

3次元構造計算機が多層間光結合

長谷川 誠

静岡大学 工学部 知能情報工学科

hasegawa@scibu2.shizuoka.ac.jp

コンピュータの高速化/並列化と共に光相互結合網は大きな役割を果たすことになる。従来の光結合3次元構造アーキテクチャでは隣接処理機能層間だけの直接相互結合を実現していたが、これを拡張して遠く離れた処理機能層間の直接光結合をも可能とした方式について示す。本方式においては、さまざまな相互結合パターンを簡単に実現でき、問題をマッピングする自由度が非常に高い。また、3次元空間の中への効率よい埋め込みが可能である。その実現方法は、対象とする場所によっては意外と簡単な方法が存在する。

キーワード: 光相互結合網, 3次元構造アーキテクチャ, VLSIアーキテクチャ, マルチチップモジュール, プリント配線基板。

Warping the Multi-layer Barrier by the Optical Interconnection

Makoto HASEGAWA

Comp. Sci. Dept. Fac. of Eng. SHIZUOKA Univ.

hasegawa@seibu2.shizuoka.ac.jp

We propose an Optical Warp architecture for interconnecting the far separated functional layers in 3-dimensional structured computer system. Many useful interconnecting topologies and algorithms are easily mapped on this architecture, e.g. trees, torus, hyper-trees.

Keywords: optical interconnection, 3-dimensional structured architecture, VLSI architecture, multi-chip module, printed wiring board

1. まえがき

単体プロセッサのクロック周波数向上に限界が見えはじめている中で、将来の高速コンピュータには多かれ少なかれ何らかの形で並列プロセッサを取り入れざるを得ない。すでに、数値計算分野に関しては幾つかの商用超並列マシンが登場し、同程度の性能ならば従来のベクタマシンよりも一桁小さなコストで実現され、また性能の上限に関してはより大きなスケールビリティを示している。さらに、次の10年間に予想される動画像や静止画像の扱いの一般化と普通の人々への普及は、そのような超並列プロセッサに確かな活躍の場を提供することになるだろう。並列処理の全面的な実用化にはまだ暫くかかるとしても、少なくとも部分的な実用化と一般化ははさほど遠くない将来に期待できるに違いない。

そのような並列処理システムにおいて、性能を決定づける本質的な重要性を担っているのが機能ユニット間のインターコネクション・ネットワークである[OKUG91]。現時点で、商用化されている超並列コンピュータはいくつか存在するがどれもこの部分の構成に大きな努力が投入されている[KITS90][TMC93]。

現在の超並列マシンは数十台から一千台程度の要素プロセッサから構成されているが、配線結合によるインターコネクションネットワークの実現としてはほぼ限界に近い困難性を抱えている。これをさらに数万台から数百万台のプロセッサを一体化したものへと進化させようと考えたとき、インターコネクションネットワークにおける困難性はより厳しいものとなり新しい結合方式を必要とすることになるだろう。それに対処するために、超並列コンピュータのインターコネクションネットワークに光結合を導入しようという動きが急速に盛り上がりつつある[MPPOI94]。

このようなインターコネクション・システムとして光結合の採用を提案した例として3次元

構造コンピュータ[HASE86a][HASE86b][HASE88]が存在し高速フーリエ変換やソーティングの問題に対して良好な性能を期待できることが判っている。しかし、この3次元構造コンピュータにおいては隣接機能層間のみ相互結合を実現していた。これは、高速フーリエ変換やソーティングのような大量の処理データが存在しパイプライン型の処理が可能な場合には特に不都合は生じないが、一般的な問題を解こうとするとプログラムを考えるにあたってかなりの工夫を必要とすることとなる。このことから、3次元構造コンピュータは特定問題に対する専用マシンとしての性格が強い物となっていた。しかし、相互結合を隣接層間だけに限定しなければならない本質的な制約条件が存在するわけではない[HASE86c]。

隣接機能層相互間での通信を行うだけに留まらず、さらに遠く隔てられた処理層間を直接光結合するために、それぞれの機能層を構成する基板に光透過性を与えることを提案する。これによって、それまでは障壁となっていた複数の機能層を自由に通り返けて遠く離れた要素プロセッサの間を直接相互結合する。

