元データベース1aを再編成する処理を実行させるためのものである。このデータベース再編成プログラムを実行するコンピュータは、レコード複写手段3とレコード書き込み手段5とを有する。

20

【0021】

レコード複写手段3は、複写元データベース1aの再編成要求に応じて、複写元データベース1a内のページを順次選択し、選択したページ内のレコードを複写先データベース2aに複写すると共に、複写元データベース1a内のページ毎の複写の進捗状況をページ状態管理テーブル4に設定すると、レコード書き込み手段5は、更新レコードの書き込み要求に応じてページ状態管理テーブル4を参照し、複写元データベース1a内の書き込み対象ページが未複写の状態であれば複写元データベース1a内の該当ページに対して更新レコードを書き込み、書き込み対象ページが複写済の状態であれば複写元データベース1aと複写先データベース2aとに更新レコードを書き込む。

30

【0022】

このようなデータベース再編成プログラムを実行するコンピュータによれば、複写元データベース1aの再編成要求が入力されると、レコード複写手段3により複写元データベース1a内のページを順次選択され、選択されたページ内のレコードが複写先データベース2aに複写される。さらに、レコード複写手段3により、複写元データベース1a内のページ毎の複写の進捗状況がページ状態管理テーブル4に設定される。その後、更新レコードの書き込み要求が入力されると、レコード書き込み手段5により、ページ状態管理テーブル4が参照され、複写元データベース1a内の書き込み対象ページが未複写の状態であれば複写元データベース1a内の該当ページに対して更新レコードが書き込まれる。また、書き込み対象ページが複写済の状態であれば複写元データベース1aと複写先データベース2aとに更新レコードが書き込まれる。

40

【0023】

また、上記課題を解決するために、複数のページそれぞれに対して格納すべきレコードが決められているデータベースを再編成するデータベース再編成方法において、レコード複写手段が、前記複写元データベースの再編成要求に応じて、複写元データベース内のページを順次選択し、選択した前記ページ内のレコードを複写先データベースに複写すると共に、前記複写元データベース内のページ毎の複写の進捗状況をページ状態管理テーブルに設定し、レコード書き込み手段が、更新レコードの書き込み要求に応じて前記ページ状態管理テーブルを参照し、前記複写元データベース内の書き込み対象ページが未複写の状態であれば前記複写元データベース内の該当ページに対して前記更新レコードを書き込み

50

、前記書き込み対象ページが複写済の状態であれば前記複写元データベースと前記複写先データベースとに前記更新レコードを書き込む、ことを特徴とするデータベース再編成方法が提供される。

【0024】

このようなデータベース再編成方法によれば、複写元データベースの再編成要求が入力されると、レコード複写手段により複写元データベース内のページが順次選択され、選択されたページ内のレコードが複写先データベースに複写される。さらに、レコード複写手段により、複写元データベース内のページ毎の複写の進捗状況がページ状態管理テーブルに設定される。その後、更新レコードの書き込み要求が入力されると、レコード書き込み手段により、ページ状態管理テーブルが参照され、複写元データベース内の書き込み対象ページが未複写の状態であれば複写元データベース内の該当ページに対して更新レコードが書き込まれる。また、書き込み対象ページが複写済の状態であれば複写元データベースと複写先データベースとに更新レコードが書き込まれる。

10

【0025】

また、上記課題を解決するために、複数のページそれぞれに対して格納すべきレコードが決められているデータベースを再編成するデータベース再編成装置において、前記複写元データベースの再編成要求に応じて、複写元データベース内のページを順次選択し、選択した前記ページ内のレコードを複写先データベースに複写すると共に、前記複写元データベース内のページ毎の複写の進捗状況をページ状態管理テーブルに設定するレコード複写手段と、更新レコードの書き込み要求に応じて前記ページ状態管理テーブルを参照し、前記複写元データベース内の書き込み対象ページが未複写の状態であれば前記複写元データベース内の該当ページに対して前記更新レコードを書き込み、前記書き込み対象ページが複写済の状態であれば前記複写元データベースと前記複写先データベースとに前記更新レコードを書き込むレコード書き込み手段と、を有することを特徴とするデータベース再編成装置が提供される。

20

【0026】

このようなデータベース再編成装置によれば、上記データベース編成プログラムを実行するコンピュータと同様の処理が実行される。

【発明の効果】

【0027】

以上説明したように本発明では、ページ単位で複写を行い、書き込み対象ページが未複写の状態であれば複写元データベース内の該当ページに対して更新レコードを書き込み、複写済の状態であれば複写元データベースと複写先データベースとに更新レコードを書き込むようにした。そのため、データベースへのアクセスを伴う業務を停止せずに再編成が実現できると共に、システム障害時には複写元データベースに基づいてデータベースを容易に復旧できる。

30

【発明を実施するための最良の形態】

【0028】

以下、本発明の実施の形態を図面を参照して説明する。

まず、実施の形態に適用される発明の概要について説明し、その後、実施の形態の具体的な内容を説明する。

40

【0029】

図1は、実施の形態に適用される発明の概念図である。図1の例では、第1の記憶手段1に設けられた複写元データベース1aを再構成し、第2の記憶手段2内に再構成された複写先データベース2aを生成するものとする。複写元データベース1a内の記憶領域は、複数のページに分割されている。そして、各ページに格納すべきレコードが、インデックス1bで定義されている。同様に、複写先データベース2a内の記憶領域は、複数のページに分割されている。そして、各ページに格納すべきレコードが、インデックス2bで定義されている。また、複写元データベース1aと複写先データベース2aとに対してレコードの更新処理が行われると、更新処理の内容がログとして保存される。

50

## 【 0 0 3 0 】

なお、図 1 の例では、インデックス 1 b , 2 b において、レコード番号 ( R m : m は 0 以上の整数 ) とページ番号 ( R n : n は 0 以上の整数 ) との対応関係が示されている。各ページには、インデックス 1 b , 2 b で対応付けられたレコード番号以下であり、且つ、ページ番号が 1 つ小さいページに対応付けられたレコード番号より大きなレコード番号を有するレコードが格納される。

## 【 0 0 3 1 】

レコード複写手段 3 は、複写元データベース 1 a の再編成要求に応じて、複写元データベース 1 a 内のページを順次選択する。そして、レコード複写手段 3 は、選択したページ内のレコードを複写先データベース 2 a に複写する。また、レコード複写手段 3 は、複写元データベース 1 a 内のページ毎の複写の進捗状況をページ状態管理テーブル 4 に設定する。

10

## 【 0 0 3 2 】

ページ状態管理テーブル 4 には、複写元データベース 1 a 内のページ毎に、複写の状態が設定されている。複写の状態としては、未複写 ( 図中「未」と示す )、複写中 ( 図中「中」と示す )、及び複写済 ( 図中「済」と示す ) がある。

