

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第5329756号  
(P5329756)

(45) 発行日 平成25年10月30日(2013.10.30)

(24) 登録日 平成25年8月2日(2013.8.2)

(51) Int.Cl. F I  
G06F 12/00 (2006.01) G06F 12/00 501M

請求項の数 12 (全 14 頁)

|               |                               |           |   |
|---------------|-------------------------------|-----------|---|
| (21) 出願番号     | 特願2006-538352 (P2006-538352)  | (73) 特許権者 | 502303739                                       |
| (86) (22) 出願日 | 平成16年10月28日 (2004.10.28)      |           | オラクル・インターナショナル・コーポレーション                         |
| (65) 公表番号     | 特表2007-510231 (P2007-510231A) |           | アメリカ合衆国カリフォルニア州94065レッドウッド・シティー, オラクル・パークウェイ500 |
| (43) 公表日      | 平成19年4月19日 (2007.4.19)        |           |   |
| (86) 国際出願番号   | PCT/US2004/036122             | (74) 代理人  | 100064746                                       |
| (87) 国際公開番号   | W02005/043414                 |           | 弁理士 深見 久郎                                       |
| (87) 国際公開日    | 平成17年5月12日 (2005.5.12)        | (74) 代理人  | 100085132                                       |
| 審査請求日         | 平成19年9月12日 (2007.9.12)        |           | 弁理士 森田 俊雄                                       |
| 審査番号          | 不服2012-15492 (P2012-15492/J1) | (74) 代理人  | 100083703                                       |
| 審査請求日         | 平成24年8月9日 (2012.8.9)          |           | 弁理士 仲村 義平                                       |
| (31) 優先権主張番号  | 10/697,070                    | (74) 代理人  | 100096781                                       |
| (32) 優先日      | 平成15年10月29日 (2003.10.29)      |           | 弁理士 堀井 豊  |
| (33) 優先権主張国   | 米国 (US)                       |           |   |

最終頁に続く

(54) 【発明の名称】 データベースにおける空間使用の追跡

(57) 【特許請求の範囲】

【請求項1】

データベースにおける空きスペースを判断するための方法であって、

第1のデータベースサーバにおいて、或る時間に、前記データベースに関連付けられる空きスペースの量を反映する第1の組の空間使用データを記憶するステップと、

前記第1のデータベースサーバにおいて、前記或る時間の後に前記第1のデータベースサーバによって前記データベースに加えられる変更に基づいて、前記第1の組の空間使用データを更新するステップと、

1つ以上の第2のデータベースサーバによって前記データベースに加えられる変更の結果として、前記データベースに関連付けられる空きスペースの量を反映するように、前記1つ以上の第2のデータベースサーバによって前記データベースに加えられる変更に基づいて、前記1つ以上の第2のデータベースサーバにおいて更新される第2の組の空間使用データを前記1つ以上の第2のデータベースサーバから前記第1のデータベースサーバが検索するステップとを含み、

前記第1のデータベースサーバは、前記第2の組の空間使用データで前記第1の組の空間使用データを更新して、前記或る時間の後に前記第1のデータベースサーバおよび前記1つ以上の第2のデータベースサーバによって前記データベースに加えられるすべての変更の結果として、前記データベースに関連付けられる最新の空きスペースの量を反映する更新された第1の組の空間使用データを生成し、前記方法はさらに、

前記更新された第1の組の空間使用データに基づいて前記データベースにおける空きス

10

20

ペースが設定可能なしきい値を上回ったかどうかを判断するステップを含む、方法。

【請求項 2】

前記第 1 の組の空間使用データおよび前記第 2 の組の空間使用データは各々、前記データベースのいくつかの部分の各々における空きスペースの量を反映する、請求項 1 に記載の方法。

【請求項 3】

前記第 1 の組の空間使用データを記憶する前記ステップは、前記第 1 の組の空間使用データのサブセットを記憶するステップを含み、前記サブセットは、前記データベース上の前記第 1 のデータベースサーバによって実行されるトランザクションに関連付けられる、請求項 1 に記載の方法。

10

【請求項 4】

前記第 1 の組の空間使用データを記憶する前記ステップは、前記第 1 の組の空間使用データを生成するために前記データベースを調べるステップを含む、請求項 1 に記載の方法。

【請求項 5】

前記第 2 の組の空間使用データを検索する前記ステップは、設定可能な期間が終了したことを判断するステップを含み、前記設定可能な期間は、前記判断するステップが最後に実行されてから、前記 1 つ以上の第 2 のデータベースサーバから前記第 2 の組の空間使用データを検索する前に待機すべき時間の量を示す、請求項 1 に記載の方法。

【請求項 6】

前記データベースにおける空きスペースが設定可能なしきい値を上回ったことを示す警告を発するステップをさらに含む、請求項 1 に記載の方法。

20

【請求項 7】

前記検索するステップ、前記更新するステップ、および前記判断するステップは、前記判断するステップが最後に実行されてから設定可能な時間の量が経過した後、順々に繰返され得る、請求項 1 に記載の方法。

【請求項 8】

前記データベースの前記いくつかの部分の各々は、前記データベースに関連付けられるテーブルスペースである、請求項 2 に記載の方法。

【請求項 9】

前記いくつかの部分の各々は、前記データベースに関連付けられるファイルである、請求項 2 に記載の方法。

30

【請求項 10】

前記第 1 のデータベースサーバは、前記第 2 の組の空間使用データで前記第 1 の組の空間使用データを更新するために、前記或る時間の後に前記 1 つ以上の第 2 のデータベースサーバによって前記データベースに加えられる変更に対応する前記第 2 の組の空間使用データにおけるエントリだけを考慮する、請求項 1 に記載の方法。

