

論文の内容の要旨

論文題目 Online Data Placement Reorganization for Parallel Database Systems
(和訳 並列データベースにおける実行時データ再配置手法に関する研究)

氏名 フィーリフル ヒシャム サイド

近年データベースシステムの高性能化に対する要求は著しく高く、並列データベースシステムの導入が進みつつある。並列化により、多大なる性能向上が期待されるものの、ノード間の負荷の偏りにより必ずしもノード数に比例した性能を得ることは容易ではない。当該問題は、ノード間でデータの再配置を行うことにより改善されるが、全ノードにわたっての大量なデータの授受が発生する為一般にその負荷は小さくなく、通常処理効率を向上させるべくオフライン処理がなされることが想定されている。これに対し、最近の電子商取引サイト等では24時間の運用が通常化しており、システムの無停止性は必須要件となりつつある。このような背景から、本論文では、通常のサービス処理と並行してデータ再配置の為に処理を行うオンラインデータ再配置技術について新しい手法を提案するものである。

並列データベースシステムでは各ノード間におけるデータアクセス負荷が異なると、一部のノードにアクセスが集中するため並列処理効率が低減する。そこで、並列データベースを構築するにあたっては、あらかじめ、各ノード間でのアクセス負荷が均等になるようデータの配置を行う。しかしながら、データアクセス特性はデータベースの利用目的により異なるが、多くは時間とともに変化し、データの再構成を行う必要が生じる。特にオンラインサービス処理では繰り返されるデータの更新により頻繁にデータ再配置を行う必要がある。さらに通常のサービス処理を停止せず並行してデータ再配置を行う必要があるため、データの再配置コストを最小にしスループットを維持しなくてはならない。また、オンライン処理では効率よくアクセスを行うために通常インデクスを構築しているため、データ再配置とともにインデクスの変更コストも考慮する必要がある。以上より、並列データベースシステムにおけるオンラインデータ再配置技法として、1)インデクス再構築技法、

2) データアクセス特性の変化への対応、3) データ再配置処理パラメタ、4) データ再配置における構成技法の4点について論じる。

第一にインデクス再構築について論じる。オンライン処理におけるインデクスの再構築はインデクス木構造の多くの枝にまたがって行くと非常にコストがかかる。そこで、提案するデータ再配置技法では、各ノード間での再配置データの粒度をインデクスの変更コストを最小にするため、対象となるデータを含むインデクス木構造の最小の枝を転送の単位とする。この結果、ノード間データ転送は、元のノードにおける枝刈りおよびその配下のデータのバルク転送、転送先での枝の接合となり、インデクスの変更コストを最小限に押さえられる。しかしながら、データの動的再配置が収斂しない場合にデータ転送単位を比較的大きなサイズで行うと性能が逆に大きく劣化することとなる。すなわち、大きなデータ粒度における収束しないデータ再配置方式ではインデクス再構築コストおよびデータ転送コストが大幅に高くなり、システムの性能向上が得られないばかりか逆に性能低下を招く恐れがあるため、再配置のオーバーヘッドコストを考慮し、システム処理として一定の性能を保証しなくてはならない。

そこで、データ再配置コストとのトレードオフを考慮し、データマイグレーションが収束することを保証できるデータ再配置決定方式を提案する。その結果、マイグレーション粒度として選択されたインデクス枝が決まるため、再配置の収束が保証されることになる。本方式により、再配置要求が生じた場合に、処理エレメント間での負荷分散が可能なかぎり均等かつ高速に行われることとなる。

第三に、データ再配置時のデータ配置のチューニングに関して論じる。データ再配置処理性能をあらゆるパラメタを導入することで、各サービス処理が要求するデータ再配置処理を可能とする。たとえば、データ再配置時の性能低下を調整するために、処理スピードパラメタを導入する。このパラメタを設定することにより、1ステップの再配置あるいは複数ステップの再配置を行うなどの調整が可能となる。このように、パラメタを導入することで、多くのデータ再配置手法を選択することが可能となり、容易にオンラインデータ再配置処理を導入することが可能となる。

第四にデータマイグレーション処理に関する設計方針に関して論じる。ここでは、リニアおよびリング構成の場合のデータマイグレーションについて検討を行う。リニア構成は初期データ配置方式として、また、レンジ分割方式に基づくマイグレーション方式として良く知られた方式であり、すでに多くの処理技法が提案されている。そこで、より多くのデータマイグレーション方針を可能とするリング構成を導入する。その結果、より再配置コストをおさえることが出来る。実験結果から、双方の構成とも、処理負荷を均等にするといく観点からは同等であるが、再配置コストはリング構成の場合がはるかに良いことが確認できた。この結果より、オンラインデータ再配置コストに対して、データ移動時のノード間構成も大きな要素のひとつであることが確認できた。

以上、本論文は並列データベースシステムにおけるオンラインデータ再配置技法に関して検討を行い、高性能な新しい処理技法を提案し、提案された処理技法の有効性について実証した。