

高並列推論エンジンPIEの LSI版スイッチングユニットの設計

1D-9

高橋 栄一, 小池 汎平, 山内 宗, 田中 英彦
(東京大学工学部)

1. はじめに

我々は、高並列推論エンジンPIEのネットワークとして、負荷分散適応網の使用を予定している[1]。既に、PIEEE[2]の結合網をMSI、SSIを使用して試作した[3]。今回、さらに大規模なPIEを構築する際の基本構成部品として使えるよう、SU(スイッチングユニット)をゲートアレイ化した。このSUチップは、4×4のクロスバスイッチ、データ幅8ビットで、ルータの動くマスタモードとスイッチングのみを行うスレーブモードの2つのモードで動作する。また、負荷分散と相手指定の2つの通信方法をサポートする。使用したデバイスは、6000ゲート、179ピンのゲートアレイである。本稿では、このSUについて報告する。

2. SUチップの概要

SUチップの概要について、以下に述べる。

2.1 特徴

① 通常の行先指定の通信方法の他に、自動負荷分散機能を有している。これは、空いている出力ポートから入力されている行先プロセッサの負荷情報を比較して、その負荷量が最小の行先へ自動的にルーティングするものである。

② ルーティング機能を持つマスタモード、あるいは、スイッチング機能のみのスレーブモードで動作させることができる(図1)。1チップは4×4のクロスバスイッチで、データは8ビットであり1つのチップをマスタ、他をスレーブとして用いることにより任意のビット幅の4×4クロスバスイッチを構成できる。

③ さらに多段構成も可能で、これにより大規模なネットワークが構成できる。

④ スwitchingは半二重・回線交換であり、同期通信、及び、非同期通信、マルチキャストが可能である。

⑤ アービトレーションは回転式で、1クロック、あるいは2クロックで行える。

2.2 構成

図2に示すように、SUチップは5つのブロックから構成されている。

① クロスバスイッチ

4入力4出力のクロスバスイッチで、制御線DIRにより双方向半二重通信が可能である。未接続の出力ポートからは、そこに接続可能な全ての行先プロセッサの中で最小の負荷情報が入力されており、それらの中で最小の負荷情報を未接続の入力ポートからさらに前段のSUへ転送する。

② ルーティング部

行先指定時は、そのアドレスをデコードして接続されるべき出力ポートに対する接続要求を、負荷分散時は、負荷量最小の出力ポートに対する接続要求信号を作成してアービトレーション部へ伝える。

③ アービトレーション部

同時発生した接続要求に対し、回転式の公平なアービトレーションを行う。接続要求は、各段が同期動作を行う場合は1クロックで、非同期で動作する場合でも2クロックでアービトレーションされる。

④ 負荷モニタ

未接続の出力ポートから入力される行先プロセッサの負荷情報を比較して、負荷量最小の行先プロセッサに接続できる出力ポートを決定する。

⑤ コミュニケーション制御部

次段に対する制御信号を作成する。

SUチップの各ポートは、次のような信号線で構成されている。

- ① データ(8ビット)・・・双方向
- ② REQ・・・接続要求
- ③ LREQ・・・負荷分散接続要求
- ④ DIR・・・データの転送方向指定
- ⑤ REL・・・接続解除
- ⑥ STB・・・ストローブ
- ⑦ ACK・・・アクノリッジ

また、他に次のような制御ピンを持つ。

- ① RESET・・・リセット
- ② CLOCK・・・クロック
- ③ STAGE(2ビット)・・・何段目かの指定
- ④ CHNODE・・・マスタ/スレーブ
- ⑤ ARMODE・・・アービトレーションを1クロック/2クロックで行う
- ⑥ IH・・・クロック・インヒビット
- ⑦ CAO-CAE、CBO-CBE、CCO-CCE、CDO-CDE・・・接続情報をマスタチップからスレーブへと伝えるためのピン
- ⑧ その他・・・内部のフリップフロップをスキャンするためのピン

また、表1に各ブロックのゲート数を示す。

ブロック名	ゲート数
クロスバスイッチ	896
アービトレーション部	2240
ルーティング部	580
負荷モニタ	545
コミュニケーション制御部	112

表1 各ブロックのゲート数

2.3 動作

行先指定通信と負荷分散通信を行うことができ、それぞれ同期/非同期、1対1/マルチキャストもサポートしている。

① 行先指定通信

データ線の下位8ビットに行先のアドレスを出力し、LREQをインアクティブ、REQをアクティブにする。そして、行先プロセッサからACKが返ってきたら、REQをインアクティブにする。これで2つのプロセッサ間是一本の回線で接続された状態となる。データの転送は、DIRで転送の方向を指定し、STB(同期式)、あるいは、STBとACK(非同期式)を用いてを行う。マルチキャストを行う場合は、接続要求を出してACKを待つという動作を必要なだけ繰り返す。RELをアクティブにすると、接続が解除される。

