

マルチステージネットワーク:PBSF (Piled Banyan Switching Fabrics)

天野 英晴† 藤川義文†¹

†慶應義塾大学 理工学部

マルチプロセッサ用の結合網のクラスの一つ SSS(Simple Serial Synchronized) 型ネットワークアーキテクチャ(SSS-MIN)に適した新しいマルチステージネットワーク PBSF(Pile Banyan Switching Fabrics)を提案する。PBSFはbanyan網(omega網)を立体的に重ね、対応するエレメント間に結合線を設けた構造を持つ。ある層のbanyan網で衝突によって進路を塞がれたパケットは下の層に送られ、再びその層でルーティングを続ける。この網は同じくbanyan網の縦列接続からなり、SSS-MIN用のネットワークとして検討されていたTBSF(Tandem Banyan Switching Fabrics)よりも通過率、通過時間の両方で優れた性能を持つ。PBSFを用いたSSS-MINは、約5分の1のチップ数で絶対性能においても従来型のMINを上回る。

The Piled Banyan Switching Fabrics: A novel multistage network for parallel machines

Hideharu Amano† Yoshifumi Fujikawa†

† Faculty of Science and Technology, Keio University

A high throughput multistage network called the Piled Banyan Switching Fabrics (PBSF) is proposed for a class of network methodology (Simple Serial Synchronized network architecture) used in multiprocessors. PBSF consists of banyan (omega) networks which are piled up and connected each other. A packet which loses its way by the conflict is transferred vertically to the banyan network in the next lower layer and routed horizontally again. Both the throughput and latency of the PBSF is better than those of the Tandem Banyan Switching Fabrics (TBSF) which consists of banyan networks connected in tandem. The performance of the SSS multistage network based on the PBSF is superior to that of the conventional MIN for multiprocessors with almost 1/5 of the chip number or cost.

¹現在、日立製作所に勤務

1 はじめに

多段結合網 (Multistage Interconnection Network:MIN)[1] は中規模 (数百、数千プロセッサ) の並列計算機におけるプロセッサとメモリ間の結合機構として、長年研究が続けられてきた。従来のMINでは、メモリに対するアクセスは、パケットの形でスイッチングエレメント (多くは 2×2) 間をビットパラレル (8~64bit) に転送される。MINのトポロジはOmega網などのブロッキング網で、エレメントスイッチ内で衝突が起きた場合、片方のパケットはエレメントスイッチ内のバッファに格納される。このようなMINは交信負荷の偏りに対する性能低下が大きく、特に、プログラム中に同一メモリ番地へのアクセスが集中すると (Hot Spot)、ネットワーク全体が飽和し性能が大幅に低下する現象 (Tree Saturation) が指摘されている [2]。また、動作の複雑さと、ピンネックにより、高密度実装に困難が伴う。

そこで我々は、単純なエレメントから構成される、高い通過率を持つ結合網により、逐次的に同期して入力されたパケットを交換する方式、SSS(Simple Serial Synchronized)-MINを提案した [3][4]。そして、この方式に適合したいくつかのネットワークポロジ [5][10] を検討し、チップの試作も行なうとともに [6]、この構成のMINに基づくプロトタイプの実装を行なっている。しかし、今までに検討したネットワークポロジを用いた場合、SSS-MINは性能価格比では従来型のMINを大きく上回るものの、ユニフォームトラフィックにおける絶対性能では従来型のMINにやや劣っていた [4]。

本報告では、今までSSS-MIN用のネットワークポロジとして検討したBDOC[5]、TBSF[10]よりも高い性能を得ることのできるネットワークポロジ Piled Banyan Switching Fabrics (PBSF) を提案する。PBSFを用いることにより、絶対性能についても従来型MINに匹敵もしくはこれを上回るSSS-MINを構成することが可能になる。

2 SSS-MINの動作

2.1 基本動作

SSS-MINにおいて全パケットはプロセッサに接続された入力バッファから、フレーム信号に同期して1bit単位でMINへ入力される。交換網は 2×2 のスイッチングエレメントを多数用いて通過率の高い網を構成する。各エレ

