

データ駆動計算機 E D D E N の通信制御機構

6W-3

三浦宏喜 田中一行 川口正樹 清水雅久 森 憲敬
三洋電機(株) 情報通信システム研究所

1. はじめに

我々は、実用的な並列処理計算機の実現に向けて、データ駆動計算機 E D D E N (Enhanced Data Driven E N g i n e)を開発中である。E D D E Nは、通信制御機構をも含んだ要素プロセッサを1チップのCMOS-LSIで実現し、これを多数台接続するものであり、ベクトル演算機構の導入による定型的な構造体処理の高速化、効率的発火制御機構の導入などの特長を有している^{[1][2]}。

本稿では、E D D E Nのプロセッサ結合網の概要、プロセッサ間通信制御方式などについて述べる。

2. プロセッサ結合網の概要

E D D E Nのプロセッサ結合網の条件として、①通信ポート数が4以下、②プロセッサ間距離が小さい、③セルフルーティングが容易、④デッドロック回避が可能、⑤一様構造である、などを考えた。このうち、②についてはプロセッサ数に対するオーダーの値よりもプロセッサ数1024台以下での実際の数値を考慮した。また、画像処理、科学技術計算などのいわゆるデータパラレル型の処理により良く適応させるために⑤を重視した。更に、プロセッサ間あるいはボード間の配線などの実装面での容易性をも重視した結果、上記の全ての条件を満たす結合網としてトラス結合網を採用し、各要素プロセッサ(PE)には、複数語構成のバケットを行き先へ向けて最短距離でセルフルーティングする機能を持たせることにした。プロセッサ間の通信リンクは全て双方向とし、チップのピン数の節約のために同一のリンクを双方向で共用することにした。

3. 通信制御方式

(1) 基本的なデータ転送方式

結合網は、基本的にカスケード接続された複数のラッチによって構成され、2相式クロックによって偶数番目のラッチと奇数番目のラッチにデータが交互にラッチされる。各段では、現在保持しているデータが次段に転送されるまで新たなラッチ動作を許可しない。即ち、1語単位で自己同期式のデータ転送を行う、エラスティック転送路となっている。これによって、固定長バケットのみならず、構造体などの可変長データの通信をも可能にしている。

なお、E D D E Nにおけるバケット構成は、プロセッサ内部では1語構成であり、プロセッサ間通信の際は、構造体データを除いては4語構成の固定長としている。

(2) 双方向通信制御方式

双方向通信の優先度制御は次の様に行う。

- ① 2つのバケットが時差をもって衝突した時は、先に到着したバケットの転送を優先する。
- ② 全く同時に衝突した時は、固定的に片方向の転送を優先する。
- ③ 時差衝突、同時衝突によって待たされたバケットは、次のサイクルでは優先して転送する。

(3) 基本ルーティング方式

図1に示したような $m \times n$ (m 行 n 列)の結合網において、 y 行 x 列にあるPEに (y, x) と番号付けする。各プロセッサにおいて、自身のPE番号が (q, p) 、到着したバケットの行き先が (y, x) である時、

$$\Delta x = (x - p) \bmod n \quad (|\Delta x| \leq n/2)$$

$$\Delta y = (y - q) \bmod m \quad (|\Delta y| \leq m/2)$$

を算出し、

- ① $\Delta x > 0$ ならばバケットをE(東)へ転送
- ② $\Delta x = 0$ かつ $\Delta y > 0$ ならばS(南)へ転送
- ③ $\Delta x = 0$ かつ $\Delta y < 0$ ならばN(北)へ転送
- ④ $\Delta x < 0$ ならばW(西)へ転送
- ⑤ $\Delta x = 0$ かつ $\Delta y = 0$ ならば自身へ入力

するようにルーティング制御を行う。

即ち、バケットは列番号が一致するまで東西方向のリング上を転送され、列番号が一致したPEにおいて南北方向のリングに合流して、行き先へ向けて最短距離で転送される。

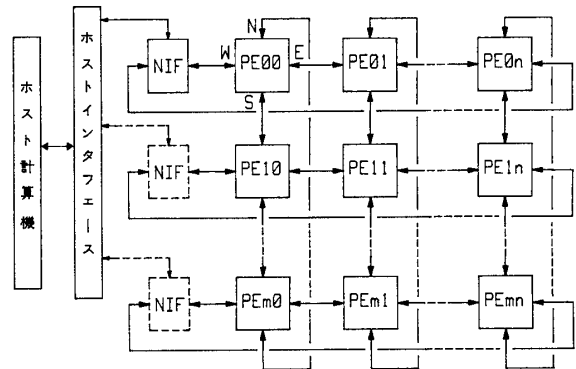


図1. E D D E Nの基本構成

The Communication mechanism for Data Driven Computer "EDDEN"

Hiroki MIURA, Kazuyuki TANAKA, Masaki KAWAGUCHI, Masahisa SHIMIZU, Noriyuki MORI
SANYO Electric Co.,Ltd.

