

# 計算機での並列処理

富士通株式会社 根尾定幸\*

## 1. 科学技術計算の高速化 (処理の高速化と並列化)

科学技術計算は、情報処理センターの重要な需要の一つである。対象となるプログラムには、研究者がプログラミングする「ユーザ開発プログラム」がある。この外には、業者提供ソフト、フリーソフトなども対象となる。本文では、これらのプログラムとユーザ開発プログラムを合わせて、ユーザプログラムと総称する。本文は、ユーザプログラムによる科学技術計算を、情報処理センターで行う場合に、遭遇する可能性の高い計算機を想定して、その範囲で並列処理を解説するものである。

本文では、計算の高速化のために計算機が用意できる基本的な方法の内、計算の処理動作の高速化と並列化について触れる。

動作の高速化の度合いは、デジタル計算機の場合、クロック周波数 (MHz) が有力な評価指標となる。クロック周波数の向上は、半導体の構造や集積度などの半導体テクノロジーに多くを依存する。が、処理命令を細かいステップに分割してデータを連続的に処理するパイプライン方式を採用すると、細かいステップを短いクロックタイムで処理できる関係から、クロック周波数を高くする方法もある。このパイプライン方式では、負荷の大きい乗除演算命令においても、クロックサイクル毎に次命令が開始されるために分割された処理ステップが並列処理されることになる。1命令処理自体には数クロックサイクルを必要としても、並列処理の効果が加わるので、命令あたりの平均クロックサイクル数 (CPI: cycle per instruction) を1程度に小さくできるため、クロック周波数の向上と合わせて、処理高速化の効果が大きくなる。

処理の並列化の技術では、「複数のパイプライン」と言う点が、スカラー計算機からベクトル計算機までの高性能計算機の共通の方法となっている。即ち、最近の高性能計算機においては、複数のパイプラインを持つプロセッサ、複数のプロセッサを持つ計算機、という階層構造を持つ処の共通性がある。各プロセッサにおいて、複数命令を同時並列処理する技術には、ベクトル計算機の方式と、WSのMPUとして発達したRISC-MPUの方式の二通りがあり、後述する。いずれの場合も、計算機内のパイプラインが空回りしないでユーザプログラムを処理できる仕組みが必要であり、要となる技術となる。この場合も評価指標は、並列処理される命令数と、それらの命令が並列処理されるのに必要なクロックサイクル数の比、に相当するCPIが使われる。このような並列処理の複合技術のおかげで、最近の高性能計算機用プロセッサのCPIは、1未満のレベルに達している。

本文では、情報処理センターの科学技術計算用の高性能計算機の特長を、並列処理をキー・ワードにして、横断的に観ようと思う。このため、計算機による並列処理の定義として、「ユーザプログラムに内在する、並列して処理しても良い関係 (並列性) を抽出して、計算機が並列に実行すること全般」を設定する。また、汎用機、ベクトル計算機、RISC-MPUサーバ、共有メモリ型並列計算機、分散メモリ型並列計算機などの情報処理センターの計算サーバとしての実績がある計算機を想定する。次章で、ここ10年程の計算機の進歩を、並列処理の側面から概観する。

注) RISC-MPU: Reduced Instruction Set Command - Macro Processor Unit

\*東京都品川区大井1-20-10 富士通大井町ビル、電話:03-3778-8281 FAX:03-3778-8107 電子メール:PDB02346@niftyserve.or.jp

なお、本文全体を通しての参考文献である「情報処理学会編 新版 情報処理ハンドブック」は、並列処理も含む計算機全般に渡って、研究レベルから実用レベルまでの事項、参考論文を網羅した良書である。

## 2. 計算機における並列処理の歴史

### 1980年代

この年代は、汎用機がセンターの主要な計算機であった時代である。後半には、ベクトル計算機の商業的成功がある。当時のセンターには高価すぎたが、90年代のセンター用計算機の並列処理技術、製品の先例、あるいは理解を助ける技術という意味で、ここに取り上げる。

汎用機で発達して、90年代にも主要な並列処理技術として受け継がれ、発展させられている技術には、プロセッサ内部の技術と、複数のプロセッサから計算機を構成する技術とがある。

