

VLSI アーキテクチャと設計自動化技術の将来

小栗 清

NTT 光ネットワークシステム研究所
〒239-0847 神奈川県横須賀市光の丘 1-1

0468-59-4232

oguri@acm.org

あらまし. 集積度のさらなる向上を生かすにはこれまでの枠組みを大きく変更しなければならない。考えられる解は2つある。第1はタイミングを抽象化して、すなわち非同期回路として、実時間に依存しない順序のみのインタフェースを持つ階層化や部品化を行うことである。こうすると設計の困難さをソフトウェアによる階層化や部品化に近づけることができる。第2は大規模なモノの均質な構成要素となり得るゲートよりも大きい何かを定義し、設計と製造をその上下に分けることである。本稿では第2の考え方に基づく新しいアーキテクチャ、プラスチック・セル・アーキテクチャ(Plastic Cell Architecture: PCA)の提案を行う。

キーワード FPGA、可変構造計算機、PCA、布線論理、セルラオートマトン、非同期回路

Breakthrough in VLSI Architecture and Design Automation Technology

Kiyoshi Oguri

NTT Optical Network Systems Laboratories
1-1 Hikarinooka Yokosuka Kanagawa 239-0847 Japan

+81-468-59-4232

oguri@acm.org

Abstract

VLSI Architecture and Design Automation Technology will have to change against for ever-increasing complexity, high frequencies, connection delay dominance, low voltages, and surging currents. We propose a new general-purpose homogeneous architecture based on the cellular automata model and programmable logic devices which divides design and manufacturing into two parts and decreases complexity extremely. It consists of a dual-structured array of cells accommodating a fixed "built-in part" and a freely programmable "plastic part". We call this composition the "Plastic Cell Architecture" (PCA).

key words FPGA, Configurable computer, PCA, Wired logic, Cellular automaton, Asynchronous circuits

1 はじめに

競争に勝つためには、良くて安いものをできるだけ短期間で開発する必要がある。“良くて安い”とは、高機能、高性能、高拡張性、低消費電力、小さい、軽い、低コストなどを意味する。これらを達成するために、VLSIの高集積化、設計の自動化、方式改良などがさらに押し進められる。新素材や新素子の発明は、大きな変化をもたらすが予測が難しいため議論から除くと、変化の中心は微細加工技術の進歩に伴うVLSIの高集積化である。

高集積化によって設計規模が増大するため、その効率化のためには、設計の抽象度の向上、階層化、設計の再利用などが必要となる。抽象度の向上はソフトウェアの設計と同様の機能に関する抽象度向上だけでなく、物理的タイミングに関する抽象度の向上も不可欠となる。現状のクロックによる順序からクロックによらない順序のレベルが要求される。一方、階層化と設計の再利用は、まず少なくともタイミング設計の方式が決定されていることが必須であり、さらにこの前提となるタイミング設計の方式によって効果が異なる。設計規模が増大すると、実時間による規範なしのタイミング設計は論外としても、すべてをクロックで制御することは次第に困難となってきている。

高集積化によって面積と動作周波数当たりの消費電力は変化しないが、トランジスタの動作速度が向上し、これを有効に使うための設計を行う、すなわち動作周波数を上昇させるのが普通なので、面積当たりの消費電力は増大してしまう。消費電力の増大は製品の価値を下げるだけでなく、放熱の限界からさらなる高集積化の妨げになる。無駄な変化による電力消費は極力抑えるべきで、クロックがその第一候補となっている。またメモリ構造は最も離れた場所へのアクセスを最も近い場所へのアクセスと同等の時間で行うために大きな電力を消費しているということが広く認識されつつある。

さらに、高集積化によって配線の単位長当たりのRCによる遅延が増大する。回路の接続関係が同じならば、配線の長さも減少し、結果として配線遅延は変化しないが、トランジスタは高速動作するので、配線遅延の割合が大きくなる。配線は論理設計とレイアウト設計に依存するため、論理設計とレイアウト設計を同時に最適化することが重要になるが、良いアルゴリズムは知られていない。

配線の遅延が論理設計において支配的になったのはこれが初めてではない。CMOSによる高集積化が進む前のECLによる高速システムの設計では、配線の距離による電磁波としての伝播遅延が支配的であった。それがCMOSによる高集積化により距離が問題となくなり、しばらくは配線の長さよりも論理素子の接続関係が遅延を決定し

