

# チップ内パケット転送に基づく高並列プロセッサの構成

藤岡 与周\*・苫米地 宣裕\*\*

## Design of a Highly Parallel Processor Based on Packet Transfer Scheme

Yoshichika FUJIOKA\* and Nobuhiro TOMABECHI\*\*

### Abstract

Register-transfer-level packet data transfer scheme is proposed for intra-chip data transfer to make complexity of interconnection switching mechanism greatly reduced. Especially, reduction of the control complexity in a parallel VLSI processor is discussed based on semi-autonomous packet routing.

**Key words:** network on chip, parallel VLSI processor, semi-autonomous packet routing, reduction of control complexity

### 1. ま え が き

これまでに、粗粒度パケット転送に基づくネットワークオンチップが提案されている<sup>1-3)</sup>。本稿では、マイクロネットワークでのチップ内データ転送をフレキシブルかつプログラマブルとするため、細粒度パケット転送方式を提案している<sup>4-5)</sup>。

ルータ構造をできるだけ簡単にするため、自律のパケットデータ転送とオフラインのスケジューリング・アロケーションの組合せに基づくプロトコルを採用しており、マイクロネットワーク内でのパケット衝突が起らない。ルータ内でパケット受信制御が自動的に行われるため、VLIW 制御メモリ容量を大幅に減少できる。効果的なパケットデータ転送を実現するため、パケット送信タイミング制御のための新しい制御モジュールを提案している。

### 2. パケットデータ転送に基づく並列 VLSI プロセッサアーキテクチャ

#### 2.1 パケットデータ転送に基づくマイクロネットワーク構造

図 1 に、複数の PE と、それらを接続するマイクロネットワークを備えた細粒度 MIMD (Multi-Instruction Multi-Data) 型並列プロセッサの提案構造をしめす。処理アルゴリズムが与えられ、それがコントロールデータフローグラフ (CDFG) で表現されており、スケジューリング・アロケーションが予め決定されていると仮定する。CDFG におけるそれぞれのノードは PE に割当てられ、エッジに対応する PE 間データ転送はマイクロネット

ワークを経由してなされる。

マイクロネットワークには、図 2 に示すようにパケット転送に用いる 2 本のデータ転送ラインがあり、一つは左から右への方向のパケット転送に用いられ、もう一つは右から左への方向のパケット転送に用いられる。2 つの PE 間のパケット転送は PE に直接接続されたルータを経由してなされる。

プロセッサの基本操作はレジスタ間データ転送に帰着されるため、もしパケットがマイクロネットワークに適切に送信されると、データ受信のための自動的なタイミング生成がなされる。このことはデータ受信のタイミン

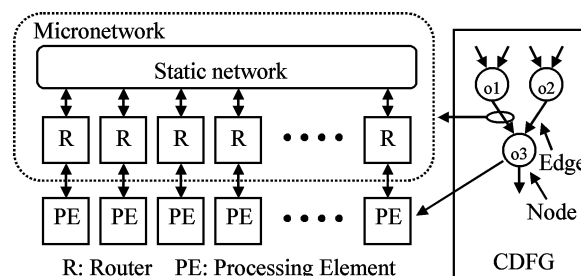


図 1 並列プロセッサアーキテクチャ

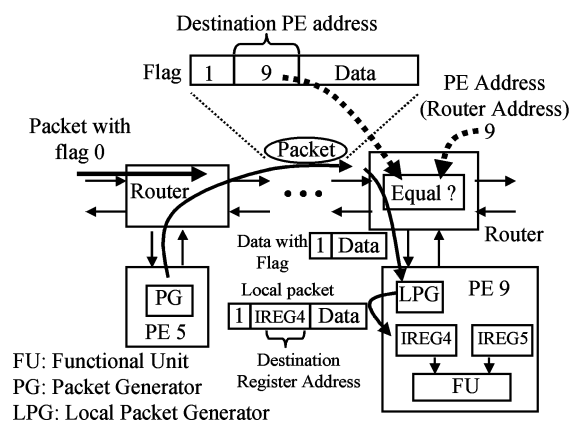


図 2 パケット転送の例

平成 19 年 1 月 5 日受理

\* システム情報工学科・助教授

\*\* システム情報工学科・教授

グ制御をマイクロプログラム中で省略できることを意味する。したがって、VLIW (Very Long Instruction Word) 制御メモリに要求されるメモリ容量を大幅に減少できる。

