

OLAP サーバにおける並列データロード処理

4 K - 2

松尾直樹* 小野晃範* 津端大輔**

*日立製作所ソフトウェア事業部 **日立システムエンジニアリング

1.はじめに

企業などの分析業務において OLAP (OnLine Analytical Processing)サーバは広く利用されている。多角的な分析を行うためには OLAP サーバは多くの次元や項目数（階層）を管理できなければならない。しかし、次元や項目数（階層）を増やせばその直積で表される多次元空間そのものが巨大化する。このため既存の OLAP サーバでは事前集約処理を含めたデータロードには数十時間、あるいはそれ以上のオーダーの時間が必要なケースがあり、実用上の課題であった。データロードを高速化するための技術としてデータベースを複数に分割し、これを複数のサーバで同時処理する技術、即ち並列化がある。この並列化によりデータロード時間を1桁向上し、データロードを数時間程度に押さえれば今までは不可能であった1日単位のデータ反映も可能となる。しかし、並列 OLAP サーバを実現するにはデータ配置の問題があった。全ての次元軸方向に関してデータが均等に分配されないと、データロード処理において特定のサーバに負荷が集中し、処理時間が長くなる点である。

この問題を解決するためハッシュブロック分割方式を開発した。更に複数ノードを持つクラスタ環境でも、あるいは単一ノードであっても高速なデータロードを実現する内部処理構造を実装した。本稿では並列 OLAP サーバ上における並列データロード処理の技術と今後の課題について述べる。

なお、本稿においては特に断らない限り、MOLAP (Multidimensional OLAP)に関して述べるものとする。

2.並列 OLAP サーバ

2.1.ハッシュブロック分割

図2. 1は地域、年度、商品という3つの次元を持ち、売上高のデータを格納するデータベースを表している。この場合、「東京地区の年ごとの売上高」や「98年の商品毎の売上高」と言った分析業務での代表的な検索処理は次元軸方向に沿ったデータアクセスを伴う。このため、次元軸方向に沿ったデータが特定のサーバ上に揃うなどのデータの配置の偏りがあると、そのサーバに負荷が集中して全体として検索が遅くなる。この問題を解決するために図2. 1に示すハッシュブロック分割を導入した。図2. 1のように多次元空間をいくつかのブロックに分割する。多次元空間上で対角面に沿って並んだブロックを並列OLAPサーバの各格納領域に均等に分割格納す

る。この際、次元の軸方向に隣り合うブロック同士を異なる格納領域に割り当てる。ハッシュブロック分割によって全ての次元軸方向に関して均等なデータの分配を行った。この結果、検索処理の各サーバへの均等な負荷分散を実現した。

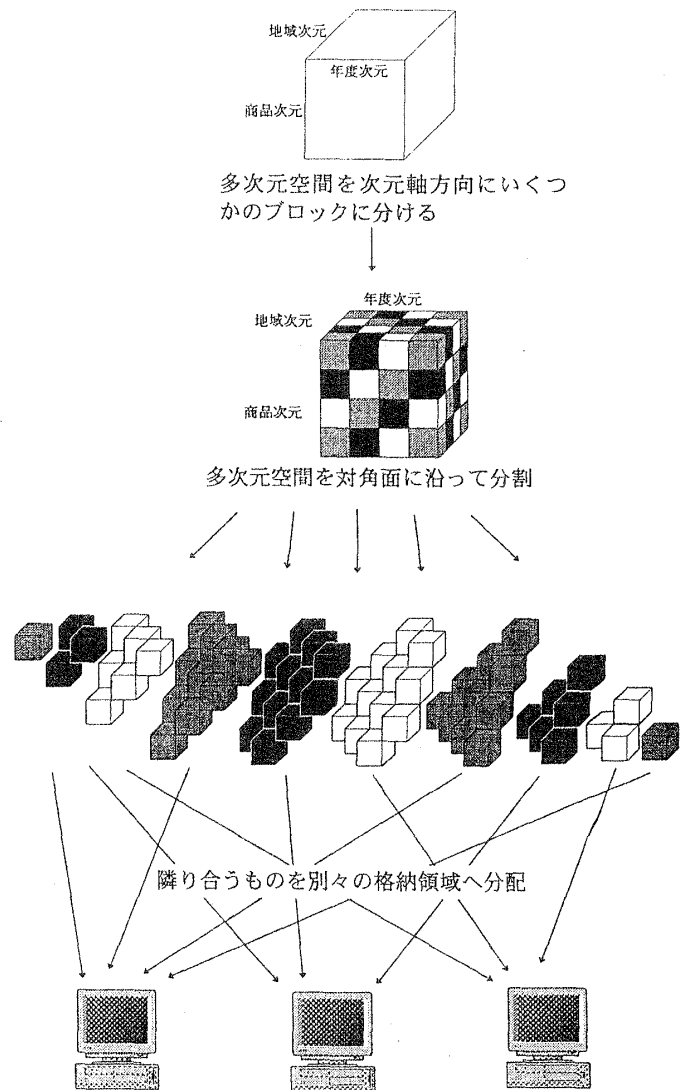


図2. 1 ハッシュブロック分割の概念図

しかし、ハッシュブロック分割は複雑なデータ配置を行うため、データロードには大きな負荷がかかる。データロード処理を高速に実行するために読み込み処理と書き込み処理を多段階に分割し、並列処理する多段階並列処理方式を開発した。

