

超大容量仮想記憶システムSTS(1)

- 概要 -

2P-1

鍵政 豊彦, 吉住 誠一, 山岡 彰, 小田原 宏明

(株)日立製作所 中央研究所

1. はじめに

計算機のメモリ・アーキテクチャは、近年の処理の大規模化と半導体技術の進展により変革がせまられている。大容量の主記憶を活用するシステムの研究<sup>[3]</sup>もすでに一部では行なわれている。超大容量仮想記憶システムSTS(Super Tera-byte System)は、メインフレーム・コンピュータを主対象とし、テラバイト(兆;  $10^{12}$ )オーダの仮想記憶ならびに実記憶の実現を目標とするものである。

2. 大容量仮想記憶実現の課題

大容量仮想記憶の実現にあたって、デマンドページング方式を基本とするシステムでは、以下の課題の解決が必要である。

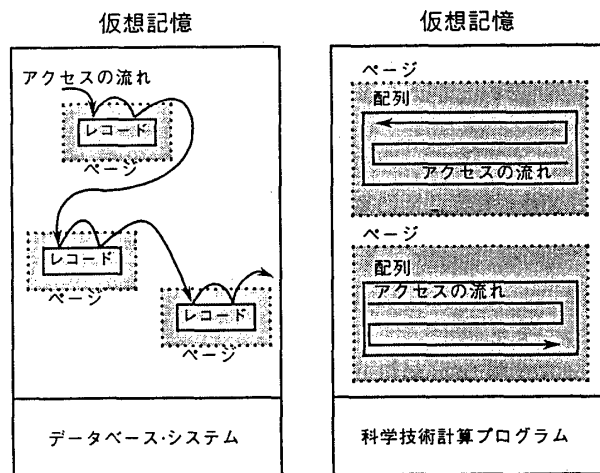
・仮想記憶管理のオーバーヘッド削減

ページフォルト、ページ計測、ページスケジュール、スワッピング処理などのCPUオーバーヘッドはページ数に比例して増大するため、単純な大容量化は逆に処理性能の低下につながる。

・アクセス特性の多様化への対応

従来の小容量仮想記憶では、データよりも命令に対するアクセス特性が強いため、ワーキングセット法やLRU法によく適合する特性となる。しかし、大容量仮想記憶ではデータのさまざまなアクセス特性が表面化し、一元的な管理は不可能となる。

図1は上記の課題の難しさを表している。データベースシステムのように、各所に散らばるレコードをアクセスする場合は、小ページサイズが有利である。



(a) 小ページサイズ有利のケース (b) 大ページサイズ有利のケース

図1. アクセス特性とページサイズ

り、科学技術計算プログラムのように、大きな配列を逐次的にアクセスする場合は、大ページサイズが有利である。ページサイズの拡大により管理オーバーヘッドの削減を図った場合、データベースシステムでは占有実記憶量が増大して実記憶不足のためにページングが多発し、性能が低下する。

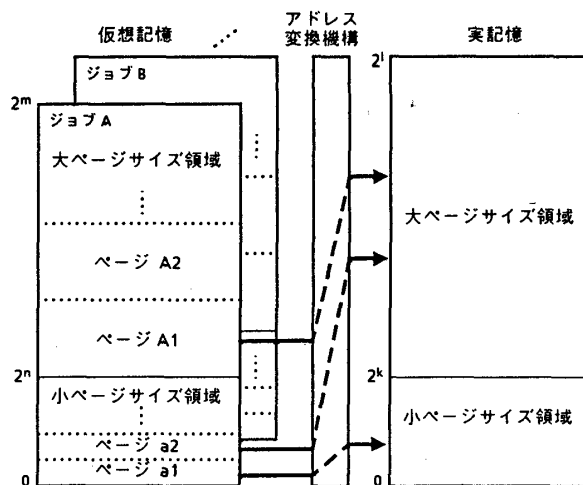


図2. 仮想/実記憶の構成

3. 大容量向き仮想記憶管理方式

STSのアーキテクチャとOSは、大容量仮想記憶を、管理単位の拡大、部分解析制御、および適応制御の3つの基本方針のもとに実現する。

仮想記憶管理の特徴方式を以下に示す。

(1) 多重ページサイズ

図2に示すように、仮想記憶、実記憶ともに管理単位であるページサイズを拡大し、ページサイズが小の領域と大の領域を設け、小仮想記憶ページを大実記憶ページにも割り当て可能とする。システム負荷のアクセス特性から小仮想記憶ページが有利な場合は、実記憶ページのみを大として管理オーバーヘッドの削減が可能である。また、この方式は従来仮想記憶システムとの互換性および移行性においても有効である。

(2) 部分解析ページ・リプレースメント

ページ・リプレースメントでは、ジョブの全実記憶ページのアクセス管理情報(エイジ, 変更ビットなど)を参照するのではなく、一部の実記憶ページの情報からジョブ全体の特性を予測して、リプレースメント対象ページを決定する方法を採り、オーバーヘッドの削減を図る。

