(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号

特許第6107429号
(P6107429)

(45)発行日　平成29年4月5日(2017.4.5)　(24)登録日　平成29年3月17日(2017.3.17)

(51)Int.Cl.		F I		
G06F 17/30	(2006.01)	G06F	17/30	330B
G06F 12/00	(2006.01)	G06F	17/30	110C
G06F 13/00	(2006.01)	G06F	12/00	512
		G06F	13/00	540E
		G06F	17/30	340D

請求項の数 9　(全 44 頁)

(21)出願番号　特願2013-113984(P2013-113984)
(22)出願日　平成25年5月30日(2013.5.30)
(65)公開番号　特開2014-232483(P2014-232483A)
(43)公開日　平成26年12月11日(2014.12.11)
審査請求日　平成28年2月26日(2016.2.26)

(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
(74)代理人　100092152
弁理士　服部　毅巖
(72)発明者　廣瀬　厚人
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
(72)発明者　川上　俊弘
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
(72)発明者　山本　明博
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

最終頁に続く

(54)【発明の名称】データベースシステム、検索方法およびプログラム

(57)【特許請求の範囲】
【請求項１】
各々に同期されるデータを有する複数のデータベースに対応し、それぞれが前記複数のデータベースの何れかに接続された複数のサーバと、
前記複数のサーバに対して、同一の検索範囲を指定した検索要求を送信する検索要求装置と、を有し、
前記複数のサーバそれぞれは、１または２以上のキーの値を含む複数のノードが木構造に連結されたインデックス木を、検索を行うサーバの台数に応じた数のキーの値を前記複数のノードのうちのルートノードが含むように生成し、
前記複数のサーバそれぞれは、受信した前記検索要求が示す検索範囲のうち当該サーバが担当する部分検索範囲を前記インデックス木に基づいて算出し、前記部分検索範囲に限定して当該サーバに接続されたデータベースからデータを検索し、検索結果を前記検索要求装置に送信する、
データベースシステム。
【請求項２】
前記部分検索範囲は、前記複数のサーバの間で互いに重複しないように算出する、
請求項１記載のデータベースシステム。
【請求項３】
前記ルートノードは、前記サーバの台数以上の数のキーの値を含み、
前記サーバの台数が変化したとき、検索を行うサーバそれぞれは、前記サーバの台数が

変化する前に使用していた前記インデックス木の再構成を抑止する、
請求項１または２記載のデータベースシステム。
【請求項４】
前記複数のサーバそれぞれは、当該サーバの優先順位を示す情報を有しており、
前記部分検索範囲は、前記サーバの台数と前記優先順位とに基づいて算出する、
請求項１乃至３の何れか一項に記載のデータベースシステム。
【請求項５】
前記検索要求装置は、前記複数のサーバのうち一部のサーバから検索結果を受信し他のサーバから検索結果を受信していないときに、前記サーバの台数が変化したことを検出すると、前記一部のサーバの検索結果を破棄して前記検索要求を再送信する、
請求項１乃至４の何れか一項に記載のデータベースシステム。
【請求項６】
前記複数のサーバのうちの第１のサーバは、前記検索要求が示す検索範囲のうち第１の部分検索範囲を算出し、前記第１のサーバに接続された第１のデータベースから、前記第１の部分検索範囲に対応する第１のデータを検索し、
前記複数のサーバのうちの第２のサーバは、前記検索要求が示す検索範囲のうち第２の部分検索範囲を算出し、前記第２のサーバに接続され前記第１のデータベースと同じデータを記憶する第２のデータベースから、前記第１の部分検索範囲に対応する前記第１のデータを無視して前記第２の部分検索範囲に対応する第２のデータを検索する、
請求項１乃至５の何れか一項に記載のデータベースシステム。
【請求項７】
各々に同期されるデータを有する複数のデータベースに対応し、それぞれが前記複数のデータベースの何れかに接続された複数のサーバと、
前記複数のサーバに対して、同一の検索範囲を指定した検索要求を送信する検索要求装置と、を有し、
前記複数のサーバそれぞれは、１または２以上のキーの値を含む複数のノードが木構造に連結されたインデックス木にアクセス可能であり、
前記複数のサーバそれぞれは、受信した前記検索要求が示す検索範囲のうち当該サーバが担当する部分検索範囲を、前記インデックス木のルートノードまたはルートノードから所定の深さ以内のブランチノードに登録されたキーの値を基準にして前記検索要求が示す検索範囲を分割することで算出し、前記部分検索範囲に限定して当該サーバに接続されたデータベースからデータを検索し、検索結果を前記検索要求装置に送信する、
データベースシステム。
【請求項８】
データベースシステムが実行する検索方法であって、
各々に同期されるデータを有する複数のデータベースに対応しておりそれぞれが前記複数のデータベースの何れかに接続された複数のサーバそれぞれにおいて、１または２以上のキーの値を含む複数のノードが木構造に連結されたインデックス木を、検索を行うサーバの台数に応じた数のキーの値を前記複数のノードのうちのルートノードが含むように生成し、
検索要求装置から前記複数のサーバに対して、同一の検索範囲を指定した検索要求を送信し、
前記複数のサーバそれぞれにおいて、受信された前記検索要求が示す検索範囲のうち当該サーバが担当する部分検索範囲を前記インデックス木に基づいて算出し、前記部分検索範囲に限定して当該サーバに接続されたデータベースからデータを検索し、
前記複数のサーバそれぞれから前記検索要求装置に検索結果を送信する、
検索方法。
【請求項９】
各々に同期されるデータを有する複数のデータベースに対応しておりそれぞれが前記複数のデータベースの何れかに接続された複数のサーバの１つとして用いられるコンピュー

タに、
１または２以上のキーの値を含む複数のノードが木構造に連結されたインデックス木を、検索を行うサーバの台数に応じた数のキーの値を前記複数のノードのうちのルートノードが含むように生成し、
同一の検索範囲を指定した検索要求を前記複数のサーバに送信する検索要求装置から、前記検索要求を受信し、
受信した前記検索要求が示す検索範囲のうち前記コンピュータが担当する部分検索範囲を前記インデックス木に基づいて算出し、前記部分検索範囲に限定して前記コンピュータに接続されたデータベースからデータを検索し、
前記部分検索範囲の検索結果を前記検索要求装置に送信する、
処理を実行させるプログラム。

【発明の詳細な説明】
【技術分野】
【０００１】
本発明はデータベースシステム、検索方法およびプログラムに関する。
【背景技術】
【０００２】
現在、データベース管理システム（ＤＢＭＳ：Database Management System）と呼ばれるソフトウェアをサーバ上で動作させ、複数のクライアントからデータベースを利用できるようにしたクライアントサーバ型のデータベースシステムが広く利用されている。クライアントは検索条件を指定した検索要求をサーバに送信し、サーバは検索条件に該当するデータをデータベースから検索してクライアントに送信する。検索条件として、ある属性（テーブルのカラムに相当）の値の範囲が指定されることがある。
【０００３】
ここで、検索要求に対するレスポンスを高速化するため、ＤＢＭＳが動作するサーバを複数設けて並列に動作させる並列データベースシステムが考えられる。並列データベースシステムの構造としては、シェアードエブリシング（ＳＥ：Shared Everything）アーキテクチャとシェアードナッシング（ＳＮ：Shared Nothing）アーキテクチャとがある。
【０００４】
ＳＥアーキテクチャでは、複数のサーバが共通のデータベースに直接アクセスする。共通のデータベースは、例えば、複数のサーバからアクセス可能な共通の記憶装置上に実現される。一方、ＳＮアーキテクチャでは、データベースのデータを予め複数のパーティションに分割しておき、各サーバは特定のパーティションにのみアクセスする。各パーティションは、例えば、他のパーティションとは異なる記憶装置上に実現される。他のパーティションのデータには、当該他のパーティションに対応する他のサーバを介して間接的にアクセスすることになる。負荷分散の観点では、ＳＥアーキテクチャよりもＳＮアーキテクチャの方がアクセス競合しにくく、スループットを向上させやすい。
【０００５】
なお、アプリケーションサーバ（ＡＰサーバ）とデータベースサーバ（ＤＢサーバ）を含み、ＡＰサーバが複数のＣＰＵ（Central Processing Unit）コアを用いてデータを検索するデータ検索システムが提案されている。このＡＰサーバは、検索条件を複数の検索条件に分割し、複数のＣＰＵコアそれぞれで動作する検索処理部に割当てる。各検索処理部は、分割された検索条件に対応するＳＱＬ文をＤＢサーバに送信する。そして、ＡＰサーバは、分割された検索条件に対応する部分的な検索結果のデータをマージする。
【先行技術文献】
【特許文献】
【０００６】
【特許文献１】特開２０１２－５９２１５号公報
【発明の概要】

【発明が解決しようとする課題】
【０００７】
１つの検索要求に対して、複数のサーバが分担して検索を行い、当該検索要求に対するレスポンスを高速化することが考えられる。各サーバの検索範囲を限定する方法としては、ＳＮアーキテクチャのように、データベースのデータを予め複数のパーティションに分割しておく方法がある。しかし、検索を行うサーバの台数は、サーバを増設する場合やサーバが故障した場合等、並列データベースシステムの運用中に変化する可能性がある。データベースを予め分割する方法では、サーバ台数が変化したときにパーティションを再構成することになり、検索範囲の再設定の負担が大きいという問題がある。
【０００８】
１つの側面では、本発明は、サーバ台数の変化が検索の並列化に与える影響を低減できるデータベースシステム、検索方法およびプログラム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】
【０００９】
１つの態様では、各々に同期されるデータを有する複数のデータベースに対応し、それぞれが複数のデータベースの何れかに接続された複数のサーバと、複数のサーバに対して、同一の検索範囲を指定した検索要求を送信する検索要求装置と、を有するデータベースシステムが提供される。複数のサーバそれぞれは、受信した検索要求が示す検索範囲のうち当該サーバが担当する部分検索範囲を算出し、部分検索範囲に限定して当該サーバに接続されたデータベースからデータを検索し、検索結果を検索要求装置に送信する。
【００１０】
また、１つの態様では、データベースシステムが実行する検索方法が提供される。この検索方法では、検索要求装置から、各々に同期されるデータを有する複数のデータベースに対応しておりそれぞれが複数のデータベースの何れかに接続された複数のサーバに対して、同一の検索範囲を指定した検索要求を送信する。複数のサーバそれぞれにおいて、受信された検索要求が示す検索範囲のうち当該サーバが担当する部分検索範囲を算出し、部分検索範囲に限定して当該サーバに接続されたデータベースからデータを検索する。複数のサーバそれぞれから検索要求装置に検索結果を送信する。
【００１１】
また、１つの態様では、各々に同期されるデータを有する複数のデータベースに対応しておりそれぞれが複数のデータベースの何れかに接続された複数のサーバの１つとして用いられるコンピュータに実行させるプログラムが提供される。プログラムを実行するコンピュータは、同一の検索範囲を指定した検索要求を複数のサーバに送信する検索要求装置から、検索要求を受信する。受信した検索要求が示す検索範囲のうちコンピュータが担当する部分検索範囲を算出し、部分検索範囲に限定してコンピュータに接続されたデータベースからデータを検索する。部分検索範囲の検索結果を検索要求装置に送信する。
【発明の効果】
【００１２】
１つの側面では、サーバ台数の変化が検索の並列化に与える影響を低減できる。
【図面の簡単な説明】
【００１３】
【図１】第１の実施の形態のデータベースシステムを示す図である。
【図２】第２の実施の形態のシステムを示す図である。
【図３】サーバ装置のハードウェア例を示すブロック図である。
【図４】システムによる検索の例を示す図である。
【図５】非インデックス検索の場合に部分検索範囲を算出する例を示す図である。
【図６】インデックスの例を示す図である。
【図７】インデックス検索の場合に部分検索範囲を算出する例を示す図である。
【図８】インデックス検索の場合に部分検索範囲を算出する例を示す図（続き）である。
【図９】システムによる更新の例を示す図である。

【図１０】サーバ装置の故障発生時の検索処理の例を示す図である。
【図１１】サーバ装置の故障発生時の検索処理の例を示す図（続き）である。
【図１２】サーバ装置の故障発生時のインデックスの更新の例を示す図である。
【図１３】システムの機能例を示すブロック図である。
【図１４】処理要求通知の例を示す図である。
【図１５】稼働サーバリストの例を示す図である。
【図１６】処理結果通知の例を示す図である。
【図１７】稼働サーバ数と応答サーバリストの例を示す図である。
【図１８】クライアント装置による検索制御の例を示すフローチャートである。
【図１９】クライアント装置による検索制御の例を示すフローチャート（続き）である。
【図２０】サーバ装置による検索制御の例を示すフローチャートである。
【図２１】サーバ装置による検索制御の例を示すフローチャート（続き）である。
【図２２】限定範囲の算出処理の例を示すフローチャートである。
【図２３】非インデックス検索の場合の限定範囲の算出処理の例を示すフローチャートである。
【図２４】クライアント装置による更新制御の例を示すフローチャートである。
【図２５】サーバ装置による更新制御の例を示すフローチャートである。
【図２６】非インデックス検索により検索されるデータの例を示す図である。
【図２７】インデックス検索により検索されるデータの例を示す図である。
【図２８】インデックス検索により検索されるデータの例を示す図（続き）である。
【図２９】システムによる検索時間の例を示す図である。
【図３０】インデックスの変形例を示す図である。
【図３１】限定範囲の算出処理の変形例を示すフローチャートである。
【図３２】サーバ装置の機能構成の変形例を示す図である。
【図３３】システムによる検索の変形例を示す図である。
【発明を実施するための形態】
【００１４】
以下、本実施の形態について図面を参照して説明する。
［第１の実施の形態］
図１は、第１の実施の形態のデータベースシステムを示す図である。
【００１５】
第１の実施の形態のデータベースシステムは、サーバ１０，１０ａを含む複数のサーバと検索要求装置２０とを有する。サーバ１０，１０ａは、例えば、検索要求装置２０からアクセスを受け付けるサーバ装置（サーバコンピュータ等）である。検索要求装置２０は、例えば、サーバ１０，１０ａにアクセスするクライアント装置（クライアントコンピュータ等）であり、ユーザが操作する端末装置であってもよい。ただし、検索要求装置２０は、サーバ装置であってもよいし、クライアント装置からの要求に応じてサーバ１０，１０ａにアクセスする装置（中継装置や中継サーバと呼ばれるもの）であってもよい。サーバ１０，１０ａと検索要求装置２０とは、例えば、ネットワークを介して通信する。ただし、検索要求装置２０が、複数のサーバの何れかと同じ装置上に実現されてもよい。
【００１６】
以下の情報処理を行うため、サーバ１０，１０ａおよび検索要求装置２０は、プロセッサおよびメモリを有してもよい。例えば、プロセッサが、メモリに記憶されたプログラムを実行する。プロセッサには、ＣＰＵ、ＤＳＰ（Digital Signal Processor）、ＡＳＩＣ（Application Specific Integrated Circuit）、ＦＰＧＡ（Field Programmable Gate Array）等が含まれ得る。複数のプロセッサの集合（マルチプロセッサ）を、「プロセッサ」と呼ぶこともある。メモリには、ＲＡＭ（Random Access Memory）が含まれ得る。
【００１７】
複数のサーバは、データベース１１，１１ａを含む複数のデータベースに対応して設けられる。複数のデータベースは、各々に同期されるデータを有し、原則として同じ内容の

データを含むことになる。各サーバは、それら複数のデータベースの何れかに接続される。サーバ１０はデータベース１１に接続され、サーバ１０ａはデータベース１１ａに接続されている。各サーバは、当該サーバに接続されたデータベースにアクセスする一方、他のサーバに接続されたデータベースにはアクセスしない。よって、サーバ１０はデータベース１１にはアクセスするが、データベース１１ａにはアクセスしない。また、サーバ１０ａはデータベース１１ａにはアクセスするが、データベース１１にはアクセスしない。

【００１８】

各データベースは、例えば、他のデータベースとは異なる記憶装置上に実現される。データベースのデータを記憶する記憶装置は、ＨＤＤ（Hard Disk Drive）等の不揮発性の記憶装置でもよいし、ＲＡＭ等の揮発性の記憶装置でもよい。１つのサーバと１つのデータベースとが「接続している」とき、当該データベースは当該サーバの筐体内に存在してもよいし、当該サーバの筐体外に存在してもよい。後者の場合、サーバとデータベースとが、スイッチ等の通信装置を介して接続されていてもよい。

【００１９】

検索要求装置２０は、サーバ１０，１０ａを含む複数のサーバに対して、検索条件として同一の検索範囲を指定した検索要求を送信する。例えば、検索要求装置２０は、サーバ１０，１０ａに対して同じ検索要求２１を送信する。検索要求２１の送信は、マルチキャストやブロードキャストであってもよい。図１に例示した検索範囲Ｘ１≦Ｃ＜Ｘ２は、属性Ｃの値がＸ１以上且つＸ２未満であるデータを検索することを意味する。

【００２０】

複数のサーバそれぞれは、検索要求装置２０から検索要求を受信すると、受信した検索要求が示す検索範囲のうち当該サーバが担当する部分検索範囲を算出する。例えば、サーバ１０は、検索要求２１が示す検索範囲Ｘ１≦Ｃ＜Ｘ２から、部分検索範囲Ｘ１≦Ｃ＜Ｔを算出する。この部分検索範囲は、属性Ｃの値がＸ１以上且つＴ未満であるデータを検索することを意味する。一方、サーバ１０ａは、検索要求２１が示す検索範囲Ｘ１≦Ｃ＜Ｘ２から、サーバ１０とは異なる部分検索範囲Ｔ≦Ｃ＜Ｘ２を算出する。この部分検索範囲は、属性Ｃの値がＴ以上且つＸ２未満であるデータを検索することを意味する。

【００２１】

複数のサーバが算出する部分検索範囲は、好ましくは、互いに重複がないようにする。部分検索範囲の重複をなくすことで、無駄な検索処理を削減できる。このような部分検索範囲は、好ましくは、サーバ間で通信を行わず、所定のアルゴリズムに従って各サーバで独立に算出できるようにする。各サーバが他のサーバとは独立に部分検索範囲を算出することで、検索開始までのオーバヘッドを小さくできる。例えば、検索を行うサーバそれぞれに対して、複数のサーバの中での優先順位を付与しておく。各サーバは、検索を行うサーバの台数に応じて、検索要求２１が示す検索範囲を複数の部分検索範囲に分割し、当該サーバの優先順位に応じて、複数の部分検索範囲の１つを選択する。

【００２２】

そして、複数のサーバそれぞれは、算出した部分検索範囲に限定して、当該サーバに接続されたデータベースからデータを検索し、部分検索範囲に対応する検索結果を検索要求装置２０に送信する。例えば、サーバ１０は、部分検索範囲Ｘ１≦Ｃ＜Ｔに該当するデータをデータベース１１から検索し、部分検索範囲Ｘ１≦Ｃ＜Ｔの検索結果１２を検索要求装置２０に送信する。サーバ１０ａは、部分検索範囲Ｔ≦Ｃ＜Ｘ２に該当するデータをデータベース１１ａから検索し、部分検索範囲Ｔ≦Ｃ＜Ｘ２の検索結果１２ａを検索要求装置２０に送信する。検索要求装置２０は、例えば、サーバ１０から受信した検索結果１２とサーバ１０ａから受信した検索結果１２ａとをマージする。

【００２３】

上記のような並列検索では、データ全体が仮想的に複数のパーティションに分割されていると言うこともできる。例えば、データベース１１，１１ａに同じデータが記憶されていても、サーバ１０からは一部の仮想パーティションのデータのみがデータベース１１に記憶されているように見える。同様に、サーバ１０ａからは、データベース１１と異なる

一部の仮想パーティションのデータのみがデータベース１１ａに記憶されているように見える。これにより、サーバ１０，１０ａが検索を分担することができる。

【００２４】

一方で、パーティションが仮想的であるため、サーバの増設や故障等により検索を行うサーバの台数が変化しても、データベース間でデータを移動しなくてもよく、パーティションを再構成する負荷を抑制できる。例えば、各サーバがサーバ台数をパラメータとするアルゴリズムに従って部分検索範囲を算出する場合、サーバ台数の変化に応じて、当該サーバが認識する仮想パーティションが自動的に調整されることになる。

【００２５】

第１の実施の形態のデータベースシステムによれば、データを検索するときサーバ１０，１０ａが異なるデータベースにアクセスするため、共通のデータベースにアクセスする場合と比べて、アクセス競合を抑制できスループットが向上しやすくなる。また、データベース１１，１１ａのデータが同期されるため、あるサーバが故障しても、残りのサーバを用いてデータ全体へのアクセスを確保でき、耐故障性が向上する。

【００２６】

また、サーバ１０，１０ａそれぞれが、検索要求装置２０から指定された検索範囲のうち担当する部分検索範囲を算出し、部分検索範囲に限定してデータを検索するため、検索処理をサーバ１０，１０ａが効率的に分担することができる。また、検索を行うサーバの台数が変化しても、データベース１１，１１ａの間でデータを移動しなくてよく、サーバ１０，１０ａに検索処理を分散させるための再設定の負担が軽減される。

【００２７】

［第２の実施の形態］

図２は、第２の実施の形態のシステムを示す図である。

第２の実施の形態のシステムは、サーバ装置１００，１００ａ，１００ｂおよびクライアント装置２００を有する。サーバ装置１００，１００ａ，１００ｂは、第１の実施の形態のサーバ１０，１０ａの一例である。クライアント装置２００は、第１の実施の形態の検索要求装置２０の一例である。クライアント装置２００は、ネットワーク３０を経由してサーバ装置１００，１００ａ，１００ｂと接続している。なお、サーバ装置の数は、２台でもよいし、４台以上でもよい。また、各サーバ装置はクラスタ化されていてもよい。

