

平成 30 年 5 月 31 日現在

機関番号：11301

研究種目：研究活動スタート支援

研究期間：2016～2017

課題番号：16H06613

研究課題名(和文) マニーチャネルメモリ時代に向けた高性能・低消費電力メモリ管理手法の創出

研究課題名(英文) High-performance and Low-Power Memory Management for the Era of Many-Channel Memories

研究代表者

佐藤 雅之 (Sato, Masayuki)

東北大学・情報科学研究科・助教

研究者番号：50781308

交付決定額(研究期間全体)：(直接経費) 2,300,000円

研究成果の概要(和文)：マニーチャネルメモリの高性能・低消費電力化を目的とし、メモリの低消費電力モードへの遷移を促進するためのアドレスマッピング手法と、その切り替えを実現するためのメモリコントローラ機構を提案した。アドレスマッピング手法は物理アドレスからチャネルアドレスを取得する際のビット数を変更することにより、アクセスされるチャネル数を限定し、性能と消費電力のトレードオフを選択可能とした。メモリコントローラ機構は、現実的なコストでアドレスマッピング手法の切り替えを管理し、アプリケーション実行中の動的に切り替え時にマッピングの一貫性を維持するデータマイグレーションを可能とした。

研究成果の概要(英文)：This project proposes the high-performance and low-power management mechanism for many-channel memories, which is realized by the address-mapping schemes that limit the number of accessed channels, and the memory controller that can switch the address-mapping schemes. The proposed address-mapping schemes change where the channel address is taken from bits of the physical address. These schemes successfully limit the number of channels accessed by executing applications, and the performance/power trade-off can be selected by switching these schemes. The proposed memory controller can dynamically switch the address-mapping schemes. It also realizes the migration of the application data to maintain the consistency of the data placement when switching the address-mapping schemes. The evaluation results show that the memory controller can be realized by the reasonable costs of hardware and data migration.

研究分野：計算機アーキテクチャ

キーワード：メモリシステム

1. 研究開始当初の背景

計算機におけるメモリシステムの性能向上は、マイクロプロセッサのメニーコア化に伴う性能向上と比較して非常に遅く、長らく計算機システムにおけるボトルネックとなってきた。この状況を打破するため、3次元積層技術に基づく High Bandwidth Memory (HBM)をはじめとした多数のメモリチャンネルを搭載するメモリ（以下、メニーチャンネルメモリ）が実用化され、次世代計算機のためのメモリシステムとして注目されている。メニーチャンネルメモリは、多数の伝送路であるチャンネルとそこに接続されるメモリバンクを搭載しており、アプリケーションを実行する際には、アプリケーションデータを全チャンネルに接続されたアレイに分散して保存しつつ、全てのチャンネルを同時に用いてデータを転送する。これにより、マイクロプロセッサへのデータ供給において高いバンド幅を実現することが可能となり、次世代アプリケーションにおいて、高い実行時性能を得られることが期待されている。

一方で、高い性能を必要としないアプリケーションではメモリシステムの電力性能比が低下する恐れがある。一般的にメモリは、アクセスをしばらく受けない場合に、アクセスを受け付けられない消費電力の低い状態（低消費電力モード）に遷移することが可能である。メニーチャンネルメモリでは多数のチャンネルを通じて多くのメモリアレイにアクセスを行い、データを転送する。このような状況化において、高い性能を維持するためにバンド幅を必要としないアプリケーションを実行すると、高いデータ転送性能が必要とされない一方で、多数のメモリチャンネルへアクセスが断続的に発生し、チャンネルに接続されているメモリバンクが低消費電力モードに遷移できない。このため、バンド幅を必要としないアプリケーションでは、メニーチャンネルメモリにおいて電力性能比がかえって低下する恐れがある。このことから、メニーチャンネルメモリにおいてアプリケーションが必要とするバンド幅を実現しつつ、性能電力比を向上させることが求められている。

2. 研究の目的

本研究の目的は、メニーチャンネルメモリに基づき高い電力効率を達成可能なメモリアーキテクチャを実現することである。そこで、アクセスがなければチャンネルとそのチャンネルに接続されているメモリバンクの低消費電力化が可能であることに着目し、アプリケーション性能の維持に必要なバンド幅を維持しつつ電力削減を可能とするデータ管理手法とその設計指針を明らかにする。

