

次世代スーパーコンピュータの国際競争力向上を目指した 光インターコネクション技術への動き

○竹内寛爾（文科省・科学技術政策研）

1. はじめに

科学技術の発展を目的としたスーパーコンピュータの開発は、日米とも技術的な優位性を確保すべく激しい競争が繰り広げられている。スーパーコンピュータ開発の真の目的は、単なるスピード競争ではなく、他の科学技術分野への貢献、さらには産業界への波及効果にある。

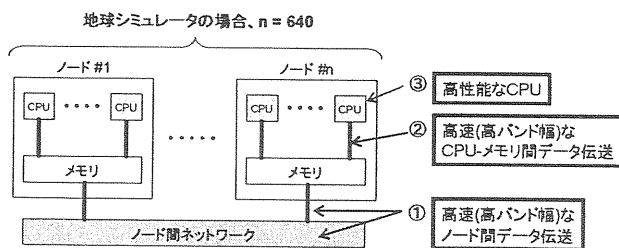
我が国は、2010 年度に 10P(ペタ)FLOPS 超級の演算能力を持つ、次世代スーパーコンピュータ「京速計算機システム」の実現を目指している。実現への課題は多いが、その一つが電気配線技術の物理限界である。その問題を解決する最有力候補として、光インターコネクションに期待が高まっている。

本稿では、スーパーコンピュータ、グリッドコンピューティング、クラスターサーバー、超大容量ルーター等を包含した超高速計算機の中で、とりわけ次世代スーパーコンピュータの取り組みを中心に、光インターコネクション技術動向について述べ、日本のとるべき方向性について考察する。

2. 電気配線方式の限界

2.1 物量における限界

スーパーコンピュータは主として CPU(中央演算処理装置)とメモリ(記憶装置)からなるノードを単位に、並列処理を行うため複数のノードで構成される(図表 1)。

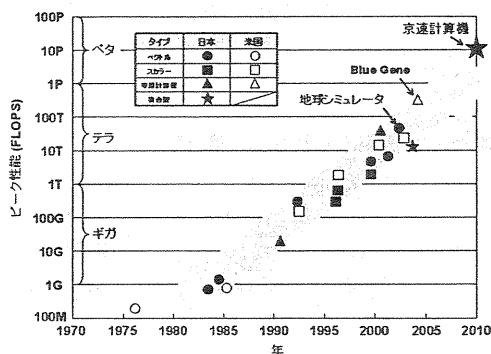


図表 1 スーパーコンピュータの基本構成例 (模式図)

(参考文献1を基に科学技術動向研究センターにて作成)

並列処理にともなうノード間のデータ通信は、一般的にスイッチにて通信相手の制御が実現され、このスイッチを含めた構成をノード間ネットワークと呼ぶ。ハードの観点からは、CPU(図表 1 の③)、CPU-メモリ間(図表 1 の②)、ノード間(図表 1 の①)のデータ転送を、それぞれ高速化することがシステム全体の高速化につながる。

近年の動きに注目すると、2002 年に我が国の地球シミュレータが LINPACK ベンチマークで 35TFLOPS を記録したが、2004 年には米国 IBM 社の BlueGene が 137TFLOPS、さらに 2005 年 11 月には同じく BlueGene が 280TFLOPS に記録を更新し、現在トップの座を維持している。我が国では、2010 年頃に「京速計算機システム」の実現によって再び世界最速を目指す計画となっている。このように、スーパーコンピュータ開発は日米の熾烈な開発競争の歴史である(図表 2)。



図表 2 日米スーパーコンピュータ開発競争の歴史

(参考資料2を基に科学技術動向研究センターにて作成)

これまでの開発競争は、ムーアの法則に従って処理能力が飛躍的に向上した CPU の存在と、メモリ容量の増大および並列化によって支えられてきた。しかし、10PFLOPS 超を目指す次世代機では、従来のアプローチだけでは到達し得ない速度領域に突入する見込みだ。

