

オーバーヘッドの少ない
ハードウェアバックアップ方式の設計

4N-1

上林 彌彦* 高倉 弘喜* 南 清一** 中村 千秋***
*九州大学工学部 **日本 I B M ***長崎大学工学部

1. まえがき

データベースシステムに対する高速化の要請は幅広くあり、そのため、ソフトウェア的ないしはハードウェア的に高速化するための研究開発が行われてきている。システムの総合的な高速性は単にデータベース演算の高速実現だけではなく、システム障害時の回復速度の改善も重要である。従来のデータベースマシンは主として前者の問題を扱ってきた。本稿では、データベースシステムの高速度化に主記憶データベースが重要であること、データベースの一部が特に使われるホットスポットデータであることに注目して、ホットスポットデータのバックアップハードウェアの設計について検討する。ホットスポットデータは更新頻度が高く、並行性を向上させるためには、データの施錠単位を小さくしなければならない。このため、ソフトウェア的な方法では並行処理制御およびバックアップのオーバーヘッドが非常に大きくなる。従って、メモリープレーンを2枚用意しバックアップ中にも処理の中断がないように設計した。本稿ではこのバックアップシステムについて、その構成とハードウェアの設計の基本的概念について述べる。

ホットスポットデータは主記憶上にあると仮定する。また、ホットスポットデータの更新は頻繁なので、この部分のバックアップをハードウェア化する。他のデータとの同期については別に議論する。バックアップシステムの構成は、図1のようになっている。図1のように本システムは、2枚のメモリープレーンから構成されており、通常は1枚のメモリーのように扱う。これらのメモリープレーンのうち1枚はホットスポットメモリープレーン(HSMP)と呼ばれ、ホットスポットデータに対する主記憶からのアクセス要求に対応し、もう1枚はバックアップメモリープレーン(BUMP)と呼ばれ、ディスクへの書き込み要求に対応する。まず図1aは、通常の動作を示す。この場合は、2枚のメモリープレーンを1枚のメモリープレーンと見なして同時に書き込み等の操作を行なう。図1bではバックアップを行なうために、BUMPは主記憶側から切り放され、2次記憶(DISK)に接続されていることを示している。この状態では、主記憶からの書き込み等はHSMPにのみ反映される。図1cではバックアップ終了後にはHSMPとBUMPとの内容が一致していない可能性があり、このためにHSMPからBUMPへデータの転送が行なわれる。また、更新処理と同時に進行するため矛盾の生じない工夫を行っている。

2. 基本的事項

2. 1 ホットスポット

データベースにおけるアクセスは、すべてのデータに対して均等にアクセスが行なわれるのではなくて、例えば全データの約10%の部分に対して、アクセスの約90%が集中する傾向がある。このように他データに比べてアクセスが集中するデータをホットスポットと呼ぶ。

2. 2 バックアップシステム

主記憶上のデータベースはなんらかの原因(例えば、停電)により失われることもありうる。従って、事故によるデータベースの消失を避けるために、定期的にデータベースを2次記憶に記憶し、事故の際には直ちに復旧回復作業を行なうようなバックアップシステムが必要となる。

2. 3 Dual-Port-RAM

通常のRAMにさらにもう1つ入出力ポートをもうけたもの。ただし、現在商品化されているもので、SRAMは2つのポートともランダムアクセスできるのに対して、DRAM(VTR用)は1つのポートではランダムアクセスであるがもう1つのポートではシリアルアクセスしかできず、またデータの転送時に制約がある。

3. バックアップシステムの構成

ここでは、すべてデータベースは主記憶上にあるとしてもよいし、一部のみ主記憶上にあるとしてもよい。しかし、

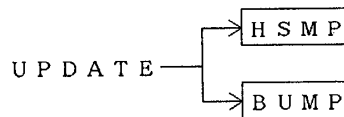


図1a 通常時

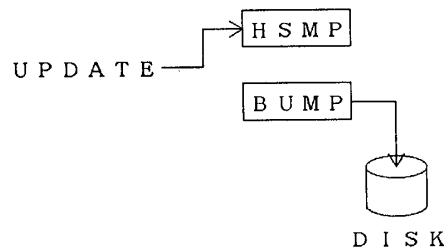


図1b バックアップ時

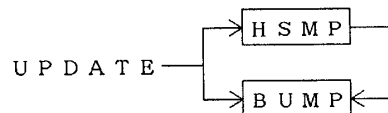


図1c データ転送時

Design of Hardware Schema to Reduce Backup Overhead

Yahiko KAMBAYASHI*, Hiroki TAKAKURA*, Seiichi MINAMI** and Chiaki NAKAMURA***

* Faculty of Engineering, Kyushu University, ** IBM Japan, *** Faculty of Engineering, Nagasaki University

