

# サービスの停止時間を短縮するソフトウェア実行環境のプリコピー移送方式

1180322 黒木 勇作 【分散処理 OS 研究室】

## 1 はじめに

広域分散システムの効率的な利用のため、アプリケーションプログラム (以下 AP と略す) の実行環境を別の計算機へ移送する際に用いられる代表的な手法として、仮想マシンを移送する方法がある。しかし、この手法では仮想マシン上の全ファイルを移送させるため、時間や移送サイズの面から非効率である。そこで、我々はファイルのアクセス情報を基にしたソフトウェア実行環境の移送方式を提案している。

## 2 提案方式

提案方式は AP のファイル群へのアクセス情報を監視する特定ステップと、特定したファイル群の転送を行う転送ステップの 2 ステップから構成される。

### 2.1 特定ステップ

AP が発行する open システムコールを十分な時間監視し、ファイル名とアクセス種別 (Read-Only, Read-Write) から転送対象となるファイルを決定する。この時対象と、影響を受ける他のファイルとの共有関係を追跡し、これらのファイルも一括して転送対象とする。

### 2.2 転送ステップ

AP を動作させたまま転送を行う「プリコピーフェーズ」と、AP を停止させて転送を行う「最終コピーフェーズ」の 2 段階で全ファイルの転送を行う。プリコピーフェーズでファイルを転送しておくことで AP 停止時間を短縮することができるが、転送済のファイルに対して更新が発生した場合、該当ファイルの再送が必要となる。本稿では、転送順序により最終コピー量にどのような変化が表れるかを明らかにする。

## 3 評価

### 3.1 評価項目

転送順序の違いによる転送結果への影響を、シミュレーションにより評価した。既存手法として文献 [1] の方式をファイルに適用したものをを用いる。既存手法では、反復回数が一定回数となっているため、本稿では 10 回に固定して比較を行う。Read-Only ファイル (RO と略す) と Read-Write ファイル (RW と略す) を用意し、RW は全体に 1:10:10:79 の比率でアクセスが発生する。

表 1 評価に用いたファイルデータセット

サイズ	R-O	R-W	R-O 合計	R-W 合計
200KB	350	3,500	70MB	700MB
400KB	300	3,000	120MB	1,200MB
800KB	200	2,000	160MB	1,600MB
1,600KB	100	1,000	160MB	1,600MB
3,200KB	50	500	160MB	1,600MB
合計	1,000	10,000	670MB	6,700MB

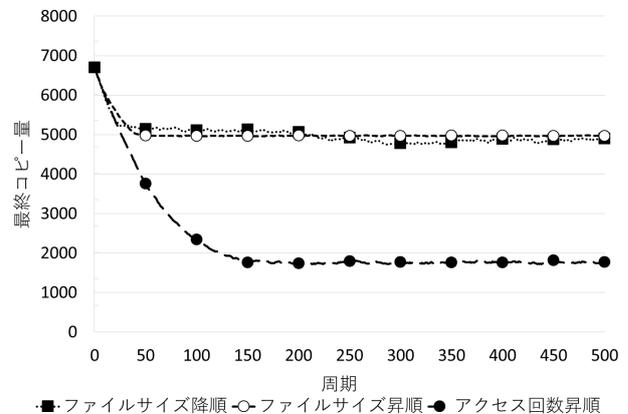


図 1 各順序の最終コピー量の推移  
転送速度は 5Mbps を想定し、Write システムコール 500 回を 1 周期として、プリコピー 500 周期で最終コピーフェーズに移行する。これを下記の転送順序により比較する。なお RO は最優先で転送するものとする。

- 既存手法
- ファイルサイズの降順 (サイズ降順と略す)
- ファイルサイズの昇順 (サイズ昇順と略す)
- アクセスされた回数の昇順 (アクセス昇順と略す)

### 3.2 評価結果と考察

評価結果を図 1 に示す。図 1 の横軸の数字はプリコピー周期数、縦軸は、最終コピー量を示している。ファイルサイズによる大きな差が見られなかったのは、再送により、プリコピー量が一定の値となった場合、プリコピー量を再送量が上回ってしまう為であると考えられる。最小最終コピー量は、既存手法が 5302MB であるのに対し、アクセス昇順が 1717MB となり、最も最終コピー量を削減することができた。これはファイルアクセスに偏りがあるため、アクセスされた回数が小さいファイルを優先的に転送することで無駄な再送を抑えつつプリコピーが行えたからであると考えられる。

## 4 おわりに

本稿では、転送順序によりファイルの転送結果にどのような影響が表れるのかを明らかにした。残された課題として、再送を抑制した状態での転送順序別の転送量の評価、実アプリケーションのファイルアクセスを用いた評価がある。

## 参考文献

- [1] 中井新太郎, 川島龍太, 斎藤彰一, 松尾啓志, “ファイルの共有関係に着目した移送するプログラムと実行環境の特定方法更新履歴に基づいたメモリページ転送順序スケジューリングによる仮想マシンライブマイグレーションの高速化”, 情報処理学会論文誌, Vol.56, No.2, pp.516-524 (2015).