

## 動的再構成可能ハードウェア上へのスケーラブルスイッチの実装 に関する検討

坪井秀幸 塩澤恒道 奥山祐市 名古屋彰

†NTT 未来ねっと研究所

〒239-0847 神奈川県横須賀市光の丘 1-1

E-mail: hideyuki@exa.onlab.ntt.co.jp

**あらまし** 動的再構成可能ハードウェアである PCA(Plastic Cell Architecture)上で回路が協調動作するために使用される通信スイッチを実装する方法について検討した。具体的には、PCA が持つ再構成要素を用いた 2 種類のスイッチの構成方法を提案し、これらのスイッチをプロトタイプチップである PCA-1 チップ上に実装してサイズおよびスループットを評価する。ここで提案する 2 種類のスイッチを組み合わせることにより PCA 上で動作するアプリケーションの特徴に適合したスケーラブルなスイッチを構成することができる。

**キーワード** 再構成可能ハードウェア, PCA, スイッチ, 通信, スケーラビリティ, スループット

## Implementation of Scalable Communication Switches on PCA

Hideyuki TSUBOI Tsunemichi SHIOZAWA Yuichi OKUYAMA and Akira NAGOYA

NTT Network Innovation Laboratories

1-1 Hikari-no-oka, Yokosuka-Shi, Kanagawa, 239-0847 Japan

E-mail: hideyuki@exa.onlab.ntt.co.jp

**Abstract** This paper shows implementation of two types of switches used for communication between objects on the reconfigurable computing architecture called Plastic Cell Architecture (PCA). Each switch uses the programmable portion and the fixed portion for performing simple functions which are the feature of PCA. We have implemented these switches on the prototype chip PCA-1 and evaluated their sizes and their communication throughputs. By combining these switches, it is possible to reconfigure the scalable communication switch to suite the feature of the application which is performed on PCA.

**Keyword** Reconfigurable Hardware, PCA, Communication Switch, Scalability, Throughput

### 1. はじめに

微細加工技術の進歩による LSI の集積度の向上はシステム全体 (または大部分) をシリコンチップに実装するシステムオンシリコンの実現を可能としている。一般に、シリコン上に実装されるシステムは複数のモジュールから構成され、各モジュールは協調して動作することでシステムに要求される機能を実現する。さらに各モジュールはさらに分割されたサブモジュールの協調動作によりモジュールの機能を実現する。このような階層を設けた実装法は、大規模化する設計対象に対して既存の設計資産の再利用やハードウェアの特徴である並列実行による負荷分散を図るために重要である。

LSI 内部の各モジュールでは、入力される処理要求に対する一連の処理を分割して実行する機能分散、予め定められた方法で同一機能を有する複数モジュールから一つのモジュールを選択して処理要求を分散させる負荷分散の組み合わせによって処理が実行される。このように並列に動作するモジュールを協調動作させることによって処理を実行することは、プロセッサによるソフトウェアの逐次的な実行に対して、性能や消費電力等の点で最適化の範囲を拡大することとなる。

モジュールの協調動作による機能分散および負荷分散効果を拡大するためには、予め定められた種類や個数のモジュールにより処理を実行するのではなく、処理の内容や進捗に適応させてモジュールの種類や数

を動的に変更可能とすることが望ましい。処理の内容に応じて変更可能な再構成可能ハードウェアとして FPGA (Field Programmable Gate Array) があるが、FPGA は少量多品種の ASIC の代替やプロトタイプ実装といった利用が大部分である。

我々は、モジュールの並列動作と静的／動的な再構成を可能とするアーキテクチャとしてプラスチック・セル・アーキテクチャ<sup>[1-3]</sup>を提案している。プラスチック・セル・アーキテクチャは動的に再構成可能で並列に動作可能な回路を生成する回路構成要素と回路構成要素上に構成された回路が他の回路と通信を行うための通信要素とを組み合わせた構造を持ち、回路構成要素上に構成された回路同士が通信によって協調して動作する特徴を持っている。また、我々はプラスチック・セル・アーキテクチャ上で協調動作するモジュールの機能や個数が動的に変化する特徴を活用した実装例<sup>[4-8]</sup>や実行環境<sup>[9]</sup>を示してきた。

