

ユーザプログラム制御階層メモリシステムに関する研究*

1B-2

牧 晋広 岡本秀輔 曾和将容†

電気通信大学大学院情報システム学研究科‡

1 はじめに

コンピュータは、基本的に高速で大容量のメモリを必要とする。しかし現実には、高速かつ大容量のメモリを用意することは困難であるので、実際には小容量で高速なメモリと大容量で低速なメモリを階層的に組み合わせて、疑似的に高速で大容量のメモリを実現している。これは、階層メモリシステムと呼ばれ、階層メモリシステムで最も用いられているメモリシステムの1つに、汎用レジスタ、キャッシュメモリ、主メモリの組み合わせて構成する、キャッシュメモリシステムがある。

キャッシュメモリシステムでは、キャッシュメモリに必要なデータが無いという、いわゆるキャッシュミスが起こり、これがコンピュータの性能を下げている。

キャッシュミスが発生する原因として、キャッシュメモリと主メモリの内容の入れ換えを一定のアルゴリズムで行っていたことがある。そのため、実行するプログラムによっては、頻繁にキャッシュミスを起こす結果となる。我々はこの問題を解決するために、実行するプログラム(便宜上処理プログラムと呼ぶ)ごとに、階層メモリ間の内容を操作する、それ専用のプログラム(メモリ操作プログラムと呼ぶ)を作製し、このメモリ操作プログラムを、処理プログラムと並列に動かすことによって、疑似的に高速大容量のメモリを構成する方法に関する研究を行ってきた。² この方式では処理プログラムが必要とするデータを、プログラム制御によって必要な時まで用意しようとするので、処理している問題が本質的にもっているデータの従属性と階層メモリシステム

が本質的にもつ性質以外の影響を受けずにデータを準備できる。

本論文では、本方式をユーザプログラム制御階層メモリシステム:UPCHMS(User Program Controlled Hierarchical Memory System)と呼び、このUPCHMSの基本原則と基本構成と簡単なプログラムについて述べる。なお、本方式ではキャッシュメモリに相当するメモリがプログラマに見えるので、キャッシュ(隠れる)という概念が使えない。そのため本研究では、キャッシュメモリに相当する階層のメモリを高速メモリと呼ぶ。

2 UCPHMSの原理

プログラムにおいて、どこの部分で、どのデータが必要になるかを熟知しているのは、そのプログラムを製作したプログラマである。しかし、キャッシュメモリシステムでは、プログラマが知っているデータ使用情報を完全には利用しない階層メモリ間のデータ転送が行われている。そのためキャッシュミスが発生していた。

そこで、本研究は、そのことに着目し、階層メモリ間のデータ使用情報を100パーセント利用することで、高速大容量なメモリを実現する。

その方法として、階層メモリ間のデータの使用情報を100パーセント利用するために、処理プログラムが必要とするデータの時間と場所を割り出し、そのデータを上部階層のメモリに用意する、メモリ操作プログラムを作製し、このメモリ操作プログラムを、処理プログラムと並列に動かすことによって、疑似的に高速大容量のメモリを構成する。

この研究は、データを扱うメモリ(これをデータメモリと呼ぶ)と命令を扱うメモリ(これを命令メモリと呼ぶ)の内の、データメモリに対する研究であり、命令メモリは、扱わない。

*Study of User Program Controlled Hierarchical Memory System

†Nobuhiro MAKI, Shusuke OKAMOTO, Masahiro SOWA

‡Graduate School of Information Systems, University of Electro-Communications, Tokyo, 182 Japan

3 UPCHMS のハードウェア構成

図1は、本研究UPCHMSのハードウェア構成である。プログラム用メモリ :IM(Instruction memory)は、プログラムを格納する為のメモリであり、データメモリ :DMは、データを格納する為のメモリである。そのデータメモリは、図のように超高速メモリ :VHM(Very high speed memory)、高速メモリ :HM(High speed memory)、主メモリ :MM(Main Memory)の3階層よりなる。高速メモリは、キャッシュメモリに相当し、超高速メモリは、レジスタに相当する。また、高速メモリは、他のメモリと独立にアドレスを持つ。PU、HU、MUは、プロセッサユニットで、以下のように働く。