## 【 0 0 3 3 】

レコード書き込み手段 5 は、更新レコードの書き込み要求に応じてページ状態管理テーブル 4 を参照し、書き込み対象のページの状態に応じて、更新レコードの書き込み対象を判定する。具体的には、レコード書き込み手段 5 は、複写元データベース 1 a 内の書き込み対象ページが未複写の状態であれば複写元データベース 1 a 内の該当ページに対して更新レコードを書き込む。また、レコード書き込み手段 5 は、書き込み対象ページが複写済の状態であれば複写元データベース 1 a と複写先データベース 2 a とに更新レコードを書き込む。さらに、レコード書き込み手段 5 は、書き込み対象ページが複写中の状態であれば複写済に遷移するのを待ち、複写中から複写済の状態に遷移したときに、複写元データベース 1 a と複写先データベース 2 a とに更新レコードを書き込む。

20

## 【 0 0 3 4 】

このようなデータベース再編成装置によれば、複写元データベース 1 a の再編成要求が入力されると、レコード複写手段 3 により複写元データベース 1 a 内のページが順次選択され、選択されたページ内のレコードが複写先データベース 2 a に複写される。さらに、レコード複写手段 3 により、複写元データベース 1 a 内のページ毎の複写の進捗状況がページ状態管理テーブル 4 に設定される。

30

## 【 0 0 3 5 】

ここで、ページ番号 ( 「 P 0 」 , 「 P 1 」 , 「 P 2 」 ) が小さい順に、ページが選択されるものとする。図 1 には、ページ番号「 P 1 」を複写しているときのページ状態管理テーブル 4 が示されている。この例では、ページ番号「 P 0 」のページが複写済、ページ番号「 P 1 」のページが複写中、ページ番号「 P 2 」のページが未複写である。

## 【 0 0 3 6 】

データベース再編成処理中に、更新レコードの書き込み要求が入力されると、レコード書き込み手段 5 により、ページ状態管理テーブル 4 が参照され、書き込み対象のページの状態が判断される。

40

## 【 0 0 3 7 】

たとえば、レコード番号「 R 3 」が更新レコード 6 a として入力された場合、複写元データベース 1 a のページ番号「 P 0 」のページが書き込み対象となる。図 1 の例では、ページ番号「 P 0 」のページは、複写済の状態である。そのため、更新レコード 6 a は、複写元データベース 1 a と複写先データベース 2 a とに書き込まれる。

## 【 0 0 3 8 】

レコード番号「 R 4 」が更新レコード 6 b として入力された場合、複写元データベース 1 a のページ番号「 P 1 」のページが書き込み対象となる。図 1 の例では、ページ番号「 P 1 」のページは、複写中の状態である。そのため、該当ページの状態が複写済になるの

50

を待ち、更新レコード 6 a が複写元データベース 1 a と複写先データベース 2 a とに書き込まれる。

【 0 0 3 9 】

レコード番号「 R 5 」が更新レコード 6 c として入力された場合、複写元データベース 1 a のページ番号「 P 2 」のページが書き込み対象となる。図 1 の例では、ページ番号「 P 2 」のページは、未複写の状態である。そのため、更新レコード 6 c は複写元データベース 1 a に書き込まれる。

【 0 0 4 0 】

このようにして、レコードの書き込みを停止することなく、データベースの再編成が行われる。なお、ページが複写中の場合、書き込みの待機時間が発生するが、1 ページの複写処理は短時間で実行できるため、業務を停止させる程の影響はない。しかも、複写済のページに対する更新の場合、複写元データベース 1 a に対しても書き込みが行われるため、ジョブキャンセル時等のデータベースの復旧が可能となる。

10

【 0 0 4 1 】

以下、本発明の実施の形態を具体的に説明する。

図 2 は、本発明の実施の形態のシステム構成例を示す図である。本発明の実施の形態では、複数のホストコンピュータ 1 0 0 , 2 0 0 , 3 0 0 , 4 0 0 と複数の端末装置 2 1 , 2 2 , . . . とがネットワーク 1 0 を介して接続されている。ホストコンピュータ 1 0 0 , 2 0 0 , 3 0 0 , 4 0 0 には、ストレージデバイス 5 0 0 , 6 0 0 が接続されている。また、ホストコンピュータ 1 0 0 と他のホストコンピュータ 2 0 0 , 3 0 0 , 4 0 0 とは、高速の通信回線で互いに接続されている。ホストコンピュータ 2 0 0 , 3 0 0 , 4 0 0 は、その通信回線を介して、ホストコンピュータ 1 0 0 内の共有メモリにアクセスできる。

20

【 0 0 4 2 】

ホストコンピュータ 1 0 0 , 2 0 0 , 3 0 0 , 4 0 0 は、端末装置 2 1 , 2 2 , . . . からの要求に応じて、ストレージデバイス 5 0 0 , 6 0 0 内に構築されたデータベースに対するランザクションを実行する。また、少なくとも 1 つのホストコンピュータは、データベースを再構成するための処理機能を有している。本実施の形態では、ホストコンピュータ 1 0 0 がデータベースの再構成を行うものとする。ホストコンピュータ 1 0 0 は、データベースの再構成を行うために、以下の様なハードウェア構成を有している。

30

【 0 0 4 3 】

図 3 は、データベースの再構成を行うホストコンピュータのハードウェア構成例を示す図である。ホストコンピュータ 1 0 0 は、CPU (Central Processing Unit) 1 0 1 によって装置全体が制御されている。CPU 1 0 1 には、バス 1 0 1 a を介して RAM (Random Access Memory) 1 0 2 、ハードディスクドライブ (HDD : Hard Disk Drive) 1 0 3 、グラフィック処理装置 1 0 4 、入力インタフェース 1 0 5 、ストレージデバイスインタフェース 1 0 7 、通信インタフェース 1 0 8 、およびホストコンピュータ間通信インタフェース 1 0 9 が接続されている。

【 0 0 4 4 】

RAM 1 0 2 には、CPU 1 0 1 に実行させる OS (Operating System) のプログラムやアプリケーションプログラムの少なくとも一部が一時的に格納される。また、RAM 1 0 2 には、CPU 1 0 1 による処理に必要な各種データが格納される。HDD 1 0 3 には、OS やアプリケーションプログラムが格納される。

40

【 0 0 4 5 】

グラフィック処理装置 1 0 4 には、モニタ 1 1 が接続されている。グラフィック処理装置 1 0 4 は、CPU 1 0 1 からの命令に従って、画像をモニタ 1 1 の画面に表示させる。入力インタフェース 1 0 5 には、キーボード 1 2 とマウス 1 3 とが接続されている。入力インタフェース 1 0 5 は、キーボード 1 2 やマウス 1 3 から送られてくる信号を、バス 1 0 1 a を介して CPU 1 0 1 に送信する。

【 0 0 4 6 】

50

共有メモリ106は、他のホストコンピュータ200, 300, 400との間で共有可能な記録媒体である。共有メモリ106としては、たとえば、バッテリーでバックアップされた半導体メモリが使用される。データベースの再編成時には、共有メモリ106内に、データベース内のブロック毎の状態を示すページ状態管理テーブルが格納される。

