

マルチコア積層システム Cube-2 の実装と評価

松下 悠亮† 小島 拓也† 門本 淳一郎† 黒田 忠広† 天野 英晴†

† 慶應義塾大学理工学部情報工学科

1 はじめに

科学研究費 S「ビルディングブロック型計算システムの研究」はチップ間無線通信用インターフェース誘導結合 ThruChip Interface(TCI) 技術を用いて、レゴブロックのように様々な機能のチップを組み合わせる大規模システムを構築する技術の確立を目指している。ここでは、最初のプロトタイプ Cube-1[1] の実装経験を基に、プロトタイプ 2 号機である Cube-2 を開発した。本稿では、GeysersCM と CC-SOTB2 の実機評価に基づく Cube-2 のクロックレベルのシミュレーション評価を報告する。

2 ビルディングブロック型計算システム Cube-2

Cube-2 は AI 画像認識システムを応用対象とするビルディングブロック型計算システムのプロトタイプである。Cube-2 の特徴は、異なった種類のアクセラレータを接続して一つのシステムとして利用できる点にある。積層の様子を図 1 に示す。

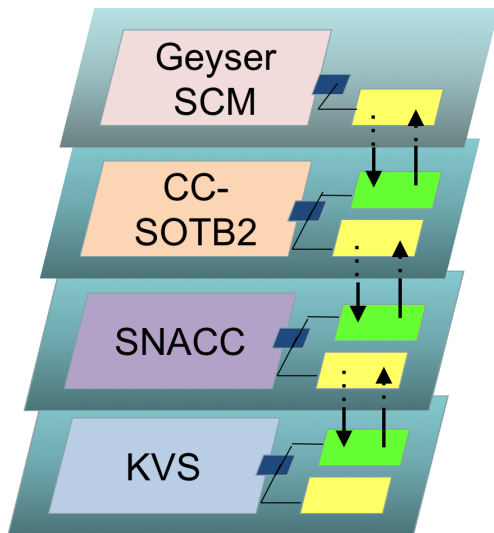


図 1: Cube-2 の概念図

Cube-2 は以下のチップから構成される。各チップは独自の特徴を持っている。

- 省電力用プロセッサ GeysersCM: MIPS R3000 の演算ユニットに細粒度パワーゲーティングの機能を搭載したプロセッサ。Linux OS が動作可能である。

る。低リーク電流を実現するために Standard Cell Memory[2] を搭載している。

- CC-SOTB2: Cool Mega Array(CMA) アーキテクチャを実装した低電力粗粒度再構成可能プロセッサ [3]。画像フィルタなどの処理を低電力で実行できる。
- SNACC: Convolutional Neural Network(CNN) 用アクセラレータ。SIMD 型で専用命令を持つアクセラレータを 4 コア搭載している [4]
- KVS: Key Value Store 型の過去のデータを蓄積及び利活用する近似計算用アクセラレータ [5]
- SMTT: チップ上に複数のチップを積層可能な同期型の共有メモリ [6]。

各チップは三次元方向に積層するために、無線通信用の誘導結合 ThruChip Interface(TCI) を搭載している。Cube-2 では、Geysers がホスト CPU の役割を担い、複数のチップを制御する。ホストで受け取った画像を CC-SOTB2 で前処理し、SNACC で深層学習による画像認識を行い、KVS によりこれをデータベースに格納する操作を各チップで行う。この構成以外にも、目的に応じてアクセラレータの枚数を変化させ、より電力効率・性能の良いシステムを構築することができる。

3 実装

Cube-2 の実装環境および設計仕様をそれぞれ表 1、2 に示す。

表 1: 実装環境

シミュレータ	Cadence NC-Verilog
プロセス	Renesas SOTB 65 nm / LPT-8
ライブラリ	Synopsys Milkyway 2013.12-SP2
論理合成	Synopsys Design Compiler 2013.12-SP2
配置配線	Synopsys IC Compiler 2013.03-SP2

3.1 SOTB プロセス

実装には SOTB(Silicon on thin buried oxide) プロセス [7] を用いた。SOTB は、FD-SOI の一種であり、10nm 程度の極薄酸化膜層上にトランジスタを形成する。ボディバイアス電圧を広い範囲で制御することが

表 2: 実装環境

チップサイズ	6mm × 3mm
TCI	2.5Gb/s 240μm × 240μm
CPU	MIPS R3000
Cache	4KW 2way sep.
TLB	16-entry shared
供給電圧	0.3-1.2V
動作周波数	50-100MHz
チップサイズ	6mm × 3mm

でき、リーク電力と性能のトレードオフを調節することが可能である。これにより、積層するチップ間で生じる特性ばらつきを抑えることが可能となっている。

3.2 TCI

各チップは無線 3 次元積層のためのインターフェースとして TCI(ThruChip Interface)[8] を備えている。TCI の受信部と送信部を合わせて重ねることでチップ間の通信を可能としている。

