

階層的ファイル自動管理システムの設計†

藤 井 実†† 浅 井 清††

磁気ディスク、大容量記憶装置 (MSS)、磁気テープの3種の記憶媒体を用いた、簡素で効率的なファイル管理システムとその設計に必要な一連の手法を提案する。

本論文で筆者らは、ユーザ・ファイルについて、

(1) ファイル連続参照特性、ファイル分布が人間の記憶モデルにおける記憶保存関数で非常に良く近似できることを示し、記憶保存分布密度関数を新しく定義した。これを(3)で利用した。

(2) 磁気ディスク、MSS、磁気テープの損益分岐ラインの一計算方式を示した。

(3) ファイル需要を損益分岐ラインによって各記憶媒体に振分けるなど、ファイル管理システムにおける最適装置構成を決定する一手法を示した。

(4) ファイル管理の階層化、在庫管理を応用した空き領域管理方式、損益分岐ラインを使ったファイル移動制御などを採用した、簡素で実用的、効率的なファイル自動管理の一方式を提案した。

1. はじめに

計算機で使用されるファイルは、参照頻度の高い期間はよりアクセスの速い記憶媒体に保存され、参照頻度が低くなればよりコストの安い記憶媒体に移されるのが望ましい。

今日、ユーザのファイル使用に課金制、割当制などの規制を設けていない計算機センタの多くでは、磁気ディスクのスペース管理が非常に大きな問題となっている。すなわち、TSS、バッチ・ジョブの増加に伴い、磁気ディスクのスペースはすぐに空きがなくなる傾向にあり、使用頻度の低いファイルを磁気テープに移すなどして空きスペースを確保している。しかし、これらの作業は非常に人手のかかるわずらわしいものになってきている。

本論文は、この問題を解決するため、磁気ディスク、大容量記憶システム (MSS)、磁気テープの3種の記憶媒体を用いたユーザ・ファイルの自動管理システムとその設計方法を考察したものである。筆者らが追求した要件は、(i)ファイル管理の大部分を自動化できること、(ii)簡素であること、(iii)ファイルを参照頻度と大きさに応じて最も経済的な記憶媒体に移動できること、(iv)ファイル需要に応じた最適装置構成を設定できることである。

(i)の自動化は、最近発表されたIBMのソフトウェアHSM等^{1)~3)}によっても実現されている。しかし、

HSM等は(ii)~(iv)を満たしていない。この点は8で述べる。(iii)の要件には評価基準、たとえば磁気ディスク、MSS、磁気テープの損益分岐ラインが必要である。MSSの適用領域に関しては文献4)、5)に概説されている。損益分岐ライン作成は、Lum et al.⁶⁾、藤井⁷⁾で試みられている。(iv)は、ユーザ・ファイルを対象としたものは従来経験的に決められてきているようである。

本論文は、上記の要件を満たすファイル管理システムの設計を現実のデータの上に立って追求し、以下の新しい考え方を提案した。

(1) ファイルの連続参照特性が記憶保存関数の2次の式、ファイル・サイズ分布、参照頻度分布がともに記憶保存関数の3次の式で非常によく説明できることを示し、記憶保存分布密度関数を新しく定義した。これらの結果を最適装置構成の決定に利用した。

(2) 磁気ディスク、MSS、磁気テープの損益分岐ラインの一計算方式を示した。

(3) 大きさと参照頻度の同時分布関数を定義できるファイル需要であれば、(2)で作成した損益分岐ラインによってこの需要を各記憶媒体に振分けることができるなど、ファイル管理システムにおける最適装置構成を決定する一手法を示した。

(4) ファイル管理の階層化、在庫管理を応用した空き領域管理方式、損益分岐ラインを使ったファイル移動制御などを採用した、簡素で実用的、効率的なファイル自動管理の一方式を提案した。

本論文では、主にディスク・ファイルのデータについて一連の手法を説明するが、テープ・ファイルについても同様の分析が可能であり、実際のシステムを作

† A Design of Hierarchical, Automatic File Management System by MINORU FUJII and KIYOSHI ASAI (Computing Center, Japan Atomic Energy Research Institute).

