

仮想メモリの返却忘れ検出機構の試作

7P-3

吉木安里

(富士通株式会社)

1 はじめに

仮想記憶方式のオペレーティングシステムには、仮想メモリ（以後単にメモリと呼ぶ）を動的に獲得／返却するための機能が装備されている。メモリは、通常、いくつかの枠に区切られ、各枠の大きさは、オペレーティングシステムの仕様やシステム構築時のパラメタで決定される。獲得したメモリの返却を忘ると、枠内の獲得可能な量が減少する。要求量が獲得できなければ、要求元の処理は続行できなくなる。計算機システムはパニックに陥ることもある。

獲得したメモリは、一般に、要求元の責任で返却しなければならない。オペレーティングシステムは、ジョブの終了時にそのジョブが獲得したメモリを自動的に返却する機能を備えているが、長時間にわたって走行し続けるジョブのメモリは、勝手に返却しないからである。プログラムは、以下の誤りをしばしば犯す。

- (1) 獲得したメモリを全く返却しない。
(2) 獲得したメモリを返却するが、返却量が獲得量より小さい。

本機構は、このような返却を忘れられているメモリとその要求元の情報を検出し、記録する。

2 実現方法

本機構の構成を図2に示す。1)~4)の各部を2.1~2.4に説明する。

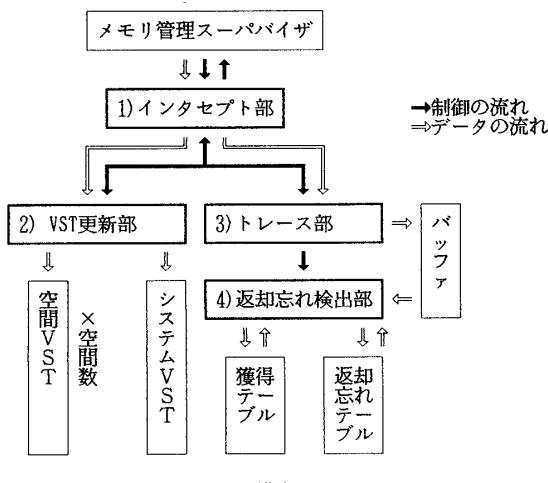


図2 各部の構成

2.1 インタセプト部

メモリを管理するスーパバイザとインターフェースを持ち、獲得／返却処理が成功するたびに、次の情報を収集する。

- 獲得／返却の区別
 - 獲得／返却アドレス
 - 獲得／返却サイズ
 - 要求元モジュール名とモジュール内オフセット
 - 要求元のジョブ名
 - 獲得／返却時刻

これらの情報をパラメタとして VST更新部とトレース部を呼び出す。本機構は、試作であること、およびオーバヘッドを分散するため、スープライザの改造は極力抑えることとした。

2.2 VST更新部

VSTは、メモリの各枠内の現時点の獲得量(U)と現時点までの獲得量のピーク値(M)を記録するテーブルであり、ジョブ別のVST(空間VST)とシステム全体のVST(システムVST)の2種類がある。空間VSTは実行中のジョブ数分だけ作成し、そのジョブの獲得量とピーク値を記録する。システムVSTはシステムに1つだけ作成し、各ジョブで共通に利用される枠の獲得量とピーク値、およびピークに達したときのジョブ名を記録する。OSTV/R4 MSP(非アドレス拡張)を例にしたVSTを図2.2に示す。

なお、0クリア可能なカラムを用意することにより、そのジョブの任意区間のピーク値も計測できるようになっている。

システム枠		ジョブ枠			
SQA	CSA	LSQA	PLSQA	REGION	J
UMC	UMC	UMC	UMC	UMC	J

システム査

Uカラム： 現時点の各領域の獲得量
 Mカラム： 現時点までの Uカラムのピーク値
 Jカラム： 空間VST はその空間のジョブ名
 Cカラム： システムVST はピークに達したときのジョブ名
 Oカラム： 0クリア可能な Mカラム

図2.2 VSTの構成

2.3 トレース部

トレース部は、インタセプト部で収集した情報のシステム領域のバッファへの時系列な記録処理とジョブ領域のバッファへの転送処理に分かれる。システム領域に獲得できるバッファには、限度があることおよびこの後で説明する返却忘れ検出のオーバヘッドがシステム運用に影響を与えないようにするために、バッファを2つに分割し、処理を切り離した。転送処理から後の処理は、メモリ管理スーパバイザと非同期に動作する。なお、どちらのバッファも一杯になると上書きし、再利用される。トレース部の構成を図2.3に示す。