本稿では、プラスチック・セル・アーキテクチャ上に構成された回路が協調動作するために必要となる通信手段の実装方法について述べる。以下 2. ではプラスチック・セル・アーキテクチャの概要およびその上に構成された回路間での通信方法について述べる。3. では回路間の通信手段として 2 種類のスイッチを実装する方法について述べる。4. ではプロトタイプチップである PCA-1 チップ上にこれらのスイッチを実装して評価した結果について述べる。

## 2. プラスチック・セル・アーキテクチャ

プラスチック・セル・アーキテクチャの概要とその上に構成される回路間での通信方法について述べる。

### 2.1. PCA の概要

プラスチック・セル・アーキテクチャ<sup>[1-3]</sup> (Plastic

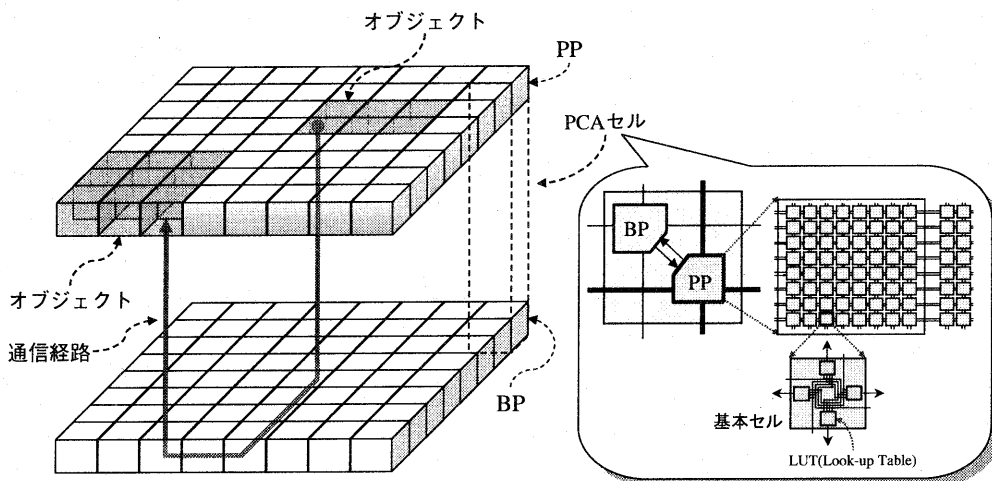


図 2.1 Plastic Cell Architecture(PCA)

Cell Architecture: 以下 PCA と略記)の構成を図 2.1 に示す。PCA は相互に接続された均一なセル(以下 PCA セルと呼ぶ)で構成され、各 PCA セルは可変部(以下 PP と呼ぶ)と組み込み部(以下 BP と呼ぶ)とから成る。

PP は書き換えが可能なプログラマブル論理の単位であり、隣接した PCA セルの PP と相互に接続され、論理ゲートや記憶素子から成る論理回路を構成することができる。1 個以上の隣接した PP に回路情報を書き込むことで、当該 PP 上に同一の PCA セルに含まれる BP を入出力とする機能オブジェクトを定義することができる。また、PP は LUT で構成されることから、データを記憶するメモリオブジェクトとして定義することもできる。

BP は予め定義された固定機能からなり、隣接した PCA セルの BP と相互に接続されている。BP は外部入力および PP 上から命令を受信して動作する単純な固定回路であり、通信を行うための経路設定、PP の読み出し／書き込み、および PP 上に構成された回路の初期化のための命令を実行する。

以下、プロトタイプ LSI である PCA-1 チップ<sup>[13]</sup>に基づいて具体的に説明する。PP は 4 つの 4 入力 1 出力の LUT で構成されるベーシックセルが  $8 \times 8$  個で構成され、各 LUT の入力は隣接したベーシックセルの出力と接続されている。PP を機能オブジェクトとして使用する場合、LUT に書き込む構成情報に応じて各 LUT は、配線、状態を保持する記憶 (FF)、隣接した LUT の出力を入力として論理演算結果を出力する論理ゲートとなる。PP をメモリオブジェクトとして使用する場合、4bit  $\times 1024$  word のメモリとなり、PP 単位での読み出し、書き込み可能である。