3. 研究の方法

メニーチャンネルメモリのチャンネル数は従来と比較して飛躍的に増加する。このため、チャンネルとそこに接続されるメモリバンクの電力制御を細粒度で行うことが可能になる。一方で、多数のメモリチャンネルへアクセスが断続的に発生するため、チャンネルに接続されているメモリバンクが低消費電力モードに遷移できない。そこで、メニーチャンネルメモリにおけるチャンネルとそこに接続されるメモリバンクを電力制御可能なユニット単位（電力ドメイン）と見なし、アプリケーションの実行時性能の維持に必要なメモリバンド幅の実現に十分な電力ドメインのみを有効化しつつ、不要なドメインを待機電力モードに移行させる手法の実現を目指した。具体的には以下のような取り組みを行った。

(1) アドレスマッピング方式に基づくアクセスチャンネル数制御方式の検討

本取り組みでは電力ドメインに対するデータアクセスパターンを制御するためにアドレスマッピングに着目した。アドレスマッピングはデータの物理アドレスに基づき、メモリモジュールの物理的な位置へデータを保存するか決定する。これを利用し、指定した数のチャンネルのみにアクセスを制限できるアドレスマッピング方式について検討を行った。

(2) アドレスマッピング方式切り替えにおけるデータマイグレーション手法の検討

(1)に示したアドレスマッピング方式でアクセスされるチャンネル数を指定する事ができた場合、アプリケーションやその実行状況に合わせてチャンネル数を制御するためには、アドレスマッピング方式の切り替えが必要である。アドレスマッピング方式を切り替え可能とする場合、アドレスマッピング方式の変更に伴い、実際にチャンネルに割り当てられているデータのマイグレーションが必要である。

本取り組みではアドレスマッピング方式の切り替えを管理するための機構を検討した。この機構では、実行されているアプリケーションのアクセス挙動履歴に基づいて各メモリ領域へのアクセス時に必要なメモリバンド幅を予測し、各メモリ領域に最適なチャンネル数のアドレスマッピング方式を判断する。また、アドレスマッピング方式の切り替え時はデータ配置に齟齬が発生しないようデータを移動・管理する必要がある。このような機構について、その実現に必要な要件を明らかにし、その消費電力削減効果と管理オーバーヘッドのトレードオフを明らかにした。

4. 研究成果

(1) 問題点の明確化

本研究で対象とする問題点として、メモリーチャンネルメモリを含むシステムで高いバンド幅を必要としないアプリケーションが実行された場合、各チャンネルが断続的にアクセスを受けるため低消費電力モードに遷移する事が困難である点が挙げられる。

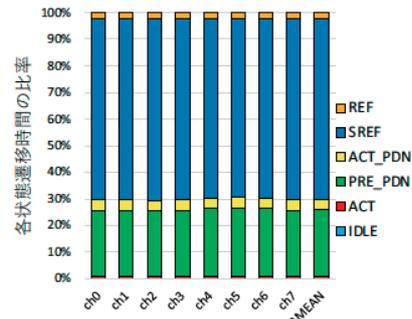
本問題点に関する検証実験を行った。実験結果を図1に示す。図1は異なる2つのアプリケーション(a)namdと(b)gobmkにおいてメモリーチャンネルメモリの各チャンネルがどの状態に遷移していたかを表す。横軸は各チャンネルと全チャンネルの平均を表し、縦軸は実行時間中に各チャンネルが遷移した電力モードの時間割合を示す。図1に挙げた2つのアプリケーションはどちらも高いバンド幅を必要とせず、実効バンド幅も低いアプリケーションである。しかし、namdは最も深い低消費電力(SREF)モードに長く遷移する事ができた一方で、gobmkにおいてSREFモードに遷移していた時間はわずかである。以上から、高いバンド幅を必要としないアプリケーションであっても、低消費電力モードに遷移する事ができない場合があることを明らかとなった。