【0047】

ストレージデバイスインタフェース107は、ストレージデバイス500, 600に接続されている。ストレージデバイスインタフェース107は、CPU101からの指示に従って、ストレージデバイス500, 600に対するデータの入出力を行う。

【0048】

通信インタフェース108は、ネットワーク10に接続されている。通信インタフェース108は、ネットワーク10を介して、端末装置21, 22, ...との間でデータの送受信を行う。

【0049】

ホストコンピュータ間通信インタフェース109は、他のホストコンピュータ200, 300, 400に接続されている。ホストコンピュータ間通信インタフェース109は、他のホストコンピュータ200, 300, 400との間でデータの送受信を行う。たとえば、ホストコンピュータ間通信インタフェース109は、他のホストコンピュータ200, 300, 400に対して共有メモリ106に格納されている情報を送信する。

【0050】

以上のようなハードウェア構成によって、本実施の形態の処理機能を実現することができる。なお、図3には、ホストコンピュータ100のハードウェア構成を説明したが、他のホストコンピュータ200, 300, 400も同様のハードウェア構成で実現することができる。ただし、ホストコンピュータ200, 300, 400には、共有メモリは搭載されていなくてもよい。

【0051】

以下、図2, 3に示す構成のシステムにおけるデータベース再編成処理について詳細に説明する。

図4は、本実施の形態を実現するための機能を示すブロック図である。図4には、ストレージデバイス500内に構築されているデータベースを、ストレージデバイス600内に再編成する場合の例が示されている。

【0052】

ストレージデバイス500には、インデックス510、複写元のデータベース520、更新前ログ530、及び更新後ログ540が含まれる。データベース再編成を実行する以前は、ストレージデバイス500内の各要素を用いて端末装置21, 22, ...に対するサービスが提供される。

【0053】

インデックス510には、データベース520内に設けられた各ページに格納すべきレコードの識別番号(レコード番号)が定義されている。データベース520は、複数のページに分割された記憶領域を有している。各ページには、インデックス510で対応付けられたレコードが格納される。更新前ログ530には、データベース520内のレコードが更新された際の更新前のレコードの内容が記録される。更新後ログ540には、データベース520内のレコードが更新された際の更新後のレコードの内容が記録される。

【0054】

ストレージデバイス600には、インデックス610、複写先のデータベース620、更新前ログ630、及び更新後ログ640が含まれる。なお、ストレージデバイス600内の各要素は、データベース再編成処理が開始された際に生成される。

【0055】

インデックス610には、データベース620内に設けられた各ページに格納すべきレコードの識別番号(レコード番号)が定義されている。データベース620は、複数のページに分割された記憶領域を有している。各ページには、インデックス610で対応付け

10

20

30

40

50

られたレコードが格納される。更新前ログ630には、データベース620内のレコードが更新された際の更新前のレコードの内容が記録される。更新後ログ640には、データベース620内のレコードが更新された際の更新後のレコードの内容が記録される。

【0056】

ホストコンピュータ100は、複写トランザクション処理部110、ページ状態管理テーブル120、およびオンライントランザクション処理部130を有する。

複写トランザクション処理部110は、データベースの再編成処理を行う。具体的には、複写トランザクション処理部110は、データベース520の再編成要求を受け付けると、ストレージデバイス600内に、前述のインデックス610、複写先のデータベース620、更新前ログ630、及び更新後ログ640を生成する。このとき、複写先のデータベース620の記憶容量が複写元のデータベース520の記憶容量よりも大きくなるように、インデックス610内の情報を設定する。そして、複写トランザクション処理部110は、データベース520内のレコードをページ単位でデータベース620に複写する。その際の各ページの状態は、ページ状態管理テーブル120に設定される。

10

【0057】

ページ状態管理テーブル120には、データベース再編成中に、複写元のデータベース520内の各ページの状態が登録される。ページの状態は、「未複写」、「複写中」、「複写済」がある。

【0058】

オンライントランザクション処理部130は、端末装置21, 22, ...からの要求に応じてデータベース操作のトランザクションを実行する。なお、データベース再構成処理中においては、オンライントランザクション処理部130は、ページ状態管理テーブル120を参照し、アクセス対象とすべきデータベースを判断する。

20

【0059】

次に、インデックス510とデータベース520との関係について詳細に説明する。

図5は、インデックスとデータベースとのデータ構造を示す図である。インデックス510には、各ページのページ番号に対応付けて、そのページに格納すべきレコードのレコード番号が定義されている(ページ番号をPm、レコード番号をRnで表している)。

【0060】

効率的なトランザクション処理を行うために、レコード番号の小さいレコードほど、ページ番号の小さいページに割り当てられる。そこで、インデックス510では、各ページのページ番号に対応付けて、そのページに格納されるレコードに対するレコード番号の最大値が設定されている。

30

【0061】

図5の例ではインデックス510において、ページ番号「P0」に対応付けてレコード番号「R3」が設定されている。これは、ページ番号「P0」のページに対して、レコード番号が「R0x」~「R3x」(xは0以上の整数)のレコードが登録されることを意味する。

【0062】

また、ページ番号「P1」に対応付けてレコード番号「R4」が設定されている。これは、ページ番号「P1」のページに対して、レコード番号「R4x」のレコードが登録されることを意味する。

40

【0063】

ページ番号「P2」に対応付けてレコード番号「R6」が設定されている。これは、ページ番号「P2」のページに対して、レコード番号が「R5x」、「R6x」のレコードが登録されることを意味する。

【0064】

なお、データベース520がネットワークデータベース(NDB)として利用される場合、各レコードには、木構造で互いに関連付けられた複数のデータが含まれる。この場合、データベース520でのデータ間の関連付けは、複写先のデータベース620で再現さ

50

れる。

【0065】

データベース520には、複数のプライムページを格納するためのプライムページ格納領域521と複数のオーバフローページを格納するためのオーバフローページ格納領域522とが設けられている。プライムページは、インデックス510に設定されている各ページ番号に対応する記憶領域である。各プライムページには、インデックス510で定義されたレコード番号のレコードのみが格納できる。

【0066】

オーバフローページは、プライムページの容量不足によって、レコード番号に対応するプライムページに格納できなかったレコードを格納するための記憶領域である。オーバフローページにレコードが格納された場合、そのレコードのレコード番号に対応するプライムページ内に、オーバフローページへのポインタが設定される。そのポインタを参照することで、レコードが格納されたオーバフローページを特定できる。

10

【0067】

このように、各プライムページには、格納可能なレコードのレコード番号が予め指定されている。そして、1つのプライムページにレコードが格納しきれなくなると、レコードがオーバフローページに格納される。

【0068】

しかし、オーバフローページに格納されるレコードの量が増加すると、データベース520へのアクセス効率が悪化する。すなわち、任意のレコードにアクセスする場合、オンライントランザクション処理部130は、まずインデックス510を参照して、レコードが登録されているプライムページを判断する。そして、オンライントランザクション処理部130は、該当するプライムページからアクセス対象のレコードを検索する。

