

## 細粒度並列計算機お茶の水1号—基本構想—

6G-1

平木 敬 松本 尚 稲垣 達氏 大津 金光 戸塚 米太郎 中里 学

東京大学大学院理学系研究科情報科学専攻

## 1 はじめに

超大型汎用計算機の性能向上が飽和の兆候をみせ始め、それとともにコスト/パフォーマンスの増大が顕著となってきた現在、逐次または並列度の低い密結合汎用計算機を置き換えるものとしての高並列計算機ひいては超並列計算機に対する要求が高まりつつある。しかしながら、領域分割等の自明な方式で容易に高並列性を引き出せる科学技術計算機と異なり、汎用超並列計算機においてはループや関数内部レベル、ブロックレベル、領域分割レベル等問題の持つ多様な並列性を可能な限り秩序立てて引き出す必要がある。

粗粒度で容易に並列化が可能な部分は、並列実行するプロセスの台数と、相互のデータ交換に必要な相互結合網のバンド幅さえあれば、かなりのスピードアップが達成される。その結果、粗粒度による並列化が困難な部分が全体の実行時間のボトルネックとなる。この部分に関しても並列処理を適用して高速化を達成するためには、細粒度の並列性を活用が不可欠である。また、汎用並列システムとして高性能を発揮するためには、粗粒度および細粒度レベルの並列性の抽出による高速化とならび、要素処理装置内部におけるパイプライン並列またはファンクション・ユニットレベルの並列性を可能な限り抽出して高速化を行なうことが基本性能向上のために必要な条件である。

従来、高並列並列処理システムに関しては、粗粒度方向で並列度を引き出すことが主要な研究目標であった。これは、処理装置台数全体を満たす限り粗粒度な並列処理は効率良く並列実行が可能であるという事実に基づく。しかしながら、超並列レベルの処理を行なうためには、自明な並列性の抽出だけでは十分でなく、問題の持つ全てのレベルに渉る並列性の利用が課題である。上記3レベルの並列処理形態は独立に並列性を引き出すものでなく、問題に内在する並列性を分配する。従って、より効率的な並列処理システム構築には低レベルの並列性をいかに効率的に引き出すかが重要な課題である。(図1)

## 2 お茶の水プロジェクトの概要

現在東京大学理学系大学院平木研究室で研究・開発を進めているお茶の水計算機シリーズは、上記前提に則り3レベルの並列性の効率的利用を図るアーキテクチャ研究および最適化コンパイラ研究のためのテストベッドである。お茶の水計算機では、

Basic concept of a general purpose fine-grained parallel processor Ochanomiz-1

Kei Hiraki, Takashi Matsumoto, Tatsushi Inagaki, Yonetaro Totsuka, Gaku Nakazato and Kanemitsu Ootsu

Department of Information Science, Faculty of Science, the University of Tokyo

7-3-1 Hongo, Bunkyo-ku, Tokyo, 113, Japan

E-mail: hiraki@is.s.u.tokto.ac.jp

データ分割等自明な並列化が不可能である細粒度並列処理の効率化を中心課題とし、それを粗粒度並列の単位として並列実行する基本アプローチ、すなわちクラスタ化した基本アーキテクチャを対象とする。

お茶の水計算機シリーズの最初であるお茶の水1号は、超並列計算機の構成要素を目的とした、1クラスタの細粒度並列計算機である。ここでは、粗粒度レベルの並列性に分割した後のプログラムブロックから、細粒度並列処理技術を用いて実行性能を向上させることが課題となるとともに、粗粒度処理の1要素として大域データ通信および大域同期処理を効率的に実現するメカニズムが課題である。

お茶の水1号で対象とする細粒度並列処理に関する問題点は次のものである。

データ依存関係に基づく同期 DOACROSS 構文に代表されるデータ依存関係に基づく横方向の同期をもつプログラムブロックに含まれる局所的並列性の利用は、DOACROSS ループを複数の処理装置で実行することにより実現される。これを効率的に実現するためにはデータ依存関係に基づく同期をループ間依存変数の授受により実現することが必要である。

ループ構造内の細粒度同期処理 DOACROSS 構文においても、ループ間依存変数が多様な場合には個々の変数による同期により実行制御を行なう場合は同期オーバーヘッドが大きく現れる。このような場合には、ループ構造を更にサブブロック化して elastic barrier 等の重複を許す効率的バリア機構を用いることが必要である。

高機能同期構造体の実現 クラスタ化した超並列計算機システムでは、超並列システム特有のクラスタ間通信遅延を隠蔽/短縮するため、多重スレッド環境で実行を行なうことが不可欠である。しかしながら、通信および多重スレッド環境をすべて割り込み等の手段を用いて実現すると、多大なオーバーヘッドが現出する。高機能同期構造体はメモリを構成する全ての語に同期機能を持たせることにより、同期に関するオーバーヘッドを減少させることを目的としている。従来、この種の同期構造体は低コストの多重環境を必要とするため主にデータフロー計算機と組み合わせて用いられてきた。効率的クラスタ処理を行なうためには、スレッド処理を行なう要素処理装置に対してオーバーヘッドを生じることのない高機能同期構造体の実現が必要である。

最適共有メモリプロトコルの実現 多様な並列構造を持つプログラムを階層化したメモリアーキテクチャ上で効率的に実現するためには、アーキテクチャでサポートされた共有メモリ機構が不可欠である。しかしながら、プログラムにおける様々な要請を単一の共有メモリプロトコルで実

現することは非常に困難であり、大きなオーバーヘッドを生む。キャッシュプロトコル切替機構は多様なプロトコルをプログラムから切替えて使用することにより、共有メモリ実現のオーバーヘッドおよびバスまたはネットワークのトラフィックを削減することを目的としたものである [1]。

