2N-3　可変構造型並列計算機のネットワーク・アーキテクチャ

蒲池恒彦　村上和彰　福田晃　末吉敏則　富田眞治
(九州大学)

1 はじめに

現在我々は，あらゆる結合形態が実現可能な「可変構造」の相互結合網で128台のPE（Processing Element）を接続するマルチプロセッサ・システム「可変構造型並列計算機」を開発している.[1]「可変構造」の相互結合網は，128×128のクロスバー網を時分割多重化したもので多重化クロスバー網（MC-net：Multiple Crossbar network）と呼んでいる.

本稿ではMC-netの実現方法について述べる.

2 MC-netの設計思想

可変構造型並列計算機の相互結合網には，以下の能力が要求される.

① PE間の結合形態に制限がない非閉塞網であること
② ルーティング制御（経路選択制御）が容易であり，回線設定のオーバヘッドが小さいこと
③ 結合形態が可変（reconfigurable）であること

①，②の要件を満たすため相互結合網としてクロスバー網を採用し，さらに後述するようにクロスバー網本来の機能に加え，時分割多重化を図ることにより「可変構造性（reconfigurability）」を持たせた.

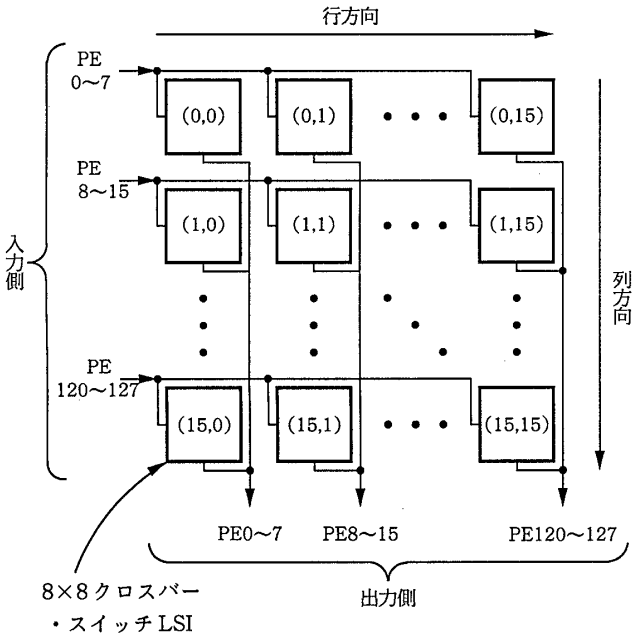


図1. MC-netの基本構成

3 MC-netの構成

MC-netの基本構成を図1に示す. 128×128のクロスバー網は，8×8のクロスバー・スイッチLSI 256個で平面分割し，行方向（16個），列方向（16個）にそれぞれバス接続して構成している. クロスバー・スイッチLSIは現在CMOSゲート・アレイにより開発中であり，その信号線の構成を図2に，諸元を表1に示す.

4 動作モード

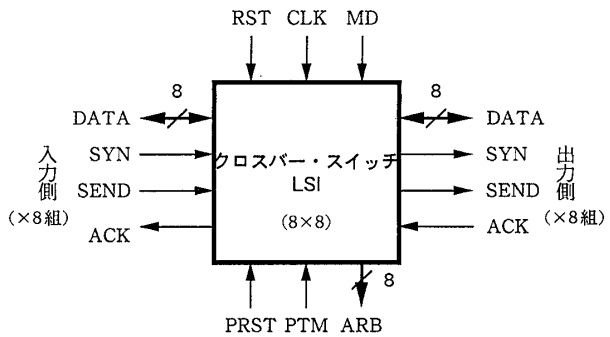
MC-netはデマンドモード，プリセットモードと呼ぶ2つのモードで動作可能である. 以下各モードの具体的な制御方法について述べる.

4.1 デマンドモード

プログラム実行時にランダムに生起するPEからの接続要求に対して回線の設定を動的に行うモードである.

