

世界最速のFPGA ソーティングアクセラレータの初期検討

臼井 琢真[†] 眞下 達[†] 松田 裕貴[†] 小林 諒平[†] 吉瀬 謙二[†]

[†]東京工業大学 大学院情報理工学研究科

1 はじめに

ソーティングは、画像処理、データ圧縮、データベース処理など、多くのアプリケーションにとって非常に重要な計算カーネルであり、常に高速化が求められている。その手段として、FPGA を用いたアクセラレータの使用が挙げられる。FPGA アクセラレータは、アプリケーションに特化した演算パイプラインとデータ供給機構を実現する回路をFPGA上に実装することによって、アプリケーションによってはCPUやGPUと比較して高い演算性能を達成することができる。

本稿では、まずFPGAで実現するソーティングの基本的なデータパスを紹介し、次に世界最速のFPGA ソーティングアクセラレータの構想を説明し、その性能を見積もる。

2 FPGA に適したソーティングデータパス

本章では、FPGA に適した基本的なアルゴリズムであるソーティングネットワークとマージソートツリーを紹介する。

2.1 ソーティングネットワーク

ソーティングネットワーク [1] とは、配線と比較交換器で構成される、数列をソーティングするための数理モデルである。配線は値を伝播し、比較交換器は2本の配線を入力、伝播されてきた値の大小比較し、一方に小さい方の値を、他方に大きい方の値を出力する。この一例を図1に示した。ソーティングネットワークにおいては入力系列に問わずソートの手順が予め定まっているため、図中の各枠を互いに独立に実行できる。このため、FPGA に実装する際にはこの間にパイプラインレジスタを置くことのみで、複数のデータ系列を並列にソートできる。このため、ソーティングネットワークはFPGA で高速なソーティングを行うためのコンポーネントとして非常に好ましい。

ソーティングネットワークは比較交換器の接続方法を変更することで、バブルソート、バイトニックソート、奇偶マージソートなどといった様々なソーティングアルゴリズムを実現できる。[1]にて様々なソーティングネットワークをFPGAに実装した際のロジック使用量等が詳細に議論されており、奇偶マージソートを採用したものが最も高効率とされている。このため我々のアプローチにおけるソーティングネットワークにも奇偶マージソートを用いている。

2.2 マージソートツリー

前述のソーティングネットワークでは、大規模なデータを対象にする場合にその長さに応じて入出力幅を大きくしなければならず、回路規模が膨大になってしまう。このため、大規模なデータのソーティングに対応するた

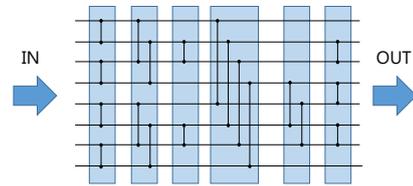


図 1: 8 要素に対する Batcher の奇偶マージソーティングネットワーク。

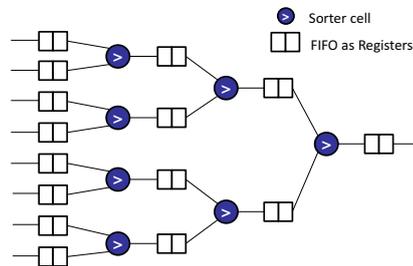


図 2: マージソートツリー。

めにはソーティングネットワークとは別のアプローチを取る必要があることとする。

マージソートは高速なソーティングアルゴリズムとして広く用いられ、このハードウェア実装としてマージソートツリーが存在する。これを図2に示した。“Sorter cell”は2つの入力を比較・選択し出力する組み合わせ回路である。マージされるソート済みデータ系列数が入力数より多い場合、これを繰り返し用いることで大規模なデータのソーティングにも対応できる。我々の過去の研究においても、FACE[2]はマージソートツリーを複数実装しメモリバンド幅使用効率を高めたことでVC707において3.4GHz動作のCore i7-4770と比較して最大10.06倍の性能向上を達成している。

しかしマージソートツリーには、特にマージの最終段階において扱う要素の幅がボトルネックになり、メモリの性能を最大限活用できないといった欠点がある。また、FPGAのハードウェア資源には限りがあるため一度にマージする系列の数にも限界がある。従ってこの問題を解決することを世界最速のFPGA ソーティングアクセラレータを実現する際の課題とする。

3 世界最速のFPGA ソーティングアクセラレータの構想

本章では、はじめに前述した問題点解決のための我々のアプローチを示し、次にその性能やハードウェア使用量の見積もりを行う。近年次世代メモリの研究・開発が盛んであるため、メモリバンド幅の制約は無視することにする。現在世界最速のFPGA ソーティングアクセラレータ [3] が64bit要素を扱っているため、我々は64bit要素のソーティングが [3] よりも高速になることを目指す。

High-speed Sorting using Portable FPGA Accelerator
Takuma USUI[†], Susunu MASHIMO[†], Yuki MATSUDA[†],
Ryohei KOBAYASHI[†], and Kenji KISE[†]
[†]Graduate School of Information Science and Engineering,
Tokyo Institute of Technology

3.1 ベースとなる FPGA アクセラレータ: FACE

我々は既に, FACE[2] という高速な FPGA ソーティングアクセラレータを設計・実装しており, 前述の通り高い性能を達成している. このハードウェアは更にオープンソースにて公開されており, 性能モデルも示されている. よって, 我々の設計するソーティングアクセラレータのベースとして以上のソーティングアクセラレータを採用することにする.