【請求項 11】

データベースにおける空間の使用を判断するための命令の 1 つ以上のシーケンスを記憶するコンピュータ読取可能な不揮発性媒体であって、1 つ以上のプロセッサによる前記命令の 1 つ以上のシーケンスの実行により、前記 1 つ以上のプロセッサに請求項 1 から 10 のいずれかに記載の方法のステップを実行させる、コンピュータ読取可能な不揮発性媒体。

40

【請求項 12】

1 つ以上のプロセッサによる実行により、前記 1 つ以上のプロセッサに請求項 1 から 10 のいずれかに記載の方法のステップを実行させる、プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

発明の分野

50

この発明は、データベースにおける空間の使用の決定に関する。

【背景技術】

【0002】

発明の背景

データベース内に記憶された使用可能なデータは、1つ以上のテーブルスペースに論理的に記憶され得、特定のテーブルスペースに関連付けられる1つ以上のデータファイルに物理的に記憶され得る。テーブルスペースは、テーブルまたはテーブル区画などのデータベースオブジェクトのためのデータを記憶するのに用いられる1つ以上のデータファイルの組である。データベースメタデータは、テーブルスペースと、テーブルスペースがデータを記憶するデータベースオブジェクトとを規定する。データベース管理者はテーブルスペースを用いて、以下のうちの1つ以上を行ない得る。すなわち、データベースに記憶されたデータのためのディスクスペース割当の制御、データベースユーザのための特定のスペース割当分の割当、オンラインまたはオフラインで個々のテーブルスペースを獲得することによるデータの可用性の制御、部分的なデータベースバックアップまたは回復動作の実行、および性能を向上させるためのデータ記憶アクセス装置の割当、である。

10

【0003】

データファイル（または単純に「ファイル」）は、特定のテーブルスペースに関連付けられ、データを物理的にテーブルスペースに記憶する。データファイルは、1つ以上の論理記憶装置で構成される。たとえば、データファイルは1つ以上のエクステンツ（extents）で構成されてもよい。各エクステンツは1つ以上のデータブロックの論理的な群である。データブロックは、データファイル内にありデータベースデータを記憶するのに用いられる連続した記憶空間の割当である。テーブルスペースにおける各データブロックのサイズは同じである。データブロックの典型的なサイズは2 K、4 K、8 Kまたは16 Kである。

20

【0004】

データベースオブジェクトに割当てられないデータファイル内の記憶空間は空きスペースである。いずれのオブジェクトにも割当てられないエクステンツの数はファイルの第1のブロックに記憶される。テーブルスペースにおけるすべてのデータファイルにわたる空きスペースの量を合計することにより、テーブルスペースにおいて空きスペースがどれだけ利用可能であるかを決定することができる。

30

【0005】

テーブルスペースは決まったサイズであり得る。データベース管理者は、新しいテーブルスペースを作成するとき、新しいテーブルスペース内にデータがどれだけ記憶されるか、またはテーブルスペース内に記憶されるデータの増加率を知ることができない。それにもかかわらず、データベース管理者は、テーブルスペースの作成時に当該テーブルスペースについての決まったサイズを選択する。

【0006】

テーブルスペース内に記憶されるデータの量が或るサイズまたはしきい値を上回る場合、テーブルスペースが空間を使い果たしてしまわないことを確実にするために、テーブルスペースのサイズを大きくすることが有利である。加えて、或る最小レベルの空きスペースをテーブルスペースに存在させることにより、最適性能でのテーブルスペースへのデータの記憶が確実に実行される。

40

【0007】

現在、データベースにおいて利用可能な空きスペースの量は、(a)データベースにテーブルスペースがいくつあるか、(b)データベースにおける各テーブルスペースにデータファイルがいくつあるか、および(c)データベースにおける各テーブルスペース内の各データファイルにデータブロックがいくつあるか判断するためにデータベースを定期的に調べることにより、データベースサーバによって決定される。データファイルは、当該データファイル内の割当てられたデータブロックの数についての情報を含む。データベースにおける各テーブルスペース内の各データファイルのための空きスペースは、データベ

50

ースにおけるテーブルスペース内の空きスペースの量が特定のしきい値を下回るかどうか判断するために合計される。データベースにおける空きスペースの量を決定するこの技術は、以後、データベースの「ポーリング」と称される。

【発明の開示】

【発明が解決しようとする課題】

【0008】

データベースにおけるテーブルスペース内の空きスペースの量が特定のしきい値を下回る場合、データベースサーバはデータベース管理者に警告を発し得る。警告が受取られると、データベース管理者はテーブルスペースのサイズを大きくして、そのテーブルスペースにおける空きスペースを増やし得る。

10

【0009】

この方策は、不所望には、テーブルスペース内の空きスペースの量を決定するためにかなりの量の時間と計算リソースとを必要とする。たとえば、テーブルスペースが100個のデータファイルを含んでいたとしても、テーブルスペースにおける空きスペースの量が最後に決定されてから割当てられたかまたは割当て解除されたデータブロックには10個のデータファイルしか関連付けられず、テーブルスペースにおいて空きスペースの量が最後に決定されて以来当該データファイルのうちの90個が変化しなかったとしても、上述の方策は、依然として、テーブルスペースにおける空きスペースを決定するために100個のデータファイルの各々を調べることを必要とするだろう。

【0010】

したがって、当該技術において、先の方策に関連付けられる問題を招くことなくデータベースにおける空間の使用を決定する必要性がまだ対処されないまま存在する。

20

【0011】

この部分に記載される方策は、実行可能な方策であるが、必ずしも以前に考案されたかまたは実行された方策であるとは限らない。したがって、特に指定のない限り、この部分に記載された方策がいずれも、単にこの部分に含まれることによるのみ先行技術として認められるものであると考えられるべきではない。