プロセッサ内部の並列処理技術として、パイプライン、VLIWを取り上げる。

パイプラインは先に述べた様に、クロック周波数の向上と実効的並列処理の効果があり、RISC-MPUを特徴付ける基本技術の一つであり、ベクトル計算機を含めて今日の高性能計算機の共通の基本技術となっている。

VLIW (very long instruction word) は、複数の命令を1複合命令として1回の発行制御で済ませる。クロックサイクル毎に複合命令が1個発行されて、複数の命令が並列処理されることになる方式で、ベクトル計算機に採用された技術として、後述する。

複数プロセッサの構成技術には、コ・プロセッサ (co-processor)、マルチ・プロセッサ (multi-processor) がある。前者は、主プロセッサの処理の一部を、従属する専用プロセッサに任せる方式で、異なる内容の並列処理を行う方式である。最近では、高性能の描画処理ボードを内蔵した画像処理WSが相当する。大規模市場を狙う汎用プロセッサとは違い、専用プロセッサは市場規模が限られるため、開発投資に見合う効果が得られるような分野で、特殊目的の計算機で生き残る技術と思われる。

マルチ・プロセッサは、同じプロセッサを複数個、共有メモリに接続した構造のものである。MPU (micro-processor-unit) が汎用プロセッサとして低コストで大量生産される今日では、計算機製品の性能レンジを拡大する手法として普及が進むと思われる。80年代の汎用機と、最近の並列計算機とは、共有メモリに対してSMP (symmetric multi processor) である点と、OSは1個 (ブロック分割可能なVirtual Machineを含む) であることの両方は同じである。が、メモリとプロセッサの接続技術には、大きな違いがあり、後述する。

ベクトル計算機は、科学技術計算のプログラムの中の「ステップ数に比べ計算負荷が大きいループ計算の持つ並列性、規則性の高い繰り返し演算」を高速に処理することを目的とした計算機である。代表的な技術には、第一に演算パイプライン (演算処理のパイプライン化)、第二にVLIW、第三に大規模データ供給機構 (大容量の演算用ベクトルレジスタ設計と、主記憶と大量データを高速で入出力するロード/ストア処理のパイプライン化)、第四に自動ベクトル化コンパイラ (並列処理に関する高度な静的解析能力と並列化オブジェクト自動生成)、の4件がある。後の3件は、並列処理のハード機能を空回りさせないための技術と位置づけられる。

VLIW方式による並列処理では、プログラムの論理性を壊さない並列性保証が、方式の前提となる。この並列性保証は、ハードウェア・ソフトウェアからなる計算機システムとしての保証で、コンパイラが並列性保証付きの並列化オブジェクトを生成することに委ねられる。この関係は、VLIW設計の際に予め調整、考慮される要件である。オブジェクト実行中の並列性保証はノーチェックで、ひたすら計算性能を稼ぐ仕組みになっている。このため、コンパイラの静的解析能力と高並列化オブジェクトの生成能力とが、重要な技術課題となる。プログラム全体を一括して最適な並列性を抽出するために、サブルーチンを主プログラム内に自動展開して静的解析する「インライン展開」もコンパイラのユーザ向け機能として開発されている。さらに性能向上を図るための「動的解析の補助ツール」もユーザに提供されるようになっている。

ベクトル計算機処理は、ループ計算と大規模数値計算においては、今日なお強力で使い易い実績技術となっているが、特に日米メーカーの国際的な競争で生み出されたコンパイラ進化の成果は、90年代の各種の高性能計算機用コンパイラに受け継がれている。筆者は、小規模なWS用プログラムをベクトルコンパイラのメッセージに従って簡単な記述修正をしたプログラムが、元のWSで倍以上の早さ

で処理されることを体験して、ベクトル・コンパイラの静的解析力とチューニング用コメントが、人の気配りを上回ることを実感した。なお、当時のベクトル計算機は、ハードウェアの全てが最高の部品で構成されたため、価格は汎用機の数倍と高価であったが、何をやっても一番速い。特に数値計算は桁違いに速かったため、最高性能の計算機として、この機種がスーパーコンピュータの名で呼ばれることがあった。

