大規模再構成可能データパスにおけるオンチップ・ ネットワーク・アーキテクチャの検討

島崎, 慶太 九州大学大学院システム情報科学府

長野, 孝昭 九州大学大学院システム情報科学府

本田, 宏明 九州大学情報基盤研究開発センター

メディプー,ファラハド 九州大学情報基盤研究開発センター

他

https://hdl.handle.net/2324/8698

出版情報:情報処理学会研究報告. ARC, [計算機アーキテクチャ]. 2007 (55), pp.115-120, 2007-06. Information Processing Society of Japan

バージョン:

権利関係:ここに掲載した著作物の利用に関する注意 本著作物の著作権は(社)情報処理学会に帰属し ます。本著作物は著作権者である情報処理学会の許可のもとに掲載するものです。ご利用に当たっては 「著作権法」ならびに「情報処理学会倫理綱領」に従うことをお願いいたします。 社団法人 電子情報通信学会 THE INSTITUTE OF ELECTRONICS, INFORMATION AND COMMUNICATION ENGINEERS 信学技報 TECHNICAL REPORT OF IEICE.

大規模再構成可能データパスにおける オンチップ・ネットワーク・アーキテクチャの検討

島崎 慶太[†] 長野 孝昭[†] 本田 宏明^{††} ファラハドメディプー^{††} 井上 弘士^{†††}

村上 和彰***

† 九州大学大学院システム情報科学府 †† 九州大学情報基盤研究開発センター ††† 九州大学大学院システム情報科学研究院

あらまし Large Scale Reconfigurable Data Path (LSRDP)は、二次元アレイ状に配置した多数の演算器を搭載し、演算器の種類と演算器 間のネットワークを再構成可能とするデータパスをもつプロセッサアクセラレータである.LSRDPにおいて、演算器数と演算器間のネットワー ク構成の間には面積に関してトレードオフの関係が存在する.本稿ではLSRDPに量子化学計算の二電子積分の初期積分部分を実装し、クロス バースイッチにて演算器行間ネットワークを実装する場合の検討を行った.その結果、各演算器を他の9個の演算器と接続した場合、LSRDP全 体の面積が最小となることが明らかになった.

On-chip Network Architecture for Large Scale Reconfigurable Datapath

Keita SHIMASAKI[†], Takaaki NAGANO[†], Hiroaki HONDA^{††}, FarhadMehdipour^{††}, Koji INOUE^{†††}, and Kazuaki MURAKAMI^{†††}

† Graduate School of Information Science and Electrical Engineering, Kyushu University
 †† Research Institute for Information Technology, Kyushu University
 ††† Graduate School of Information Science and Electrical Engineering, Kyushu University

Abstract Large Scale Reconfigurable Data Path (LSRDP) is a data path type processor accelarator. On the LSRDP, enormous Floating Point number processing Units (FPUs) are arranged as 2-dimensional array, and each FPU and FPU network is reconfigurable. There is a trade-off relation about the area size between the number of FPUs and network configuration for the LSRDP. In this research, the LSRDP area size is estimated under condition that the initial integral part of the quantum chemistry two electron integral calculation is implemented and the crossbar switch is assumed to implement the network connecting each FPU array. As a result, it was obtained that each FPU in an array is connected with the nine FPUs in next array for the minimized LSRDP area size.

1. はじめに

近年,学術研究,ライフ・サイエンス,自動車業界における 構造解析等の分野で,高度な科学技術計算のためのハイ・パ フォーマンス・コンピューティング(HPC:High Performance Computing)の必要性が高まっている.現在,HPCの分野で 主流となっている計算機システムは,汎用プロセッサを用いた スカラー型並列計算機やクラスタシステムである.TOP500[1] における性能ランキングにおいても,そのほとんどが汎用プロ セッサによる構成である.

一方で,汎用プロセッサにアクセラレータを付加した計算機

システムについても研究されている.アクセラレータは,汎用 プロセッサに対するコプロセッサとして動作し,非常に高い演 算性能を持つ.また,アクセラレータの多くは低消費電力に設 計されているという利点もある.HPC 分野において重要視さ れる電力あたりの性能が非常に高く,高速な計算機システムを 構築する上での選択肢として有効なものといえる.

実際に,アクセラレータに関する研究・開発は盛んに行われている.ClearSpeed 社[2]のアクセラレータボードである DualCSX600 PCI-X Board を装備した東京工業大学の計算機 である TSUBAME[3] や,東京大学において開発された HPC 向けのアクセラレータチップである GRAPE-DR プロセッサ[4] などがある.