【0012】

この発明は、同様の参照番号が同様の要素を指す添付の図面において、限定のためではなく例示のために示される。

30

【発明を実施するための最良の形態】

【0013】

発明の詳細な説明

データベースにおける空間の使用を決定するための方法および装置を説明する。以下の記載においては、説明する目的で、この発明を完全に理解できるようにするために多数の特定の詳細が述べられる。しかしながら、これらの特定の詳細なしにこの発明が実施可能であることが明らかとなるだろう。他の場合には、周知の構造および装置は、この発明を不必要に曖昧にすることを避けるためにブロック図の形で示される。

【0014】

アーキテクチャ概要

40

図1は、この発明の実施例に従ったデータベース管理システム100を示すブロックネットワーク図である。データベース管理システム100は、データベースにおける各テーブルスペース内の各データファイルを調べることなくデータベースの空間の使用を正確に決定するのに用いられ得る。加えて、この明細書中に記載されるとおり、データベース管理システム100を用いて他の利点が達成され得る。図1のデータベース管理システム100は、データベース110、データベースサーバ120Aおよび120B、ならびに通信リンク130および132を含む。

【0015】

データベース110などのデータベースは、電子情報を永続的に記憶するためのコンピュータ化された機構である。データベースの非限定的な具体例は、リレーショナルデータ

50

ベース、オブジェクト指向データベース、多次元データベース、コンピュータの分散クラスタにおけるデータベース、およびサーバブレードのグリッドにおけるデータベースを含む。データベースサーバの分散クラスタは、引用によりこの明細書中に援用される米国特許第6,353,836号にさらに詳細に説明される。1組のデータベースサーバが実行されているグリッドが、引用によりこの明細書中に援用される米国仮特許出願連続番号第60/500,050号にさらに詳細に説明される。

【0016】

データベースサーバ120Aおよび120Bなどのデータベースサーバは、1組の統合されたソフトウェア構成要素と、プロセッサ上で当該1組の統合されたソフトウェア構成要素を実行するためのメモリおよびプロセスなどの計算リソースの割当ての組合せであり、この場合、ソフトウェアと計算リソースとの組合せはデータベースを管理するのに用いられる。データベース管理の機能の中でも、データベースサーバは、データベースへのアクセスを管理しかつ容易にし、データベースクライアントによるデータベースへのアクセスの要求を処理する。データベースサーバのクライアントは他のデータベースサーバを含み得る。2つのデータベースサーバ、すなわち、データベースサーバ120Aおよび120Bだけが図1に図示されているが、いくつかのデータベースサーバがデータベース110に動作可能に接続されてもよい。

10

【0017】

データベースサーバ、たとえばデータベースサーバ120Aは、空間使用データおよび更新プロセスを含み得る。空間使用データ、たとえば空間使用データ122Aおよび122Bは、データベースに関連付けられる空きスペースの量を反映するデータを指す。一実施例においては、空間使用データは、空間使用データが記憶されているデータベースサーバによってデータベースに加えられ変更に基づいて更新される。空間使用データは、「データベースにおける空間の使用の決定(“DETERMINING THE USAGE OF SPACE IN A DATABASE”)」と題された以下の段落においてより詳細に説明される。

20

【0018】

更新プロセス、たとえば更新プロセス124Aおよび124Bは、この明細書中で用いられるとおり、(a)データベース110および1つ以上のデータベースサーバから空間使用データを検索し、(b)更新プロセスが当該検索された空間使用データで配置されているデータベースサーバに記憶された空間使用データを更新することのできる1つ以上のソフトウェア構成要素の組を指す。更新プロセスの動作は、「データベースにおける空間の使用の決定」と題された以下の段落においてさらに詳細に説明される。

30

【0019】

通信リンク130は、データベース110とデータベースサーバ、たとえばデータベースサーバ120Aまたはデータベースサーバ120Bとの間でデータの交換を行なう如何なる媒体または機構によっても実現され得る。通信リンク132は、データベースサーバ間、たとえば、データベースサーバ120Aとデータベースサーバ120Bとの間でデータの交換を行なう如何なる媒体または機構によっても実現され得る。通信リンク130および132の例は、ローカルエリアネットワーク(LAN)、ワイドエリアネットワーク(WAN)、イーサネット(登録商標)もしくはインターネット、または1つ以上の地上、衛星もしくは無線のリンクなどのネットワークを含むがこれらに限定されない。

40

【0020】

機能概要

一実施例に従うと、データベース内の1つ以上のテーブルスペース内における空間の使用は、この発明の実施例を用いて決定され得る。第1の組の空間使用データは第1のデータベースサーバに記憶される。一実施例においては、第1の組の空間使用データは、データベース内における1つ以上のテーブルスペースに関連付けられる空きスペースの量を反映する。一実施例においては、第1の組の空間使用データは、第1のデータベースサーバによってデータベースに加えられ変更に基づいて更新される。たとえば、第1の組の空間使用データは、データベースオブジェクトに割当てられたり割当て解除されたりするデ

50

ータブロックの数に基づいて更新されてもよい。

【0021】

一実施例においては、第1の組の空間使用データは、データベースを調べ、公知の技術によってデータベースの各テーブルスペース内における空間の使用を決定することによって得ることができる。

【0022】

次いで、第2の組の空間使用データが、1つ以上の他のデータベースサーバから検索される。第2の組の空間使用データはまた、データベースに関連付けられる空きスペースの量を反映する。一実施例においては、第2の組の空間使用データは、第2の組の空間使用データを記憶する特定のデータベースサーバによってデータベースに加えられる変更に基づいて更新される。第2の組の空間使用データは、第1のデータベースサーバにある更新プロセスを用いて検索され得る。

10

【0023】

次いで、第1の組の空間使用データは、第1のデータベースサーバにおいて第2の組の空間使用データで更新される。その後、第1のデータベースサーバは、更新された第1の組の空間使用データに基づいてデータベースにおける空間の使用を評価し得る。設定可能な量の時間が終了した後、(第1のデータベースサーバであり得る)データベースサーバは、第2の組の空間使用データを取得し、局所的に記憶された1組の空間使用データを更新し、当該更新された1組の空間使用データに基づいてデータベースにおける空間の使用を評価するプロセスを繰返し得る。

