

審査の結果の要旨

氏名 石井 康雄

クロック周波数を増加させることなく性能を向上させるにはパイプラインストール時間の削減による命令レベル並列性の利用が有効である。特に、ハードウェアの予測機構を用いる投機実行はパイプラインストールを削減する有効な手段として知られている。現在では、多くの目的の異なる予測手法がひとつのプロセッサに実装されて性能向上に貢献している。しかし、そうした別々の予測機構は冗長な実行時情報を収集することでコストを増加させ、お互いの予測が干渉することで性能の低下を招いている。

本論文では、複数の目的の異なる予測手法を効果的に統一することで上記問題を解決し、さらなる性能向上を目指すUnified Predictionアーキテクチャを提案している。Unified Prediction アーキテクチャは、複数の異なる機能を持つ予測機構を統一して単一の予測機構を構成する考えに基づいて設計されたアーキテクチャの総称である。この予測機構では統一されたデータ構造に実行時情報を格納し、統一されたそれぞれの予測機能は共通化された実行時情報に基づいて予測を行う。また、それぞれの予測機能はお互いの予測アルゴリズムの妨げとならないように再設計し、複数の予測機構が協調動作できるようにする。統一された予測機構の実現例として、データ依存によるパイプラインストールを削減するUnified Memory Optimizingアーキテクチャと制御依存によるパイプラインストールを削減するUnified Branch Predictionアーキテクチャを提案し、シミュレータにより性能およびエネルギー消費量を評価している。

本論文は6章から構成されている。第1章および第2章は本論文の導入部および背景説明である。ここでは、パイプラインストールを削減する既存の予測機構について概観し、多くの目的の異なる予測手法がひとつのプロセッサに実装されて性能向上に貢献していることを示している。そして、そうした別々の予測機構が似た実行時情報を収集することでハードウェアコストを増加させ、予測アルゴリズムがお互いに干渉することで性能の低下を招くことも指摘している。

第3章において、本論文の一つ目の貢献であるUnified Memory Optimizing (UMO) アーキテクチャについて述べている。最近のDDRメモリデバイスでは、その構造の性質からメモリ領域を連続してアクセスしないと高速データ転送が行えない。複数のコアを有するCPUにおいて、各コア上で動作しているプログラムがメモリ領域からデータを連続してアクセスしていたとしてもメモリシステムには各コアからの要求が集まるため、不連続にメモリ領域をアクセスすることになり高速データ転送ができないことになる。この問題点に対して、メモリアクセススケジューリング、データプリフェッチ、先進的なキャッシュ置き換え方式を統一したUnified Memory Optimizing (UMO) アーキテクチャを提案

している。UMOアーキテクチャでは、複数のメモリアクセス強化のための予測機構を各コア上に実装したUnified Memory Subsystem Controller (UMSC)に統一している。UMSCにはDRAM-Aware PrefetchingとPrefetch-Aware Cache Line Promotionという2つの手法を実現することで、高い性能を実現している。さらに、UMOアーキテクチャの実装として、Map-based Unified Memory Subsystem Controller (MUMSC)を提案している。MUMSCはメモリアクセスマップと呼ぶコスト効率のよいメモリアクセス履歴管理方式の上に、データプリフェッチやキャッシュ置き換えアルゴリズムが実装可能で、このデータ構造を利用することで高いコストパフォーマンスを実現する。SPEC CPU2006ベンチマークプログラムを用いてMUMSCを既存システムと比較した結果、性能が17.8%改善しメモリスシステムの消費エネルギーを6.1%削減した。

第4章は2つ目の貢献であるUnified Branch Prediction (UBP)アーキテクチャを提案している。プログラムにおける分岐にはif文やloop文などで現れる単一のターゲットアドレスをもつ分岐(Monomorphic分岐)と、オブジェクト指向言語などのメソッド呼び出しのためにメソッドテーブルを参照して分岐するような間接分岐(実行時に複数のターゲットアドレスを持つPolymorphic分岐)がある。従来の予測機構ではそれぞれの分岐予測に対して異なる手法が提案されてきた。これら2つの予測に対応するために、アドレス予測機構を統一するUnified Branch Prediction (UBP)アーキテクチャを提案している。UBPには、共通化された実行履歴情報テーブルと専用化された予測論理を持つ。予測論理のコストは実行履歴テーブルと比較して小さいため、高いコスト性能が実現できる。UBPアーキテクチャの実装としてBimode Cascading Branch Target Buffer (BCBTB)を提案している。BCBTBは、広く普及しているMonomorphic分岐向けのBTBとPolymorphic分岐向けのアルゴリズムであるITTAGEを融合させた予測機構で、従来のBTBとほとんど同じコストで高い予測性能を実現する。BCBTBを対象としたシミュレーションの結果、SPEC JVM2008で性能が7.1%改善し、消費エネルギーが4.5%改善することが確認された。第5章で関連研究を述べ本手法の新規性を主張し、第6章において本論文をまとめている。

このように従来個別に設計されていた予測機構を統一することにより、複数の予測機構による競合をなくすだけでなく、予測アルゴリズムの精度も改善し、また、予測機構の実装コストが削減できることを、アーキテクチャを設計し既存予測機構と性能、消費エネルギーの2つの観点で比較し、その有効性を示した。本研究の成果は、計算機アーキテクチャの分野に多大な貢献を行っており、高性能計算機システムの発展に顕著な貢献をしたといえる。

よって本論文は博士(情報理工学)の学位請求論文として合格と認められる。