

ミニ・ミニ・スーパーコンピュータ ー作ってみようー

村上佳久

筑波技術大学 障害者高等教育研究支援センター 障害者基礎教育研究部 (視覚障害系)

要旨: スーパーコンピュータ (以下: スパコン) の超小型タイプを製作し, スパコンとは何かについて原寸大で教育できるような教材化を行い, 検証した。実際, 作製した超小型スパコンは, 東京工業大学の大型スパコンである "Tsubame 2.0" とソフトウェア面で互換となるように設計し, 将来的には発展が見込めるようなものが実現できた。

キーワード: Tsubame 2.0, GPGPU, Tesla, CUDA, Xeon Phi

1. はじめに

現在世界のスパコンの順位は "Top500"[1] として, 2013年6月17日に発表された順位は以下のようになっている。

- 一位: 中国「天河2号」
- 二位: アメリカ「タイタン」
- 三位: アメリカ「セコイア」
- 四位: 日本「京」

これらのスパコンは, 科学技術計算や大規模気象シミュレーション, 理論核実験や製薬における薬効の分析などに幅広く活用されており, スパコンの活動する領域は年々大きくなってきている。しかし従来のスパコンは専用のハードウェアの利用していたため, その利用法には専用のプログラムを習得する必要がある, 汎用性にはほど遠いものであった。しかし, 昨今のスパコンでは, 汎用性が増し, 誰でも利用できる体制が整いつつある。例えば, 東京工業大学の大型スパコンである "Tsubame"[2] は, 「みんなのスパコン」として, 東京工業大学の学生は全員が利用できる環境にある。

一方, 筑波技術大学のような小規模の大学では, 自前でのスパコン導入は不可能である。そこで, 本研究では, 他大学 (例えば東京工業大学) の大型スパコンのミニミニタイプを製作し, できるだけソフトウェア互換のシステムで, スパコンのハードウェア学習とソフトウェア学習の両面に渡って利用できるようなシステムを教材化し, 開発したので報告する。

2. 現在のスパコンの現状

はじめに現在のスパコンの現状について, 幾つかの種類を整理分類すると, 次の3つのスタイルに分かれる。

- a) 特殊で高速な CPU を開発し利用する
- b) 低速な CPU を非常に多数使用する
- c) GPU を物理計算プロセッサとして多数使用する

このような, 幾つかの種類スパコンが登場した背景には, 従来の専用型スパコンに比べて, 最近では非常に高速な計算が求められる一方で, 消費電力や維持費などの諸費用も桁違いに大きくなって来たことが一因と思われる。このため, 2010年代以降のスパコンでは, 消費電力と機器冷却費用, 維持管理費なども含めて考慮されるようになり, エネルギー消費効率のよいスパコンランキングである, "Green500"[3] も発表されるようになり, 導入費用に対する計算能力以外にエネルギー効率も求められるようになってきた。したがって, このようなスパコンのスタイルは, 各々に利点があり今後ともこのような分類で発展していくものと考えられている。

2. 1 特殊で高速な CPU を開発し使用

2000年以前のアメリカのクレイ社や日本における「地球シミュレータ」[4], 現在の理化学研究所の「京」[5]などがこのカテゴリーに分類される。

「地球シミュレータ」では, 汎用科学技術計算機として構築され, NEC製の "SX-5" と呼ばれるスパコンを基に1ノードCPU8個で, 640ノード5,120個のCPUで構築され, 2002年6月に35.86TFlopsで世界一となった。後に, NECのスパコン "SX-9/E" ベースに更新され, CPU8個の計算ノードを160台つないで, 2009年3月に131TFlopsとなり, システムアップした。

一方, 「京」では, 計算速度が毎秒1京回(10P FLOPS)として設計され, 1ラック102ノードで, 1ノードが1CPU。これを864ラックつないで, 88,128個のCPU (SPARK SPARC 64 VIIIfx) からなり性能は

10.51P FLOPS である。

これらのスパコンは専用のスパコンを開発するため非常に高額な費用が掛かるため、現在は、ほとんど行われていない。日本でも 2009 年の政府の事業仕分けにおいて大きな問題となり、実際の価格性能比でも維持費の面でもほとんどメリットはなく、今後はこのスタイルは、専用の目的のみの機能に絞って存続していくものと思われる。