†† 日本原子力研究所計算法センタ

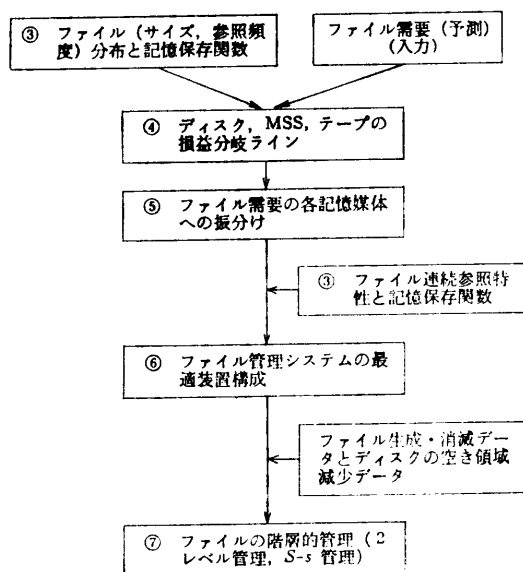


図 1 ファイル管理システムの作成過程
Fig. 1 A construction process of the file management system.

成するときには両方のデータを考慮しなければならない。

2. システム設計の概要

図 1 にファイル管理システムの作成過程を示す。図のブロック中の番号は本論文の章番号に対応する。

ファイル・サイズ分布、参照頻度分布がともに記憶保存関数で近似できることなどを使うと、ファイル需要を④で示す損益分岐ラインで磁気ディスク、MSS、磁気テープの各記憶媒体に振分けることができる。また、ファイル連続参照特性が記憶保存関数で近似できることを利用すると、MSS のステージング・ディスク* の最適台数が決定でき、先の振分けと合わせるとファイル需要に対する最適装置構成が得られる。ファイル生成・消滅データ、空き領域減少データを調べ、これらを参考にすると在庫管理⁸⁾を応用した S-s (補充点-消去点) 空き領域管理方式が設定できる。そして、通常の記憶階層間のファイル移動はユーザに任せ、磁気ディスク、MSS のどちらか一方の空き領域が下限値以下になったときに全体調整が働く 2 レベルのファイル管理を考える。この管理方法を採用すると筆者らの計算センタの例では、全体調整は半月～1 か月に 1 回程度行えばよく、損益分岐ラインを使った効率

* MSS では、ファイルはカートリッジ・テープに保管され、参照されるときに MSS 配下の一時的ディスクに移され、一時ディスク上でアクセスされる。この一時ディスクをステージング・ディスクという。

的なファイル移動制御が可能となる。この制御は、指数平滑法⁸⁾を使うことにより、各ファイルについて 2 個の収集データの追加でよく、容易に実用化できる。

3. ファイル連続参照特性と記憶保存関数、ファイル分布と記憶保存関数

一般に今日のファイルを参照するかはランダムでない。前日の仕事にひきずられて同じファイルを参照することが多い。

では、どの程度過去の参照に影響されているであろうか。もし、この参照特性がわかるならば、その特性をうまく利用することによってオーバーヘッドの少ない、簡素なファイル管理システムが設計できるはずである。

筆者らはこのファイル参照の連続性を調べている過程で、ファイル参照間隔分布が Wickelgren の示した記憶保存関数⁹⁾で非常によく近似できることを発見した。ここでは、その近似式と意味について考察する。さらに一見記憶とは何の関係もないように思えるファイル・サイズ分布、ファイル参照頻度分布もこの Wickelgren の記憶保存関数で非常によく近似できることを示す。

Wickelgren の記憶保存関数と本論文で新しく定義した記憶保存分布密度関数について先に述べる。

3.1 Wickelgren の記憶保存関数

記憶モデルの研究においては、人間を 1 つの情報処理系とみなし、図 2 に示すように記憶を短期と長期の 2 つの成分に分離する考え方が広く認められてきたようである^{10), 11)}。また、最近では記憶を連続的に取扱うために、記憶の保持量 (あるいは忘却率) と時間の関係を実験的に調べ、これを数式表現した研究なども行われている^{9), 10)}。