20

【0024】

結果として、更新された第1の組の空間使用データが、データベースのための空間の使用が最後に決定されて以来データベースになされたすべての変更を反映するので、データベースにアクセスすることなく、第1のデータベースサーバにおけるデータベースのための空間の使用を正確に決定することができる。さらに、この発明の実施例を用いると、データベースにおける空間の使用を決定する効率が高められる。というのも、データベースのために空間の使用が最後に決定されて以来変更のなかったデータベースの部分についての情報にアクセスしなくても、データベースのための空間の使用が正確に決定されるからである。

【0025】

データベースにおける空間の使用の決定

図2は、この発明の実施例に従った、データベースにおける空間の使用を決定する機能ステップを示すフローチャートである。一実施例に従うと、データベース管理システム100における各データベースサーバは、図2に示されるステップを実行することができるが、説明を簡単にするために、図2に示される機能ステップは、図2のステップを実行する単一のデータベースサーバの観点から説明される。

30

【0026】

ステップ202において、第1の組の空間使用データはデータベースサーバに記憶される。たとえば、データベースサーバ120Aは、空間使用データ122Aを記憶することによってステップ202を実行し得る。データベースサーバは、非持続性記憶装置(たとえば揮発性メモリ)または持続性記憶装置(たとえばファイルサーバもしくはデータベース)に第1の組の空間使用データを記憶することによってステップ202を実行し得る。

40

【0027】

図3は、この発明の実施例に従ったデータベースサーバ304における記憶された空間使用データ302の実体図である。図3は、データベースの部分、たとえばデータファイルまたはテーブルスペースにおける空きスペースの量を反映する空間使用データ302を示す。他の実施例は、他の方法で空きスペースの量を表わし得る。たとえば、空きスペースの量が、データベースの一部分のサイズをその部分に現在記憶されているデータの量と比較することによって計算され得るので、他の実施例は、データベースの複数の部分のサイズが容易に確認可能であればデータベースの当該複数の部分の使用された空間の量を記

50

憶し得る。

【 0 0 2 8 】

一実施例においては、第 1 の組の空間使用データは、データベースに接触し、公知の技術、たとえば、各テーブルスペースにおける各データファイルの第 1 のデータブロックを調べてそのデータファイルに関連付けられる空きスペースを決定しかつ結果を合計することによりデータベースの空間の使用を空間使用データにおいて特定される粒度のレベルに決定することによって、取得され得る。たとえば、図 2 に図示のとおり、データベースサーバ 1 2 0 A は、公知の技術を用いて通信リンク 1 3 0 を介してデータベース 1 1 0 と通信することにより、データベース 1 1 0 から空間使用データ 1 2 2 A を取得し得る。

【 0 0 2 9 】

一実施例においては、ステップ 2 0 2 において記憶された第 1 の組の空間使用データは、第 1 の組の空間使用データを記憶するデータベースサーバによってデータベースに加えられた変更に基づいて連続的に更新される。図 3 が示すとおり、各々のトランザクションがデータベースに対してデータベースサーバ 3 0 4 によって開始された後、空間使用データ 3 0 2 は、データベースによってトランザクションが処理された後にデータベースに空きスペースがどれだけ残っているかを反映するよう更新され得る。トランザクションがデータブロックを割当てるかまたは割当て解除する度に、空間使用データ 3 0 2 が更新され得る。たとえば、トランザクションが、テーブルスペース 1、データファイル 1 における 1 つ以上の新しいデータブロックを割当てする場合、図 3 に図示のとおり、そのトランザクションが処理された後に残る空きスペースの量が、空間使用データ 3 0 2 に記録されるだろう。別の例においては、トランザクションが、テーブルスペース 2、データファイル 1 における 1 つ以上のデータブロックを割当て解除する場合、トランザクションが処理された後に残っている空きスペースの量が空間使用データ 3 0 2 ( 図示せず ) に記録されるだろう。

【 0 0 3 0 】

データベースサーバがトランザクションを処理した後に残るデータベースの一部分に関連付けられる空きスペースの量は、たとえば、当該トランザクションに関連付けられるデータファイルにおける第 1 のデータブロックをデータベースサーバが調べることによって決定され得る。各データファイルに空きスペースがどれだけ残っているかについての情報は、データファイルの第 1 のデータブロックに保持される。各データファイルに空きスペースがどれだけ残っているかについての情報は、トランザクションが各データファイルに残っている空きスペースの量を変更する度に更新される。

【 0 0 3 1 】

図 3 における空間使用データ 3 0 2 を記憶するテーブルにおける各々の列は、システム変更番号 ( 以後 S C N とする ) に関連付けられる。S C N は、データベースプロセスをおのおの処理するデータベースによって割当てられる番号である。インクリメント方式でデータベースが S C N を割当てるので、より最近のトランザクションよりも古いトランザクションには、データベースによって、より低い S C N が割当てられるだろう。空間使用データ 3 0 2 の一部分に関連付けられる S C N は、空きスペースの量の変更を空間使用データのその部分に最後にもたらしたトランザクションの S C N 番号に対応する。

【 0 0 3 2 】

空間使用データ 3 0 2 は、トランザクションが多数のレベルの粒度で処理された後、データベースに空きスペースがどれだけ残っているかを記録し得る。たとえば、図 3 に示される空間使用データ 3 0 2 は、各トランザクションによってどのテーブルスペースが影響を受けたか、各トランザクションによってどのデータファイルが影響を受けたか、および、各データファイルに空きスペースがどれだけ残っているかについての情報を記録する。他の実施例は、より低いレベルの粒度で情報を記録し得るが、たとえば、この発明の実施例は、特定のデータファイルにおけるどのデータブロックがトランザクションによって変更されたか、および、そのデータブロックにおいて空きスペースがどれだけ利用可能であるかについての情報を記録する空間使用データ 3 0 2 を用い得る。この発明の他の実施例