たので、多段論理式を最適化することが回路を最適化することとなり、論理設計の自動化が進んだ。また論理ゲートを構成単位としてもLSIの全体が扱える大きさであったこと、LSI内を単一のタイミング抽象度で扱えたことも論理設計の自動化が成功した理由である[1]。

以上説明したように、集積度のさらなる向上を生かすにはこれまでの枠組みを大きく変更する必要があることが分かる。規模の増大を単なる階層設計でまかなおうとすると、構成要素の大きさが大きくなり、次第に均質な構造ではなくなってくる。たとえばHW/SW協調設計による解決では様々な構成のメモリ、プロセッサ、バス、DSP、専用回路などが構成要素となり、設計の自動化は困難になると思われる。特にハードウェアでは機能的動作と物理的動作が共に正しく設計できる必要があるために、階層設計がソフトウェアの場合よりも難しい。さらに、仮に設計は階層的に行えたとしても、VLSIを階層的に作っていくことはできないため、階層設計はソフトウェアの場合ほど徹底できない。

それではどうするか、考えられる解は2つある。第1はタイミングを抽象化して、すなわち非同期として、実時間に依存しない順序のみをインタフェースとする階層化や部品化を行うことである。こうすると設計の困難さをソフトウェアによる階層化や部品化に近づけることができる。第2は大規模なモノの均質な構成要素となり得るゲートよりも大きい何かを定義し、設計と製造をその上下に分けることである。広い意味でプロセッサとメモリはこの何かに相当するが、この何かによって布線論理が構成できることが望ましい。以下にこの第2の考え方に基づく新しいアーキテクチャ(プラスチック・セル・アーキテクチャ(Plastic Cell Architecture: PCA))について述べる。

2 プログラム論理と布線論理

ある論理的機能を実現しようとする時、さまざまな方法を探ることができる。もっとも手軽なのはCPU(Central Processing Unit)とメモリによるプログラム論理である。CPUとメモリを用意しておけば物理的構成を変えなくて様々な機能を実現することができる。論理回路により直接、機能を実現する布線論理を使えばより高性能なものを得ることができる。しかし、物理的にものを作る行程が必要となり、手軽というわけにはいかない。さらに布線論理による専用ハードウェアの性能は、単一のCPUの性能向上に数年で追い越されてしまう場合が多い。

専用ハードウェアがCPUに対して大幅な性能向上をなし得ないのはメモリを使用しているからである。アドレスによる読み出し/書き込みサイクルで特徴づけられるメモ

りは非常に単純な機能ながら柔軟性に富み、これこそがノイマンアーキテクチャの本質である。CPUはそのメモリを最大限有効に使えるように作られた専用マシンである。従ってメモリを使っている限り、問題毎に専用マシンを設計してもCPUを大幅に越えることはできない。

それでは性能や柔軟性においてCPUとメモリを越えるためにはどのような構造を考えれば良いのであろうか。

アプリケーションをプログラム論理に展開する場合と布線論理に展開する場合の共通点と相違点を図1にまとめた。まず、コンパイラや論理合成系によって生成されたメモリイメージ、回路イメージを実際のメモリやASIC、FPGAに構成するところが、当然ではあるが、各々、記憶機能、工場、記憶機能と異なる。もう一つ重要な相違点がある。それはプログラム論理ではメモリ上のデータ構造を手続きの実行に伴って動的に構成できる点である。複雑な情報処理はデータの表現を生成、操作することによって行われるため、この相違点は本質的である。