ここでは、Wickelgren の示した記憶保存関数に着

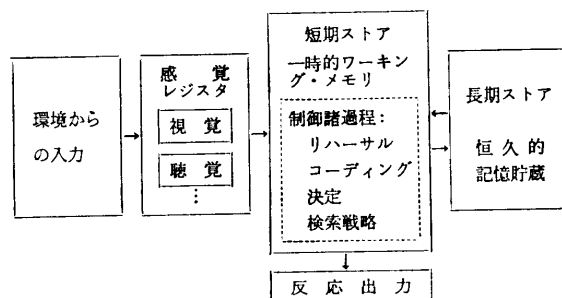


図 2 Atkinson & Shiffrin の記憶モデル^{10), 11)}
Fig. 2 The model of memory system by Atkinson & Shiffrin.

目する。彼は、忘却は時間と干渉によって起こり、記憶の痕跡は記憶の強さ (Strength) ともろさ (Fragility) という2つの動的性質によって説明づけられるとした。そして、 m : 記憶の強さ、 f : 記憶のもろさ、 π : 干渉 (Interference) 過程に対するパラメータ (音声入力等における場合、類似の新しい入力によって古い記憶は忘れ去られる)、 η : 記憶の強さの時間による減衰パラメータ、 μ : 記憶のもろさの時間による減衰パラメータ、 λ : 最初の記憶の強さ、 ρ : 最初の記憶のもろさとする、次の2つの微分方程式によって記憶現象がよく近似されると述べている。

$$\frac{dm}{dt} = -\eta f m - \pi m, m(t=0) = \lambda > 0 \quad (3.1)$$

$$\frac{df}{dt} = -\mu f^2, f(t=0) = \rho > 0 \quad (3.2)$$

(3.2) の解を (3.1) に代入すると

$$m(t) = \lambda(1 + \beta t)^{-\phi} e^{-\pi t} \quad (3.3)$$

となる。ここで、 $\beta = \mu\rho > 0, \phi = \eta/\mu > 0$ である。Wickelgren はこの式の $(1 + \beta t)^{-\phi}$ が10~20秒から2年ぐらまでを、 $e^{-\pi t}$ が10~20秒以下の記憶痕跡の減衰データをよく説明すると述べている。

3.2 記憶保存分布密度関数

筆者らは、干渉項 ($e^{-\pi t}$) を考慮しない場合の記憶保存関数 (3.4) を分布密度関数として使用することを試みた。すなわち、記憶保存曲線の形状をもつ密度関数を定義する。

$$m(t) = \lambda(1 + \beta t)^{-\phi} \quad (3.4)$$

この式が密度関数として定義できるのは、 $\int_0^\infty m(t) dt = 1$ なる正規化条件を満たすときで、これは

$$\phi > 1, \lambda = (\phi - 1) \cdot \beta \quad (3.5)$$

の関係を満たすときである。

次に (3.4) の密度関数をもつ分布の平均値 $\bar{t}$ は、 $\bar{t} = \int_0^\infty t \cdot m(t) dt$ に (3.5) を代入すると、

$$\begin{cases} \phi > 2 \text{ のとき, } \bar{t} = \frac{1}{(\phi - 2) \cdot \beta} \\ \phi \leq 2 \text{ のとき, } \bar{t} = \infty \end{cases} \quad (3.6)$$

となる。

以上より、筆者らは (3.4) に (3.5) の条件を代入し、 $n = \phi, \lambda = \beta, x = t$ と変数を置換えた (3.7) で示す $m(x)$ を n 次の記憶保存分布密度関数と定義し、文中では以後これも記憶保存関数という用語で使用する。

$$m(x) = (n - 1) \lambda (1 + \lambda x)^{-n} \quad (3.7)$$

表1 ファイルの連続参照特性と記憶保存関数
Table 1 Continuous file references and memory retention function.

