

7M-7

SSS-MIN アーキテクチャに基づくマルチプロセッサプロトタイプ SNAIL*

笹原 正司 寺田 純 小椋 里 周 洛 天野 英晴†

慶應義塾大学 理工学部†

1 はじめに

我々は現在までに大規模マルチプロセッサ構成を目指し、その基本となる高速ネットワークとして SSS-MIN を提案してきた [1]。この SSS-MIN アーキテクチャに基づいたネットワークチップは既に実装が行なわれており、アーキテクチャの評価を行なうため、マルチプロセッサテストボードの設計および実装を行っている。

2 SNAIL の設計方針

近年、共有バスとスヌープキャッシュを用いた共有メモリ型計算機は既に商品化の段階に達しているが、このような形態の場合、特に近年の高速プロセッサは、高速なメモリアクセスを要求するため、バスの混雑がネックとなり、接続可能なプロセッサ数は4個程度になってしまう。このようなことから、大規模密結合マルチプロセッサの構成には高速なプロセッサ・メモリ間結合網を欠かすことはできない。従来のネットワークはパラレルでデータ転送を行ない、さらにバケットバッファを持つものが多いため、その構造は複雑であり、高速動作性に乏しい。SSS-MIN では、高速な結合網を実現するため、同期シリアル転送とすることでスイッチの構造および制御を簡単化し、さらに同期入力であることを利用して内部をパイプライン化するなど、特に高速・高実装密度を目指して開発された。この SSS-MIN を用いたマルチプロセッサ SNAIL(SSS-MIN Network Architecture Implementation) は、将来の SSS-MIN を用いた大規模分散共有メモリ型並列計算機設計の第一歩として、以下の点を目標に実装が行なわれている。

- SSS-MIN アーキテクチャの評価環境
- スケジューラビリティ・デバッガビリティの高い計算機設計
- 将来の大規模並列計算機のプロセッサクラスタとして活用できる

SSS-MIN はきわめて規則的な動作を行なうため、プロセッサと同期して動作させることでシステム全体を規則的に動作させることができる。この動作の規則性は通常のプログラミングに際して高いデバッガビリティ

を提供するだけでなく、インストラクションレベルでのスケジューリングを可能にするため、特に細粒度並列処理の分野への活用を狙っている。

また、SSS-MIN はメッセージコンバイン機能など、マルチステージネットワークの短所を隠蔽する機能が実装されており、これらの機能がどの程度スループットの向上に作用するかを測定する環境として実装を行なう。

さらに、大規模システムのノードとするため、その基板面積ができるだけ小さくなるように実装を行なう。具体的には Xilinx 社の LCA(Logic Cell Array) を用いることで基板面積の縮小を図っている。

3 SNAIL の設計

3.1 構成

SNAIL の全体の構成を図1に示す。プロセッサモジュール・メインメモリ間をネットワークで結合する。この構成で最高16プロセッサから成るシステムを構成できる。

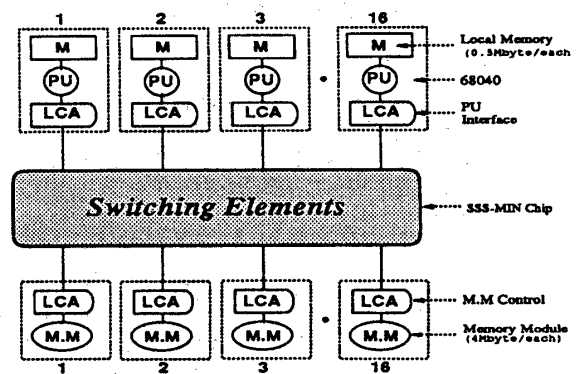


図1: SNAIL の構成

3.2 プロセッサ部

プロセッサ部の構成を図2に示す。図のようにプロセッサにはMC68040を用いており、各プロセッサは512Kbyteのローカルメモリを持っている。このプロセッサ部は、プロセッサ周辺回路、ネッ