20

【0069】

ここで、アクセス対象のレコードが見つからない場合、オンライントランザクション処理部130は、プライムページに登録されているポインタを参照して、関連付けられたオーバフローページを判断する。そして、オンライントランザクション処理部130は、該当するオーバフローページからアクセス対象のレコードを検索する。

【0070】

このように、オーバフローページに格納されたページにアクセスする場合、オンライントランザクション処理部130において2段階の検索処理が行われる。そのため、アクセスの処理効率が低下する。

30

【0071】

そこで、オーバフローページに格納されたレコードの量が増加した場合、プライムページの容量を拡張するように、データベース再編成が行われる。データベースを再編成する際には、データベース520内の各レコードがページ単位でデータベース620に複写される。その際、データベース520内のプライムページの状態が、ページ状態管理テーブル120で管理される。

【0072】

図6は、ページ状態管理テーブルのデータ構造例を示す図である。ページ状態管理テーブル120には、データベース520内のプライムページのページ番号に対応付けて、そのページの状態が設定されている。状態には、未複写(図中「未」と示す)、複写中(図中「中」と示す)、及び複写済(図中「済」と示す)がある。未複写の状態とは、データベース520からデータベース620へのレコードの複写処理が行われていないことを示す。複写中の状態とは、データベース520からデータベース620へのレコードの複写中であることを示す。複写済みの状態とは、データベース520からデータベース620へのレコードの複写が完了したことを示す。

40

【0073】

なお、本実施の形態では、オンライントランザクションを継続して実行しながらデータベース再編成処理を行うことができる。以下、データベース再構成処理と、データベース

50

再構成中のオンライントランザクションについて具体的に説明する。

【0074】

図7は、データベース再編成処理の手順を示すフローチャートである。以下、図7に示す処理をステップ番号に沿って説明する。

〔ステップS11〕複製トランザクション処理部110は、ストレージデバイス600にインデックス610を作成する。このインデックス610では、プライムページの数を増加させる。すなわち、データベース520のプライムページよりも多数のプライムページ格納領域621について定義される。プライムページの数が増加することにより、データベースの容量が増加する。

【0075】

〔ステップS12〕複製トランザクション処理部110は、移行先ボリュームを決定する。本実施の形態では、ストレージデバイス600を移行先ボリュームとする。

〔ステップS13〕複製トランザクション処理部110は、ページ状態管理テーブル120の全てのページの状態を「未複製」にする。

【0076】

〔ステップS14〕複製トランザクション処理部110は、データベース520のプライムページから、最もページ番号が小さい未複製のページを選択する。

〔ステップS15〕複製トランザクション処理部110は、ページ状態管理テーブル120におけるステップS14で選択したプライムページの状態を、「複製中」に変更する。

【0077】

〔ステップS16〕複製トランザクション処理部110は、複製元のデータベース520内のステップS14で選択したプライムページ内のレコードを、ストレージデバイス600のデータベース620に複製する。この際の複製先のプライムページは、ステップS11で作成したインデックス610に従う。また、複製トランザクション処理部110は、ステップS14で選択したプライムページからポインタで指し示されたオーバフローページに格納されたレコードも、同様にデータベース620に複製する。

【0078】

〔ステップS17〕複製トランザクション処理部110は、ステップS14で選択されたプライムページ及びポインタで関連付けられたオーバフローページに格納されていた全てのレコードを複製したか否かを判断する。ページ内の全てのレコード複製が完了したら、処理がステップS18に進められる。複製していないレコードがある場合、処理がステップS16に進められ、該当レコードの複製が行われる。

【0079】

〔ステップS18〕複製トランザクション処理部110は、ページ状態管理テーブル120におけるステップS14で選択したプライムページの状態を、「複製済」に変更する。

【0080】

〔ステップS19〕複製トランザクション処理部110は、データベース520内の全てのページに対する複製処理が完了したか否かを判断する。全てのページの複製が完了した場合、処理が終了する。複製していないページがある場合、処理がステップS14に進められ、該当ページの複製が行われる。

【0081】

このように、複製トランザクション処理部110がページ単位で複製を行い、ページ状態管理テーブル120における複製対象のページの状態を適宜更新する。これにより、ページ状態管理テーブル120を参照することで、オンライントランザクション処理部130が各プライムページの状態を認識することができる。オンライントランザクション処理部130が、プライムページの状態を認識することにより、データベース再構成中であっても、データベースに対するレコード更新のトランザクションが可能となる。

【0082】

10

20

30

40

50

図 8 は、データベース再構成中のレコード更新処理の手順を示すフローチャートである。以下、図 8 に示す処理をステップ番号に沿って説明する。

【ステップ S 2 1】オンライントランザクション処理部 1 3 0 は、インデックス 5 1 0 を参照し、処理対象のレコードが格納されているプライムページを判断する。そして、オンライントランザクション処理部 1 3 0 は、ページ状態管理テーブル 1 2 0 を参照し、該当するプライムページの状態が複写中か否かを判断する。複写中であれば、処理がステップ S 2 2 に進められる。複写中でなければ、処理がステップ S 2 3 に進められる。

【0 0 8 3】

【ステップ S 2 2】アクセス対象のプライムページが複写中の場合、オンライントランザクション処理部 1 3 0 は、所定時間待機後処理をステップ S 2 1 に進め、再度プライムページの状態を判断する。

10

【0 0 8 4】

【ステップ S 2 3】アクセス対象のプライムページが複写中でなければ、オンライントランザクション処理部 1 3 0 は、複写元のデータベース 5 2 0 から更新対象のレコードをリードする。

【0 0 8 5】

【ステップ S 2 4】オンライントランザクション処理部 1 3 0 は、ステップ S 2 3 でリードしたレコードの内容を更新し、データベース 5 2 0 にライトする。

【ステップ S 2 5】オンライントランザクション処理部 1 3 0 は、ページ状態管理テーブル 1 2 0 を参照し、更新対象のプライムページの状態が複写済か否かを判断する。複写済みであれば、処理がステップ S 2 6 に進められる。複写済みでなければ、処理が終了する。

20

【0 0 8 6】

【ステップ S 2 6】オンライントランザクション処理部 1 3 0 は、複写先のデータベース 6 2 0 から更新対象のレコードをリードする。

【ステップ S 2 7】オンライントランザクション処理部 1 3 0 は、ステップ S 2 6 でリードしたレコードの内容を更新し、データベース 6 2 0 にライトする。その後、処理が終了する。

【0 0 8 7】

このように、更新対象となるプライムページの状態に応じて、アクセスするデータベースを判断することで、データベース再編成中であってもオンライントランザクションが可能となる。

30

【0 0 8 8】

図 9 は、データベース更新処理とオンライントランザクションとが同時に実行された場合の処理を示す概念図である。システム管理者からの操作入力等によってデータベース 5 2 0 の再編成要求が出されると、ストレージデバイス 6 0 0 内にインデックス 6 1 0 とデータベース 6 2 0 とが生成される。データベース 6 2 0 は、複数のプライムページを格納するプライムページ格納領域 6 2 1 と複数のオーバフローページを格納するオーバフローページ格納領域 6 2 2 を含んでいる。