10

20

30

40

50

は、より高いレベルの粒度で情報を記録し得るが、たとえば、この発明の実施例は、各トランザクションによってどのテーブルスペースが影響を受けたか、および、各テーブルスペースにおいて空きスペースがどれだけ残っているかについての情報を記録するだけである空間使用データ302を使用し得る。したがって、この発明の実施例は、空間使用データに情報を記録するための如何なるレベルの粒度にも限定されない。

**【0033】**

ステップ204においては、データベースサーバは、1つ以上の他のデータベースサーバから第2の組の空間使用データを検索する。第2の組の空間使用データは、第2の組の空間使用データの一部を保持するデータベースサーバによってデータベース上で実行されたトランザクションの結果として、データベース内に空きスペースがどれだけ残されているかを反映するよう各データベースサーバにおいて更新される。図2を参照すると、データベースサーバ120Aは、データベースサーバ120Bから空間使用データ122Bを検索することによってステップ204を実行し得る。一実施例においては、データベース管理システム100における各データベースサーバは、データベース管理システム100における互いのデータベースサーバから第2の組の空間使用データを得ることによってステップ204を実行する。

10

**【0034】**

一実施例においては、データベースサーバの更新プロセスは、ステップ204において1つ以上のデータベースサーバから第2の組の空間使用データを検索し得る。たとえば、更新プロセス124Aは、通信リンク132を介してデータベースサーバ120Bと通信して空間使用データ122Bを検索することによってステップ204を実行し得る。

20

**【0035】**

ステップ206において、第1の組の空間使用データは、第2の組の空間使用データで更新される。たとえば、更新プロセス124Aは、ステップ204において検索された空間使用データ122Bで空間使用データ122Aを更新し得る。一実施例においては、データベース管理システム100における各データベースサーバはステップ206を実行する。

**【0036】**

一実施例においては、第2の組の空間使用データは、第1の組の空間使用データと組合せられて、更新された第1の組の空間使用データを生成する。空間の使用の最新の図を簡単なものにするために、更新された第1の組の空間使用データは、各データファイルのためのデータファイルにおける空間の使用を変更した最大(すなわち最近の)SCNの関連付けだけを保持し得る。言い換えれば、SCN(「より古いSCN」)が、データファイル内の空間の使用に影響を及ぼした別のトランザクションよりも古いトランザクションに関連付けられる場合、そのより古いSCNは、更新された第1の組の空間使用データには反映されない。

30

**【0037】**

ステップ208において、データベースにおける空間の使用は、ステップ206において更新された第1の組の空間使用データに基づいて評価される。一実施例においては、ステップ208は、空間使用データ122Aに基づいてデータベース110に対する空間の使用を評価することにより、データベースサーバ120Aによって実行され得る。一実施例においては、データベース管理システム100における各データベースサーバがステップ208を実行する。

40

**【0038】**

以下にさらに詳細に説明されるように、この発明の実施例は、図2に示される1つ以上のステップを繰り返し実行し得る。一実施例においては、ステップ208を実行する際に、ステップ208が最後に実行されたのと少なくとも同じくらい新しい空間使用データの部分だけが分析される必要がある。たとえば、或る時間Tでステップ208が実行され、テーブルスペースに関連付けられる空きスペースが100MBであると決定されたとする。ステップ208が10分後に(時間T+10分で)実行される場合、時間Tの後に空間使

50



用データに記録されたそのテーブルスペースに対するトランザクションに対応する空間使用データにおけるエントリだけが考慮される必要がある。というのも、時間Tの後のそれらのエントリだけが、そのテーブルスペースにおける空間の使用についての現在の決定に反映されないからである。一実施例においては、空間使用データにおける各エントリに関連付けられるSCNを用いて、ステップ208が最後に実行されて以来空間使用データにおけるどのエントリが変わったかまたは更新されたかを決定し得る。また、予め処理された最大のSCNよりも大きいSCNに関連付けられるエントリだけが、ステップ208の現在の実行の際に考慮される必要がある。ステップ208の実行後、処理はステップ210に進む。

#### 【0039】

ステップ210において、データベースの空きスペースまたはそのいずれかの部分が1つ以上の設定可能なしきい値を上回るかどうかについての判断がなされる。一実施例においては、データベースサーバ120Aは、ステップ208において決定されたデータベース110の空きスペースまたはそのいずれかの部分（たとえばテーブルスペース）が設定可能なしきい値を上回るかどうか判断することによって、ステップ210を実行し得る。設定可能なしきい値は、いずれかのレベルの粒度、たとえばデータブロック、データファイル、テーブルスペース、またはデータベース全体に対して設定され得る。各々の設定可能なしきい値はさまざまな態様で表現され得るが、たとえば、設定可能なしきい値は、総容量の百分率として、または空きスペースの特定された割当てとして表現され得る。各々の設定可能なしきい値は、データベースサーバの管理者またはデータベースサーバによって設定され得る。データベースサーバが設定可能なしきい値を設定する実施例においては、データベースサーバは、データベースの総容量、データベースに記憶される情報の種類、データベースの予想される増加率、またはクラスタもしくはグリッドにおける別のデータベースのしきい値を含めて、さまざまな要因を用いて設定可能なしきい値を決定し得る。

#### 【0040】

ステップ210の判断がノーである（設定可能なしきい値を上回らない）場合、処理はステップ212に進む。ステップ212において、データベースサーバは設定可能な期間待機する。設定可能な期間は、ステップ204に進む前に待機すべき時間の量を示す。一実施例においては、ステップ212における設定可能な時間の量は10分である。待機すべき設定可能な時間の量が生じたことをデータベースサーバが検出した後、処理はステップ204に進む。ステップ212が実行される度に、データベース管理システム100における別のデータベースサーバがステップ204を実行し得る。たとえば、最初にステップ204が実行される時、データベース120Aが当該ステップを実行し得、2度目にステップ204が実行される時、データベースサーバ120Bが当該ステップを実行し得る、等である。したがって、データベースのための空間の使用は、ステップ202の実行後、データベース管理システム100における単一のデータベースサーバがデータベースにおける空間の使用をポーリングすることなく連続的に監視され得る。