2. 2 低速な CPU を非常に多数使用

一般用のパソコンに使われていた CPU 開発において、動作周波数を上げて発熱量が多くなり性能が向上しないとあった状況が、2000 年以降続いてきた。そこで、この事態を解決する方法の 1 つとして、SONY のゲーム機である "PS3" において "cell" と呼ばれる 1CPU 当たり 9 個のコアを有する CPU を開発し、1 個のコアの性能はそれほどでもないが、8 個のコアを利用すると性能が向上するようなシステムが考えられた。9 個のコアがあるのに 8 個のコアしか使わないのは、製造上の歩留まりを考慮したためである。これを嚆矢として、多数のコアを利用するメニーコアタイプの CPU がスパコンにも利用されるようになってきた。この例が、IBM 系と Intel 系である。

IBM 系の場合は、"Blue Gene" と呼ばれるシステムで、PowerPC 系の CPU を非常に多数使用したシステムである。1 つ 1 つは高性能ではないが、省電力性に優れ、最高で数十万個の CPU を搭載し、並列処理させることでパフォーマンスを得ている。2004 年の Blue Gene/L は、CPU 数 37,768 個で、70.72T Flops に達し、その当時世界最速の「地球シミュレータ」を抜いて世界一となった。最新の Blue Gene/Q では、CPU 数は、約 99,000 個で、2012 年 6 月に 16.32P Flops でその時点で世界最速であった。

Intel 系の場合は、"Xeon Phi" と呼ばれるシステムで、1 つの PCI/e x16 用拡張スロットに 60 コアの Xeon プロセッサと共有メモリなどを搭載し演算処理が可能とした。これを基となる Dual Xeon CPU 搭載のラックマウントサーバに 3 台ほど組み込み、非常に多数の Xeon CPU でシステムを構成する。1 台のラックマウントサーバ当たり換算して 4T FLOPS の計算能力があり、これを多数並列接続するとスパコンとして運用可能である。実際に現在スパコンランキング "Top500" の一位である「天河 2 号」は、この方式で構築されている。

この CPU を多数利用するタイプでは、価格性能比が非常に高く、今後ともこの種のシステムが増えてくるものと思われる。

2. 3 GPU を物理計算プロセッサとして多数使用

GPU は、画像処理専門の IC で、動画信号生成専用 IC である。動画像の実時間内での生成は、高負荷な演算能力が要求されるが、その多くが定式化された単純演算の繰り返しなので専用ハードウェア化に向いており、1990 年代以降に 3D 描画性能が劇的に向上した。それに伴い行列演算を中心とした演算機の色が濃くなり、2000 年代に入ると、プログラブル・シェーダへ移行し、グラフィック・レンダリングのみならず、他の数値演算にも利用する事が可能となってきた。

GPU の浮動小数点演算能力は 2010 年頃に 2T FLOPS をオーバーした一方で、CPU の浮動小数点演算能力は向上せず、2010 年代で数 G FLOPS 台に留まっている。GPU は構成が単純で浮動小数点演算などの効率がよく、GPU 専用メモリのバンド幅も広く、主な用途がパソコンゲームと動画再生で、ゲームをしないユーザーにとっては、持てあまし気味の機能で注目されている。この余剰機能を汎用計算に利用するのが、GPGPU である。

GPGPU (General-purpose computing on graphics processing units) とは、GPU の演算資源を画像処理以外の目的に応用する技術のことである。この技術は、Windows パソコンの発達とともに進化し、特に DirectX 9.0 の登場以降、nVIDIA による GPGPU 専用のプログラム統合開発環境 "CUDA" や、AMD の "ATI Stream" が現われ、GPGPU 活用の幅が広がっている。AMD 社は "AMD FireStream" を nVIDIA 社は "Tesla" を投入し、GPU を GPGPU 化する傾向はここ数年、非常に進化している。特に nVIDIA 社の "Tesla" を GPGPU として利用するスパコンは多数に上り、東京工業大学の "Tsubame" もこのタイプである。

現在の "Tsubame" は、1 台のラックマウントサーバに CPU2 基とメインメモリ、GPGPU 用の "Tesla" を 3 台搭載して、これを 1 ノードとして 1,408 ノードとサブシステム 34 ノードで構成され、1.1P FLOPS (ピーク性能: 2.4P FLOPS) を誇っている。