② 負荷分散通信

LREQ、REQ共にアクティブにして、行先プロセッサからのACKを待ち、それからREQをインアクティブにして、データの転送を行う。他は同様である。

2.4 予測性能

SUチップの動作速度は、クロックを25MHzとして、以下のように見積ることができる(カッコ内はアービトラジョンを1クロックで行う場合)。

- ・行先指定の経路設定 113 (73) ns
- ・負荷分散時の経路設定 111 (71) ns
- ・接続解除 51 ns
- ・負荷情報の伝搬遅延 39 ns
- ・データ転送遅延 15 ns

2.5 使用デバイス

使用したゲートアレイは、富士通のCMOS、UHBシリーズ(高速、大規模、多ピンタイプ)である。内部ゲート数は、2入力NAND換算で約7000ゲートであるが、本チップで使用したのは入出力バッファに1000、内部ロジックに4300ゲートである。パッケージはPGA(ピングリッドアレイ)179ピンを用いた。

2.6 構成例

SUチップを4つ(マスタ1、スレーブ3)使用すると、

- ・ネットワーク制御線 4 (REQ、LREQ、REL、DIR)
- ・双方向データ線 32
- ・順方向通信制御線 13
- ・逆方向通信制御線 4

だけの信号をもった4x4のクロスバスイッチを構成することができる。これをさらに多段構成することにより、さらに大規模なオメガ網を構成することができる。

3. おわりに

これまでスタンダードロジックで作成していた回路をゲートアレイ化することによって、

- ① システムの小型化
- ② 省電力化
- ③ 配線遅延の減少による速度の向上
- ④ 組立の省力化と高信頼性化

を図ることができた。

現在、今回紹介したチップを用いて64x64オメガ網の実装の検討を進めている。これについては、別途報告する予定である。

<参考文献>

- [1] 坂井、小池、田中、元岡、“動的負荷分散を行う相互結合網の構成”、情処論、Vol.27, No.5, 1986.
- [2] 小池、山内、野田、田中、“高並列推論エンジン実験環境PIEEE -全体構成-”、第34回情処全大、4P-3, 1987.
- [3] 山内、小池、野田、田中、“高並列推論エンジン実験環境PIEEE -自動負荷分散ネットワーク-”、第34回情処全大、4P-5, 1987.
- [4] 富士通半導体デバイス CMOSゲートアレイデザインマニュアル UHBシリーズ

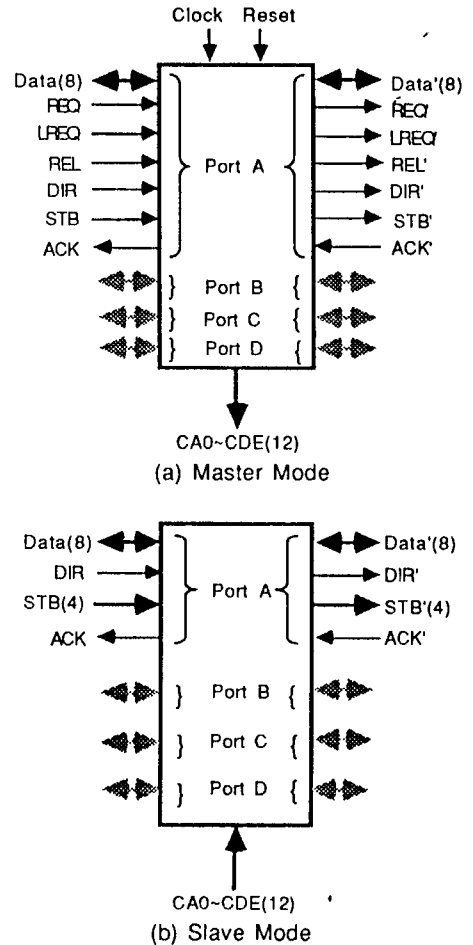


図1 マスタモード/スレーブモード

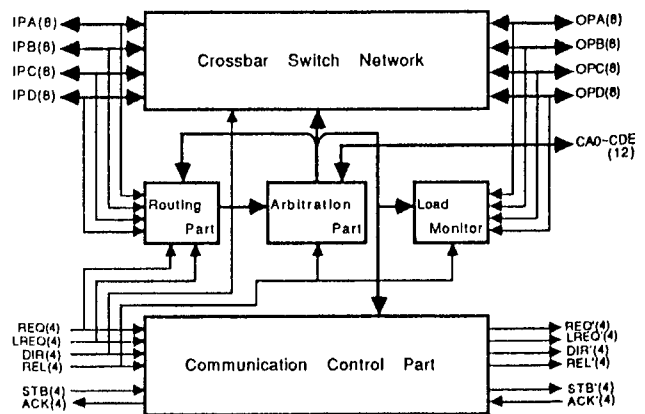


図2 SUチップの構成