社団法人 情報処理学会 研究報告
IPSJ SIG Technical Reports

2007－ARC－173 (20)
2007／6／1

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 Reconfigurable 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．

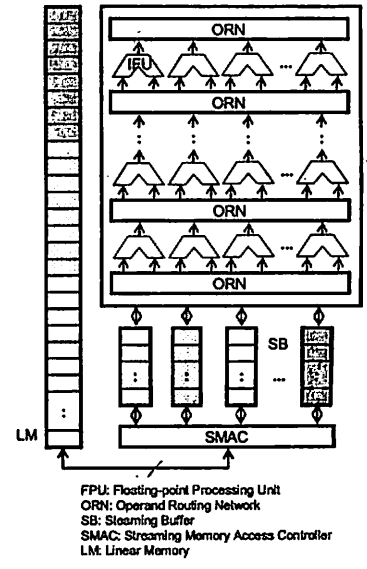

図 1　LSRDP の概観

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

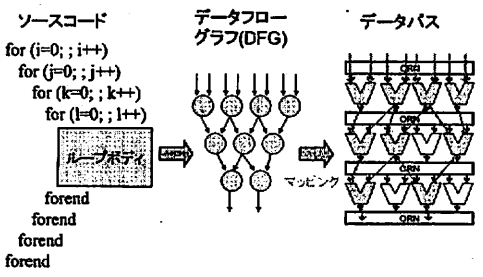

図 2 アプリケーションの実装手順

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

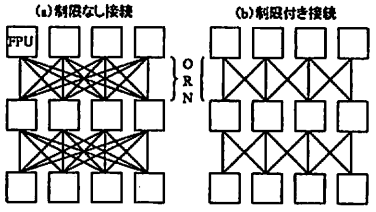

図 3 ORN の構成例

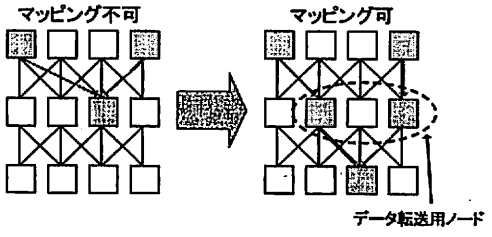

図 4 FPU 間の接続に制限があるときのマッピング

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．

$$M \times (\text{一行の } FPU \text{ アレイの面積} + ORN \text{ の面積}) \quad (1)$$

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．

$$N \times (F\_w \times F\_h) \quad (2)$$

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

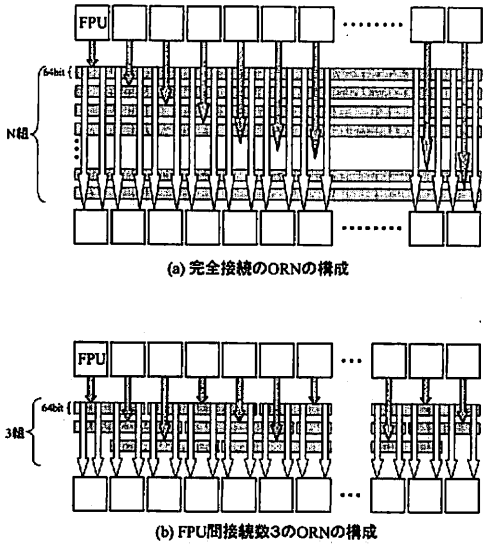

図 5 構成の違う ORN のレイアウト略図

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．

$$n \times 64 \times (MI\_w + MI\_s) \quad (3)$$

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．

$$\{n \times 64 \times (MI\_w + MI\_s)\} \times (N \times F\_w)\} \quad (4)$$

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．

$$M \times N \times F\_w \times \{F\_h + n \times 64 \times (MI\_w + MI\_s)\} \quad (5)$$

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．

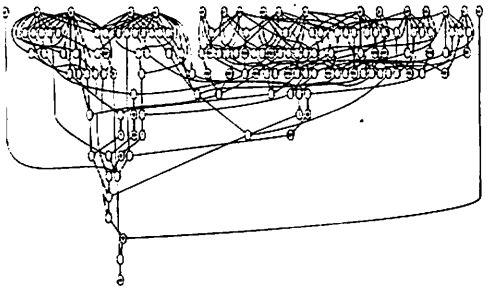

図 6　初期積分計算の DFG

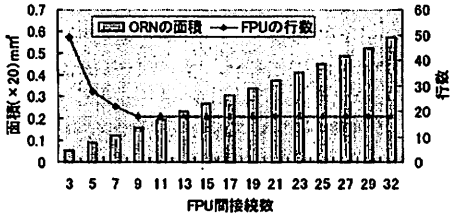

図 7　各 ORN の構成における ORN の面積とマッピング可能な FPU の行数

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 $0.57 \times 20mm^2$ 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．

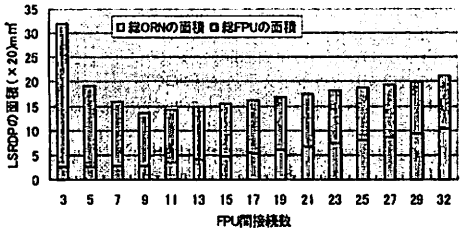

図 8　各 ORN の構成におけるマッピング可能な LSRDP の面積

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

[1] 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.