メントの構造は非常に単純であり、基本的にはパケットの1bit分の記憶のみを行なう。このため、網全体は交換機能を持ったシフトレジスタとして働き、パケットは交換網の段数分の遅延の後、出口から1bitずつ出力される。

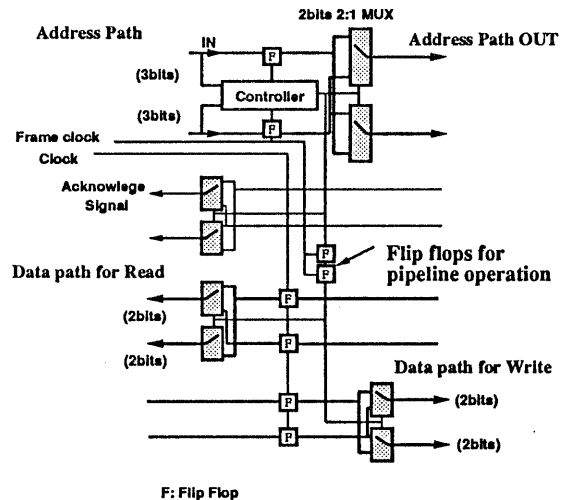


図 1: Structure of a switching element

図1に各スイッチングエレメントの構成を示す。エレメントはアドレス、応答信号 (ACK/NAK)、データ (双方向) の3つのブロックから構成されており、各部には専用の信号線とマルチプレクサが用意されている。

SSS-MINにおいて、スイッチの状態はアドレスパケット内のルーチングタグにより決定される。2つの入力パケットが同一の出口に向かった場合、衝突が発生し、希望の方向に進めなかったパケットの conflict bit がセットされる。以降、この bit がセットされたパケットは他のパケットの進行を妨害しない。アドレスパケットの先頭がMINの出口 (メモリ側) に達した時、すべてのエレメントの状態が決定し、入出力間に論理的なパスが設定される。このパスをトレースと呼ぶ。応答信号、入出力データはすべてこのトレースに沿って送られる。このため、図

1のようにSSS-MINのエLEMENTはアドレスパケット転送路に対してのみコントローラが用意され、他の部のマルチプレクサは全てこの信号により設定される。このことにより、複雑なハードウェアを要するコントローラが1つで済み、ハードウェア量が大幅に単純化されている。

応答信号はアドレスパケットの先頭がMINの出口(メモリ側)に達した時に、即座にプロセッサに戻される。この信号はELEMENT内で記憶されずに、マルチプレクサと信号線の遅延だけで伝搬する。一方、データパケットの転送は次のフレームのアドレスパケットの転送とオーバーラップされる。この動作をパイプライン化サーキットスイッチングと呼ぶ。

2.2 高速化手法

2.2.1 トップビットルーティング

大多数のMINではスイッチの状態を決めるヘッダには行き先の番号が直接用いられ、各ステージでは自分の段数に相当するbitをチェックして状態を決める。SSS-MINでは各ELEMENTでの遅延を減らすため、各段で状態決定するのに用いたbitをけずっていき、常に先頭のbitのみで状態決定していく。現在、本大学で実装され、動作が確認されている試作チップ[6]では性能の向上のため、アドレス転送線路は3bit化されている。このうち1bitは行き先の番号、残りの2bitはconflict bit, data-valid bitに用いられ、1クロックでELEMENTの状態を決定する。

2.2.2 ステージの飛び越し

アドレスパケットのヘッダにより一度トレースが設定されてしまえば、アドレスパケットの残りも、データパケットはすでに設定された経路をたどっていけば良い。このため、応答確認信号(ACK/NAK)同様、各ステージで記憶を行う必要はなく、動作速度に応じていくつかのステージを飛び越して転送することができる。試作チップでは4ステージ(つまりチップ全体)を飛び越して転送を行う。このことにより、網通過のための遅延時間を大幅に低減できる。

2.2.3 Message Combining 機能

Message Combining機能は同一宛先のパケットをELEMENT内で結合し、1つのパケットとして扱うことにより、Hot stopに対するアクセス集中に伴うTree saturationを回避する方法である[7]。この機能の効果はシミュレーシ