【００２８】

クライアント装置２００は、例えば、デスクトップ型コンピュータやノート型コンピュータ等、ユーザが操作するコンピュータである。クライアント装置２００は、アプリケーションプログラムを実行する。クライアント装置２００が実行するアプリケーションプログラムは、サーバ装置１００，１００ａ，１００ｂの有するデータベースのデータを用いてアプリケーションプログラムを実行する。クライアント装置２００は、例えば、データベースのデータの更新や検索を要求する通知をサーバ装置１００，１００ａ，１００ｂへ一斉送信する。なお、クライアント装置２００は、他の端末装置にサービスを提供するアプリケーションサーバとして用いられてもよい。

【００２９】

サーバ装置１００，１００ａ，１００ｂは、データベースを管理するサーバコンピュータである。サーバ装置１００，１００ａ，１００ｂそれぞれは、個別にデータベースを有する。また、各サーバ装置のデータベースのデータは、サーバ装置間で同期がとられた状態である。また、サーバ装置１００，１００ａ，１００ｂそれぞれは、クライアント装置２００の要求に応じて、データベースのデータへのアクセス処理を実行し、実行結果をクライアント装置２００に応答する。

【００３０】

なお、各サーバ装置のデータベースは、ＨＤＤにデータを配置するディスク型データベースでもよいし、主記憶上にデータを配置して高速にアクセスを行うインメモリ型データベースでもよい。また、同期がとられたデータベースそれぞれは、各サーバ装置に対応付けられて外部に接続されたデータベースサーバ装置に記憶されていてもよい。

【００３１】
全サーバ装置のうち、１台のサーバ装置の種別には“現用”が割当てられ、他のサーバ装置の種別には“待機１”、“待機２”等が割当てられる。以下、種別が“現用”であるサーバ装置を“現用系のサーバ装置”と記載し、種別が“現用”以外であるサーバ装置を“待機系のサーバ装置”と記載する場合がある。
【００３２】
現用系のサーバ装置は、更新処理および検索処理を行う。更新処理には、例えば、データベースのデータの登録や削除等、データベースのデータを更新する処理を言う。検索処理は、データベースからデータを検索する処理を言う。また、現用系のサーバ装置は、更新処理の実行後、他のサーバ装置に自己のデータベースの更新後のデータと同期をとるよう要求する。また、各サーバ装置は、並列にデータベースからデータを検索する。
【００３３】
現用系のサーバ装置が故障した場合、待機系のサーバ装置の何れか１台のサーバ装置（例えば、種別が“待機１”であるサーバ装置）が現用系のサーバ装置となる。
図３は、サーバ装置のハードウェア例を示すブロック図である。サーバ装置１００は、プロセッサ１０１、ＲＡＭ１０２、ＨＤＤ１０３、画像信号処理部１０４、入力信号処理部１０５、ディスクドライブ１０６および通信インタフェース１０７を有する。これらのユニットは、サーバ装置１００内でバス１０８に接続されている。
【００３４】
プロセッサ１０１は、プログラムの命令を実行する演算器を含むプロセッサであり、例えばＣＰＵである。プロセッサ１０１は、ＨＤＤ１０３に記憶されているプログラムやデータの少なくとも一部をＲＡＭ１０２にロードしてプログラムを実行する。なお、プロセッサ１０１は複数のプロセッサコアを備えてもよい。また、サーバ装置１００は、複数のプロセッサを備えてもよい。また、サーバ装置１００は、複数のプロセッサまたは複数のプロセッサコアを用いて並列処理を行ってもよい。また、２以上のプロセッサの集合、ＦＰＧＡやＡＳＩＣ等の専用回路、２以上の専用回路の集合、プロセッサと専用回路の組み合わせ等を「プロセッサ」と呼んでもよい。
【００３５】
ＲＡＭ１０２は、プロセッサ１０１が実行するプログラムやプログラムから参照されるデータを一時的に記憶する揮発性メモリである。なお、サーバ装置１００は、ＲＡＭ以外の種類のメモリを備えてもよく、複数個の揮発性メモリを備えてもよい。
【００３６】
ＨＤＤ１０３は、ＯＳ（Operating System）やファームウェアやアプリケーションソフトウェア等のソフトウェアのプログラムおよびデータを記憶する不揮発性の記憶装置である。なお、サーバ装置１００は、フラッシュメモリ等の他の種類の記憶装置を備えてもよく、複数個の不揮発性の記憶装置を備えてもよい。
【００３７】
画像信号処理部１０４は、プロセッサ１０１からの命令に従って、サーバ装置１００に接続されたディスプレイ４１に画像を出力する。ディスプレイ４１としては、ＣＲＴ（Cathode Ray Tube）ディスプレイや液晶ディスプレイ等を用いることができる。
【００３８】
入力信号処理部１０５は、サーバ装置１００に接続された入力デバイス４２から入力信号を取得し、プロセッサ１０１に通知する。入力デバイス４２としては、マウスやタッチパネル等のポインティングデバイス、キーボード等を用いることができる。
【００３９】
ディスクドライブ１０６は、記録媒体４３に記録されたプログラムやデータを読み取る駆動装置である。記録媒体４３として、例えば、フレキシブルディスク（ＦＤ：Flexible Disk）やＨＤＤ等の磁気ディスク、ＣＤ（Compact Disc）やＤＶＤ（Digital Versatile Disc）等の光ディスク、光磁気ディスク（ＭＯ：Magneto-Optical disk）を使用できる。ディスクドライブ１０６は、プロセッサ１０１からの命令に従って、記録媒体４３から

読み取ったプログラムやデータをＲＡＭ１０２またはＨＤＤ１０３に格納する。

【００４０】

通信インタフェース１０７は、ネットワーク３０等のネットワークを介して他の情報処理装置（例えば、クライアント装置２００等）と通信を行う。

なお、サーバ装置１００はディスクドライブ１０６を備えていなくてもよく、専ら他の端末装置から制御される場合には、画像信号処理部１０４や入力信号処理部１０５を備えていなくてもよい。また、ディスプレイ４１や入力デバイス４２は、サーバ装置１００の筐体と一体に形成されていてもよい。

【００４１】

なお、サーバ装置１００ａ，１００ｂおよびクライアント装置２００も、サーバ装置１００と同様のハードウェアを用いて実現できる。

図４は、システムによる検索の例を示す図である。サーバ装置１００は、サーバ制御部１１０およびデータベース１２０を有する。サーバ装置１００ａは、サーバ制御部１１０ａおよびデータベース１２０ａを有する。サーバ装置１００ｂは、サーバ制御部１１０ｂおよびデータベース１２０ｂを有する。サーバ装置１００の種別に“現用”が割当てられ、サーバ装置１００ａの種別に“待機１”が割当てられ、サーバ装置１００ｂの種別に“待機２”が割当てられている。

【００４２】

データベース１２０，１２０ａ，１２０ｂは、クライアント装置２００が用いるデータを記憶する。データベース１２０，１２０ａ，１２０ｂのデータは、同期がとられた状態である。データベース１２０，１２０ａ，１２０ｂではそれぞれ、データが１または２以上のテーブルにより管理されている。

【００４３】

サーバ制御部１１０，１１０ａ，１１０ｂは、テーブルに記憶されているデータへのアクセス処理を実行し、処理結果をクライアント装置２００に送信する。検索処理の処理結果には、検索されたデータや検索件数を示す情報が含まれる。更新処理の処理結果には、更新処理の成否を示す情報が含まれる。

【００４４】

また、サーバ制御部１１０，１１０ａ，１１０ｂは、クライアント装置２００から検索処理を要求されたとき、まず、自サーバ装置の種別や検索に参加するサーバ装置の数や検索条件の示す検索範囲等に基づいて部分検索範囲を算出する。部分検索範囲は、検索条件の示す検索範囲のうち各サーバ装置が検索を担当する検索範囲である。また、部分検索範囲は、サーバ装置間で範囲が重複しないように算出される。そして、サーバ制御部１１０，１１０ａ，１１０ｂは、算出した部分検索範囲について検索処理を実行する。

【００４５】

また、サーバ制御部１１０，１１０ａ，１１０ｂは、クライアント装置２００から更新処理を要求されたとき、自サーバ装置の種別に基づいてデータの更新を行うか否か判定する。

【００４６】

クライアント装置２００は、アプリケーションソフトウェア２１０およびクライアント制御部２２０を有する。アプリケーションソフトウェア２１０は、データベース１２０，１２０ａ，１２０ｂのデータを用いて所定の情報処理を実行する。

【００４７】

クライアント制御部２２０は、サーバ装置１００，１００ａ，１００ｂにデータの検索や更新等を要求し、各サーバ装置から処理結果を受信する。そして、受信した処理結果をマージし、アプリケーションソフトウェア２１０に出力する。

【００４８】

以下、第２の実施の形態のシステムによる検索の例をステップ番号に沿って説明する。

まず、アプリケーションソフトウェア２１０は、クライアント制御部２２０にデータベースからのデータの検索を要求する（Ｓ１）。次に、クライアント装置２００は、サーバ

装置１００，１００ａ，１００ｂに、データベースからのデータの検索要求を一斉送信する。検索要求の一斉送信は、ブロードキャストやマルチキャストとして実行してもよい。例えば、テーブル名が“Ｔ０１”であるテーブルについてデータの検索を要求する（Ｓ２，Ｓ２ａ，Ｓ２ｂ）。以下、テーブル名が“Ｔ０１”であるテーブルを“テーブル（Ｔ０１）”と記載する場合がある。

【００４９】

次に、サーバ制御部１１０，１１０ａ，１１０ｂそれぞれは、自己のサーバ装置の種別と検索に参加するサーバ装置の数とに基づき部分検索範囲を算出する。例えば、サーバ制御部１１０は部分検索範囲として“部分検索範囲＃１”を算出し、サーバ制御部１１０ａは部分検索範囲として“部分検索範囲＃２”を算出し、サーバ制御部１１０ｂは部分検索範囲として“部分検索範囲＃３”を算出する。そして、サーバ制御部１１０，１１０ａ，１１０ｂそれぞれは、自己のテーブルから算出された部分検索範囲についてデータを検索する（Ｓ３，Ｓ３ａ，Ｓ３ｂ）。

【００５０】

次に、サーバ制御部１１０，１１０ａ，１１０ｂそれぞれは、検索結果をクライアント装置２００に送信する（Ｓ４，Ｓ４ａ，Ｓ４ｂ）。そして、クライアント制御部２２０は、サーバ装置１００，１００ａ，１００ｂから受信した検索結果をマージし、マージした検索結果をアプリケーションソフトウェア２１０に通知する（Ｓ５）。

【００５１】

次に、図５～８を用いて、インデックスを用いて検索する場合とインデックスを用いずに検索する場合において、部分検索範囲を算出する例について説明する。以下、インデックスを用いて検索することをインデックス検索と記載する場合がある。また、インデックスを用いずに検索することを非インデックス検索と記載する場合がある。

【００５２】

サーバ制御部１１０は、インデックス検索により検索処理を実行する場合と、非インデックス検索により検索処理を実行する場合とがある。インデックス検索により実行する場合は、例えば、検索条件に含まれているテーブルの特定の列に対応するインデックスが存在する場合が挙げられる。この際、インデックスには、Ｂ－Ｔｒｅｅインデックス等の木構造のインデックスが用いられる。非インデックス検索により実行する場合は、例えば、検索条件に含まれているテーブルの特定の列に対応するインデックスが存在しない場合が挙げられる。

【００５３】

図５は、非インデックス検索の場合に部分検索範囲を算出する例を示す図である。サーバ制御部１１０は、非インデックス検索の場合、クライアント装置２００が指定した検索条件の示す検索範囲を、検索に参加するサーバ装置の数で分割することで部分検索範囲を算出する。

【００５４】

例えば、クライアント装置２００に、サーバ装置１００，１００ａが接続されている。この状態で、クライアント制御部２２０が、テーブル（Ｔ０１）から“１０＜Ｃ０１＜１００”の検索条件を満たすデータの検索を、サーバ装置１００，１００ａそれぞれに要求したとする。“Ｃ０１”は、テーブル（Ｔ０１）の有する列である。

【００５５】

このとき、検索条件の下限値である“１０”と上限値である“１００”の中間の値は“５５”であるため、“１０＜Ｃ０１＜５５”が、部分検索範囲＃１としてサーバ制御部１１０により算出される。また、“５５≦Ｃ０１＜１００”が、部分検索範囲＃２としてサーバ制御部１１０ａにより算出される。

【００５６】

図６は、インデックスの例を示す図である。各サーバ装置は、データベースのテーブルからデータを検索する際、検索条件に含まれる列に対応するインデックスを用いる。テーブル１２１は、データベース１２０のデータを格納する。テーブル１２１は、列名が“Ｃ

０１”である列および列名が“Ｃ０２”である列を有する。以下、列名が“Ｃ０１”である列を“列（Ｃ０１）”と記載する場合がある。

【００５７】

サーバ装置１００は、例えば、テーブル１２１の列（Ｃ０１）が特定の検索条件を満たすデータをテーブル１２１から検索するとき、列（Ｃ０１）の値をキーに設定したインデックス１３１を用いる。なお、各キーには、複数の列の値の組み合わせが設定されてもよい。

【００５８】

インデックス１３１は、データベース１２０に記憶されている。インデックス１３１は、テーブル１２１に格納された列（Ｃ０１）の値の集合に対応して一意に生成される。そのため、同期がとられた複数のデータベースからは、同一のインデックスが生成される。すなわち、サーバ装置１００以外の各サーバ装置も、テーブル１２１と同期がとられたテーブルを有するため、インデックス１３１と同一のインデックスを有することになる。

【００５９】

インデックス１３１には、例えば、Ｂ－Ｔｒｅｅインデックスが用いられる。Ｂ－Ｔｒｅｅインデックスは、木構造のデータ構造であるＢ木を用いたインデックスである。

インデックス１３１は、キー＃１，＃２を含むルートノード、キー＃３を含むブランチノード、キー＃４を含むブランチノード、キー＃５を含むリーフノード、キー＃６を含むリーフノード、キー＃７を含むリーフノード、キー＃８を含むリーフノードを有する。

【００６０】

ルートノードは、木構造の頂点にあたるノードである。ルートノードは、１または２以上のキーおよび、２以上のポインタを有する。ルートノードのポインタ１つは、ブランチノード１つを指し示す。ルートノードはルートブロックと呼ばれることもある。ブランチノードは、ルートノードとリーフノードとの間にある中間のノードである。ブランチノードは、１または２以上のキー、および２以上のポインタを有する。ブランチノードのポインタ１つは、他のブランチノードまたはリーフノード１つを指し示す。ブランチノードはブランチブロックと呼ばれることもある。リーフノードは、木構造の終端にあたるノードである。リーフノードは、１または２以上のキーおよび木構造の終端を示す情報（例えば、“ＮＵＬＬ”）が設定された２以上のポインタを有する。リーフノードはリーフブロックと呼ばれることもある。

【００６１】

各ノードには、最大でｋ個のキーおよびｋ＋１個のポインタが、ポインタを先頭に交互に配置される。例えば、インデックス１３１において、ルートノードには、ポインタ＃１１、キー＃１、ポインタ＃１２およびキー＃２が順番に配置されている。また、キー＃３を含むブランチノードは、ポインタ＃１３、キー＃３およびポインタ＃１４が順番に配置されている。また、キー＃４を含むブランチノードには、ポインタ＃１５、キー＃４およびポインタ＃１６が順番に配置されている。

【００６２】

各ノードのキーの値は、昇順に並んでいる。例えば、インデックス１３１において、キー＃１には“２４”が設定され、キー＃２には“４６”が設定されている。

各キーは、キーに対応するレコードへのポインタを有する。例えば、キー＃１はレコード（Ｃ０１＝２４）へのポインタを有し、キー＃４はレコード（Ｃ０１＝３６）へのポインタを有し、キー＃７はレコード（Ｃ０１＝２９）へのポインタを有する。

【００６３】

ポインタ＃１１，＃１２，＃１３，＃１４は、１つ下の階層のノードを指し示す。各ポインタの指し示す１つ下のノードには、ポインタの前のキーより大きい値、かつ、ポインタの後ろのキーより小さい値のキーが配置される。

【００６４】

例えば、ポインタ＃１１は、キー＃３を含むブランチノードを指し示し、キー＃３にはキー＃１の値より小さい“１３”が設定されている。ポインタ＃１２は、キー＃４を含む

ブランチノードを指し示し、キー＃４にはキー＃１の値より大きくキー＃２の値より小さい“３６”が設定されている。また、ポインタ＃１３はキー＃５を含むリーフノードを指し示し、キー＃５にはキー＃３の値よりも小さい“６”が設定されている。ポインタ＃１４はキー＃６を含むリーフノードを指し示し、キー＃６にはキー＃３の値よりも大きい“１７”が設定されている。ポインタ＃１５はキー＃７を含むリーフノードを指し示し、キー＃７にはキー＃４の値よりも小さい“２９”が設定されている。ポインタ＃１６はキー＃８を含むリーフノードを指し示し、キー＃８にはキー＃４の値よりも大きい“４０”が設定されている。

【００６５】

このように、Ｂ－Ｔｒｅｅインデックスでは、キーの値の範囲が、ルートノードからリーフノードに向かって階層的に分割される。１つ下の階層のノードを指し示すポインタは、キーの値の範囲１つに対応していると言える。図６の例の場合、ルートノードにおいて、列（Ｃ０１）の値が“Ｃ０１＜２４”，“２４”，“２４＜Ｃ０１＜４６”，“４６”を含む複数の範囲に分割されていると言える。また、ブランチノードにおいて、“Ｃ０１＜２４”の範囲が更に“Ｃ０１＜１３”，“１３”，“１３＜Ｃ０１＜２４”に分割され、“２４＜Ｃ０１＜４６”の範囲が更に“２４＜Ｃ０１＜３６”，“３６”，“３６＜Ｃ０１＜４６”に分割されていると言える。インデックス１３１のルートノードのキーを部分検索範囲の境界値とすることで、各サーバ装置が担当する部分検索範囲の検索負荷（例えば、検索されるレコードの数）が、ほぼ均等になるものと期待できる。

【００６６】

１つの値を指定したインデックス検索は、例えば、次のように行う。

まず、ルートノードについて、検索する値と一致するキーがルートノードに含まれているか判定する。検索する値と一致するキーがある場合、当該キーのポインタの指し示すレコードを抽出する。

【００６７】

検索する値と一致するキーがない場合、ルートノードについて、検索する値より大きいキーのうち最も左側（最も手前）にあるキーを特定する。検索する値より大きいキーが特定された場合、特定されたキーの１つ前のポインタを選択する。検索する値より大きいキーが特定されなかった場合、末尾のポインタを選択する。

【００６８】

次に、選択したポインタが指し示すブランチノードについて、ルートノードと同様に、検索する値と一致するキーが当該ブランチノードに含まれているか判定する。検索する値と一致するキーがある場合、当該キーのポインタの指し示すレコードを抽出する。検索する値と一致するキーがない場合、ルートノードと同様、着目しているブランチノードについて、検索する値より大きいキーのうち最も左側にあるキーを特定する。検索する値より大きいキーが特定された場合、特定されたキーの１つ前のポインタを選択する。検索する値より大きいキーが特定されなかった場合、末尾のポインタを選択する。次に、選択したポインタの指し示すノードが１つ下の階層のブランチノードの場合、１つ上の階層のブランチノードと同様に、レコード抽出し、あるいは、ポインタを選択する。

【００６９】

選択したポインタが指し示すノードがリーフノードの場合、そのリーフノードについて、検索する値と一致するキーがリーフノードに含まれているか判定する。検索する値と一致するキーがある場合、当該キーのポインタの指し示すレコードを抽出する。検索する値と一致するキーがない場合、例えば、検索条件に該当するレコードが検索されなかったとして終了する。

【００７０】

例えば、レコード（Ｃ０１＝２９）をインデックス検索する場合、まず、ルートノードについて、“２９”より大きい最も左側にあるキー＃２が特定され、そのキーの１つ前のポインタ＃１２が選択される。これは、検索範囲を“２４＜Ｃ０１＜４６”に絞ることを意味する。次に、ポインタ＃１２が指し示すブランチノードについて、“２９”より大き

いキー＃４が特定され、キー＃４の１つ前のポインタ＃１５が選択される。これは、検索範囲を“２４＜Ｃ０１＜３６”に絞ることを意味する。次に、ポインタ＃１５が指し示すリーフノードについて、値が“２９”であるキー＃７が特定され、特定されたキー＃７のポインタが指し示すレコード（Ｃ０１＝２９）が抽出される。

【００７１】

インデックスを更新する場合、ルートノードからリーフノードに向かって、削除するキーを含むノードまたは追加するキーを格納すべきノードを探す。例えば、レコード（Ｃ０１＝２９）が削除された場合、キー＃１を含むルートノード、キー＃４を含むブランチノードを経由して検索されたキー＃７を削除する。また、例えば、その後、レコード（Ｃ０１＝３１）が登録されたときに“３１”であるキーを何れかのノードに追加する場合、キー＃１を含むルートノード、キー＃４を含むブランチノードを経由して、ポインタ＃１５が指し示すリーフノードに値が“３１”であるキーを追加する。