## 2.2 セミオートノマスパケットルーティング

細粒度並列処理における頻繁なパケット転送のためには多くのルータが要求されるため、ルータ内の制御機構はできるだけ単純でなければならない。オフラインのスケジューリング・アロケーションがセミオートノマスパケット転送に用いられる。最適問題はパケット転送が起こらないという制約下で全体の処理時間を最小化することである。もしこの制約が満たされるなら、ルータ内のバッファキューメモリが不要になる。

## 2.3 パケットフォーマット

セミオートノマスパケットルーティングに基づくことにより、1ビットのフラグ、宛先アドレスと1個のデータからなり、優先順位など他のヘッダ情報が不要である、単純なパケットフォーマットを定義できる。フラグはパケットが有効なものか否かを表し、パケット送信制御を自律的にするために効果的に採用されている。

## 2.4 階層的パケット転送

階層的に、PE内でもパケット転送が利用される。PE間およびPE内の両方の階層に関するパケット転送の例を図2に示す。

[手続き1] 宛先 PE アドレス 9 とデータを含むパケットがパケットジェネレータ (PG) で生成され、それが PE5 から送出される。PE5 からのフラグが有効であり、かつ隣のルータからのフラグが無効であるため、PE5 からのパケットが自律的に PE9 へ送られる。

[手続き2] パケットはルータ間をパイプライン方式で転送される。宛先アドレスと各ルータアドレスの比較がなされ、それらが等しくない場合パケットは隣のルータへ転送される。

[手続き3] 宛先 PE アドレスとルータアドレス 9 が等しいため、パケットは PE9 に転送され、フラグ情報が PE 内データ転送でのパケット送信のきっかけとして用いられる。

[手続き4] PE9 では、新しいローカルパケットがローカルパケットジェネレータ (LPG) で生成され、IREG4 に送られる。

## 3. プロセッシングエレメントの構成

図3はPEの構成を示しており、レジスタファイルがプログラマブルディレイユニット (PDU) 内に中間結果

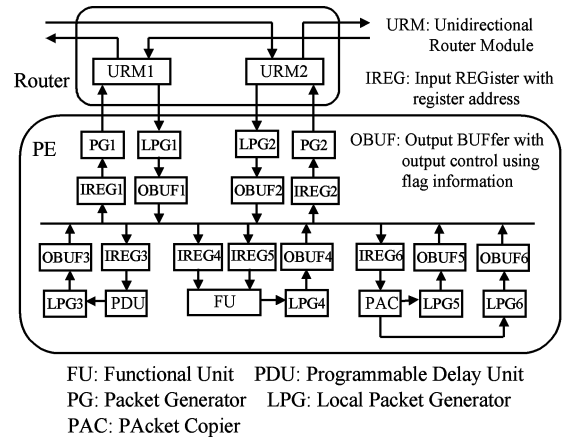


図3 パケット転送のためのルータとPEの構成

の一時記憶のために備えられている。最も重要な制御機能はPDUからパケットをセミオートノマスパケット転送法に基づく指定されたタイミングで送出することである。PDUでは、遅延制御がプログラムされた遅延情報に従ってなされる。

## 3.1 プログラマブルディレイユニット

図4は複数のレジスタモジュールを備えたPDUの構造を示している。それぞれのレジスタモジュール内には、データの保存と遅延制御それぞれのために、ひとつのレジスタとプリセットブルダウンカウンタ (PDC) が備えられている。もし有効なフラグがデータとともにPDUに到着すれば、以下に示す遅延制御がなされる。