直前参照日 t	全ファイル に占める割合	既存ファイル に占める割合 $\beta(t)$	指数分布近似 $e(t)$	記憶保存関数 $m(t)$
新規作成	14.4%	0.0%		
1 日前	48.5	56.7	57.3%	60.0%
2 "	10.5	12.2	18.5	11.4
3 "	5.3	6.2	10.5	6.4
4 "	3.4	4.0	5.9	4.0
5 "	2.4	2.8	3.4	2.8
6 "	1.6	1.9	1.9	2.1
7 "	1.5	1.7	1.1	1.5
8 "	1.0	1.2	0.6	1.3
9 "	1.0	1.2	0.3	1.0
10 "	0.7	0.8	0.2	0.8
11~20 日前	5.1	5.9	0.3	4.0
21 日以上前	4.6	5.4	0.0	4.7

ここで、 $n > 1$ である。 $n > 2$ のときは平均 $\bar{x} = 1/[(n - 2) \cdot \lambda]$ が成立つ。(3.7) の λ は、先の記憶の強さとは関係のないパラメータである。

3.3 ファイル連続参照特性と記憶保存関数

ファイル参照の連続性をみるためにファイル参照間隔日数について調べた。すなわち、今日参照されたファイルの直前の参照は何日前であったかを調べた。筆者らの計算センタにおける昭和53年11月の約2万件のデータより表1の結果を得た。

これを整理すると、今日新規作成されたファイルと既存ファイルとの割合は、

$$\begin{cases} \text{新規作成 (NEW)} & \alpha = 14.4\% \\ \text{既存 (OLD)} & 1 - \alpha = 85.6\% \end{cases}$$

である。このうち、既存ファイル全体を100%とし、ファイルが何日前に参照されたものであるかを示す確率密度 $\beta(t)$ とこれを指数分布で近似させた $e(t)$ を表1に示す。

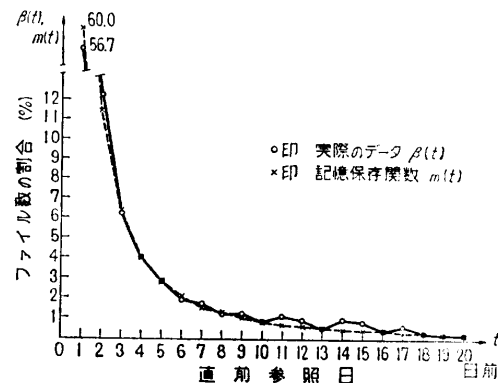


図3 ファイル連続参照分布と記憶保存関数
Fig. 3 Distribution of continuous file references and memory retention function.

もし、今日のファイルを参照するかを選択がランダムであれば、ファイルの参照間隔分布は指数分布となるはずである。しかし、データ $\beta(t)$ と指数分布で近似した $e(t)$ は、 t が大きくなるとまったく合わない。

ところが、ファイル参照と人間の記憶参照の類似性に着目し、記憶に関する文献の中から見つけた Wickelgren の記憶保存関数を適用したところ非常によい近似が得られた。

(3.7) で示した記憶保存関数の 2 次の式を使って、次のように近似式 $m(t)$ を定義する。

$$\begin{cases} m(t) = \int_{t-0.5}^{t+0.5} (1+t)^{-2} dt & (t \geq 2) \\ m(1) = \int_0^{1.5} (1+t)^{-2} dt \end{cases} \quad (3.8)$$

表 1、図 3 に示すようにこの $m(t)$ がファイル参照間隔の分布を非常によく近似する。しかし、(3.8) からわかるように $t \geq 2$ においては記憶保存関数の 2 次の式によく合うが、 $t=1$ (前日参照) の割合は $\int_{0.5}^{1.5} (1+t)^{-2} dt$ でなく、 $\int_0^{1.5} (1+t)^{-2} dt$ で近似できる値となっている。これを記憶モデルに対比させると、

前日参照——短期記憶 (連続参照)
2 日以上前の参照——長期記憶 (非連続参照)

と考えられる。前日参照ファイルはユーザの記憶の作業領域 (短期記憶) のワクにとどまっていると考えられ、直前参照日が t 日前 ($t \geq 2$) のファイルの参照割合は、一度長期記憶として記憶された事象の t 日後の記憶保存割合に等しいと考えられる。

