



## SC06 レポート

この 10 年で大きく変わったコンピューティング・プラットフォームの構成において、現在は、標準化され広く流通しているコンポーネントを効率よく組み合わせることで、製品化されたプラットフォームが、HPC 分野でも主流となっています。もちろん、特別に設計され、特定の目的のために最適化されたシステムの価値は、依然として高いものがあります。しかし、今後 10 年の HPC システムの動向を考えた場合、また今後予想されるワークロードや利用モデルにおいてのプラットフォームの進化を考えた場合、標準コンポーネントを利用した HPC プラットフォームが大きな位置を占めることは間違いありません。このような HPC プラットフォームの方向性として、弊社では、「HP2C」というコンセプトを提唱し、そのようなコンセプトを実現するためのプラットフォーム事例を積極的に紹介しています。

米国フロリダ州タンパで開催されたハイパフォーマンスコンピューティング、ネットワーキングおよびストレージ関連技術のための年次会議である「Supercomputing 2006」でも、弊社が HP2C システムとしてご紹介している NEXXUS4000 と FUSION1200 をインテル社のブースで展示され、デモが行われています。

インテル社は、この SC06 の開催前に、「Intel Boosts Super-Computing Efforts」というニュースリリースを行って、その中で次の 3 点を特に強調して説明しています。

1. **QUAD コアプロセッサ**
2. **クラスタリングのための新しいソフトウェアツール群**
3. **ラックとコストを低減する新しい HPC システムの構築ブロック**

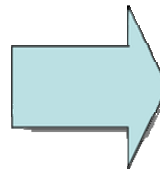
このレポートでは、上記の 3 点に関して、NEXXUS と FUSION がどのような位置づけにあるかをブース展示の写真と一緒にご紹介し、HPC プラットフォームの今後の動向を考えたいと思います。

## QUAD コアプロセッサ

1996年にインテルが、7264 個の Pentium® Pro プロセッサを搭載したスーパーコンピューターを開発し、1.06 テラフロップスの記録 (Linpack\* ベンチマーク) を打ち立てることで破られました。このシステムは、57 台のキャビネットで構成され、その後、9200 個のプロセッサ、84 台のキャビネットに拡張され、最終的な理論最大性能は、1.8 テラフロップとなりました。同じピーク性能を実現するのに、現在の QUAD コアプロセッサでは、44 個のプロセッサで十分であり、1500ft<sup>2</sup> (140m<sup>2</sup>) の設置面積は、16ft<sup>2</sup> (1.5m<sup>2</sup>) で十分です。消費電力も、800,000 Watts から 10,000 Watts となっています。



1996年にインテルが、7264 個の Pentium® Pro プロセッサを搭載したスーパーコンピューターを開発し、1.06 テラフロップスの記録 (Linpack\* ベンチマーク) を打ち立てることで破られました。このシステムは、57 台のキャビネットで構成され、その後、9200 個のプロセッサ、84 台のキャビネットに拡張され、最終的な理論最大性能は、1.8 テラフロップとなりました。



同じピーク性能を実現するのに、現在の QUAD コアプロセッサでは、44 個のプロセッサで十分であり、1500ft<sup>2</sup> (140m<sup>2</sup>) の設置面積は、16ft<sup>2</sup> (1.5m<sup>2</sup>) で十分です。消費電力も、800,000 Watts から 10,000 Watts となっています。

既に、マイクロプロセッサは複数の演算コアを実装したデュアルコアプロセッサが一般化しています。今回のSC06では、4つの演算コアを搭載したクアッドコア インテル Xeon プロセッサ5300番台の発表にあわせて、クアッドコアプロセッサを搭載したデル社製のサーバ(写真左)を展示しています。クアッドコア インテル Xeon プロセッサ5300 番台は、64 ビット・コンピューティングに対応したインテル Core マイクロアーキテクチャーを採用し、大容量の8MB L2 キャッシュ(2つのコアにつき1つの4MB共有L2 キャッシュ)が統合されたプロセッサであり、システム性能を大きく向上させると同時にプロセッサの消費電力を低減させることが可能となり、DPサーバの密度と消費電力あたりの性能の向上を実現しています。



クアッドコアプロセッサを利用したサーバの展示では、9月のIDFでも紹介されたFUSION 1200(右側写真)も展示されています。FUSION 1200は、「共有メモリー並列計算(SMP)」システムであり、コンパクトな筐体に48コアを搭載した共有メモリアーキテクチャのシステムとなり、ブースでは、48WayのSMPサーバとして、その性能をデモしています。従来のRISCベースのSMPシステムでは各社が独自のチップセットとインターコネクトなどを開発することが一般的でした。このFUSION 1200は、プロセッサ・チップセットとしては、インテルのDPプラットフォームを利用し、そのDPプラットフォームの持つ高いフロントサイド・バス(FSB)、FB-DIMMの高いスループットと大容量のメモリー、そして、PCI Express(PCIe)の高速性と高い信頼性を最大限に活用し、スケーラブルなSMPシステムの構築を可能としています。

