

博士論文概要

論文題目

組込みシステム設計のための正確かつ
高速なキャッシュ構成シミュレーションに
関する研究

Exact and Fast Cache Configuration
Simulation for Embedded Systems

申請者

多和田	雅師
Masashi	TAWADA

情報理工学専攻 情報システム設計研究

2014年11月

(受理申請する部科主任会開催年月を記入)

近年の LSI 生産, 設計技術の向上に伴いプロセスルールの微細化が進んでいる。微細化による小型化, 低電力化の成果で組込みシステムは世界中の至る所で使用され, 必要不可欠な存在となった。微細化が進みスケーリング則によりプロセッサの演算の処理速度が向上している。一方, 記憶装置のアクセス速度も向上しているが演算の処理速度に追いついていないため, 記憶装置のアクセス速度がボトルネックとなり演算全体の速度は向上しづらくなっている。そこで演算の処理速度と記憶装置のアクセス速度の差を埋めるため, 小容量で高速なキャッシュメモリによる記憶装置の階層化が導入されている。しかし, キャッシュメモリのサイズやメモリブロックの管理などの構成方法は応用によって異なり, 応用ごとに最適な構成方法を選択する必要がある。例えば, キャッシュメモリが十分な容量を持たない構成の場合には記憶装置のアクセス速度が要求性能に達しない可能性があり, またキャッシュメモリが必要以上に容量の大きい構成の場合には性能は満たしても消費電力や面積が余分にかかる可能性がある。要求性能と消費電力, 面積に対して適切なキャッシュ構成を設計する必要がある。組込みシステムのプロセッサアーキテクチャは汎用のプロセッサアーキテクチャと異なり動作するアプリケーションの種類が限定されているのでキャッシュ構成を特化することができる。そこで, 組込みプロセッサ向けに最適なキャッシュ構成を求める問題を考え, アクセス性能を満たしながら容量の小さいキャッシュの構成法を求める。このためには, 種々の容量のキャッシュ構成に対してそのアクセス速度を正確に評価することが重要である。キャッシュ構成のアクセス性能は, 通常キャッシュヒット率を指標として評価されるので, ここでもそれを用いる。キャッシュヒット率は, アクセスされたデータがキャッシュ上に存在する確率であり, アクセス数とキャッシュヒット/ミス回数から計算できる。あるキャッシュ構成でアプリケーションが動作したときのキャッシュヒット/ミス回数が推定できれば性能の良さの指標にすることができる。

キャッシュヒット/ミス回数を推定する手法は大きく分けて 2 つある。1 つはキャッシュの性質からキャッシュヒット/ミス回数を見積もる手法であり, もう 1 つは実際にメモリアクセスをシミュレーションしてキャッシュヒット/ミス回数を数え上げる手法である。前者の手法は見積もりにかかる時間は短いが誤差が大きく, 後者の手法は正確だが時間がかかる。本論文では正確さを重視して後者の手法を対象とし, その高速化の研究を行った。あるアプリケーションが実行される時, プロセッサからメインメモリへのアクセス要求の順序は, 命令の実行順序を変更しないインオーダー実行ならばキャッシュ構成に依存しない。よってプロセッサとアプリケーションが与えられたとき, キャッシュ構成に無関係にプロセッサ上でアプリケーションが動作した際のメモリアクセストレースを得ることができる。このメモリアクセストレースをもとに, キャッシュセット数やブロックサイズ, 連想度などのパラメータを組み替えた複数のキャッシュ構成でシミュレー

シオンし, キャッシュヒット/ミス回数を数え上げることでキャッシュ構成の評価が可能である. これをトレースベースのキャッシュ構成シミュレーションと呼ぶ. キャッシュの構成の場合の数は非常に大きくなるので, 設計時にキャッシュ構成シミュレーションするのは時間がかかる. 設計期間を短縮し, 設計コストを削減するためには, キャッシュ構成シミュレーションを高速化する必要がある.