このファイル参照間隔の分布特性をファイルの連続参照特性とよぶと、これは 1 つの記憶保存関数で表現できる。これを 6.1 で使う。

3.4 ファイル・サイズ分布と記憶保存関数

図 4 に筆者らの計算センタにおける磁気ディスク・ファイルのファイル・サイズ分布を指数分布密度関数と (3.7) の 3 次の記憶保存関数で近似した曲線を示す。ファイル・サイズ (x) の平均は $1/\lambda_1 = 31 \times 13 \text{ kB}$ (キロバイト) である。

$$\text{指数分布密度関数: } f_0(x) = \lambda_1 e^{-\lambda_1 x} \quad (3.9)$$

$$\text{記憶保存関数: } f_1(x) = 2\lambda_1(1+\lambda_1 x)^{-3} \quad (3.10)$$

図 4 に示すように指数関数で近似すると、特にファイル・サイズの大きいところでは全く適合しない。ところが、(3.10) の記憶保存関数を適用すると、 10^{-1} ~

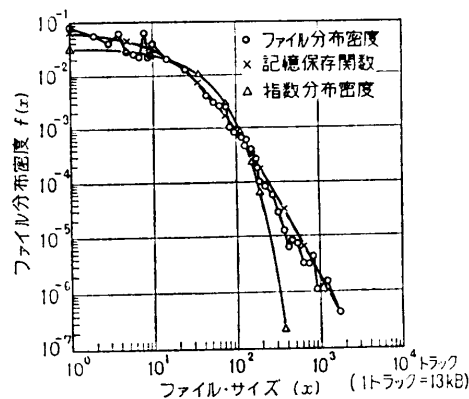


図 4 ファイル・サイズ分布と記憶保存関数
Fig. 4 File size distribution and memory retention function.

10^{-6} の範囲にわたって非常によく近似できる。

3.5 ファイル参照頻度分布と記憶保存関数

図 5 にファイル参照頻度分布と記憶保存関数による近似曲線を示す。ファイル参照頻度 (y) の平均は $1/\lambda_2 = 10.86$ 回/月 である。

$$\text{記憶保存関数: } f_2(y) = 2\lambda_2(1+\lambda_2 y)^{-3} \quad (3.11)$$

3.6 ファイル・サイズとファイル参照頻度の関係

図 6 に筆者らの計算センタにおけるファイル・サイズと各ファイル・サイズごとに算出した月平均ファイル参照回数の関係 (約 1 万件のデータより) を示す。 $\bar{y}(x)$ をファイル・サイズ x (kB) のファイルの月平均ファイル参照回数とし、この関係を次の式で近似した。

$$\bar{y}(x) = \frac{1}{\lambda_2(x)} = 3 + 7 \log_{10} x \quad (3.12)$$

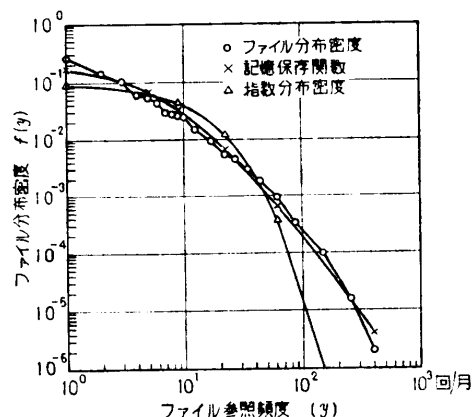


図 5 ファイル参照頻度分布と記憶保存関数
Fig. 5 File reference distribution and memory retention function.

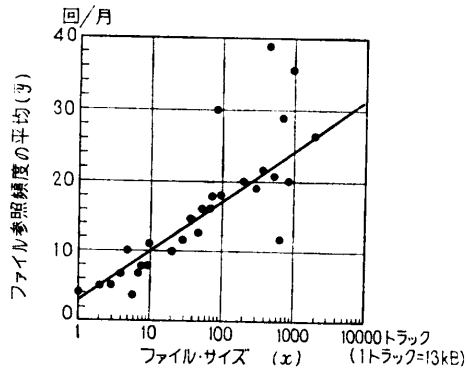


図 6 ファイル・サイズと参照頻度
Fig. 6 File size and file references.