PU : 一般の処理を実行するプロセッサユニットで、超高速メモリのみを対象に処理を行う。

HU : 超高速メモリと高速メモリにアクセスすることができるプロセッサユニットで、超高速メモリと高速メモリ間のデータの転送を行なう。また、簡単な演算機能と演算用のレジスタメモリ (HUR) を持つ。

MU : 高速メモリと主メモリをアクセスすることができるプロセッサユニットで、高速メモリと主メモリ間のデータの転送を行なう。また、簡単な演算機能と演算用のレジスタメモリ (MUR) を持つ。

それぞれのプロセッサユニットは、独立したプログラム用メモリを持ち、それらのプログラムは、それぞれのプロセッサ用のプログラム用メモリに格納される。

4 UPCHMS のソフトウェア構成

ここでは、以下のような、 $\sum_{k=0}^5 f(k)$ を行うプログラムを例に説明を行う。まずは、シリアルプログラムで、ベースのプログラムを製作し、それを、UPCHMS のプログラムに変換する。

シリアルのベースプログラム

```

move 0,r1
move 0 r2
loop  great r1,5
      branch EXIT
      add r1,DATA,r4
      load r4,r3

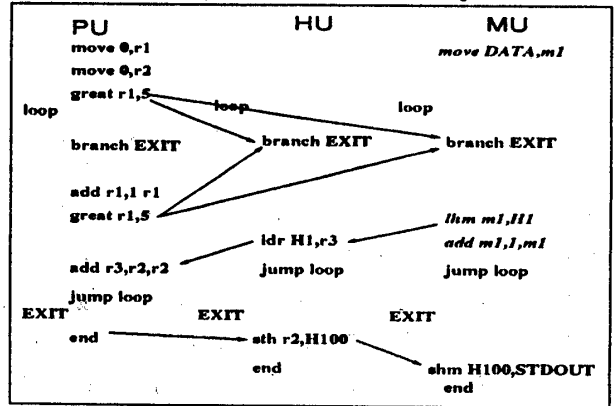
```

```

add r2,r3,r2
add r1,1,r1
jump loop
EXIT  store r2,STDOUT
end

```

これをUPCHMSが、最適に動作するように変換を行うと、以下のようなプログラムになる。



3つのプログラムは、左側から、PU、HU、MUのプログラムである。また、矢印は、命令間の先行関係を示す。

このプログラムの実行時間を調べて見ると、同容量のキャッシュメモリを持つシステムに比べ、実行時間ので1.14倍の性能が得られ、超高速メモリにデータがないためにPUが、処理を停止する時間は、存在しなかった。

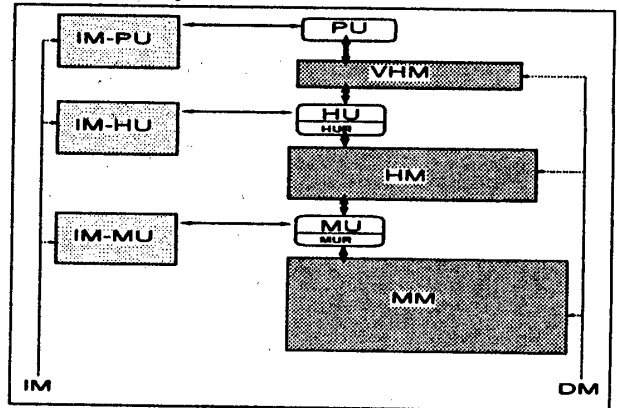


図1 UPCHMS のハードウェア構成
参考文献

1. J.Hennessy & D.Patterson, 'COMPUTER ARCHITECTURE: A QUANTITATIVE APPROACH' MORGUN KAUFMAN PUBL
2. 佐藤 正樹, 'キャッシュメモリの操作明示化に関する研究', 名古屋工業大学卒業論文 (1989)
3. 牧 晋広, '階層化メモリの能動的構成法', 平成4年度名古屋工業大学卒業論文