

# 記憶階層システムにおけるプリロードセット学習 制御方式とキャッシュ付きディスクサブシステムへの適用

## 5K-3

山本 彰\*, 坪井俊明\*\*\*, 北嶋弘行\*, 本間繁雄\*\*, 野沢正史\*\*

\*(株)日立製作所システム開発研究所 \*\*\*(株)日立製作所小田原工場 \*\*\*\*(株)日立マイコンシステム

### 1. はじめに

計算機システムの主たる2次記憶系であるディスクサブシステムの高速化のために、近年、制御装置内に設けたディスクキャッシュを利用した記憶階層化技術が注目されている。ディスクキャッシュ付き制御装置においては、キャッシュ内に格納したデータがCPUからアクセスされると(ヒット)、ディスク装置への直接アクセスが必要なくなり、高速な入出力動作が可能となる。

一般に、記憶階層システムを用い、高速化を実現するには、近い将来にアクセスされる確率の高いデータを下位層から上位層に、プリロードすることが重要である。しかし、これまで、実用化されてきたディスクキャッシュの基本的な管理体系は、トラック単位のLRU(Least Recently Used)管理であった。具体的には、CPUからアクセスされたデータがキャッシュ内に存在しない場合(ミス)、LRU管理にしたがってキャッシュ割当て単位(スロット)が決定され、アクセスされたデータを含むトラック上のデータ、あるいは、アクセスされたデータとそのトラック上の後のデータを、決定されたスロットにロードするというものである。

LRU管理は、局所参照性の存在を前提にし、仮想記憶システム一般に用いられている管理体系である。しかし、データのアクセスには必ずしも、局所参照性は存在しないことが提言されている。例えば、以上のような管理体系では、アクセスパターンに局所性のないシステム(例えばランダムアクセス)にディスクキャッシュを適用すると、アクセスデータの周辺データのロード処理により、かえって性能が劣化するという問題があった。したがって、本講演では、あるデータがアクセスされた後、近い将来アクセスされる確率の高いデータのセットを学習し、それらのデータを上位記憶階層へのプリロードセットとするプリロードセット学習制御方式の提案と、キャッシュ付きディスクサブシステムへの適用について報告する。さらに、実サイトのアクセスパターンを用いた提案方式の評価結果を示す。

### 2. プリロードセット学習制御方式

表1に本講演で用いる記号を定義する。記憶階層システムで、アクセスされるデータ単位をレコードとする。本講演で提案するプリロードセット学習制御方式においては、レコードxのプリロードセットを、(1)、(2)式により定義する。(1)式で定義されるレコードのセットは、レコードxがアクセスされたとき、近い将来(時間:τ)アクセスされる確率の高いレコードのセットである。さらに、(1)式でSxに属すると定義されたレコードがアクセスされた後、アクセスされる確率の高いレコード、例えばレコードzが存在すれば、レコードzも、レコードxがアクセスされた時、プリロードしておいた方が、記憶階層の下位層へのアクセス回数を減少できる。したがって、(2)式を満たすレコードもSxに属するように定義した。

$$\text{Prob}(A(Ry, t_2)/A(Rx, t_1), t_2 - t_1 \leq \tau) \geq \alpha \Rightarrow Ry \in Sx \quad (1) \quad (Ry \in Sx) \wedge (Rz \in Sy) \Rightarrow Rz \in Sx \quad (2)$$

従来、仮想記憶システムで用いられてきたLRU管理などの方式は、直前に参照されたデータほど、近い将来参照される確率が高いということ、すべてのデータに関し無条件に仮定していた。これに対し、提案方式では、各データ対応に、近い将来アクセスされる確率の高いデータのセットそのものを直接学習する形をとる。

### 3. ディスクサブシステムへの適用上の課題

以上の考え方を、ディスクサブシステムに適用するためにはいくつかの課題がある。以下その内容を説明する。

[学習で取得する情報量の削減] 現在の大型ディスクサブシステム内のレコード数は、億に近い単位となるため、各レコードごとのプリロードセットを単純に学習すると、取得情報量が膨大になるという課題がある。

[τの決定方法] (1)式を満たすプリロードセットを求めるためには、τを何らかの形で決定する必要がある。

### 4. 適用方式

図1に本講演で提案するプリロードセット学習制御の適用方式の概要を示す。提案方式では、学習制御で取得する情報量を削減するため、各レコードのプリロードセットを学習す

表1. 記号の定義

<p>Ri:レコードi, <math>\text{PROB}(A(b, t_2)/A(a, t_1), t_2 - t_1 \leq \tau)</math> : aが時刻t1にアクセスされた後、τ時間内にbがアクセスされる確率</p> <p>Si, Sa:レコードi, アクセス部のプリロードセット</p> <p>Da, Df, Dr:トラックのアクセス部、フロント部、リア部</p>
---

Preload Set Learning Control in Memory Hierarchy Systems and Its Application for Cached Disk Subsystems  
Akira YAMAMOTO\*, Toshiaki TUBOI\*\*\*, Hiroyuki KITAJIMA\*, Shigeru HONMA\*\*, Masafumi NOZAWA\*\*  
\*Systemes Development Laboratory, Hitachi, Ltd., \*\*Odawara Works, Hitachi, Ltd., \*\*\*Hitachi Microcomputer System, Ltd.