(2) チャンネル数を制御するアドレスマッピング手法の提案

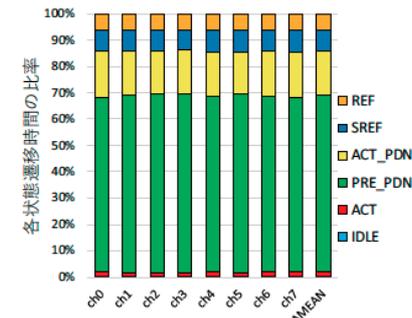
図2にチャンネル数を制御するアドレスマッピング手法の概要を示す。本手法ではアクセスされるチャンネル数を限定するために、アドレスマッピングを利用してメモリ内への適切なデータ配置を実現する方法を用いる。本手法ではアクセスチャンネル数を調節するために、チャンネルアドレスを以下に示す算出方法に変更する。まず、チャンネルアドレスの上位nビットを物理アドレスの最上位nビットから得る。次に、チャンネルアドレスの下位mビットを物理アドレスの第11ビット目から第(11+m)ビット目までで得られるビット列と、第20ビット目から第(20+m)ビット目までで得られるビット列の排他的論理和より得る。

一般に、データの参照局所性に基づき、物理アドレスの下位ビットは上位ビットと比較して値が変わりやすい。このため本手法のようにチャンネルアドレスを算出すると、その下位mビットは上位nビットよりも値が変化しやすい。結果として、メモリアクセスは2mチャンネルに集中する傾向を示す。この時、mとnの比を変更することにより、アクセスチャンネル数を調節可能である。ここで、 $2(n+m)$ はメモリの持つチャンネルの数と等しい。

本手法によりアクセスチャンネル数を減らした時の効果を図3に示す。図3はアクセスチャンネル数を変更したときのエネルギー遅延積とアプリケーションの実効バンド幅を示す。折れ線グラフはエネルギー遅延積を示し、8チャ



(a) namd



(b) gobmk

図1：各チャンネルにおける電力モード遷移時間の比率

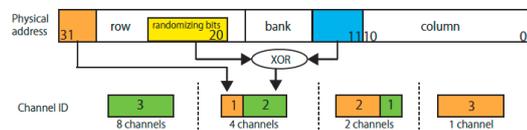


図2：チャンネル数を制御するアドレスマッピング手法の概要

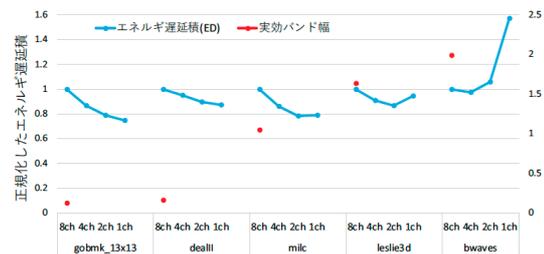


図3：チャンネル数とエネルギー遅延積の関係

ネルの場合で正規化されている。赤い点は8チャンネルの場合のアプリケーションの実効バンド幅を示している。図3より、実効バンド幅が小さい gobmk や dealII といったアプリケーションでは、アクセスチャンネル数を減らすほどエネルギー遅延積が減少している。これらのアプリケーションは演算律速であり、高いバンド幅を必要としない。そのため、アクセスチャンネル数を削減しても性能に影響がな

かった。したがって、実効バンド幅が小さいアプリケーションでは、本提案により、アクセスチャネル数を減らすことで、性能を低下させずにエネルギー効率を改善できることが明らかとなった。

一方で、実効バンド幅が高いアプリケーションである bwaves では 1 チャネル使用時には 57.3%, 2 チャネル使用時には 5.7% のエネルギー遅延積の増加がみられた。また, milc と leslie3d において, 2 チャネルまではエネルギー遅延積が減少しているが, 1 チャネル使用時には 2 チャネル使用時よりも増加している。これらのアプリケーションはメモリ律速であり, アクセスチャネル数が減ると性能維持のために必要なデータ転送性能が得られない。そのため, 実行時間が増加し, 消費エネルギーも増加した。

(3) アドレスマッピング切り替えと動的データマイグレーションを可能とするメモリコントローラ機構

本機構は図 4 に示すように, モードテーブルとアクセスカウンタ, コントローラで構成される。モードテーブルは物理アドレス毎のアクセスモード (メモリ上でどのアクセスマッピング手法で管理されているかを示す) を保存するテーブルである。アクセスカウンタは指定された間隔毎の読み込みアクセス回数を計数する。また, コントローラは現在のアクセスモードを保存しておくモードレジスタを備えている。そして, アクセスモードの更新のタイミングでは, アクセスカウンタから読み込みアクセス回数を参照し, モードレジスタを適切なアクセスモードに更新する。コントローラは LLC から送られてきた要求やメモリからの応答に応じて適切な処理を行う。