[手続き1] レジスタモジュールのアドレスと待ちクロックステップ数がウェイトクロックステップROM (WCSR)により生成される。

[手続き2] 入力データがレジスタモジュール内のレジスタに保存され、待ちクロックステップ数がPDCにセットされる。

[手続き3] 待ちクロックステップ数のカウントダウン

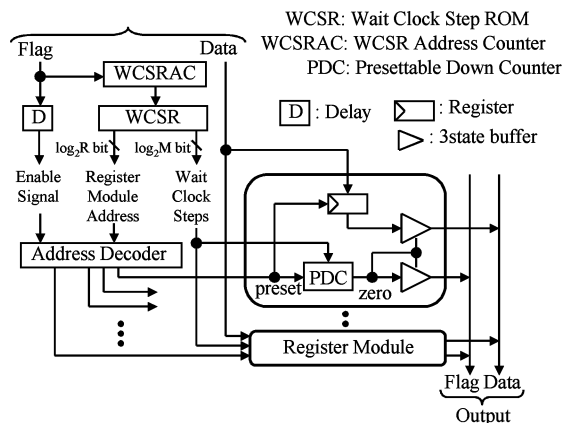


図4 プログラマブルディレイユニット

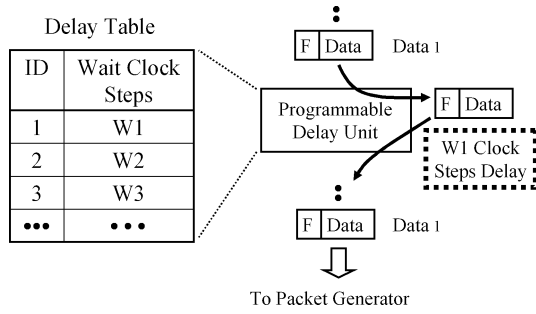


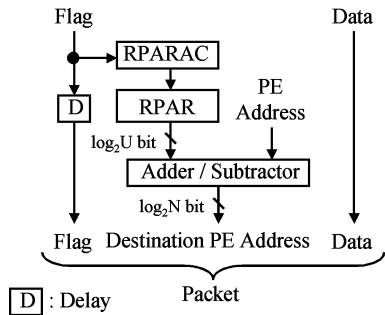
図5 PDUの動作

が開始される。もしカウントが零になると、データは PDU から LPG に有効フラグとともに送られる。

以上の手続きにより、図5に示すように自律的に遅延操作とデータ送出が行われる。

### 3.2 パケットジェネレータ

図6は有効フラグ入力後すぐにパケットを生成する PG を示している。また、図7にその動作概念を示す。制御メモリ容量を減少するために、あらかじめプログラムされた相対 PE アドレスが導入されており、それが相対 PE アドレス ROM (RPAR) に保存されている。この相対 PE アドレスはもしパケットデータ転送方向が右方向であれば PE アドレスに加算される。一方、もしそれが左方向であれば、PE アドレスから減算される。



RPAR: Relative PE Address ROM  
RPARAC: RPAR Address Counter

図6 パケットジェネレータ

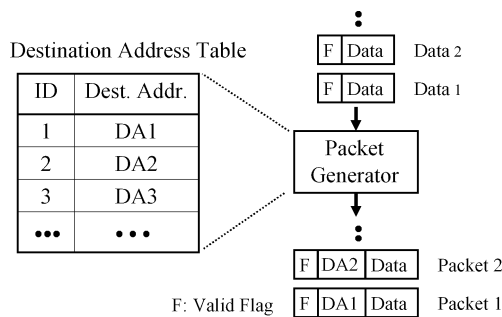


図7 PGの動作

## 4. 評価

提案する VLSI プロセッサと従来の VLIW 制御に基づく VLSI プロセッサとの制御メモリ容量を比較してみる。提案する VLSI プロセッサの制御メモリ容量  $Mp$  は次式で与えられる。

$$Mp = P \log_2 U + 3X + D(\log_2 R + \log_2 M). \quad (1)$$

最初の項は PE 間パケット転送に用いられるパケットの総数  $P$  と相対 PE アドレスのビット長  $\log_2 U$  との積であり、ここで  $U$  はアロケーションで決定される PE 間パケット転送の最大距離である。第2項は PE 内パケット転送に用いられるパケットの総数  $X$  と宛先レジスタアドレスのビット長 3 との積である。第3項は PDU に保存されるデータの総数  $D$  と、レジスタモジュールアドレスのビット長  $\log_2 R$  とスケジューリングで決定される最大待ちクロックステップ数ビット長  $\log_2 M$  の和との積である。