PCA-1 チップでサポートされている BP が解釈実行する 12 種類の命令を表 2.1 に示す。これらの命令を組

み合わせることによってチップ外部およびチップ上に構成された機能オブジェクトから任意の PCA セルの PP に対する構成情報やデータの書き込みおよび読み出し、および以下 2.2 で述べるオブジェクト間の通信が実現される。

表 2.1 PCA-1 チップでサポートされている BP の命令

命令	機能	記事
Clear	メッセージの最後	経路解除
Open	PP との接続オープン	
Close	PP との接続クローズ	
Cl/f	機能オブジェクトの書き込み	
Cl/m	メモリオブジェクトの書き込み	
CO	PP の内容を読み出す	
COCl	CI 命令付きの CO 命令	
West	以後のデータを西へ	経路設定
North	以後のデータを北へ	経路設定
East	以後のデータを東へ	経路設定
South	以後のデータを南へ	経路設定
PP_out	以後のデータを PP へ	経路設定

## 2.2. PCA におけるオブジェクト間の通信

PP 上に定義された各オブジェクトは通信を行う場合、経路設定情報を出力した後にデータを出力する。経路設定情報は送信先となるオブジェクトへの相対位置を経路設定命令(West, North, East, South, PP\_out)の列で指定したものである。この時、送信先オブジェクトへ到達するまでの各 BP では経路の設定毎に経路設定命令が消費され、BP は経路が設定された入力に到着した後続の Clear 命令以外の命令およびデータを設定された経路へ出力する動作を繰り返すので、送信先オブジェクトへはメッセージのみが送信されることになる。各 PCA セル内の BP は 5 方向(West, North, East, South, PP)からの入力要求を矛盾無く 5 方向に出力するためのアービトレーション機能を有しており、経路設定命令で指定された方向への経路が既に設定されている場合には、設定されている経路が解除されるまで後続の経路の設定やデータは設定済みの経路上にバッファリングされて待たされる。設定された経路は設定されている経路上を伝播する Clear 命令を受信することで解除される。

PCA における動的再構成は、PP を書き換えることによってオブジェクトの機能や数を変更する方法と、オブジェクト間の通信経路の書き換えによって機能を変更する方法との 2 つがある。一般に経路の書き換えによる機能の変更は PP の書き換えに比べてオーバーヘッドが低く抑えられる利点はあるが、予め必要な種類および数のオブジェクトを PCA 上に構成しておく必要がある。現実には PCA デバイスの性能(PP 書き換え時間や経路設定時間等)と許容されるオーバーヘッドとのトレードオフによって上記の 2 つの方法が組み合わされて使用される。

以下では、PCA の動的再構成の特徴を生かし、オブジェクトが協調して動作するための通信経路の実現方法としてオブジェクト間の通信を行うスイッチの実装方法を検討した。

## 3. 通信スイッチの構成

ここでは PCA 上に構成された N 個の並列に動作する機能オブジェクトが N 個の任意オブジェクトへ通信を行う  $N \times N$  構成のスイッチを実装する方法について述べる。第一の方法は BP に予め設定された経路を PP 上に構成された機能オブジェクトが順次選択してメッセージを送信する方法であり、第二の方法は送信元が通信要求発生毎に随時通信先までの経路を BP に設定する方法である。

### 3.1. PP スイッチ

PP スイッチは、先頭部分に送信先を指定するヘッダ情報を付加したメッセージを入力し、PP 上の機能オブジェクトがヘッダ情報を参照して送信先を選択するセルフルーティングスイッチ(self-routing switch)である。このようなスイッチは  $M (\geq 2)$  個の入出力ポートを持つ  $M \times M$  スイッチを複数使用することでスケラブルに  $nM \times nM$  スイッチに拡張可能である<sup>[10]</sup>。M 個の入出力ポートを持つセルフルーティングスイッチとして動作する機能オブジェクトを PP スイッチプリミティブと呼び、このプリミティブを図 3.1 に示すように組み合わせることで構成されるスイッチを PP スイッチと呼ぶ。

