

CMP におけるキャッシュメモリ階層構成の違いによる電力性能評価

安西 由行[†] 大津 金光[†] 横田 隆史[†] 馬場 敬信[†][†]宇都宮大学大学院工学研究科情報工学専攻

1 はじめに

近年の計算機では CMP (Chip Mult Processor) が主流となっている。従来のシングルプロセッサでは性能向上を動作周波数の向上により達成してきたが、CMP では並列処理を行うことで性能向上を行っている。そのため、CMP はシングルプロセッサよりも性能あたりの消費電力効率に優れているアーキテクチャである。また、トランジスタの微細化によりプロセッサ内に搭載できるキャッシュメモリが大容量となっている。そして、キャッシュメモリの階層構成はプロセッサの進化と共に変化してきた。CMP のキャッシュメモリ階層構成には、各々のプロセッサコアがキャッシュメモリを持つ非共有型と、プロセッサコア間でキャッシュメモリを共有する共有型とがある。CMP でのキャッシュメモリの共有型と非共有型を比較しプロセッサ実行速度の性能評価を行う研究 [1] はあるが、電力性能について言及している研究はない。しかしながら、現在のプロセッサは高性能かつ省電力なものが求められており、電力性能も重要である。なぜならば、電力を消費すると熱が発生し、冷却のためにさらに電力を消費する悪循環に陥る。また、発熱により故障も発生する可能性もあり、プロセッサにおいて省電力化は重要である。

そこで本研究では、CMP のキャッシュメモリにおけるプロセッサコア間の共有・非共有の違いによる比較検証を行う。評価はシミュレータ SESC 上でキャッシュメモリ容量、キャッシュメモリ階層構成を変え、並列化を施したアプリケーションを用いて行う。比較には実行速度と消費電力量の両方を考慮した EDP (Energy Delay Product) を用い、電力性能の良いキャッシュメモリ階層構成の検討を行う。

2 CMP におけるキャッシュメモリ階層構成

本研究ではプロセッサコアが各々キャッシュメモリを持つ非共有型とプロセッサコア間で共有する共有型について検討を行う。この構成により各プロセッサコアが使用可能なキャッシュメモリ容量とプロセッサコア間のデータの受け渡しによる電力性能に与える影響を検討する。2 次キャッシュで共有・非共有の比較と 3 次キャッシュで共有・非共有の比較を行う。どの階層構成でも 1 次キャッシュメモリとして小容量かつ高速であるデータキャッシュ・命令キャッシュを持つものとする。また、3 次キャッシュにて共有・非共有の比較を行う場合には 2 次キャッシュを各プロセッサコアが持つものとする。

非共有型と共有型の特徴を述べる。非共有型は図 1 のように各プロセッサコアがキャッシュメモリを持つ階層構成である。この階層構成のメリットは 1 つのプロセッサコアあたりのキャッシュメモリ容量が (共有型と合計値が同じ場合) 小さいためキャッシュアクセスによるレイテンシや消費電力は低くなる。デメリットは

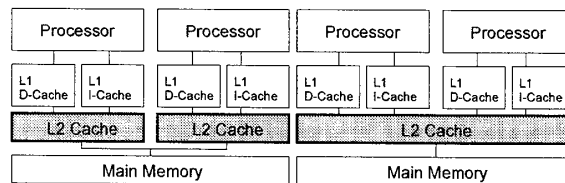


図 1: 非共有型

図 2: 共有型

プロセッサコア間のデータの受け渡しに主記憶を介することによる性能低下、1 つのプロセッサコアが使用できるキャッシュメモリ容量が (共有型と比べ) 少ない点が挙げられる。

共有型は図 2 のように各プロセッサコア間でキャッシュメモリを共有する階層構成である。この階層構成のメリットは 1 つのプロセッサコアが最大で利用できるキャッシュメモリ容量が大きい、プロセッサコア間でデータの受け渡しが高速なキャッシュメモリ内で行うことができる点である。デメリットとしてはキャッシュメモリ容量が大きいためキャッシュアクセスによるレイテンシや消費電力が増加する点が挙げられる。