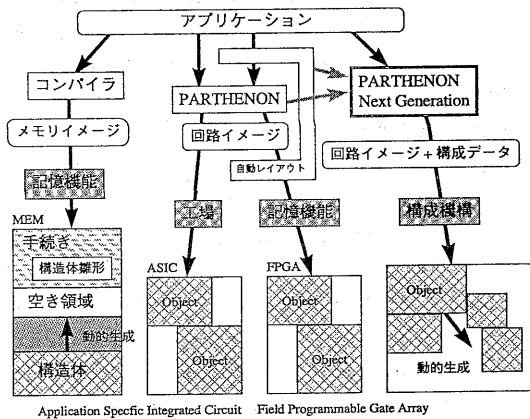


図1: プログラム論理と布線論理の共通点と相違点

3 汎用情報処理システムの二重構造

動的構成、すなわちある初期状態から出発して、後から新しい機能や構造を追加できるということは、構成の変更を行える部分(可変部)が存在し、かつその構成操作のための基本機能があらかじめ備え付けられている(本能)ということである。これを汎用情報処理システムの二重構造と呼ぼう。CPUとメモリの例では、メモリの内容が前者に相当し、回路として作られたCPUとメモリの機能が後者に相当する。

布線論理のみによって、この汎用情報処理の二重構造を形成するために、図2の構造を導入する。図2は布線論理を変更可能な部分(可変部)とあらかじめ特定の機能が組み込

まれている部分(本能)にわけて形成することを意味する。本能としては、少なくとも可変部を任意の布線論理として構成できるだけの機能がなければならない。

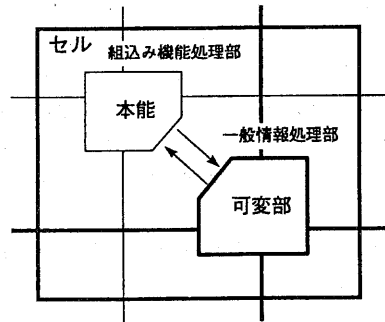


図2: プラスティックセル

図2の構造が単独で存在するものとする、構成処理は逐次的なものとなってしまふ。そこで、構成処理そのものにも並列性を導入するために、図2の構造が複数個連結されたものを考える。可変部については、可変部そのものを任意の結合網とできるので均一な結合でよい。また本能についても可変部が一樣である以上、特別な構造を導入することはできない。よってここではVLSIの作りと相性がよいと思われる2次元のメッシュ結合を選ぶ(図3)。

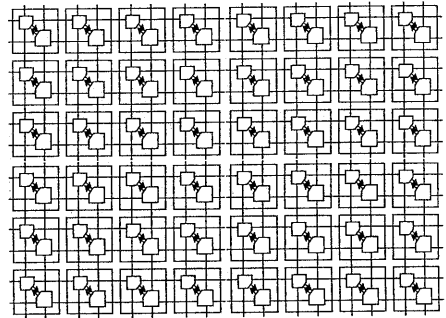


図3: プラスティックアレイ

本能の構成機能により、可変部は機能、サイズ共に任意の布線論理回路に構成できる。一旦構成された布線論理回路は他の部分をさらに構成したり、すでに構成された布線論理回路と通信し、連携して処理を進めることができる。すなわち、図2、図3の構造により、布線論理と並列処理の枠組みの中で汎用情報処理システムの二重構造を作ることができる。

図2、図3による汎用情報処理システムをプラスチックセルアーキテクチャ(Plastic Cell Architecture: PCA)と名付ける。Plasticは「形成力のある」という意味で用い

た。以後、図2をプラスチックセル(または単にセル)、図3をプラスチックセルアレイ(または単にセルアレイ)と呼ぶ。

プラスチックセルアーキテクチャは規則正しい構造を持つため、VLSI向きであり、メモリに匹敵する集積化の可能性もある。さらに機能の構成を製造後に行うため、製造欠陥が与える影響をより少なくでき、一層の高集積化に対応できる。また2次元のメッシュは容易に折り畳むことができ、LSIの3次元方向への拡張も可能である。

4 プラスチックセルアーキテクチャの設計

4.1 オブジェクトとメッセージ通信

動的に生成される布線論理回路をオブジェクトと呼ぶ。オブジェクトを構成するセルの可変部は互いに直接結合されて回路を形作る。オブジェクトの境界が定義できるためにはオブジェクトは境界を越えて可変部の結合を持つてはならない。オブジェクト間の通信は本能を介してのみ可能である。本能は可変ではないので通信はあらかじめ決められたメッセージ通信の形態をとる(図4)。

