

# 3次元積層技術による次世代ベクトルキャッシュの設計と評価

船矢祐介<sup>1</sup>, 永岡龍一<sup>1</sup>, 江川隆輔<sup>2</sup>, 滝沢寛之<sup>1</sup>, 小林広明<sup>2</sup>

<sup>1</sup> 東北大学大学院情報科学研究科, <sup>2</sup> 東北大学サイバーサイエンスセンター

{funaya, nagaoka}@sc.isc.tohoku.ac.jp, {egawa, tacky, koba}@isc.tohoku.ac.jp

概要: 将来のベクトルプロセッサが抱える演算性能に対するメモリバンド幅低下の問題に対し, ベクトルキャッシュ機構が提案されている. しかし従来の半導体加工技術の下では, チップ面積の制約から小規模なキャッシュメモリしか搭載できないという問題がある. そこで本論文では, 3次元積層技術を用いてこれまでの制約から脱却し, 次世代ベクトルプロセッサのための大容量キャッシュメモリを設計し, 評価を行う.

## 1 はじめに

ベクトルプロセッサは, 大規模, かつ高精度なシミュレーションが必要な科学技術分野において広く用いられている. しかし近年, チップ面積やI/Oピン数の制限などから高性能化, 低消費電力化が困難となってきている. そこで, ベクトルプロセッサが現在抱えているこれらの問題を解決する手法として, 本論文では3次元積層技術に着目する. 3次元積層技術とは, 複数のシリコンレイヤを垂直方向に積層し, レイヤ間を非常に低レイテンシな配線によって接続する技術である. 3次元積層技術を用いることで, デバイスの集積密度, 配線の自由度が向上し, 新たな3次元回路の実現によりベクトルプロセッサの抱える様々な問題点の解決が期待できる. 本論文では, 特に将来のベクトルプロセッサが抱える演算性能に対する相対的なメモリバンド幅 (Bytes/Flop, B/F) 低下の問題に対し, 3次元積層技術を用いたベクトルキャッシュを設計し, それを搭載したベクトルプロセッサの評価を行う.

## 2 3次元積層型ベクトルキャッシュの設計

3次元積層技術では, レイヤ間を接続する Through Silicon Vias (TSV) はチップ上の配線に比べて非常に短く RC 抵抗も小さい. そのため, これまでチップ上で2次元方向に結合されていた機能ブロック同士を TSV を用いて垂直方向に配置・配線することで, 配線長が大幅に短縮され, レイテンシ, および消費エネルギーの減少が期待できる. また, 近年では, TSV の実装密度が非常に向上し, 積層したレイヤ間を高いバンド幅で接続することが可能となってきている. さらに, チップ上に複数のレイヤを積層することで, 1チップに搭載可能なリソース量の増大も期待できる. これらの特徴を生かし, 3次元積層技術をベクトルアーキテクチャに適用する方法はいくつか考えられるが, B/F が低下しつつあるベクトルプロセッサにおいては, まずメモリバンド幅の増強に取り組むことが重要である. 現在, ベクトルプロセッサのメモリバンド幅を補う手法としてベクトルキャッシュ機構 [1] が提案され

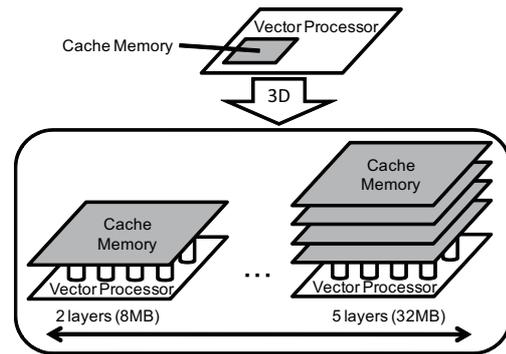


図1 キャッシュメモリ積層型ベクトルプロセッサ

ている. キャッシュに再利用性の高いデータを保持し演算器に供給することで, メインメモリと併せて高いデータ供給能力を実現できる. さらに, メインメモリへのアクセス数減少に伴う消費エネルギー削減の効果も期待できる. しかし, 現在の2次元設計ではチップ面積の制約によって小規模なキャッシュしか搭載できないという問題がある.

そこで本論文では, 3次元積層技術を用いてキャッシュのみで構成されるシリコンレイヤ(メモリレイヤ)を積層することで, ベクトルキャッシュの大容量化を実現する. また, メモリレイヤとプロセッサ間は TSV で接続し, 高いメモリバンド幅を実現する. プロセッサの構成を図1に示す. この構成では, 積層するメモリレイヤを増やすことによってキャッシュ容量を増やすことができる. さらに, キャッシュ容量増加に伴うレイテンシと消費エネルギー増加の問題に対し, 本論文ではキャッシュの3次元分割手法 [2] を適用する. この手法は, 3次元積層技術を用いてキャッシュを分割・積層する手法であり, キャッシュ内部の配線の一部を TSV で置き換えることにより長配線を削減し, キャッシュアクセス速度の向上と消費エネルギーの削減が期待できる. 3次元分割手法を用いたキャッシュ構成を図2に示す. ベクトルプロセッサでは, メインメモリを多数のバンクに分割し, データをインターリーブ転送することで高いメモリバンド幅を実現していることから, その利点を生かすためにキャッシュも多数のサブ

