

可変構造型並列計算機のメッセージ通信ユニット

2N-4

森眞一郎 村上和彰 福田晃 末吉敏則 富田眞治  
(九州大学)

1 はじめに

我々は、128台のPE (Processing Element) を、128×128のクロスバー網 (MC-netと呼ぶ) で接続し、PE間結合形態を可変とした可変構造型並列計算機を開発中である。本システムでのPE間通信はメッセージベースであるが、さらに密結合マルチプロセッサ構成のメモリイメージもサポート可能である。本稿では、本システムにおけるPEのメッセージ通信ユニットについて述べる。

2 PE構成

図1にPEの構成を示す。PEは、

- ①プロセッサ・ユニット (PU : Processor Unit)
- ②SMWアクセス・ユニット (SAU : Shared-memory-window Access Unit)
- ③メッセージ通信ユニット (MCU : Message Communication Unit)
- ④メモリ・ユニット (MU : Memory Unit)

の4つのユニットで構成される。PUおよびSAUはMCUを介してMC-netの入力および出力ポート(それぞれ1個)に接続される。PUおよびSAUからはMCUは通常のI/Oデバイスと等価に見える。

3 MCUの目的

MCUの目的は、以下の通りである。

- I) 論理チャンネルの導入による、物理入力ポートの仮想化 : PE台数分の論理的なI/Oチャンネル, SMWアクセスチャンネルを設けることにより、PU/SAUにMC-netとのインターフェースを意識させない。
- II) メッセージ・バッファリングおよびキューイングのオフロード : チャンネル毎にデータバッファとキューを用意することによりPU/SAUの負荷を軽減する。
- III) 放送 (broadcast) 機能のサポート : 放送用チャンネルを設けることにより、PUに同一内容のメッセージを繰り返し転送させる手間を省く。
- IV) 将来の拡張性 : 汎用マイクロプロセッサを装備することで、OSのメッセージ交換機能の一部を将来オフロードすることを可能にする。

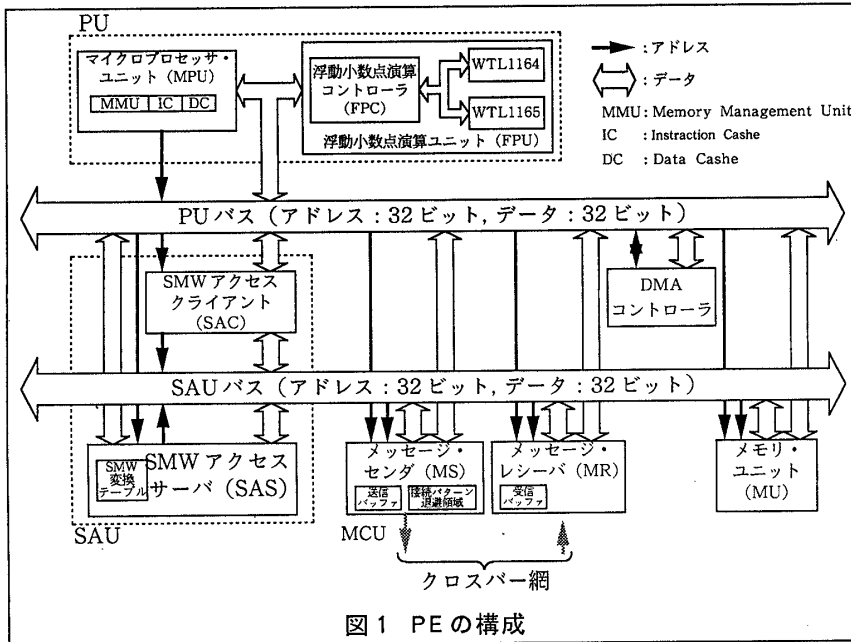


図1 PEの構成

4 MCUの機能

MCUは

- ①メッセージ・センダ (MS : Message Sender)
- ②メッセージ・レシーバ (MR : Message Receiver)

から成る。MCUが処理するメッセージの種類としては、

- ①データ送出メッセージ : 1~256バイトの任意長データの送出
- ②割り込み要求メッセージ : PE間割り込みをサポートするためのメッセージ
- ③SMWアクセス・メッセージ : 特定PEの共有メモリ・ウィンドウに対す

Message Communication Unit for a Reconfigurable Parallel-Processor

Shinichirou MORI, Kazuaki MURAKAMI, Akira FUKUDA, Toshinori SUEYOSHI, and Shinji TOMITA  
Kyusyu University

