

テクニカルノート

UNIX システムにおけるマイグレーション機能の有効性について

末 永 正[†] 中 山 仁^{††} 大 西 淑 雅^{††}

本論文では、大規模集合教育に使用されている UNIX システムにおけるユーザファイルの参照パターンを分析した。分析の結果から、メインフレームシステム同様 UNIX システムにおいてもマイグレーション/リコール機能によってディスクスペースの有効利用が図れることを確認した。

Efficiency of File Migration Facility on a UNIX System

TADASHI SUENAGA,[†] HITOSHI NAKAYAMA^{††} and YOSHIMASA OHNISHI^{††}

This paper describes the analysis of user-file access patterns. It is based on the data on the use of user-files on a large scale UNIX system for education. From the result of analysis, we reconfirmed that the file migration facility could also work for the availability of disk space not only on a mainframe system but on a UNIX system.

1. はじめに

近年、UNIX システムに代表される分散処理システムにおいて、ファイルを複数のマシンに分散配置することによって、各マシンの有効利用を図る方式の研究が進められている^{1),2)}。しかし、頻繁にファイルの更新が行われるユーザファイルについては、バックアップなどの管理上、集中的な配置の方が取り扱いやすい。

筆者らは、メインフレームでのマイグレーション/リコール（アクセスされそうにないファイルを低速大容量装置に移動し、必要なときに呼び戻す）機能を UNIX システムでも実現すべく、九州工業大学情報科学センターにおいて、大規模集合教育に使用されている UNIX システムでのファイル利用状況を5か月間にわたって、予備調査を行った。本稿では、これらの収集データに基づく分析結果について報告する。

2. ファイルの使用状況分析

2.1 調査対象システム

図1は調査したシステムのハードウェア構成を示したものである。使用しているソフトウェアは表1に示すとおりである。なお、このシステムには、大学在籍の学生および教員のうち、飯塚キャンパスの情報工学部所属のすべてが登録されており、その数は2381であった。また、1ユーザが作成できるファイル容量が12Mバイトに制限されている環境で調査を行った。

2.2 ファイル参照パターンの分析

図2は、ディレクトリなどの特殊ファイルを除く全ユーザファイル（約50万個）に対する日々の参照関係を調べ、ファイルの参照される平均確率が最後に参照した日からの経過日数によって変化する様子を示したものである。アクセス後の日数経過とともに、再度アクセスされる確率は急激に減少しているのが分かる。この現象はある程度予想されたものであるが、単調減少になっていないという特徴がある。調査対象システムが教育用であるために、授業時間割からくる周期性（7日ごとの増加）は7日と14日目に顕著に現れている。しかし、これより長くなるとその周期性が明確ではない。

ユーザファイルのほとんどがその所有者によってア

[†] 近畿大学九州工学部

Kinki University in Kyushu

^{††} 九州工業大学情報科学センター

Information Science Center of Kyushu Institute of Technology

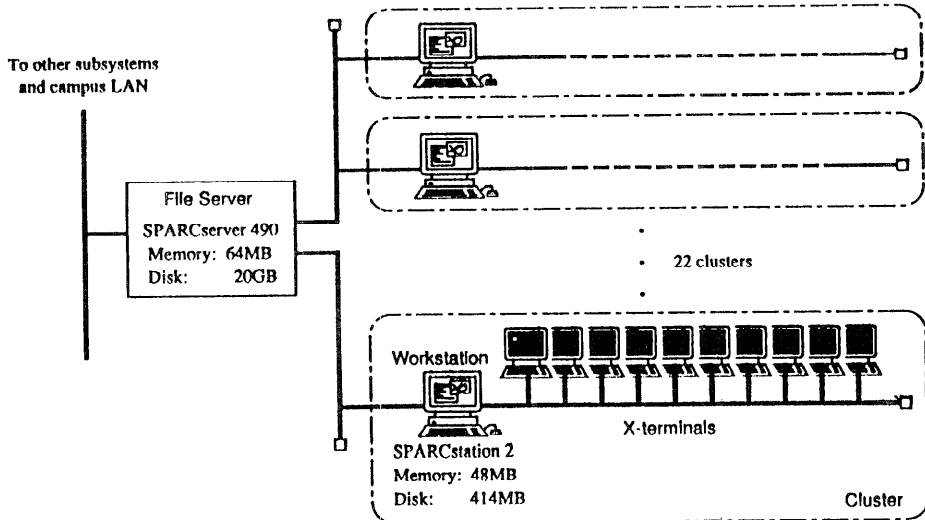


図 1 システムの概観
Fig. 1 Overview of the system.

表 1 ソフトウェア構成
Table 1 Software configuration.