ン等で確認されているが、従来型のMINでこの機能を実現するためには6倍から32倍のハードウェアを要する[2]ため、今までに実現された例はない。

しかし、SSS-MINではシリアル転送の特性を生かして簡単に実現できる。図2(a)に読みだし同士の結合の実現例を示す。各ELEMENTはアドレスパケット通過時に入力された2つのパケットの全アドレスを比較する。そして、両方のアドレスが完全に等しければ、データ転送路のマルチプレクサを図2(b)に示すようにBackward Broadcastingモード(B)にセットし、全体としてTree状のトレースを形成する。読み出されたデータはこのトレースにしたがってマルチキャストされ、同時に各プロセッサに到着する。この方法では、比較とマルチキャストがビットシリアルに行われるため、付加回路が簡単で、高速性を害することもない。

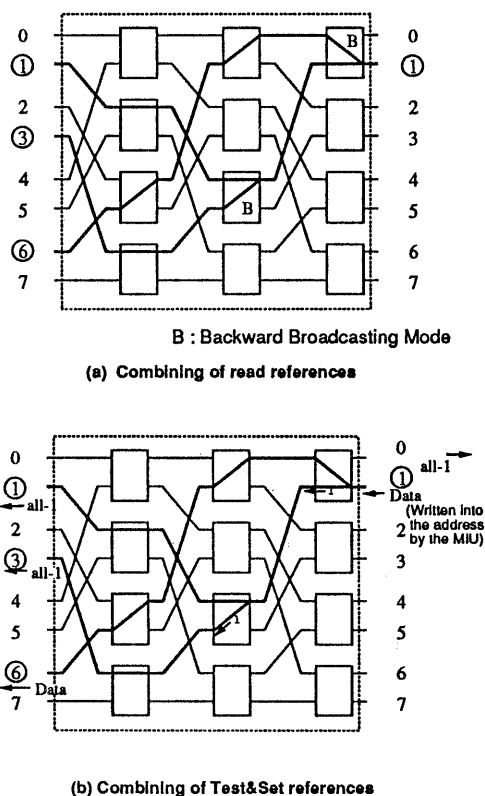


図 2: Bit serial message combining

この方法では Test&Set の結合も実現されている。図 2 同様にアドレス通過時に各エレメントは Tree 状のトレースが形成されるが、データの転送時片方のデータ転送路からは強制的に H レベル (all-1) がプロセッサ側に戻される。このため、ただひとつのプロセッサだけが、メモリからのデータを受けとり、残りのプロセッサは all-1 のデータを受けとる。読み出し直後、メモリコントローラによってメモリの値は all-1 とされるので、全体として Test&Set 操作の結合が実現される。試作チップではこの 2 つの形式の Message Combine 機能が実装されている。

3 ネットワークトポロジ

3.1 TBSF (Tandem Banyan Switching Fabrics)

SSS-MIN では、パケット同士の衝突が起きた場合、入力バッファから再送する必要があるが、性能低下の大きな原因となる。このため、SSS-MIN で用いるネットワークトポロジは従来型の MIN でよく用いられる Omega 網等のブロッキング網よりも高い通過率が要求される。一般に通過率の高い網はスイッチングエレメントを数多く必要とするが、SSS-MIN のスイッチングエレメントの構造が簡単であるためある程度の数は許容できる。我々は、当初 Batcher-banyan 網を基にした BDOC (Batcher Double Omega network with Combining) について検討した。この網はノンブロッキングであり、Message Combining 可能なパケットはすべて Combine される等理論的には興味深い性質を持つが、転送遅延時間が大きすぎ、実用的ではなかった [5]。次に我々は、B-ISDN 用に提案された TBSF (Tandem Banyan Switching Fabrics) [10] を検討した。TBSF は、図 3 に示すように Banyan 網 (Omega 網) を次々と接続し、各網の出口にバイパス路を設けた構造を持つ。宛先に到着したパケットはバイパス路によりそのままメモリモジュールに送られ、衝突により到着しなかったパケットのみが次の段の Banyan 網に入力される。