【００７２】

Ｂ－Ｔｒｅｅインデックスは、１つのノードが有するポインタの数の最大値をｊとすると、ブランチノードの有するポインタは少なくともｊ／２となる特徴を有する。例えば、１つのノードが有するポインタの数を最大“２”とすると、各ブランチノードは少なくとも１つのポインタを有することとなる。これにより、リーフノードの階層の深さが均等化され、インデックス１３１を用いたデータの検索回数が最大でも、対数オーダーの階層の深さ以内の回数となる。よって、データの検索速度が安定する。

【００７３】

なお、インデックス１３１には、例えば、Ｂ＊Ｔｒｅｅインデックスや Ｂ＋Ｔｒｅｅインデックス等、他の木構造のインデックスを用いてもよい。

図７は、インデックス検索の場合に部分検索範囲を算出する例を示す図である。各サーバ装置は、検索に参加するサーバ装置の数が２以上である場合は、ルートノードのキーの数が“検索に参加するサーバ装置の数－１”になるようにインデックスを生成する。検索に参加するサーバ装置の数が１である場合、上記の方法でインデックスを生成するとルートノードのキーの数が０になるため、ルートノードのキーの数が１以上の任意の数となるようにインデックスを生成する。また、各サーバ装置のサーバ制御部は、次のように、部分検索範囲を算出する。

【００７４】

まず、各サーバ装置のサーバ制御部は、自己のサーバ装置の種別に対応するルートノードのキーが境界値となるように、仮想パーティションに含まれるデータの範囲を算出する。仮想パーティションは、検索に参加するサーバ装置の数でデータベースのデータを分割したときのデータの集合である。データベースのデータは複数のサーバ装置の間で同期されている（複数のデータベースが同じデータを含んでいる）ことから、このパーティションはデータベースを物理的に分割したものではなく、仮想的に分割したものであると言える。各サーバ装置は、データベースのデータを検索する際、仮想パーティションに含まれるデータに限定して検索する。すなわち、検索処理の際、各サーバ装置のデータベースには、仮想パーティションの範囲に限定してデータが格納されているように見える。以下、仮想パーティションに含まれるデータの範囲を“限定範囲”と記載する場合がある。そして、各サーバ装置のサーバ制御部は、算出した限定範囲と、検索条件の示す検索範囲との重複する範囲を部分検索範囲として算出する。

【００７５】

例えば、図７では、クライアント装置２００に、サーバ装置１００，１００ａが接続されている。サーバ装置１００，１００ａそれぞれは、同期がとられたテーブル（Ｔ０１）を有する。また、インデックス１３１はサーバ装置１００に記憶され、インデックス１３１ａはサーバ装置１００ａに記憶されている。

【００７６】

また、インデックス１３１，１３１ａそれぞれのルートノードは、“２４”が設定されているキーを有する。検索に参加するサーバ装置の数は２であるため、インデックス１３

１，１３１ａそれぞれのルートノードのキーの数は、“２－１＝１”となる。

【００７７】

この状態で、テーブル（Ｔ０１）について、“１０＜Ｃ０１＜１００”の検索条件を満たすデータの検索を、クライアント制御部２２０がサーバ装置１００，１００ａへ要求したとする。すると、サーバ制御部１１０，１１０ａにより、“Ｃ０１＜２４”および“２４≦Ｃ０１”の限定範囲が算出される。例えば、“Ｃ０１＜２４”である限定範囲は、サーバ制御部１１０により算出され、“２４≦Ｃ０１”である限定範囲はサーバ制御部１１０ａにより算出される。

【００７８】

そして、算出された限定範囲と要求された検索条件の示す検索範囲との重複する範囲が、部分検索範囲としてサーバ制御部１１０，１１０ａにより算出される。例えば、要求された検索条件の示す検索範囲と、算出された“Ｃ０１＜２４”の重複する“１０＜Ｃ０１＜２４”が、部分検索範囲＃１としてサーバ制御部１１０により算出される。また、要求された検索条件の示す検索範囲と、算出された“２４≦Ｃ０１”の重複する“２４≦Ｃ０１＜１００”が、部分検索範囲＃２としてサーバ制御部１１０ａにより算出される。

【００７９】

なお、インデックス検索の場合における検索範囲の算出方法では、クライアント装置２００から要求された検索条件の示す検索範囲が、ルートノードのキーで分割された検索範囲に含まれない場合がある。例えば、図７のインデックス１３１を用いて検索条件の示す“３０＜Ｃ０１＜１００”の検索範囲を分割する場合、検索条件の示す検索範囲と、サーバ装置１００が算出した限定範囲（“Ｃ０１＜２４”）とが重複しないため、サーバ装置１００の担当する部分検索範囲が存在しないことになる。

【００８０】

この場合、少なくとも１つのキーが検索条件の示す検索範囲に含まれるブランチノードを選択し、選択したブランチノードのキーを境界値として部分検索範囲を算出してもよい。このブランチノードとしては、検索条件に該当するキーを有するブランチノードのうち、最もルートノードに近いもの（ルートノードからの深さが最も小さいもの）を選択することが好ましい。例えば、“３０＜Ｃ０１＜１００”の検索条件を指定した場合、ルートノードのキーは検索条件に含まれないため、少なくとも１つのキーが検索条件に含まれているブランチノードを選択する。その結果、“３６”が設定されたキーを含むブランチノードが選択される。そして、“３６”を境界値とした“３０＜Ｃ０１＜３６”および“３６≦Ｃ０１＜１００”の部分検索範囲が算出される。

【００８１】

なお、選択するブランチノードについて、検索条件の示す検索範囲に含まれるキーの数が多いほど、多くのサーバ装置に空でない部分検索範囲が割当てられることになる。そのため、検索に参加するサーバ装置が３以上あるとき、選択するブランチノードには、検索条件の示す検索範囲内のキーが多く含まれることが好ましい。ただし、検索範囲が過度に小さい部分検索範囲に分割されないように、選択するブランチノードを、ルートノードから所定の深さ以内にあるものに限定してもよい。

【００８２】

図８は、インデックス検索の場合に部分検索範囲を算出する例を示す図（続き）である。

例えば、図８では、クライアント装置２００に、サーバ装置１００，１００ａ，１００ｂが接続されている。サーバ装置１００，１００ａ，１００ｂそれぞれは、同期がとられたテーブル（Ｔ０１）を有する。インデックス１３１，１３１ａ，１３１ｂそれぞれは、テーブル（Ｔ０１）の列に対応するインデックスである。また、インデックス１３１はサーバ装置１００に記憶され、インデックス１３１ａはサーバ装置１００ａに記憶され、インデックス１３１ｂはサーバ装置１００ｂに記憶されている。

【００８３】

インデックス１３１，１３１ａ，１３１ｂそれぞれのルートノードは、“２４”が設定

されているキーと、“４６”が設定されているキーとを有する。検索に参加するサーバ装置の数は3であるため、インデックス１３１，１３１ａ，１３１ｂのルートノードのキーの数は、“３－１＝２”となる。

【００８４】

この状態で、テーブル（Ｔ０１）について、“１０＜Ｃ０１＜１００”の検索条件を満たすデータの検索を、クライアント制御部２２０がサーバ装置１００，１００ａ，１００ｂそれぞれへ要求したとする。すると、サーバ制御部１１０，１１０ａ，１１０ｂにより、“Ｃ０１＜２４”、“２４≦Ｃ０１＜４６”および“４６≦Ｃ０１”の限定範囲が算出される。例えば、“Ｃ０１＜２４”の限定範囲はサーバ制御部１１０により算出され、“２４≦Ｃ０１＜４６”である限定範囲はサーバ制御部１１０ａにより算出され、“４６≦Ｃ０１”である限定範囲はサーバ制御部１１０ｂにより算出される。

【００８５】

そして、算出された限定範囲と要求された検索条件の示す検索範囲との重複する範囲が、部分検索範囲としてサーバ制御部１１０，１１０ａ，１１０ｂにより算出される。例えば、要求された検索条件の示す検索範囲と、算出された“Ｃ０１＜２４”の重複する“１０＜Ｃ０１＜２４”が、部分検索範囲＃１としてサーバ制御部１１０により算出される。また、要求された検索条件の示す検索範囲と、算出された“２４≦Ｃ０１＜４６”の重複する“２４≦Ｃ０１＜４６”が、部分検索範囲＃２としてサーバ制御部１１０ａにより算出される。また、要求された検索条件の示す検索範囲と、算出された“４６≦Ｃ０１”の重複する“４６≦Ｃ０１＜１００”が、部分検索範囲＃３としてサーバ制御部１１０ｂにより算出される。

【００８６】

図５で説明した方法により、部分検索範囲を算出すると、特定の範囲にデータが偏って存在する場合、部分検索範囲の間で検索負荷が均等にならない場合がある。図６～８で説明したように、各サーバ装置は、Ｂ－Ｔｒｅｅインデックスのルートノードのキーを境界値とすることで、部分検索範囲の間の検索負荷をほぼ均等化できる。

【００８７】

図９は、システムによる更新の例を示す図である。サーバ装置１００のデータベース１２０は、インデックス１３１を記憶している。同様に、サーバ装置１００ａのデータベース１２０ａは、インデックス１３１ａを記憶し、サーバ装置１００ｂのデータベース１２０ｂは、インデックス１３１ｂを記憶している。ここでは、現用系のサーバ装置のみがクライアント装置２００からの更新要求に応答し、待機系のサーバ装置は更新要求に応答しない。以下、第２の実施の形態のシステムによる更新の例をステップ番号に沿って説明する。

【００８８】

まず、アプリケーションソフトウェア２１０は、クライアント制御部２２０にデータの更新を要求する（Ｓ１１）。クライアント制御部２２０は、サーバ装置１００，１００ａ，１００ｂにデータベースのデータの更新要求を一斉送信する（Ｓ１２，Ｓ１２ａ，Ｓ１２ｂ）。

【００８９】

次に、サーバ制御部１１０，１１０ａ，１１０ｂそれぞれは、自サーバ装置の種別に基づき更新処理を行うか判定する。具体的には、自サーバ装置が現用系である場合、更新処理を行うと判定し、自サーバ装置が待機系である場合、更新処理を行わないと判定する。図９では、現用系のサーバ装置であるサーバ装置１００がデータベース１２０のデータについて更新処理を実行する（Ｓ１３）。

【００９０】

次に、サーバ制御部１１０は、待機系の各サーバ装置にデータベース１２０のデータと同期をとるように要求する。同期の要求を受信した各サーバ装置は、自己のデータベースのデータについて、データベース１２０のデータと同期をとる。例えば、サーバ装置１００ａはデータベース１２０ａのデータについて同期をとり（Ｓ１４）、サーバ装置１００

ｂは、データベース１２０ｂのデータについて同期をとる（Ｓ１４ａ）。

【００９１】

次に、各サーバ装置は、同期がとられたデータベースのテーブルの特定の列に対応するインデックスを更新する。具体的には、サーバ装置１００はデータベース１２０のテーブルの特定の列に対応するインデックス１３１を更新し（Ｓ１５）、サーバ装置１００ａはデータベース１２０ａのテーブルの特定の列に対応するインデックス１３１ａを更新し（Ｓ１５ａ）、サーバ装置１００ｂはデータベース１２０ｂのテーブルの特定の列に対応するインデックス１３１ｂを更新する（Ｓ１５ｂ）。

【００９２】

次に、サーバ制御部１１０は、更新結果をクライアント装置２００に送信する（Ｓ１６）。そして、クライアント制御部２２０は、受信した更新結果をアプリケーションソフトウェア２１０に出力する（Ｓ１７）。

【００９３】

なお、図９では、サーバ装置１００は、データベース１２０，１２０ａ，１２０ｂの同期の後に更新結果をクライアント装置２００に送信したが、データベース１２０，１２０ａ，１２０ｂの同期の前に送信してもよい。また、インデックス１３１，１３１ａ，１３１ｂの更新は、サーバ装置１００がクライアント装置２００に更新結果を送信した後に行われてもよい。

【００９４】

次に、図１０～１２を用いて、現用系のサーバ装置が故障した場合の動作について説明する。

図１０は、サーバ装置の故障発生時の検索処理の例を示す図である。図１０では、図４と同様のシステム構成である場合について説明する。この場合、サーバ装置１００，１００ａ，１００ｂそれぞれでは、自サーバ装置の種別に応じて部分検索範囲＃１，＃２，＃３の何れかが算出される。この状態で、サーバ装置１００が故障したとする。

【００９５】

図１１は、サーバ装置の故障発生時の検索処理の例を示す図（続き）である。サーバ装置１００が故障すると、サーバ装置１００，１００ａ，１００ｂの集合がサーバ装置１００ａ，１００ｂの集合に縮退して、クライアント装置２００からの要求を処理することになる。このとき、種別が“待機１”であるサーバ装置１００ａの種別が“現用”に変更され、種別が“待機２”であるサーバ装置１００ｂの種別が“待機１”に変更される。また、サーバ装置１００ａ，１００ｂそれぞれには、自サーバ装置の種別に応じて部分検索範囲＃１，＃２の何れかが算出される。検索を行うサーバ装置の数が変わったことに伴って、各サーバ装置が認識する仮想パーティションが変わり、限定範囲が変わる。よって、同じ検索範囲を指定した検索要求が受信された場合であっても、各サーバ装置は、サーバ装置１００が故障する前とは異なる部分検索範囲を算出することになる。

【００９６】

図１２は、サーバ装置の故障発生時のインデックスの更新の例を示す図である。図１２では、図１０と同様のシステム構成である場合について説明する。この場合、検索に参加するサーバ装置の数は３台であるため、各サーバ装置のインデックスのルートノードのキーの数は、“３－１＝２”である。また、各サーバ装置のインデックスは、図１２の上側のインデックス１３１ａのような状態である。

【００９７】

この状態で、サーバ装置１００が故障した場合、クライアント装置２００と接続されている検索に参加するサーバ装置の数は、サーバ装置１００ａ，１００ｂの２台となる。そのため、各サーバ装置は、ルートノードのキーの数が“２”から“２－１＝１”となるようインデックスを更新する。その結果、各サーバ装置のインデックスは、例えば、図１２の下側のインデックス１３１ａのように更新される。

【００９８】

また、故障したサーバ装置１００が復旧した場合、検索に参加するサーバ装置の数は２

台から３台に戻るため、各サーバ装置は、ルートノードのキーの数が“１”から“２”となるようインデックスを更新する。その結果、各サーバ装置のインデックスは、例えば、図１２の下側のインデックス１３１ａから、図１２の上側のインデックス１３１ａのように更新される。

【００９９】

なお、インデックスの更新は、ノード間のキーの配置を組み替えることで実現されてもよいし、インデックスに対応するデータベース上のデータに基づいて再構築することで実現されてもよい。

【０１００】

図１３は、システムの機能例を示すブロック図である。クライアント装置２００は、アプリケーションソフトウェア２１０、クライアント制御部２２０および稼働サーバ情報記憶部２３０を有する。稼働サーバ情報記憶部２３０は、クライアント装置２００が備えるＲＡＭやＨＤＤに確保された記憶領域として実現できる。アプリケーションソフトウェア２１０およびクライアント制御部２２０は、クライアント装置２００が備えるプロセッサが実行するプログラムのモジュールとして実現できる。

【０１０１】

アプリケーションソフトウェア２１０については、図４で説明したため、説明を省略する。稼働サーバ情報記憶部２３０は、各サーバ装置へ送信した検索要求に対して、応答があったサーバ装置の一覧情報が格納される応答サーバリストを記憶する。また、稼働サーバ情報記憶部２３０は、検索に参加するサーバ装置の数を示す情報を記憶する。

【０１０２】

クライアント制御部２２０の例について、図４で説明していない点について説明し、図４で説明した点については説明を省略する。クライアント制御部２２０は、処理要求部２２１および実行結果制御部２２２を有する。処理要求部２２１は、１または２以上のテーブルについて、データの更新や検索の要求をアプリケーションソフトウェア２１０から取得する。また、処理要求部２２１は、取得した要求に基づいて処理要求通知を生成する。処理要求通知は、データベースのデータへのアクセス処理（例えば、更新処理や検索処理）の要求を示す通知である。処理要求通知には、データベースのデータへのアクセス処理に関する情報が設定される。処理要求部２２１は、生成した処理要求通知を複数のサーバ装置（サーバ装置１００等）に一斉送信する。

【０１０３】

実行結果制御部２２２は、処理結果通知をサーバ装置から受信する。処理結果通知は、処理要求部２２１が送信した処理要求通知に対する処理結果を示す通知である。受信した処理結果通知が更新処理の処理要求通知に対応するものである場合、実行結果制御部２２２は、１台のサーバ装置のみ（例えば、現用系のサーバ装置）から処理結果通知を受信する。そして、処理結果をアプリケーションソフトウェア２１０に出力する。

【０１０４】

また、受信した処理結果通知が検索処理の処理要求通知に対応するものである場合、実行結果制御部２２２は、１または２以上のサーバ装置から処理要求通知を受信する。そして、実行結果制御部２２２は、受信した処理要求通知に含まれる検索結果をマージし、マージされた検索結果をアプリケーションソフトウェア２１０に出力する。

【０１０５】

サーバ装置１００は、サーバ制御部１１０、データベース１２０および稼働サーバリスト１４０を有する。サーバ制御部１１０およびデータベース１２０の例について、図４で説明していない点について説明し、図４で説明した点については説明を省略する。データベース１２０および稼働サーバリスト１４０は、サーバ装置１００が備えるＲＡＭ１０２やＨＤＤ１０３に確保された記憶領域として実現できる。サーバ制御部１１０は、サーバ装置１００が備えるプロセッサ１０１が実行するプログラムのモジュールとして実現できる。データベース１２０は、第１の実施の形態のデータベース１１，１１ａの一例である。

【０１０６】
データベース１２０は、インデックス情報記憶部１３０を有する。インデックス情報記憶部１３０は、データベース１２０に記憶されているテーブル（例えば、テーブル１２１等）の特定の列に対応するインデックス（例えば、インデックス１３１等）を記憶する。
【０１０７】
稼働サーバリスト１４０は、稼働中のサーバ装置の識別情報のリストである。稼働サーバリスト１４０には、サーバ装置の識別情報が稼働優先度の高い順に記憶される。稼働優先度とは、予め定義されたサーバ装置間の優先順位であり、異なるサーバ装置には異なる稼働優先度が設定される。第２の実施の形態のシステムにおいて、稼働優先度が１位のサーバ装置が現用系のサーバ装置となる。また、稼働優先度が２位以下のサーバ装置が待機系のサーバ装置となる。待機系のサーバ装置が２以上ある場合、“待機系１”，“待機系２”，・・・のように、それら待機系のサーバ装置の間でも優先順位が定義されることになる。現用系のサーバ装置が故障した場合は、待機系のサーバ装置のうち稼働優先度が最も高いサーバ装置が現用系のサーバ装置となる。また、待機系１のサーバ装置が故障した場合は、待機系２のサーバ装置が待機系１に繰り上がる。すなわち、ある優先順位のサーバ装置が故障すると、それ以下のサーバ装置の優先順位が１つずつ繰り上がる。
【０１０８】
サーバ制御部１１０は、処理内容判定部１１１、データベース制御部１１２およびシステム管理部１１３を有する。
処理内容判定部１１１は、クライアント装置２００から処理要求通知を受信する。受信した処理要求通知が更新処理を要求するものである場合、次の処理を実行する。
【０１０９】
処理内容判定部１１１は、自己が現用系のサーバ装置であるとき、処理要求通知に含まれる更新処理をデータベース制御部１１２に要求する。処理内容判定部１１１は、自己が待機系のサーバ装置であるとき、更新処理を要求しない。
【０１１０】
また、受信した処理要求通知が検索処理を要求するものである場合、処理内容判定部１１１は、処理要求通知に含まれる検索条件、検索に参加するサーバ装置の数、自己の稼働優先度、および、インデックス情報記憶部１３０に記憶されているインデックスのルートノードに基づき、自サーバ装置が担当する部分検索範囲を算出する。そして、算出された部分検索範囲について検索処理をデータベース制御部１１２に要求する。
【０１１１】
データベース制御部１１２は、処理内容判定部１１１からの要求に応じて、データベース１２０のデータに対し、更新処理や検索処理を行う。また、データベース制御部１１２は、他のサーバ装置からテーブルの同期要求通知を受信する。同期要求通知には、同期をとるテーブルを示す情報が含まれる。データベース制御部１１２は、受信した同期要求通知に含まれるテーブルについて、同期要求通知の送信元のサーバ装置の有する同じテーブル名のテーブルと同期をとる。そして、データベース制御部１１２は、同期がとられたテーブルの特定の列に対応するインデックス（例えば、インデックス１３１）を更新する。
【０１１２】
システム管理部１１３は、各サーバ装置が正常に稼働しているか（故障していないか）確認するため、各サーバ装置に対し応答を要求する応答要求通知を送信する。システム管理部１１３は、応答要求通知に対する応答がない場合、その応答要求通知の送信先のサーバ装置を、故障したものと判断して稼働サーバリスト１４０から削除する。そして、システム管理部１１３は、削除後の検索に参加するサーバ装置の数に基づいて、インデックス情報記憶部１３０に格納されているインデックスを更新する。
【０１１３】
また、システム管理部１１３は、故障していたサーバ装置が復旧した場合、復旧したサーバ装置を稼働サーバリスト１４０に登録する。そして、システム管理部１１３は、登録後の検索に参加するサーバ装置の数に基づいて、インデックス情報記憶部１３０に格納さ