## クラスタリングのための新しいソフトウェアツール群

並列アプリケーションの開発には、多くの試行錯誤とプログラムの実行と検証、デバッグといった作業が必要になります。そのプログラムについては、その実行性能の向上を図るためのソースコードの書き換えなどの作業を、プログラムの開発中も開発後も行う必要があります。このような作業を効率良く、また短時間で 行うには、優れた開発環境が必要です。また、そのプログラム開発をより容易に実行するために OpenMP などの共有メモリプログラミングをクラスタ環境にまで拡張した Cluster OpenMP なども利用可能となっています。クラスタリングのための新しいソフトウェアツールのデモも NEXXUS 上で実行されていました。



Intel Cluster Tools には、以下のようなツールやツールキットが含まれています。

- Intel Cluster Toolkit
  - ・ Intel MPI Library
  - ・ Intel Trace Analyzer and Collector
  - ・ Intel Math Kernel Library Cluster Edition
- Cluster OpenMP for Intel Compilers

共有メモリ用の API である OpenMP を MPI のようなメッセージパッシングを利用することなく、クラスタ環境で利用できるようにするのが、Cluster OpenMP です。Cluster OpenMP では、陽的なメッセージ通信を行うことなく、また、ノード内とノード間のプログラミングモデルや API を変えることなくプログラミングを行うことを可能とします。

Cluster OpenMP では、通常のクラスタシステムに HW の変更を加えることなく、ソフトウェアによって分散共有メモリのサブシステムを構築し、クラスタ全体を一つの共有メモリのシステムとして、OpenMP を同様の共有メモリとしてのプログラミングを可能とします。

インテル クラスタ・ツールはクラスタ上でのアプリケーション開発を効率化し、た最新のプロセッサ、プラットフォーム技術に対応したソフトウェアを効率よく開発できます。クラスタを利用することで、より低い導入コストの実現は可能ですが、その上で生産性の高いアプリケーション開発が出来ないと、TCO の低減を図ることは困難です。

## ラックとコストを低減する新しい HPC システムの構築ブロック

クアッドコアプロセッサを利用することで、高密度で低消費電力の HPC プラットフォームを構築することは、より容易になります。しかし、HPC プラットフォームでは、一般のデータセンターで利用されるよりも高いメモリバンド幅がシステムに要求されます。また、HPC システムでは、非常に大規模なシステムを構築し、HPC システムは、数千プロセッサを同時に利用し、数テラバイトのメモリを利用するような高度なシュミレーションが行われます。そのようなシステムでは、更に高密度でのシステムの実装が求められます。

このような HPC システムの要求に対応するために、インテルは幾つかの高密度実装用のサーバボードを提案しています。今回の SC06 では、それらの最新サーバボードの展示も行われています。

このような HPC サーバボードとして製品化された最初のサーバボードは、‘Caretta’ と呼ばれる UP サーバボード インテル サーバボード SE7230CA1-E です。このサーバボードは、あえて 1 ソケット構成を採用することで、コア当たりのメモリー帯域幅やネットワーク帯域幅を確保すると同時に、既存の ATX フォームファクターをちょうど半分に切ったようなコンパクトさを実現しています。この ‘Caretta’ の後継サーバボードとして、インテル Core2 プロセッサ・ファミリ (開発コード名: Conroe) やクアッドコア・プロセッサをサポートしたサーバボードとして、S3000PT (開発コード名: Port Townsend) が、SC06 では紹介されています。さらに、SC06 では、DP サーバボードにも関わらず、S3000PT とほぼ同じ大きさのサーバボード (開発コード名: Atoka) が発表になりました。このサーバボードを利用した場合、1U サイズのサーバ内に、4 プロセッサ、16 コアが搭載可能となります。デュアルコアプロセッサを搭載した今までの 1U サーバに対して、4 倍のプロセッサコアの搭載が可能となります。これらのサーバボードは、HPC システムが必要とする絶対性能、コスト、設置面積、消費電力の低減などを図るために新しいプラットフォームの構成ブロックとして利用可能となっています。