しかしながら,アクセラレータには問題点もある.アクセラ レータは一般にチップ上に多数の演算器を配置し並列計算を行 うことで高い演算性能を実現する.そのため,データ供給のた めに非常に大きなメモリバンド幅を必要とする.しかし,現在 主記憶として使われる DRAM の速度は低速であり,十分なメ モリバンド幅を確保できない.このため,アクセラレータの高 い演算性能が抑えられてしまう(メモリウォール問題)[6][7]. このような問題に対して,キャッシュメモリなどのオンチップ メモリを使用することにより対処している.しかし,複雑な計 算では中間結果が大量に生じ,オンチップメモリにデータが収 まらなくなってしまうことがありうる.

そこで,筆者ら研究グループでは,メモリアクセス回数の増 大を抑え,かつ高い演算性能を実現するアクセラレータとして, 大規模再構成可能データパス(LSRDP:Large Scale Reconfigurable Data Path)を提案している.LSRDPは,チップ上に 浮動小数点数演算器(FPU:Floating-point Processing Unit) を多数並べ,それらをプログラマブルなスイッチと配線で接続 したものである.データを演算器間で直接受け渡すことができ るため,中間結果をメモリを介することなく転送でき,メモリ アクセスを削減することができる.特に,データ依存関係が深 い複雑な計算では相対的にメモリアクセス回数を少なくするこ とができるため,有利であるといえる.

LSRDPにおいて,演算器間を接続するオンチップネットワークの構成に関する検討は十分に行う必要がある.多くの演算器を相互接続する場合,配線面積がチップ面積に対して支配的になってしまう可能性がある.逆に,相互に接続する演算器の数を少なくすると,配線面積は小さくすることができるが,演算器間でのデータの受け渡しに制限がされるため,アプリケーションが実装できない場合がある.LSRDPではこのような場合,演算器をデータ転送用として利用することにより,直接接続されていない演算器同士でもデータのやりとりができるように対処している.しかし,データ転送用に演算器を使用するため,アプリケーションを実装するためにより多くの演算器が必要となり,結果としてより大きなチップ面積が必要となる可能性がある.

そこで本稿では,LSRDP にアプリケーションが実装できる という制約のもと,どのような相互接続網の構成が面積に関し て有利であるのかを検討する.そのために,構成の違う複数の 相互接続網に対してアプリケーションの実装に必要な演算器数 を求め,LSRDP の面積を求めた.

本稿の構成は,以下の通りである.2節で大規模再構成可能 データパスについて述べ,3節でLSRDPの面積の見積もり 方法について説明し,4節で相互接続網の構成に具体的にパラ メータを与え,面積の評価を行う.最後に5節で,本稿のまと めと今後の課題について言及する.

2. 大規模再構成可能データパス

2.1 概 要

大規模再構成可能データパス (LSRDP:Large-Scale Reconfig-



図1 LSRDP の概観

urable DataPath) は多数の演算器 (FPU:Floating-Point Unit) と,それらを相互接続するネットワーク (ORN:Operand Routing Network) を搭載し, FPU の演算内容と ORN 上の FPU 間 接続を再構成可能としたデータパスプロセッサである.LSRDP は多数の演算を並列実行することによって,高い演算性能を実 現する.さらにデータ依存関係にあるデータをメモリを介する ことなく演算器間で直接受け渡すことにより,演算量の増加に 伴うメモリアクセス回数の増加を抑制することが可能である.

2.2 ハードウェア構成

LSRDP の概観を図1に示す.LSRDP は演算器を二次元ア レイ状に並べた構成となっている.以後,横に並んだ演算器ア レイを行,縦に並んだ演算器アレイを列と呼ぶ.LSRDP は科 学技術計算を対象としているため,演算器を浮動小数点演算器 (FPU)としている.

本稿において,演算器間の相互接続網 (ORN) は図1のよう に演算器の行間の接続網を指すものとする.演算器間のデータ の受け渡しには以下の制約があるとする.

- 隣接行間では, ORN を経由して演算器間でデータの受け 渡しをする.
- 非隣接行間では、それらの間に存在する演算器を利用して データを受け渡しできる。
- 同一行でのデータの受け渡しはできない.
- すべての演算器は一方向からデータを入力し一方向へ出力 する.