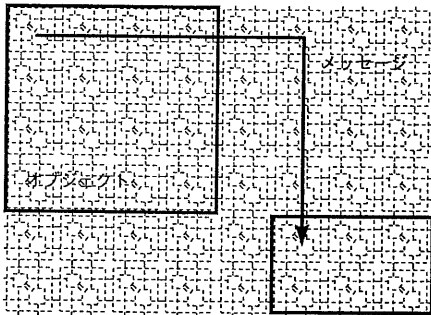


図4: オブジェクトとメッセージ

4.2 オブジェクトの利用に必要な処理

複雑な並列動作は既に構成されたオブジェクトが、テンプレートから新しいオブジェクトを生成し、利用し、削除することによって行われる。その手順は次の様になる。

1. まず必要な空き領域を探索し確保する。
2. 構成情報を注入して損傷が発生しないように回路を構成する。
3. 回路が完成したら、誤動作しないように回路を初期化する。
4. 本能とオブジェクトとの間の通信ゲートを開く。
5. 処理をメッセージとして送付する。

6. 解放メッセージを注入し領域を開放する。

4.3 経路指定によるメッセージの伝播

経路情報を共有する複数の情報をまとめて一つのメッセージとし、メッセージは処理の単位である命令から構成されるものとする。メッセージは

- 経路設定情報
- メッセージ本体
- 経路設定解除命令

から構成される。経路設定情報はメッセージの後続部分をどのように伝えるかをセルの本能部分に記憶させるもので、経路設定命令(west命令、north命令、east命令、south命令)からなる。セルの本能部分は一旦経路を設定されるとすべての情報をそのまま経路設定された次のセルに送り続ける。すなわち、シフトレジスタとして動作するようになる。経路設定命令は経路を設定するセルで消費され、次々とシフトレジスタを構成していく。経路設定解除命令(clear命令)は経路設定されたセルでは消費されずに伝播されて次々とシフトレジスタを構成していたセルの本能部分を初期化する。

複数メッセージのデッドロックを避け易くするため、各セルの本能は図5に示すように独立に動作可能なwest、north、east、southの4つの入力ポートを持ち、それぞれは4つの出力ポートへメッセージを転送できるものとする。出力ポートが競合したときは待ちが発生する。この構成によって直交する2つのメッセージなどを扱うことが可能となる。メッセージの経路をeast-southであるものとwest-northであるものに制限すると、循環待機条件が成立しなくなるのでとりあえず、デッドロックを回避することができる。もっと効率の良い方法の発見は今後の課題である。

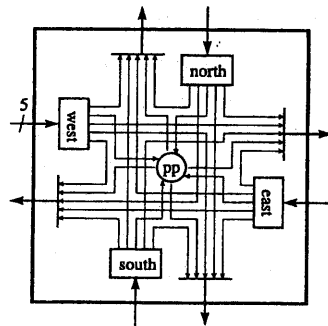


図5: 本能部分のポート構造

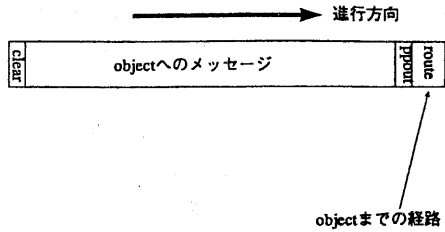


図 9: オブジェクトへのメッセージ

命令 命令は各メッセージの説明に現れた、clear 命令、pp_open 命令、pp_close 命令、config_in 命令、config_out 命令、west 命令、north 命令、east 命令、south 命令、pp_out 命令の 10 命令である。これらのために 5 ビットを使用する。clear 命令のために 1 ビットを使用し、残りの 9 命令を 4 ビットにエンコードする。こうすれば構成データに clear 命令を除く任意のパターンが使えるようになり、都合がよい。

4.5 可変部の構成

従来の FPGA は論理関数を LUT(Look Up Table) すなわち小規模のメモリで実現する場合であってもフリップフロップと配線を専用に持ち、これらを接続するセレクタやトランスファゲート、そしてそれらを構成するメモリから構成される。また、LUT には構成情報の格納のためのデコーダと LUT を構成するためのデコーダを個別に設けることが多く、また誤った構成情報から回路を守るための付加回路なども必要である。これらは必要なものではあるが、その機能が使われないときは全く無駄になってしまい、メモリの様な均質な構造でもない。このためこれまでの FPGA はメモリやゲートアレイに比べると集積量や速度に於いて大きく劣っている。このため現状ではプロトタイピングや