本論文ではトレースベースのキャッシュ構成シミュレーションの高速化を目的とする. キャッシュメモリには LRU (Least Recently Used) や FIFO (First In First Out), PLRU (Pseudo LRU) ベースキャッシュがある. FIFO/PLRU ベースキャッシュは LRU ベースキャッシュよりもハードウェアコストが低いため組み込みシステムのプロセッサアーキテクチャとの親和性が高い. そこでまずシングルコアプロセッサにおける FIFO/PLRU ベースキャッシュを対象にキャッシュ構成シミュレーションを高速化する. また, 近年組み込みシステムでも用いられるようになりつつあるマルチコアプロセッサアーキテクチャのキャッシュについてトレースベースのキャッシュ構成シミュレーションの高速手法を示す. マルチコアプロセッサのための高速なキャッシュ構成シミュレーションはこれまで例がない.

本論文は 6 章で構成される.

第 1 章では本論文の研究背景および目的, 概要を説明する.

第 2 章では関連研究を紹介する. 最初にキャッシュメモリの説明とキャッシュメモリを構成する際のパラメータ, キャッシュ構成シミュレーションを説明する. 次に関連研究としてキャッシュシミュレーション高速化手法を紹介する. シングルコアプロセッサアーキテクチャのキャッシュを対象とする手法として Janapsatya らの手法や CRCB 手法, CRCB-U 手法, Haque らの手法がある. Janapsatya らの手法や CRCB 手法, CRCB-U 手法は LRU ベースキャッシュの性質 Inclusion Property を利用してキャッシュ構成シミュレーションを高速化する手法である. Haque らの手法は FIFO ベースキャッシュの性質 Intersection Property を利用してキャッシュ構成シミュレーションを高速化する手法である.

第 3 章ではシングルコアプロセッサアーキテクチャにおける FIFO/PLRU ベースキャッシュを対象にキャッシュ構成シミュレーションの高速化手法を提案, 評価する. 最初に第 2 章で紹介した LRU ベースキャッシュを対象とするキャッシュ構成シミュレーション高速化手法 CRCB 手法が FIFO/PLRU ベースキャッシュ構成シミュレーションで動作するか検証し, 動作するために満たすべき条件を考察し, 適用可能であることを証明する. 次に FIFO/PLRU ベースキャッシュの性質を用い, キャッシュ構成シミュレーションの高速化手法を考察する. FIFO/PLRU ベースキャッシュで異なるキャッシュ構成の間に成立する関係性を考える. 小さな連想度のキャッシュセットにアクセスする動作が大きな連想度のキャッシュセットにアクセスする動作と等しいとき, 小さなキャッシュセットは大きなキャッシュセットに「含まれる」関係にあると定義する. この関係にあるとき大きな連

想度のキャッシュセットでシミュレーションすることで小さな連想度のキャッシュセットをシミュレーションしたのと同じ結果となり，シミュレーションを高速化できる．この関係にあるキャッシュセットを統合するデータ構造とこれを扱うアルゴリズムを提案する．最後に提案手法の計算機実験を行う．提案手法を全探索手法と比較した結果，最大 249 分の 1 にシミュレーション時間を削減できた．

第 4 章ではシングルコアプロセッサアーキテクチャにおける FIFO ベースキャッシュを対象に，キャッシュ構成シミュレーションのさらなる高速化手法を提案し評価する．最初に FIFO ベースキャッシュの性質からキャッシュ構成シミュレーション高速化手法を考察する．FIFO ベースキャッシュは LRU ベースキャッシュや PLRU ベースキャッシュと異なりキャッシュヒットが発生してもキャッシュセットの内容に変更がないという性質がある．この性質によりあるキャッシュ構成でキャッシュヒットが起きると判定できるならそのキャッシュ構成での探索処理を省略できる．キャッシュヒットが判定できるキャッシュセットを考察する．小さな連想度のキャッシュセットに存在するデータすべてが大きな連想度のキャッシュセットに存在するとき，小さなキャッシュセットは大きなキャッシュセットに「緩く含まれる」関係にあると定義する．この関係にあるとき小さな連想度のキャッシュセットでキャッシュヒットが発生すれば大きな連想度のキャッシュセットでもキャッシュヒットが発生するので，探索処理を省略できシミュレーションを高速化できる．この関係にあるキャッシュセットを統合するデータ構造とこれを扱うアルゴリズムを提案し，計算機シミュレーションで評価したところ，提案手法を第 3 章の手法と比較した結果，最大 44% シミュレーション時間を削減できた．