#### 【0041】

一実施例においては、ステップ210の判断がイエスである（設定可能なしきい値を上回る）場合、一実施例においては、処理がステップ214に進む。ステップ214において、データベースのために空間の再利用がスケジュールされる。一実施例においては、ステップ214は、データベース110のための空間の再利用をスケジュールすることにより、データベースサーバ120Aによって実行され得る。

#### 【0042】

別の実施例においては、ステップ210の判断がイエスである（設定可能なしきい値を上回る）場合、一実施例においては、処理がステップ216に進む。ステップ216において、データベースにおけるテーブルスペース内の空間の使用が設定可能なしきい値を上回ったことをデータベース管理者に示す警告を発する。一実施例においては、ステップ216は、データベース110におけるテーブルスペース内の空間の使用が設定可能なしきい

10

20

30

40

50

値を上回ったことを示す警告をデータベースサーバ120Aがデータベース管理者に発することによって実行され得る。

【0043】

この発明の実施例は、ステップのシーケンスをさまざまな順序で実行し得る。たとえば、実施例は、図2に示されるのとは異なる順序で1つ以上のステップを実行し得るか、または、図2に示される1つ以上のステップを並行して実行し得る。たとえば、この発明の実施例は、ステップ212および214を並行して、または順々に実行し得る。したがって、図2に示されるステップのシーケンスは単に例示的なものであり、この発明の実施例は、図2に示される特定のステップのシーケンスに限定されない。

【0044】

この発明の実施例に従うと、データベース管理システム100におけるいくつかのデータベースサーバが、図2に示されるステップのシーケンスを実行してもよい。具体的には、データベースサーバの各々が図2のステップの各々を実行し得るが、データベース管理システム100における特定のデータベースサーバがステップ204を実行し、次いで、その特定のデータベースサーバが、データベース管理システム100内における他のデータベースサーバの各々から第2の組の空間使用データを検索する。

【0045】

データベース管理システム100における特定のデータベースサーバがクラッシュした場合、当該クラッシュしたデータベースサーバが、当該データベースサーバが動作不能になったのと同時に当該クラッシュしたデータベースサーバの活動を反映するよう動作可能になった後、データベース管理システム100における各データベースサーバに記憶された空間使用データを再度初期設定する必要があるかもしれない。空間使用データは、データベース管理システム100における各データベースサーバにおいて図2のステップを実行することによって再度初期設定されてもよい。

【0046】

ステップ206において得られた更新された第1の組の空間使用データ122Aは、データベース110のために空間の使用が最後に決定されて以来データベース110になされたすべての変更を反映するので、ステップ202の最初の実行後、空間の使用を決定するためにデータベース110をポーリングすることなく、データベース110のためにデータベースサーバ120Aによって空間の使用が正確に決定され得る。さらに、この発明の実施例を用いる場合、データベース100のために空間の使用が最後に決定されて以来変更のなかったデータベース110の部分についての情報を調べなくても、データベース100のための空間の使用が正確に決定される。これらの利点は、データベースサーバ120Aおよびデータベース管理システム100の両方の時間と計算リソースとを有利に節約する。

【0047】

ハードウェア概要

図4は、この発明の実施例が実現され得るコンピュータシステム400を示したブロック図である。コンピュータシステム400は、情報の通信を行なうためのバス402または他の通信機構、および、情報を処理するための、バス402に結合されるプロセッサ404を含む。コンピュータシステム400はまた、プロセッサ404によって実行されるべき情報および命令を記憶するための、バス402に結合されるランダムアクセスメモリ(RAM)または他の動的記憶装置等のメインメモリ406を含む。メインメモリ406はまた、プロセッサ404によって実行されるべき命令の実行中に一時的数値変数または他の中間情報を記憶するために用いられ得る。コンピュータシステム400はさらに、プロセッサ404のための命令および静的情報を記憶するための、バス402に結合される読出専用メモリ(ROM)408または他の静的記憶装置を含む。情報および命令を記憶するための磁気ディスクまたは光ディスク等の記憶装置410が設けられ、バス402に結合される。

【0048】

10

20

30

40

50

コンピュータシステム400は、コンピュータユーザに情報を表示するための、陰極線管(CRT)等のディスプレイ412にバス402を介して結合され得る。英数字および他のキーを含む入力装置414がバス402に結合されて、情報およびコマンド選択をプロセッサ404に伝達する。別の種類のユーザ入力装置は、方向情報およびコマンド選択をプロセッサ404に伝達し、さらにはディスプレイ412上でのカーソルの動きを制御するための、マウス、トラックボールまたはカーソル方向キー等のカーソル制御416である。この入力装置は、典型的には、2つの軸、すなわち第1の軸(たとえば、x)と第2の軸(たとえば、y)とにおいて2自由度を有し、これによって、装置が平面で位置を特定することが可能となる。

**【0049】**

この発明は、ここで説明される技術を実現するためのコンピュータシステム400の用途に関する。この発明の1つの実施例に従うと、これらの技術は、メインメモリ406内に含まれる1つ以上の命令の1つ以上のシーケンスをプロセッサ404が実行することに対応して、コンピュータシステム400によって実行される。このような命令は、記憶装置410等の別のコンピュータ読取可能な媒体からメインメモリ406へと読出され得る。メインメモリ406内に含まれる命令シーケンスを実行すると、プロセッサ404がここで説明されるプロセスステップを実行する。代替的な実施例では、ソフトウェア命令の代わりに、またはソフトウェア命令と組合せてハードワイヤード回路を用いてこの発明を実現することもできる。したがって、この発明の実施例は、ハードウェア回路とソフトウェアとの特定のいずれかの組合せに限定されない。