【0 0 8 9】

40

その後、複写トランザクション処理部 1 1 0 によって、複写元のデータベース 5 2 0 から複写先のデータベース 6 2 0 へのレコードの複写が開始される。レコードの複写は、1 つのプライムページ毎に生成される複写トランザクション 1 1 1 により実行される。

【0 0 9 0】

複写トランザクション 1 1 1 がスタートすると、ページ状態管理テーブル 1 2 0 内の複写対象のプライムページ（ページ番号が小さいプライムページから順に、処理対象となる）の状態が複写中に変更され、複写が開始する（ステップ S 3 1）。図 9 の例では、ページ番号が「P 0」のプライムページの状態が複写中（図中「中」と示す）に変更されている。

【0 0 9 1】

50

次に、複写トランザクション111では、複写元のデータベース520内の複写対象のプライムページ、及びそのプライムページからポインタで関連付けられたオーバフローページから、順次レコードが読み出される(ステップS32)。図9の例では、プライムページからレコード番号「R1」、「R2」のレコードが読み出され、オーバフローページからレコード番号「R3」のレコードが読み出される。

【0092】

次に、読み出されたレコードの内容が、インデックス610に従って複写先のデータベース620に書き込まれる(ステップS33)。図9の例では、複写先のデータベース620に対するインデックス610において、レコード番号「R1」、「R2」、「R3」のレコードは、それぞれページ番号「P0」、「P1」、「P2」に対応付けられている。したがって、レコード番号「R1」のレコードはページ番号「P0」のプライムページに書き込まれ、「R2」のレコードはページ番号「P1」のプライムページに書き込まれ、「R3」のレコードはページ番号「P2」のプライムページに書き込まれる。

10

【0093】

複写先のデータベース620に複写元のデータベース520の「P0」のプライムページのレコードが全て格納されると、ページ状態管理テーブル120における該当ページの状態が、複写中から複写済(図中「済」と示す)に変更される。これにより、1つのプライムページに対する複写トランザクションが終了する。

【0094】

複写トランザクション処理部110は、他のプライムページ(「P1」、「P2」)に対する同様の複写トランザクションを繰り返し実行する。これにより、複写元のデータベース520の先頭ページから最終ページまでのレコード複写が実施される。

20

【0095】

ここで、オンライントランザクション処理部130が端末装置21からレコードの更新の処理要求を受け取ると、オンライントランザクション131が開始される。オンライントランザクション131では、複写元のデータベース520にレコードが追加される(ステップS41)。また、ページ状態管理テーブル120が参照され、書き込み対象のプライムページの状態が複写済の場合、複写先のデータベース620にレコードが追加される(ステップS42)。なお、複写先のデータベース620にレコード追加をする際のページ番号は、複写先のデータベース620のインデックス610に基づいて判断される。

30

【0096】

ページ状態が複写中の場合は、排他制御によりオンライントランザクション131を待機させる。そして、複写が完了しページ状態が複写済になると、複写先のデータベース620にレコードが追加される。

【0097】

ページ状態が未複写の場合は、複写先のデータベース620へのレコードの格納処理は行われない。すなわち、複写元のデータベース520だけにレコードが格納される。この場合、複写トランザクション111により複写元のデータベース520から複写先のデータベース620にページの内容が複写され、論理的な等価性が保証される。

【0098】

このように、再編成単位を最小単位のページ単位にしたことで、オンラインサービスを停止することなくデータベース再編成が可能となる。しかも、複写済のページに対して更新を行う場合、複写元のデータベース520と複写先のデータベース620との双方にレコードが書き込まれる。このとき、オンライントランザクション処理部130により、書き込まれたレコードの内容を含むログが更新後ログ540に記録される。更新後ログ540にレコードが記録されることで、ホストコンピュータ100が異常終了等によりデータベース520の内容が不完全なものとなっても、データベース520の内容を容易に復旧することができる。

40

【0099】

さらに、本実施の形態では、レコード更新時において、オンラインで使用している複写

50

元のデータベース520とオンライン再編成対象の複写先のデータベース620とへのレコードの書き込みを、同一トランザクションで処理している。これにより、ジョブキャンセル等があっても、元のデータを保証できる。

【0100】

なお、上記の処理機能は、コンピュータによって実現することができる。その場合、ホストコンピュータが有すべき機能の処理内容を記述したプログラムが提供される。そのプログラムをコンピュータで実行することにより、上記処理機能がコンピュータ上で実現される。処理内容を記述したプログラムは、コンピュータで読み取り可能な記録媒体に記録しておくことができる。コンピュータで読み取り可能な記録媒体としては、磁気記録装置、光ディスク、光磁気記録媒体、半導体メモリなどがある。磁気記録装置には、ハードディスク装置(HDD)、フレキシブルディスク(FD)、磁気テープなどがある。光ディスクには、DVD(Digital Versatile Disc)、DVD-RAM(Random Access Memory)、CD-ROM(Compact Disc Read Only Memory)、CD-R(Recordable)/RW(ReWritable)などがある。光磁気記録媒体には、MO(Magneto-Optical disk)などがある。

10

【0101】

プログラムを流通させる場合には、たとえば、そのプログラムが記録されたDVD、CD-ROMなどの可搬型記録媒体が販売される。また、プログラムをサーバコンピュータの記憶装置に格納しておき、ネットワークを介して、サーバコンピュータから他のコンピュータにそのプログラムを転送することもできる。

【0102】

プログラムを実行するコンピュータは、たとえば、可搬型記録媒体に記録されたプログラムもしくはサーバコンピュータから転送されたプログラムを、自己の記憶装置に格納する。そして、コンピュータは、自己の記憶装置からプログラムを読み取り、プログラムに従った処理を実行する。なお、コンピュータは、可搬型記録媒体から直接プログラムを読み取り、そのプログラムに従った処理を実行することもできる。また、コンピュータは、サーバコンピュータからプログラムが転送される毎に、逐次、受け取ったプログラムに従った処理を実行することもできる。

20

【0103】

(付記1) コンピュータに、複数のページそれぞれに対して格納すべきレコードが決められているデータベースを再編成処理を行わせるデータベース再編成プログラムにおいて、

30

前記コンピュータを、

前記複写元データベースの再編成要求に応じて、複写元データベース内のページを順次選択し、選択した前記ページ内のレコードを複写先データベースに複写すると共に、前記複写元データベース内のページ毎の複写の進捗状況をページ状態管理テーブルに設定するレコード複写手段、

更新レコードの書き込み要求に応じて前記ページ状態管理テーブルを参照し、前記複写元データベース内の書き込み対象ページが未複写の状態であれば前記複写元データベース内の該当ページに対して前記更新レコードを書き込み、前記書き込み対象ページが複写済の状態であれば前記複写元データベースと前記複写先データベースとに前記更新レコードを書き込むレコード書き込み手段、