2.2.ノード間並列データロード処理

並列 OLAP サーバではデータロード時に各サーバ上で

複数のスレッドを起動し、処理を実行する。図2. 2にこの概要を示す。

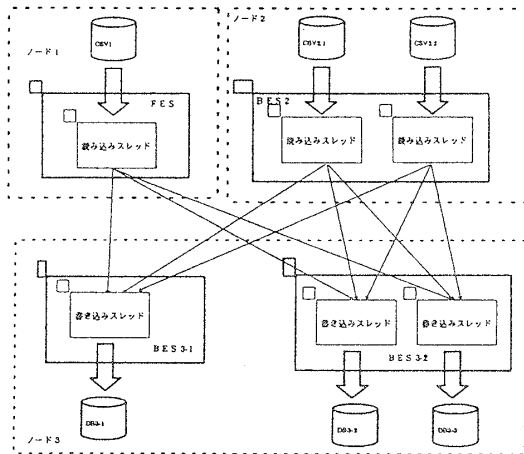


図2. 2 スレッド構成図

図2. 2において、FES、BESとは以下の機能を持つサーバである。

- (a) FES (Front End Server) クライアントからの要求受付、返信。図2. 2のように入力データのファイルからの読み込み処理も実行。
- (b) BES (Back End Server) データベースアクセスおよび入力データのファイル入力処理を実行。

図2. 2に示すようにスレッドを起す単位を入力データの格納されているファイル毎に1個、データの格納領域毎に各1個とした。サーバとスレッドは1:n ($n \geq 0$)の関係にある。これらのスレッドは各々が独立に処理を行う。これにより各読み込みスレッドと書き込みスレッドの処理を並列化した。

また、図2. 2の書き込み側のスレッドでもバッファリングを行い、データベースへの書き込み回数の削減を行うと同時にデータの書き込み処理に起因する通信処理の待ち時間を縮小した。こうした多段階並列処理方式により高速なデータロードが実現できた。

2.3. ノード内並列データロード処理

データベースに格納処理を行うサーバと格納領域の対応を利用者が指定可能とした。即ち、1サーバに複数個の格納領域を割り当てることが出来る。また、サーバプロセスは入力データの読み込み処理も実行する。並列OLAPサーバは読み込み処理と書き込み処理の各々に対応してスレッドを起動する。読み込みスレッドと書き込みスレッドが同じプロセス上で処理を実行するため、スレッド間通信によって高速なデータの受け渡しができた。並列OLAPサーバはこの1サーバ複数BES方式によりプロセッサの処理能力を高いレベルまで引き出すことができた。これにより、複数台サーバのクラスタ構成だ

けでなく、高い処理能力を持つSMPマシン上においても高速なデータロードを実現した。

2.4. スケーラビリティ

並列OLAPサーバによるデータロード処理のスケラビリティを図2. 4に示す。

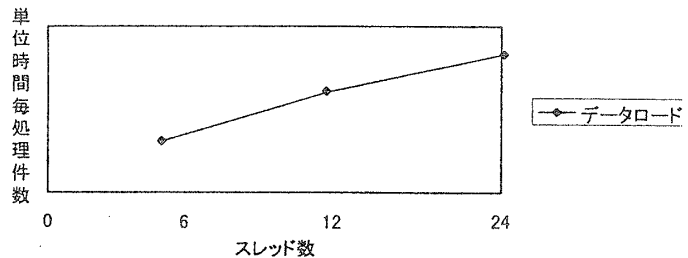


図2. 4 データロード処理性能

図2. 4は単位時間毎の処理件数を表すグラフである。ただし、図2. 4ではスレッド数12のケースは2ノード4CPU構成で、スレッド数24のケースは4ノード8CPU構成にて実測した。このためノードをまたがる通信処理などの影響も若干出ているが、ほぼスレッド数に比例した処理性能となることを出荷製品版にて確認した。

3. まとめと今後の課題

本稿では並列OLAPサーバの概要とデータロードの処理方式を述べた。本稿の提案により並列OLAPサーバ上での高速なデータロードが実現できた。

今後の課題としてはハッシュブロック分割実行時の分割幅の最適化を行い、データベースサイズを縮小することにより、検索とデータロードの一層の高速化を図ることである。

備考

本稿は情報処理振興事業協会(IPA)における「創造的ソフトウェア育成事業」の一環として技術開発された内容を含む。

参考文献

- [1] 大畑他：「創造的ソフトウェア育成事業 大規模データ処理用並列多次元分析サーバソフトの開発 機能仕様書」, IPA(8情高第216号)
- [2] 「Open Middleware Report」(株)日立製作所ソフトウェア事業部(1999 VOL.9)
- [3] 並列OLAPサーバホームページ
<http://www.hitachi.co.jp/Prod/comp/soft1/cosmicube/index.html>