れているインデックスを更新する。なお、システム管理部１１３は、サーバ装置が増設された（新しいサーバ装置が追加された）とき、故障していたサーバ装置が復旧したときと同様の処理を行ってもよい。

【０１１４】

次に、図１４～１７を用いて、第２の実施の形態によるシステムが用いるテーブルまたは通知について説明する。

図１４は、処理要求通知の例を示す図である。処理要求通知５１は、データベースのデータへのアクセス処理の要求を示す通知である。処理要求通知５１は、クライアント装置２００から、複数のサーバ装置に一斉送信される。

【０１１５】

処理要求通知５１は、制御情報、並列フラグ、種別、テーブル、列および条件の項目を有する。制御情報の項目には、処理要求通知５１に含まれる文字の文字数や文字コード等、処理要求通知５１に対する処理を制御するための通知制御情報が設定される。

【０１１６】

並列フラグの項目には、検索処理を２以上のサーバ装置が並列に実行するか否かを示す情報が設定される。例えば、図４のようにサーバ装置１００，１００ａ，１００ｂが並列に検索処理を実行する場合、並列フラグの項目には“ＴＲＵＥ”が設定される。一方、１つのサーバ装置（例えば、現用系のサーバ装置）のみが検索処理を実行する場合は、“ＦＡＬＳＥ”が設定される。

【０１１７】

種別の項目には、クライアント装置２００が宛先のサーバ装置に要求する処理の種別を示す情報が設定される。例えば、検索処理を要求する場合、種別の項目には、“検索”が設定される。また、データの追加を要求する場合、種別の項目には、“追加”が設定される。また、データの削除を要求する場合、種別の項目には、“削除”が設定される。また、データの更新を要求する場合、種別の項目には、“更新”が設定される。

【０１１８】

テーブルの項目には、検索処理や更新処理の対象となるテーブルを識別する情報が設定される。列の項目には、テーブルの項目に設定されたテーブルについて、データを抽出する列（“カラム”という場合がある）または書換える列を識別するための情報が設定される。条件の項目には、検索処理や更新処理の対象とするレコードを限定するための条件が設定される。

【０１１９】

例えば、テーブル（Ｔ０１）において“１０＜Ｃ０１＜１００”の検索条件を満たすレコードの列（Ｃ０１）および列（Ｃ０２）の値の検索をクライアント装置２００が要求する場合、テーブルの項目には“Ｔ０１”が設定される。列の項目には“Ｃ０１，Ｃ０２”が設定され、条件の項目には“Ｃ０１＞１０　ＡＮＤ　Ｃ０１＜１００”が設定される。

【０１２０】

また、例えば、テーブル（Ｔ０１）において列（Ｃ０１）の値を“２０”から“１０”に更新する処理をクライアント装置２００が要求する場合、テーブルの項目には“Ｔ０１”が設定される。列の項目には“Ｃ０１＝１０”が設定され、条件の項目には“Ｃ０１＝２０”が設定される。

【０１２１】

なお、検索条件を満たすレコードについて、全ての列を取得するときは、列の項目に“＊”を設定するようにしてもよい。

また、処理要求通知は、種別、テーブル、列および条件の項目の代わりに、検索処理や更新処理を示すＳＱＬ文を含んでもよい。

【０１２２】

図１５は、稼働サーバリストの例を示す図である。稼働サーバリスト１４０は、稼働中のサーバ装置の識別情報を格納するリストである。

稼働サーバリスト１４０は、サーバの項目を有する。サーバの項目には、サーバ装置を

識別するための識別子が設定される。サーバの項目に設定される識別子は、サーバ装置の稼働優先度が高いほど上に設定される。稼働優先度に基づいてサーバ装置の種別が判定される。例えば、サーバ装置の種別は、稼働優先度が一番高い順に“現用”、“待機１”、“待機２”と判定される。以下、識別子が“ＳＶ＃Ａ”であるサーバ装置を“サーバ装置（ＳＶ＃Ａ）”と記載する。

【０１２３】

例えば、稼働サーバリスト１４０に“ＳＶ＃Ａ”、“ＳＶ＃Ｂ”、“ＳＶ＃Ｃ”の順に上から登録されているとする。この場合、サーバ装置（ＳＶ＃Ａ）の稼働優先度は“１”となり、サーバ装置（ＳＶ＃Ｂ）の稼働優先度は“２”となり、サーバ装置（ＳＶ＃Ｃ）の稼働優先度は“３”となる。そして、サーバ装置（ＳＶ＃Ａ）の種別は“現用”となり、サーバ装置（ＳＶ＃Ｂ）の種別は“待機１”となり、サーバ装置（ＳＶ＃Ｃ）の種別は“待機２”となる。

【０１２４】

なお、稼働サーバリスト１４０は、稼働優先度を示す情報を設定する項目やサーバ装置の種別を示す情報を設定する項目を、サーバの項目の他に有してもよい。また、サーバ装置の識別子として、サーバ名を用いてもよいし、ＩＰ（Internet Protocol）アドレス等のネットワークアドレスを用いてもよい。

【０１２５】

図１６は、処理結果通知の例を示す図である。処理結果通知５２は、クライアント装置２００が送信した処理要求通知５１に対する処理結果を示す通知である。

処理結果通知５２は、制御情報、直接指定フラグおよび処理結果の項目を有する。

【０１２６】

制御情報の項目には、検索に参加するサーバ装置の数や、処理結果通知５２の送信元のサーバ装置の稼働優先度や、他の制御情報（例えば、通知の文字数や文字コード等）が設定される。サーバ装置の数や稼働優先度は、処理結果通知５２を送信したサーバ装置が送信時点で認識しているものである。例えば、検索に参加するサーバ装置の数が“３”であり、処理結果通知５２の送信元のサーバ装置の稼働優先度が“１”であり、他の制御情報が“情報Ａ”である場合、制御情報の項目には、“３：１：情報Ａ”が設定される。

【０１２７】

直接指定フラグの項目には、インデックス検索を行う場合において、直接指定により検索されたか否かを示す情報が設定される。例えば、データベースのデータが直接指定により検索された場合、“ＴＲＵＥ”が設定される。また、範囲指定により検索された場合、直接指定フラグの項目には“ＦＡＬＳＥ”が設定される。直接指定とは、“Ｃ０１＝１０”等のように、検索条件において、列の値の範囲ではなく列の特定の値を直接指定することである。一方、“１＜Ｃ０１＜１０”のように、検索条件において列の値の範囲を指定することを範囲指定という。

【０１２８】

処理結果の項目には、処理結果通知５２の送信元のサーバ装置による処理結果が設定される。処理結果通知５２の送信元のサーバ装置が検索処理を実行した場合、処理結果の項目には、検索された列のデータが設定される。例えば、列（Ｃ０１）に“２０”が設定され、列（Ｃ０２）に“ａａ”が設定されたレコードと、列（Ｃ０１）に“２５”が設定され、列（Ｃ０２）に“ｂｂ”が設定されたレコードが検索されたとき、検索結果が検索結果の項目には、“（２０，ａａ），（２５，ｂｂ）”が設定される。データが検索されなかった場合は、処理結果の項目は空になるか、データが検索されなかったことを示す情報（例えば、“データ無し”）を含む。また、処理結果が異常である場合、処理結果の項目には、処理結果が異常であることを示す情報（例えば、“異常終了”）が設定される。

【０１２９】

また、処理結果通知５２の送信元のサーバ装置が更新処理を実行した場合、処理結果の項目には、更新処理の成否を示す情報が設定される。

図１７は、稼働サーバ数と応答サーバリストの例を示す図である。稼働サーバ数２３１

は、検索に参加するサーバ装置の数を示す情報である。応答サーバリスト２３２は、処理結果通知５２の送信元のサーバ装置を示すリストである。稼働サーバ数２３１および応答サーバリスト２３２は、稼働サーバ情報記憶部２３０に記憶される。

【０１３０】

稼働サーバ数２３１には、１個目の処理結果通知５２が示す検索に参加するサーバ装置の数が登録される。クライアント装置２００は、２個目以降の処理結果通知５２を受信したとき、２個目以降の処理結果通知５２の制御情報の項目に含まれる検索に参加するサーバ装置の数を示す情報と、稼働サーバ数２３１とを比較する。比較結果が一致しない場合、クライアント装置２００は、同じ処理要求通知５１に対する全サーバ装置からの処理結果通知５２が揃う前に、サーバ装置の故障や復旧により検索に参加するサーバ装置の数が変化したと判定する。クライアント装置２００は、サーバ装置の数が変化したと判定したとき、受信済みの処理結果を破棄し、処理要求通知５１を再度複数のサーバ装置に一斉送信する。

【０１３１】

また、クライアント装置２００は、検索の処理要求通知５１に対する処理結果通知５２を受信したとき、受信した処理結果通知５２の送信元のサーバ装置の稼働優先度を、応答サーバリスト２３２に登録する。サーバ装置の稼働優先度は、処理結果通知５２の制御情報の項目に含まれる。

【０１３２】

次に、図１８～２３を用いて、データベースのデータの検索処理について説明する。

図１８は、クライアント装置による検索制御の例を示すフローチャートである。以下、図１８に示す処理をステップ番号に沿って説明する。

【０１３３】

（ステップＳ２１）処理要求部２２１は、アプリケーションソフトウェア２１０から検索を要求される。検索の要求には、検索対象となるテーブルを示す情報、検索対象となるテーブルの列を示す情報および検索条件を示す情報等が含まれる。アプリケーションソフトウェア２１０が出力する上記の情報は、ＳＱＬ文として記述されていてもよい。

【０１３４】

（ステップＳ２２）処理要求部２２１は、アプリケーションソフトウェア２１０からの検索の要求に基づいて、次のように処理要求通知５１を生成する。

処理要求部２２１は、処理要求通知５１の制御情報の項目に通知制御情報を設定する。また、処理要求部２２１は、並列フラグの項目に、複数のサーバ装置に並列して検索処理をさせるかを示す情報を設定する。並列して検索処理をさせるかを示す情報は、稼働サーバ情報記憶部２３０に予め記憶されていてもよいし、アプリケーションソフトウェア２１０から指定されてもよい。また、処理要求部２２１は、種別の項目に“検索”を設定する。また、処理要求部２２１は、テーブルの項目に、検索の要求に含まれるテーブルを示す情報を設定する。また、処理要求部２２１は、列の項目に、検索の要求に含まれるテーブルの列を示す情報を設定する。また、処理要求部２２１は、条件の項目に、検索の要求に含まれる検索条件を示す情報を設定する。

【０１３５】

そして、処理要求部２２１は、生成した処理要求通知５１を複数のサーバ装置（例えば、サーバ装置１００，１００ａ，１００ｂ）に一斉送信する。

（ステップＳ２３）実行結果制御部２２２は、現在時刻を処理開始時刻として一時的に格納する。処理開始時刻の格納場所には、例えば、稼働サーバ情報記憶部２３０等の記憶領域を用いる。

【０１３６】

（ステップＳ２４）実行結果制御部２２２は、ステップＳ２２で送信した処理要求通知５１に対する１つの処理結果通知５２を、１つのサーバ装置から受信したか判定する。処理結果通知５２を受信した場合は、処理をステップＳ２５へ進める。処理結果通知５２を受信していない場合は、処理をステップＳ３２へ進める。

【０１３７】
（ステップＳ２５）実行結果制御部２２２は、処理結果通知５２が異常終了を示すか判定する。異常終了を示すか否かは、例えば、処理結果通知５２の処理結果の項目に“異常終了”が設定されているかで判断する。処理結果通知５２が異常終了を示す場合、処理をステップＳ３６へ進める。処理結果通知５２が異常終了を示さない場合、処理をステップＳ２６へ進める。
【０１３８】
（ステップＳ２６）実行結果制御部２２２は、処理結果通知５２の処理結果の項目に設定されている検索されたデータをクライアント装置２００の備えるＲＡＭ等の記憶領域に一時的に格納する。
【０１３９】
（ステップＳ２７）実行結果制御部２２２は、受信した処理結果通知５２が処理要求通知５１に対する１つ目の応答である場合、その処理結果通知５２が示すサーバ装置の数を稼働サーバ数２３１として登録する。稼働サーバ数２３１は、１つ目の処理結果通知５２を送信したサーバ装置が送信時点で認識していた、検索に参加するサーバ装置の数を意味する。また、実行結果制御部２２２は、処理結果通知５２の送信元のサーバ装置の稼働優先度を示す情報を応答サーバリスト２３２に登録する。送信元のサーバ装置の稼働優先度を示す情報は、処理結果通知５２の制御情報の項目に含まれる。
【０１４０】
（ステップＳ２８）実行結果制御部２２２は、処理結果通知５２の直接指定フラグの項目が“ＴＲＵＥ”であるか判定する。直接指定フラグが“ＴＲＵＥ”である場合、処理をステップＳ３９へ進める。直接指定フラグが“ＦＡＬＳＥ”である場合、処理をステップＳ２９へ進める。
【０１４１】
（ステップＳ２９）実行結果制御部２２２は、ステップＳ２２で一斉送信した処理要求通知５１において、並列フラグの項目が“ＴＲＵＥ”であるか判定する。並列フラグが“ＴＲＵＥ”である場合、処理をステップＳ３０へ進める。並列フラグが“ＦＡＬＳＥ”である場合、処理をステップＳ３９へ進める。
【０１４２】
（ステップＳ３０）実行結果制御部２２２は、検索に参加する全サーバ装置から処理結果通知５２を受信済みであるか判定する。全サーバ装置から処理結果通知５２を受信したかは、例えば、応答サーバリスト２３２に登録されているサーバ装置の数が、稼働サーバ数２３１と一致するかにより判断する。全サーバ装置から処理結果通知５２を受信した場合、処理をステップＳ３８へ進める。少なくとも１つのサーバ装置から処理結果通知５２が受信されていない場合、処理をステップＳ３１へ進める。
【０１４３】
（ステップＳ３１）実行結果制御部２２２は、ステップＳ２３で設定した処理開始時刻と現在時刻との差（処理要求通知５１を送信してからの経過時間）が閾値未満か判定する。閾値は、例えば、クライアント装置２００の備えるＲＡＭ等の記憶領域に格納されている。時刻差が閾値未満である場合、処理をステップＳ３２へ進める。時刻差が閾値以上である場合、処理をステップＳ３７へ進める。
【０１４４】
（ステップＳ３２）実行結果制御部２２２は、処理結果通知５２の受信待ちの状態であるか判定する。処理結果通知５２の受信待ちの状態であるかは、ステップＳ２４で処理結果通知５２を受信したかにより判断する。処理結果通知５２の受信待ちの状態である（ステップＳ２４で処理結果通知５２を受信しなかった）場合、処理をステップＳ３４へ進める。処理結果通知５２の受信待ちの状態でない（ステップＳ２４で処理結果通知５２を受信した）場合、処理をステップＳ３３へ進める。
【０１４５】
（ステップＳ３３）実行結果制御部２２２は、処理要求通知５１に応じて行われる並列

検索の途中で、検索に参加するサーバ装置の数が変化したか判定する。
検索に参加するサーバ装置の数は、サーバ装置の故障、サーバ装置の復旧、サーバ装置の追加等によって変化する。実行結果制御部２２２は、今回受信した処理結果通知５２が示すサーバ装置の数が、稼働サーバ数２３１と一致しないとき、検索に参加するサーバ装置の数が変化したと判定する。これは、今回受信した処理結果通知５２の送信時点で認識されていたサーバ装置の数が、１つ目の処理結果通知５２の送信時点で認識されていたサーバ装置の数から変化したことを示している。ただし、実行結果制御部２２２は、復旧したサーバ装置を示す情報を含むコマンドを、第２の実施の形態のシステムの管理者から受信したとき、サーバ装置の数が変化したと判定してもよい。

【０１４６】

検索に参加するサーバ装置の数が変化した場合、処理をステップＳ３５へ進める。検索に参加するサーバ装置の数が変化していない場合、処理をステップＳ３４へ進める。
（ステップＳ３４）実行結果制御部２２２は、一定時間（例えば、１０ミリ秒または１００ミリ秒）経過するのを待つ。一定時間は、第２の実施の形態のシステムのユーザに設定されてもよいし、クライアント装置２００の備えるＨＤＤ等の記憶領域に予め記憶されていてもよい。そして、処理をステップＳ２４へ進め、次の処理結果通知５２を受信したか判定する。

【０１４７】

（ステップＳ３５）実行結果制御部２２２は、ステップＳ２６で格納した受信済みのデータをクリアする。また、実行結果制御部２２２は、稼働サーバ数２３１および応答サーバリスト２３２に登録されている情報を全てクリアする。そして、実行結果制御部２２２は、処理をステップＳ２２へ進める。これにより、処理要求通知５１が再送される。なお、実行結果制御部２２２は、処理要求通知５１の再送後に古い処理要求通知５１に対応する処理結果通知５２を受信した場合は、その処理結果通知５２を無視する。

【０１４８】

図１９は、クライアント装置による検索制御の例を示すフローチャート（続き）である。以下、図１９に示す処理をステップ番号に沿って説明する。
（ステップＳ３６）実行結果制御部２２２は、要求された検索が異常終了した旨をアプリケーションソフトウェア２１０に通知する。なお、通知後に処理要求通知５１に対応する処理結果通知５２（未受信分の処理結果通知）を受信した場合は、その処理結果通知５２を無視する。また、実行結果制御部２２２は、ステップＳ２６で格納した受信済みのデータをクリアする。また、実行結果制御部２２２は、稼働サーバ数２３１および応答サーバリスト２３２に登録されている情報を全てクリアする。アプリケーションソフトウェア２１０は、異常終了に対応する処理を実行する。そして、クライアントの検索制御を終了する。

【０１４９】

（ステップＳ３７）実行結果制御部２２２は、要求された検索がタイムアウトした旨をアプリケーションソフトウェア２１０に通知する。なお、通知後に処理要求通知５１に対応する処理結果通知５２（未受信分の処理結果通知）を受信した場合は、その処理結果通知５２を無視する。また、実行結果制御部２２２は、ステップＳ２６で格納した受信済みのデータをクリアする。また、実行結果制御部２２２は、稼働サーバ数２３１および応答サーバリスト２３２に登録されている情報を全てクリアする。アプリケーションソフトウェア２１０は、タイムアウトに対応する処理を実行する。そして、クライアントの検索制御を終了する。

【０１５０】

（ステップＳ３８）実行結果制御部２２２は、ステップＳ２６で格納した複数のサーバ装置からのデータをマージする。例えば、実行結果制御部２２２は、複数のサーバ装置から受信したレコードのリストを連結する。

【０１５１】

（ステップＳ３９）実行結果制御部２２２は、マージしたデータをアプリケーションソ

フトウェア２１０に通知する。また、実行結果制御部２２２は、稼働サーバ数２３１および応答サーバリスト２３２に登録されている情報を全てクリアする。アプリケーションソフトウェア２１０は、通知されたデータを用いて処理を実行する。

【０１５２】

図２０は、サーバ装置による検索制御の例を示すフローチャートである。図２０～２３で説明する処理は、サーバ装置１００が実行しているものとする。また、図２０～２３では、直接指定による検索条件で非インデックス検索の要求の場合、処理要求通知５１の並列フラグの項目には、“ＦＡＬＳＥ”が設定されているものとする。以下、図２０に示す処理をステップ番号に沿って説明する。

【０１５３】

（ステップＳ４１）処理内容判定部１１１は、クライアント装置２００から処理要求通知５１を受信する。処理内容判定部１１１は、処理要求通知５１の種別の項目が“検索”であることを確認する。

【０１５４】

（ステップＳ４２）処理内容判定部１１１は、受信した処理要求通知５１の条件の項目を参照し、クライアント装置２００から要求された検索条件の示す検索範囲を確認する。

（ステップＳ４３）処理内容判定部１１１は、現在時刻を処理開始時刻に設定する。処理開始時刻は、例えば、サーバ装置１００が備えるＲＡＭ１０２に記憶される。

【０１５５】

（ステップＳ４４）処理内容判定部１１１は、検索に参加するサーバ装置の数（以下、検索サーバ装置数ａ）を取得する。検索に参加するサーバ装置の数は、例えば、稼働サーバリスト１４０に登録されているサーバ装置の数をカウントすることで取得する。

【０１５６】

（ステップＳ４５）処理内容判定部１１１は、サーバ装置１００の稼働優先度を取得する。サーバ装置１００の稼働優先度（以下、稼働優先度ｂ）は、例えば、図１５のように、稼働サーバリスト１４０に記憶されているサーバ装置１００の順位により取得できる。

【０１５７】

（ステップＳ４６）処理内容判定部１１１は、検索サーバ装置数ａおよび稼働優先度ｂに基づいて、処理結果通知５２に設定する制御情報を生成する。ただし、制御情報の生成は、処理結果通知５２を送信する直前に行ってもよい。

【０１５８】

（ステップＳ４７）処理内容判定部１１１は、受信した処理要求通知５１の並列フラグの項目が“ＴＲＵＥ”であるか判定する。並列フラグが“ＴＲＵＥ”である場合、処理をステップＳ４９へ進める。並列フラグが“ＦＡＬＳＥ”である場合、処理をステップＳ４８へ進める。

【０１５９】

（ステップＳ４８）処理内容判定部１１１は、稼働優先度ｂが“１”であるか判定する。稼働優先度ｂが“１”である場合（サーバ装置１００が現用系である場合）、サーバ装置１００のみで検索を行うと判定し、処理をステップＳ５２へ進める。稼働優先度ｂが“１”以外である場合（サーバ装置１００が待機系である場合）、他の１つのサーバ装置のみで検索を行うと判定し、処理をステップＳ６５へ進める。