インテルのブースでは、S3000PT を搭載したサーバやブレードが展示されています。また、S3000PT と搭載したサーバは、インテルのブースだけでなく、他の出展企業のブースでも紹介されています。インテルのブースでは、Ever Case Technology 社のブレードが展示されていました。この製品は、4U のサイズで、10 枚の S3000PT の搭載が可能となっています。また、S3000PT 搭載のシステムとしては、Ciara Technologies 社の NEXXUS 4820PT も展示されています。この NEXXUS 4820PT については、「パーソナルクラスタ」として、後で説明します。



S3000PTベースのサーバは複数のベンダーが1Uや4U(ブレード)サイズの製品を展示  
S3000PTを利用したサーバに関しては、インテルのホームページでも一覧表が公開されています。



インテルのブースでは展示されていませんでしたが、Atoka サーバボードも注目を集めています。大きさは、ほとんど、S3000PTと同じで、デュアルコア、クアッドコアのインテル Xeon プロセッサを2つ搭載可能となっています。また、このサーバボードは、ボード上に、InfiniBandを搭載し、カードなどを利用することなく、直接スイッチを介しての接続が可能となります。HPC サーバプラットフォームでは、InfiniBandは、インターコネクトはもちろん、ストレージなどの接続にも利用されており、GbE と同じようにオンボード上に実装されることで、よりスケーラブルなクラスタシステムの構築が可能となります。



- Atoka(左側)と S3000PT(右側)のサイズの比較
- Atoka は非常に高密度で実装しており、従来の1Uサイズのサーバの筐体内に、16コアの実装も可能
- Atokaでは、ボード上に InfiniBandを実装している(写真では、実際にIBケーブルを接続)



## パーソナルクラスタ

以上、'Intel Boosts Super-Computing Efforts'としたプレスリリースの内容とインテルのブースの内容を説明してきました。この他に、今回のインテルのブースでは、「パーソナルクラスタ」の展示がなされています。「パーソナルクラスタ」は、HPC システムの新しい潮流であり、今回の SC06 のインテルのブースでも大きく紹介されています。

「パーソナルクラスタ」は、従来のサーバをラックに搭載したクラスタではなく、デスクサイドに設置可能な筐体に複数のプロセッサを搭載し、より容易にクラスタを利用することを目的として、各社から販売されている製品です。インテルのブースでは、この「パーソナルクラスタ」として、Tyan の Typhoon PSC と Ciara Technologies の NEXXUS 4820 が展示されています。

Tyan の「パーソナルクラスタ」システム(写真右側)であり、非常にコンパクトな筐体に 40 コアを搭載することを可能としています。このような「パーソナルクラスタ」システムは、大規模な解析を行うことに関して、より高い柔軟性とその適用の可能性を大幅に拡張します。パーソナルコンピュータが計算機の利用に多くの進化と革新を可能としたように、スーパーコンピュータの新しい可能性を示すものです。デモでは、環境予測のシミュレーションの結果が示されています。このような高度はシミュレーションのこのようなコンパクトな筐体の計算機で十分に可



能であることが示されています。

NEXXUS では、インテルの新しい HPC システムの構成ブロックとなる S3000PT(シングルソケット)を搭載することで、コストの低減とまた非常にコンパクトな形での新しいクラスタ環境の構築が可能なが示されていました。また、NEXXUS では、先にご紹介した Atoka ベースのブレードの搭載も予定しています。その場合、デスクサイド(に設置可能なサイズ)に 80 コアのプロセッサを搭載可能となります。このレポートの最初にご紹介した 1996 年に構築されたシステムとほぼ同等のシステムをデスクサイドで利用できることとなります。

パーソナルクラスタは、スーパーコンピュータをデスクサイドに設置するというよりも一般に利用されているパーソナルコンピュータ(PC)の機能を大幅に拡張するシステムとして、考えることが必要です。その意味では、Windows CCS による PC 環境とのインテグレーションやより広範囲な利用方法(エクセルや MATLAB など)が可能となります。



また、マイクロソフト社のブースに展示された NEXXUS では、Windows CCS のデモ も行われており、パーソナルクラスタの利用方法の様々なデモとなっています。

## まとめとして

二極分化する HPC システムの利用用途やマーケットに対して、従来には無かったコンセプトや新技術の展開も可能となっています。マイクロプロセッサのマルチコア化が急速に進み、従来、HPC 分野に対して製品を提供してこなかったマイクロソフトの Windows Compute Cluster Server 2003 の発表など、HPC システムを取り巻く環境は大きく変化しています。そのような環境に応じた HPC システムは、従来のシステムの延長として、システムを考えるだけでなく、新しい発想と技術によって、大きく変わる可能性もあります。そのような可能性が今回の SC06 では示され、HPC プラットフォームのパラダイムシフトの可能性も示されています。