これらの階層構成の違いによる主な影響として使用可能なキャッシュメモリ容量、プロセッサコア間のデータ受け渡しにかかる時間、キャッシュアクセスによるレイテンシ・消費電力量の違う点が挙げられる。そのため、低消費電力であるがプロセッサコア間のデータ受け渡しが低速な非共有型と、データ受け渡しは高速であるが消費電力の高い共有型の電力性能の比較を行う。

3 評価

CMP におけるキャッシュメモリ階層構成を変更して評価を行うことで階層構成の違いによる影響を調べ、電力性能の良いキャッシュメモリ階層構成を検討する。

評価指標には EDP を用いる。EDP は実行サイクル数と消費電力量の積であり、EDP の値は低い方が電力性能が高い。

3.1 評価環境

評価にはマイクロプロセッサシミュレータ SESC[2] を用いた。SESC はイリノイ大学が開発したシミュレータであり、CMP における電力と性能評価を行うことができる。

キャッシュメモリ階層構成ごとに用いるキャッシュメモリ容量の合計値は同じ値で行う。例えば、プロセッサコア数が 4 でキャッシュメモリ容量が 8MB の場合、非共有型は 2MB のキャッシュメモリを各プロセッサコアが個別に持つ。共有型は 8MB のキャッシュメモリを全てのプロセッサコアで共有する。

キャッシュメモリのアクセスレイテンシ見積もりには CACTI (ver 5.0)[3] を用いた。評価パラメータを以下の表 1 に示す。キャッシュメモリアクセスレイテンシは表 2 に示す。アクセスレイテンシは動作周波数が 3.0 GHz より 1 cycle = 0.33 nsec として算出した。プロセッサコア数 8、キャッシュメモリ容量 8、16、32MB にて評価を行う。なお、3 次キャッシュメモリにて共有・非共有の比較を行う場合に用いる 2 次キャッシュメモリ容量は 256KB で行う。

Power performance evaluation of cache memory organizations in CMPs

[†]Yoshiyuki Anzai, Kanemitsu Ootsu, Takashi Yokota, Takano Baba

Department of Information Science, Graduate school of Engineering, Utsunomiya University (†)

表 1: 評価パラメータ

動作周波数	3.0 GHz
テクノロジーサイズ	45 nm プロセス
データキャッシュ	32KB ブロックサイズ 64B
命令キャッシュ	4-way レイテンシ 2 cycle
(1次キャッシュ)	置換法 LRU Write Back
2次・3次	ブロックサイズ 64B
キャッシュ	4-way(1MB 未満)
(共有・非共有型	16-way(1MB 以上)
の場合)	置換法 LRU Write Back
メインメモリ	レイテンシ 300 cycle (100ns)

表 2: キャッシュメモリのアクセスレイテンシ

size(B)	256K	512K	1M	2M	4M	8M	16M	32M
latency	3	4	6	7	9	12	17	23

(cycles)

3.2 評価アプリケーション

評価はベンチマークアプリケーションである MediaBench の mpeg2decode、複数のアプリケーションを組み合わせたプログラムにて行う。

mpeg2decode には並列化を施して評価を行う。並列化は実行頻度の高い関数 Reference_IDCT に対して行う。このアプリケーションではプロセッサコア間のデータ受け渡しにより電力性能に与える影響を検証する。この関数は3重ループが2つ含まれており、ループ反復回数は全て8回である。並列化はこの3重ループの1番外側のループに対して処理分割を行う。最も内側のループ内で処理を行い、この処理で用いる変数はプロセッサコア間で共有する。3重ループと3重ループの間にはデータの整合性を保つためバリア同期を行う。入力データは画像サイズ 1920 × 1080 pixel、10フレームの動画を用いる。

複数のアプリケーションを組み合わせた評価アプリケーションには SPEC ベンチマークアプリケーションから go、mcf、twolf、art、mesa を MediaBench から cjpeg、djpeg、mpeg2decode の8つのアプリケーションを用いる。このアプリケーションでは使用可能なキャッシュメモリ容量の影響とプロセッサコア間でデータの受け渡しを行わない場合について検証する。入力データには SPEC アプリケーションは train データセット、cjpeg と djpeg は 4000 × 3000 pixel の画像、mpeg2decode は上記で用いた動画を用いる。