様々な結合トポロジーとアルゴリズムを容易にマッピングできる非常にフレキシブルなインターコネクション・ネットワークが実現でき、それによって新たに構成可能となるコンピュータがどのようなものであるのかを示す。このことにより、対象とする問題に適合したフレキシブルな結合パターンの実現が容易となり、分散メモリ型の一般的な超並列マシン[JOHN91]としての利用法も当然考えられるようになるのである。

2. 概念と構造

光結合を採用することによりコンピュータの高性能化をはかることが可能であるとの認識が徐々に広まりつつある。理論的見地や実用的観

点からの多くの研究が存在し[AE88][AE94]、広域相互結合に関して光結合方式は配線結合に対する優位性を確かなものにしつつある[HASE87]。しかし、これまではそこで考えられていたのは配線結合に比べれば距離は短縮でき駆動遅延も小さいものの、あくまでも隣接機能層間の相互結合にしかすぎなかった(図1)[MPPOI94]。

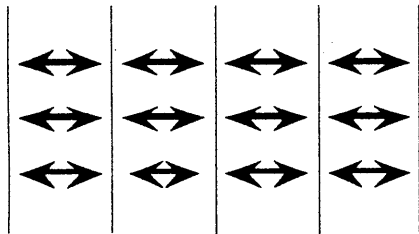


図1 隣接した機能層間だけの光結合

これに対して、図2に示すように各機能層に光を通過させるための機能を持たせることで、これまで実現されていた領域の広域通信よりもさらに拡大した範囲にまで光相互結合の効果を及ぼすことが可能となる。

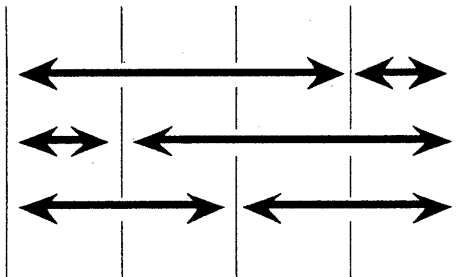


図2 機能層を突き抜けての光結合

もちろん、これを実現するために全面的に真の光透過性を各機能層へ与えることは極めて困難なことであるが、ある程度の制約を我慢するならば意外と簡単な実現方法が存在する。たと

えば、各処理機能層を構成する基板に光を通過させるための穴をうがつことを考えることができる。現在の技術水準に照らしてみると、これはさほど困難なことではない[ITOH89][SHM94]。

3. 実現方法

いかにして実現するかを考えるにあたって重要なのは、コンピュータの実装のどのレベルを対象とするかを意識することである。それは、それぞれの領域毎に考慮すべき制約事項が異なってくるからである。ここでは、まずChip間の相互結合とボード間配線とについて考える。

ボード間配線の方に関しては、プリント配線基板(Printed Wiring Board)とマルチ・チップ・モジュール(Multi-Chip Module)とを考慮の対象とする。

プリント基板のレベルでの実現方法

光透過用に空ける穴の大きさは極めて小さな物で構わない。むしろ、その大きさは多数の機能層を重ね合わせたときに機械的工作精度の限界から来る各層毎の位置ずれのために穴を通る光が遮られてしまうことがないようにと配慮して定めなければならない。プリント基板の場合には、現在の技術水準からすると直径0.1mm程度もあれば十分であろう[ITOH89]。これくらいのサイズならば、プリント基板上の配線用グリッドの中間の領域におさめることはさほど難しいことではない。

さらに、もしプリント基板の基材自体を光透過性を有するようを実現できるならば、いちいち光透過用の穴を空ける必要性は全くなってしまう。現在のプリント基板は不透明であるがこれは本質的制約ではない。強度と絶縁性さえ確保できるならば材料選択の自由度は大きい

から[ITOH89]、光透過性を有するプリント基板の実現はさほど難しい物ではないであろう。

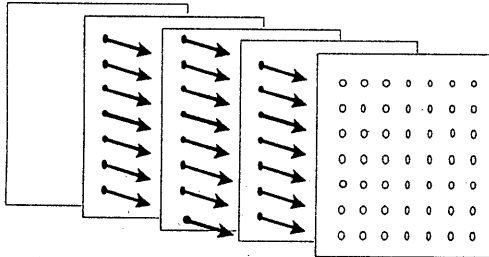


図 3 穴あきボードによる実現(プリント基板/MCM)