**【0050】**

ここで用いられる「コンピュータ読取可能な媒体」という用語は、命令をプロセッサ404に与えて実行させることに関与するいかなる媒体をも指す。このような媒体は、不揮発性媒体、揮発性媒体および送信媒体を含むがそれらに限定されない多くの形をとり得る。不揮発性媒体は、たとえば、記憶装置410等の光ディスクまたは磁気ディスクを含む。揮発性媒体は、メインメモリ406等のダイナミックメモリを含む。送信媒体は、バス402を含むワイヤを含んだ、同軸ケーブル、銅ワイヤおよび光ファイバを含む。送信媒体はまた、電波および赤外線データ通信中に生成されるような音波または光波の形をとり得る。

**【0051】**

コンピュータ読取可能な媒体の一般的な形は、たとえば、フロッピー(登録商標)ディスク、フレキシブルディスク、ハードディスク、磁気テープ、その他のいかなる磁気媒体、CD-ROM、その他のいかなる光媒体、パンチカード、紙テープ、穴のパターンを備えたその他の物理的な媒体、RAM、PROM、EPROM、FLASH-EPROM、その他いかなるメモリチップまたはカートリッジ、後に述べる搬送波、またはコンピュータがそこから読取ることができるその他いかなる媒体をも含む。

**【0052】**

さまざまな形のコンピュータ読取可能な媒体は、1つ以上の命令の1つ以上のシーケンスをプロセッサ404に搬送して実行させることに関与し得る。たとえば、命令は、最初にリモートコンピュータの磁気ディスク上で搬送され得る。リモートコンピュータは、命令をそのダイナミックメモリにロードし、モデムを用いて電話線上で命令を送信し得る。コンピュータシステム400にとってローカルなモデムは、電話線上のデータを受信し、赤外線送信機を用いてデータを赤外線信号に変換し得る。赤外線検出器が、赤外線信号で搬送されたデータを受信し得、適切な回路がデータをバス402上に置き得る。バス402は、データをメインメモリ406へと搬送し、ここから、プロセッサ404が命令を取出し実行する。メインメモリ406が受取った命令は、プロセッサ404による実行の前または後に記憶装置410に随意に記憶され得る。

**【0053】**

コンピュータシステム400はまた、バス402に結合される通信インターフェイス418を含む。通信インターフェイス418は、ローカルネットワーク422に接続される

10

20

30

40

50

ネットワークリンク 420 に対する双方向データ通信結合を提供する。たとえば、通信インターフェイス 418 は、統合サービスデジタル通信網 (ISDN) カードまたはモデムであってもよく、データ通信接続を対応する種類の電話線に提供し得る。別の例として、通信インターフェイス 418 は、互換性のある LAN にデータ通信接続をもたらすローカルエリアネットワーク (LAN) カードであり得る。ワイヤレスリンクも実現され得る。このようないずれの実現例においても、通信インターフェイス 418 は、さまざまな種類の情報を表わすデジタルデータストリームを搬送する電気信号、電磁信号または光信号を送受信する。

【0054】

ネットワークリンク 420 は、典型的には、1つ以上のネットワークを介して他のデータ装置にデータ通信をもたらす。たとえば、ネットワークリンク 420 は、ローカルネットワーク 422 経由でホストコンピュータ 424 またはインターネットサービスプロバイダ (Internet Service Provider) (ISP) 426 が作動させるデータ装置に接続をもたらし得る。ISP 426 は次いで、現在一般に「インターネット」428 と呼ばれるワールドワイドパケットデータ通信網を介してデータ通信サービスを提供する。ローカルネットワーク 422 およびインターネット 428 はともに、デジタルデータストリームを搬送する電気信号、電磁信号または光信号を用いる。さまざまなネットワークを経由する信号と、ネットワークリンク 420 上にあり、コンピュータシステム 400 との間でデジタルデータを搬送する通信インターフェイス 418 経由の信号とは、情報を移送する例示的形態の搬送波である。

【0055】

コンピュータシステム 400 は、ネットワーク、ネットワークリンク 420 および通信インターフェイス 418 を介してメッセージを送信し、プログラムコードを含むデータを受信し得る。インターネットの例では、サーバ 430 は、インターネット 428、ISP 426、ローカルネットワーク 422 および通信インターフェイス 418 を介してアプリケーションプログラムのために要求されたコードを送信し得る。

【0056】

受信されたコードは、受信されたときにプロセッサ 404 によって実行され得、および/または後の実行のために記憶装置 410 もしくは他の不揮発性記憶装置に記憶され得る。このようにして、コンピュータシステム 400 は搬送波の形でアプリケーションコードを得ることができる。

【0057】

上述の明細書では、この発明の実施例を実現例ごとに異なり得る多数の特定の詳細を参照して説明してきた。したがって、この発明が何であるか、およびこの発明を目指して出願人が何を意図しているかを独占的に示す唯一のものが、この出願から発生して特有の形態をとった請求項の組である。特有の形態であるこのような請求項は、以降の任意の補正を含んで発生する。このような請求項に含まれる用語に対してこの明細書で明示されたいかなる定義も、請求項で用いられているような用語の意味を支配するものとする。したがって、請求項に明示的に記載されていない限定、要素、特性、特徴、利点または属性は、このような請求項の範囲を決して限定しない。したがって、明細書および図面は限定的な意味ではなく例示的な意味で捉えられるべきである。

