

研究室紹介

新しいコモディティ・コンピューティングのインフラの波に乗って

ベーナディ
アブドゥハン
Bernady O. APDUHAN

九州産業大学 情報科学部 知能情報学科
Faculty of Information Science, Kyushu Sangyo University
bob@is.kyusan-u.ac.jp, <http://www.is.kyusan-u.ac.jp/~bob/>

1. はじめに

ここでは、コモディティ・クラスタコンピューティングについて簡単にご紹介することにします。「コモディティ」は、一般には日用品とか商品などの意味ですが、この分野では、一般に普及している汎用計算機をネットワークでつないで並列コンピューティングを行おうというのがその考えです。最近急速に成長している分野で、近い将来高性能コンピューティングを支配することになるだろうと注目されています(図1)。

スーパーコンピュータや超並列コンピュータなどの高性能コンピュータは、高度に計算密度の高い問題を解いたり、その他の現実世界のアプリケーションを実行する強力な手段であることが知られています。しかしながら、こういったマシンは何千万ドルもするのです。わずかな予算の研究者達には途方もない値段です。

コンピュータを結ぶという考えは新しいものではありません。1950年代、60年代にはアメリカ空軍が真空管コンピュータのネットワークを作り上げたし、1980年代半ばにはDEC社が中規模のVAXミニコンピュータを統合してより大きなシステムを作り上げ、クラスタという用語を作り出しました。ワークステーションクラスタは、一般にはミニコンピュータクラスタほどの性能はなかったが、パソコンクラスタよりも速かったので、すぐどの研究機関でも使われるようになりました[1]。

1990年代始めに安価なパソコンが出現しネットワークの相互接続費用が下がったので、これまでよりずっと安価なコンピューティングインフラを作り出そうという研究者が現れました。つまり、安価な汎用パソコンをつなぎ、これらの普通の計算機を大変大きな問題に取り組みせるためのソフトウェアを描こうとしたのです。そしてLinux OSのようなソフトウェアの進歩によって、PCクラスタへの道が開かれました[2]。クラスタを定義しますと、相互接続された独立計算機の集合からなる並列あるいは分散型の処理システムで、協調して統一したコンピューティング資源として働くものと言えます[3]。

2. プログラミング環境

標準のプログラミングツールやユーティリティが利用できるので、クラスタは並列処理プラットフォームに代わるものとして実用的です。しかし、複数の独立計算機を協調して動かすためのプログラミング手法が重要です。最も一般のツールを見てみましょう

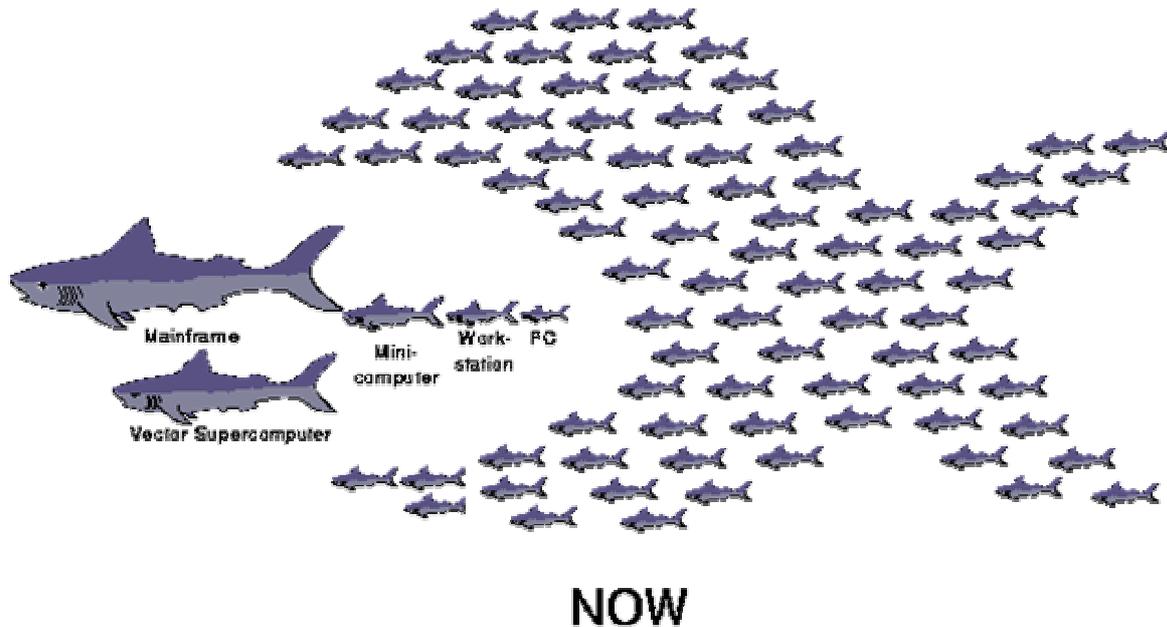
2.1 メッセージパッシングシステム (PVMとMPI)