40

として機能させることを特徴とするデータベース再編成プログラム。

【0104】

(付記2) 前記コンピュータを前記レコード書き込み手段として機能させる際、前記書き込み対象ページが複写中の状態であれば複写済に遷移するのを待ち、複写中から複写済の状態に遷移したときに、前記複写元データベースと前記複写先データベースとに前記更新レコードを書き込むことを特徴とする付記1記載のデータベース再編成プログラム。

【0105】

(付記3) 前記コンピュータを前記レコード複写手段として機能させる際、前記再編成要求を受け取った際に、前記ページ状態管理テーブルにおいて全てのページの状態を未

50

複写とし、複写を開始したページの状態を複写中に変更し、複写が完了したページの状態を複写済にすることを特徴とする付記1記載のデータベース再編成プログラム。

【0106】

(付記4) 前記複写元データベースは、格納すべきレコードの番号が予め規定されている複数のプライムページと、前記プライムページに格納できなかったレコードが格納される複数のオーバフローページとを有しており、

前記コンピュータを前記レコード複写手段として機能させる際、前記プライムページを順次選択し、選択された前記プライムページと、選択された前記プライムページに格納できなかったレコードが格納された前記オーバフローページとの複写処理を行うことを特徴とする付記1記載のデータベース再編成プログラム。

10

【0107】

(付記5) 前記コンピュータを前記レコード複写手段として機能させる際、選択された前記プライムページの複写処理と、選択された前記プライムページに格納できなかったレコードが格納された前記オーバフローページの複写処理とを、同一トランザクションで実行することを特徴とする付記4記載のデータベース再編成プログラム。

【0108】

(付記6) 前記コンピュータを前記レコード複写手段として機能させる際、前記プライムページを選択したとき、選択された前記プライムページの状態を複写中に変更し、選択された前記プライムページ及び選択された前記プライムページに格納できなかったレコードが格納された前記オーバフローページとの複写処理が共に完了したとき、選択された前記プライムページの状態を複写済に変更することを特徴とする付記4記載のデータベース再編成プログラム。

20

【0109】

(付記7) 複数のページそれぞれに対して格納すべきレコードが決められているデータベースを再編成するデータベース再編成方法において、

レコード複写手段が、前記複写元データベースの再編成要求に応じて、複写元データベース内のページを順次選択し、選択した前記ページ内のレコードを複写先データベースに複写すると共に、前記複写元データベース内のページ毎の複写の進捗状況をページ状態管理テーブルに設定し、

レコード書き込み手段が、更新レコードの書き込み要求に応じて前記ページ状態管理テーブルを参照し、前記複写元データベース内の書き込み対象ページが未複写の状態であれば前記複写元データベース内の該当ページに対して前記更新レコードを書き込み、前記書き込み対象ページが複写済の状態であれば前記複写元データベースと前記複写先データベースとに前記更新レコードを書き込む、

30

ことを特徴とするデータベース再編成方法。

【0110】

(付記8) 複数のページそれぞれに対して格納すべきレコードが決められているデータベースを再編成するデータベース再編成装置において、

前記複写元データベースの再編成要求に応じて、複写元データベース内のページを順次選択し、選択した前記ページ内のレコードを複写先データベースに複写すると共に、前記複写元データベース内のページ毎の複写の進捗状況をページ状態管理テーブルに設定するレコード複写手段と、

40

更新レコードの書き込み要求に応じて前記ページ状態管理テーブルを参照し、前記複写元データベース内の書き込み対象ページが未複写の状態であれば前記複写元データベース内の該当ページに対して前記更新レコードを書き込み、前記書き込み対象ページが複写済の状態であれば前記複写元データベースと前記複写先データベースとに前記更新レコードを書き込むレコード書き込み手段と、

を有することを特徴とするデータベース再編成装置。

【0111】

(付記9) コンピュータに、複数のページそれぞれに対して格納すべきレコードが決

50

められているデータベースを再編成処理を行わせるデータベース再編成プログラムを記録したコンピュータ読み取り可能な記録媒体において、

前記コンピュータを、

前記複写元データベースの再編成要求に応じて、複写元データベース内のページを順次選択し、選択した前記ページ内のレコードを複写先データベースに複写すると共に、前記複写元データベース内のページ毎の複写の進捗状況をページ状態管理テーブルに設定するレコード複写手段、

更新レコードの書き込み要求に応じて前記ページ状態管理テーブルを参照し、前記複写元データベース内の書き込み対象ページが未複写の状態であれば前記複写元データベース内の該当ページに対して前記更新レコードを書き込み、前記書き込み対象ページが複写済の状態であれば前記複写元データベースと前記複写先データベースとに前記更新レコードを書き込むレコード書き込み手段、

として機能させることを特徴とするデータベース再編成プログラムを記録したコンピュータ読み取り可能な記録媒体。

【図面の簡単な説明】

【0112】

【図1】実施の形態に適用される発明の概念図である。

【図2】本発明の実施の形態のシステム構成例を示す図である。

【図3】データベースの再構成を行うホストコンピュータのハードウェア構成例を示す図である。

【図4】本実施の形態を実現するための機能を示すブロック図である。

【図5】インデックスとデータベースとのデータ構造を示す図である。

【図6】ページ状態管理テーブルのデータ構造例を示す図である。

【図7】データベース再編成処理の手順を示すフローチャートである。

【図8】データベース再構成中のレコード更新処理の手順を示すフローチャートである。

【図9】データベース更新処理とオンライントランザクションとが同時に同時に実行された場合の処理を示す概念図である。

【図10】従来の一般的なデータベース再構築方法を示す図である。

【図11】アクセスを停止せずにデータベース再編成を実行する方式を示す図である。

【符号の説明】

【0113】

- 1 第1の記憶手段
- 1 a 複写元データベース
- 1 b インデックス
- 2 第2の記憶手段
- 2 a 複写先データベース
- 2 b インデックス
- 3 レコード複写手段
- 4 ページ状態管理テーブル
- 5 レコード書き込み手段
- 6 a , 6 b , 6 c 更新レコード