この網は厳密にはノンブロッキングではないが、同一宛先のパケットを複数通過させることができるため、ノンブロック網より高い通過率を持ち、交信の偏りに対しても強靱である。TBSF は B-ISDN 用の交換機に用いる場合、各段の Banyan 網から時間がずれて出力されたパケットを描える部分が問題になるが、SSS-MIN として用いた場合、到着順にメモリのアクセスを行えばよいので、逆にメモリ利用率を上げることができ利点に転じる。

TBSF は接続する Banyan 網数を増やすと通過率は向上するが、通過時間が増やした網の分長くなる。このため、システムによって最適の段数が存在する。この最適段数は、プロセッサ数、メモリアクセスの種類と頻度、メモリアクセス時間により影響を受ける。16 から 256 プロセッサのシステムと 50-100nsec 程度のメモリアクセスタイムを仮定すると、最適段数は 2-4 段程度になる [4]。

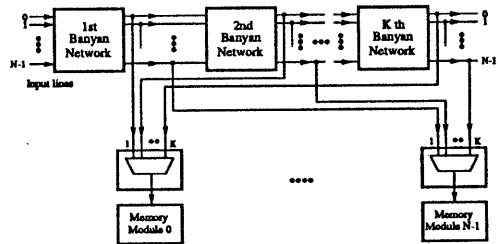


図 3: Tandem Banyan Switching Fabrics

3.2 TBSF の問題点

SSS-MIN では、スイッチングエレメントの構造の簡便化と LSI のピンネック問題の解決により、高い実装密度と、高速なクロックの利用が可能になる。しかし、ビットシリアル転送は転送容量、遅延時間の双方で従来型に比べ、基本的なハンディがある。今、従来型の MIN に比べ 4 倍高速なクロックが使用可能であるとする。それでも 32bit パラレルの転送を行う従来型に比べ転送容量で 8 倍のハンディがある。SSS-MIN はビットシリアル転送が基本だがピンネックとのバランスを考えつつ、任意の bit 幅を選ぶことができる。ここで、2bit パラレルの転送を行うと、ハンディは 4 倍になる。

この 4 倍のハンディは、ビットシリアル Message Combining、パイプライン化サーキットスイッチング、トップビットルーティング、ステージ飛び越し等の技術によりある程度補うことが可能である。これらの技術は SSS-MIN が本来持つ、構造、動作の簡単さを損なわずに高速性を実現する点に特徴がある。

しかし、2-4 段の TBSF を用いた SSS-MIN は従来型と比べハードウェア要求量は少ないが (1/10-1/16 程度)、様々な条件下におけるシミュレーションの結果、ユニフォームトラフィックにおける性能が劣ることがわかった [4]。これは以下の原因によると考えられる。

- TBSF を用いた SSS-MIN では衝突により正しくルーティングされなかったパケットは次の段の MIN に入力されるが、その段で衝突が起きる前に行なわれたルーティングは全て無駄になる。これに対し、従来型の MIN では衝突したパケットはエレメント内のパケットバッファに格納されるため、そのエレメントまでのルーティングが有効に働く。
- TBSF では接続段数が多いと、網の通過時間が長くなる。この通過時間は同一メモリモジュールに対するアクセスが複数あれば、メモリモジュールのアクセス時間に当てられるが、交信の集中が少なければ、大多数のメモリモジュールでは複数アクセスは起こらず、無駄になる。

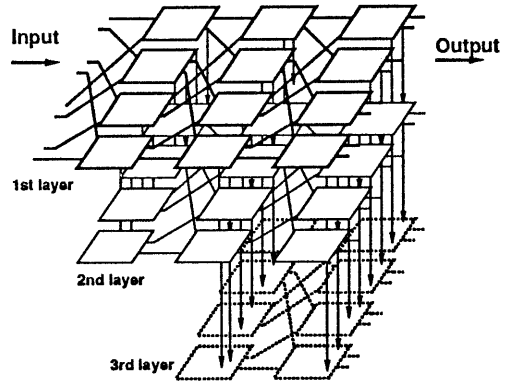


図 4: Piled Banyan Switching Fabrics

TBSF を用いた SSS-MIN では、エレメント内にバッファを持たない代わりに、衝突した結果希望の出力に到達できなかったパケットが次の段に送られる。TBSF の Banyan 網一段は全体として 1 入力につき 1 つのパケットを格納するバッファ(シフトレジスタ)と考えることができる。衝突に際しては、パケットは時間をずらして、独立なバッファ(次の段)に格納されることになる。衝突時までのルーティングが無駄になる点も、接続段数が多くなると通過時間が長くなる点も、時間的にずらして独立なバッファに入れるという TBSF の結合に起因する。PBSF(Piled Banyan Switching Fabrics) はこのような TBSF の問題点を解決するために考案された。