【０１６０】

（ステップＳ４９）処理内容判定部１１１は、インデックス情報記憶部１３０に記憶されているインデックス、ステップＳ４２で確認した検索範囲、検索サーバ装置数ａ、および稼働優先度ｂに基づいて、クライアント装置２００から検索条件として指定された列の値について限定範囲を算出する。詳細は、図２２で説明する。

【０１６１】

（ステップＳ５０）処理内容判定部１１１は、ステップＳ４２で確認した検索範囲と、算出された限定範囲の重複する範囲が存在するか判定する。重複する範囲が存在する場合、処理をステップＳ５１へ進める。重複する範囲が存在しない場合、処理をステップＳ６

５へ進める。
【０１６２】
（ステップＳ５１）処理内容判定部１１１は、図７～８で説明したように、ステップＳ４２で確認した検索範囲と、限定範囲とが重複する範囲を部分検索範囲として算出する。
なお、インデックス検索と非インデックス検索のうち、非インデックス検索においては、図２３で説明するように、実質的に、処理要求通知５１で指定される検索範囲を分割することで限定範囲を算出している。この場合、算出される限定範囲は指定される検索範囲と明らかに重複しているため、処理内容判定部１１１は、ステップＳ５０の処理を実行しなくてもよい。また、サーバ装置１００の稼働優先度が１位でなく末尾でもない場合、図２３の処理に従って算出される限定範囲は、そのまま部分検索範囲とみなすことができる。その場合、処理内容判定部１１１は、ステップＳ５１の処理を実行しなくてもよい。
【０１６３】
（ステップＳ５２）データベース制御部１１２は、並列検索が行われる場合（ステップＳ４７のＹＥＳ）、ステップＳ５１で算出した部分検索範囲に限定して、データベース１２０からデータを検索する。また、データベース制御部１１２は、サーバ装置１００のみが検索を行う場合（ステップＳ４８のＹＥＳ）、クライアント装置２００から受信された処理要求通知５１に従って、データベース１２０からデータを検索する。検索するテーブルは、処理要求通知５１のテーブルの項目を確認する。データを抽出する列は、処理要求通知５１の列の項目を確認する。そして、処理をステップＳ６１へ進める。
【０１６４】
図２１は、サーバ装置による検索制御の例を示すフローチャート（続き）である。以下、図２１に示す処理をステップ番号に沿って説明する。
（ステップＳ６１）システム管理部１１３は、検索に参加するサーバ装置の数が変化したか判定する。検索に参加するサーバ装置の数は、サーバ装置の故障、サーバ装置の復旧、サーバ装置の追加等により変化する。
【０１６５】
システム管理部１１３は、サーバ装置の故障を次のように判定する。システム管理部１１３は、定期的に各サーバ装置に対し応答要求通知を送信する。送信先のサーバ装置のアドレスは、例えば、サーバ装置１００の備えるＨＤＤ１０３等の記憶領域に格納されている。システム管理部１１３は、応答要求通知に対する応答がない場合、送信先のサーバ装置が故障したと判定する。
【０１６６】
また、システム管理部１１３は、復旧したサーバ装置を示す情報を含むコマンドを、第２の実施の形態のシステムの管理者から受信したとき、サーバ装置が復旧したと判定する。コマンドには、復旧したサーバ装置の識別子が含まれる。ただし、システム管理部１１３は、故障したサーバ装置にも応答要求通知を定期的に送信し、応答が得られたときに当該サーバ装置が復旧したと判定してもよい。
【０１６７】
検索に参加するサーバ装置の数が変化していない場合、処理をステップＳ６５へ進める。検索に参加するサーバ装置の数が変化した場合、処理をステップＳ６２へ進める。
（ステップＳ６２）システム管理部１１３は、稼働サーバリスト１４０を更新する。例えば、ステップＳ６１でサーバ装置の故障を検出したとき、システム管理部１１３は、故障と判定されたサーバ装置を稼働サーバリスト１４０から削除する。また、例えば、ステップＳ６１でコマンドを受け付けたとき、システム管理部１１３は、当該コマンドから復旧したサーバ装置の識別子を確認し、確認した識別子を稼働サーバリスト１４０の末尾に登録する。
【０１６８】
（ステップＳ６３）データベース制御部１１２は、図２０のステップＳ４３で設定した処理開始時刻と、現在時刻との差（処理要求通知５１が受信されてからの経過時間）が閾値より大きいか判定する。時刻差が閾値より大きい場合、処理をステップＳ６４へ進める

。時刻差が閾値以下である場合、処理をステップＳ４４へ進める。なお、このサーバ側の閾値は、図１８のステップＳ３１のクライアント側の閾値とは異なってもよい。このサーバ側の閾値は、ステップＳ３１のクライアント側の閾値より小さい（クライアント装置２００よりもサーバ装置１００の方が先にタイムアウトする）ことが望ましい。

【０１６９】

（ステップＳ６４）データベース制御部１１２は、タイムアウト通知をクライアント装置２００に送信する。そして、サーバの検索制御を終了する。

（ステップＳ６５）データベース制御部１１２は、要求された検索が直接指定による検索処理か判定する。直接指定による検索処理である場合、処理をステップＳ６６へ進める。範囲指定による検索処理である場合、処理をステップＳ６８へ進める。

【０１７０】

（ステップＳ６６）データベース制御部１１２は、図２０のステップＳ５２で検索処理を実行したか判定する。検索処理を実行した場合、処理をステップＳ６７へ進める。検索処理を実行していない場合、クライアント装置２００へ応答せずに（処理結果通知５２を送信せずに）、サーバの検索制御を終了する。

【０１７１】

（ステップＳ６７）データベース制御部１１２は、処理結果通知５２の直接指定フラグの項目に“ＴＲＵＥ”を設定する。そして、処理をステップＳ６９へ進める。

（ステップＳ６８）データベース制御部１１２は、処理結果通知５２の直接指定フラグの項目に“ＦＡＬＳＥ”を設定する。

【０１７２】

（ステップＳ６９）データベース制御部１１２は、処理結果通知５２の制御情報の項目に、図２０のステップＳ４６で生成した制御情報を設定する。また、データベース制御部１１２は、処理結果通知５２の処理結果の項目に、図２０のステップＳ５２の検索処理により検索されたデータの集合を設定する。検索条件または部分検索条件に該当するレコードが１件もなかった場合、このデータの集合は空集合になる。その場合、処理結果通知５２の処理結果の項目は、空になるか、データがない旨の情報を含む。そして、データベース制御部１１２は、処理結果通知５２をクライアント装置２００に送信する。

【０１７３】

図２２は、限定範囲の算出処理の例を示すフローチャートである。図２２の処理は、ステップＳ４９で実行される。以下、図２２に示す処理をステップ番号に沿って説明する。

（ステップＳ８１）処理内容判定部１１１は、検索処理がインデックス検索か判定する。すなわち、処理要求通知５１において検索範囲が指定されている列について、インデックス１３１が作成されているか判定する。検索処理がインデックス検索である場合、処理をステップＳ８３へ進める。検索処理が非インデックス検索である場合、処理をステップＳ８２へ進める。

【０１７４】

（ステップＳ８２）処理内容判定部１１１は、ステップＳ４２で確認した検索範囲、検索サーバ装置数ａおよび稼働優先度ｂに基づいて、列の値の限定範囲をインデックス１３１を用いずに算出する。すなわち、処理内容判定部１１１は、検索サーバ装置数ａに応じて検索範囲を分割する。詳細は、図２３で説明する。

【０１７５】

（ステップＳ８３）処理内容判定部１１１は、稼働優先度ｂが“１”であるか判定する。稼働優先度ｂが“１”である場合（サーバ装置１００が現用系である場合）、処理をステップＳ８４へ進める。稼働優先度ｂが“１”以外である場合（サーバ装置１００が待機系である場合）、処理をステップＳ８５へ進める。

【０１７６】

（ステップＳ８４）処理内容判定部１１１は、限定範囲を“ｐ＜Ｒ（１）"と算出する。ｐは、処理要求通知５１の検索条件に記載された列の値を示す変数であり、以下同様とする。“Ｒ（ｑ）”は、ルートノードに含まれるキーのうち、左から（前方から）ｑ番目

のキーを示し、以下同様とする。例えば、１番目のルートノードのキーは、“Ｒ（１）”と示される。そして、限定範囲算出を終了する。

【０１７７】

（ステップＳ８５）処理内容判定部１１１は、稼働優先度ｂの値が最大か判定する。具体的には、処理内容判定部１１１は、稼働優先度ｂが検索サーバ装置数ａと一致するか判定する。稼働優先度の値が最大である場合（サーバ装置１００の稼働優先度が検索に参加するサーバ装置の中で最も低い場合）、処理をステップＳ８６へ進める。稼働優先度の値が最大でない場合、処理をステップＳ８７へ進める。

【０１７８】

（ステップＳ８６）処理内容判定部１１１は、限定範囲を“Ｒ（稼働優先度ｂ－１）≦ｐ"と算出する。そして、限定範囲算出を終了する。

（ステップＳ８７）処理内容判定部１１１は、限定範囲を“Ｒ（稼働優先度ｂ－１）≦ｐ＜Ｒ（稼働優先度ｂ）"と算出する。

【０１７９】

図２３は、非インデックス検索の場合の限定範囲の算出処理の例を示すフローチャートである。図２３の処理は、前述のステップＳ８２で実行される。以下、図２３に示す処理をステップ番号に沿って説明する。

【０１８０】

（ステップＳ９１）処理内容判定部１１１は、図２０のステップＳ４２で確認した検索範囲において、上限値ｍａｘと下限値ｍｉｎとの差ｘを算出する。

（ステップＳ９２）処理内容判定部１１１は、差ｘを検索サーバ装置数ａで割った商ｙを算出する。

【０１８１】

（ステップＳ９３）処理内容判定部１１１は、稼働優先度ｂが“１”であるか判定する。稼働優先度ｂが“１”である場合（サーバ装置１００が現用系である場合）、処理をステップＳ９４へ進める。稼働優先度ｂが“１”以外である場合（サーバ装置１００が待機系である場合）、処理をステップＳ９５へ進める。

【０１８２】

（ステップＳ９４）処理内容判定部１１１は、限定範囲を“ｐ＜ｍｉｎ＋稼働優先度ｂ＊ｙ"＝"ｐ＜ｍｉｎ＋ｙ"と算出する。そして、限定範囲算出を終了する。

（ステップＳ９５）処理内容判定部１１１は、稼働優先度ｂの値が最大か判定する。稼働優先度ｂの値が最大である場合（サーバ装置１００の稼働優先度が検索に参加するサーバ装置の中で最も低い場合）、処理をステップＳ９６へ進める。稼働優先度ｂの値が最大でない場合、処理をステップＳ９７へ進める。

【０１８３】

（ステップＳ９６）処理内容判定部１１１は、限定範囲を“ｍｉｎ＋（稼働優先度ｂ－１）＊ｙ≦ｐ"と算出する。そして、限定範囲算出を終了する。

（ステップＳ９７）処理内容判定部１１１は、限定範囲を“ｍｉｎ＋（稼働優先度ｂ－１）＊ｙ≦ｐ＜ｍｉｎ＋稼働優先度ｂ＊ｙ"と算出する。

【０１８４】

次に、図２４～２５を用いて、データベースのデータの更新処理について説明する。

図２４は、クライアント装置による更新制御の例を示すフローチャートである。以下、図２４に示す処理をステップ番号に沿って説明する。

【０１８５】

（ステップＳ１０１）処理要求部２２１は、アプリケーションソフトウェア２１０から更新を要求される。更新の要求には、更新対象となるテーブルを示す情報、更新対象となる列およびその値を示す情報、更新処理の種別を示す情報（追加、更新または削除）および更新対象のレコードの条件が含まれる。アプリケーションソフトウェア２１０が出力する上記の情報は、ＳＱＬ文として記述されていてもよい。

【０１８６】

（ステップＳ１０２）処理要求部２２１は、アプリケーションソフトウェア２１０からの更新の要求に基づいて、処理要求通知５１を生成する。処理要求通知５１は、処理要求部２２１により次のように生成される。

【０１８７】

処理要求部２２１は、処理要求通知５１の制御情報の項目に、通知制御情報を設定する。また、処理要求部２２１は、種別の項目に、要求された更新処理の種別を設定する。また、処理要求部２２１は、テーブルの項目に、更新対象のテーブルを示す情報を設定する。また、処理要求部２２１は、列の項目に、更新対象の列およびその値を示す情報を設定する。また、処理要求部２２１は、条件の項目に、更新対象のレコードの条件を示す情報を設定する。

【０１８８】

そして、処理要求部２２１は、生成した処理要求通知５１を複数のサーバ装置（例えば、サーバ装置１００，１００ａ，１００ｂ）に一斉送信する。

（ステップＳ１０３）実行結果制御部２２２は、現在時刻を処理開始時刻として一時的に格納する。

【０１８９】

（ステップＳ１０４）実行結果制御部２２２は、複数のサーバ装置の何れか１つから処理結果通知５２を受信したか判定する。何れか１つのサーバ装置から処理結果通知５２を受信した場合、処理をステップＳ１０８へ進める。何れのサーバ装置からも処理結果通知５２を受信していない場合、処理をステップＳ１０５へ進める。

【０１９０】

（ステップＳ１０５）実行結果制御部２２２は、ステップＳ１０３で設定した処理開始時刻と、現在時刻との差（処理要求通知５１を送信してからの経過時間）が閾値未満か判定する。時刻差が閾値未満である場合、処理をステップＳ１０６へ進め、処理結果通知５２を待つ。時刻差が閾値以上である場合、処理をステップＳ１０７へ進める。

【０１９１】

（ステップＳ１０６）実行結果制御部２２２は、一定時間（例えば、１０ミリ秒または１００ミリ秒）経過するのを待つ。一定時間は、第２の実施の形態のシステムのユーザに設定されてもよいし、クライアント装置２００の備えるＨＤＤ等の記憶領域に予め記憶されていてもよい。そして、処理をステップＳ１０４へ進め、処理結果通知５２を受信するのを待つ。

【０１９２】

（ステップＳ１０７）実行結果制御部２２２は、要求された更新がタイムアウトにより終了した旨をアプリケーションソフトウェア２１０に通知する。アプリケーションソフトウェア２１０は、タイムアウトに対応する処理を実行する。そして、クライアントの更新制御を終了する。

【０１９３】

（ステップＳ１０８）実行結果制御部２２２は、処理結果通知５２の処理結果の項目を参照し、ステップＳ１０２で送信した処理要求通知５１に対応する更新処理が成功したか確認する。実行結果制御部２２２は、確認した更新処理の結果をアプリケーションソフトウェア２１０に通知する。

【０１９４】

図２５は、サーバ装置による更新制御の例を示すフローチャートである。図２５で説明する処理は、サーバ装置１００が実行しているものとする。以下、図２５に示す処理をステップ番号に沿って説明する。

【０１９５】

（ステップＳ１１１）処理内容判定部１１１は、クライアント装置２００から処理要求通知５１を受信する。処理内容判定部１１１は、処理要求通知５１の種別の項目が更新処理の種別（追加、更新または削除）を示す情報であることを確認する。

【０１９６】

（ステップＳ１１２）処理内容判定部１１１は、検索サーバ装置数ａを取得する。
（ステップＳ１１３）処理内容判定部１１１は、稼働優先度ｂを取得する。
（ステップＳ１１４）処理内容判定部１１１は、検索サーバ装置数ａおよび稼働優先度ｂに基づいて、処理結果通知５２に設定する制御情報を生成する。ただし、制御情報の生成は、処理結果通知５２を送信する直前に行ってもよい。

【０１９７】

（ステップＳ１１５）処理内容判定部１１１は、稼働優先度ｂが“１”であるか判定する。稼働優先度が“１”である場合（サーバ装置１００が現用系である場合）、処理をステップＳ１１６へ進める。稼働優先度が“１”以外である場合（サーバ装置１００が待機系である場合）、処理をステップＳ１２０へ進める。

【０１９８】

（ステップＳ１１６）データベース制御部１１２は、ステップＳ１１１で受信した処理要求通知５１に基づいて、更新処理を実行する。更新するテーブルは、処理要求通知５１のテーブルの項目を確認する。更新するレコードは、処理要求通知５１の条件の項目に基づいて検索する。

【０１９９】

（ステップＳ１１７）データベース制御部１１２は、更新後のテーブルについて、同期を要求するための同期要求通知を他のサーバ装置それぞれへ送信する。例えば、同期要求通知には、データベース１２０の更新履歴が含まれる。更新履歴は、データベース制御部１１２がデータベース１２０に対して行った操作を示すコマンド（ＳＱＬ文でもよい）を含んでもよいし、更新後のデータを含んでもよい。

【０２００】

（ステップＳ１１８）データベース制御部１１２は、更新したテーブルの特定の列に対応するインデックスがインデックス情報記憶部１３０に記憶されている場合、そのインデックスを更新する。ただし、テーブルの更新内容によっては、インデックスを更新しなくてもよい場合がある。

【０２０１】

（ステップＳ１１９）データベース制御部１１２は、処理結果通知５２の制御情報の項目に、ステップＳ１１４で生成した制御情報を設定する。また、データベース制御部１１２は、処理結果通知５２の処理結果の項目に、ステップＳ１１６の更新処理の結果を示す情報（例えば、更新成功を示す情報）を設定する。

【０２０２】

データベース制御部１１２は、処理結果通知５２をクライアント装置２００に送信する。そして、サーバの更新制御を終了する。
（ステップＳ１２０）データベース制御部１１２は、受信した処理要求通知５１に応じた更新処理は行わず、現用系のサーバ装置に更新処理を任せる。そして、データベース制御部１１２は、同期要求通知を現用系のサーバ装置から受信する。

【０２０３】

（ステップＳ１２１）データベース制御部１１２は、同期要求通知が示すデータについて、データベース１２０を現用系のサーバ装置が有するデータベースと同期させる。例えば、データベース制御部１１２は、受信した同期要求通知が示す更新履歴に基づいて、現用系のサーバ装置が行ったものと同様の操作をデータベース１２０に対して行い、現用系のデータベースの状態を再現する。

【０２０４】

（ステップＳ１２２）データベース制御部１１２は、更新したテーブルの特定の列に対応するインデックスがインデックス情報記憶部１３０に記憶されている場合、そのインデックスを更新する。ただし、テーブルの更新内容によっては、インデックスを更新しなくてもよい場合がある。

【０２０５】

次に、図２６～２８を用いて、第２の実施の形態のシステムにより検索されるデータ例

について説明する。
図２６は、非インデックス検索により検索されるデータの例を示す図である。クライアント装置２００は、サーバ装置１００およびサーバ装置１００ａと接続している。サーバ装置１００はデータベース１２０を有し、サーバ装置１００ａはデータベース１２０ａを有する。データベース１２０，１２０ａそれぞれは、テーブル（Ｔ０１）を有する。テーブル（Ｔ０１）は、列（Ｃ０１）および列（Ｃ０２）を有する。
【０２０６】
以下、データベース１２０，１２０ａに格納されたレコードを示す情報を“列（Ｃ０１）の値，列（Ｃ０２）の値”のように記載する。例えば、データベース１２０，１２０ａにおいて、列（Ｃ０１）に“２０”が設定され、列（Ｃ０２）に“ｂｂ”が設定されているレコードは、“２０，ｂｂ”と記載される。
【０２０７】
データベース１２０，１２０ａは、“３，ａａ”、“２０，ｂｂ”、“２５，ｃｃ”、“７５，ｄｄ”および“２００，ｅｅ”を含むテーブル（Ｔ０１）を有する。ただし、データベース１２０，１２０ａは、列（Ｃ０１）に対応するインデックスを有していない。
【０２０８】
この状態で、クライアント装置２００は、テーブル（Ｔ０１）から“１０＜Ｃ０１＜１００”を満たすレコードを検索するよう、サーバ装置１００，１００ａに要求するものとする。すると、サーバ装置１００，１００ａは、テーブル（Ｔ０１）に対し、非インデックス検索により次のようにデータを検索する。
【０２０９】
サーバ制御部１１０は、クライアント装置２００から指定された検索範囲、検索に参加するサーバ装置の数（＝２台）およびサーバ装置１００の稼働優先度（＝“１”）に基づいて、“１０＜Ｃ０１＜５５”を部分検索範囲＃１として算出する。そして、サーバ制御部１１０は、部分検索範囲＃１に限定してデータベース１２０からデータを検索する。その結果、“２０，ｂｂ”および“２５，ｃｃ”が抽出され、サーバ装置１００の検索結果１１４としてクライアント装置２００に送信される。
【０２１０】
サーバ制御部１１０ａは、クライアント装置２００から指定された検索範囲、検索に参加するサーバ装置の数（＝２台）およびサーバ装置１００ａの稼働優先度（＝“２”）に基づいて、“５５≦Ｃ０１＜１００”を部分検索範囲＃２として算出する。そして、サーバ制御部１１０ａは、部分検索範囲＃２に限定してデータベース１２０ａからデータを検索する。その結果、“７５，ｄｄ”が抽出され、サーバ装置１００ａの検索結果１１４ａとしてクライアント装置２００に送信される。
【０２１１】
クライアント装置２００は、サーバ装置１００の検索結果１１４（“２０，ｂｂ”および“２５，ｃｃ”）と、サーバ装置１００ａの検索結果１１４ａ（“７５，ｄｄ”）とをマージする。ここでは、検索結果１１４のレコードと検索結果１１４ａのレコードとを連結して１つのリストにする。マージされた検索結果２１１は、検索要求に対する応答としてアプリケーションソフトウェア２１０に提供される。
【０２１２】
図２７は、インデックス検索により検索されるデータの例を示す図である。図２７～２８では、図２６で説明した内容と同様の点について、説明を省略する。データベース１２０は、列（Ｃ０１）に対応するインデックス１３１を有し、データベース１２０ａは、列（Ｃ０１）に対応するインデックス１３１ａを有する。インデックス１３１，１３１ａそれぞれのルートノードのキーには“２４”が設定されている。
【０２１３】
この状態で、クライアント装置２００は、図２６の場合と同様に、“１０＜Ｃ０１＜１００”を満たすレコードを検索するよう、サーバ装置１００，１００ａに要求するものとする。すると、サーバ装置１００，１００ａは、テーブル（Ｔ０１）に対し、インデック