LSRDP への入力データの供給はストリーミング・バッファ (SB: Streaming Buffer)により行われる.科学技術計算では, 大規模行列計算のように大量のデータに対して同様の処理を繰



図 2 アプリケーションの実装手順

り返すことが少なくない.よって,繰り返される処理の部分を LSRDP にデータパスとして実現しておき,SB にメモリから 絶え間なくデータを供給し続け,パイプライン処理をすること によりメモリ・アクセス・レイテンシを隠蔽することができる と可能となる.

2.3 アプリケーションの実装方法

LSRDP へのアプリケーションの実装は、ソースプログラム から、データパスの構成を作り出すことにより行う. アプリケー ションの実装手順を図 2 に示す。

(1) ソースプログラムを解析し、コア計算部のループボ ディーをデータフローグラフ (DFG: Data Flow Graph) に変 換する.

(2) ORN によって DFG のデータ依存関係が保たれるように DFG の各接点を LSRDP の各 FPU に割り当てる.

これにより、データパスの構成情報を作り出す。DFG の各接 点を各 FPU に割り当てる作業をマッピングと呼ぶ。

2.4 配線自由度とハードウェアコストのトレードオフ

LSRDP でアプリケーションを実行するためには,マッピン グが必要である.マッピングは,LSRDPのORNの構成を制約 として行われる.その制約により,マッピング可能(アプリケー ション実装可能)なLSRDPのFPUの行数は変化する.つま り,ORNの構成によってアプリケーション実装可能なLSRDP の面積は異なる.

図3にORNの構成例を示す.図3(a) は各 FPU から隣接す る行の FPU 全てに接続した場合であり,図3(b) は各 FPU か ら隣接する行の FPU のうち,3つ(下1/左1/右1)にだけ接 続した場合である.本稿で検討するORNの構成はこのように, 各 FPU が隣接する行の FPU に接続している数により区別す る.(a)の方を完全接続のORN と呼び,(b)の方を FPU 間接 続数3のORN と呼ぶようにする.FPU 間接続数に制限があ る(b)のような場合には,配線数の少なさから完全接続よりも ORNの面積は小さくなると予想される.しかし,ORN だけで はアプリケーションのデータ依存関係を維持することが難しく なる.このような場合,FPU をデータ転送用として使うこと によりデータ依存関係を維持することができるが,より多くの 行の FPU が必要になる可能性がある.このことを,図4を用 いて説明する.

図 4 の例では, ORN は FPU 間接続 3 のものであり, 両端の



図 4 FPU 間の接続に制限があるときのマッピング

FPUの出力を次の行の FPUの入力として与えることができな い.そのため, FPU2 つをデータ転送用として使用することで マッピングを行う.ただし,ORN が完全接続の構成の場合と 比べると,1行多くの FPU アレイが必要となってしまう.一般 に,FPU 間の接続に制限が無いほどマッピングに必要な FPU の行数は少なくなると考えられる.しかし,FPU 間を多くの配 線で接続するほど ORN の面積は増大する.つまり,ORN の 面積とアプリケーションをマッピング可能とする FPU の行数, つまりは FPU の総数であり FPU の総面積の間にはトレード オフ関係があるといえる.このトレードオフ関係によりアプリ ケーション実装可能な LSRDP の面積は ORN の構成により異 なると考えられる.

3. 大規模再構成可能データパスの面積

LSRDP の面積は,演算器 (FPU) と相互接続網 (ORN) の面 積の合計により求める. LSRDP の演算器の行数を M とする と,LSRDP の面積は以下の式で表される.

 $M \times (- fon FPU アレイの面積 + ORN の面積)$ (1)

3.1 FPU

FPU の面積は, FPU の横幅 F₋w と縦幅 F₋h の積 F₋w × F₋h で表すことにする.LSRDP の演算器の列数を N とすると, 一行あたりの FPU アレイの面積を以下の式で表される.

$$N \times (F_{-}w \times F_{-}h) \tag{2}$$

3.2 ORN

ORN はクロスバスイッチにより実現する.ORN のレイア ウトの概略を図 5 に示す.(a)の完全接続の場合,各 FPU の 出力は次の行のすべての FPU の入力となる可能性があるため, 図のように左端から右端までバスを伸ばす必要がある.そのた め,FPU の列数が N の場合,N 組のバスを縦に並べた構成と





図 5 構成の違う ORN のレイアウト略図

なる.(b)の FPU 間接続数3の場合,各 FPU の出力は次の行 の FPU3 つにしか入力されず,各 FPU の出力のバスは次の行 の FPU3 つ分にしか伸ばす必要はない.(a)では存在しなかっ たスペースを,他の FPU の出力のバスに使うことができ,(a) よりも面積を小さくすることができる.この例では,3組のバ スを縦に並べた構成となっている.他の構成の ORN に関して も同様にレイアウトすることができる.よって,ORN の面積 は FPU 間接続数に比例するといえる.