第 5 章ではマルチコアプロセッサアーキテクチャのキャッシュを対象にキャッシュ構成シミュレーションの高速化手法を提案，評価する．最初にマルチコアプロセッサキャッシュが持つキャッシュコヒーレンシプロトコルを説明する．キャッシュコヒーレンシプロトコルはマルチコアプロセッサの各プロセッサのキャッシュ間で一貫性を保つために，キャッシュ内のデータに状態を結びつけて矛盾が発生しないように管理するプロトコルである．キャッシュコヒーレンシプロトコルの状態はキャッシュメモリ上のデータにアクセスする手順を定めている．マルチコアプロセッサキャッシュの複数のキャッシュ構成を 1 つのデータ構造で表現するには，キャッシュコヒーレンシプロトコルが障害となる．そこで少ないデータ構造で複数のキャッシュ構成を表現するためにキャッシュコヒーレンシプロトコルの状態に複数の状態を同時に表現するメタ状態を導入する．そしてメタ状態を用いて複数のキャッシュ構成を表現するデータ構造とこれを扱うアルゴリズムを提案する．最後に提案手法の計算機実験を行う．提案手法を全探索手法と比較した結果，最大 80% のシミュレーション時間を削減できた．

第 6 章では本論文のまとめを述べる．

早稲田大学 博士（工学） 学位申請 研究業績書

氏名 多和田 雅師 印

(2015年2月現在)

種 類 別	題名、 発表・発行掲載誌名、 発表・発行年月、 連名者（申請者含む）
a. 論文:学術 誌原著論 文 講演:国際 会議	<p>① ○M. Tawada, M. Yanagisawa, and N. Togawa, "A high-speed trace-driven cache configuration simulator for dual-core processor L1 caches," IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, E96.A, no. 6, pp. 1283-1292, Jun. 2013.</p> <p>② ○M. Tawada, M. Yanagisawa, and N. Togawa, "Speeding-up exact and fast FIFO-based cache configuration simulation," IEICE Electronics Express, vol. 8, no. 14, pp. 1161-1167, Aug. 2011.</p> <p>③ ○M. Tawada, M. Yanagisawa, T. Ohtsuki, and N. Togawa, "Exact, fast and flexible L1 cache configuration simulation for embedded systems," IPSJ Transactions on System LSI Design Methodology, vol. 4, pp. 166-181, Apr. 2011.</p> <p>④ S. Matsuno, M. Tawada, M. Yanagisawa, S. Kimura, T. Sugibayashi, and N. Togawa, "Energy consumption evaluation for two-level cache with non-volatile memory targeting mobile processors," IEIE Transactions on Smart Processing & Computing, vol. 2, no. 4, pp. 226-239, Aug. 2013.</p> <p>⑤ M. Tawada, M. Yanagisawa, and N. Togawa, "A fast trace-driven heterogeneous L1 cache configuration simulator for dual-core processors," in Proc. SASIMI 2013, R4-10s, pp. 259-260, Sapporo, Japan, Oct. 2013.</p> <p>⑥ M. Tawada, M. Yanagisawa, T. Ohtsuki, and N. Togawa, "Exact and fast L1 cache configuration simulation for embedded systems with FIFO/PLRU cache replacement policies," in Proc. VLSI-DAT 2011, pp. 247-250, Hsinchu, Taiwan, Apr. 2011.</p> <p>⑦ (ポスター発表) M. Tawada, M. Yanagisawa, T. Ohtsuki, and N. Togawa, "Exact and fast L1 cache configuration simulation for embedded systems for FIFO/PLRU-based caches," in Proc. ASP-DAC 2011 Student Forum, Makuhari, Japan, Jan. 2011.</p>

