

## 【論文審査の要旨】

不揮発半導体メモリの高速度大容量化に伴い、不揮発メモリがディスクだけでなく主記憶にも活用され、主記憶を最終記録媒体とするインメモリデータベースが広く使われるようになる。また、クラウドサービスでの情報サービスの展開が急速に高まっており、近い将来、このような高速なデータベースをクラウドサービスとして提供することが広まると考えられる。

インメモリデータベースにおいては、データベースマネジメントシステム(DBMS)のクエリ最適化で用いられるコスト計算ではIOコストよりも、CPUコストの割合が大きくなる。一般に、コスト計算の精度が低いと誤ったジョイン方式を選択してしまい、クエリの処理時間が増大する問題がある。そこで、適切なジョイン方式が切り替わるクエリのフィルタリング条件(選択率)を高精度で計算することができるCPUコスト計算方式を提案することが研究の目的である。また、クラウドサービスでは様々な世代のCPUが使用されるため、コスト計算式を異なる世代のCPUごとに用意するのは運用作業量の増大につながる。そこで、コスト計算式を異なるCPUに容易に適用可能にすることも本研究の目的とする。

そこで、CPUの処理時間を正確に見積るために、CPUの動作をモデル化する。そのために、CPUアーキテクチャにおける命令パイプラインの動作、およびメモリアクセスに着目したコスト計算式を提案する。ジョイン処理のコストはデータベースの検索処理時間に相当し、データキャッシュを参照する時間、命令キャッシュミスが発生してから命令を取得するまでの時間、分岐予測が失敗してから回復するまでの時間の3種類に分類できる。これらの要素の処理時間を、実際にジョイン処理を実行しているときにCPU搭載のパフォーマンスモニタで計測した統計情報をもとに計算する。統計情報としては命令パイプラインが停止しているサイクル数、実行命令数、キャッシュメモリのヒット回数等があり、これらの統計情報を回帰分析することで計算式を得る。さらに、クエリは多種多様な記述があるため、クエリのアクセスパスにおいて繰り返し出てくる要素で処理を分割し、各要素ごとに部品化したコスト計算式を作成する。これらのコスト計算式を組み合わせることで様々なクエリのコスト計算を可能にする。

一方、本研究で提案したコスト計算式は、各キャッシュメモリや主記憶といったメモリ階層ごとに部品化されているため、異なる世代のCPUに適用するために統計情報の再計測ではなく、統計情報を収集したCPUと世代の異なるCPUで変更した内容(キャッシュメモリ容量や連想度)に対する効果を、変更があった要素の効果(キャッシュミス率低下など)を反映することで対応できる構造になっている。

提案したコスト計算方式の効果を検証するために、実際にコスト計算式を作成し、2 表から 5 表のジョイン処理のコスト計算の精度評価したところ、従来の方法と比較して約 90%の改善、チューニングをした従来方法との比較でも 70%以上の改善が可能になることが分かった。また、異なる世代の CPU で構築したコスト計算式を提案方法で更新した結果、評価したクエリの 66% で精度向上が可能であることが分かった。以上の結果から、本提案コスト計算方式を DBMS に適用することで高精度な CPU コスト計算が可能になり、DBMS の処理性能向上に貢献できると期待されるため、本研究の成果は工学的に重要な意義があると考えられる。よって、本論文は博士（工学）の学位を授与するに十分な価値があるものと認められる。

（最終試験又は試験の結果）

本学の学位規則に従い、最終試験を行った。公開の席上で論文発表を行い、学内外の教員による質疑応答を行った。また、論文審査委員により本論文及び関連分野に関する試問を行った。これらの結果を総合的に審査した結果、専門科目についても十分な学力があるものと認め、合格と判定し