ソフトウェア種類	名 称
UNIX operating system	Sun OS (ver. 4.1.1.3 JLE 1.1.3)
Window system	X (ver. 11 Release 5)
Compilers	Fortran, Pascal, C
Document tools	Jlatex, tgif+, gnuplot
Editor	Nemacs
News reader	gnus
Mail reader	mh-e
Kana-Kanji translater	Egg

アクセスされると仮定して、図 2 と、アクセスしたユーザの数との関係を調べた結果、ユーザ数との相関は低いことが分かった。このことから、図 2 では、オブジェクト生成 (make) や文字列検索 (grep) などのソフトウェアツールの支援を得た。複数のファイルへの一括アクセスによる影響も現れているように推測される。

なお、図 2 における参照確率の推移でファイルのサイズによる違いはみられなかった。

2.3 必要なディスクスペース分析

2.2 節の結果は、ファイルは使用されてから数日以内に再び使用されないかぎり、将来にわたって使用されない可能性が極めて高いことを示している。したがって、参照日から一定期間過ぎたファイルは、必要なときにアクセスさえできれば、マイグレーションされリコールに多少時間がかかるとしても、ユーザ

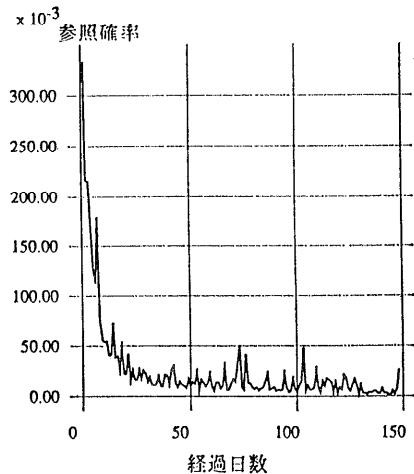


図 2 参照確率
Fig. 2 Reference probability.

に与える影響は少ないといえる。ただし、ファイル参照時のリコール処理にあたっては、複数ファイルの一括参照が効率よく行えるような配慮が必要となろう。

図 3 は、ファイルサイズを加味した LRU (Least Recently Used) 手法^{3),4)}によって、長期未参照ファイルをディスクから別の場所に移した場合に、参照時にディスク上に存在しない比率 (ミスヒット率) をシミュレートしたものである。横軸は、実ディスク上にあるファイルの記憶容量と、システムに存在するすべてのファイルの記憶容量との比を示し、縦軸にはミス

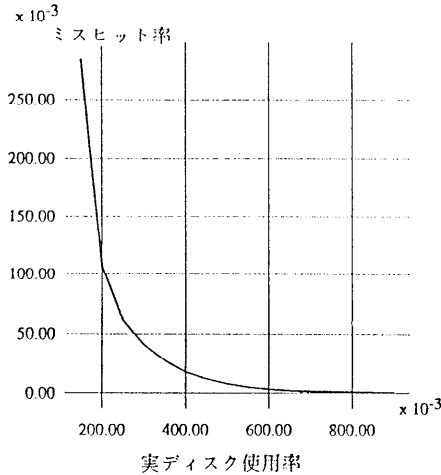


図3 ミスヒット率のシミュレーション
Fig. 3 Simulation of miss hit ratio.

ヒット率を示す。図3から、実ディスク上の記憶容量が全体の50%もあれば、ミスヒット率は極めて低く、実用上問題ないことが分かる。

3. おわりに

以上、情報処理教育のためのUNIXシステムにおいて、ユーザファイルの参照パターンを分析した。これらの結果は、メインフレームでの木構造でないファイルシステムにおけるファイル参照パターンと似た特徴を示しているが、UNIX特有のものとして、複数ファイルの一括参照による影響が現れているように思われる。筆者らはこれらの結果を踏まえて、UNIXシステムにマイグレーション/リコール機能を実装するための設計を進めており、具体化した時点であらためて報告したい。

参考文献

- 1) Liu Sheng, O. R.: Optimization of File Migration Policies in Distributed Computer Systems, *Comput. Oper. Res.*, Vol. 19, No. 5, pp. 335-351 (1992).

- 2) Hurley, R. T., Yeap, S. A., Wong, J. W. and Black, J. P.: Potential Benefits of File Migration in a Heterogeneous Distributed File System, *Proc. ICCI '93*, pp. 123-127 (1993).
- 3) Lawrie, D. H., Randal, J. M. and Rarton, R. R.: Experiments with Automatic File Migration, *Computer*, Vol. 15, No. 7, pp. 45-55 (1982).
- 4) 景川耕宇, 武富 敬, 末永 正: 2レベル階層化補助記憶システムにおける最適制御方式について, *情報処理学会論文誌*, Vol. 28, No. 5, pp. 516-524 (1987).

(平成6年9月1日受付)
(平成6年11月17日採録)



末永 正 (正会員)

1949年生。昭和47年九州大学工学部電子工学科卒業。九州大学大型計算機センター、同情報処理教育センターを経て、平成元年九州工業大学情報科学センター助教授、平成3年近畿大学九州工学部教授、現在に至る。計算機システムの性能評価に関する研究に従事。九州大学工学博士。電子情報通信学会会員。



中山 仁 (正会員)

1961年生。1986年九州大学工学部情報工学科卒業。1988年同大学院修士課程修了。同年九州工業大学情報科学センター助手。分散処理システム、関数型言語システムの研究に従事。ソフトウェア科学会会員。



大西 淑雅 (正会員)

1966年生。1989年九州工業大学工学部情報工学科卒業。同年九州工業大学情報科学センター助手。分散システム、計算機ネットワークの研究に従事。