3.2 複数の要素を扱うマージネットワーク

世界最速の FPGA アクセラレータの実現にあたってはまず, 最終ノードにおけるボトルネックを解決しなければならない. [3] においては複数の要素を一度にマージするためのマージネットワークを提案している. 図3に1サイクルに複数要素を出力するためのソートセルを示した. 図中の FIFO 内の系列は全てソート済みであることに留意して頂きたい. FIFO 内先頭要素の結果に従いデキューされた要素とフィードバックされた要素を比較して並べ替えることによって, 1 サイクル中に複数の要素を出力することができる. これにより前述のボトルネックを緩和することができる.

3.3 多数の入力に対応するための省ロジック化

マージソートツリーでは最終ノードから平均して1サイクルに一定数の要素が出力されるため, 各層で平均して高々1個のソートセルしか動いていない. このため, ソートセルを各層で1つにし, FIFO をレジスタではなく Block RAM で実装することでロジックの大幅な削減を行うことが可能となる. この手法の概念を図4に示した. 図2の各レジスタが BRAM にまとめられ, ソートセルを Block RAM のアドレスに割り当てることで仮想化する. これはマージソートツリーの way の増大に大きく貢献する. 前述した複数の要素を扱うソートセルと併用する場合, フィードバック要素をレジスタに格納しているが, 同様に Block RAM にまとめることができる.

3.4 検討するソーティングアーキテクチャの性能考察

[4] ではベースとなるアーキテクチャにおいて複数の要素を扱うソートセルを用いた場合の性能が定式化されている. 今回はメモリについては考えていないため, ソーティングにかかるサイクル数 C について以下の式で推定できる. ただし, N は総データ数, M は複数要素を扱うソートセルが1サイクルで出力できる要素数, k はマージソートツリーの入力数, α はデータ系列のバッファリングにかかるサイクル数である. 複数要素を扱うソートセルにおいては入力はソートされている必要があるため, ソーティングネットワークでソートされる要素数は M とする.

$$C = \sum_{i=1}^{\log \frac{N}{M}} \left(\frac{N}{M} + \frac{N}{Mk^i} (3\log_2 k + 1) + k\alpha \right)$$

よって, 例として8要素を1サイクルでマージするソートセルを用いた 8192-way マージソートツリーを実装した場合, 512M 要素のマージに 6.74×10^7 サイクルかかる. これを [3] と同じ 200MHz で動作させると 0.673 秒かかる計算になり, スループットは 6.083GB/s を達成する. [3] は実測値で 3.1GB/s, メモリ性能を考えない理論値でも 4.8GB/s のスループットであるため, 我々の構想するシステムは世界最速となる可能性を持つ.

一方で, 複数要素を1サイクルで扱えるソートセルを用いてマージソートツリーを BRAM を用いて実装し

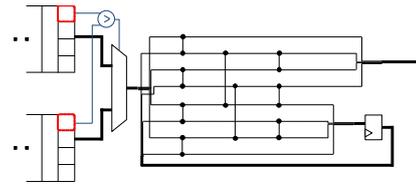


図3: 1サイクルに複数の要素を出力するソートセル.

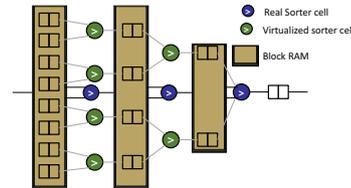


図4: マージソートツリーの省ロジック化手法.

た場合, 各仮想ソートセルに対する入力数の合計分とフィードバックデータ格納分だけ必要となる. よって M 個の要素を1サイクルでマージできるソートセルを L 個用いて $k = 2^{L\text{-way}}$ のマージソートツリーを構成した場合, i 層目では必要な BRAM 容量は1要素 $l\text{bit}$ としたときにフィードバックデータ格納分も含めて $lM \times (2 \times 2^{i+1} + 2^{i-1}) = 5lM \times 2^{i-1}$ であるため, 必要な BRAM 容量 $S(\text{bit})$ は

$$S = \sum_{i=1}^L (5lM \times 2^{i-1}) = 5lM(k-1)$$

となる. 前述の例だと 20.0Mbit 必要となる. しかし現存する FPGA である, 36kbit BRAM を 1030 個持つ Virtex-7 XC7VX485T を用いた場合合計 37.1Mbit の BRAM を持つため BRAM 容量に余裕がある. よって我々の構想する世界最速の FPGA ソーティングアクセラレータを実装できる可能性があることが示される.

4 まとめ

本稿において我々は, 世界最速の FPGA ソーティングアクセラレータの実現に向けたアプローチと見積もりを示した. 今後の課題として, より詳細なアーキテクチャの検討, 設計, 実装が挙げられる.

参考文献

- [1] K. E. Batcher. Sorting networks and their applications. In *Proceedings of the April 30-May 2, 1968, Spring Joint Computer Conference*, AFIPS '68 (Spring), pp. 307-314, New York, NY, USA, 1968. ACM.
- [2] R. Kobayashi and K. Kise. Face: Fast and customizable sorting accelerator for heterogeneous many-core systems. In *Embedded Multicore/Many-core Systems-on-Chip (MC-SoC), 2015 IEEE 9th International Symposium on*, pp. 49-56, Sept 2015.
- [3] Jared Casper and Kunle Olukotun. Hardware acceleration of database operations. In *Proceedings of the 2014 ACM/SIGDA International Symposium on Field-programmable Gate Arrays*, FPGA '14, pp. 151-160, New York, NY, USA, 2014. ACM.
- [4] 小林諒平, 吉瀬謙二. Fpga を用いた世界最速のソーティングハードウェアの実現に向けた試み (リコンフィギャラブルシステム). 電子情報通信学会技術研究報告 = IEICE technical report : 信学技報, Vol. 115, No. 109, pp. 65-70, jun 2015.