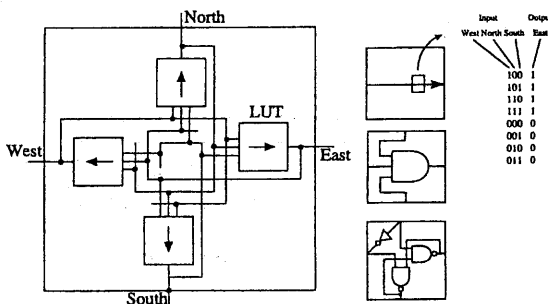


図 10: トライステート出力を持つ LUT だけからなる可変部のセルの構造

エミュレーションなどのような傍流で利用されることが多く、その将来性に疑問も持たれている。ここに、そもそも FPGA をもっと均質な構造にすることはできないのかという疑問がある。図 10 にトライステート出力を持つ LUT だけからなる可変部のセルの構造を示す。LUT の内容を設定することによりセルは結線や論理ゲートそしてラッチなどに構成することができる。これを図 5 の構造に近づけた図 11 はまさに LUT だけからなる構造であるとともに、LUT の利用率が高い。これを Sea of LUTs と呼ぶ。LUT である小規模メモリへの構成情報の書き込み読み出しはシフトレジスタ構成で行う。Sea of LUTs の構成ではメモリと同様の集積度が期待できる。この結果、可変部をメモリオブジェクトとして使うときは完全にメモリと同様の集積度を使うことができる。一方可変部を布線論理オブジェクトとして使うときも従来 FPGA よりも高い LUT 集積度を使うことができる。性能については今後の検討を必要とするが、規則正しい構造はすべてに於いて有利であると考えている。

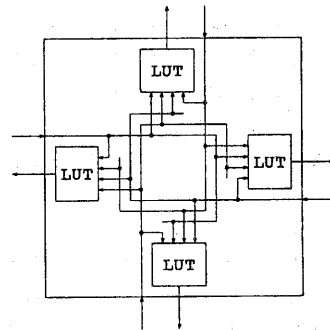


図 11: LUT だけからなる可変部のセルの構造

4.6 非同期アーキテクチャ

可変部ではレジスタやラッチを含むすべての布線論理を LUT のみで形成する。このためレジスタやラッチの配置位置や方向の自由度は高いが、グローバルクロック線を用意することができない。LUT を通過する信号をクロックに用いることは全く現実的でないため、可変部は非同期回路とせざるを得ない。しかしながら PCA は 2 次元のアーキテクチャであり、その拡張性を自然に生かす、地理的に広がりを持つ通信と処理をカバーするという将来像を考えると、グローバル通信は非同期とせざるを得ず、本能、可変部ともに非同期とすることは必然かもしれない。

PCA では回路を自動合成することが前提であるため、出来るだけトップダウンに構成できる非同期回路アーキテクチャを採用する。同期回路からグローバルクロック信号を

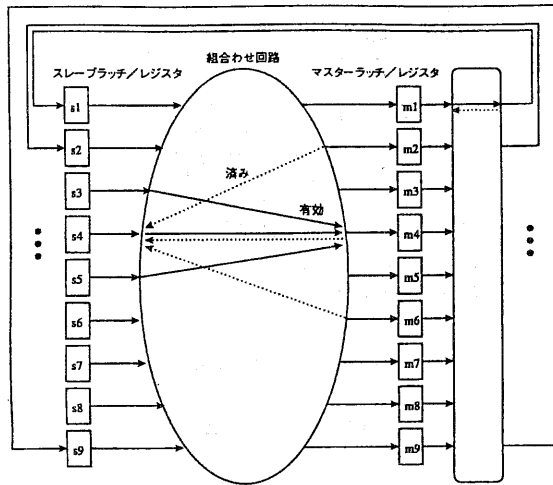


図 12: 非同期回路の構成

削除し、それを等価なタイミング信号に置換えることにより非同期回路を構成する方法を以下に示す。

ここでは Sutherland のマイクロパイプラインとは違ってレジスタや組み合わせ回路の性質は動作時に変化しないものとする (レジスタが組み合わせ回路の一部にならない)。そうするとレジスタには有効データが存在しているか存在していないかのいずれかになる。