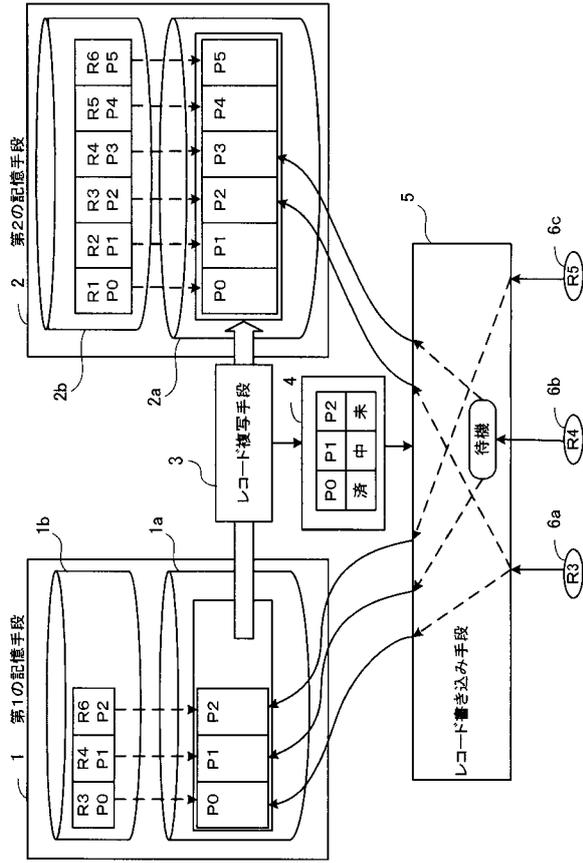
10

20

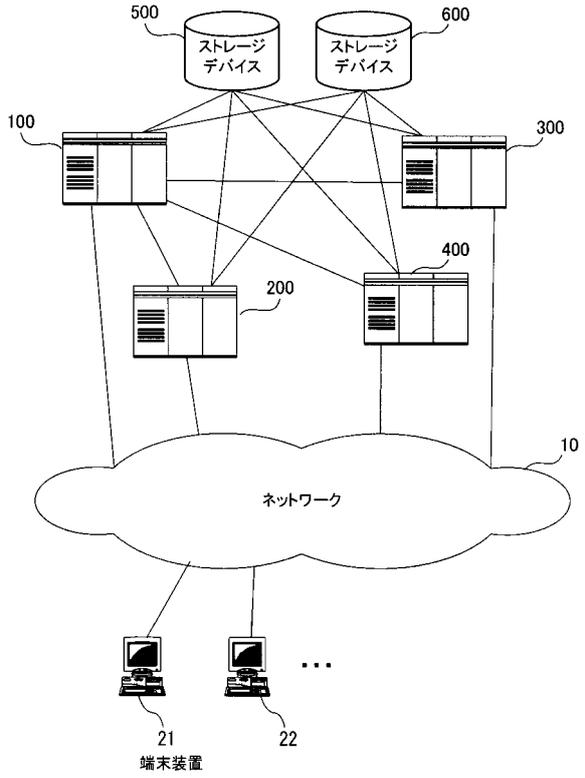
30

40

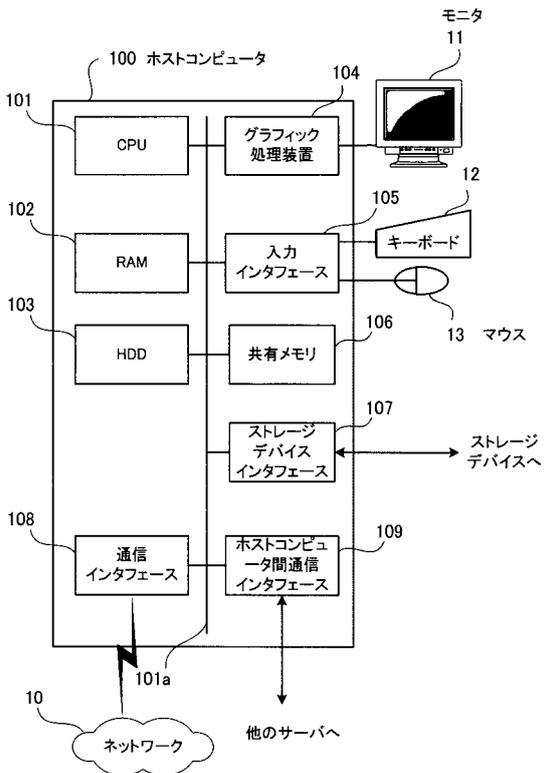
【図1】



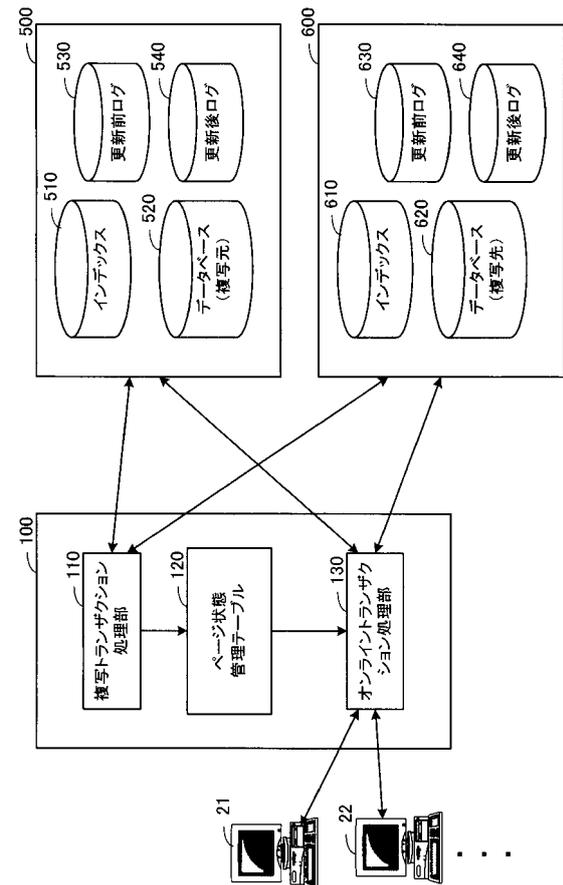
【図2】



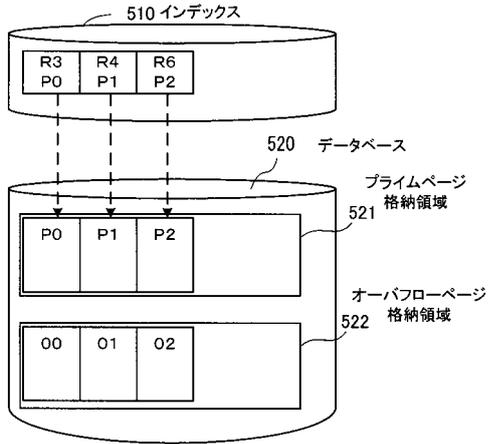
【図3】



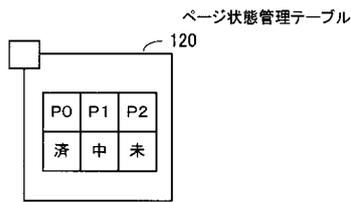
【図4】



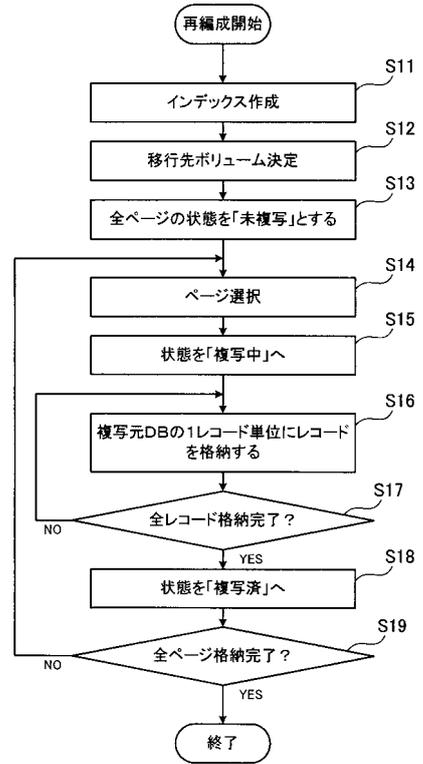
【図5】



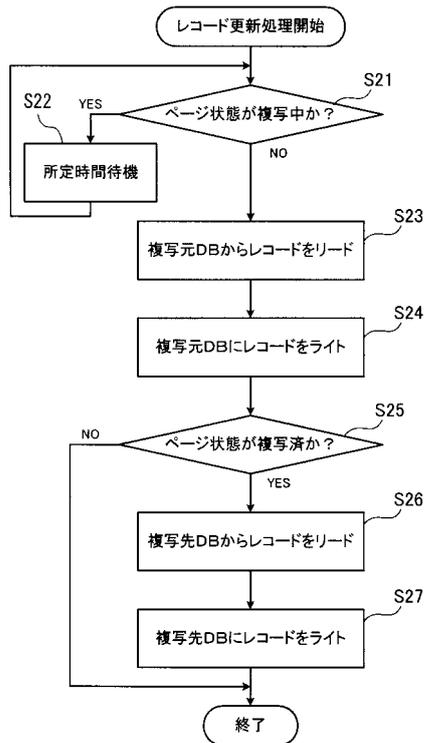