次に、従来の VLIW 制御に基づくプロセッサを評価してみる。図8にパイプラインバスアーキテクチャに基づく PE 構造を示す。VLIW 制御法は2種類がある。ひとつは図9に示すようにそれぞれのクロックステップ毎にオンオフ制御信号が与えられる方法である。この場合、制御メモリ容量  $Mv$  は次式で与えられる。

$$Mv = 13SN + D(\log_2 R + \log_2 M) \quad (2)$$

ここで、 $S$  はクロックステップ総数である。制御メモリ容量  $Mv$  は制御信号がそれほどしばしば変化しない場合、メモリ容量が無駄になる。

別の方法は、制御信号がオンになるようなクロックステップ間隔の情報を使うことである。制御信号がクロックステップ  $C1$  でオンになり、再びそれがオンになるもっとも近いクロックステップが  $C2$  であるとする。データ転送のための  $C1$  と  $C2$  のクロックステップ間隔がプログラムされる。制御信号がそれほどしばしば変化しない場合、この方法は制御メモリ容量の減少に有用である。PE 内の制御信号発生器 (CSG) が図10に示すようにプログラムされたクロックステップ間隔に応じたそれぞれの制御信号生成に用いられる。この場合、制御メモリ容量  $Mb$  は次式で与えられる。