### 1990年代

90年代は、WS発展型サーバとも呼ぶべき、RISC-MPUサーバが、情報処理センターの「スカラー計算処理の主力」として加わった。即ち、RISC-MPUは、半導体テクノロジーの技術革新を効果的に反映して性能向上とコスト・ダウン両面で革新的な進化を続け、MPU単体サーバとしてあるいは、スカラー並列計算機の単位プロセッサとして、情報処理センターに登場した。また、日本政府のスーパーコンピュータ調達規定の裏返しとして、もはやスーパーコンピュータとは呼べないが、ループ計算のコストパフォーマンスに優れたコンパクトな「ベクトル計算機」が、スカラー計算機と使い分ける形で情報センターに採用されるようになった。計算用サーバ多様化の時代を迎えたのである。これは、ユーザプログラム特性と計算機特性の組み合わせを、ユーザ自身が選択する時代になったことを示している。以下、RISC-MPU、並列計算機ハードウェア、並列処理OSの順に述べる。ベクトル計算機については、80年代の項を参照頂きたい。

### RISC-MPU

RISC-MPUは、簡単命令、複雑な制御がLSIに組み込まれて、小キャッシュメモリとの組み合わせ、クロックの高速化で性能を稼ぐと言う構造になっている。複雑な制御には、パイプライン制御、同時に複数の命令を発行する制御、動的スケジューリング制御などがある。いずれも並列処理の動的制御の機構である。コンパイラの並列オブジェクト生成とMPUの動的制御の両方の機能によって、ユーザ開発プログラムの自動並列化処理が実用化されている。

なお、スケジューリングとは、計算処理の過程でハードウェアの一部が空き状態になった時に、並列性を侵さない範囲でオブジェクト命令の順序を変えてハードウェアの空回りあるいは原因となるメモリアクセス競争を無くすことなどである。プログラマーによる高度なチューニング技術の一つであるが、RISC-MPUは、局所的解析力を持つハードウェアの制御機能として、動的チューニング機能を備えていることに特長がある。局所的解析故の限界があるためハードウェアが処理した後で、並列性チェックを行う機能も併せてハードウェア化されていて、無駄な処理をした結果に終わることもあるが、プログラム論理性は保証されており、平均的には計算高速化の効果がある。

代表的なRISC-MPUを比較すると、「小キャッシュと2本の高速クロック・パイプライン」、「大キャッシュと多数の低速クロック・パイプライン」の二つの製品傾向が見られる。

高速クロックのタイプには、スーパー・パイプライン方式のMIPS社のR-MPU、多分岐技術で最速クロックが特長のDEC社の $\alpha$ -MPUがある。

スーパー・パイプライン方式は、通常のパイプラインの並列処理の効果と、細分化された命令ステップに対応するクロックサイクルをさらに細分化して、クロック高速化の効果とを出している。 $\alpha$ チップの方式は、普通のパイプラインでは逐次処理される「レジスタ書き込みと次処理ステップ」を、分岐して並列処理するなど、クロックサイクルの短縮工夫が特長である。

なお、老舗の技術を革新して、互換性を維持した新設計により計算性能での他社追撃を開始したSUN社のULTRA-SPARC-MPUもこの区分に入る。

大キャッシュのタイプには、MByte級キャッシュと、他社の倍以上のパイプラインを備えて高性能化を図ったHP社のPA-RISC-MPUがある。

各社が登録しているSPEC-BMTの登録データによると、高速クロックタイプMPUの場合でも外部キャッシュ増設による性能向上を読み取ることが出来る。ベクトル機の項で述べた様に、高速計算機でのハードウェアを空回りさせないことの重要性との関係で興味深い。計算の高性能化と言う計算機側の事情とともに、ユーザプログラムの計算規模の拡大などによる、「計算機が参照するデータ容量が大きい」場合が、増えると考えられる。今後は、MByte級キャッシュとMPUの組み合わせによる製品が多くなるだろう。

以上に例示したMPUは、1チップLSIのものである。複数チップ構造のIBM社のPower-MPUは、パソコンから高並列計算機までに使われている。

## 並列計算機

複数のMPUで、科学技術計算のユーザプログラムを処理する場合の期待効果は、ユーザプログラム単体の処理時間の短縮（単体処理の向上）と、多数のユーザプログラム全体を処理する時間の短縮（スループット処理の向上）の、二つである。スループット処理の向上においては、多数の計算機をLAN接続するクラスターシステムのイメージと重なる点もある。が、並列計算機の場合、製品固有の高速ネットワーク（LANに比べて桁違いに速いデータ転送性能）、プロセッサ間の自動負荷分散、統一的なシステム管理機能などが備わっており、ハードウェアを空回りさせないのに役立つ実効環境と、運用管理の負荷軽減環境などがユーザの評価対象となると思われる。