【図面の簡単な説明】

【0058】

【図 1】この発明の実施例に従った空間使用システムを示すブロックネットワーク図である。

【図 2】この発明の実施例に従ったデータベースにおける空間の使用を決定する機能ステップを示すフローチャートである。

【図 3】この発明の実施例に従った記憶された空間使用データの実体図である。

【図 4】この発明の実施例が実現され得るコンピュータシステムを示すブロック図である。

。

10

20

30

40

50

【 図 1 】

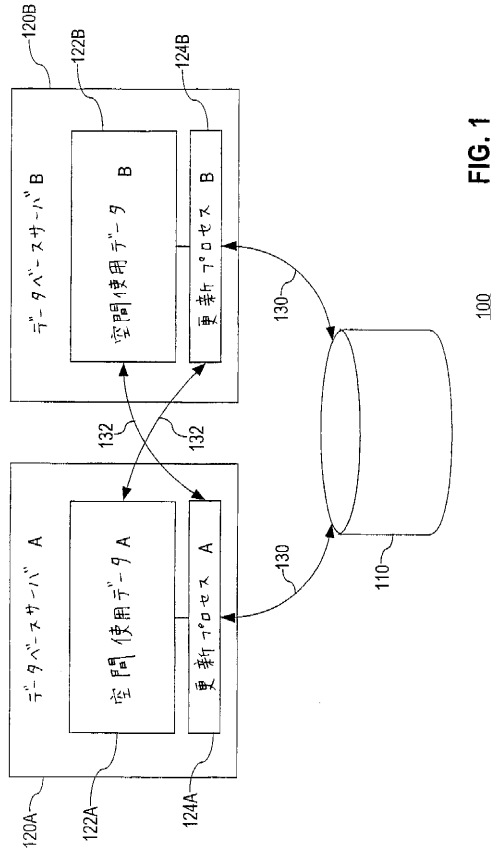


FIG. 1

【 図 2 】

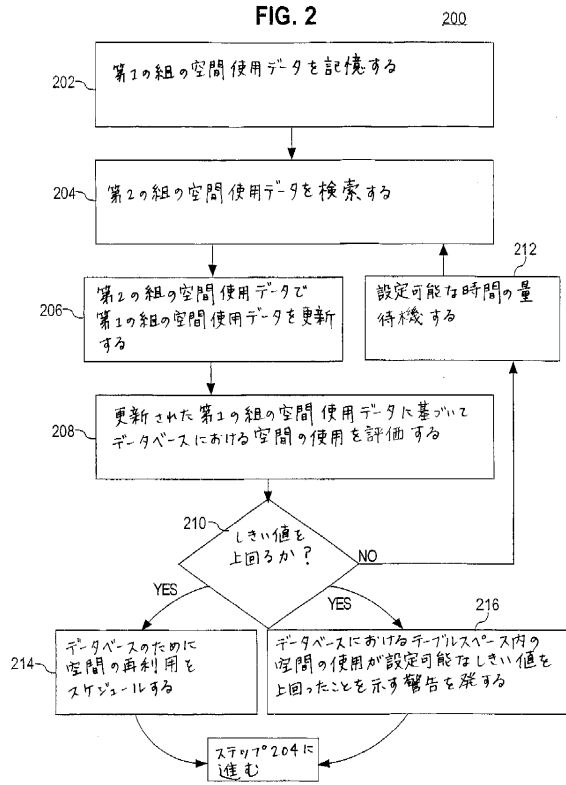


FIG. 2

【 図 3 】

データベースサーバ

| 空間使用データ |           |                |                |
|---------|-----------|----------------|----------------|
| テーブル番号  | データファイル番号 | 割当されたデータブロックの数 | システム変更番号 (SCN) |
| 1       | 1         | 40             | 230            |
| 1       | 2         | 20             | 231            |
| 2       | 5         | 30             | 232            |
| 3       | 10        | 10             | 240            |

FIG. 3

【 図 4 】

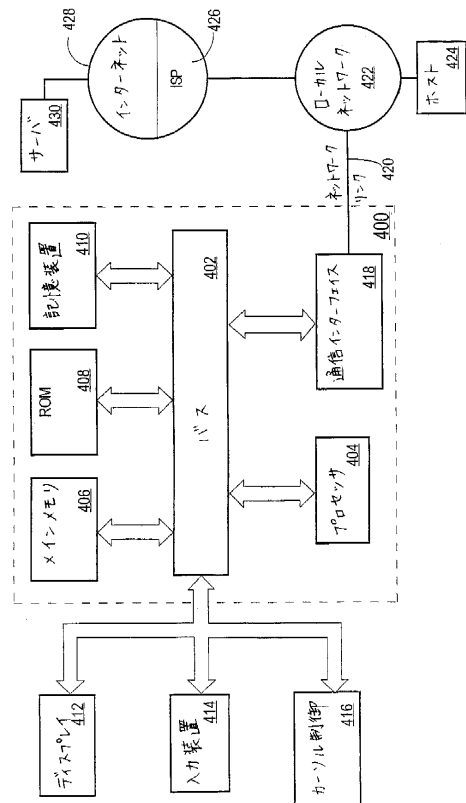


FIG. 4

---

フロントページの続き

(74)代理人 100098316

弁理士 野田 久登

(72)発明者 ハンキン, キース・アラン

アメリカ合衆国、94306 カリフォルニア州、パロ・アルト、カリフォルニア・アベニュー、133、シー・202

合議体

審判長 清水 稔

審判官 五十嵐 努

審判官 稲葉 和生

(56)参考文献 特開平9-330254(JP,A)

高橋信頼, オープンシステムの運用: 自動化・省力化で, 成長するシステムを支援, 日経オープンシステム, 日経BP社, 1995年6月15日, 第27号, p. 276~283

Frank Schmuck et al., GPFS: A Shared-Disk File System for Large Computing Clusters, Proceedings of the Conference on File and Storage Technologies (FAST '02), 米国, USENIX Association, 2002年1月28日, p. 231-244, ISBN: 1-880446-03-0, URL, <http://www.almaden.ibm.com/StorageSystems/projects/gpfs/Fast02.pdf>