

## 制御ビットを持つユニファイドキャッシュ構成の研究\* ～ 分岐予測の観点からの考察～

3H-8

池田吉朗† 安里 彰‡ 鈴木和宏‡ 木村康則‡  
(株)富士通研究所† 新情報富士通研‡

### 1. はじめに

集積度が向上するとより多くのトランジスタをキャッシュに費やせるようになる。この時キャッシュエントリ数を増やすことの他に、何らかの制御ビットをキャッシュに付加することでそれらを利用することができる。

本稿では特に、これを分岐予測に使うことを考える。制御ビットとして一次キャッシュにブロック単位で分岐予測情報を付加し、それらの情報をユニファイド型の二次キャッシュにまで反映させることで予測的中率の向上を試みる。

### 2. 制御ビットを持つキャッシュ構成

本稿では以下のようなキャッシュ構成を前提とする。

一次キャッシュは命令キャッシュ(以下 I-キャッシュ)とデータキャッシュ(以下 D-キャッシュ)が独立したセパレート型で、共に容量を 16KB、連想度を 2 とする。制御ビットを二次キャッシュにまで反映させる上で都合が良いように I-キャッシュと D-キャッシュのブロックサイズは 8 ワード (1 ワードは 32 ビット) に統一する。二次キャッシュ(以下 L2 キャッシュ)はユニファイド型で容量は 4MB、ダイレクトマップとし、ブロックサイズは 16 ワード、8 ワードのサブブロック単位で制御ビットを持たせる。制御ビットは I/D-キャッシュではブロック当たり 16 ビット、L2 キャッシュではサブブロック当たり 8 ビットを割り当てる。

I-キャッシュ上の制御ビットとしては、分岐予測に一般に使われる 2 ビットの制御情報を採用する。分岐予測情報を専用のテーブルで集中的に管理するのではなくキャッシュブロック単位で管理することで、予測情報の競合がブロック内の命令間に限定される。I-キャッシュの各ブロックには 16 ビットの制御ビットがあるので 8 個の予測情報を持つことになる。

以上の構成を前提に、単体プロセッサのトレースベースシミュレータ Paratool [1] を用いて SPECint92 の各ベンチマーク実行時の分岐予測的中率を測定する。

### 3. 分岐予測方式

#### 3.1. 分岐予測情報の状態遷移

分岐予測情報は分岐結果に従って 4 つの状態 (Strongly Taken、Likely Taken、Likely Not Taken、Strongly Not Taken) を遷移し、その遷移方法は UltraSPARC に従う。L2 キャッシュに制御情報を反映させるため、ブロック置換に伴って予測情報の書き戻しが必要になるが、その時は図 1 のに示すように L2 キャッシュ上では 2 ビット値を 1 ビット値に縮退させて格納する。図中の実線は分岐結果による遷移、破線は置換による遷移

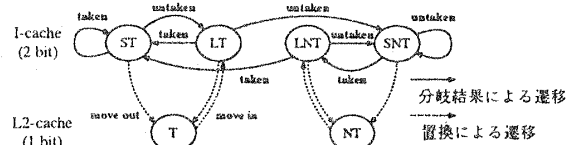


図 1: 予測情報の状態遷移

線は置換に伴う遷移を表す。I-キャッシュへの復帰時は LT と LNT に遷移する方が ST と SNT に遷移するよりの中率が 0.5~1% 程度良く、2 ビットをそのまま L2 キャッシュに格納した場合とほとんどかわらない。

#### 3.2. 予測情報の選択

分岐命令は自分が属するブロック内にある 8 個の予測情報の中から 1 個を選択し、それに基づいて予測を行わなければならない。その選択方法を図 2 に示し

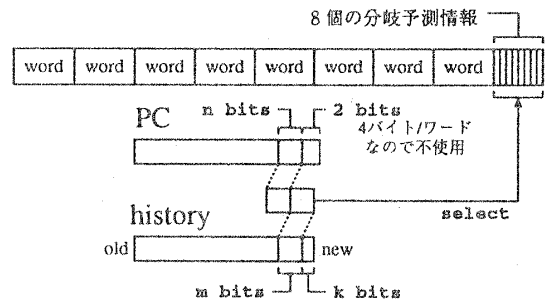


図 2: 予測情報の選択方法

た。history はグローバルな分岐履歴を表し、これの下位 k+1 目から m ビットと、PC の下位 3 ビット目から n ビットを用いて予測情報を選択する。グローバルな履歴を使うことで他の分岐命令との相関が利用できる。ここでは予測情報は 8 個なので m+n=3 でなければならない。n が大きいほど競合が小さく、m が大きいほど他の分岐命令との相関が利用できる。m+n=3 になる 4 つの組合せ (m,n)=(0,3), (1,2), (2,1), (3,0)

\* Study of branch prediction using functional unified secondary cache

† Yoshiroh IKEDA

‡ Akira ASATO, Kazuhiro SUZUKI, Yasunori KIMURA

† FUJITSU LABORATORIES LTD.