3.3 PBSF(Piled Banyan Switching Fabrics)

PBSF は 4 に示すように Banyan(Omega) 網を各層に用いる三次元構造を持つ。最上層と最下層を除く層のスイッチングエレメントは水平方向の入出力を 2 つずつ、垂直方向を 2 つずつ、計 4 入力 4 出力を持つ。パケットは基本的に水平方向に進み、衝突が起こって進めなくなると下の層のネットワークに送られる。

パケットはまず最上層のネットワークに入力される。最上層において水平方向に進む 2 つのパケットがあるエレメントの出力リンクで競合すると、片方のパケットは希望の方向に送られ、敗れたパケットはひとつ下の層のエレメントに、1 クロック(厳密には半クロック)の遅延を伴って送られる。上層から送られたパケットは下の階層を水平方向に進んでくるパケットと競合し、ここで敗れたパケットはさらに下の階層に 1 クロックの遅延の後送

られる(実際は垂直方向が優先権を持つ)。最下層のネットワークで競合に敗れるか途中の層で垂直方向に送ることのできなかった場合は衝突を示す conflict bit がセットされ、TBSF 網同様、以後他のパケットを妨害しない。

網の通過時間を経過した後、PBSF の各層から層が下がるにつれて 1 クロックずつ遅れて、パケットが出力される。最上層以外の網の出力からは到着できなかったパケットが出力される可能性があり、ここでは TBSF 網同様、conflict bit のチェックが行なわれ、トレースを用いた応答機構により入力バッファに転送の失敗が通知され、パケットの再送が行なわれる。TBSF 同様メモリシステムには層数分のパケットが到着する可能性がある。TBSF とは異なり、これらのパケットはほぼ同時に到着するので、メモリシステム内にバッファを設け、これらを蓄えて順にアクセスする。アクセス回数が多くフレーム時間内にすべてのデータを返せない場合、次のフレームで到着するパケットが失敗と見なされ、再送される。

PBSF の各スイッチングエレメントを図 5 に示す。この構造では垂直方向の線が、各エレメントの出力線に対して用意されているのが特徴的である。上層での競合に敗れたパケットは希望の出力に対する垂直方向の線を通して半クロック分の遅延で下の層に送られ、その層での水平方向の出力線を確保する。ここで、さらに半クロック(上層に比べると 1 クロック)遅れて水平方向にパケットが到着し、コントローラ部で進路が判定されるが、目的の出力線がすでに上層からのパケットで占拠されていた場合、垂直方向の線を通して下層に送られる。すなわちこの方式は垂直方向優先である。水平方向からのパケ

ト2つ、垂直方向(上層)からのパケットがすべて同一出力線を目指した場合、上層からのパケットが目的の出力線に進み、水平方向からのパケットのうちのひとつが選ばれ、下層に送られる。残った1つのパケットは conflict bit がセットされ、希望しない方の出力に送られる。

垂直方向の線をスイッチングエレメントの入力部に受けエレメントを 4×4 のクロスバにするとハードウェア量が TBSF 用の 2×2 のエレメントに対し約4倍になってしまう。図5の方法は3つのパケットが集中した場合に1つのパケットがダメージを受けるが、ハードウェア要求量は1.5-1.8倍程度で済む。さらに、1つのエレメントについて垂直方向(上層)の入力からのパケットが2つ同時に送られることがないので、パケットの集中による消失や正常にルーティングされているパケットへの妨害が生じない。