$$Mb = 2P \log_2 M + 2X \log_2 M + D(\log_2 R + \log_2 M). \quad (3)$$

スケジューリングおよびアロケーションされた CDFG の制御メモリ容量の比較を検討する。処理時間に対応するクロックステップ数  $S$  は次式で与えられる。

$$S = Q(W + V) \quad (4)$$

ここで、 $W$ ,  $V$ ,  $Q$  はそれぞれ次のとおりである。

$W$ : 各ノードのクロックステップ数

$V$ : 並列 PE 間データ転送の平均クロックステップ数

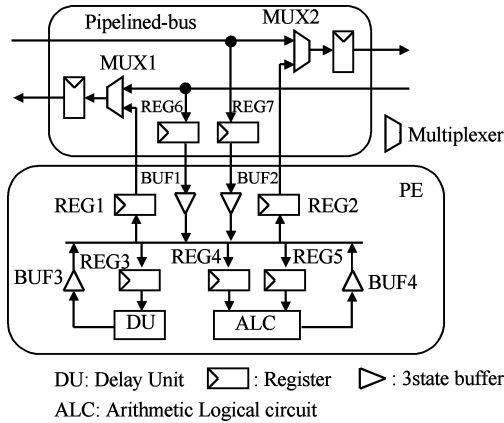


図8 従来のデータ転送用パイプラインバスアーキテクチャ

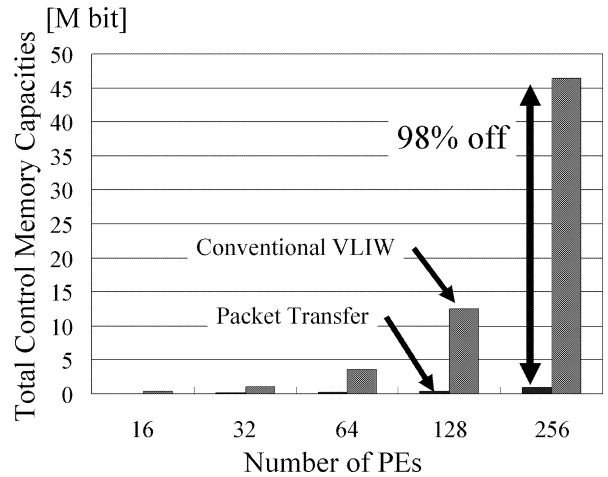


図11 パケット転送方式とクロックステップベース VLIW 方式の制御メモリ容量の比較

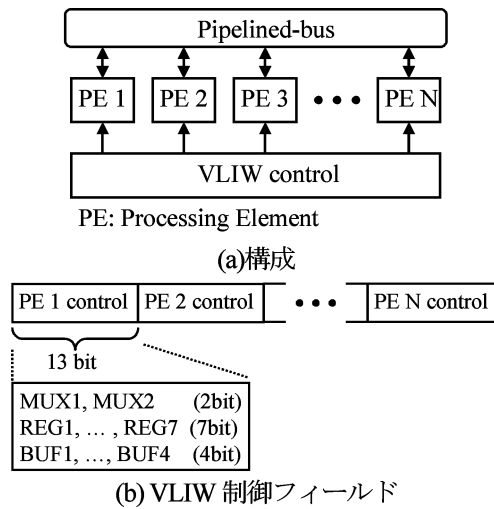


図9 クロックステップベース VLIW 方式パイプラインバスアーキテクチャ

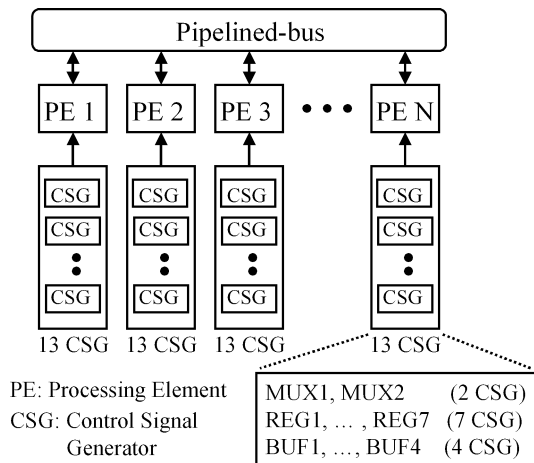


図10 制御信号ベース VLIW 方式パイプラインバスアーキテクチャ

$Q$ : 処理の繰り返し数。ここで、ひとつの処理はひとつのノードの処理と PE 間データ転送に  $W+V$  クロックステップを要する。

図 11, 表 1 と表 2 は  $P=E$ ,  $U=N/2$ ,  $X=4E$ ,  $D=E$ ,

表1 パケット転送方式とクロックステップベース VLIW 方式の制御メモリ容量

$N$	$Q$	$S$ [c.s.]	$M_p$ [bit]	$M_v$ [bit]	$M_v/M_p$
64	20	840	34560	712960	20.6
64	100	4200	204800	3596800	17.6
256	20	2760	158720	9251840	58.3
256	100	13800	896000	46361600	51.7
1024	20	10440	716800	139284480	194.3
1024	100	52200	3993600	696832000	174.5

表2 パケット転送方式と制御信号ベース VLIW 方式の制御メモリ容量

$N$	$Q$	$S$ [c.s.]	$M_p$ [bit]	$M_b$ [bit]	$M_b/M_p$
64	20	840	34560	116480	3.4
64	100	4200	204800	806400	3.9
256	20	2760	158720	578560	3.6
256	100	13800	896000	3507200	3.9
1024	20	10440	716800	2764800	3.9
1024	100	52200	3993600	16281600	4.1

$R=Q/2$ ,  $M=S/2$ ,  $E=NQ$ ,  $W=10$ ,  $V=U$  である場合の制御メモリ容量  $M_p$ ,  $M_v$ ,  $M_b$  の比較を示している。実際の処理応用では、提案アーキテクチャにより制御メモリ容量の大幅な減少が期待できる。

## 5. む す び

提案する VLSI アーキテクチャは、不規則なデータ転送がしばしば発生するような場合に、制御メモリ容量の減少に非常に有用である。従って、提案するマイクロネットワークアーキテクチャ内のより多くの PE を同一面積

のチップ内に備えることができ、並列処理能力を大幅に改善できる。

#### 参考文献

- 1) P.P. Pande, C. Grecu, A. Ivanov, R. Saleh and G.D. Micheli, "Design, synthesis and test of networks on chips," *IEEE Design and Test of Computers*, Vol. 22, No. 5, pp. 404-413, (2005).
- 2) K. Goossens, J. Dielissen and A. Radulescu, "AEthereal Network on Chip: Concepts, Architectures, and Implementations," *IEEE Design and Test of Computers*, Vol. 22, No. 5, pp. 414-421, (2005).
- 3) S.J. Lee, K. Lee and H. J. Yoo, "Analysis and Implementation of Practical, Cost-Effective Network on Chips," *IEEE Design and Test of Computers*, Vol. 22, No. 5, pp. 422-433, (2005).
- 4) Y. Honma, M. Kameyama, Y. Fujioka and N. Tomabeichi "VLSI Architecture Based on Packet Data Transfer Scheme and Its Application," *Proc. of 2005 IEEE Int. Symp. on Circuits and Systems*, pp.1786-1789, (2005).
- 5) Y. Fujioka, N. Tomabeichi and M. Kameyama : "Functional-Unit-Level Packet Data Transfer Scheme for A Highly Parallel VLSI Processor," *Proc. of Int. Conf. on Computers and Devices for Communication*, CD-ROM, pp.9-13. (2006).