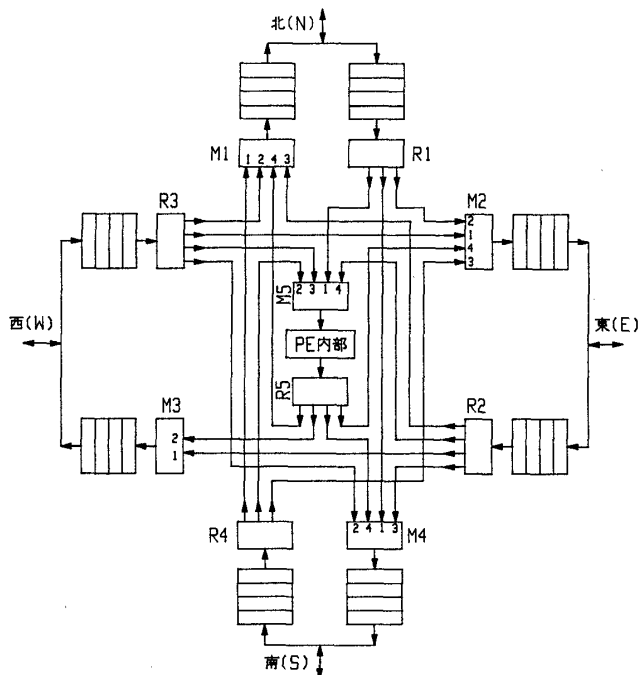


図2. 通信制御部の構成

(4) デッドロックの回避

EDDENで採用したようなトラス結合網におけるデッドロックとしては、①網上で複数のパケットがループを形成し、このループに新たなパケットを合流させようとした時に生じるデッドロック、②双方向通信リンク上を転送されているパケットが、その転送の終了を待っているパケットが原因で、転送の途中で前に進めなくなってしまった時に生じるデッドロックが考えられる。

①については、東西南北の各通信ポート間にそれぞれ2系統の通信路を設けて結合網をループフリーにすれば防ぐことができる。しかし、トラス網の場合、この方法ではハードウェア規模が大きくなりすぎる。

②については、当然ながらパケットの途中で転送が停止しないことを保証すれば防止できる。

また、前述のルーティングアルゴリズムでは、PEに南北方向から入力されたパケットが東西方向に合流して転送されることがない。従って、パケットによるループは1行または1列すべてのPEを巡るもののみである。

そこでEDDENでは、東西南北の各通信ポートに1パケット(4語)の入力バッファおよび出力バッファを設けて②を防止する。更に、東西方向または南北方向に通過するパケットを、他の方向から合流しようとするパケットやPE内部から出力されるパケットよりも優先して転送することによって①を防止することにした。

これによって、固定長パケットに関してはデッドロックを完全に回避している。また、構造体パケットの通信に起因するデッドロックについては、デッドロック防止のためのセマフォを導入しており、これによって通信を排他制御することによって回避できる。

4. 構成

図2に要素プロセッサの通信制御部の構成を示す。図においてR1～R5は分岐回路(ルータ)であり、基本的には前述のルーティングアルゴリズムに従ってパケットをルーティングする。M1～M5は優先度付きマルチプレクサであり、図の数字で示される優先度(1が最優先)でパケットを合流させる。東西南北各ポートが保持する入力バッファ、出力バッファは4段のラッチがカスケード接続された構成であり、前述のような停止制御により、1語単位でエラスティックとなっている。隣接プロセッサとの通信、及び分岐・合流の方向制御はパケット単位で行い、ヘッダの到着で方向を選択し、テイルの転送を終了するまで同一のパケットを選択し続ける。

5. 付加的機能

(1) 網の外部とのインターフェース

網の外部行きのパケットはネットワークインターフェース(NIF)の直近のPEに向けて通常のルーティングによって転送した後、そこでPE内部に入力せずに東ポートに出力する。これによってPE内部の処理に余計な負荷がかかるのを防いでいる。NIFでは外部行きパケットを分岐させて外部へ出力するとともに外部からのパケットを東西方向のリングに合流させて入力する。PEチップは初期設定によって、このNIFとしても動作できるように設計しているためNIF用のLSIを新規に開発する必要はない。

(2) フレキシビリティ

外部から、図1におけるm、nを初期設定することによって結合網サイズを可変にすることができる。但し、m、nは、2のべき乗に限っている。また、図1において南北方向の通信リンクのみを接続するとリング結合網として動作が可能であり、システムの規模に応じて柔軟な構成がとれる。更に、前述のルーティングアルゴリズムにおいてモジュロをとらないモードに設定することによって、2次元メッシュ結合網をも実現できる。

6. おわりに

以上、EDDENのプロセッサ結合網の概要、プロセッサ間通信方式、ルーティング方式、デッドロックの回避方式、構成、フレキシビリティなどについて述べた。今後は、実機の開発を急ぐとともに、クラスタ型結合網への拡張などについても検討する予定である。

末筆ながら、本研究を御指導、御支援して下さい関係各位に厚く感謝します。

参考文献

- [1] 三浦他:「高並列データ駆動計算機EDDENの概要」, 情処会38回全国大会論文集2T-2.
- [2] 田中他:「データ駆動計算機EDDENにおける発火制御とカラー管理」, 情処会39回全国大会論文集.