コンパイルには SESC 用 gcc コンパイラ (ver 3.4.4) を用い、最適化オプションは -O3 を使用する。

3.3 評価結果

図3に mpeg2decode、図4に複数のアプリケーションの EDP 結果を示す。図では縦軸に EDP、横軸にプロセッサ構成、棒軸の色にてキャッシュメモリ容量の違いを表している。プロセッサ構成は非共有型・共有型を L2 ならば2次キャッシュにて L3 ならば3次キャッシュメモリにて比較を行うことを表す。

図3の結果では、非共有型より共有型の EDP が低く良い結果である。これはプロセッサコア間のデータ受け渡しに非共有型はメインメモリを介すためであり、これは非共有型のリードミス回数が共有型よりも大幅に多いことからわかる。また、2次キャッシュと3次キャッシュの共有型を比較すると3次キャッシュの構成が良い結果を示している。サイクル数はほぼ同じであるが、消費電力量は3次キャッシュ共有型が2次キャッシュ共有型より少ない結果になったためである。

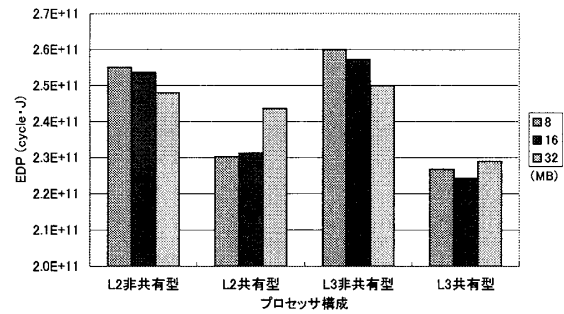


図 3: mpeg2decode: EDP

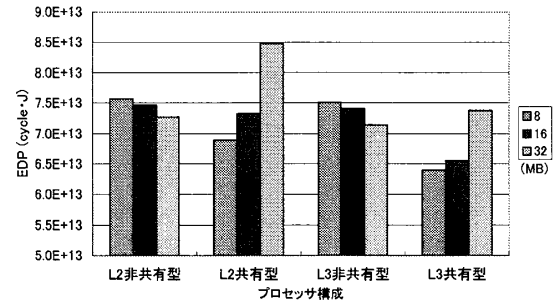


図 4: 複数のアプリケーション組み合わせ: EDP

図4の結果では、共有・非共有の違いにおいてキャッシュメモリ容量の違いで良い構成が異なる。これは共有型のキャッシュメモリ容量が大きくなると消費電力量も増加するためである。2次キャッシュと3次キャッシュの構成を比較すると非共有型には差が見られないが、共有型では3次キャッシュの構成が2次キャッシュの構成より良い結果を示している。これは3次キャッシュの構成には256KBの2次キャッシュがあることで大容量であるが消費電力が大きい3次キャッシュへのアクセス数が減ったためである。

これらの結果より3次キャッシュによる共有型が電力性能が良いといえる。

4 おわりに

本研究では CMP のキャッシュメモリにおけるプロセッサコア間の共有・非共有の違いによる比較検証を行い、電力性能の良いキャッシュメモリ階層構成の検討を行った。評価を行った結果、3次キャッシュによる共有型が電力性能が良いと結論付けた。今後の課題として他の幅広い分野での現実的なアプリケーションによる評価と他のキャッシュメモリ階層構成による評価が挙げられる。

謝辞

本研究は、一部日本学術振興会科学研究費補助金（基盤研究 (B)18300014、同 (C)19500037、同 (C)20500047）および宇都宮大学重点推進研究プロジェクトの援助による。

参考文献

- [1] L. Hsu, R. Iyer, S. Makineni, S. Reinhardt, D. Newell, "Exploring the cache design space for large scale CMPs", ACM SIGARCH Computer Architecture News, pp.24-33, 2005
- [2] J. Renau, B. Fraguera, J. Tuck, W. Liu, M. Prvulovic, "SESC: SuperEScalar Simulator", <http://sesc.sourceforge.net>, Jan. 2005
- [3] S. Thoziyoor, N. Muralimanohar, NP. Jouppi, "CACTI 5.0", <http://www.hpl.hp.com/techreports/2007/HPL-2007-167.pdf>, Oct. 2007