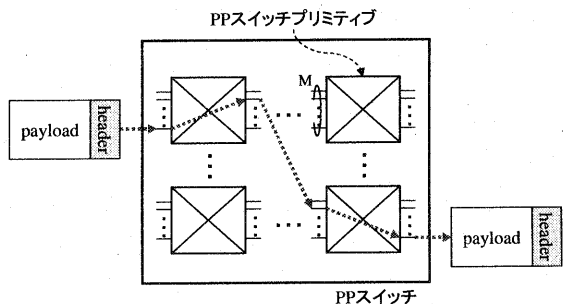


図 3.1 PP スイッチ

PP スイッチはオブジェクト間の通信経路を通信毎に再設定することなく送信元がヘッダを書き換えることで送信先を変更できる利点があるが、スイッチの規模が拡大するとスイッチのサイズや通信のレイテンシが増大する欠点があると考えられる。

### 3.2. BP スイッチ

2.2 で述べた BP のアービトレーション機能を利用して構成するのが BP スイッチである。図 3.2 に示すように、送信元オブジェクト側に送信先に対応する BP スイッチプリミティブを設け、送信元オブジェクトが送信時にセレクタ回路等で 1 つを選択してメッセージを

送信する。送信要求が発生すると BP スイッチプリミティブは送信先への経路設定命令に引き続いて受信したメッセージを出力し、送信が終了すると Clear 命令を出力して経路設定を解除する。

BP スイッチは通信の経路設定が完了すればスイッチのサイズに影響されることなく、固定回路で PP 上の回路より高速に動作する BP の本来の通信能力に近い速度でメッセージの送信が行える利点があるが、送信元からの通信要求発生毎に通信経路を設定するため経路設定命令列をヘッダとして付加するオーバーヘッドが欠点となる。

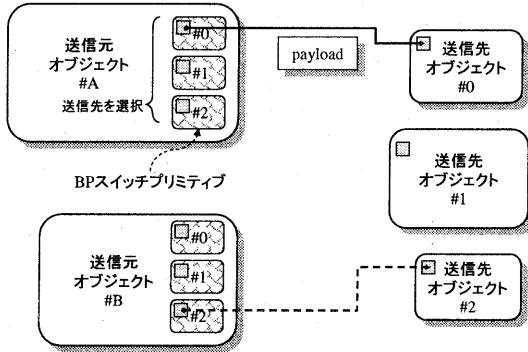


図 3.2 BP スイッチ

#### 4. 実装および評価

前述の 3 で述べた PP スイッチプリミティブおよび BP スイッチプリミティブを用いて  $8 \times 8$  スイッチを構成し、PCA-1 チップ上に実装し、実装サイズおよびスループットを評価した。

##### 4.1. PP スイッチの実装

$2 \times 2$  の PP スイッチプリミティブを PP 上に実装した結果を図 4.1 に示す。PP スイッチプリミティブは 3 つの機能オブジェクトからなり、実装サイズは 8PCA セルである。入力側の 2 つの機能オブジェクトと出力側の機能オブジェクトとに分離したのは以下の理由による。

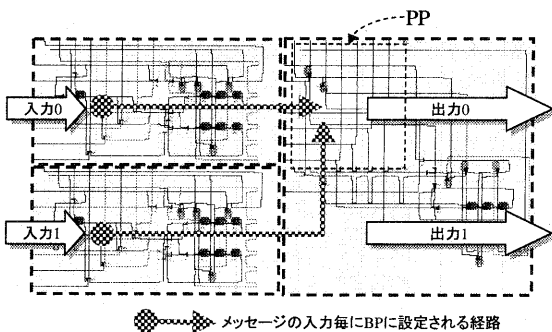


図 4.1 PP スイッチプリミティブの実装例

a) BP の持つアービトレーション機能およびバッファリング機能によって、2 つの入力に同時にメッセージが到着した場合でも矛盾無く待ち合わせが行われる。

b) 入力側の機能オブジェクトおよび出力側の機能オブジェクトを組み合わせることによって任意のサイズの PP スイッチプリミティブを構成することが可能となる。

$2 \times 2$  の PP スイッチプリミティブを用いて構成される  $8 \times 8$  のスイッチの構成を図 4.2 に示す。 $8 \times 8$  のスイッチは  $2 \times 2$  の PP スイッチプリミティブを 4 行  $\times$  3 列に配置し、それらの各出力はスイッチ構成時に BP 経路で予め定められた入力に接続されている。各列の出力側のオブジェクトは列毎に異なる条件でメッセージ内のヘッダ値を参照して何れの出力ポートに送信するかを決定し、受信したメッセージを転送する。