3. ミニミニスパコン

3. 1 ミニミニスパコンの検討

前章でスパコンの現状を整理したが、視覚障害者も利用可能なミニミニスパコンについて検討する。

3 種類あるスパコンのうち、a) は、大規模構築事例

なので事実上不可能である。そこで、b)や c)について検討する。b)の IBM 系の"Blue Gene"は、大規模構築事例なのでこれも事実上不可能である。b)の INTEL 系は、"Xeon Phi"を手に入れば構築可能であるが、秋葉原等の一般市場でも個人が入手することは不可能のようである。専門のベンダー向けには出荷されており、卓上のミニミニスパコンとして各社から販売されている。したがって、これらを購入して教材化することも考えられる。

c)の GPGPU に関しては、nDIVIA 社の"Tesla"が秋葉原等の一般市場でも容易に入手可能なため、個人でもミニミニスパコンの構築が可能である。そこで、本研究では、c)の GPGPU を活用した nDIVIA 社の"Tesla"を用いたミニミニスパコンを製作し教材化することとした。

3. 2 ミニミニスパコンの設計

教材化に当たって、価格も考慮しなければならない。GPGPU 用の nDIVIA 社"Tesla"は、1 台 20 万円程度するため非常に高価である。また、この"tesla"は電力を多量に消費するため、1 台当たり設置可能な"Tesla"の数には制限がある。これは主として、電源の問題である。さらに主計算を行うプラットフォームとしてマザーボードに CPU を何個搭載するかという問題である。この、価格・電力・CPU 個数・GPGPU 個数を検討し、設計を行う。

教材化に当たっては、次の 2 つのことを考慮する。1 つは、汎用性の問題である。ミニミニスパコンには様々な用途があるが、特定用途向けのミニミニスパコンもある。例えば、ハイビジョン・ビデオ編集や 3D アニメーション作製中の実動画編集や自動車設計などにおけるシミュレーションなど大規模スパコンは必要ではないが、一般の高性能パソコン（ワークステーション）では、少し能力が不足するような場合である。このようなミニミニスパコンは汎用性があり、特定のソフトウェアとともにその能力を発揮し、汎用科学技術計算やシミュレーションも可能である。

もう一つは、大規模スパコンの 1 ノード程度の小規模ミニミニスパコンである。こちらの方は、汎用性業務と言うよりも純粋に大規模スパコンのプログラム学習用として機能し、ミニミニスパコンで書かれたプログラムが、大規模スパコンに導入しても動作するように設計する。

以上から、2 種類のミニミニスパコンを設計し構築する。

3. 2. 1 汎用性ミニミニスパコンの設計

この設計では、消費電力を 400W 程度として、通常のワークステーション級として設計する。

CPU として、Core i7 を 1 個とメインメモリ 32GB で構築する。GPGPU として nDIVIA 社の"Tesla"を 1 台、通常の GPU として同じ nDIVIA 社の"Quadro"を導入する。この GPU と GPGPU の組合せは、nDIVIA 社の「Maximus テクノロジー」を活用するためである。

一般にパソコン用グラフィック API には、Microsoft 社の Direct X と Khronos グループが策定する OpenGL がよく知られている。nDIVIA 社の「Maximus テクノロジー」は、主計算を CPU に、グラフィックス API を OpenGL として"Quadro"GPU に、並列計算用の CUDA プログラムを GPGPU 用の"Tesla"に適宜振り分けるユニファイド・ドライバであり、対応する業務用ソフトウェアに対して、CPU, GPU, GPGPU を最適化することが可能となる。例えば、3D アニメーションソフトにおいて、デザインとシミュレーションを一体化して作業することが可能となり、リアルタイムで業務が行える。この"Tesla"と"Quadro", "Core i7"の組合せで、消費電力は常時最大稼働時で約 400W 程度となり、電源容量としては、約 2 倍の 750W を用意する。

マザーボードは、Intel Z77 チップセットのもので、"Tesla"と"Quadro"が最高速度で動作可能な PCI/e x16 が 2 本利用できるものを選択する。