またすべての因果関係を正しく評価するために、有効データは使われて次の有効データを生成するものとする。すなわち、有効データが使われもせず消失したり、使われたのに存在し続けたり、無から突然出現したりすることはないものとする。

さらに、有効データが生成された (到着した) ことは信号の変化でしか伝えられないので、簡単のため、有効データが生成される前には必ず、有効データが存在していない時間があることとし、この有効データの「なし・あり」をそのまま有効データ到着の信号とする。

有効データなしの時間を設けることにより、有効データが存在しない時に限って有効データを生成するという簡単なメカニズムで (有効データを消失させる) 上書きを防げる。

このような原則に基づく非同期回路構成方式を 2 相式という。

さて、同期回路ではレジスタに設定された有効データを数クロックに渡って利用することがある。有効データであることを到着でしか表現できない非同期回路では、このような場合、同期回路の 1 つのレジスタに対し少なくとも 2 つのレジスタを用意して、その間で有効データをループさ

せなければならない。

同期回路のレジスタをタイミングレベルで詳細に見るとマスターラッチとスレーブラッチからなっているので、簡単のため、非同期へ変換した回路もすべてマスターラッチとスレーブラッチの構造を持ち、処理を行う組み合わせ回路はスレーブラッチの出力からマスターラッチの入力までの間にあるものとする。この構造を図 12 に示す。

以下ラッチとレジスタを区別する必要のない時はレジスタという。2 相式の非同期回路ではレジスタ上の有効データのありなしが交互に配置されたときに最大性能となるので、初期化信号のハイレベルにより、すべてのマスターレジスタを有効データ存在、すべてのスレープレジスタを有効データ不在の状態とし、初期化信号の立下がりにより、すべてのスレープレジスタに有効データを生成し、すべてのマスターレジスタの有効データを消去することとする。初期化信号の変化は同時ではないのでこのような工夫を行う。

グローバルクロックに代るタイミング規則を例を使って説明する。s3, s4, s5 のデータが加工されて m4 へ書込まれる場合は、s3, s4, s5 へ有効データが到着したこと、加工が終了したこと、m4 が消去されたことのすべてを待って書込む。この書込み信号は済み情報として左へ送られる。s4 の有効データを m2, m4, m6 が利用していたとすると、s4 の有効データは m2, m4, m6 からの済み信号を待って消去する。

2 線式を使う場合は組み合わせ回路内の信号がすべて無効状態へクリアされたことを待つ必要があるのもう少し複雑となる。

レジスタをグラフの節、有効データ利用の関係をグラフの枝とするとときグラフが連結であることと、順序に関し元の同期回路と同じ動作をすることを証明できる。またこの構成はデッドロックフリーであることを証明できる。さらに 2 線式の論理回路はユニートであるので論理単純化にこの性質を有効に使うことができる。

5 関連する研究

プラスチックセルアーキテクチャは汎用情報処理システムの二重構造を、まず可変部を布線論理とし、次にこれを 2 次元のメッシュに敷き詰めて得られたものであり、FPGA の技術とセルラオートマトンの技術を結合させたものになっている。その生い立ちから当然のことではあるが、セル 1 個のものは可変部をメモリとすればノイマン型コンピュータとなり、セルアレイから可変部をのぞくとセルラオートマトンとなる。セルアレイの可変部をメモリとするとメッシュ結合網のマルチプロセッサとなる。可変部

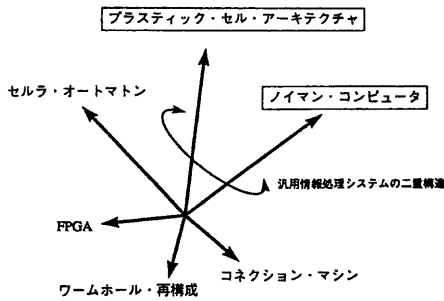


図 13: 関連する技術

から本能への制御バスを除き、本能を単なる可変部の構成バスとすると FPGA となる。可変部を布線論理としてこれを集中制御により一斉に同一のものに構成する場合に限定するとコネクションマシンとなる。