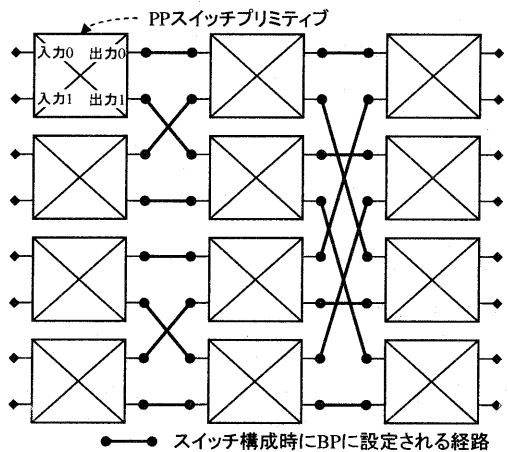


図 4.2  $8 \times 8$  の PP スイッチ

##### 4.2. BP スイッチの実装

BP スイッチプリミティブはメッセージを受信する毎に送信先までの経路命令列を出力する必要がある。このような機能の実装方法として経路命令列の長さ以上の状態数を持つステートマシンを PP 上に実装する方法もあるが、ここでは経路命令列を繰り返し格納するメモリオブジェクトと命令列によらず同一の回路となる機能オブジェクトとの組み合わせで実装した。機能オブジェクトは送信要求としてメッセージを受信すると、経路命令列を格納しているメモリオブジェクトの内容を読み出して経路命令列を出力した後、受信したメッセージを出力する。このように、メモリオブジェクトとの組み合わせで実装したのは以下の理由による。

a) 経路命令列の内容や長さを変更する場合、PP 上に構成される回路の変更に比べてメモリ内容の変更は

容易であり、メモリオブジェクトに格納されている値を変更することで、マンハッタン距離で約1000PCAセルの距離までにある任意のオブジェクトを送信先として指定することが可能である。

b) 経路命令列の長さが数十程度以上であれば、PP上に数十状態以上を持つステートマシンを構成する実装に比べて、メモリから命令列を読み込んで出力する実装の方が高速に動作する。

BPスイッチプリミティブは図4.3に示すように機能オブジェクトおよびメモリオブジェクトからなり、実装サイズは3PCAセルである。

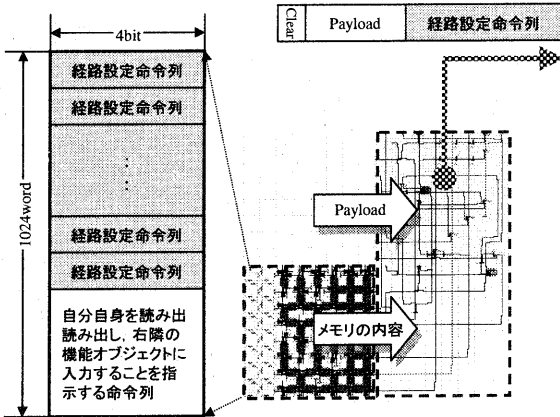


図 4.3 BP スイッチプリミティブの実装例

BPスイッチプリミティブを用いて8×8のスイッチを構成するためには送信元となる各オブジェクトにBPスイッチプリミティブを8個配置する必要がある。また、各BPスイッチプリミティブから同一の送信先への経路はどれかのBPで経路設定時に競合することになるが、スイッチのスループットを高めるためにはこのような競合箇所を広範囲のBPに分散することが望ましい。一方、BPの通信が競合している箇所にスイッチとは無関係の他の通信経路が設定されると逆にスイッチの性能を大幅に低下させる危険性がある。このような理由からBPスイッチを構成する際にはBPでの経路設定の競合が発生する範囲を極小化し、その範囲も実装サイズに含めることとした。一般に、N×NのBPスイッチを構成するための競合箇所はN×(N-1)PCAセルでカバーすることができる。

### 4.3. 実装サイズの評価

8×8および64×64のスイッチをPCA-1チップ実装するために使用するPCAセル数を表4.1に示す。64×64のPPスイッチを構成するためにはPPスイッチプリミティブを32行×6列に配置することになるが、各PPプリミティブで使用されるヘッダがメッセージの先頭であることを保障するためには、3列目の出力側オブジェクトから出力されたメッセージの最初のヘッダを

