

超大容量仮想記憶システムSTS(2)

2P-2

- 多重ページサイズ方式の評価 -

高橋 喜久雄, 鍵政 豊彦, 森 利明, 松田 芳樹

(株)日立製作所 中央研究所

1. はじめに

大容量仮想記憶時代の問題の一つとして、オペレーティングシステム(OS)の仮想記憶管理オーバーヘッドがある。デマンドページング方式では、OSで管理するページ数が、仮想記憶量および実記憶量に比例して増大し、その管理オーバーヘッドが膨大なものになる。さらに、実記憶においては、その増加と共にシステム性能が低下すると言った逆現象が危惧される。STS(Super Terabyte System)では、この解決の一方法としてページサイズの拡張によりページ数を減らす手段を採った。さらに、ジョブのメモリアクセス特性や互換性に対する実ページサイズ拡張の影響を考慮し、複数サイズの仮想/実ページが同一システム内に共存する多重ページサイズ方式を採用した。STS実験システムを試作し仮想/実ページサイズを種々変化させ多重ページサイズ方式の実験評価をしたので、その結果を報告する。

2. 多重ページサイズ方式の概要

(1) 仮想/実記憶の構成

本実験システムにおける多重ページサイズ方式では、図1に示すような仮想/実記憶の構成となっている。仮想記憶は2GBを境として4KBの仮想ページで管理する領域と、拡張ページサイズで管理する領域とに2分割されている。

一方、実記憶もシステム構成時に定まる、あるアドレスを境に、同様に2分割されている。さらに、それぞれの領域は、以下の様に2分割管理され、合計4領域より実記憶を構成する。

(a) 同一サイズ割り当て専用領域

自領域を構成する実ページと同一サイズの仮想ページのみを割り当てる専用領域(図の4KB専用領域と拡張専用領域)として管理する。

(b) 異種ページサイズ割り当て可能領域

実ページサイズによらず、いずれのサイズの仮想ページも割り当て可能な領域として管理する。また、ページ毎のアクセス情報(参照ビットや更新ビットなど)を含むキー記憶は、実ページ単位に実装されている。

(2) ページ割り当て方式の概要

上記の仮想/実記憶構成において以下に示す方式により仮想ページに実記憶領域を割り当てる。

(a) 各専用領域

仮想/実ページサイズが等しく仮想/実ページを1対1に単純に割り当てする。

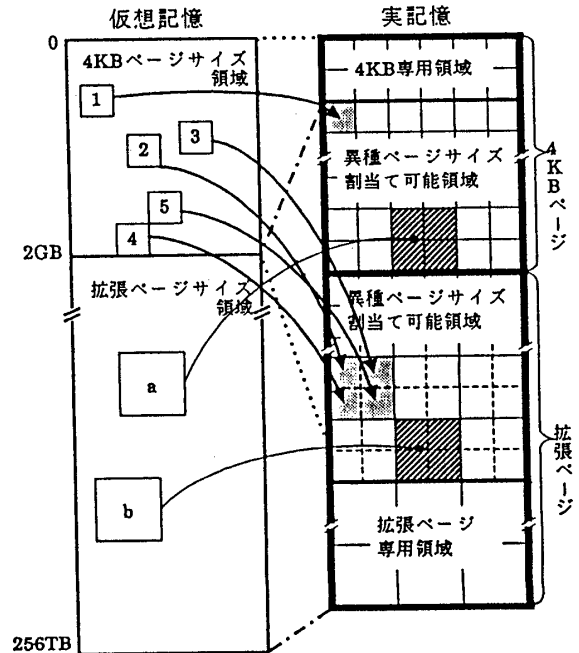


図1 STS実験システムの仮想/実記憶構成

(b) 異種ページサイズ割り当て可能領域

本領域においても、例えば、図中仮想ページ1やbのように仮想/実ページサイズが等しい場合には、前記と同様に仮想/実ページを1対1に単純に割り当てする。ページサイズが異なる場合には、以下のように実施する。

(i) 4KBの仮想ページを拡張実ページに割り当てる場合には、拡張実ページ内を4KBサイズにn分割して、各々の分割領域を仮想ページ単位に割り当てる。この時、同一拡張実ページ内には同一ジョブかつ同一キーの仮想ページのみ割り当てる。すなわち、拡張実ページ上にページ2が最初に割り当てられた後、この拡張実ページは、当該ジョブのローカルな未使用実記憶領域として管理される。

以降、当該ジョブでページ割り当てが発生すれば、このジョブのローカル未使用実記憶領域から優先して、ページ3, 4, 5が順次割り当てられる。

(ii) 図1の拡張ページaに示すように、拡張仮想ページを4KB実ページサイズ領域に割り当てる場合は、連続するn個の4KB実ページ

ージを一個の拡張実ページと考え、これを割り当てる。

以上述べた異種仮想／実ページサイズ間の相互割り当てを可能とするため、本実験システムでは、まず、実記憶領域を拡張仮想ページサイズ単位に管理し、その内部をさらに4KB仮想ページサイズ単位に分割管理する。以上の様に仮想および実記憶ともに拡張したシステムを仮想記憶拡張システムと呼ぶことにする。