PBSF は TBSF に比べ次のような利点を持つ。

- N 入力のネットワークにおいて、PBSF の層数 TBSF の段数とともに K とすると、TBSF の通過時間である $K \times \log_2 N$ に比べ PBSF の通過時間は $K-1 + \log_2 N$ となり、 $1/K$ に近くなる。
- TBSF では、衝突したパケットは次の網に送られ、衝突するまでのルーティングは全て無駄になるが、PBSF では衝突するまでのルーティングは下の層で有効に働く。

すなわち、TBSF ではフレーム時間分ずらして次の網に送ったパケットを PBSF では1クロック分だけずらして、下の層の網に送ることになる。下層の網は、従来型の MIN がエレメント内に持つパケットバッファと同様の効果がある。このことにより、通過率、通過時間両方の改善が期待できる。

一方、PBSF は TBSF に比べ、エレメントのハードウェア量が増加する。また、3次元方向の層はすべてが同一チップ上に実装される必要がある。しかし、図5の実装方法を用いればハードウェア量の増加はさほど大きくない。TBSF の試作チップ [6] では16入力16出力の Banyan 網が約9000ゲートで実装されており、3層程度の PBSF は現在の技術で十分実装可能である。

4 性能評価

SSS-MIN(250MHz 動作、データアドレス分離各2bit 平行、4ステージ飛び越し、入力バッファ数3)と従

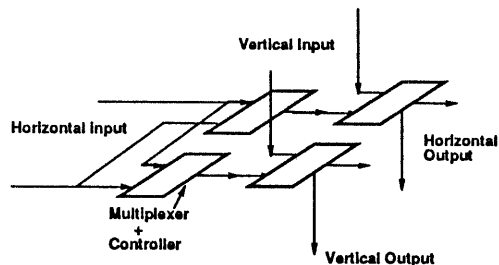


図5: Switching element of PBSF

来型 MIN(50MHz 動作、40bit 平行(8bit はヘッダ)、各エレメントに4パケット分のバッファ)の性能を確率モデルに基づくシミュレーションにより評価し、比較した。TBSF に関しては、Stanford 大と共同で開発したチップ [8] と SSS-MIN 専用の試作チップ [6] に基づいてパラメータを設定し、従来型は RP3[11]、Cedar[12] に用いられた MIN を参考に設定した。Message Combine 機能は SSS-MIN のみが持っているとする。これは SSS-MIN 用 TBSF 試作チップでは20%のハードウェア量の増加で Message Combine 機能を実現できたのにもかかわらず、従来型の MIN では実現例がないためである。

プロセッサは50MIPSの性能を持ち、20nsec 単位で一定の確率で共有メモリにアクセス(アドレス32bit、データ64bit)を発生する。多くの MIN に基づくマルチプロセッサ同様、各プロセッサはローカルメモリと MIN に対するライトスルーキャッシュを持ち、共有データのみ共有メモリに置くとする。キャッシュのコンシステンシーはコンパイラにより埋め込まれたコードにより維持される。プロセッサの MIN に対するアクセスはノンブロックアクセス(同期命令等、確率 P_n で発生)とブロックアクセス(書き込み、プリフェッチ等、確率 P_b で発生)に分けられる。プロセッサ-MIN 間のバッファがフルの場合は、ノンブロックアクセスでもプロセッサは待たされる。メモリは DRAM を想定し、サイクル時間は30-90nsec とした。MIN の性能はプロセッサが待たされないで有効に働く時間の割合(プロセッサ稼働率)で評価する。

図6a,b に256プロセッサ、メモリサイクル時間70nsec のシステムにおける PBSF, TBSF の banyan 網数とプロセッサ稼働率の関係を示す。アクセスは各メモリモジュールに均等に分布しているとする。TBSF は banyan 網数が多いと通過率は増大するが、その分だけ通過時間が大

さくなるため、最適値が存在する。ここで設定した条件ではメモリのサイクル時間である 70nsec が Banyana 網の通過時間である 40nsec ($8 \times 5nsec$) より大きいので、網数を多くした効果が少なく、最適値は 2 となる。PBSF は網数が増えた場合の通過時間の増加が 1 クロックであるので、網数が 2 以上になると稼働率の変動は少ない。この条件では最適値 2 または 3 であるが、両者共にさほど稼働率は変わらない。最適値同士を比較すると PBSF の稼働率は TBSF より 6-7% 程度高くなる。