より望ましくは、プリント配線基板上的配線そのものやVLSIのパッケージまでをも透明化できるならば、相互結合の自由度は飛躍的に高まるのであるが、こちらの方の実現はやはりまだまだ困難な課題であろう。

マルチチップモジュールのレベルでの実現方法

次にマルチチップモジュール(multi-chip module)のレベルにおいての実現方法を考えてみることにしよう。これは高性能コンピュータの実装方法として非常に一般的に用いられている必須ともいえる手法であるので[BOGA89]、このレベルにおいても直接相互結合が可能となる意義は非常に大きい。

現在使用されているマルチチップモジュールには、MCM-C(ceramic), MCM-D(deposited), MCM-L(laminated)の3種類が存在する。MCM-Lはプリント配線基板と類似の技術であるので、前項に述べたのと同様の手法が適用できるに違いない。MCM-Dの分類の中にはシリコンを基板材料として用いる方法がある[HOND84][SHM94]。この場合、シリコンに望み通りの穴を開けることは必ずしも容易ではないが、一方でこのシリ

コンにはある特定の波長の光に関してはほとんど透明であるという望ましい性質を備えている。このことを有効に活かすことができる。MCM-Cはこれまでの超大型計算やスーパーコンピュータの実装に使われてきた高い実績の実装方法であるが、相手がセラミックスであるだけに、焼成する以前の軟らかい時点であらかじめ光透過用の穴を開けておく以外の手法はとりにくい。

チップ間結合

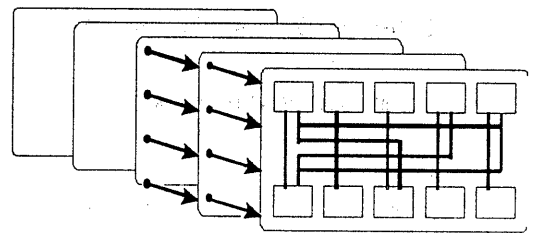


図 4 シリコンの光透過領域を利用

シリコンの光透過波長領域を使っでのチップ間相互結合を行うことができる。この透過波長領域を使うというアイデアの公表に関しては[LACY94]に先を越されてしまい大変に残念であるが、彼らは隣接機能層間の相互結合のみしか頭の中にない。わざわざ光結合を使う真の価値は、遠く離れた要素プロセッサ間を直接相互結合できる点にあるのだという本質的な点に気づいてはいない。この点で、彼らとは大きく立場が異なる。

シリコンチップ上での配線領域は数十パーセント以上にも上る。また、チップからの入出力のためのボンディングパッドやドライバの占めている面積は決して無視できる物ではない。そういった部分を置き換えて、相互結合のための専用領域として用いることが可能であろう。このチップ・レベルにおいても透明な配線が実現できると大きな自由度を手にすることが可能

となり大変に有利となる。しかし、実現は非常に困難であろうと予測される。

4. 実現可能なトポロジー

さまざまな有用な結合トポロジーがマッピング可能である。たとえば、ツリー、トーラスなどが容易に実現できる。このことから、これらを対象に作られたアルゴリズムをそのまま流用することが可能となる。しかも、ツリーは2進木ではなく多進木の形となる。

ツリー構造のマッピング例

ツリーの実現の具体例をみると、2次元の埋め込みを3次元に拡張しただけの非常に直感的な形で実現できてしまう。

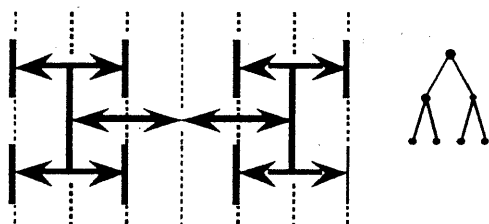


図5 ツリー構造をマッピングした例

トーラス構造のマッピング例

もちろん、トーラス構造ならば特別な工夫をしなくとも素直にマッピングするだけで実現可能なのだが、この構成で強力なのは、従来の隣接ノードよりもはるかに多数のノードを隣り合ったノードしてみなしてしまうことが可能な点だろう。このことを元にして、いくつかのアルゴリズムの高速化に関して大変に有利な立場を期待することができる。