計算機間にメッセージを受け渡す形で処理を行うシステムです。メッセージパッシングライブラリによって、分散メモリマシンのための効率的な並列プログラムを書くことができます。このライブラリは、データのペケットを送受信するのと同様、メッセージ受渡しの環境を開始したり、構成するといったルーチンを提供します。

PVMはメッセージ受渡しのライブラリを持った環境であり、ハイエンド・スーパーコンピュータレベルからワークステーションクラスタレベルまでのシステム上で並列アプリケーションを走らせるのに使用可能です。一方、MPIはメッセージパッシング仕様で、明示的なメッセージ受け渡しを使って分散メモリ並列コンピューティングのための標準となるように設計されました。このインターフェイスは、実用的でポータブル、効率的で柔軟性のあるメッセージパッシングのための標準を作ろうとするものです。MPIとPVMライブラリはFortran77、Fortran90、ANSI C、C++、mpiJavaに利用可能です。

2.2 分散共有メモリ (DSM) システム

メッセージパッシング方式の問題点は、共有メモリプログラミングシステムに比べてプログラムが複雑で難しい点です。共有メモリシステムは、全てのメモリを1つの塊として扱い、すべてのプロセッサに接続する方式で、シンプルで一般的なプログラミングモデルを提供しますが、問題はスケイラビリティ(大規模性)にあり、台数が増えると効率が悪くなることです。それに代わる費用効率のよい解決法は、分散メモリシステム上にDSMシステムを構築することです。これは、シンプルで一般的



NOW

図1 コンピュータの「食物連鎖」(現在および未来)

なプログラミングモデルを持ち、分散メモリシステムと同じスケラビリティを示します。

DSM は共有可変のプログラミングを可能にし、ソフトウェア、ハードウェアのいずれかに実装されます。ソフトウェア DSM システムの特徴は、普通通信インターフェイスの上部に別の層として作られ、仮想ページ、オブジェクト言語タイプは共有のユニットであることなど、アプリケーションの特徴の利点を十分に利用している点です。ハードウェア DSM システムの特徴は、性能がよりよいこと、ユーザとソフトウェアの層に負担がないこと、共有の粒度が高いこと、キャッシュ結合シェーマの拡張、ハードウェアの複雑性です。

2.3 スレッド

スレッドはユニプロセッサ、マルチプロセッサマシンの並行するプログラミングのためによく使われているプログラミング・パラダイムです。マルチプロセッサシステム上で、スレッドは本来、全ての利用可能なプロセッサを同時に利用するために使われました。これは重なった計算や通信のアプリケーションの非同期的な動作を活用するものです。マルチスレッドのアプリケーションはユーザインプットに対してより早い応答を提供し、より早く動作します。スレッドの生成は安価で管理しやすく、スレッドは、親のプロセスアドレス空間内に生成されるので、共有の変数を用いて通信します。

3. クラスタコンピューティング環境 DSM

1990年代の初め、私の元の研究室、九州工業大学知能情報工学科情報処理機構講座では、10Mbps のイーサネットによって結ばれたワークステーションのネットワーク上