削除する必要があるため、この処理を行うヘッダ削除オブジェクトを各出力毎に追加する必要がある。ヘッダ削除オブジェクトの実装サイズは1PCAセルである。

表 4.1 スイッチの実装サイズ(PCAセル数)

	8×8	64×64
PPスイッチ	96 (8×4×3)	1,600 (8×32×6+64×1)
BPスイッチ	248 (3×8×8+8×7)	16,320 (3×64×64+64×63)

表4.1に示した比較では、BPスイッチに比べてPPスイッチは少ないPCAリソースを使用して実装される。ただし、ここでは全ての送信元オブジェクトと全ての送信先オブジェクトとの通信を可能とすることを前提としているが、以下のような場合についてはBPスイッチの実装サイズを大幅に削減することができる。

- 各送信元オブジェクトが必ずしも全ての送信先オブジェクトへの通信を行う必要が無い場合。
- 比較的長いサイクルで送信先を変更すればよく、送信先オブジェクトの経路を格納しているメモリオブジェクトの内容の書き換えで対応できる場合。
- 負荷分散のように、メッセージの送信毎に送信先は変わるが、送信先のオブジェクトを予め決定した順序に選択すればよい場合。

### 4.4. スループットの評価

8×8スイッチをPCA-1チップに実装し、以下に示す各条件でスループットの評価を行った。各条件についてメッセージのpayload長を4~64[Byte]の範囲で変化させて測定を行った。ここで、Payload長は、ヘッダや経路設定命令列を除いたメッセージのデータサイズである。

- 1つの送信元から特定の送信先にデータを連続して送信した場合のスループット。
- 8つ全ての送信元から同一の送信先にデータを連続して送信した場合のスループット。

条件Iの測定結果を図4.4に示す。BPスイッチについては送信元と送信先との間のマンハッタン距離dを

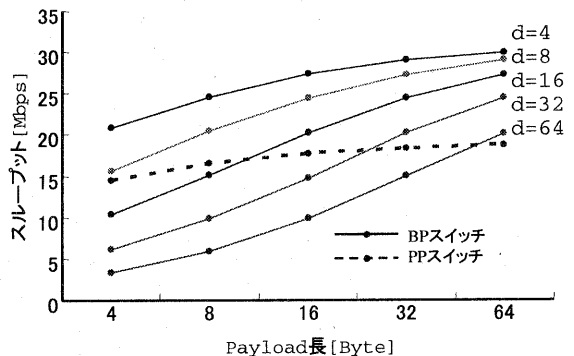


図 4.4 条件 I での測定結果

変化させた場合についても測定した。

PPスイッチはPayload長に対して安定したスループットを保証するが、BPスイッチはdが8以上またはPayload長が64[Byte]以上の領域でPPスイッチのスループットを上回る。このようなことから、アプリケーションの要求スループットやPayloadサイズに適應させてスイッチの構成を変更したり、これらを組み合わせて構成したりすることは有効であると考えられる。

条件Iはスイッチ内部で競合が発生しない場合の1ポート当たりの基本性能であり、8×8のPPスイッチでは図4.1に示す構成から測定スループット値の4倍がスイッチの最大スループットとなる。一方BPスイッチでは各送信元が異なる送信先に経路を設定する場合にはBPでの競合の生じない経路設定が可能であることから測定スループット値の8倍が最大スループットとなる。

条件IIの測定結果を図4.5に示す。BPスイッチの場合、BPの経路設定が競合する箇所までの経路設定オーバーヘッドは他の7個の送信元からのデータ転送時間内に行われるのであればスループットに影響しないので、測定では8個の送信元から送信先への平均マンハッタン距離を約22とした。PPスイッチは条件Iの場合と同様にPayload長に対して安定したスループットとなっている。BPスイッチではPayloadサイズが8[Byte]の時に最大スループットを示しているが、これは送信するデータの大部分がBP上にバッファリングされた状態となることによる。

