

学位論文 博士(工学)

大規模組み合わせ回路による
演算器アレイを用いたアクセラレータに関する研究

2014 年度

慶應義塾大学大学院理工学研究科

小崎 信明

主 論 文 要 旨

| | | | | |
|---|-------|---|-----|-------|
| 報告番号 | 甲 乙 第 | 号 | 氏 名 | 小崎 信明 |
| 主 論 文 題 目： 大規模組み合わせ回路による演算器アレイを用いたアクセラレータに関する研究 | | | | |
| <p>近年、スマートフォンやコンパクトデジタルカメラなどモバイル端末の多機能化に伴い、モバイル端末に対してはこれらの処理の高速化とバッテリー駆動時間長期化の両立が求められるようになった。これらの要求を満たすために、低電力で高性能なアクセラレータである粗粒度動的再構成プロセッサ(CGDRP)が注目されている。CGDRPはアレイ状に配置された演算素子(PE)とメモリから構成され、演算素子の接続網や演算内容を毎クロック切り換えることで様々な画像処理アプリケーションを実現する。しかし、動的再構成やPE間の同期をとるためのクロックツリーで多くの電力を消費しており、電力効率に改善の余地がある。</p> <p>そこで、同等の性能をより低消費電力で実現するアーキテクチャとしてCool Meg Array(CMA)を提案する。CMAは大規模な組み合わせ回路によるPEアレイ、マイクロコントローラ、データメモリの3つのモジュールで構成する。PEアレイは再構成する頻度をアプリケーション毎とするが、比較的大規模化することで柔軟性を維持する。また、PEアレイからメモリを排除することでクロックツリーをマイクロコントローラに集約でき、クロックツリーによる電力オーバーヘッドに悩まされずPEアレイを拡張できる。CGDRPではパイプライン段数を増やす際に演算毎の遅延時間の差が性能オーバーヘッドになっていた。CMAでは組み合わせ回路を用いることでこのオーバーヘッドを削減し、効果的に性能を制御できる。</p> <p>本論文では8×8サイズのPEアレイをもつCMA-1を提案し、アプリケーションの実装手法と制御手法の最適化を行う。次にPEアレイ接続網の最適化により面積効率を高める。最後にオフロードエンジン型とコプロセッサ型のCMAを比較し、より電力効率の高いアーキテクチャを明らかにする。CMA-1にアプリケーションを実装する際、一つの入力データに対する演算をPEアレイの列方向に配置し、直結網を優先的に用いることでPEアレイの遅延を最小化できる。また、PEアレイに与えるデータ幅を8bitから24bitに拡張してアプリケーションを実装することでPEアレイ使用率を高めつつコントローラの負荷を下げる。これにより、CMA-1は最大で2.72 GOPS/11.2 mW(247 MOPS/mW)の電力効率を達成した。これは同じプロセスで同じサイズに実装されたCGDRPであるMuCCRA-3の10倍の電力効率で、世界でも最高レベルの電力効率を達成した。</p> <p>また、PEアレイへの供給電圧を下げると電力効率が悪化する課題を解決するために、ウェーブパイプラインを適用する。そのために、PEアレイの遅延時間をモデル化し、アプリケーション構成情報から動作周波数を算出可能にした。ウェーブパイプラインの適用により、PEアレイへの供給電圧が0.45Vの時の電力効率を最大で2.1倍改善し、CMA-1のPEアレイへの供給電圧が0.6V以下の領域ではウェーブパイプラインの適用で電力効率を改善できることを明らかにした。次に、PEアレイ接続網の最適化を行った。PEアレイの行に対して同じ定数を供給する定数パスを設けることで、柔軟性を損なうことなく配線資源を削減し、PEアレイの面積を22%、電力を23%削減した。</p> <p>最後にオフロードエンジン型とコプロセッサ型のCMAを比較し、コプロセッサ型のCMAは面積を24.8%削減し、性能を2.7倍改善でき、システム全体で消費される電力が少ないと判明した。これにより、65 nm CMOSプロセスを用いた小規模なシステムではコプロセッサ型のアクセラレータの方が効率が良いことを明らかにした。</p> | | | | |

SUMMARY OF Ph.D. DISSERTATION

| | | |
|---|---|---|
| School Science for Open and Environmental Systems | Student Identification Number 81246795 | SURNAME, First name OZAKI, Nobuaki |
| Title A Study of Accelerator Using Large-Scale PE Array Consisting of Combinational Circuit. | | |
| Abstract <p>Recently, mobile devices like smart phone have become multifunctional. High performance and prolongation of time driven by a battery are required along with it. Coarse Grained Dynamically Reconfigurable Processors (CGDRPs) have been paid considerably attention as an accelerator. CGDRP has an array of processing elements (PEs) and memory modules. Some CGDRPs provide multiple sets of configuration data called hardware contexts, and execute application programs by switching them in one or a few clock cycles. Since the data flow graph of the target computation can be directly mapped and executed with the numerous PEs in the PE array, the required performance can be achieved. However, CGDRPs still have energy can be saved. One is the frequent dynamically reconfiguration of the PE operations and their interconnection. Another is for storing of intermediate results in PE registers. The large clock tree distributed in the PE array for the dynamic reconfiguration and intermediate storing also consumes a lot of power.</p> <p>Therefore, I proposed a highly power efficient media accelerator architecture called Cool Mega Array or CMA. CMA architecture consists of a large PE array using combinational circuit, micro controller for data management, and the data memory. PE array reconfigures application by application, and maintains flexibility by using large scale array. PE array can be extended without being bothered decrease of energy efficiency because of the clock tree localized to micro controller. CGDRPs are bothered by difference of delay of each PE operations for increasing the pipeline stage, on the other hand, CMA can reduce this overhead and control performance easily by using combinational circuit.</p> <p>In this thesis, I propose CMA-1 which has 8×8 PE Array, and optimize implementation method of application and control method. When implementing an application to CMA, place operations for single data in the vertical direction of the PE array and use direct-link preferentially in order to reduce delay of PE array. Additionally, I have extended the bit width of input data to PE array and reduced the frequency of data fetch in order to reduce workload of micro controller. As a result, CMA-1 has achieved 2.72 GOPS/11.2 mW (247 MOPS/mW). This result is 10 times the power efficiency of MuCCRA-3 implemented in the same area using same 65nm CMOS process, and highest level power efficiency in the world. Wave-pipelining improves power efficiency 2.1 times when the supply voltage to PE array is lowered below 0.6V. Next, I optimize connection network of PE array. 22% area and 23% power reduction were achieved without impairing the flexibility by providing links that share the same constant value for the PEs in the same row in PE array.</p> <p>Then, I compared offload engine style CMA and co-processor style CMA. As a result, co-processor style CMA reduced 24.8% area, and achieved 2.7 times improvement of performance. Therefore, I revealed that co-processor stile accelerator is suitable for small-scale system using 65nm CMOS process in order to enhance power efficiency.</p> | | |