

YAWARA: メタレベル計算原理に基づく柔構造計算機システム ソフトウェア構成

大津 金光 斉藤 盛幸 高濱 健吾 横田 隆史 馬場 敬信†
宇都宮大学工学部情報工学科‡

1 はじめに

現実社会のますます高まる高性能計算機システムへの要求に答えるために、我々は、新しい計算原理「メタレベル計算原理」と、それを実現するための新しい計算機システム「柔構造計算機システム」の研究開発を行なっている^[1]。本稿では、本システムにおいて「メタレベル計算原理」の実現上中心的な役割を果たすソフトウェアの構成について述べる。

本ソフトウェアシステムは、柔構造計算機ハードウェア^[2]上でアプリケーションプログラムをソフトウェア(以下、SW)およびハードウェア(以下、HW)の双方で実行するためのコード生成と、実行時プロファイル情報収集コードの生成、収集されたプロファイル情報の解析とそれを基にした最適化処理、アプリケーションコード実行全体の制御を基本的な動作として行なう。その上で、プロファイル戦略や最適化戦略を変化させながら、プログラム実行時の挙動を正確に把握し、その中から最適化に有用な情報を見付け出し、その結果をプログラムコード最適化に反映することで、高い実行性能を目指すものである。

2 基本概念

本システムに関して重要な概念として、(1)SWとHWによる並列投機実行、(2)HWマルチスレッド処理、(3)実行時プロファイル処理、の3つが挙げられる。

2.1 SW/HW 並列投機実行

本システムでは、従来型のプロセッサ上で行なわれていた投機実行を、さらに推し進めて、アプリケーションコードをHW化したレベルで投機実行を行なうことで、従来の投機実行方式では達成困難な更なる高速化を目指す。そのために、我々はアプリケーションコードをSWとHWの双方で実行する構成を採った。アプリケーションコードを適切なサイズに分割したコード片をHWで投機実行することで性能向上を図る一方、従来型のプロセッサ同様のアプリケーションの逐次実行をSWで行なうことで、HWでの投機実行失敗時に正しい実行結果を保証する。

SWおよびHWコードはスレッドを単位としてその実行を管理され、SWで実行されるものをSWスレッド、HWで実行されるものをHWスレッドと呼ぶ。HWスレッドは、アプリケーションコード中で最も実行され

る可能性が大きいと考えられるトレースに沿って投機実行を行なう。HWスレッドが投機実行に成功した場合、そこでアプリケーションの実行が高速化される。その際、正しい実行結果を保証するために実行されるSWスレッドは不要となるため、その実行を中断する。実行を中断したSWスレッドは、投機実行に成功したHWスレッドの終了地点の次位置から実行を再開する。HWスレッドが投機実行に失敗した場合、その実行結果を破棄し、SWスレッドの実行結果をプログラムの実行結果として反映する。

2.2 HWマルチスレッド処理

柔構造計算機ハードウェア上では、1個のSWスレッドと、複数個のHWスレッドが並列実行される。HWスレッドは利用可能なHW資源がある限りいくつでも実行させることが出来、HW化自体による演算レベルの並列性と共に、スレッドレベル並列性を活用することで、従来のマルチスレッド処理方式より一段上の並列度を抽出可能である。

HWスレッドは最も実行される可能性が大きいトレースを投機実行するが、HWスレッドで実行中のコードの次のコード片をさらに別のHWスレッドで投機実行することで、マルチスレッド実行を行なう。これは単一パススレッド実行モデルに相当する。

また、実行される可能性が高いトレースが複数個存在する場合、それらを全てHWで投機実行しておき、後で正しいトレースの結果を実行結果として採るといふ戦略のマルチスレッド実行も可能である。これは複数パススレッド実行モデルに相当する。

以上のような、マルチスレッド実行の際に採りうる戦略を柔軟に変更可能であるため、アプリケーションの特性に応じて最適な並列処理戦略を採ることが可能であり、それによって高速化を図ることが出来る。

2.3 実行時プロファイル処理

アプリケーションコードはHWスレッドによる投機実行により高速化されるが、この投機実行が頻繁に失敗する場合、HWコードが最適な状態でないと考えられる。その場合、HWコードをより最適なものに変更する必要があるが、そのために実行時のプロファイル情報の収集が必要不可欠である。しかし一方で、既に最適な状態である場合にはプロファイル情報の収集は不要な作業であり、そのためのハードウェアは性能上寄与しない。この場合、そのハードウェア資源をアプリケーションの実行に割り当てた方が性能上有利であると考えられる。

そこで、本システムでは、プロファイル情報の取得

YAWARA: A Flexible Computer System based on Meta-Level Computation Model - Software Organization -

† Kanemitsu Ootsu, Moriyuki Saitou, Kengo Takasumi, Takashi Yokota and Takanobu Baba

‡ Department of Information Science, Faculty of Engineering, Utsunomiya University