(株)富士通研究所の試算によれば、ピーク性能 3PFLOPS のスーパーコンピュータを同社の現行製品技術にて実現しようとする、ノード間の電気ケーブル(同軸ケーブル)が 20,000km、空調を含めた消費電力は 30MW、設置面積が 8,500m² にも達するという。これは、ノード間の電気ケーブルの距離だけで地球半周に相当し、電力消費量は小さな発電所が必要な規模となる。実際に目指す「京速計算機システム」の性能は 10PFLOPS 超であり、本試算よりもさらに大きな障壁となるのは間違いない。このように、次世代スーパーコンピュータを従来の電気配線方式のみによって実現することは、物量面の観点から極めて困難であると言わざるを得ない。

2.2 データ伝送における限界

高速化のポイントは CPU の性能、CPU・メモリ間の伝送帯域、ノード間の伝送帯域の向上となる。CPU 単体の高速化は言うまでもないが、「京速計算機システム」クラスの演算能力を実現するにはシステム上、CPU へのデータ供給のための伝送を如何に高速に行うかが性能を決める大きな要因となる。

ところが、従来の電気配線方式ではその特性上、伝送速度を上昇させようすると伝送距離を短くせざるを得ないという制約がある。伝送距離の制約はボード設計、システム設計にも大きな制約を与える。電気配線方式のままトータルの伝送容量を向上させるには、並列チャンネル数を増加するか、多値化によってチャンネルあたりの伝送容量を増加させる、といった手法が考えられる。しかし、これらの手法は LSI の回路規模を著しく増大させ、消費電力増大を招く結果となる。この他にも基板設計の負荷が著しく増加するなど、電気配線方式は、チャンネルあたり 10Gbps(ビット/秒)を超えることが難しいと考えられている¹⁾。

地球シミュレータの場合でさえ、1 ノード内の CPU・メモリ間の総信号本数は約 20,000 本であったが、地球シミュレータよりも 2 桁以上高速になる次世代機を考えると、CPU の大幅な並列化は避けられない。したがって、その実現は電気配線方式のみでは規模的にも実装的にも困難を極めると考えられる。

3. 光インターコネクションとは

3.1 光インターコネクションの定義

このような電気配線方式の問題を解決する最有力候補が、光通信技術を用いた光インターコネクションである。ここでは超高速計算機への適用に注目し、長距離接続とは区別してチップ内から装置間までの数十 m 以下の距離の接続を「光インターコネクション」と定義し、議論の対象とする。

3.2 光インターコネクションの特徴

光インターコネクションは電気配線方式と比較して、次のような特徴を有する。

- (1) 伝送距離と関係なく高速伝送が可能
- (2) 高密度な多重化および自由度の高い実装が可能
- (3) 上記(1)、(2)から物量の削減が可能

このように、光インターコネクションは伝送帯域を大幅に増加させるとともに物量を削減することができる。

3.3 これまでの取り組み

これまで日本は、光通信分野の黎明期から半導体レーザーの室温連続発振、光ファイバの実用化等、現在に至るまで光通信分野で世界的に大きな役割を果たしてきた。過去の国家プロジェクトは、そのほとんどが長距離・大容量伝送への光通信システムを目指した開発であった。光インターコネクション関連では、リアルコンピューティングワールドを目指した次世代情報基盤技術のごく一部で、要素技術開発がいくつか取り組まれた程度であった。

ところが近年、次世代スーパーコンピュータや超大容量ルーターのような装置の内部通信に必要なテクノロジーとして期待が高まったことを契機に、メトロ系、アクセス系からなる光通信に加えて光インターコネクション技術にも研究開発の重きが置かれ始めようとしている。