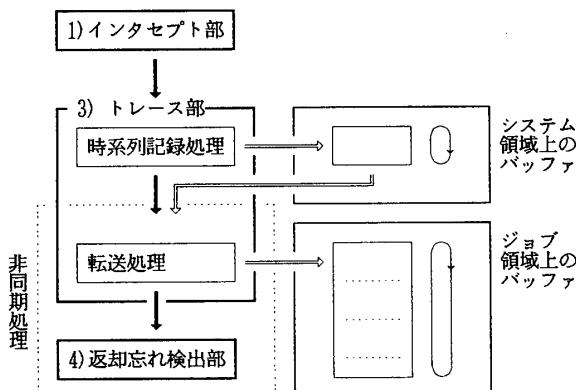


図2.3 トレース部とバッファの構成

2.4 返却忘れ検出部

ジョブ領域に転送されたトレース結果を入力し、獲得の場合には、1回の獲得に付き1つの獲得テーブルを作成する。獲得テーブルには次の情報を記録する。

- 獲得アドレス
- 獲得サイズ
- 要求元モジュール名とモジュール内オフセット
- 要求元ジョブ名
- 獲得時刻

検索オーバヘッドを小さくするため、枠とアドレスでハッシュし、かつアドレス順に連結する(図2.4-1)。返却の場合には、返却パターン(図2.4-2)に従って獲得テーブルの削除/更新/分割を行う2)。更新/分割の場合には、獲得アドレスとサイズは返却後の値とし、また、要求元モジュール名とモジュール内オフセットおよびジョブ名は、返却の要求元に置き換える。時刻は返却時刻に置き換える。こうして、獲得されたままの情報が残る。

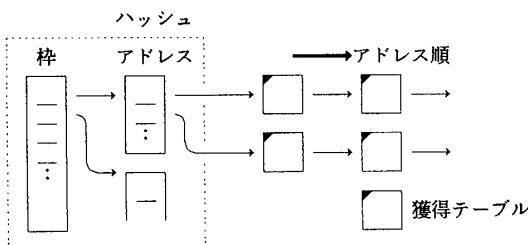


図2.4-1 獲得テーブルの管理方法

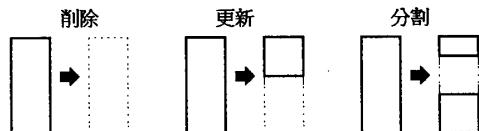


図2.4-2 返却のパターン

転送されたトレース結果の処理がすべて終わった時点で、返却忘れの候補を選び出す。返却残りのサイズが同じ、かつ、要求元が同じ、かつ、回数が多いものが原因となって、やがて領域不足につながる可能性が高いという経験に基づき、次の条件を満足する獲得テーブルを抽出する。

- (1) 更新/分割されず、かつ、要求元モジュール名とモジュール内オフセットおよび獲得サイズが同じで獲得時刻が異なる獲得テーブルが2つ以上ある獲得テーブル。ただし、今回の周期で新たに作成した獲得テーブルは、抽出の対象外である。
- (2) 更新/分割されており、かつ、要求元モジュール名とモジュール内オフセットおよび獲得(残り)サイズの同じ獲得テーブルが2つ以上ある獲得テーブル。

抽出した獲得テーブルは、返却忘れテーブルとして獲得テーブルと同様の方法で管理し、追加する。ただし、同じテーブルがすでにある場合には、追加せず、検出回数に1を足す。返却忘れテーブルは、検出回数の情報が格納されるほかは、獲得テーブルと同じである。獲得テーブルを削除/更新/分割すると、返却忘れテーブルも削除/更新分割する。

検出回数が多く、かつ、古い返却忘れテーブルが返却忘れの有力候補となる。

3 おわりに

本機構により、次の効果が得られると推測される。

- 本機構を装備したシステムで、プログラムを実行してみることにより、返却忘れないか検証できるようになる。返却忘れテーブルを時系列に、また、特定のジョブだけを選択して印刷するプログラムを作ることで、使いやすくなる。
- 返却忘れによるシステム異常終了の原因調査の時間が大幅に短縮される。
- 本機構の途中経過として作成されるVSTやトレース情報を用いた分析プログラムを作ることにより、返却忘れの検出だけでなく、ジョブやコマンド等のメモリ使用状況の分析や品質向上に役立てることができる。

参考文献

- [1] 吉本、高橋：オペレーティングシステムの性能評価と監視の効率化、情報処理学会第33回全国大会
- [2] 村井、住田、大庭、加藤：システム性能監視へのエキスパートシステムの適用、情報処理学会第35回全国大会