早稲田大学 博士（工学） 学位申請 研究業績書

種 類 別	題名、 発表・発行掲載誌名、 発表・発行年月、 連名者（申請者含む）
c. 講演： 国内学会 （査読付）	<p>⑧ 多和田雅師, 柳澤政生, 戸川望, “2 コアプロセッサを対象とする正確で高速なヘテロ L1 キャッシュシミュレーション,” 情報処理学会 DA シンポジウム 2012, vol. 2012, no. 5, pp. 85-90, 2012 年 8 月.</p> <p>⑨ 多和田雅師, 柳澤政生, 戸川望, “複数のキャッシュ構成を同時に表現するデータ構造とこれを用いた高速で正確な 2 コアキャッシュシミュレーション,” 第 25 回 回路とシステムワークショップ, pp. 414-419, 2012 年 7 月.</p> <p>⑩ 多和田雅師, 柳澤政生, 大附辰夫, 戸川望, “FIFO と PLRU をキャッシュ置換ポリシーとする高速なキャッシュ構成シミュレーション手法,” 情報処理学会 DA シンポジウム 2010, vol. 2010, no. 7, pp. 63-68, 2010 年 9 月.</p>
c. 講演： 国内学会 （査読無）	<p>⑪ 多和田雅師, 柳澤政生, 戸川望, “2 コアアーキテクチャを対象とするトレースベース キャッシュシミュレーションの精度評価,” 情報処理学会 ETNET 2013, vol. 2013-SLDM-160, no. 15, 2013 年 3 月.</p> <p>⑫ 多和田雅師, 柳澤政生, 戸川望, “2 コアプロセッサ L1 キャッシュ構成の正確で高速なシミュレーション手法,” 情報処理学会 ETNET 2012, vol. 2012-SLDM-155, no. 3, 2012 年 3 月.</p> <p>⑬ 多和田雅師, 柳澤政生, 戸川望, “2 コアプロセッサアーキテクチャを対象とする正確なキャッシュ構成シミュレーションの高速化に対する一考察,” 2011 年電子情報通信学会 ソサイエティ大会, A-3-11, p. 85, 2011 年 9 月.</p> <p>⑭ 多和田雅師, 柳澤政生, 大附辰夫, 戸川望, “柔軟な置換ポリシーをもつ 2 階層キャッシュの正確で高速なシミュレーション手法,” 電子情報通信学会 VLD 研究会, VLD2010-118, pp. 13-18, 2011 年 3 月.</p> <p>⑮ 多和田雅師, 柳澤政生, 大附辰夫, 戸川望, “FIFO をキャッシュ置換ポリシーとする正確なキャッシュ構成シミュレーションの高速化,” 電子情報通信学会 デザインガイア 2010, VLD-2010-64, pp. 55-60, 2010 年 11 月.</p>

早稲田大学 博士（工学） 学位申請 研究業績書

種 類 別	題名、 発表・発行掲載誌名、 発表・発行年月、 連名者（申請者含む）
e. その他 （賞） （日本学 術振興会 科学研究 費補助金）	2013年8月 情報処理学会 SLDM 研究会 優秀発表学生賞. 2013年11月 情報処理学会, 電子情報通信学会 デザインガイアポスター賞. 日本学術振興会特別研究員奨励費, “世界最速を実現するメニーコアプロセッサの正確なキャッシュ構成シミュレーション技術,” 2012-2014年度, 総額270万円 (2012年度: 90万円, 2013年度: 90万円, 2014年度: 90万円).