4 評価

本稿では、4 種類の画像処理用アプリケーション alpha, sepia, sf, gray を使用してシミュレーションによる評価を行った。本稿で比較する対象は GeysersCM1 コアと、GeysersCM と CC-SOTB2 を積層した Cube-2 のプロトタイプである。現状、GeysersCM + CC-SOTB2 の積層チップは積層チップの仕上がり日の都合上届いていないため、GeysersCM と CC-SOTB2 の実チップ評価に基づく、クロックレベルでのシミュレーション評価を行った。各アプリケーションでシミュレーションを行った時の実行クロック数を図 2 に、消費エネルギーを図 3 に示す。

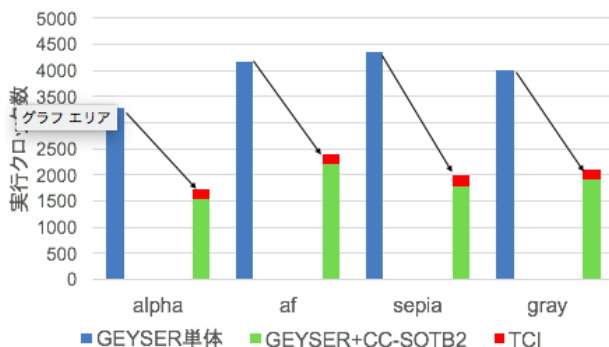


図 2: GEYSER 単体と Cube-2 の実行クロック数

結果、GeysersCM 単体と比較して、GeysersCM + CC-SOTB2 の積層チップでは、平均 48% の実行クロック数削減に成功した。しかし、TCI の電力性能がボトルネックとなり、GeysersCM + CC-SOTB2 の消費エネルギーは GeysersCM 単体と比較して約 2 倍になる

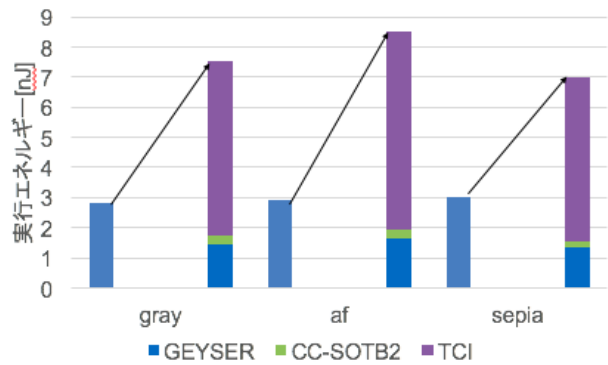


図 3: GEYSER 単体と Cube-2 の消費エネルギー

という結果が生じた。

5 結論

本稿では、チップ間無線接続を利用したビルディングブロック型計算システムのプロトタイプとして、Cube-1 に代わる新たなプロトタイプ Cube-2 の実装と評価を行った。各種チップを積層することで、高速化を可能にした。結果として、平均 48% の実行クロック数が削減されており、性能を向上させることができた。今後は、GeysersCM+CC-SOTB2 の実チップによる評価や他のチップを組み替えてのシミュレーション評価・実チップ評価を行い、電力効率や性能の面から Cube-2 の最適な構成を追求したい。

参考文献

- [1] N. Miura and etal, "A Scalable 3D Heterogeneous Multicore with an Inductive ThruChip Interface," IEEE Micro, Vol.33, No.6, pp.6-15, 2013.
- [2] Y.Yoshida and K.Usami, "Energy Efficient Standard Cell Memory with Optimized Body-Bias Separation in Silicon on Thin BOX(SOTB)," IEICE Trans. on Fundamentals of Electronics, Communications and Computer Sciences, pp.2785-2796, Dec. 2017.
- [3] 増山 隼一朗他, "超低消費電力再構成可能アクセラレータ cc-sotb2 の実装と評価," 信学技報, 第 116 巻, pp.61-66, Nov. 2016 .
- [4] 高田 遼他, "スケーラブルなディープラーニング向けアクセラレータチップの設計と評価," 研究報告システム・アーキテクチャ(ARC), vol.2016-ARC-223, pp.1-6, Nov. 2016 .
- [5] Y.Tokusashi and H. Matsutani, "A multilevel nosql cache design combining in-nic and in-kernel caches," Proc. of HOT Interconnect24, vol.63, pp.60-67, Aug. 2016.
- [6] 寺嶋 爽花他., "ツインタワー用共有メモリチップの開発," 信学技報, 第 117 巻, pp.43-48, Nov. 2017 .
- [7] Y. Morita and et.al., "Smallest Vth variability achieved by intrinsic silicon on thin BOX (SOTB) CMOS with single metal gate," 2008 Symposium on VLSI Technology, pp.166-167, June 2008.
- [8] Y. Take and et.al., "3-d noc with inductive-coupling links for building-block sips," IEEE Transactions on Computers(TC), vol.63, pp.748-763, Mar. 2014.