また、仮想記憶の構成はそのまま、実記憶構成のみ、そのページサイズを拡張した場合の実験システムでは、図1の仮想／実記憶構成が変わるだけであり、4KBの仮想ページを拡張実ページに割り当てる方式は前記と同様である。この様なシステムを実記憶拡張システムと呼ぶことにする。

(3) ページリプレースメント方式の概要

ページ計測とページスチール処理は、仮想ページサイズとそれが割り当てられている実ページサイズの内、大きい方のページサイズ毎に実施し、処理ページ数を減らしオーバーヘッドの低減を図った。

3. 実験結果

上記方式のOS及び実験機を試作し、ジョブのメモリ・アクセス・パターンにより、逐次型、ランダム型、混在型メモリアクセスジョブ等をベンチマークモデルとして実測評価した。

(1) 実記憶拡張システムの実測結果

表1に、混在型メモリアクセスジョブ(実記憶量1.5GB、クイックソートジョブを10多重実行)を例に各拡張ページサイズ毎の性能比較を示す。

本実験システムのような方式を採った場合、ページ計測処理も、ページスチール処理も、実ページサイズ大ほどオーバーヘッド小となる。

しかし、ページサイズ大ほど、より多くの仮想ページが同一の実ページ上に混在し、そのためLRU精度が低下しページング率を上げ、性能が低下する。また、拡張実ページ上に1個でも4KB仮想ページが割り当てられると、実ページ内の、残りの未割り当て領域は、ジョブのローカルな空き領域となる。このため、システム全体での実記憶利用効率が、実ページサイズ大ほど低下し、ページング率を上げる(表1の実ページ内分割損)。実測ではこれらの要因がバランスする64KB実ページで最良の結果を示した。

(2) 仮想記憶拡張システムの実測結果

表2に逐次型のデータ参照を主とするジョブ(逐次型メモリ・アクセスジョブ)とランダム・アローブジョブ(ランダムメモリアクセスジョブ)の仮想ページサイズによる性能比較を示す。逐次型ではページサイズ大ほどページング率が下がり性能が向上する。ランダム型では、

ページサイズ拡張とともに、急激にページング率が上がり性能が低下する。仮想ページサイズの拡張は、実ページサイズの拡張に比べ、予想した以上に、ジョブのメモリ・アクセス・パターンに敏感である。

(3) 評価結果

(a) 実ページサイズ拡張は前記のように、性能に影響する要因がバランスする最適サイズが存在し、仮想ページサイズの拡張とは異なり、そのサイズによる性能差は、ほとんど無く、実記憶増加に共なう管理オーバーヘッドを抑止するのに効果的である。

(b) 仮想ページサイズ拡張は、管理オーバーヘッドの大幅低減が可能である。しかし、そのサイズによる性能への影響は、予想以上に、ジョブのメモリ・アクセス特性に敏感であり、アクセス特性に対応した、例えば、参考文献[2]のアクセス特性ページングなどが必要であることが分かった。

(c) 多重ページサイズ方式は、同一システム内に、複数の仮想／実ページサイズを、ある程度混在させる事ができ、前記のような課題(アクセス特性)に適応が容易であり効果的である。

4. おわりに

STS実験システムを試作して、多重ページサイズ方式の有効性を確認した。今後は、本方式の適応化の拡充を進めたい。

参考文献

- [1] 鍵政ほか：超大容量仮想記憶システム(1) - 概要 - (本大会予稿集)
- [2] 森ほか：超大容量仮想記憶システム(3) - 利用技術の評価 - (本大会予稿集)

表1 混在型メモリ・アクセス・ジョブの実ページサイズによるシステム性能

実ページサイズ (KB)	相対アローブレムCPU利用率	相対スーパーバイザCPU利用率	相対ページ計測オーバーヘッド	相対ページスチールオーバーヘッド	相対ページイン数	拡張実ページ内分割空き損(MB)
4	1	1	1	1	1	-
16	1.02	0.84	0.37	1.56	0.96	0.4
64	1.03	0.70	0.30	0.60	0.73	1.9
256	1.03	0.75	0.26	0.41	1.14	4.9
1024	1.01	0.90	0.22	0.82	1.29	27.0

表2 仮想ページサイズによるシステム性能

仮想ページサイズ (KB)	逐次型メモリ・アクセス・ジョブ			ランダム型メモリ・アクセス・ジョブ		
	相対アローブレムCPU利用率	相対スーパーバイザCPU利用率	相対ページイン数	相対アローブレムCPU利用率	相対スーパーバイザCPU利用率	相対ページイン数
16	1	1	1	1	1	0
64	1.72	0.27	0.250	1.51	1.41	0
256	2.00	0.11	0.062	0.034	0.106	1
1024	2.13	0.10	0.015	0.009	0.066	3.1