‡ RWCP Multi-Processor Computing Fujitsu Laboratory

について、それぞれを採用した時の的中率を  $k=0$  として測定した結果を図3の(a)~(d)に示す。

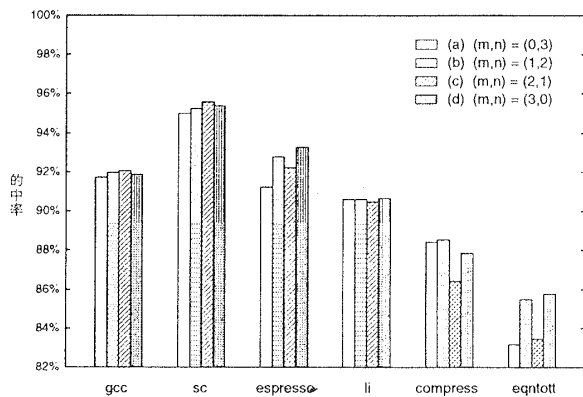


図3: 予測情報の選択方法の違いによる的中率の差

予測情報の選択にグローバルな分岐履歴を使用する場合、履歴の中で予測対象の分岐命令の結果に対して最も相関の大きな部分を使用するのが理想である。これは、図2の  $k$  に適当な値を与えることに相当する。 $k$  の値として0,1,2,3,4を採用した時の的中率を  $m=2$ ,  $n=1$  として測定した結果を図4に示す。 $k=1$ ~2の時、つまりグローバル履歴の最新の1~2ビットを無視して古い履歴を使った時の方が良い結果が得られている。その理由としては、ループを制御する分岐命令などは *taken* が続き、最後に *untaken* になって参照されなくなり置換されるというパターンが多いことが挙げられる。

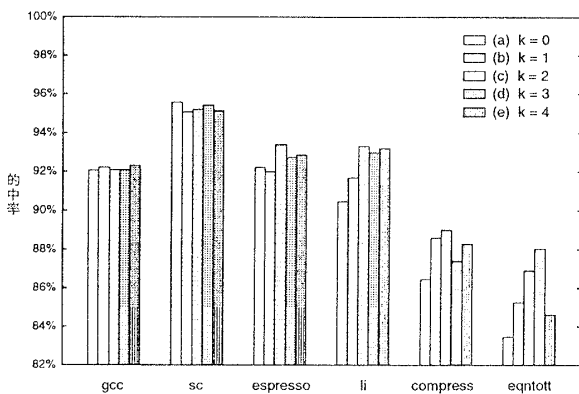


図4: 履歴ビットの取り方の違いによる的中率の差

#### 4. 他の手法との比較

本稿で述べた方法を評価するため、比較対象として通常の2ビットの予測情報のテーブル、gshare [2] についても測定した。これらを用いる分岐予測テーブルの容量は、命令キャッシュへの制御ビットの追加によるハード量の増加分(8Kビット)に等しくしている。ともにタグは使用せず、ダイレクトマップとした。また、制御ビットの情報をL2キャッシュにまで反映させない場合についても調べた。制御ビットを

使う手法では  $m=2$ ,  $n=1$ ,  $k=2$  とした。結果を示す図5中の(a)~(d)は以下を表す。

(a) 2ビット予測情報のテーブル

(b) I-キャッシュにのみ制御ビット(予測情報)を付加

(c) I-キャッシュとL2キャッシュの両方に制御ビットを付加(図4(c)に相当)

(d) gshare

L2キャッシュにも制御ビットを付加することでの的中率が向上し、ベンチマークによってはgshareに匹敵する的中率が得られる。しかしグローバルな分岐履歴を2ビットしか使っていないために、他の分岐命令との相関が十分に利用できない分gshareよりは劣ってしまう。

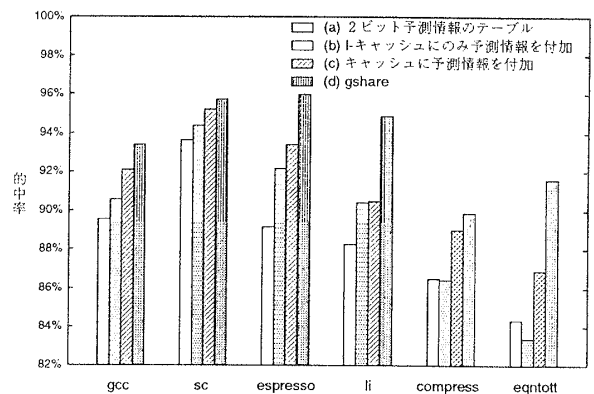


図5: 分岐予測的中率の比較

#### 5. まとめ

制御ビットとしてI-キャッシュにブロック単位で分岐予測情報を付加し、それをユニファイド構成のL2キャッシュにまで反映させる手法を評価した。L2キャッシュにまで予測情報を持ち込むことでの中率は1~3%程度向上し、ベンチマークによってはgshareに匹敵する効果が得られた。

また、データに関しては制御ビットとして、D-キャッシュ滞在中の参照回数を数えるカウンタを採用し、これを用いてブロック置換の有無を制御することを考えている。

本稿では制御情報として分岐予測情報を試行したが、制御ビットとして付加する最適な情報が何であるかをさらに検討する必要がある。

#### 参考文献

- [1] 志村, 西本, 江口, 木村. "スーパースカラプロセッサの性能評価—paratool—". 情報処理学会アーキテクチャ研究会, 10, 1993.
- [2] S. McFarling. "Combining branch predictors". Tech. Note TN-36, DEC WRL, June 1993.