をHWスレッドとして実行し、アプリケーションを実行するHWスレッドと同列に扱うことで、システム上のHW資源の有効活用を図る。プロファイル処理が不要な場合は、再構成可能ブロックの資源を全てアプリケーション実行のためのHWスレッドに割り当て、プロファイル処理が必要な場合は、再構成可能ブロックの資源の一部をプロファイル情報収集処理を行なうHWスレッドに割り当てることで、HW資源を有効に利用する。

3 システム構成

図1に本ソフトウェアシステムの全体構成を示す。本システムはアプリケーションの実行前に動作する静的最適化システムSOS (Static Optimization System) と、実行中に動作する動的最適化システムDOS (Dynamic Optimization System) の大きく二つのサブシステムから構成される。特に、DOSはアプリケーションの実行と並行して、メタレベルプロセッサ^[2]上で動作するソフトウェアであり、メタレベル計算原理を実現する主体的存在である。

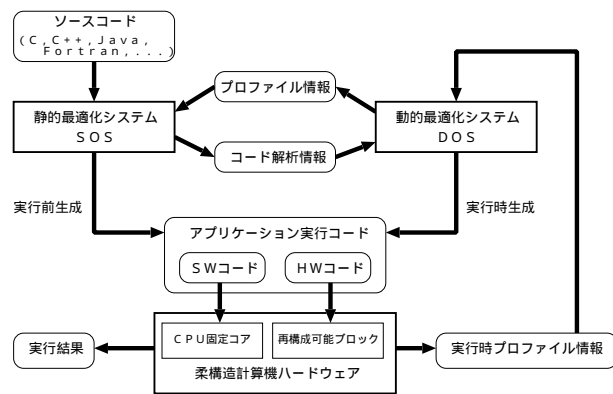


図1. ソフトウェアシステム全体構成

3.1 SOS

SOSはアプリケーションを記述したソースコードを入力とし、柔構造計算機ハードウェア上で実行するアプリケーション実行コードを実行前に生成する。この実行コードは、従来型プロセッサであるCPU固定コア上で実行されるSWコードと、再構成可能ブロックで実行されるHWコードの二つのコードから構成される。

実行コード生成の際、後述のDOSからのプロファイル情報が利用可能であれば、その情報を、最も実行可能性の高いトレースを選択するなどの最適化処理に活用する。DOSからのプロファイル情報が利用できない場合には、何らかのヒューリスティクスやソースコードへのユーザノテーション情報により、トレースを決定して、それに沿った形でHWコードを生成する。この場合、最適なコードになっているとは限らないが、DOSにより実行時に最適化されることで、最適なコードへと変更されていく。

また、SOSはDOSが実行時最適化処理を行なうため

に必要な、制御フロー情報やデータフロー情報などのアプリケーションに関する解析情報を提供する。これにより、DOSの実行時処理の負荷を低減する。

3.2 DOS

DOSはアプリケーション実行中に、プロファイル情報を収集しつつ、その内容を解析し、実行時最適化を行なう。柔構造計算機ハードウェア上では、最初、SOSにより生成された初期のアプリケーション実行コードが実行されるが、HWコードによる投機実行の失敗が頻出した場合にプロファイル処理を開始する。

DOSは、プロファイル情報収集コードをHWコードとして生成し、これをHWスレッドとして実行した結果を取り込むことでプロファイル情報を取得する。再構成可能ブロック上では、アプリケーション実行用のHWスレッドとプロファイル用のHWスレッドが並列に実行されることになるが、それぞれのスレッドへの資源の割り当てもDOSによって行なわれる。プロファイル情報が必要である状況は、アプリケーション実行に割り当てているHW資源が性能に寄与していないことを意味するので、この状況では、プロファイル用のHWスレッドに優先的に資源を割り当てる。

最適化に必要なプロファイル情報が取得でき次第、DOSはプロファイル情報を解析し、最適化戦略を決定、それに基づいたHWコードの生成作業に入る。例えば、プロファイル情報により、単一パスでマルチスレッド実行を行なうか、複数パスで実行を行なうかの戦略の決定などを行なう。新たに生成されたコードを実行することで、HWスレッドの投機実行の成功率を上げる。

コード生成および最適化処理の際、前述のSOSから提供されるアプリケーションコードに関する解析情報を使うことで実行時の負荷を低減し、その分、最適化本来の処理にメタレベルプロセッサの計算資源を回すことが出来るため、より複雑な最適化処理を行なうことが可能となる。

4 おわりに

本稿では、我々が研究開発を行なっている「メタレベル計算原理」を実現する上で、中心的な役割を果たすソフトウェアシステムの構成について述べた。

謝辞 本研究は、一部日本学術振興会科学研究費補助金(基盤研究(B)14380135、同(C)14580362、若手研究14780186)の援助による。

参考文献

- [1] 馬場敬信, 横田隆史, 大津金光 “YAWARA: メタレベル計算原理に基づく柔構造計算機システム —概要—” 情報処理学会第65回全国大会 6G-1, 2003.
- [2] 横田隆史, 大津金光, 馬場敬信 “YAWARA: メタレベル計算原理に基づく柔構造計算機システム —ハードウェア構成—” 情報処理学会第65回全国大会 6G-2, 2003.