図 5 において入力側の配線の幅を MI_w, 配線間隔を MI_s とすると, バス 1 組の幅は, 64 × (MI_w + MI_s) となる.こ れより ORN の縦の長さは以下の式で表せる.

$$n \times 64 \times (MI_w + MI_s) \tag{3}$$

ここで,nとは FPU 間接続数のことで,図5の(a) はNで (b) は3である.ORN の横の長さは FPU1 行の横の長さに依 存する.なぜなら,FPU の入力データの配線面積と FPU の 面積を比べた場合,後者の方が大きいからである.以上より, ORN の面積は以下の式で表せる.

$$\{n \times 64 \times (MI_w + MI_s)\} \times (N \times F_w)\}$$

$$\tag{4}$$

式(3.1)に式(3.2)と式(3.4)を代入すると以下のようになる .

 $M \times N \times F_w \times \{F_h + n \times 64 \times (MI_w + MI_s)\}$ (5)

以上の式より LSRDP の面積を求める.

4. 評 価

本節では,構成の違う ORN をもつ LSRDP に具体的なアプ リケーションをマッピングすることにより,LSRDP の面積を 求め,ORN の構成に関して検討する.今回実装するアプリケー ションは,LSRDP での実行に適したデータ依存関係の深い計 算を選択した. 4.1 面積見積もりにおける各種パラメータの決定 4.1.1 FPU

本稿で検討する LSRDP に搭載する FPU のパラメータとして, GRAPE-DR の PE を選択した.GRAPE-DR の PE の詳 細については以下のとおりである.

- 90nm CMOS テクノロジを使用
- 0.6mm 角

本稿では,この PE を 1 行あたり 32 個並べることにする. つまり,LSRDP における FPU の列数は 32 である.30×0.6 で,LSRDP の一辺は約 20mm となる.20mm という大きさ は,実際にチップを製作するのに現実的な数字である.実際に, 最先端のグラフィック処理用途のアクセラレータでは,20mm 四方を超える面積のものも製品化されている.3節の式に与え るパラメータは以下の通りである.

$$N = 32$$

$$F_-w = 0.6(mm)$$

$$F_h = 0.6(mm)$$

4.1.2 ORN

ORN の面積は,第3節の式より求まる.実際にレイアウト設計をして,図5のようなORN が設計可能であることを確かめた. レイアウト設計は,CADENCE 社の Virtuoso Layout Editor を使用して行った.プロセステクノロジは,ASPLA(Advanced SoC Platform)社の90nmCMOS テクノロジを使用し,デザ インルールに違反しないようにした.ASPLA90nmCMOS テ クノロジの特徴は以下のとおりである.

- 最小加工寸法 90nm
- 6 層銅配線

本設計では, ORN の面積が最小となるようレイアウトを行った. つまり, デザインルールが定める最小寸法により設計をした.3節の式に与えるパラメータは以下のとおりとなる.

 $MI_w=0.14(um)$

 $MI_{-s} = 0.14(um)$

4.2 二電子積分計算の実装

本評価では,実装するアプリケーションとして,二電子積分 計算における初期積分計算を用いる.二電子積分計算は,非経 験的分子軌道法計算における重要な処理であり,全計算時間の 95%以上を占める.二電子積分計算の解法には小原のアルゴリ ズム[9]を採用した.初期積分計算は以下の特徴を有する.

- 4 重ループ構造の最内ループにて処理されるため,繰り返し連続実行される.
- データ依存関係がある多くの演算が存在する.
- 入力データが 17 個,出力データが1 個と少数である.



図 6 初期積分計算の DFG



図 7 各 ORN の構成における ORN の面積とマッピング可能な FPU の行数

このような特徴により,多数の FPU を二次元アレイに配置 した LSRDP アーキテクチャにて効率的に実行することができ る.初期積分計算の DFG を図 6 に示す.入出力ポートを含む DFG のノード数は 141 であり、最大幅は 51,深さは 14 であ る。内部演算構成として、四則演算が 99,符号反転が 3、べき 乗が 15、逆数が 1、開平計算が 1、指数計算が 2、誤差関数計 算が 1 となっている。

本稿では,対象アプリケーションの実装に必要となる FPU 行数が最小になるように人手でマッピングする.なお,各 FPU は除算,剰余算,平方根などの特殊演算すべてを含めた演算が 行えると仮定した.