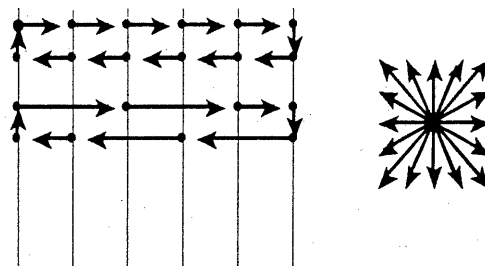


図6 トーラスをマッピングした例。隣接ノードを増やすことが可能

5. 何が改善されるのか?

この方式を採用することで多くの利点を期待することができるが、ここではツリー構造のマッピングが可能となることによって実現できる内容を中心としてみていくこととする。そのような改善が期待される事項として、ブロードキャスト、グローバルサムおよびその一族、同期操作などがまず上げられる。

Broadcast

ある一つのノードから他の全てのノードへ向けてある同一データを送りつけるブロードキャストの操作は、ツリーの頂点からリーフに向かってデータを流し込むことによって達成することができる。これに要する時間は対数オーダーで終了する。

$$t = O(\log N)$$

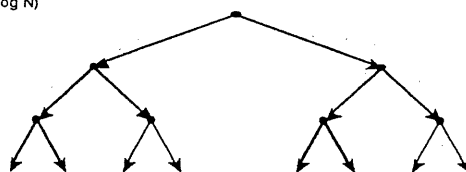


図7 ブロードキャストの例

GLOBAL SUM

グローバルサムは、全ノードのデータの総和をとる操作である。必ずしも頻繁に出現するわけではないが、十分な速度が確保できない場合にはこの操作がボトルネックとなりかねないので注意を要する。

ブロードキャストの時とはちょうど逆にリーフ・ノードから順番にデータを吸い上げながら加算を繰り返していけばよい。したがって、この演算も対数オーダの時間で実現することができる。

$$t = O(\log N)$$

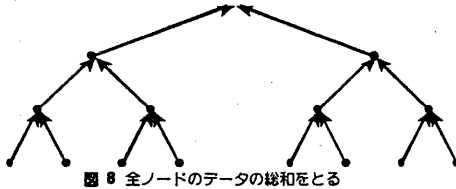


図 8 全ノードのデータの総和をとる

GLOBAL SUM が高速に実行できるようになると、その仲間たちも同様の操作によって実現することができる。そのような演算としては次のような物が挙げられる。

- Global Max/全ノードの最大値
- Global Min/全ノードの最小値

同期操作

同期操作に要する時間は超並列コンピュータの性能を大きく左右することになる。ここでは、ツリー構造がマッピングできることから、所要時間は $t = O(\log N)$ であり、しかも 2 進木のみならず多進木を簡単にマッピングできるので、そのことによる性能改善効果も大いに期待することが可能である。

3次元空間への効率よい埋め込み

このようなシステムを考える上で見落としはならないことの一つは、3次元空間の中へ効率の良い埋め込みが可能な方式であるか否かである。配線結合の場合に比べれば有利とはいえ、光結合においても結合距離の増加は性能の低下をもたらさざるを得ない。このことから、システムの全体をコンパクトに3次元空間へ詰め込むことが可能な方式が求められることになるのである。

たとえ、相互結合網それ自体としての性能には勝るとしても、この点に関して十分ではない方式には問題が残る。意外なことではあるが、もっとも理想的と考えられがちな完全結合グラフも、単純な実現方法ではこの問題を満足することができない。相互結合のみを用いられる専用の空間を中心部分に用意せざるを得ないために、全体としての外形は非常に大きな物となってしまう、このために結果として、要素プロセッサ間の物理的距離は必ずしも小さくはならないのである。