に分散共有メモリモデルに基づいた分散処理環境のプロトタイプを開発しました。そして、この DSE を Windows 95/NT で走るパソコンのクラスタ、100 Mbp のイーサネットで結合した Linux プラットフォーム、ATM スイッチに移植しました。また、ハイエンドのコンピューティングプラットフォームである IBM SP2 マシンにも移植して実験を行いました。この SP2 は AIX OS のパワーワークステーションのクラスタで、低レベルアプリケーションプログラミングインターフェイス (LAPI) を持った専用の IBM-SP スイッチで相互接続されたものです [4]。

これらの実験により、通信オーバーヘッド、ネットワークポロジの効果、オペレーティングシステムやプログラミングモデルやアプリケーションがシステム性能に与える影響などについて明らかにしてきました。

4. クラスタの現在とこれからの動向

クラスタシステムの初めの世代が作られて以来、クラスタは急速に伸びてきました。新しいクラスタはより低いコストでより高速の処理を実現し、性能曲線を常に上昇させてきました。大手の計算機会社は今ではクラスタを大きな計算需要を持つビジネスに売り込んでいます。2002年の11月に発表された世界の最速コンピュータのトップ500に、パソコンのクラスタ、ワークステーションクラスタ、SMP (Symmetric Multi-Processor) が入っています。トップ500のスーパーコンピューティングのウェブサイトはLinuxクラスタを世界で5位の強力な計算機に挙げています [5]。

クラスタはデータベースやファイルサーバ、ウェブサーバやシミュレーションアプリケーションに使われる以外に、広範なグリッドコンピューティング技術の分野での

計算リソースとしても考えられています [6]。クラスタシステムは発電所網のような「計算用グリッド」の一部になるかもしれません。そうすれば、ユーザは私たちが今電気を使うように簡単に計算処理能力を使うことができるようになるでしょう。グリッドの概念の背後にある実際の問題は、「ダイナミックな多機関による仮想組織において、協調してリソースを共有し、問題を解決する」ことです。関心を集めている共有とは、元々ファイル交換ではなく、むしろ計算機やソフトウェア、データ、その他のリソースに直接アクセスすることであり、それは現在産業や科学工学の分野で起こっている協同の問題解決やリソース管理という範囲で必要とされています。

プロセッサ、メモリ、ディスク、ネットワークなど新しい機器が利用できるようになると、それらはデスクトップやサーバードに統合されやすいので、クラスタはそれらの進歩から最も利益を受けやすい並列システムであると言えます。そのようなハードウェア技術の進歩を受けて、それを支援するソフトウェアシステムやツールの機能の充実や成熟が重要とされてきました。コモディティクラスタは技術革新の道 (OpenMP[7], Java[8], VIA[9], Infiniband[10] など) を開き、新しい構成要素の出現に最も素早く対応するものだと言えるでしょう。

とりわけ、コモディティクラスタの概念は力を与えるものです。特権的な少数だけのものだった高いレベルの計算処理を、廉価な並列処理システムでリソースの少ない人々にも利用可能にするのです。

我々はクラスタコンピューティングについて研究活動を続け、この分野の諸問題によりよい解決方法を探り、クラスタが様々な分野で幅広く利用されるように貢献したいと考えています。また、クラスタコンピューティングの教育も行い、クラスタコンピューティングを並列コンピューティングの主流にしていこうと考えています。

◇ 参 考 文 献 ◇

- [1] <http://now.cs.berkeley.edu/>
- [2] W.W. ハーグローブ, F.M. ホフマン, T. スターリング, "中古パソコンでスーパーコンピューターを作る", 日経サイエンス 2001 年 11 月号。
- [3] <http://www.ieeefcc.org>
- [4] Y. Shimono, B.O. Apduhan, I. Arita, Y. Ohnishi, "Evaluating the Performance of a DSM Cluster with Improved Communication Subsystem", Proc. ICOIN-15, pp. 561-567, Jan. 2001
- [5] <http://www.top500.org>
- [6] F. Berman, G. Fox, T. Hey, "Grid Computing: Making the Global Infrastructure a Reality", John Wiley and Sons Ltd., 2003.
- [7] <http://www.openmp.org>
- [8] <http://www.javasoft.com/>
- [9] <http://www.viarch.org/>
- [10] <http://www.infinibandta.org>