図 6a,b 上には従来型の MIN を用いた場合の稼働率を直線で示す。SSS-MIN の最適値と比較すると、PBSF は従来型よりやや稼働率が高く、TBSF ではやや低くなる。図 6b はブロッキングアクセスの負荷がやや重い場合であるが、図 6a とほぼ同様の傾向が見られる。

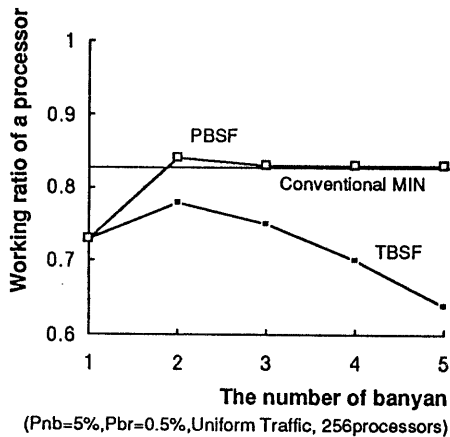


図 6a: 網数対プロセッサ稼働率 (Pbr=0.5%)

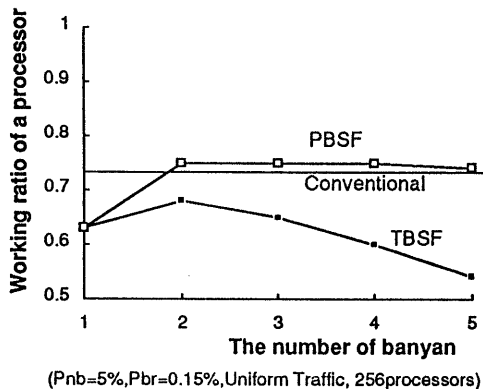


図 6b: 網数対プロセッサ稼働率 (Pbr=1.5%)

図 7 にブロッキングアクセスの一部が Hot Spot に集中した場合を示す。PBSF、TBSF 共に 2 段の結果を用いている。PBSF、TBSF 共に Combine 機能の効果により従来型に比べ稼働率の低下はわずかである。以上の評価結果より、現実的なパラメータの範囲では、PBSF は 2 層でも十分効果があることがわかる。PBSF は層数が多いと複数のチップを用いて大規模なネットワークを構成する場合、ピンネックの問題が起きるが、2 層ならばさほど問題は起きず、チップの集積度が大きければ、むしろ TBSF よりもチップ数も少なくて済む。

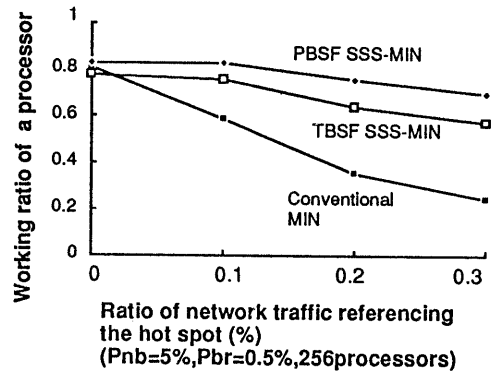


図 7: HOT SPOT を含む場合の稼働率

5 おわりに

PBSF を用いることにより、SSS-MIN は、現実的なパラメータの範囲でユニフォームトラフィック下の絶対性能についても、従来型 MIN に比べ優位に立つことができる。PBSF は TBSF に比べ全体としてのハードウェア量はやや大きくなるが、2 層で十分な性能を得ることができるため、チップ数はむしろ減る場合が多い。

PBSF のアイデアは大変単純であり、トポロジだけに注目すると以下のネットワークと類似点がある。

- Banyan 網を並列に並べた Parallel Banyan Switching Fabrics [14] [15]
- エレメントのサイズを倍にして冗長線を与えた冗長ネットワーク [13]

このうち、Parallel Banyan Switching Fabrics は、トラフィックを単純に分割して、互いに全く独立の Banyan 網

に入れる構成で、TBSF よりも効率の悪い網であることが知られている [9]。ここで提案する PBSF では重なっている網間のパケット転送が可能であるため、はるかに有利である。