4.3 面積見積もり結果

各 ORN の構成における ORN の面積ならびに初期積分計算 のマッピングに必要な FPU の行数として図 7 の結果を得た. 面積の数値軸で×20 をしているが,この 20 は FPU1 行の長 辺の長さのことである.つまり,3節の式における $F_w \times N$ で ある.

以上の結果をもとに,各ORNの構成における初期積分計算 がマッピング可能なLSRDPの面積を計算したものが図8であ る.面積の数値軸で×20しているのは前述しているのと同じ 理由による.

4.4 考 察

図 7 に各 ORN の構成における面積ならびにマッピング可能 な FPU アレイの行数を示す.

各 ORN の構成における ORN の面積に関しては, FPU 間 接続数の増加に比例して ORN の面積も大きくなっている.こ



図 8 各 ORN の構成におけるマッピング可能な LSRDP の面積

れは配線の面積が ORN の面積に対して支配的であることを意味している.完全接続のときの面積は 0.57 × 20mm² となっており,縦の長さは FPU 一辺と同じほどの大きさになっている. 配線面積のみにチップの面積を費やすことはトランジスタの集積度の点からも好ましくない.したがって,LSRDPのORNとしては,できるだけ FPU 間接続数が少ないものが望ましい.

次に,初期積分計算のマッピングに必要な FPU の行数につ いてだが,FPU 間接続数が完全接続から接続数9までは,行 数の変化なしにマッピングできている.接続数7からは少しず つ必要な行数が増え,接続数3のときには約3倍まで増えてい る.これはアプリケーション(DFG)の形状に依存するものと 考えられる.

図8より,LSRDP 全体の面積という評価指標において,二 電子積分の初期積分計算をマッピング可能な ORN の構成とし ては,FPU 間接続数が9のものが最適といえる.

5. おわりに

本稿では,LSRDPの構成要素であるORN に着目し,FPU やORNの面積を見積もり,アプリケーションを実装すること によって,できるだけ面積が小さくなるLSRDPの構成を検討 した.その結果,二電子積分の初期積分計算をアプリケーショ ンとして選択した場合には,ORNの構成としてFPU間接続 数が9の時面積が最小となることを明らかにした.

今後は,他の様々な科学技術計算に対しても同様の実験を行 い,アーキテクチャの設計空間を探索する必要がある.そのた めに,ORNの制約をパラメータとしたマッピングツールの開 発を行う.また,アプリケーションの特徴がマッピングに与え る影響の解析も課題である.最終的な目標である LSRDP アー キテクチャの決定,評価に向けて以上の課題に取り組む.

謝辞 日頃から御討論頂いております九州大学安浦・村上・ 松永・井上研究室ならびにシステム LSI 研究センターの諸氏に 感謝します.本チップの設計は東京大学大規模集積システム設 計教育研究センターを通し、Cadence ツールを用いて行われた ものである。なお,本研究は一部,科学技術振興機構戦略的創 造研究推進事業 CREST ならびに科学研究費補助金 (若手研究 A:課題番号 17680005) による.

文 献

- TOP500 Supercomputer Sites , http://www.top500.org/
- [2] ClearSpeed 社, http://www.clearspeed.com/
- [3] TSUBAME グリッドクラスタ(TGC), http://www.gsic.titech.ac.jp/ ccwww/tgc/
- [4] Grape-DR Project , http://grape-dr.adm.s.u-tokyo.ac.jp/
- [5] 野澤哲生, "512 個の演算器を集積 東京大学などが LSI 開発 1TFLOPS のボードを 2007 年に発売",日経エレクトロニクス, No.941, pp.36-37, 2006 年 12 月 18 日.
- [6] A. Salsbury, F. Pont, and A. Nowatzyk, "Missing the memory wall: The Case for Processor/ Memory Integration," Proc. of ISCA '96, pp. 90-101, May, 1996.
- [7] D. Burger, J. R. Goodman, and A. K "agi, "Memory Bandwidth Limitations of Future Microprocessors," Proc. of ISCA '96, pp. 78-89, 1996.
- [8] CADENCE 社 http://www.cadence.com/
- [9] S. Obara and A. Saika, "General recurrence formulas for molecular integrals over Cartesian Gaussian function," J. Chem. Phys. Vol98 no.3, August 1988.