条件IIの測定値は各スイッチの最小スループットとなり、送信オブジェクトから見たスループットの平均値は何れのスイッチの場合も1/8となる。

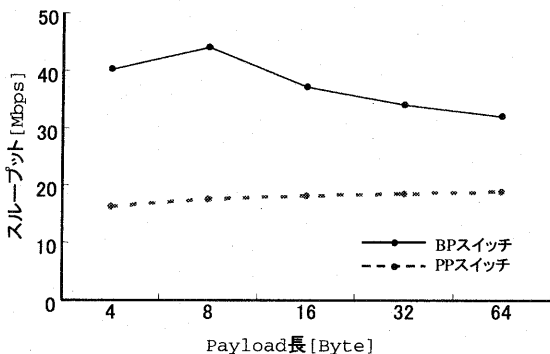


図4.5 条件IIでの測定結果

## 5. おわりに

再構成可能ハードウェアであるPCA上に2種類のスイッチを実装してそのスループットを評価した結果について述べた。本稿で述べたPPスイッチおよびBPスイッチは実装サイズとスループットに関して特長があ

る。PCA上にアプリケーションを実装する際には、そのアプリケーションの動的な振る舞いに適合するようにこれら2種類のスイッチを適宜組み合わせた多段接続網とすることによって各スイッチの特徴を生かしたスケーラブルなスイッチを構成することができる。

今回の実装サイズおよびスループット評価はPCAの実現性確認のために試作したLSIであるPCA-1上で行ったが、本検討結果を反映させてPCAの持つスケーラビリティを十分に活用してアプリケーションが実装できるようにBPの命令やBP-PP間のインタフェースの改良を検討している。

## 文 献

- [1] 永見康一, 塩澤恒道, 小栗清, "自律再構成可能アーキテクチャ", 設計自動化研究会, 87-3(1998).
- [2] Kouichi Nagami, Kiyoshi Oguri, Tsunemichi Shiozawa, Hideyuki Ito and Ryusuke Konishi, "Plastic Cell Architecture: A Scalable Device Architecture for General-Purpose Reconfigurable Computing," IEICE Transactions on Electronics, vol. E81-C, No. 9, pp.1431-1437, September 1998.
- [3] H. Ito, R. Konishi, H. Nakada, K. Oguri, A. Nagoya, N. Imlig, K. Nagami, T. Shiozawa and M. Inamori: "Dynamically Reconfigurable Logic LSI - PCA-1," Proc. of 2001 Symposium on VLSI Circuits, pp. 103-106, Jun. 2001.
- [4] T. Shiozawa, K.Oguri, K.Nagami, H.Ito and R.Konishi, "A Hardware Implementation of Constraint Satisfaction Problem Based on New Reconfigurable LSI Architecture," Proc of 8th International Workshop of Field-Programmable Logic and Applications (FPL'98), pp426-430, Aug. 1998.
- [5] 塩澤恒道, Norbert Imlig, 永見康一, 小栗清, "Communicating Logicによる汎用情報処理機構の実現", 信学会コンピュータシステム研究会, 99-91(1999).
- [6] T. Shiozawa, N. Imlig, K.Nagami, K.Oguri, A. Nagoya, and H. Nakada, "An Implementation of Longest Prefix Matching for IP Router on Plastic Cell Architecture," Proc. of 10th International Workshop of Field-Programmable Logic and Applications (FPL2000), pp805-809, Aug. 2000.
- [7] N. Imlig, T. Shiozawa, R. Konishi, K. Oguri, K. Nagami, H Ito, M. Inamori and H. Nakada: "Programmable Dataflow Computing on PCA," IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, Vol. E83-A, No. 12, pp. 2409-2416, Dec. 2000.
- [8] N. Imlig, T. Shiozawa, K. Nagami, Y. Nakane, R. Konishi, H. Ito, and A. Nagoya, "Scalable space/time-shared stream-processing on the run-time reconfigurable PCA architecture," Proc of 8th Reconfigurable Architectures Workshop (RAW2001), April 2001.
- [9] 中根良樹, 永見康一, 塩澤恒道, Norbert Imlig, 名古屋彰, 小栗清, "自律再構成可能アーキテクチャ PCA における実行時資源管理の方法", pp.67-72, 情処学会 DA シンポジウム 2001.
- [10] 青木利晴, 青山友紀, 濃沼健夫, "広帯域 ISDN と ATM 技術", 電子情報通信学会, 1995.