冗長ネットワークの中には、効果的に用いれば 2 層の PBSF と同様の効率を達成できる網が存在する。しかし、スイッチの構造とルーティング法が複雑で、高速動作が困難であり、ハードウェア量も大きい。逆に PBSF は、その冗長性を利用して耐故障性を上げることができる可能性があり、今後の研究課題である。

PBSF は、SSS-MIN 形式において網内にバッファを持たせた方法としても考えることができる。すなわち、従来型の MIN では既にエレメント内にバッファを持つため、PBSF を用いても効果は少ない。また、B-ISDN 用の高速パケット交換機としては、網の通過時間よりも通過率が問題となる。PBSF は通過率に関しては TBSF と大差ない。このように PBSF は SSS-MIN にとっては大変優れたネットワークトポロジであるが、他の用途ではさほど効果を発揮しない。これが、このような単純なアイデアが今まで注目されなかった理由であろう。

参考文献

- [1] G.Broomell, and J.R.Heath, "Classification Categories and Historical Development of Circuit Switching Topologies," ACM Computing Surveys, Vol.15, No.2, Jun. 1983
- [2] G.F.Phister, and V.A.Norton, "Hot Spot Contention and Combining in Multistage Interconnection Networks," IEEE Trans. on Comput. vol. c-34, No.10, Oct. 1985.
- [3] 天野,Gaye,"SSS (Simple Serial Synchronized) スイッチングアーキテクチャに基づく並列計算機," 信学報 CPSY91-11(SWoPP91), Jul 1991.
- [4] H.Amano, L.Zhou, K.Gaye,"SSS (Simple Serial Synchronized)-MIN: A novel multi stage interconnection architecture for multiprocessors," To appear in Proc. of IFIP Congress 92, Sept. 1992.
- [5] K.Gaye, H.Amano, "A Batcher Double Omega network with Combining," IEICE Trans. on information and systems, Vol.E75-D No.3 May, 1992.
- [6] 周、天野, "SSS-MIN 用プロトタイプチップの実装," 情報処理学会第 44 回全国大会
- [7] A.Gottlieb, R.Grishman, C.P.Kruskal, K.P.Mcauliffe, L.Rudolf, and M.Snir, "The NYU Ultracomputer - Designing an MIMD Shared Memory Parallel Computer", IEEE Trans.on Comput. vol. c-32, No.2, Feb.1983.
- [8] F.Chiussi, H.Amano, F.A.Tobagi, "A 0.8 μ m BiCMOS Sea-Of-Gates implementation of the Tandem Banyan Fast Packet Switch," Proc. of IEEE Custom Integrated Circuits Conference May, 1991.
- [9] F.A.Tobagi, T.Kwok, "The Tandem Banyan Switching Fabric: a Simple High-Performance Fast Packet Switch," Stanford University Technical Report No.CSL-TR-90-451, Sept. 1990.
- [10] F.A.Tobagi, T.Kwok, "The Tandem Banyan Switching Fabric: a Simple High-Performance Fast Packet Switch," Proc. INFOCOM91, Apr. 1991.
- [11] G.F.Phister, et al, "The IBM Research Parallel Processor Prototype (RP3): Introduction and architecture," in Proc. IEEE 1985 Int. Conf. Parallel Processing, Aug. 1983.
- [12] J.Konicek, et al, "The Organization of the Cedar System, in Proc. 1991 Int. Conf. Parallel Processing, Aug. 1991.
- [13] D.P.Agrawal, G.B.Adams III, H.J.Siegel, "A Survey and Comparison of Fault-Tolerant Multistage Interconnection Networks," Computer, June. 1987, pp.14-27.
- [14] C.P.Kruskal, M.Snir, The performance of multistage interconnection networks for multiprocessors," IEEE Trans. on Computers, Vol.C-32, No.12, pp.1091-1098, Dec. 1983.
- [15] M.Kumar, J.R.Jump, "Performance of unbuffered shuffle-exchange networks," IEEE Trans. on Computers, Vol.C-35, No.6, pp.573-577, Jun. 1986.