ス検索により次のようにデータを検索する。

【０２１４】

サーバ制御部１１０は、インデックス１３１、検索に参加するサーバ装置の数（＝２台）およびサーバ装置１００の稼働優先度（＝“１”）に基づいて、限定範囲＃１として“Ｃ０１＜２４”を算出する。サーバ装置１００からは、データベース１２０に、“３，ａａ”および“２０，ｂｂ”を含み、“２５，ｃｃ”、“７５，ｄｄ”および“２００，ｅｅ”を含まない仮想パーティションが形成されているように見える。

【０２１５】

サーバ制御部１１０ａは、インデックス１３１ａ、検索に参加するサーバ装置の数（＝２台）およびサーバ装置１００ａの稼働優先度（＝“２”）に基づいて、限定範囲＃２として“２４≦Ｃ０１”を算出する。サーバ装置１００ａからは、データベース１２０ａに、“２５，ｃｃ”、“７５，ｄｄ”および“２００，ｅｅ”を含み、“３，ａａ”および“２０，ｂｂ”を含まない仮想パーティションが形成されているように見える。

【０２１６】

図２８は、インデックス検索により検索されるデータの例を示す図（続き）である。

サーバ制御部１１０は、クライアント装置２００から指定された検索範囲と、限定範囲＃１とが重複する範囲“１０＜Ｃ０１＜２４”を、部分検索範囲＃１として算出する。そして、サーバ制御部１１０は、部分検索範囲＃１に限定してデータベース１２０からデータを検索する。その結果、“２０，ｂｂ”が抽出され、サーバ装置１００の検索結果１１５としてクライアント装置２００に送信される。

【０２１７】

サーバ制御部１１０ａは、クライアント装置２００から指定された検索範囲と、限定範囲＃２とが重複する範囲“２４≦Ｃ０１＜１００”を、部分検索範囲＃２として算出する。そして、サーバ制御部１１０ａは、部分検索範囲＃２に限定してデータベース１２０ａからデータを検索する。その結果、“２５，ｃｃ”および“７５，ｄｄ”が抽出され、サーバ装置１００ａの検索結果１１５ａとしてクライアント装置２００に送信される。

【０２１８】

クライアント装置２００は、サーバ装置１００の検索結果１１５（“２０，ｂｂ”）と、サーバ装置１００ａの検索結果１１５ａ（“２５，ｃｃ”および“７５，ｄｄ”）とをマージする。ここでは、検索結果１１５のレコードと検索結果１１５ａのレコードとを連結して１つのリストにする。マージされた検索結果２１２は、検索要求に対する応答としてアプリケーションソフトウェア２１０に提供される。

【０２１９】

図２９は、システムによる検索時間の例を示す図である。図２９では、同期がとられたデータを格納している３台のサーバ装置（１台の現用系と２台の待機系）を有するシステムにおいて、クライアント装置が検索要求を送信してから検索結果を取得するまでの間の時間について説明する。

【０２２０】

図２９の上側は、第２の実施の形態のシステム以外のシステムにより、データを検索する場合の処理時間の例を示す図である。図２９の上側で説明するシステムでは、現用系のサーバ装置のみが検索処理を実行する。時間Ｔ１は、現用系のサーバ装置による検索処理の時間である。時間Ｔ２は、現用系のサーバ装置からクライアント装置に検索結果を送信するための通信時間である。このシステムによるクライアント装置上のアプリケーションソフトウェアから見た検索のレスポンス時間は、“Ｔ１＋Ｔ２”となる。

【０２２１】

図２９の下側は、第２の実施の形態のシステムにより、データを並列検索する場合の処理時間の例を示す図である。時間Ｔ３は、３台のサーバ装置それぞれによる検索処理の時間である。時間Ｔ４は、３台のサーバ装置それぞれからクライアント装置に検索結果を送信するための通信時間である。また、時間Ｔ５は、クライアント装置が３台のサーバ装置から受信した検索結果をマージする時間である。この場合、第２の実施の形態のシステム

によるクライアント装置上のアプリケーションソフトウェアから見た検索のレスポンス時間は、“Ｔ３＋Ｔ４＋Ｔ５”となる。

【０２２２】

ここで、第２の実施の形態のシステムでは、同じ検索処理を３台のサーバ装置が並列して実行しているため、時間Ｔ３は、時間Ｔ１の約１／３となる。また、検索されるデータの総量は第２の実施の形態のシステム以外のシステムと変わらないため、時間Ｔ４は時間Ｔ２とほぼ同じと考えられ、また、時間Ｔ３と比べて十分に短いと考えられる。また、データをマージする処理の時間は、時間Ｔ３と比べると、非常に短いと考えられる。そのため、第２の実施の形態のシステムにおけるレスポンス時間は、図２９の上側で説明したシステムの場合と比べ、時間Ｔ１の２／３だけ短縮される。また、各サーバ装置の検索の負荷が大きい場合は、レスポンス時間が全体として１／３程度になると期待できる。

【０２２３】

第２の実施の形態のシステムによれば、クライアント装置２００は、サーバ装置１００，１００ａ，１００ｂに対して、同じ検索条件を指定してデータの検索を要求する。次に、データの検索を要求された各サーバ装置は、自己の稼働優先度や、検索に参加するサーバ装置の数等に基づいて、検索条件の示す検索範囲のうち自サーバ装置が担当する部分検索範囲を算出し、算出した部分検索範囲に限定して検索処理を行い、その検索結果をクライアント装置２００に送信する。そして、クライアント装置２００は、サーバ装置１００，１００ａ，１００ｂから受信した検索結果をマージし、マージした検索結果をアプリケーションソフトウェアに提供する。これにより、クライアント装置２００から見たレスポンス時間を短縮できる。すなわち、データの検索を高速化できる。

【０２２４】

また、データを検索するときサーバ装置１００，１００ａ，１００ｂが異なるデータベースにアクセスするため、共通のデータベースにアクセスする場合と比べて、アクセス競合を抑制できスループットが向上しやすくなる。また、データベース１２０，１２０ａ，１２０ｂのデータが同期されているため、あるサーバが故障しても、残りのサーバを用いてデータ全体へのアクセスを確保でき、耐故障性が向上する。

【０２２５】

また、各サーバ装置が担当する部分検索範囲は、重複しないように算出される。これにより、各サーバ装置による無駄な検索処理を削減できる。

また、各サーバ装置は、部分検索範囲を算出する際、Ｂ－Ｔｒｅｅインデックス等の木構造のインデックスを用いる。これにより、部分検索範囲の間で検索の負荷（例えば、部分検索範囲に属するデータの量）がほぼ均等になることが期待できるため、各サーバ装置による検索処理の時間が均等化される。よって、検索処理の時間の差により生じる遅延が抑制され、本システムによるデータの検索時間がより短縮される。

【０２２６】

［第３の実施の形態］

次に、第３の実施の形態を説明する。前述の第２の実施の形態との違いを中心に説明し、第２の実施の形態と同様の事項については説明を省略する。第３の実施の形態のシステムでは、各サーバ装置の限定範囲（検索処理を担当する範囲）を算出するときに用いるインデックスの構造が、第２の実施の形態のものと異なる。

【０２２７】

例えば、第２の実施の形態のシステムでは、インデックスのルートノードのキーの数を“検索に参加するサーバ装置の数－１”としている。そのため、サーバ装置の故障や復旧により検索に参加するサーバ装置の数が変化する度に、インデックスを更新することになる。

【０２２８】

そこで、第３の実施の形態では、検索に参加するサーバ装置の数が変化しても更新が不要になるようなインデックスを生成することにする。ただし、第３の実施の形態では、検索に参加するサーバ装置の数は、“１～システムの運用当初のサーバ装置の数”の範囲内

で変化するものとする。すなわち、サーバ装置の数が運用当初よりも増えない限り、サーバ装置の故障や復旧が生じても、インデックスを再構成しなくてよいようにする。

【０２２９】

図３０は、インデックスの変形例を示す図である。インデックス１３２は、図６等で説明した第２の実施の形態のインデックス１３１の変形例である。インデックス１３２は、システムの運用当初の検索に参加するサーバ装置の数を“ｘ”とすると（以下、同様とする）、“ｘ”が２以上である場合、ルートノードのキーの数が「（“１”から“ｘ”までの自然数の最小公倍数）－１」となるように生成される。“ｘ”が１の場合は、上記の方法でインデックスを生成するとルートノードのキーの数が０になるため、ルートノードのキーの数が１以上の任意の数となるようにインデックスを生成する。

【０２３０】

例えば、システムの運用当初の検索に参加するサーバ装置の数が“３”とすると、インデックスのルートノードのキーの数は、「（“１”、“２”、“３”の最小公倍数）－１」により“６－１＝５”となる。

【０２３１】

このように、ルートノードのキーの数を「（“１”から“ｘ”までの自然数の最小公倍数）－１」とすることで、運用当初のサーバ装置のうちの任意の台数が故障しても、インデックスを生成し直さなくてよくなる。

【０２３２】

例えば、インデックス１３２は、ルートノードに“４”、“１１”、“１８”、“２５”および“３２”の５つのキーが設定されている。この状態で、３台のサーバ装置で検索処理を並列して実行する場合、図３０の上側のように、“ｐ＜１１”、“１１≦ｐ＜２５”および“２５≦ｐ”の限定範囲が算出される。また、この状態で、３台のサーバ装置の何れか１台が故障した場合、検索に参加するサーバ装置の数は２台となり、図３０の下側のように、“ｐ＜１８”および“１８≦ｐ”限定範囲が算出される。

【０２３３】

このように、“ルートノードのキーの数＋１”は、“１”からシステムの運用当初のサーバ装置の数までの全ての自然数で割りきれる。このため、検索に参加するサーバ装置が任意の数だけ故障しても、ルートノードのキーによって分割される区間を、稼働中のサーバ装置に均等に割り振ることができる。よって、インデックス１３２を再構成しなくても、稼働中の複数のサーバ装置の間で、検索の負荷をほぼ均等にすることが可能となる。

【０２３４】

図３１は、限定範囲の算出処理の変形例を示すフローチャートである。図３１の処理は、第２の実施の形態の限定範囲算出（図２２の処理）に代えて、前述のステップＳ４９において実行される。第２の実施の形態との違いは、ステップＳ８１とステップＳ８３の間にステップＳ８１ａが追加される点と、ステップＳ８４，Ｓ８６，Ｓ８７に代えてステップＳ８４ａ，Ｓ８６ａ，Ｓ８７ａが実行される点である。以下、ステップＳ８１ａ，Ｓ８４ａ，Ｓ８６ａ，Ｓ８７ａについて説明する。

【０２３５】

（ステップＳ８１ａ）処理内容判定部１１１は、“（ルートノードのキーの数＋１）／検索サーバ装置数ａ”により変数“ｎ”の値を算出する。

（ステップＳ８４ａ）処理内容判定部１１１は、限定範囲を“ｐ＜Ｒ（ｎ）"と算出する。そして、限定範囲算出を終了する。

【０２３６】

（ステップＳ８６ａ）処理内容判定部１１１は、限定範囲を“Ｒ（（稼働優先度ｂ－１）＊ｎ）≦ｐ"と算出する。そして、限定範囲算出を終了する。

（ステップＳ８７ａ）処理内容判定部１１１は、限定範囲を“Ｒ（（稼働優先度ｂ－１）＊ｎ）≦ｐ＜Ｒ（稼働優先度ｂ＊ｎ）"と算出する。

【０２３７】

図３０～図３１で説明したように、各サーバ装置は、ルートノードのキーの数が「（“

１”から“ｘ”までの自然数の最小公倍数）－１」となるように生成されたインデックス用いて限定範囲を算出し、この限定範囲から部分検索範囲を算出する。これにより、検索に参加するサーバ装置の数が変化したとき、インデックスを更新する処理が不要となるため、各サーバ装置の処理の負荷を軽減できる。

【０２３８】

なお、第３の実施の形態のシステムでは、ルートノードのキーの数を「（“１”から“ｘ”までの自然数の最小公倍数）－１」となるように生成するが、「（“１”から“ｘ”までの自然数の公倍数）－１」としても同様の効果を得ることが可能である。例えば、「“１”から“ｘ”までの自然数の公倍数」として、最小公倍数をＮ倍（Ｎは２以上の整数）したものを用いることができる。また、例えば、「“１”から“ｘ”までの自然数の公倍数」として、“１”から“ｘ”までの自然数の積を用いることができる。

【０２３９】

［第４の実施の形態］

次に、第４の実施の形態を説明する。前述の第２の実施の形態との違いを中心に説明し、第２の実施の形態と同様の事項については説明を省略する。第４の実施の形態のシステムは、複数のサーバ装置の間の更新処理と検索処理の分担が第２の実施の形態と異なる。

【０２４０】

図３２は、サーバ装置の機能構成の変形例を示す図である。第４の実施の形態のシステムは、サーバ装置１００－１，１００ａ－１，１００ｂ－１を有する。サーバ装置１００－１は、現用系のサーバ装置である。第４の実施の形態では、現用系のサーバ装置は、更新処理のみ実行し、検索処理を実行しない。そのため、サーバ装置１００－１は、インデックス（例えば、インデックス１３１ａ，１３１ｂに相当するもの）を有さなくてもよい。

【０２４１】

サーバ装置１００ａ－１，１００ｂ－１は、待機系のサーバ装置である。第４の実施の形態では、待機系のサーバ装置は、検索処理のみ実行し、更新処理を実行しない。サーバ装置１００ａ－１，１００ｂ－１は、並列に検索処理を実行することができる。サーバ装置１００ａ－１はインデックス１３１ａ－１を有し、サーバ装置１００ｂ－１はインデックス１３１ｂ－１を有する。インデックス１３１ａ－１，１３１ｂ－１それぞれは、“検索に参加するサーバ装置の数－１”の数のキーをルートノードに有する。例えば、図３２では、検索に参加するサーバ装置の数は、稼働中の待機系のサーバ装置の数である“２”であるため、ルートノードのキーの数は、“１”となる。

【０２４２】

図３３は、システムによる検索の変形例を示す図である。サーバ装置１００－１はサーバ制御部１１０－１を有し、サーバ装置１００ａ－１はサーバ制御部１１０ａ－１を有し、サーバ装置１００ｂ－１はサーバ制御部１１０ｂ－１を有する。サーバ制御部１１０－１，１１０ａ－１，１１０ｂ－１は、第２の実施の形態のサーバ制御部１１０，１１０ａ，１１０ｂに対応する。

【０２４３】

クライアント装置２００は、サーバ装置１００－１，１００ａ－１，１００ｂ－１に、同じ検索条件を指定した検索要求を送信する（Ｓ２，Ｓ２ａ，Ｓ２ｂ）。サーバ制御部１１０－１は、クライアント装置２００からデータの検索を要求されたとき、検索処理を実行せず、クライアント装置２００へ応答しない。

【０２４４】

サーバ制御部１１０ａ－１，１１０ｂ－１それぞれは、クライアント装置２００からデータの検索を要求されたとき、自己のサーバ装置の種別と検索に参加するサーバ装置の数とに基づき部分検索範囲を算出する。例えば、サーバ制御部１１０ａ－１は部分検索範囲＃１を算出し、サーバ制御部１１０ｂ－１は部分検索範囲＃２を算出する（Ｓ３ａ－１，Ｓ３ｂ－１）。そして、サーバ制御部１１０ａ－１は、部分検索範囲＃１に限定してデータベース１２０ａからデータを検索し、クライアント装置２００に検索結果を送信する（

Ｓ４ａ）。サーバ制御部１１０ｂ－１は、部分検索範囲＃２に限定してデータベース１２０ｂからデータを検索し、クライアント装置２００に検索結果を送信する（Ｓ４ｂ）。

【０２４５】

第４の実施の形態のシステムでは、現用系のサーバ装置１００－１は更新処理のみを担当し、待機系のサーバ装置１００ａ－１，１００ｂ－１は検索処理のみを担当する。これにより、現用系のサーバ装置１００－１の負荷を軽減し、データの更新および並列検索をサーバ装置１００－１，１００ａ－１，１００ｂ－１の間で適切に分担することができる。

【０２４６】

なお、前述のように、第１の実施の形態の情報処理は、サーバ１０，１０ａや検索要求装置２０にプログラムを実行させることで実現できる。第２および第３の実施の形態の情報処理は、サーバ装置１００，１００ａ，１００ｂやクライアント装置２００にプログラムを実行させることで実現できる。第４の実施の形態の情報処理は、サーバ装置１００－１，１００ａ－１，１００ｂ－１にプログラムを実行させることで実現できる。このようなプログラムは、コンピュータ読み取り可能な記録媒体（例えば、記録媒体４３）に記録しておくことができる。記録媒体としては、例えば、磁気ディスク、光ディスク、光磁気ディスク、半導体メモリ等を使用できる。磁気ディスクには、ＦＤおよびＨＤＤが含まれる。光ディスクには、ＣＤ、ＣＤ－Ｒ（Recordable）／ＲＷ（Rewritable）、ＤＶＤおよびＤＶＤ－Ｒ／ＲＷが含まれる。

【０２４７】

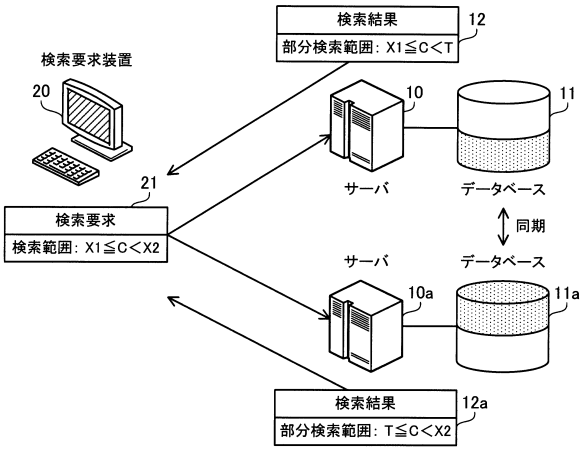
プログラムを流通させる場合、例えば、当該プログラムを記録した可搬記録媒体が提供される。コンピュータは、例えば、可搬記録媒体に記録されたプログラムを、記憶装置（例えば、ＨＤＤ１０３）に格納し、当該記憶装置からプログラムを読み込んで実行する。ただし、可搬記録媒体から読み込んだプログラムを直接実行してもよい。また、上記の情報処理の少なくとも一部を、ＤＳＰ、ＡＳＩＣ、ＰＬＤ（Programmable Logic Device）等の電子回路で実現することも可能である。

【符号の説明】

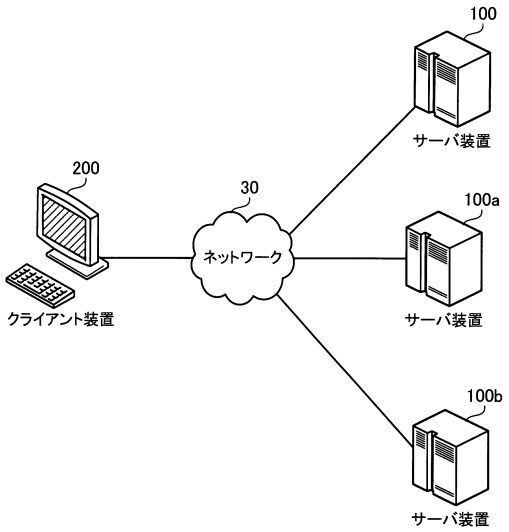
【０２４８】

１０，１０ａ　サーバ
１１，１１ａ　データベース
１２，１２ａ　検索結果
２０　検索要求装置
２１　検索要求

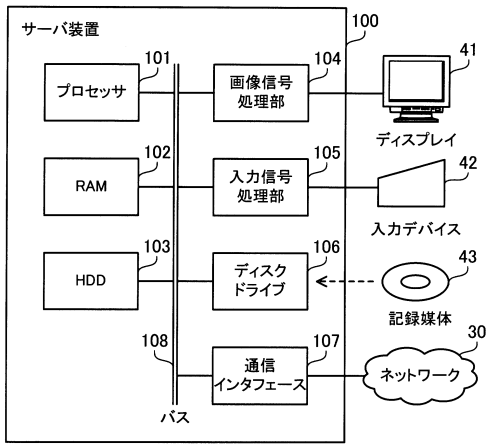
【図１】


検索結果
部分検索範囲：X1≦C＜T
12
検索要求装置
20
10
11
サーバ
データベース
21
検索要求
検索範囲：X1≦C＜X2
同期
サーバ
データベース
10a
11a
検索結果
部分検索範囲：T≦C＜X2
12a