本機構にアクセスが発生した際は, まずモードレジスタから現在のアクセスモードを取得する。次に, 過去にその物理アドレスにアクセスが行われたときのアクセスモードをモードテーブルから取得する。現在のアクセスモードと過去にアクセスされたときのアクセスモードが異なる場合は, データを現在のアクセスモードに対応する位置に移動するスワップを行う必要がある。スワップが必要な場合, そのアクセスの処理を待機し, アクセスされたデータを含むローと移動先のローを読み込む要求をメモリに送る。そして, 両方のローに保存されたデータが機構に読み込まれた後, 各データを該当するローへ書き込む要求をメモリに送る。書き込み終了後, モードテーブルの該当する物理アドレスの要素を現在のアクセスモードで更新する。そして, 待機しているアクセスの処理を再開する。

メインメモリからアクセスに対する応答が返ってきた際は, 提案機構は適切な処理を行いプロセッサに送信する。まず, モードテーブルからアクセスを行った時のアクセスモードを取得する。そして, アクセスモードが 8

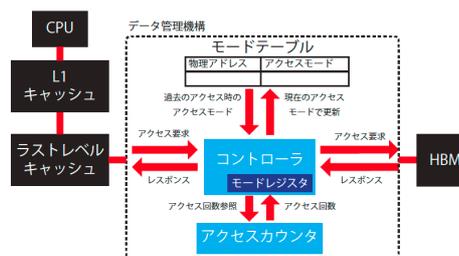


図 4: 提案したメモリコントローラ

チャンネルモード以外の場合は, 変更されていた物理アドレスを元に戻す。最後にメモリからの応答を LLC へ送信する。

ここまで述べてきた様に, 本機構では物理アドレスの変換と復元を機構内で完結させている。プロセッサは物理アドレスがアクセスの途中で変換されていることを認識せずに処理を行うことが可能である。

以下に本機構の実現に必要なハードウェアコストについて述べる。例として, 本機構がチャンネル数を 8 と 2 で切り替えられる場合を想定する。本機構ではアクセスが行われたときにのみスワップを行い, アクセスされたデータが配置されている 2 チャネル分のロー内のデータ (4kB) をまとめてスワップする。これは, アクセスモードを更新するたびにすべてのデータをスワップすると, チャンネル切り替え時のデータマイグレーションのための追加のメモリアクセス回数が膨大になるためである。さらに, アクセスの最小単位のキャッシュラインサイズ (64B) ではなく, ロー単位でまとめてスワップすることにより, スワップ回数を削減する。

このとき, 2 つのチャンネルに対しそれぞれ 2kB のアクセスが発生するが, 各メモリコントローラにより 2kB のアクセスは 32 回の 64B アクセスとして処理される。アクセス単位はロー単位となるため, スワップ時には各ローで初回の 32B アクセス以外がローバッファヒットとなり, アクセスにかかるレイテンシを短縮できる。したがって, スワップ 1 回あたりのオーバーヘッドを減らすことが可能である。また, 2 チャネル分まとめてスワップを行うことにより, 2kB のデータを 2 つ並列にスワップ可能である。そのため, メモリバンド幅を有効に活用することができ, スワップのオーバーヘッドを削減可能である。

以上のようにロー単位でスワップするため, モードテーブルにはデータがどちらのマッピング手法でアクセスされたかを 2 チャネル分のロー (4kB) 単位で保存する。メモリ容量が 4GB の場合, 物理アドレスは 32 ビットで表現可能である。したがって, 上位 20 ビットをキーとして 8 チャンネルモードの場合は 0, 2 チャネルモードの場合は 1 をテーブルに保存する。そのため, モードテーブルには 1 ビットのントリを約 79 万個格納することが必要であり, テーブルオーバーヘッドを容量で換算すると約 98.3kB である。このオーバーヘッドは現在一