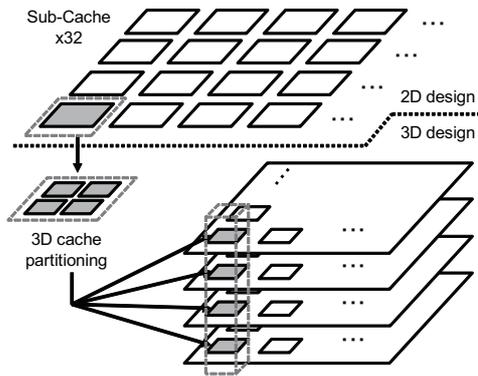


図2 3次元ベクトルキャッシュ構成

キャッシュで構成されている．このため本論文では，各サブキャッシュに3次元分割手法を適用し複数のレイヤに積層する．

### 3 性能評価

はじめに，3次元ベクトルキャッシュを搭載したベクトルプロセッサの性能評価を行う．提案するプロセッサはNEC SX-8を基本とし，ベクトルキャッシュ機構を有するベクトルプロセッサ [1] を基に設計を行う．メインメモリのメモリバンド幅は，現在のベクトルプロセッサと同等の 2B/F と，将来のベクトルプロセッサを想定した 1B/F で評価を行う．これに対し，キャッシュとプロセッサ間はTSVによって接続されているために高いメモリバンド幅が実現可能である．そこで本評価ではキャッシュのメモリバンド幅を 4B/F とする．CACTI 3.2 の試算によると，SX-8 のチップ面積約 392mm<sup>2</sup> に搭載可能なキャッシュ容量は1層当たり約 8MB である．そこで本評価では，演算器の層と合わせてプロセッサが2層（キャッシュ容量 8MB）から5層（32MB）の場合を評価する．ベンチマークには3次元FDTD法による地雷探査シミュレーションプログラム（Land Mine） [3] を使い，シミュレータにはSXシミュレータを用いた．

実効性能の評価結果を図3に，消費エネルギーの評価結果を図4に示す．それぞれの結果を2B/Fでキャッシュなしの場合の結果で正規化して示す．図3より，キャッシュ容量が8MBの場合に大きく性能向上し，16MB以上では性能向上が緩やかになることが分かる．これは，Land Mineのデータアクセスの局所性が比較的低いためである．しかし，それでもキャッシュ容量32MBの場合には実効性能が1B/Fで約10%，2B/Fで約19%向上することが分かった．また，図4より消費エネルギーはキャッシュ容量32MBの場合に1B/Fで約22%，2B/Fで約16%削減できることが分かった．以上より，設計した3次元ベクトルキャッシュにより，ベクトルプロセッサの高性能化・低消費電力化が実現可能であることが明らかになった．

次に，3次元キャッシュ分割によるレイテンシと消

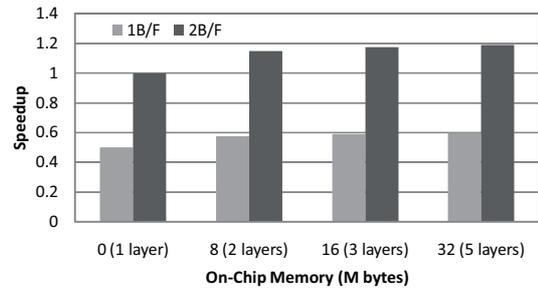


図3 実効性能 (Land Mine)

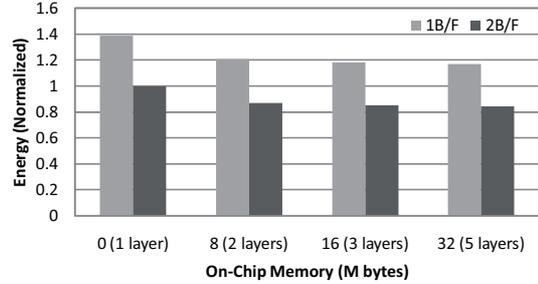


図4 消費エネルギー (Land Mine)

費エネルギーの評価を行う．実験には3DCACTI [2] を用いた．提案設計の評価の結果，3次元分割手法を用いない場合と比較し，メモリ容量の増加に伴うレイテンシの増大が32MBの場合で約44%抑えられることが分かった．また消費エネルギーの増大も，3次元分割手法を用いない場合と比較し，32MBの場合で約14%抑えられることが分かった．これにより，将来のベクトルプロセッサにおいて3次元積層技術を積極的に活用することで，レイテンシや消費エネルギーの増大を低く抑えつつキャッシュの大容量化が実現可能であることが明らかになった．

### 4 まとめ

本論文では，将来のベクトルプロセッサが抱えるB/F低下の問題に対し，3次元積層技術を用いて次世代ベクトルキャッシュの設計を行った．実アプリケーションを用いた評価の結果，提案キャッシュによるベクトルプロセッサの実効性能向上とメモリシステムの消費エネルギー削減の効果が明らかになった．さらに，3次元積層技術を積極的に用いることで，レイテンシ及び消費エネルギーの増大を最小限に抑えつつキャッシュの大容量化が実現可能であることが明らかになった．

#### 参考文献

- [1] A. Musa et al. An On-Chip Cache Design for Vector Processors. In *Proc. of the 7th MEDEA workshop*, pp. 17–24, 2007.
- [2] Y. Tsai et al. Design Space Exploration for 3-D Cache. *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, Vol. 16, No. 4, pp. 444–455, 2008.
- [3] T. Kobayashi et al. FDTD Simulation on Array Antenna SAR-GPR for Land Mine Detection. In *Proc. of SSR2003*, pp. 279–283, 2003.