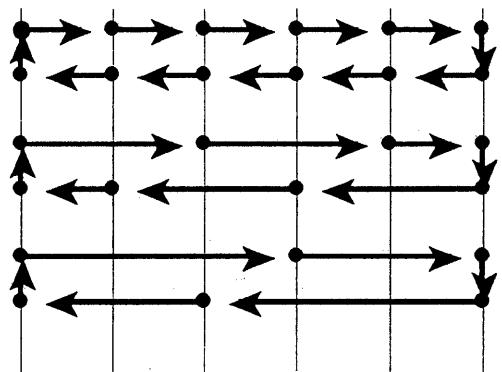


図 9 3次元構造の中にコンパクトに埋め込めることが重要である。

超多重インターリーピング

最後に、この方式の実用的価値を示すための卑近な例として、スーパーコンピュータの超多重インターリーブの実現について示すことにする。現在のスーパーコンピュータではこの部分はほとんど実装上の限界に近い状態にあるはずである。なにしろ、プロセッサからメモリへのアクセス単位が256bitあるいはそれ以上で1024wayインターリーブとかを行なうわけだから、たぶんその実装のための苦労は並大抵の物ではないであろう。接続用同軸ケーブル類の断面積だけでも大変なことになる。仮に、最新の外形1mm前後の細線同軸ケーブルを全面的に用いたとしても、ケーブルとそのコネクタをこれだけの数量用意するとなると、その専有する面積は生なかの物では済まない。このことが、コンピュータの外形を小さくして高速化に役立てようとする努力に対する障害とならざるを得ないが、これも本方式を用いれば非常にコンパクトに実現することが可能となる。

5. 結論

3次元構造コンピュータを拡張して、隣接層間だけではなく遠く離れた処理機能層間の直接光結合を可能とした方式について示した。本方式においては、問題をマッピングする自由度が非常に高く、様々な相互結合パターンを簡単に実現できる。また、3次元空間の中に効率よく埋め込むことができる。そして、実現方法は対象とする場所によっては意外と簡単である。

参考文献

- [AE88] 阿江忠:『VLSIコンピュータ』電子情報通信学会.
- [AE94] T.Ae et al.: Optical Interconnection and massively Parallel Processing, Proc. 1st International Workshop on Massively Parallel Processing Using Optical Interconnections.
- [BOGA89] E.Bogatin: High Performance Packaging of Solid-State Circuits, IEEE Journal of Solid State Circuit, Vol.24.,No.3,p.698(1989).
- [HASE86a] M.Hasegawa et al.: $AT^2=O(N \log^4 N)$, $T=O(\log N)$ Fast Fourier Transform in a Light-Connected 3-Dimensional VLSI, Proc. 13th IEEE International Symp. on Computer Architecture.
- [HASE86b] 長谷川誠,重井芳治: 光結合3次元VLSI上におけるソーティング, 電子通信学会技術研究報告 CAS86-127(1986).
- [HASE86c] M.Hasegawa: Optical Multi-Track-Bus, unpublshd (1986).
- [HASE87] 長谷川誠: VLSI上における光結合通信と配線結合通信との間での駆動遅延時間に関するトレードオフ, 電子情報通信学会論文誌 J70-D, 3, 650-651, (1987).
- [HASE88] 長谷川誠: 光結合3次元VLSIアーキテクチャ上におけるソーティング, 情報処理学会コンピュータアーキテクチャ・シンポジウム 88-3, 55-64(1988).
- [HOND84]本多: 実装革命を可能にするハイブリッドテクノロジー, M&E,工業調査会(1984).
- [ITOH89] 伊藤謹司:『プリント配線技術読本/第2版』日刊工業新聞社.
- [JOHN91] S.L.Johnson: Performance modeling of distributed memory architectures, J.parallel and Distributed Computing(1991).

[KITS90] W.D.Hillis/喜連川優著訳: 『コネクションマシン』 パーソナルメディア(1990).

[LACY94] W.S.Lacy et al.: A Fine Grain, High-Throughput Architecture Using Through-Wafer Optical Interconnect, Proc. 1st International Workshop on Massively Parallel Processing Using Optical Interconnections.

[MPPOI94] Proceedings of the First International Workshop on Massively Parallel Processing Using Optical Interconnections, IEEE Computer Society Press (1994).

[OKUG91] 奥川俊史: 『並列計算機アーキテクチャ』 コロナ社(1991).

[SHM94] ハイブリッドマイクロエレクトロニクス協会編: 『エレクトロニクス実装技術基礎講座 第4巻 実装組立技術』 工業調査会。

[TMC93] Thinking Machines Corp.: Connection Machine CM-5 Technical Summary (1993).