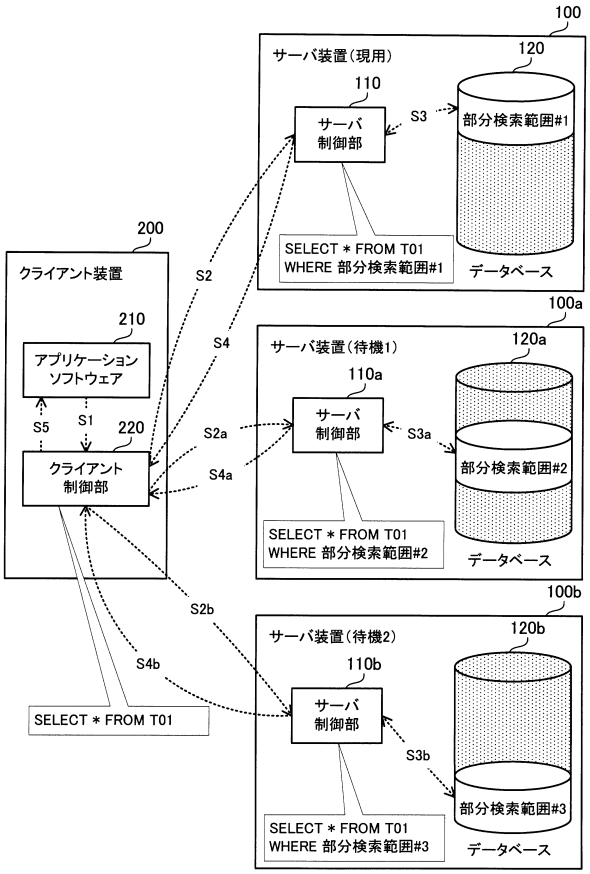
【図２】


100
サーバ装置
200
30
100a
ネットワーク
サーバ装置
クライアント装置
100b
サーバ装置

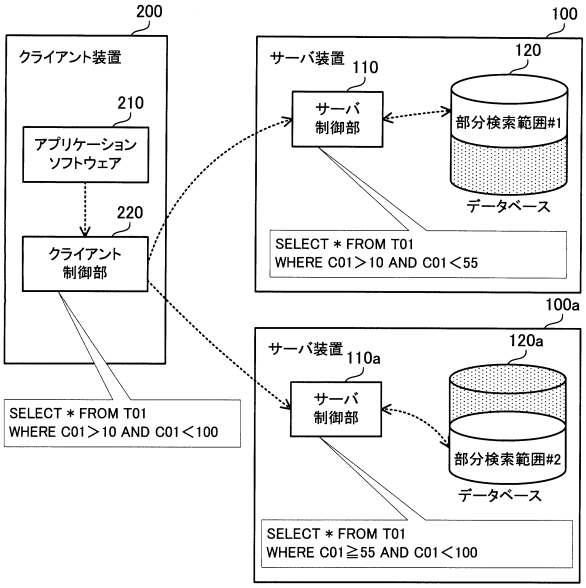
【図３】


サーバ装置
100
プロセッサ
101
画像信号処理部
104
41
ディスプレイ
RAM
102
入力信号処理部
105
42
入力デバイス
HDD
103
ディスクドライブ
106
43
記録媒体
108
通信インタフェース
107
30
ネットワーク
バス

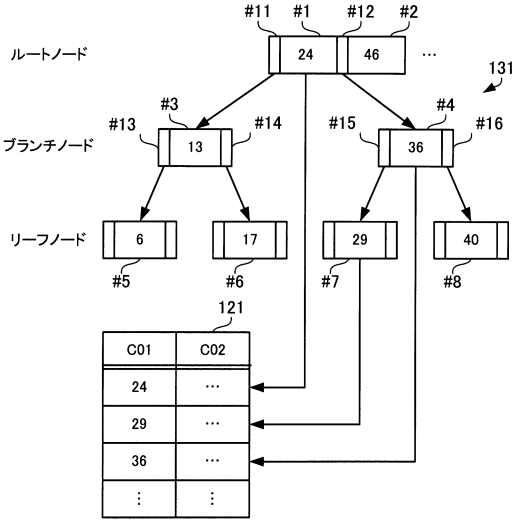
【図４】


100
サーバ装置（現用）
120
110
サーバ制御部
S3
部分検索範囲#1
SELECT＊FROM T01
WHERE 部分検索範囲#1
データベース
200
クライアント装置
S2
100a
サーバ装置（待機1）
120a
S4
210
アプリケーションソフトウェア
110a
S5
S1
220
S2a
S3a
部分検索範囲#2
クライアント制御部
S4a
SELECT＊FROM T01
WHERE 部分検索範囲#2
データベース
100b
S2b
サーバ装置（待機2）
120b
S4b
110b
SELECT＊FROM T01
S3b
部分検索範囲#3
SELECT＊FROM T01
WHERE 部分検索範囲#3
データベース

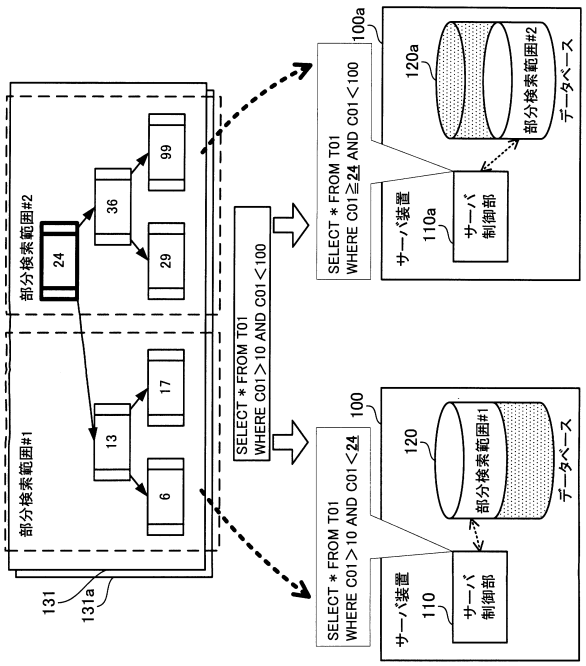
【図５】


200
クライアント装置
210
アプリケーション
ソフトウェア
220
クライアント
制御部
SELECT * FROM T01
WHERE C01＞10 AND C01＜100
100
サーバ装置
110
サーバ
制御部
120
部分検索範囲#1
データベース
SELECT * FROM T01
WHERE C01＞10 AND C01＜55
100a
サーバ装置
110a
サーバ
制御部
120a
部分検索範囲#2
データベース
SELECT * FROM T01
WHERE C01≧55 AND C01＜100

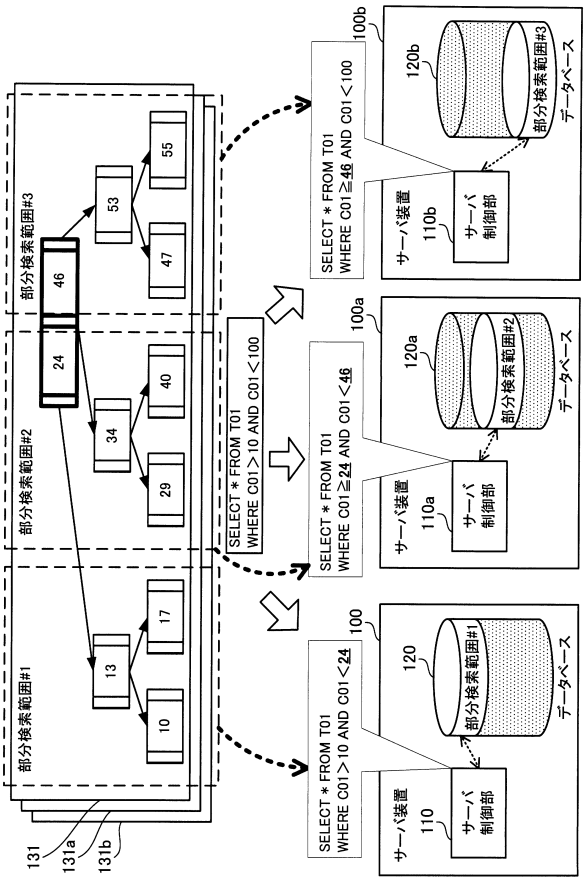
【図６】


ルートノード
ブランチノード
リーフノード
#11
#1
#12
#2
24
46
131
#3
#4
#13
#14
#15
#16
13
36
6
17
29
40
#5
#6
#7
#8
121
C01
C02
24
29
36

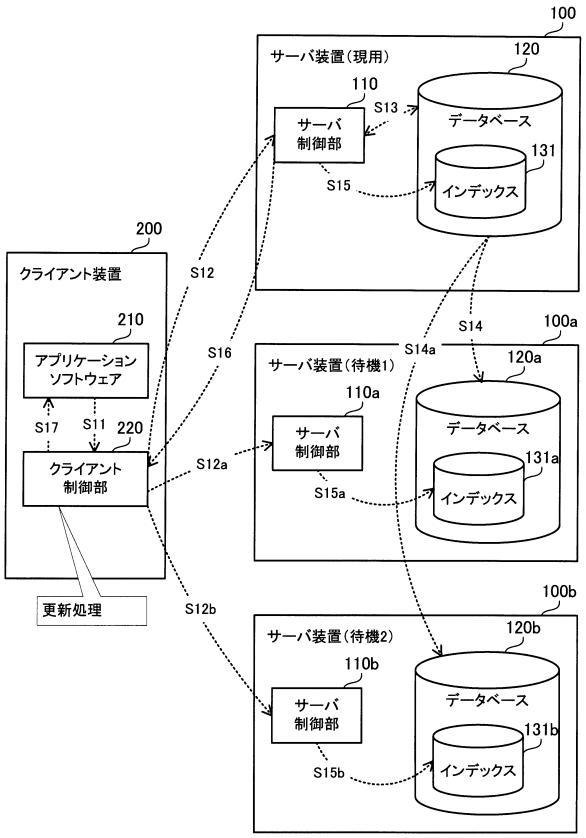
【図７】


部分検索範囲#2
部分検索範囲#1
24
36
99
29
13
17
6
131
131a
SELECT * FROM T01
WHERE C01＞10 AND C01＜100
100a
SELECT * FROM T01
WHERE C01≧24 AND C01＜100
120a
部分検索範囲#2
データベース
サーバ装置
110a
サーバ
制御部
100
SELECT * FROM T01
WHERE C01＞10 AND C01＜24
120
部分検索範囲#1
データベース
サーバ装置
110
サーバ
制御部

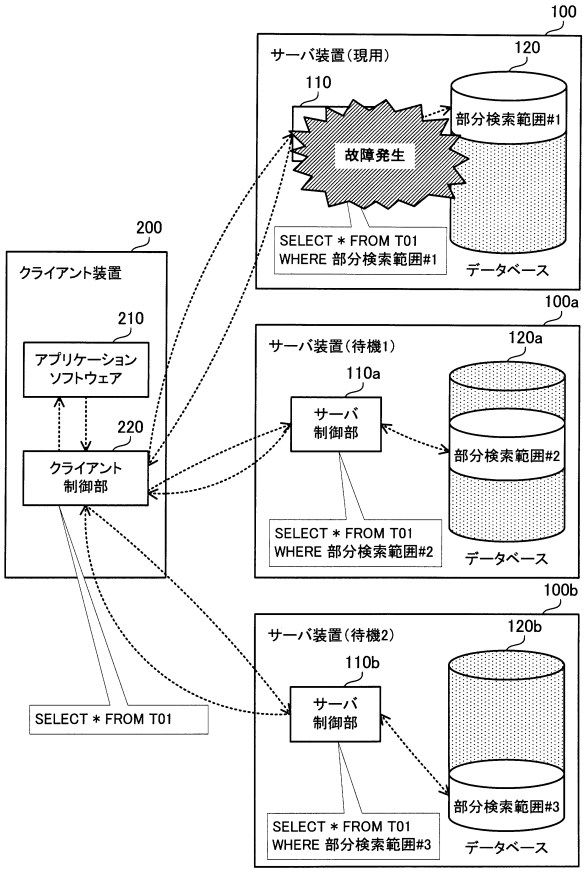
【図８】


部分検索範囲#3
部分検索範囲#2
部分検索範囲#1
46
24
53
55
47
34
40
29
13
17
10
131
131a
131b
SELECT * FROM T01
WHERE C01＞10 AND C01＜100
100b
SELECT * FROM T01
WHERE C01≧46 AND C01＜100
120b
部分検索範囲#3
データベース
サーバ装置
110b
サーバ
制御部
100a
SELECT * FROM T01
WHERE C01≧24 AND C01＜46
120a
部分検索範囲#2
データベース
サーバ装置
110a
サーバ
制御部
100
SELECT * FROM T01
WHERE C01＞10 AND C01＜24
120
部分検索範囲#1
データベース
サーバ装置
110
サーバ
制御部

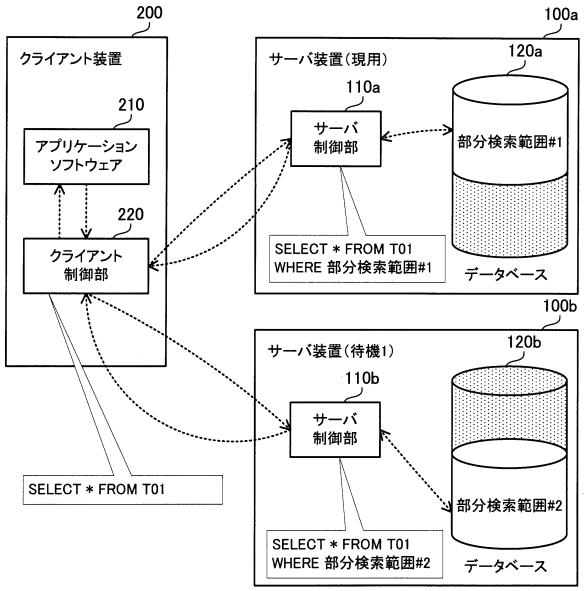
【図９】



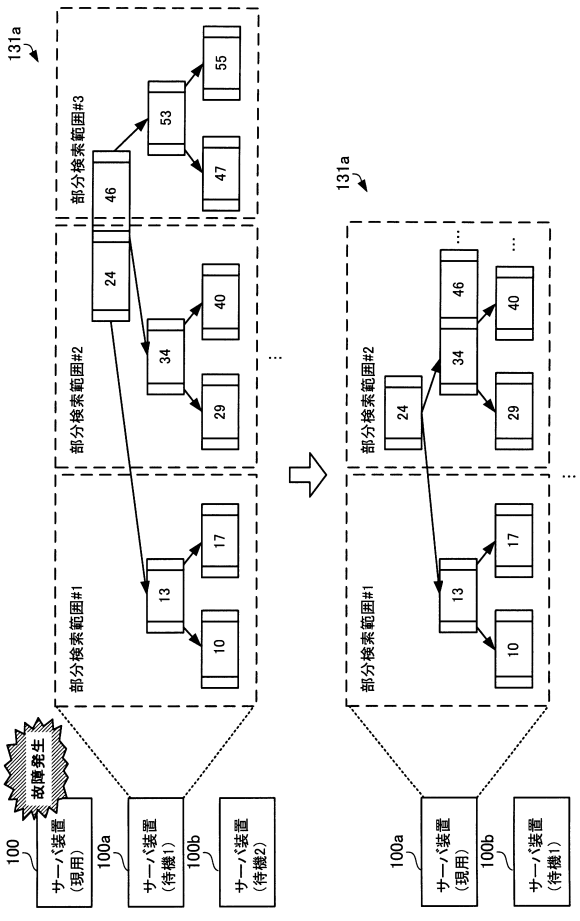
【図１０】



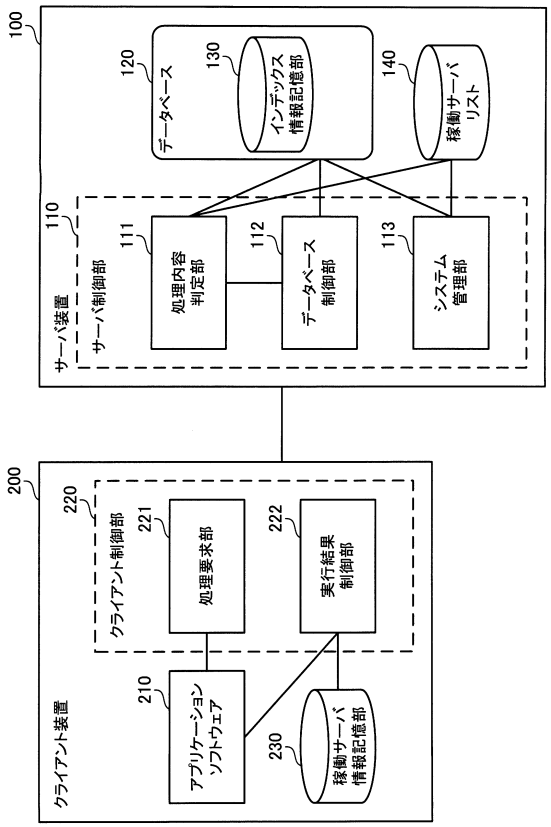
【図１１】



【図１２】



【図１３】

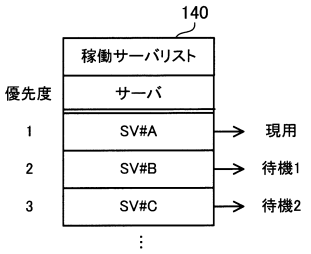


【図１４】

51

処理要求通知					
制御情報	並列フラグ	種別	テーブル	列	条件
情報A	TRUE	検索	T01	C01,C02	C01＞10 AND C01＜100

【図１５】

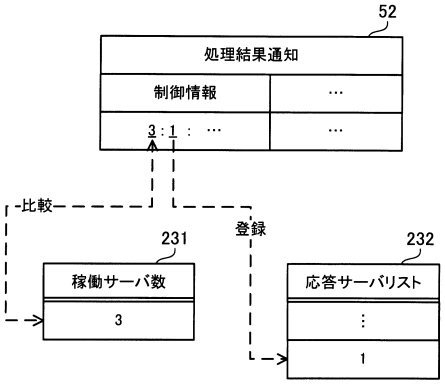


【図１６】

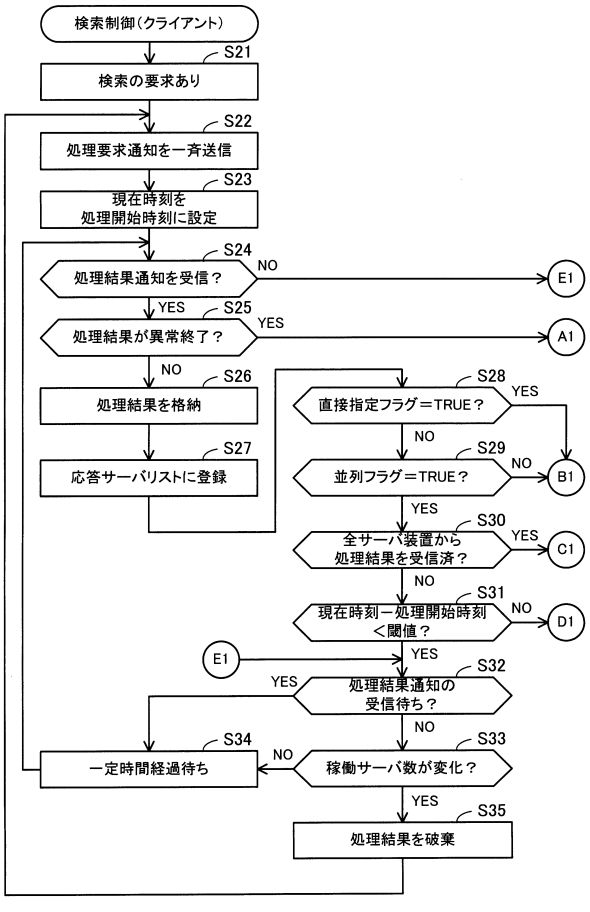
52

処理結果通知		
制御情報	直接指定フラグ	処理結果
3 : 1 :情報A	TRUE	(20, aa), (25, bb)