(1) 入力制御

PEは相手PEの宛先アドレス（ヘッダ）を7bitで指



DATA: 8ビット双方向データ線
SEND：ハンドシェイク（送信）
ACK　：ハンドシェイク（応答）
SYN　：ハンドシェイク（同期）
ARB　：外部調停要求
CLK　：クロック
RST　：リセット
MD　　：モード設定
PRST：パターンリセット
PTM　：パターン切替

図2. クロスバー・スイッチLSIの信号構成

Network Architecture of a Reconfigurable Parallel-Processor

Tsunehiko KAMACHI, Kazuaki MURAKAMI, Akira FUKUDA, Toshinori SUEYOSI and Shinji TOMITA
Kyushu University

表1. クロスバー・スイッチLSIの諸元

交換数	8×8
交換方式	回線交換
転送制御	ハンドシェイク方式
転送方向	双方向（半二重）
転送データ幅	ポート当り1バイト
調停方式 （デマンドモード時）	優先順位選択方式の 多入力非同期調停方式
放送機能	サポート
動作周波数	20MHz
パッケージ	223pin PGAパッケージ
信号線数	189

定してMC-netに接続を要求するが，各LSIはヘッダの上位4bitで独自にチップセレクトすることができない（下位3bitでLSI内の出力ポート指定）．このため外部回路（INPC：INPut Contoroler）でヘッダを一旦変換し，行方向16個のうち選択されたLSIのみに変換したヘッダを投入する．

(2) 調停

デマンドモードでは1PEに複数PE（最大128）からの接続要求が競合する可能性がある．この競合を解消するために，MC-netではその構成上次のように2レベルの調停を行っている．

① ローカル・アービタ（LSI内入力ポート間調停）

LSI内部で，8個の入力ポートの中から1個を選択する．そしてLSI外部に存在するグロバル・アービタに対して接続要求を出し，接続許可が返ってきた時点で回線を接続する．

② グロバル・アービタ（列方向LSI間調停）

列方向16個のLSIの中から1個を選択し，接続許可信号を送る．

いずれのアービタも出力ポート対応に備えられており，優先順位エンコーダを用いているが，その前段にラッチを設け，飢餓状態が生じないように考慮してある．

4.2 プリセットモード

プログラム実行前にPE間の接続形態を16パターン設定しておき，実行時にパターンを切り替えることでクロスバー網を時分割多重化するモードである．

(1) パターン切替え制御

LSI内の入力ポート対応に備えたスイッチ制御メモリ内にどの出力ポートと接続するかという制御情報を16パターン格納しておき，パターン切替え（PTM）信号を256個の全LSIに対して同時にアサートすることで順次制御情報を読み出し，MC-net全体を同期させて高速に結合パターンを切換えていく．

(2) パターンロード制御

スイッチ制御メモリ内への制御情報のロードは，PEが通常のデータ転送と類似のシーケンスで256回（16パターン×16LSI）繰り返すことで行う．

なお，プリセットモードでは競合が生じないことが保証されているので調停を行う必要はない．

このプリセットモードにより，PEは仮想リンクを一時に最大16本持つことができる．また，制御パターンをロードしなおすだけで最大接続数（次数：degree）16を越える結合形態を表現することができる．

5 実装構成

図3に実装予定図を示す．Mother Board（MB）にクロスバー・スイッチLSIを16個（4×4）実装し，これにPEボードを8枚差し込んでシステムのサブ・ユニットを構成する．更にこのサブ・ユニットを行方向，列方向にそれぞれ4個バス接続することでシステムを構築する．また，MBあたりの入出力信号線数を減少させるため，MB内に入力制御回路（INPC），及びMB内アービタを設け各MBに入出力制御を分散させた．この結果，事実上調停は3レベルになるが，ローカル・アービタ及びMB内アービタの調停時間とグローバル・アービタの調停時間はオーバーラップすることができるので，調停のためのオーバヘッドが増加することはない．

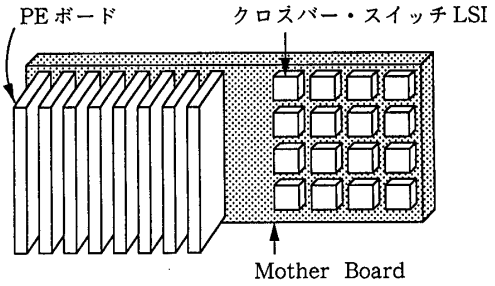


図3. 実装予定図

6 おわりに

以上，可変構造型並列計算機の相互結合網MC-netの実現方法について述べた．現在MBの実装設計中であり，昭和64年9月の完成を目指している．

参考文献

1) 村上ほか："可変構造型並列計算機のシステム・アーキテクチャ"，情報処理学会「コンピュータアーキテクチャ」シンポジウム論文集，Vol.88, No.3, pp.165－pp.174，（1988年5月）