4. 光インターコネクション技術の動向

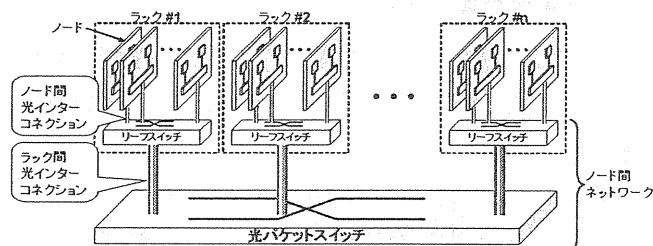
4.1 将来スーパーコンピュータ実現のための要素技術開発

我が国は 2010 年度に 10PFLOPS 超級の演算能力を持つ「京速計算機システム」の実現を目指したプロジェクト

を計画している。2005年5月、「将来のスーパーコンピューティングのための要素技術の研究開発」を開発領域として、そのうち光インターコネクション関連で次の2件が選定された³。期間は2007年までの3年間である。いずれのテーマも「京速計算機システム」の実現を念頭に置いた要素技術開発であり、企業と大学が共同で開発を行う体制となっている。

(1) ノード間ネットワークへの適用

九州大学と富士通(株)はノード同士の伝送ボトルネックを解消するため、ノード間の光インターコネクションを計画している。ノードとリーフスイッチ間を光化することにより、高速通信を可能とする(図表3)。このノードとリーフスイッチ間の光化によって、20,000km以上にもなると見積もられていた電気ケーブルの物量が1/10以下になるという。これら二つを合わせて、図表1の①で示したノード間およびノード間ネットワークの高速化を図る。



図表3 ノード間ネットワーク光インターコネクション構成例

(参考資料4を基に科学技術動向研究センターにて作成)

さらに、ラック間の光化も開発対象となっている。本方式のポイントは、現行の電気クロスバースイッチを光で置換した光パケットスイッチの導入にある。光パケットスイッチを導入すれば光のまま信号をスイッチングできるため、波長多重技術の適用によるケーブル本数の削減はもとより、一括スイッチによるスイッチ数削減、光・電気または電気・光変換モジュールが不要となり、物量面の大幅削減、消費電力の削減が期待できる。

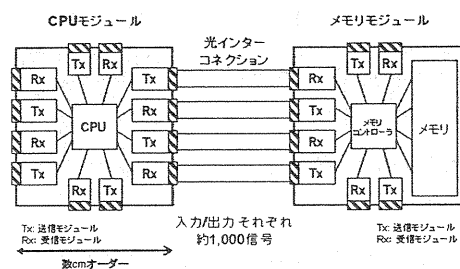
(2) CPU-メモリ間への適用

CPU性能の向上速度に比較して、CPU-メモリ間のデー

タ伝送性能の向上速度は遅く、将来的に計算機全体の実効性能がCPU-メモリ間で制限される懸念がある。

日本電気(株)と東京工業大学は、実現すれば世界初となるCPU-メモリ間の光インターコネクション技術に取り組んでいる。これは図表1に示した②の高速化を図るための技術である。開発目標はCPUあたり20Tbps以上の信号伝送能力を持つ光伝送技術の開発である。

2010年頃にはCPUあたりの処理速度が100GFLOPS以上に達する。これに対応するには、信号線あたりの伝送速度を向上させ、信号線自体の本数を削減することが望ましい。このような考えに基づき、地球シミュレータと比較して、CPU-メモリ間の伝送速度は40倍以上となる20Gbps超とし、さらにCPUあたりの信号数は半分の1,000信号の光伝送とすることで目標の20Tbpsを実現する計画となっている。



図表4 CPU-メモリ間光インターコネクションの構成図

(参考資料1を基に科学技術動向研究センターにて作成)

4.2 超大容量ルーターへの適用

インターネット上のデータは、今後映像データ等の普及により、スイッチ容量は年率約2倍の勢いで増大していく見込みとなっている⁵。ルーターはそのボトルネックとなるといわれており、電気配線方式の限界が間近に迫っている分野のひとつである。

(株)日立製作所は、スイッチファブリック当りの伝送容量が2Tbpsを超える超大容量ルーターでは、電気配線方式のままではLSIの消費電力増大、ピン数増大によるコスト高になるだろうとの見通しを示し⁵、ルーターの装置内あるいは装置間の光インターコネクションを開発中である。