*SNAIL - a Multiprocessor Prototype based on The SSS-Network Architecture

†Masashi SASAHARA, Jun TERADA, Satoshi OGURA, Luo Zhou, Hideharu AMANO

‡Keio University

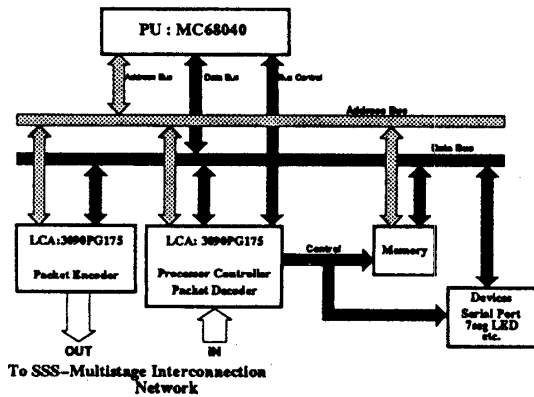


図 2: プロセッサ部の構成

トワークインターフェースが LCA2 個によって実装されている。

同期操作としては Test&Set, Fetch&Dec が用意されている。また、Data Prefetch を実装しており、インストラクションレベルでのスケジューリングを行なうことで、ネットワークによるレイテンシのない、効率の良いネットワークアクセスが可能となっている。これらの機能はメインメモリの任意の部分について行なうことができ、メモリの利用に関してフレキシビリティの高い環境を提供する。また、メインメモリはオンチップキャッシュが有効に働くよう 32bit/64bit ブロック転送が可能になっている。さらに、異なるプロセッサからの、同一アドレスへのアクセスのうち、メモリーードと Test&Set はメッセージコンバインされ、同時にデータを転送することができる。

3.3 ネットワーク

SSS-MIN アーキテクチャに適したネットワークとして、我々は現在までにオメガネットワークを基本とし、それらを直列に複数接続した TBSF (Tandem Banyan Switching Fabrics)、三次元的に構成する PBSF (Piled Banyan Switching Fabrics)[4] を提唱してきた。シミュレーションの結果では PBSF の方が比較的効率がよいことがわかっているが、ハードウェア量の増加などの点から TBSF 用のチップが実装された [2]。TBSF では、[4] により、256 入力のネットワークではオメガネットワーク 3 段の TBSF が最適であるという結果が得られている。しかし、16 入力の場合はネットワーク内での衝突ははるかに少なくなる。そこで、オメガ網を何段接続するかを決定するため、現在 ATTEMPT-0 上で稼働しているプログラムで、アクセス頻度をパラメータとした確率モデルによるシミュレーションを行なった [3]。なお、アクセス時間などのパラメータは実際のチップのデータと SNAIL 設計時のパラメータを用いている。表 1 はアプリケーション実行時間に対する共有メモリアクセスの平均待ち時間の比率を示している。

このように、いずれのアプリケーションでも 1 段と 2 段ではその差は 3% 未満の差が出る程度である。こ

表 1: プロセッサ待ち時間の割合

| プログラム | 実行時間 (μsec) | 1 段 (%) | 2 段 (%) |
|---------|--------------------------|---------|---------|
| Gauss | 2.8 | 7 | 7 |
| CG | 9.1 | 2 | 2 |
| LOGIQUE | 1.2 | 21 | 18 |
| ESIM | 2.3 | 10 | 9 |
| CHANNEL | 45.5 | 0.8 | 0.8 |

れは TBSF2 段にした場合、衝突の減少による性能向上と、2 段目をバケットが通過した場合の通過時間の増加による損失とが相殺されてしまうためである。以上の結果から、今回の実装には TBSF 1 段とした。また、今回使用する MC68040 のパフォーマンスから見て、SSS-MIN 1 個による 3bit パラレルではバンド幅が足りないため、SSS-MIN を 4 個並列に使用し、計 12 ビットパラレルの転送を行なっている。

3.4 メインメモリ

メインメモリはモジュールあたり 4Mbyte の容量を持っており、LCA 1 個によって制御されている。Test&Set, Fetch&Dec もこのコントローラ内で行なわれる。また、先に述べたキャッシュに対する転送のため、各モジュールはキャッシュブロックごとにインターリーブされている。

4 現状

現在 SNAIL は 2 プロセッサで稼働している。このボード上で動作チェックと簡単な評価を兼ねて N-Queen などの簡単なアプリケーションプログラムを動作させている。しかし、すでに 16 プロセッサシステム構成にむけ、4 プロセッサを 1 クラスタとした PCB を作成中で、現在ネットワークボードの PCB の設計にとりかかっており、まもなく 16PU で稼働を始める予定である。

謝辞

SSS-MIN チップは川崎製鉄株式会社 LSI 事業部の御協力により開発されました。深く感謝致します。

参考文献

- [1] H. Amano, L. Zhou, and K. Gaye. SSS (simple serial synchronized)-MIN: A novel multi stage interconnection architecture for multiprocessors. *Proc. of IFIP Congress 92, Sept., 1992.*
- [2] 周洛, 天野英晴. SSS-MIN 用プロトタイプチップの実装. 情報処理学会, 1992.
- [3] 鳥居, 竹本, 天野, 小椋. バス結合型並列計算機の交信用メモリの性能評価. 情報処理学会論文誌, Vol. 33, No. 3, 1992.
- [4] 天野英晴, 藤川義文. マルチステージネットワーク PBSF. 情報処理学会, 1992.