る範囲を、そのレコードが格納されたトラックに限定する。さらに、トラック上の各レコードを、アクセス部、フロント部、リア部に分類した。アクセス部は、トラック上のレコードのうち、CPUから直接アクセス対象となったレコードの集合である。フロント部は、トラック上で、アクセス部の前にあるレコードの集合、リア部はアクセス部の後にあるレコードの集合である。この場合、そのトラック上で、最初にアクセスされるレコードの集合がアクセス部となる。したがって、トラック対応に、アクセス部がアクセスされた後、フロント部、リア部がアクセスされる条件つき確率を(3)式により学習する。

$$\text{Prob}(A(Df, t_2)/A(Da, t_1), t_2 - t_1 \leq \tau) \geq \alpha \Rightarrow Df \in Sa, \text{Prob}(A(Dr, t_2)/A(Da, t_1), t_2 - t_1 \leq \tau) \geq \alpha \Rightarrow Dr \in Sa \quad (3)$$

ただし、トラック上のデータの内最初にキャッシュにロードされるレコードがアクセス部であるという点、ディスク装置が回転体であるという制約より、フロント部がSaに属する時には、リア部もキャッシュにロードする形をとる。したがって、提案方式においては、プリロードセットを、図1に示したように3つのパターンに分類する。

次に、 $\tau$ の考え方について説明する。学習範囲のある領域(適用方式ではトラック)に限定した場合、(1)、(2)式を満たすプリロードセットは、その領域単位のLRU管理を用いると求めることができる。LRU管理を用いると、ある学習範囲領域上のデータがアクセスされると、そのデータがスロット上に格納されていない場合も含め、スロットは、MRU(Most Recently Used)位置に移される。したがって、スロットが、MRUからLRUまで移動するまでの時間を $\tau$ と考えると、ある学習範囲領域の一部がキャッシュにロードされた後、その学習範囲領域全体がキャッシュから追い出されるまでに、アクセスされたデータが、(1)、(2)式を満たすプリロードセットということになる。

以上より、適用方式では、トラック上の一部データをキャッシュにロードした後、そのトラック全体をキャッシュから追い出すまでの間に、フロント部、リア部がアクセスされる条件つき確率を求め、プリロードセットを決定する。

5. 提案方式の評価

本講演で提案した制御方式を、実サイトの入出力アクセスパターンを用いたシミュレーションにより評価した。図2に、その評価結果をまとめた。本評価結果では、約75%のプリロードセットはアクセス部のみであった。したがって、従来のように、常にアクセスされたレコードの周辺レコードをキャッシュにロードすると無駄な処理を実行してしまうことになる。図2は、常にアクセス部+リア部をプリロードセットとした場合と本制御方式を適用した場合の性能比較である。この場合、提案方式の適用により約20%の効果を得られるという結果が得られた。

6. おわりに

あるデータがアクセスされた後、近い将来アクセスされる確率の高いデータを学習し、それらのデータを上位記憶階層へのプリロードセットとするプリロードセット学習制御方式の提案と、キャッシュ付きディスクサブシステムへの適用方式について報告した。実サイトの入出力パターンによるシミュレーション評価では、提案方式の効果は、約20%であった。

参考文献 1) Smith, A.J : Disk Cache-Miss Ratio Analysis and Design Consideration, ACM TOCS, Vol3, No.3, pp.161-203(1985) 2) 日立マニュアル:H-6581-C3型ディスク制御装置, 8080-2-130 3) 大須賀 : プログラムおよびデータの特性指標とその仮想記憶システムへの影響, 情報処理学会論文誌, Vol.23, No.3, (1979) pp.256-264

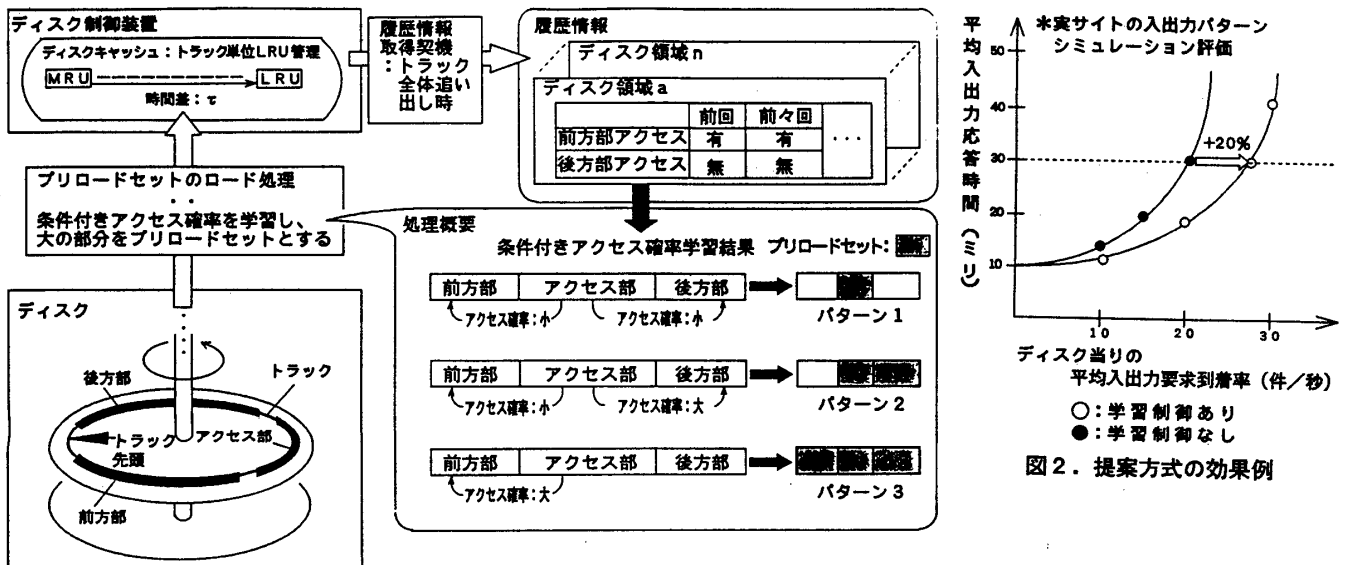


図1. プリロードセット学習制御方式の概要