(3) 適応型スワッピング

図3に示すように、ジョブ状態として、占有実記憶の全体を解放したアウト状態に加え、部分的に解放した部分アウト状態を新設する。大規模ジョブに対するスワッピングでは、必要分の実記憶のみを解放して部分アウト状態とし、残りの実記憶は必要に応じてページスチールする。ページスチールが一定量を越える場合は、ジョブ再開時にページフォルトが頻発するのを防ぐため、アウト状態にする。本方式により大規模ジョブのスワッピングI/Oによるスラッシング発生を防ぐことができる。

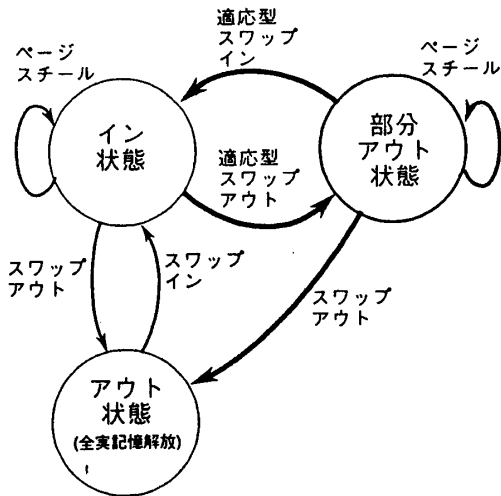


図3. 仮想記憶管理におけるジョブの状態遷移

4. 大容量仮想記憶の利用技術

仮想記憶の拡大は、応用問題の大規模化とともにメモリバウンド・アルゴリズムによる処理の高速化も可能にする。STSでは、後者の用途を支援するために、大容量仮想記憶の利用を効率化する以下の方式を開発した。

(1) アクセス特性ページング

応用プログラムから仮想記憶領域のアクセス特性をOSに通知してもらうことにより、OSが特性に応じた仮想記憶制御が可能となり、ジョブの実行が高速化される。順アクセス特性に対する先行ページングが特に効果的である。

(2) パックドシノニム・ハッシュ (PSH)

大容量主記憶の効果的利用法として、ハッシングによるデータアクセスの高速化がある。STSではさらに、大容量仮想記憶向きのパックドシノニム・ハッシュ (PSH) を開発した。シノニムの集合を同一仮想記憶ページに割り当てることによりページフォルトを削減して高速化する方式である。

参考文献

[1] 高橋 聡: 超大容量仮想記憶システムSTS (2)-多重ページサイズ方式の評価- (本大会予稿集)  
 [2] 森 聡: 超大容量仮想記憶システムSTS (3)-利用技術の評価- (本大会予稿集)  
 [3] Garcia-Molina, et.al.: A Massive Memory Machine, *IEEE Transaction on Computers*, Vol.C-33, No.5, May 1984, pp.64-73.

5. 実験システムによる評価

STSの実験システムを当社汎用計算機MシリーズおよびそのOSであるVOS3をベースに試作した。実験システムは、256TBの仮想記憶を実現し、1.5GBの実記憶を実装する。仮想/実記憶はともに、4KBの小ページサイズ領域と16,64,256,1024KBの4種のページサイズをシステム立ち上げ時に選択可能な大ページサイズ領域を有する。

STSの仮想記憶管理方式を実験システムにて評価した結果を図4に示す。各種応用プログラムごとに10多重実行し、実記憶量に比例して応用プログラムの使用仮想記憶量を拡大して、ページフォルト率がほぼ一定の環境で計測した。その結果、3節で述べた方式を無効とした場合は実記憶量の増加とともに管理CPUオーバーヘッドは増加するが、有効とした場合はオーバーヘッドはほとんど増加せず、スループットも低下しない。このことからSTSの仮想記憶管理はギガバイト、さらにテラバイト・オーダの仮想および実記憶を実現できる見通しを得た。

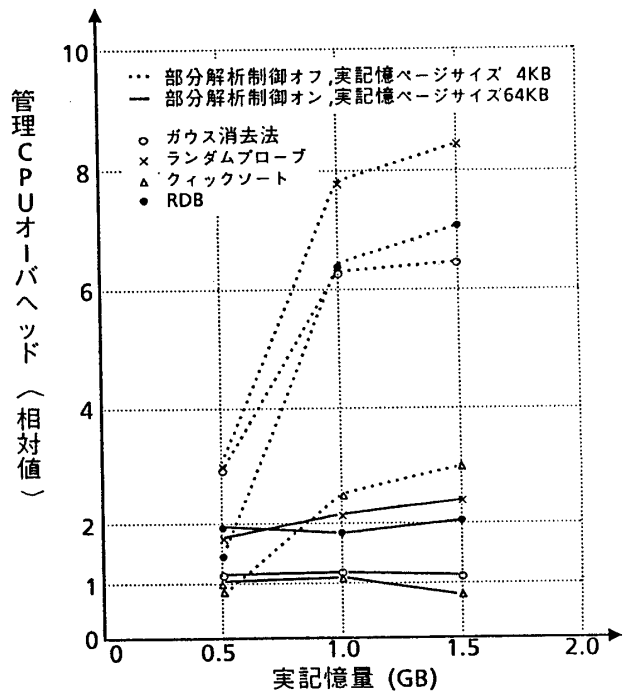


図4. 管理オーバーヘッドの削減効果

6. おわりに

大容量仮想記憶を実現する管理方式および利用技術を提案し、実験システムにて有効性を確認した。今後は、応用研究の拡充特に、データベース分野での利用技術の開発を進める。