る Read/Write アクセス・メッセージの3種がある。①と②をまとめて、I/Oメッセージと呼ぶ。I/Oメッセージに関しては、特定PE向けと放送に区別して処理を行なっている。MCUは表1に示すチャンネルをI/O領域上に提供している。

4.1 MSの機能・構成

MSの機能は

- I) PU/SAUから送られてきたメッセージをMC-netを遊ばせることなく転送すること
- II) MC-netの動作をPU/SAUに意識させないこと

である。そのためにMSはチャンネルごとに大容量の入力バッファおよび、キューを持ち、それらの管理もチャンネル単位に行なっている。MSは図2に示すように次の4つのブロックから成る。

①入力処理部：PU/SAUから送られてきたメッセージにヘッダを装着し入力バッファに格納すると同時にそ

の情報をキューに登録する

②MS制御部：キューイングがその主な働きである。将来的にはOSのネットワーク機能やMS動作のモニタ機能、MC-netモード制御などの機能も持たせる。

③MC-netインターフェース部：10Mバイト/秒の転送速度で送信バッファからMC-netへメッセージ送信を行なうためのDMA回路である。

④送信バッファ部、MS制御メモリ部：入力処理部とMC-netインターフェース部は各々独立して動作を行なうため入力バッファ、MS制御メモリはデュアルポート構成となっている。

4.2 MRの機能・構成

MRの機能は

- I) MC-netから送られて来るメッセージをMC-netを待たせることなく受信バッファに格納すること
  - II) 送られてきたメッセージのヘッダを解析し、その結果に応じて必要な処理を行なうこと
- である。MRもMSと同様4つのブロックで構成される。

①MC-netインターフェース部：MSの場合とは逆方向にMC-netから受信バッファへのメッセージ受信を行なうDMA回路である。

②MR制御部：受信バッファの高速かつ効率のよい管理、メッセージヘッダの解析などを行なう。

③出力処理部：MR制御部によるヘッダの解析結果をもとに、PU/SAUへメッセージを渡す。

④受信バッファ部、MR制御メモリ部：受信バッファはMC-netからのメッセージフローを停滞させないような構成となっている。MRにおいても各ブロックが独立に動作するため、受信バッファ、MR制御メモリはデュアルポート構成となっている。

表1 チャンネル一覧

(a) メッセージ・センダ (MS) が提供する送信チャンネル

優先順位	チャンネル名	数	送信メッセージの種類	依頼元	同期/非同期
①	SMWn アクセス・クライアント用チャンネル	128	SMW アクセス	SMW アクセス・クライアント	Read : 同期 Write : 非同期
②	対PEn送信用チャンネル	128	データ送出 (特定) 割込み要求 (特定)	MPU	非同期
③	放送用チャンネル	1	データ送出 (放送) 割込み要求 (放送)	MPU	非同期

(b) メッセージ・レシーバ (MR) が提供する受信チャンネル

優先順位	チャンネル名	数	受信メッセージの種類	依頼元
①	SMWn アクセス・サーバ用チャンネル	128	SMW アクセス	SMW アクセス・サーバ
②	対PEn受信用チャンネル	128	データ送出 (放送/特定) 割込み要求 (放送/特定)	MPU

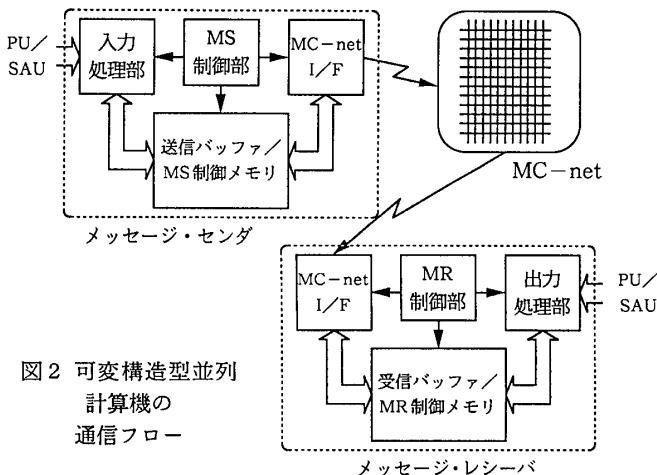


図2 可変構造型並列計算機の通信フロー

5 まとめ

メッセージ・ベースの可変構造型並列計算機において、メッセージ通信ユニットはメッセージ交換のオーバヘッドを吸収し、かつMC-netを最高転送速度で動作させることを可能にしている。

参考文献

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.165-174，(1988年5月)