さらに、HDD を大規模データ用に SSD をシステム起動用に利用して処理の高速化を図る。

3. 2. 2 小規模ミニミニスパコンの設計

この事例では、パソコンの電源容量を秋葉原等の一般市場で入手可能な最大級の 1500W として逆算して設計する。電源容量の約半分の 800W 程度を常時最大負荷で動作すると仮定すると、GPU と GPGPU が各々約 220W であるため、CPU には 130W 程度しか利用できない。サーバ用のラックマウントでは、電源容量を常時最大負荷で 1000W まで可能であるが、秋葉原等の一般市場で入手可能なものでは、常時最大負荷は 800W 程度が限界である。

CPU でコア数が多く 130W 程度のものは、"Intel Xeon E5-2600"シリーズが利用可能である。GPGPU として"Tesla"2 台、GPU として"GeForce GTX"を 1 台利用する。GPU は、Windows や Linux で動作させるため、より汎用性の高い GPU ボードを選択した。

この組合せでは、マザーボードに Intel X79 か C602,

C604 チップセットが必要となる。さらに、GPGPU2台とGPU1台が最高速度で利用可能なPCI/e x16が3本以上利用できるものを選択する。

また、非常に高熱で長時間の動作に耐性のあるシステムにするためには、システム用、データ用HDDともに高可用性のものを使用して信頼性を高める。

3. 3 ミニミニスパコンの構築

前項の設計に基づき、実際のシステム構築を行う。

3. 3. 1 汎用性ミニミニスパコンの構築

設計に基づいて、秋葉原等の一般市場で部品を調達した

M/B:ASUS P8-Z77 WS
CPU:Intel Core i7-3770
GPU:nVIDIA Quadro 600
GPGPU:nVIDIA Tesla C2050
RAM:32GB
SSD:256GB
HDD:2TB
P/S:850W Gold
OS:Windows 7 Professional 64bit

組み立てている様子を図1に示す。



図1 汎用型ミニミニスパコン

(Core i7-3770, Quadro 600, Tesla C2050 搭載)

このシステムでは、アプリケーションソフトとして"nVIDIA Maximus"対応の Adobe Premiere Pro CS6を導入した。ハイビジョン・ビデオ編集において、編集と再生をリアルタイムで行うことが可能なシステムとなる。また、ソフトウェア開発ソフトとして、

Microsoft Visual Studioを導入し、さらにnVIDIA社のGPGPU向けのC言語の統合開発環境である、CUDA (Compute Unified Device Architecture: クーダ)を導入して、アプリケーション対応とソフトウェア開発環境の両方に対応した。

3. 3. 2 小規模ミニミニスパコンの構築

M/B:ASUS P9-X79E WS
CPU:Intel Xeon E5-2670
GPU:nVIDIA GeForce 570
GPGPU:nVIDIA Tesla C2070 x2
RAM:64GB
HDD:1TB + 2TB
P/S:1250W Gold
OS:Windows 7 Enterprise 64bit
OS:SuSE Linus Enterprise Server

このシステムでは、OSにWindows 7とSuSE Linuxを導入し、ソフトウェア開発ソフトとして、Microsoft Visual Studioを導入した。さらにnVIDIA社のCUDAも導入し、WindowsとLinuxの両方の環境でソフトウェア開発環境を整えた。

組み立てている様子を図2に示す。



図2 小規模ミニミニスパコン

(GeForce GTX570, Tesla C2070 x2 搭載)

4. 教材化

本研究の課題である教材化について検討する。

はじめに、スパコンの意義である。本学のような小規模の大学では大規模なスパコンは運用できない。しかし、他大学の大規模スパコンを利用することは可能である。そこで、本学において大規模スパコンの機能と仕組み、スパコンが必要な科学技術計算や新薬開発などの創薬技術、大規模気象シミュレーションなどに

ついて学習した上で、大規模スパコンと同様のプログラム開発環境を提供し、並列処理のプログラミング学習を行うことはスパコンのプログラミングを理解する上で、極めて重要な課題である。その意味でミニミニスパコンを実際に運用する意味はある。