光ルーターは、経済産業省の「フォトニックネットワーク技術の開発」や、総務省で推進しているフォトニックネットワー

クのプロジェクトでも研究が進められており、ノード間ネットワーク光技術開発の成果、あるいは光ルーター開発の成果が相互に活用されることが期待される。

4.3 米国の研究開発

米国国防総省国防高等研究計画局(DARPA)より“Chip to Chip Optical Interconnects (C2OI)”プロジェクトが2003年より4年計画で進行している。これは文字通りチップ間のインターコネクションを目指したものである。C2OIの構成メンバーは、企業、大学、国立研究所から構成されており、産学官の連携のプロジェクトとして見てとれる。この中でも特に注目されているのがIBM社とAgilent Technologies社が共同研究している“Terabus”である。同研究だけに、4年間で3,000万ドルが出資されている。2010年までに低消費電力、低コスト、小型化を最終目標としている。

このような高速光インターコネクション技術開発は、米政府の高性能コンピューティング(High-End Computing; HEC)計画にあるハードウェアロードマップにも合致しており、近い将来、HEC実現に向けた技術となる可能性が高い。前述のBlueGeneを手がけたIBM社がTerabusの研究開発を行っていることから、今後HEC計画に深く関与していくことが予想される。

5. 今後の方向性

5.1 戦略的ロードマップの必要性

これまで我が国は、単発的にスーパーコンピュータの開発を行ってきた。単発のプロジェクトでは、要素技術開発にも継続性が望めないロードマップに基づいた長期的戦略は極めて重要である。また、これをスーパーコンピュータのロードマップだけに留めるのではなく、今後高速化が切迫してくると思われる超大容量ルーターのロードマップとも合致させ、配線技術全般に光技術が取り入れられていく大きな流れを示す戦略的なロードマップとして展望すべきであろう。

もちろん、産業界において各企業レベルで光インターコネクション技術が開発されていく必然性はあるが、直近のビ

ジネスと関係が薄ければ、利益追求が企業の理念である以上、おのずとその規模や開発スピードには限界がある。現在の日本の状況を冷静に判断すれば、国家的な取り組み無くして当該分野を優先特化しようとする日本の企業はほとんど無いだろうと思われる。

戦略的ロードマップのような継続的な計画があれば、我が国は光インターコネクション技術のような戦略的要素技術を一過性のものではなく、次世代の「京速計算機システム」、次々世代機、さらに超大容量ルーターを含む超高速計算機実現のコア技術の一つとして明確に位置づけ、長期的に推進していくことが可能となる。

5.2 将来への布石

過去のスーパーコンピュータの性能が、既に民生品レベルになっていることからわかるように、民生品のコンピュータにおいても電気配線方式が問題になるのは、もはや時間の問題である。近い将来、光インターコネクション技術とCPUは切り離せない状況が来る。現状では、民生品のコンピュータのCPUは米国勢のほぼ独占市場となっているが、光インターコネクション技術がチップ間、チップ内へと浸透していけば、従来のアーキテクチャが変わる時期が到来する。これは、市場勢力が変化する契機となるかもしれない。光インターコネクションの技術力向上は、次世代CPU、それを取り巻く周辺メモリ等の覇権争いの布石となりうる。チップ内あるいはチップ間のインターコネクションは、高い技術力が必要であり、長期的に優位性を保つことができる要素技術とも言える。光インターコネクションの技術開発は、まだ始まったばかりである。したがって現在は、戦略的ロードマップを作成し、光インターコネクション技術へ投資すべき好機であろう。

¹ 野口孝行, “次世代スーパーコンピュータのための光配線技術” 計算科学技術シンポジウム, 2005年9月

² 理化学研究所 情報基盤センターのプレゼンテーション資料:
http://accr.riken.jp/HPC/HimenoPresen/050405_Himeno.pdf

³ 文部科学省のプレスリリース:

http://www.mext.go.jp/b_menu/houdou/17/05/05052401.htm

⁴ 木村康則, “ペタフロップスコンピュータと光技術” 次世代光情報通信技術シンポジウム, 2005年9月

⁵ 西村信治, “光技術の超大容量ルーターへの適用” 次世代光情報通信技術シンポジウム, 2005年9月