**大域構造体のプリフェッチ機能の実現** クラスタ内細粒度処理では、キャッシュメモリの持つ局所性を利用した処理が、高効率な実行に不可欠である。使用するデータセットがクラスタ内に閉じていない大域処理では、プリフェッチ操作をできるだけ先行して行なうことにより、クラスタ内に局所性がある状況に変換する必要がある。しかしながら、一元的な大域アドレスを直接使うと、本来存在する処理の局所性がキャッシュメモリのアドレス空間に入らず、効率低下が発生し、また、プリフェッチを行なうプログラム部分と実行部分の同期が必要となる。従って、単純なプリフェッチではなく、大域アドレスを局所性を持ったアドレス空間に写像しつつ、同期を伴った方式でのプリフェッチの実現が必要である。

**性能評価機構の確立** ソフトウェアによる性能評価、例えばプロブイベント計数をプログラムに埋め込むことによるプログラム動作の変化は、細粒度並列処理環境では著しい物がある。また、評価に必要なパラメータを変化させること自身、実機による測定では分離することが困難であった。これらの理由により性能評価をシミュレータを用いて行なわれてきた(多くの場合、実機があるにもかかわらず)。しかしながら、並列性が現出するレベルのプログラム実行は実機で数秒以上の実行時間が必要であり、シミュレーションの限界を越える。従って、シミュレーションと同様の柔軟かつ詳細な測定機能と、遅延等のパラメータを変化可能な能力の実現が必要である。

**細粒度最適化コンパイラ** 粗粒度並列におけるプログラムブロック内部に内在する局所性を有効に引き出すためには、局所に存在する微細な並列性を生かした最適化コンパイル技法が非常に重要な役割をはたす。ここでは、既に自明な並列性が除去されているため、DOACROSS構文、多重の条件分岐構文等低並列度で容易に飽和する並列要素をクラスタ内並列度を満たすレベルで取りだし、かつオーバーヘッドを少なく最適化することが課題である。

お茶の水1号ではこれらの問題点を解決するために、FPGAを用いた柔軟なテストベッドとしての構成を取った。

お茶の水1号は、従来型の高性能マイクロプロセッサ(NEC VR4400)を要素処理装置として使用し、低オーバーヘッドの同期機構や通信機構により効率の良い細粒度並列処理が可能な密結合並列計算機システムである [2]。通信同期の支援機構として、Elastic Barrier、構造体先行フェッチ機構、メモリベースのデータ駆動同期、キャッシュプロトコル切替機構が実装されている。また、お茶の水1号をターゲットとした細粒度最適化コンパイラとして Op. 1 の研究開発を並行して行なっている [3]。

細粒度テストベッドとして、通信同期メモリアクセスに関する機構を導入したに加え、これらの機構および性能測定機構がフィールド・プログラマブル・ゲートアレイ群として実現されていることによる柔軟性が重要な特徴である。

通信同期等基本機構では、処理または通信に要する時間遅延を変化することが可能であると共に、最適ではないと考えられ

る方式との比較を実機を用いて行なうことを目的としている。具体的には、(1)大域プリフェッチ方式と単純プリフェッチ方式、(2)elastic barrier と単純バリアまたは Fuzzy barrier 方式、(3)キャッシュにおける可変プロトコルと固定プロトコル、(4)単純 full/empty 方式による同期と高性能同期構造体、(5)メモリアクセスにともなうレイテンシの変化を比較した性能評価を予定している。

お茶の水1号の構成の詳細は、引き続き発表で行なう。

### 3 おわりに

お茶の水計算機が目標とする汎用大規模並列計算機システムはお茶の水1号をクラスタとして階層化することにより実現される。お茶の水1号に引き続き研究・開発を予定しているシステムでは、クラスタ内部の効率化を更に促進するとともに、小規模ではあるがクラスタ間並列性を持ち込むことにより粗粒度と併用した場合の細粒度並列処理効率に関する課題を取り扱う予定である。

謝辞

本研究の遂行にあたり、様々な支援を頂いた(株)日本電気関係者各位に深く感謝致します。また、本研究は文部省科研費重点領域研究「超並列原理に基づく新しい情報処理体系」の一環として行なわれた。研究計画の推進に協力して頂いた研究代表者の東京大学田中英彦教授に感謝致します。

### 参考文献

- [1] Matsumoto, T. and Hiraki, K.: Dynamic Switching of Coherent Cache Protocols and its Effects on Doacross Loops. *Proc. of the 1993 ACM Int. Conf. on Supercomputing*, pp.328-337 (July 1993).
- [2] 中里 学, 大津 金光, 戸塚 米太郎, 松本 尚, 平木 敬: お茶の水1号の構成と評価. 計算機アーキテクチャ研究会報告 No.101-8, 情報処理学会, (August 1993).
- [3] 稲垣 達氏, 松本 尚, 平木 敬: 細粒度並列計算機用最適化コンパイラ: OP.1. 計算機プログラミング研究会報告, 情報処理学会, (August 1993).

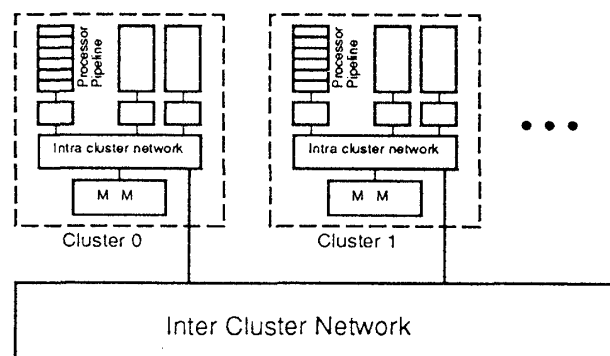


図 1: 超並列計算システムの基本構成