【図6】



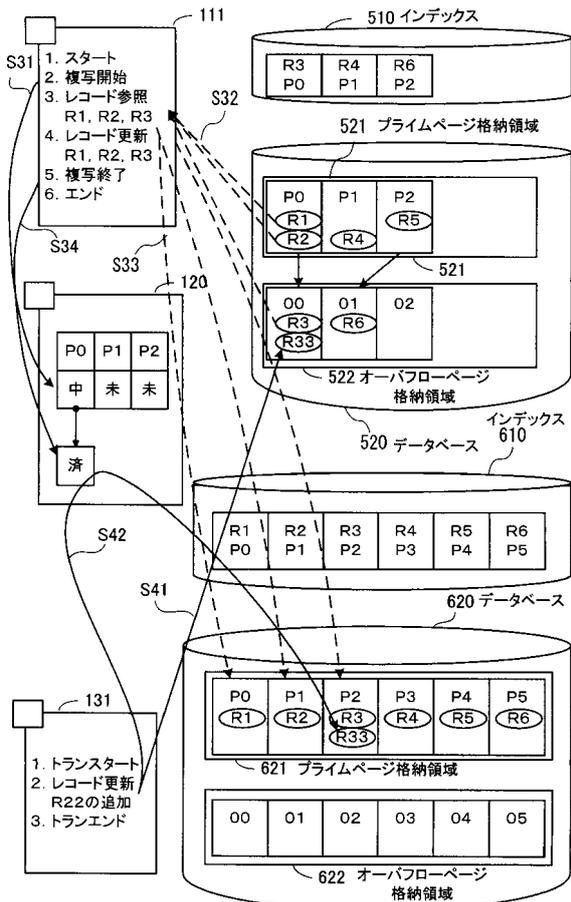
【図7】



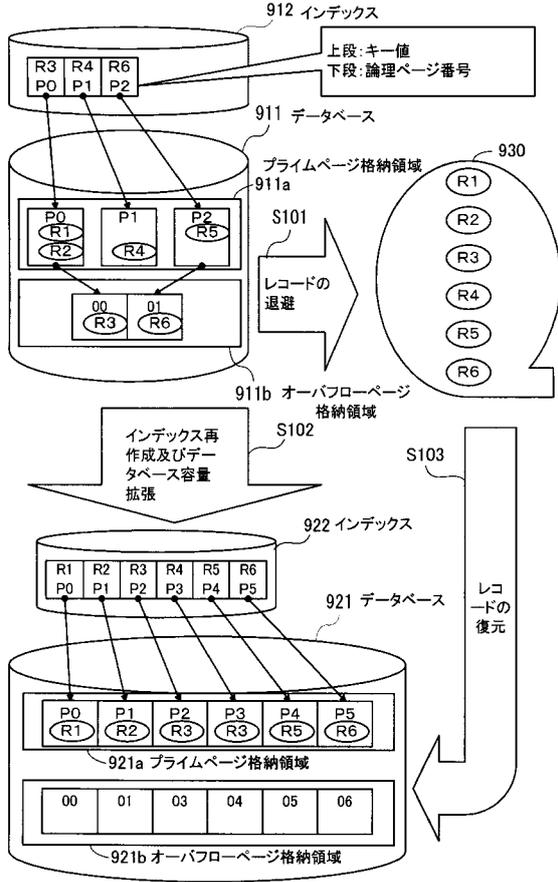
【図8】



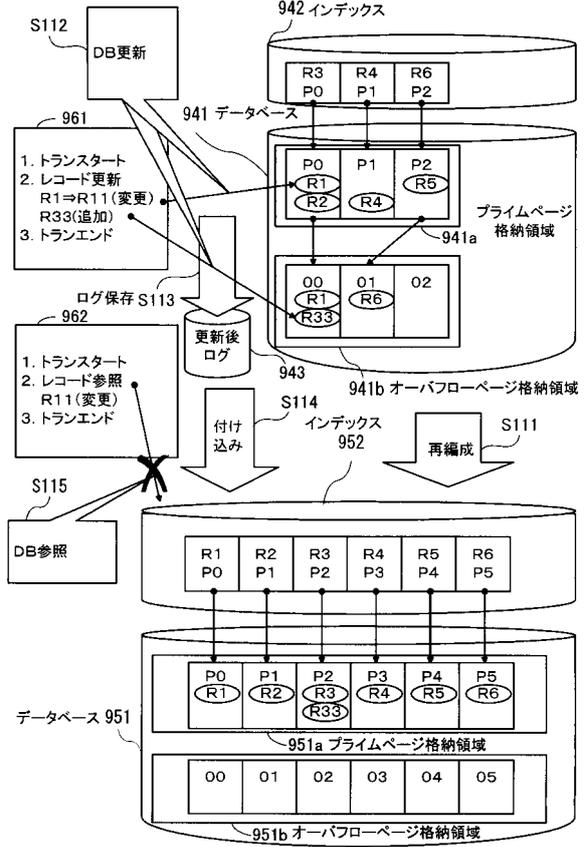
【図9】



【図10】



【図11】



---

フロントページの続き

(56)参考文献 国際公開第2004/025475(WO, A1)

特開平07-262068(JP, A)

特開平05-151037(JP, A)

特開平11-102311(JP, A)

特開2003-091449(JP, A)

(58)調査した分野(Int.Cl., DB名)

G06F 12/00

JSTPlus(JDreamII)