4. ハードウェアの構成

3. のバックアップシステムを実際にハードウェアで実現するには多くの問題がある。図1から明かなようにそれぞれのメモリープレーンは2つの入出力ポートを持つ。このため、メモリー素子にも2つの入出力ポートを持つものが希望される。これを満たすものとして当初はDual-Port-RAMを考えていた。しかし、SRAMには大容量のものがなく、またDRAMは主にその目的がVTR用のメモリーのため利用しにくいといった問題があった。さらに、DRAMにはリフレッシュやプリチャージでの待ち時間の問題もある。また、普通のRAMを用いても以下に述べる方法で2つの入出力ポートを持たせることもできる。このため本稿では普通のRAMを利用する方式を検討した。

普通のRAMは1つしか入出力ポートを持たないので、メモリープレーンに対する2つの入出力が競合することが当然起こる。図2は、2つの入出力ポートに対するアービトレーション（調停）回路の構成を示している。本研究においてはCPUは68000シリーズを使用するため、これに対応したシステムを設計する。この回路の入力線にはR/WとBSがある。R/Wはメモリーへのアクセス要求がRead要求なのかWrite要求なのかを表し、BSとはメモリーへのアクセス要求の有無を表している。一方、出力線にはSelect, CS, OE, WE, L, G, ACKがある。Selectとはアクセス要求A, Bのうちの一方のアドレス入力の選択を、CSは該当するメモリー素子の選択を、OEはメモリー素子からのデータ出力の有無を、WEはメモリー素子へのデータ書込みの有無を、Lはラッチ期間の制御を、Gは入力ゲートの制御を、ACKはデータ転送の終了をそれぞれ表す。

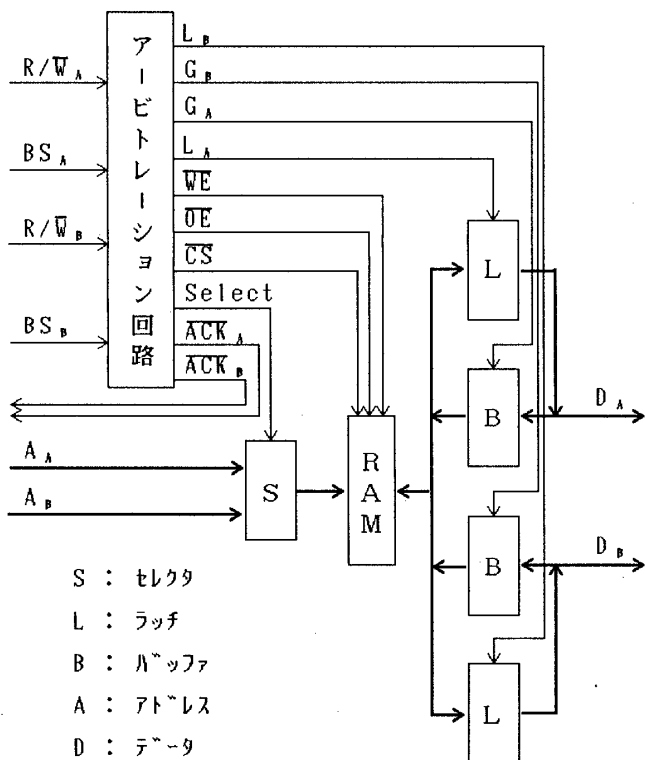


図 2

次に、データ出力（読出し）において、ラッチが必要となる理由を述べる。本システムでは、メモリー素子の1つのポートにメモリープレーンへの2つのポートを接続するため、その切り替えが行なわれる。ここで、もし1つめのポートからの出力が終了するまで2つめのポートが接続できないならば、接続を待つ時間が長くなってしまふ（図3 a）。このため、1つめのポートに対してメモリー素子からの出力をD形フリップフロップで保持し、このフリップフロップの出力からデータを読出すと、2つめのポートがすぐに接続でき、待ち時間を短縮できる（図3 b）。ただし、入力（書込み）は、書き込まれるデータが確定するまでは、そのデータに対してのもう一方からの読出し、書込みは禁止する。また、読出し中のもう一方からの書き込みも禁止する。これは、データの処理上で矛盾を生じないためである。

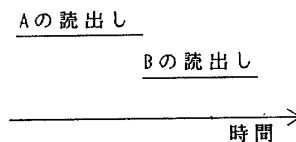


図 3 a

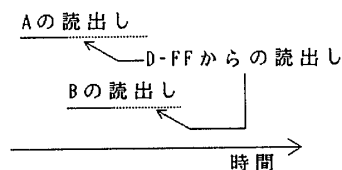


図 3 b

従って、以上のような方法を用いると、普通のRAMでも2ポートをもつメモリープレーンが構成できる。

5. まとめ

本稿は、バックアップシステムの構成について述べ、さらにそのハードウェア設計における概要を述べた。本稿の方式によると市販のものより高速で大容量のものが実現できることになる。今後はこの方法でハードウェアを実現していく予定である。

謝 辞

本研究にあたり、有益な御助言をいただいた本学 最所圭三助手、ならびに研究室諸氏に感謝いたします。

尚、本研究は文部省科学研究費試験研究によるものである。

参考文献

[1] 岡田正、"メモリーIC活用ハンドブック" CQ出版社、1989
 [2] 上林彌彦、南清一、"ホットスポットデータに対するバックアップ"、第41回電気関係学会九州支部連合大会、昭和63年