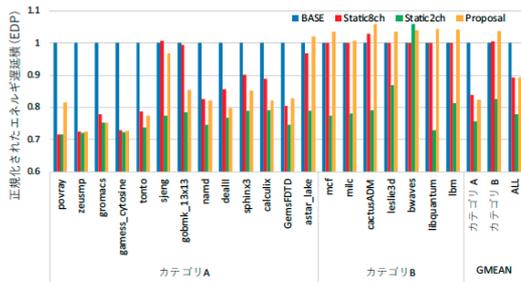


図 5：提案機構によるエネルギー遅延積

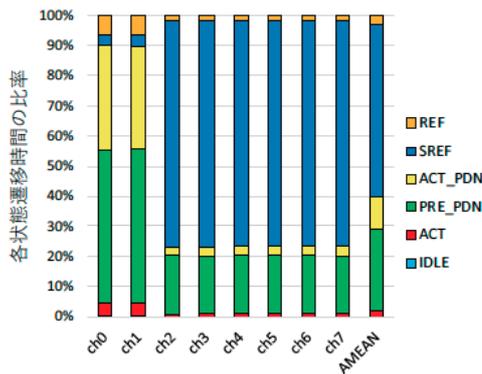


図 6：提案機構による各チャンネルにおける電力モード遷移時間（アプリケーション gobmk）

一般的に用いられているキャッシュ容量よりも小さく、影響は小さいと考えられる。

以上の機構について評価を行った。本機構を用いた場合のエネルギー遅延積の評価を図 5 に示す。縦軸は低消費電力モードを用いない場合をベースラインの結果として正規化したエネルギー遅延積を示す。各棒グラフはベースライン、8ch モードの場合、2ch モードの場合、提案機構によって 8ch と 2ch を動的に切り替えた場合である。図 5 より提案手法では平均 10.8%のエネルギー遅延積を削減している。特に、実効バンド幅が小さいカテゴリ A のアプリケーションでは、平均 17.7%のエネルギー遅延積を削減している。

また、8ch モードの場合と比較すると、カテゴリ A のアプリケーションにおいて平均 1.8%削減している。このことは、実効バンド幅が小さいアプリケーションにおいて、提案機構がセルフリフレッシュモードをより活用可能にすることで、エネルギー遅延積を改善していることを示している。

実効バンド幅が少ない一方で深い低消費電力モードに遷移できなかったアプリケーション gobmk に対してどの程度効果があったかを示すため、図 6 に提案機構による電力モード遷移時間を示す。提案機構を適用する前（図 1）ではセルフリフレッシュモードに平均 7.8%しか遷移できていない一方で、図 6 からは提

案機構を実装すると平均 57.1%の時間で遷移可能である。チャンネル 0 とチャンネル 1 では他のチャンネルへのアクセス数を削減したことによってアクセスが集中するため、セルフリフレッシュモード遷移時間が減少している。それ以外のチャンネルでは平均 75.1%の時間でセルフリフレッシュモードに遷移しており、提案機構によってセルフリフレッシュモードを活用できていることがわかる。

5. 主な発表論文等

（研究代表者、研究分担者及び連携研究者には下線）

〔雑誌論文〕（計 0 件）

〔学会発表〕（計 3 件）

(1) Takuya Toyoshima, Masayuki Sato, Ryusuke Egawa, Hiroaki Kobayashi, “An Application-adaptive Data Allocation Method for Multi-channel Memory,” In Proceedings of IEEE International Symposium on High-Performance and Low-Power Chips (COOL Chips 20), April, 2017. (査読有り)

(2) 豊嶋 拓也, 佐藤 雅之, 江川 隆輔, 小林 広明, “高バンド幅メモリのための省電力データ配置手法に関する研究,” In 東北支部大会連合大会予稿集, August 2016.

(3) Takuya Toyoshima, Masayuki Sato, Ryusuke Egawa, Hiroyuki Takizawa, Hiroaki Kobayashi, “A Power-Performance Tradeoff of HBM by Limiting Access Channels,” In Poster Proceedings of IEEE Symposium on Low-Power and High-Speed Chips (COOL Chips 19), April 2016.

6. 研究組織

(1) 研究代表者

佐藤 雅之 (Masayuki Sato)
東北大学・大学院情報科学研究科・助教
研究者番号：50781308

(4) 研究協力者

小林 広明 (Hiroaki Kobayashi)
江川 隆輔 (Ryusuke Egawa)
豊嶋 拓也 (Takuya Toyoshima)