しかし、最も重要なのはハードウェアの構築である。現在のスパコンがどのような仕組みで構築されているか、また実施のハードウェア構成がどうなっているか、さらに設計思想や構築事例などの系統だった学習もまた、スパコンを理解する上で重要な学習事項であり、大規模スパコンの一部分に過ぎないとはいえ同等の部品を有しており、視覚障害者が部品を実際に触りながら、スパコンのハードウェアの基本部分の構築を実際に経験しながら行うことは、大規模事例のハードウェアを理解する上で、設計思想や設計哲学、設計技術や構築技術、ハンドリングや冷却技術などスパコンの周辺技術も含めた総合的な学習の手助けになると思われる。そして、このハードウェアの構築と実際のプログラミングは、スパコンを理解する上で、視覚障害学生に対する教材として有用であると考えられる。

また、教材としての利用と、実際の運用や構築に関わるノウハウの習得にも教育上の効果が期待できる。今回の2つの事例では、その性能を引き出すためには電源やマザーボード、HDDの選択など一般的なパソコン組み立てとは異なった知識と理論が必要となる。

実際に、Intel Core i7搭載のパソコンで、PCI/e x16がフルレーンで2本利用できるシステムを見つけ出すことは極めて難しい。また、X79チップセットでPCI/e x16がフルレーンで4本利用できるシステムも同様である。

システム構築に当たったの制約が多いことから、本来はメーカー製のワークステーション級の機器に”Xeon Phi”やGPGPUボードを搭載し、専用ベンダーから販売されるものを選択するのが一番と思われる。しかし、学習する上では、これらのハードウェアのスペックを理解した上で、組み立てを行うと様々な困難に直面し、メモリ相性問題やHDD利用制限など一般に知られていない知識が要求される。パソコンとワークステーションでは要求される知識が異なるからである。その意味で、本研究では、ミニミニスパコンに求められる機能から必要なスペックを割り出し、システムを構築するという手法を取っている。

したがって、スパコンと言う条件を満足するための様々な知識やノウハウを学びつつ、実際に使われる部品や機材などを触りながら確認し、さらに組み立てている様子を写真に撮り各々について解説を加えながら、スパコン構築と言う流れを教材化して、最終的にソフトウェアを利用し、科学技術計算を行うようになるまでの機器を構築することが重要であると考ええる。

このように単に「ミニミニスパコンを組み立てた」と言うだけでなく、組み立ての当たった設計思想や方法、組み立て知識等を系統立って学習することにより、スパコン構築が学習教材として有用なものとなる。

5. おわりに

今回の構築事例に要した価格は決して安価ではない。GPGPU用ボードは、1台で約20万円もする高価なものである。しかし、このようなミニミニスパコンは、実際に構築してみないとわからない部分が非常に多いことに驚かされる。教科書などもなく、ノウハウ的な経験と勘に頼らざるを得ない面が多く、技術の蓄積のためには、経験を得ることが一番と思われる。したがって、コンピュータを利用するだけでなく実際に自分でシステムを構築することが、如何に重要な問題であるかを思い知らされた次第である。このようなノウハウは、教室での授業には向かないが、システムエンジニアなどの教育には、体験型学習として必要不可欠と思われる。

今回は、設計と構築面での事例のため、検証しなかったが、次の機会にこのミニミニスパコンを利用したスパコンプログラミング技術について報告する予定である。

参考文献

- [1] Top500, <http://www.top500.org/>
- [2] “Tsubame2.0”, 東京工業大学 学術国際情報センター, <http://www.gsic.titech.ac.jp/tsubame>
- [3] Green500, <http://www.green500.org/>
- [4] 「地球シミュレータ」, 独立行政法人 海洋研究開発機構, <http://www.jamstec.go.jp/es/jp/index.html>
- [5] 「京」, 独立行政法人 理化学研究所 計算科学研究機構, <http://www.aics.riken.jp/>

A Micro Supercomputer

MURAKAMI Yoshihisa

Research and Support Center on Higher Education for the Hearing and Visually Impaired,
Tsukuba University of Technology

Abstract: A micro supercomputer was created as an educational tool for visually impaired students. The effect of this educational system that used the supercomputer on visually impaired student learning was verified. A micro supercomputer was designed to become interchangeable with the supercomputer named “Tsubame 2.0” of the Tokyo Institute of Technology. We expect to develop the software and the target in the future.

Keywords: Tsubame 2.0, General-purpose computing on graphics processing unit, Tesla, Compute Unified Device Architecture, Xeon Phi