4. 磁気ディスク, MSS, 磁気テープの損益分岐ライン

4.1 評価方法と前提条件

ファイル (サイズ = x kB, 参照頻度 = y 回/月) の 1 か月あたりにかかる総費用を各記憶媒体ごとに計算し, 任意の 2 つの記憶媒体における総費用が等しくなる参照頻度曲線 $y = f(x)$ を求める. この曲線がその 2 つの記憶媒体の損益分岐ラインとなる.

評価に用いた主な前提を次に示す.

(1) ファイル属性 (シーケンシャル, ランダム), 参照形態 (オンライン, オフライン) の各比率は入力パラメータとする.

(2) テープにファイルを格納する場合, 1 ファイルが 1 本のテープを占有するものとする.

(3) テープに格納されているランダム・アクセス・ファイルを参照する場合は, 一度作業用ディスクに移してから参照し, そのファイルは 1 参照あたり 1 日作業用ディスクを使用するものとする.

(4) MSS は IBM 型の MSS¹²⁾ を対象とする. MSS の装置費は, 大容量記憶装置 (MSF), ステージング・ディスク・コントローラ (SDC), ステージング・ディスクの装置費を含む. これらの装置構成は任意に設定できるが, この損益分岐ライン算出ではこれらの装置割合を変数としていないため, 設定した装置構成 (たとえば MSF = 35 GB, SDC 1 台, ステージング・ディスク = 800 MB) から付録の入力データ, すなわち MSS 装置費 F_m (円/kB・月), MSF からファイルをステージングする割合 β_m , 平均待ち時間を考慮した MSF のアクセス時間 A_m (秒) を作成しなければならない.

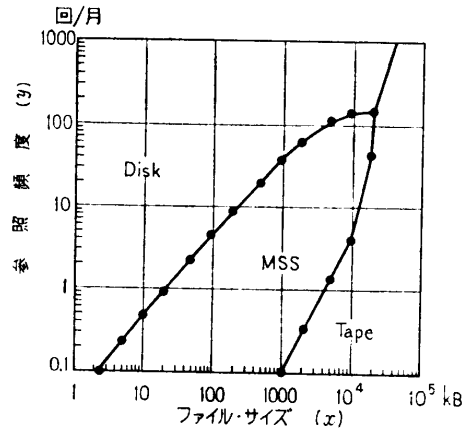


図 7 ディスク, MSS, テープの損益分岐ライン
Fig. 7 Break-even lines for disk, MSS, tape.

4.2 評価項目

各記憶媒体の総費用を算出するために考慮した評価項目を以下に示す. 損益分岐ラインの算出方法は付録, 類似の算出は文献 7) に詳述されている.

(1) 装置費: ディスク, MSS はファイル保存費, テープは装置使用費である. 各装置の使用単価は, 制御装置も含めた (装置借料 + 建屋費 + 電力費) に平均占有率, 平均稼働率を加味して算出される.

(2) 計算機使用費: ファイル Read/Write 時間に対するチャネル費, 計算機使用時間 (R/W 時間, 実行時間*) に対するシステム費 1 の 2 項目からなる. システム費 1 は, ファイル参照における計算機使用時間に対し, 演算装置費, 主記憶装置費等について付録 1 に示す評価係数を乗じて負担させるものである.

(3) 計算機使用機会損失費: ファイル・アクセス時間**によって生じる計算機使用機会損失に対する費用をシステム費 2 とした.

(4) ユーザの仕事遅れ損失費: ファイル参照におけるユーザの待ち時間を評価したものである. オンライン参照, オフライン参照における仕事遅れ損失費をそれぞれユーザ費 1, ユーザ費 2 とした.

(5) テープ関係費: テープ・ファイルの参照にのみかかる費用で, 作業用ディスク費, テープ費, オペレータ費である.

4.3 算出結果

図 7 に付録の算出方法で計算した筆者らの計算センタにおける損益分岐ラインの一例を示す. 図の損益分

* テープの場合, ランダム・アクセス・ファイルをディスクに移すときにかかるジョブ実行時間.