SRAM ベースの FPGA は外部の制御部により任意の布線論理に構成可能であり、動作中の構成を許すものも存在する。一方、外部制御ではなく内蔵の分散制御により必要に応じて構成を行うワームホール動的再構成などの研究が存在する [3][4][5][6]。また、セルラアルゴリズムを高速化するためにセルラオートマトンのセルに可変な構造を追加する研究が存在する [2]。プラスチックセルアーキテクチャと FPGA やワームホール動的再構成、可変構造セルラオートマトンとの違いは、汎用の情報処理機構であるための条件を議論しているかどうかにある。PCA の汎用情報処理モデルとしての位置付けを図 13、図 14 に示した。

6 おわりに

PCA により最初に議論した問題がどのように解決されるのかをまとめる。最初の議論では非同期化と均質な階層の導入は別の解決法としたが、結果として PCA は非同期システムであるので PCA はその両方の特徴を持つ。まず LSI の設計と製造を、繰り返し構造を持つ PCA のレベルに閉じこめたので、詳細物理レベルの設計は格段に容易になった。メモリと同様の徹底的な最適化が可能である。PCA 上の設計は論理設計もレイアウト設計もオブジェクト単位であり、LSI の規模とは無関係になった。クロックを廃したので無駄な動作による電力消費は回避される。またブロック間の非同期のインタフェースは物理的時間によらずブロック間を連携させることができ、ソフトウェアと同様の階層化部品化が可能になった。次に規則正しい構造により、従来よりもレイアウトの自由度が減少し、かつレイアウト結果の評価項目が単純化された。これは論理最適化と連携したレイアウト最適化アルゴリズムの開発がそれだ

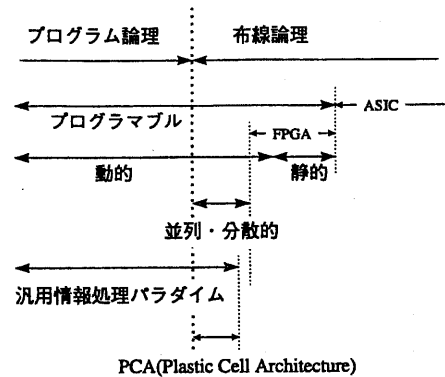


図 14: PCA の位置付け

け容易になる可能性があることを示している。

謝辞

非同期システムについては東京大学の南谷 崇教授、中村 宏助教授の指導を受けた。また本論文の内容は NTT 光ネットワークシステム研究所の永見 康一、伊藤秀之、小西 隆介、塩澤 恒道、Norbert Imlig、中田 広、稲森 稔の各氏、そして NTT コミュニケーション科学研究所の名古屋 彰氏、京都大学の中村 行宏教授、メリーランド大学の中島 和生教授との議論に基づくものである。

参考文献

- [1] Y. Nakamura, K. Oguri and A. Nagoya: "Synthesis From Pure Behavioral Descriptions," High-Level VLSI Synthesis, Edited by R. Camposano and W. Wolf, Kluwer Academic Publishers, pp.205-229, June, 1991 <http://www.kecl.ntt.co.jp/car/parthe/>
- [2] 中野博嗣, "動的可変バスを持つ並列計算機上の定数時間アルゴリズム," 情報処理学会誌 Vol. 38, No. 11, pp. 1019-1025 (Nov. 1997)
- [3] H. Schmit: "Incremental Reconfiguration for Pipelined Applications," in Proc. IEEE Workshop on FPGAs for Custom Computing Machines, pp. 47-55, April, 1997
- [4] W. Luk, N. Shirazi and P. Y.K. Cheung: "Compilation Tools for Run-Time Reconfigurable Designs," in Proc. IEEE Workshop on FPGAs for Custom Computing Machines, pp. 56-65, April, 1997
- [5] J. Burns, A. Donlin, J. Hogg, S. Singh and M. Wit: "A Dynamic Reconfiguration Run-Time System," in Proc. IEEE Workshop on FPGAs for Custom Computing Machines, pp. 66-75, April, 1997
- [6] R. A. Bittner, Jr and P. M. Athanas: "Computing Kernels Implemented with a Wormhole RTR CCM," in Proc. IEEE Workshop on FPGAs for Custom Computing Machines, pp. 98-105, April, 1997