## 並列計算機のハードウェア

共有メモリ方式と、分散メモリ方式に、大別される。共有メモリ型並列計算機は、ひとつのOS（ブロックに分割可能なVirtual Machineも含む）配下で、共有メモリと複数のプロセッサが稼働するタイプであり、接続方式の違いはあるが、汎用機マルチプロセッサと同じイメージの製品であり、SMPと呼ばれる。接続方式の違いは次の通りである。

汎用機では、小さな共有メモリを分割した全てのブロック $m$ 個と、数個から10数個の範囲のプロセッサの各データ口 $n$ 個との、それぞれを $m \times n$ 本の線で接続した構造であった。共有メモリ型並列計算機では、GByte級の大きな共有メモリのブロック $p$ 個と、数十個に達する多数のプロセッサのデータ口 $q$ 個とを、LANより格段に速い内部ネットワーク（ $p \times q$ より格段に少ない数の1~10数線の規模の高性能システムバスやクロスバー。単位線のデータ転送性能は、GByte/s以上）で接続する方式である。メモリとプロセッサ間のネットワーク内の通信の混み合いなどが新たなクリティカル・パスとなる可能性が加わった。クロスバーの方が高価であるが並列処理には、より都合の良いネットワークである。

分散メモリ型並列計算機は、メモリとプロセッサの組み合わせを単位とするPE（Processor Element）をLANよりも桁違いに速い内部ネットワークで接続するタイプであり、SPP（Scalar Pararell Processor）と、VPP（Vector Pararell Processor）とがある。PE毎に、メモリが分散していることと、OSが独立に走ることが特長となる。並列処理のためには、同期・データ転送などを含むプロセッサ間の通信処理が、OS間で処理されることとなる。プロセッサ間の通信用のライブラリとして、業者が提供するメッセージパッキングライブラリを、プログラマが使ってプログラミングする方式が、実用化されている。分散メモリ型並列計算機は、巨大プログラムの巨大計算の本命として、百から数万個のプロセッサで並列処理する発想が背景になっており、MPP（Massive Pararell Processor プロセッサ数が万程度以上の自称）とも呼ばれる。

## 並列処理のOS

共有メモリ型並列計算機において使われるOSの機能として、マルチ・プロセス、マルチ・スレッドを取り上げる。

OSは、ユーザプログラムを処理単位であるプロセスに分割して管理している。各プロセスは、他のプロセスから独立した処理単位としてその処理環境が保護されている、ために複数プロセッサに処理を割り当てることができる原理である。この原理で複数プロセスを複数プロセッサで処理できるのがマルチ・プロセスと呼ばれる並列処理OSの機能である。しかし、この方法は、プロセスの切替え毎に、保護対象の処理環境を切替える処理を伴う、と言う性能向上の阻害要因を含んでいる。

メモリ空間を共有する形で、プログラムを独立した処理単位のスレッドに分割する仕組みが考え出されて、マルチ・スレッド処理と呼ばれる。各プロセス・メモリを、大きな共有メモリ空間に共存させることに相当する発想である。必要な大きさの共用メモリ空間を用意できれば、複数スレッドを複数プロセッサで、メモリ空間書換えを省いて処理できることになる。各社とも実用化に力を入れ、出荷済または製品化計画を発表している段階である。

## 3. まとめ

様々な計算機における、並列処理の発想と技術的要素を、縦覧してきた。ここに解説した技術は、日々、新しい発想を付加し、洗練されている。年に2回を越えるペースでの新製品の発表が続いている。

プログラムの並列性を軸に述べて来たが、最後に、データ特性について触れる。

「一度使われたデータは繰り返し使われる可能性は高く、使われた近くのデータは次に使われる可能性は高い」と言う、汎用機の昔からの言葉があるが、「必要なデータは、使ったデータの近くにある可能性が高い」と言うのが、並列計算機の利用研究をみての感想である。このような特性が、分散メモリ型並列計算機の自動並列化の可能性に繋がって行くと思う。

## 参考文献

- [1] 情報処理学会編「新版 情報処理ハンドブック」 オーム出版局
- [2] SPEC-BMTのURL <http://www.specbench.org>