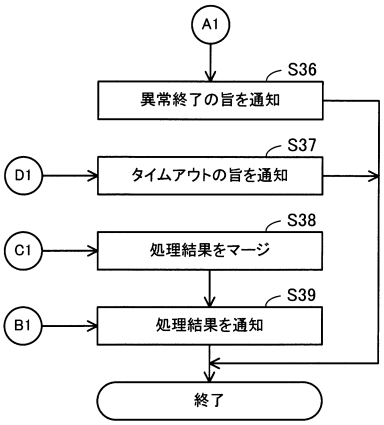
【図１７】



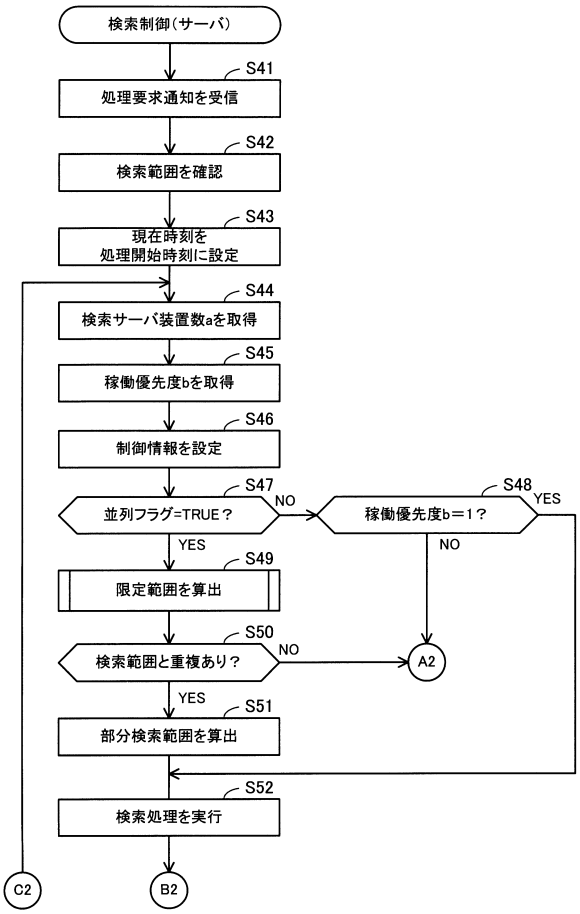
【図１８】



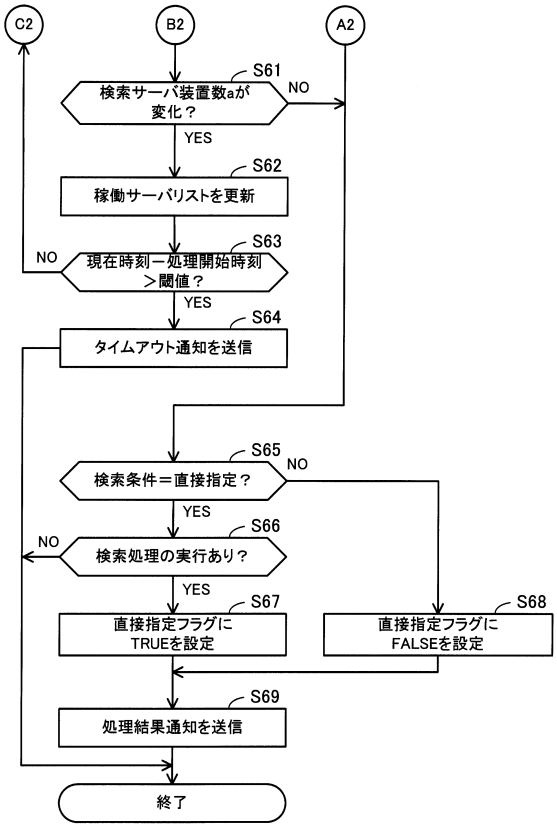
【図１９】



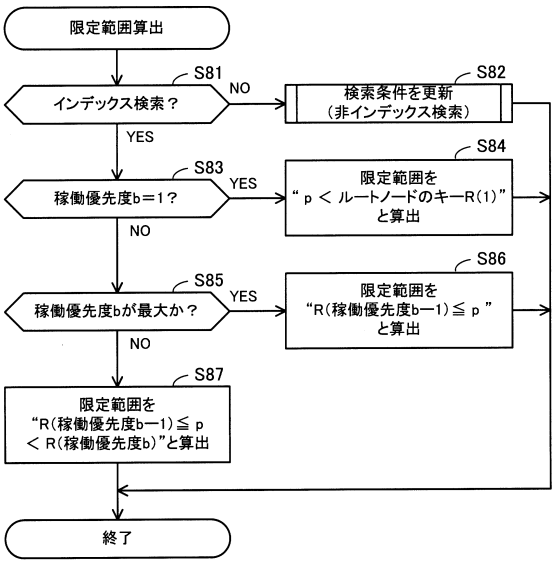
【図２０】



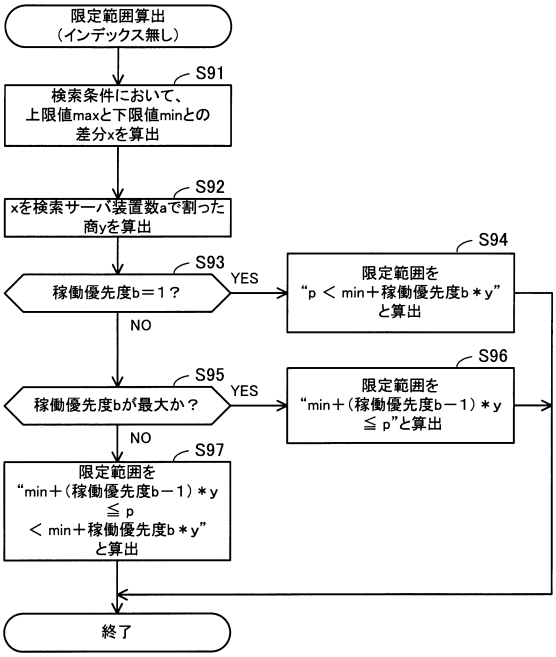
【図２１】



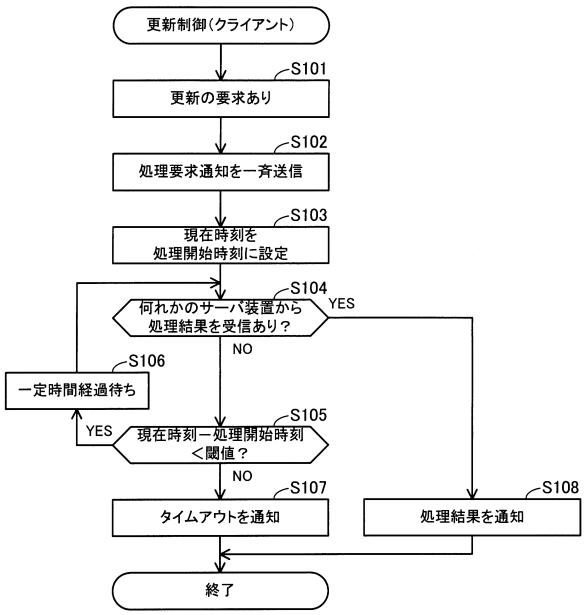
【図２２】



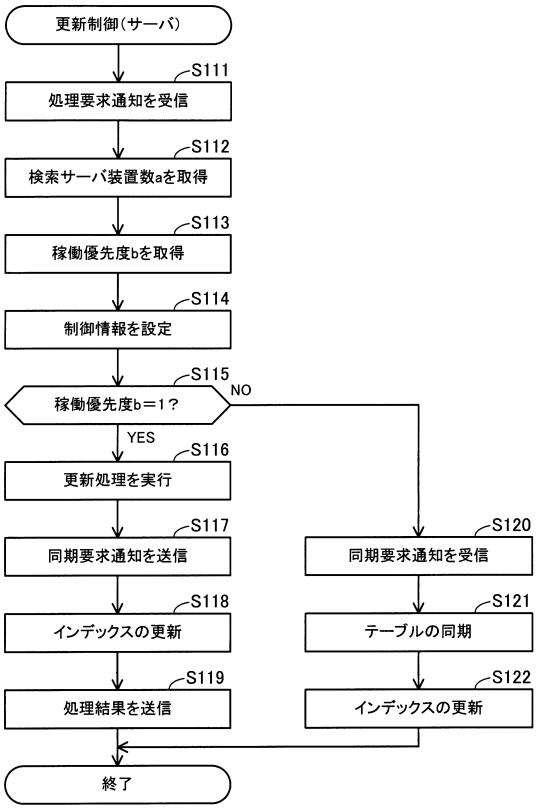
【図２３】



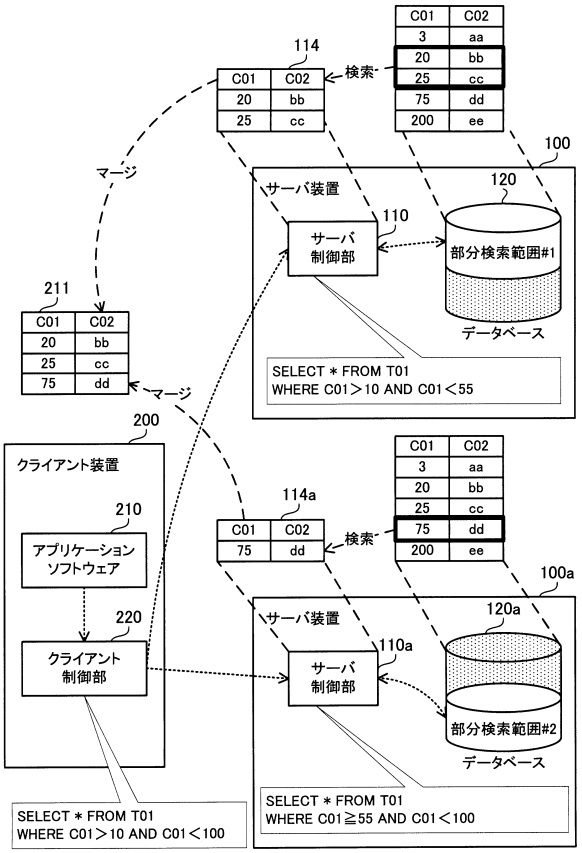
【図２４】



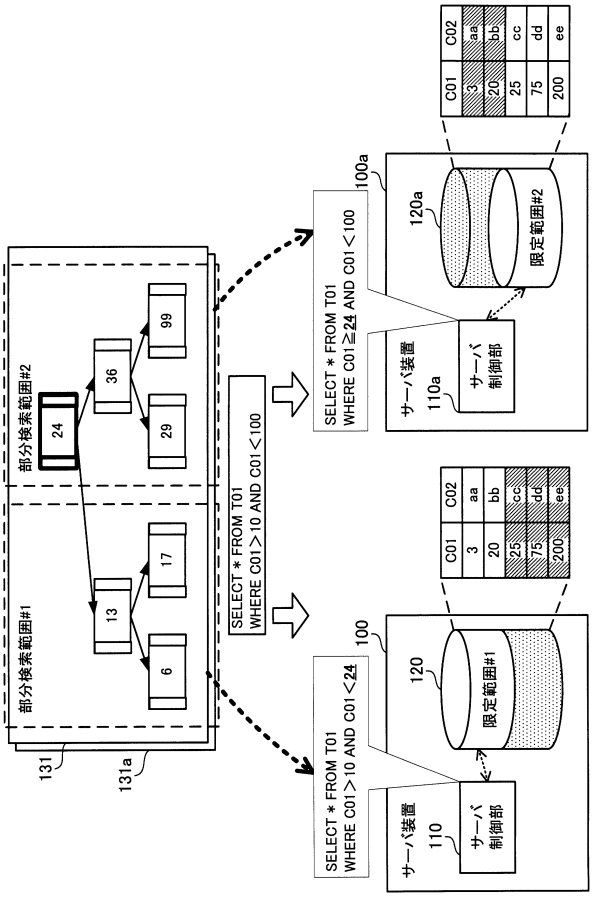
【図２５】



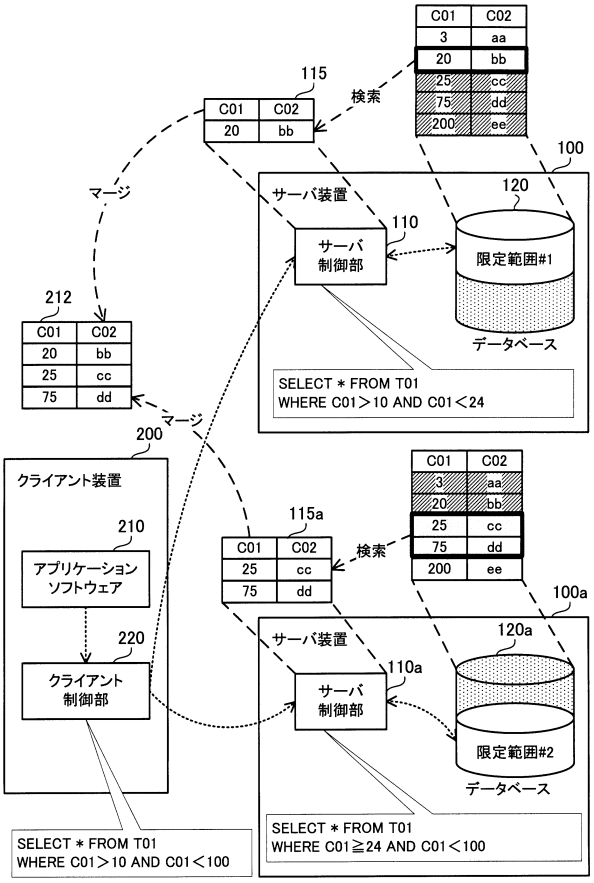
【図２６】



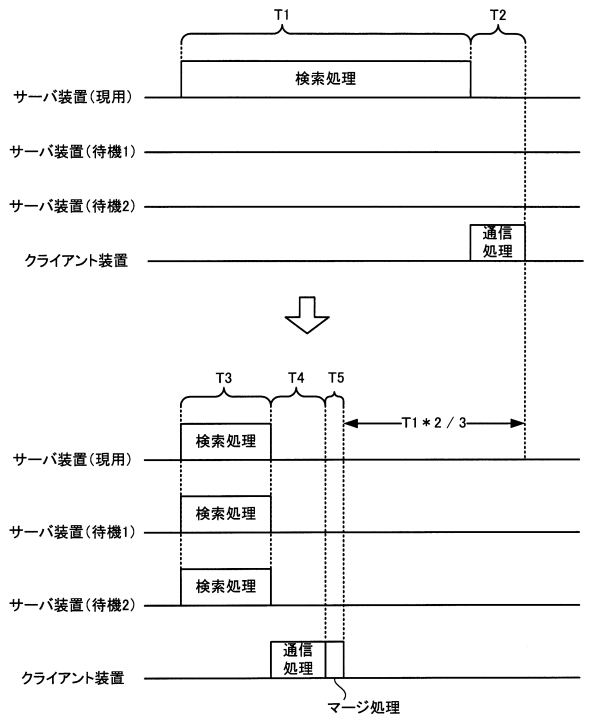
【図２７】



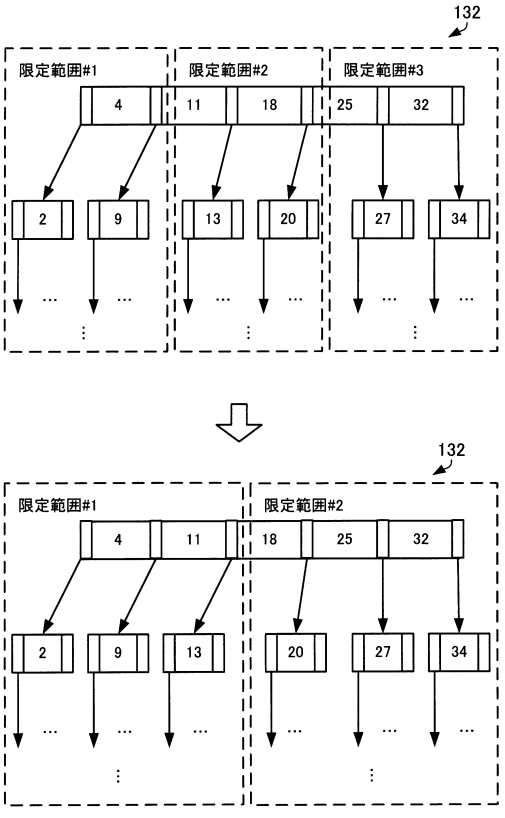
【図２８】



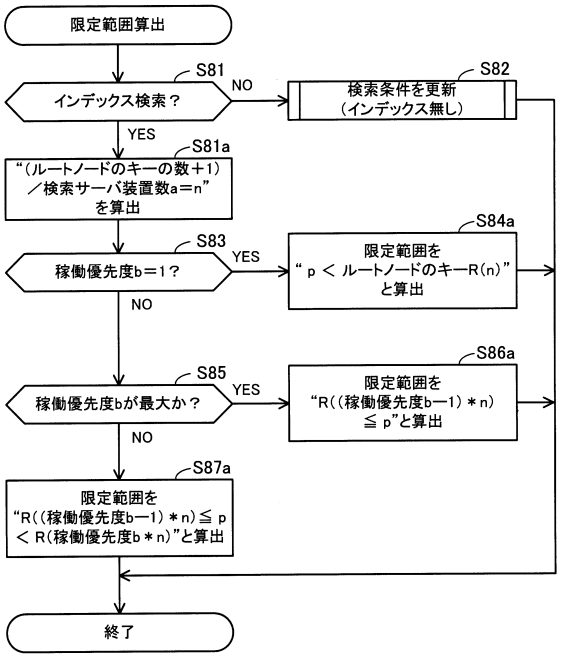
【図２９】



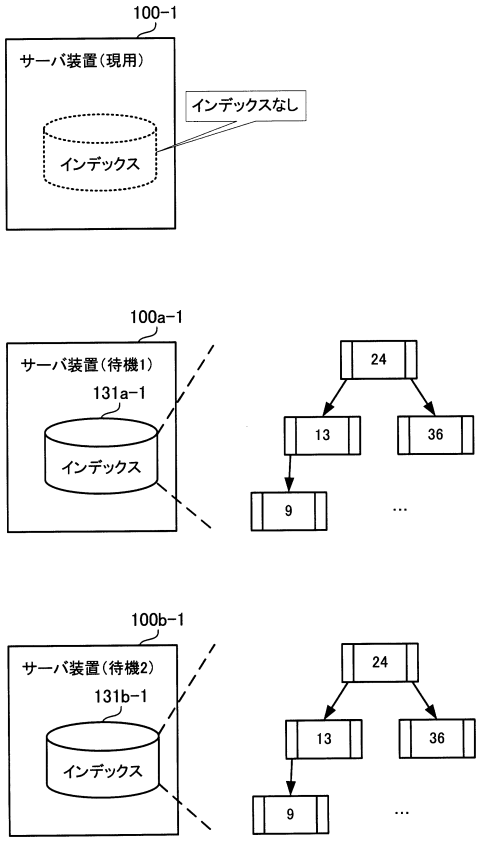
【図３０】



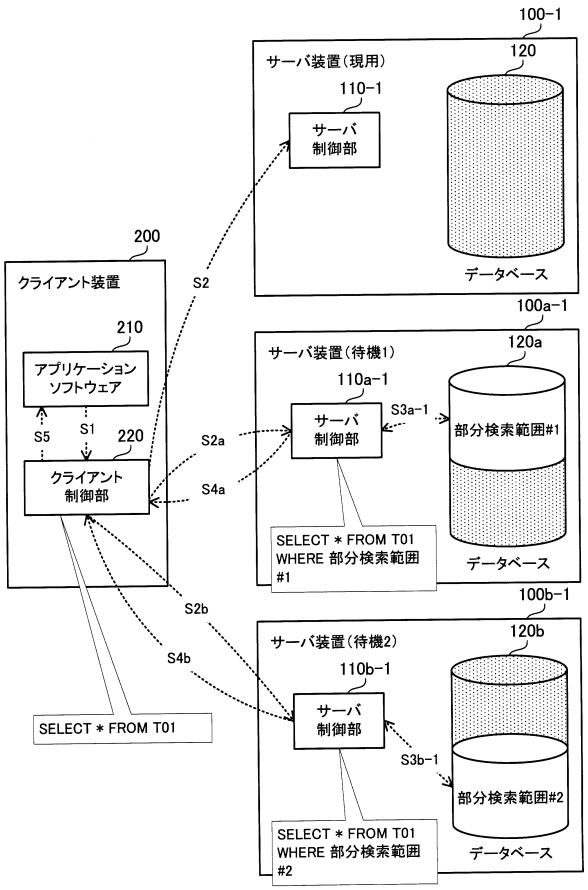
【図３１】


限定範囲算出
S81
インデックス検索？
NO
S82
検索条件を更新
（インデックス無し）
YES
S81a
“（ルートノードのキーの数＋1）
／検索サーバ装置数a＝n”
を算出
S83
稼働優先度b＝1？
YES
S84a
限定範囲を
“ｐ＜ルートノードのキーR(n)”
と算出
NO
S85
稼働優先度bが最大か？
YES
S86a
限定範囲を
“R((稼働優先度b−1)＊n)
≦ｐ”と算出
NO
S87a
限定範囲を
“R((稼働優先度b−1)＊n)≦ｐ
＜R(稼働優先度b＊n)”と算出
終了

【図３２】


100−1
サーバ装置（現用）
インデックスなし
インデックス
100a−1
サーバ装置（待機1）
131a−1
インデックス
24
13
36
9
…
100b−1
サーバ装置（待機2）
131b−1
インデックス
24
13
36
9
…

【図３３】


100−1
サーバ装置（現用）
120
110−1
サーバ
制御部
データベース
200
クライアント装置
S2
210
アプリケーション
ソフトウェア
S5
S1
220
クライアント
制御部
100a−1
サーバ装置（待機1）
120a
110a−1
S3a−1
サーバ
制御部
部分検索範囲#1
S2a
S4a
SELECT＊FROM T01
WHERE 部分検索範囲
#1
データベース
S2b
S4b
100b−1
サーバ装置（待機2）
120b
110b−1
サーバ
制御部
S3b−1
部分検索範囲#2
SELECT＊FROM T01
SELECT＊FROM T01
WHERE 部分検索範囲
#2
データベース

フロントページの続き

審査官　小太刀　慶明

(56)参考文献　特開２００８－２５０７２２（ＪＰ，Ａ）
特開平０７－１６０５５７（ＪＰ，Ａ）
米国特許出願公開第２０１０／００８２６５５（ＵＳ，Ａ１）
米国特許第０６４３９７８３（ＵＳ，Ｂ１）
米国特許出願公開第２００７／０２３９７５９（ＵＳ，Ａ１）

(58)調査した分野(Int.Cl.，ＤＢ名)
Ｇ０６Ｆ　１７／３０
Ｇ０６Ｆ　１２／００
Ｇ０６Ｆ　１３／００