** ディスクはシーク時間 + サーチ時間. MSS はステージング時間が加わる. テープはハンドリング時間である.

岐ラインは,

$$y_1 = \frac{1.604x}{0.008x + 34.9} \quad (\text{ディスク対MSS})$$

$$y_2 = \frac{0.193x - 116.0}{-0.039x + 866.0} \quad (\text{MSS対テープ})$$

$$y_3 = \frac{1.797x - 116.0}{-0.031x + 899.4} \quad (\text{ディスク対テープ})$$

である。テープ対MSS, テープ対ディスクの損益分岐ラインは, ともにテープの一番有利な使い方(シーケンシャル・アクセス, オフライン参照)の場合である。

5. ファイル需要の各記憶媒体への振分け

ユーザのファイル需要に対し, ディスク装置, MSS, テープ装置をどの程度用意すればよいかという最適装置構成の決定は重要な問題である。ここでは, 大きさと参照頻度の同時分布関数が定義できるファイル需要を4で作成した損益分岐ラインで各記憶媒体に振分ける方法を示す。

5.1 振分け方法

まず, ファイルの大きさと参照頻度の同時分布密度関数を定義する。この同時分布密度関数の定義は, 以下に筆者らの計算センタの例を使って説明する。

3で示したように, ファイル・サイズ分布, ファイル参照頻度分布の密度関数は, それぞれ

$$f_1(x) = 2\lambda_1(1 + \lambda_1 x)^{-3} \quad (5.1)$$

$$f_2(y) = 2\lambda_2(1 + \lambda_2 y)^{-3} \quad (5.2)$$

で表わされた。また, ファイル・サイズと月平均ファイル参照回数に関係も

$$1/\lambda_2(x) = \bar{y}(x) = 3 + 7 \log_{10} x \quad (5.3)$$

で表わされた。

3では, ファイル・サイズと参照頻度の関係は無視して(5.1), (5.2)を求めたが, 筆者らの計算センタのデータでは, 各ファイル・サイズごとのファイル参照頻度分布も(5.2)で近似できる。したがって, ファイル・サイズ x と参照頻度 y に関する同時分布密度関数 $f(x, y)$ は, $f_2(y|x)$ をファイル・サイズ x のファイル参照頻度分布の密度関数とすると

$$\begin{aligned} f(x, y) &= f_1(x) \cdot f_2(y|x) \\ &= 4\lambda_1 \cdot \lambda_2(x) (1 + \lambda_1 x)^{-3} (1 + \lambda_2(x) \cdot y)^{-3} \end{aligned} \quad (5.4)$$

で近似できる。

次に振分け方法を述べる。図8に示すように, 同時分布密度関数 $f(x, y)$ が定義できるファイル需要に4で作成した損益分岐ラインをあてはめると, ディスク, MSS, テープの各記憶媒体に振分けられるべきフ

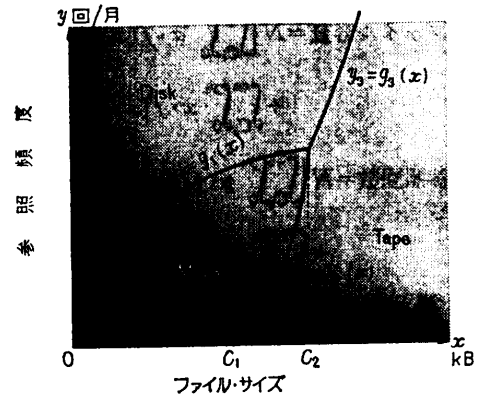


図8 損益分岐ラインによるファイル需要の振り分け
Fig. 8 File demands and break-even lines.

ファイル数, 総ファイル容量, 総参照回数が次の式で求められる。ここで, N は総ファイル数である。

(1) ディスク需要量

$$\begin{aligned} \text{ファイル数} &= N \left\{ \int_0^{C_1} \int_{y_1}^{\infty} f(x, y) dy dx \right. \\ &\quad \left. + \int_{C_1}^{\infty} \int_{y_3}^{\infty} f(x, y) dy dx \right\} \end{aligned}$$

$$\begin{aligned} \text{総ファイル容量} &= N \left\{ \int_0^{C_1} \int_{y_1}^{\infty} x \cdot f(x, y) dy dx \right. \\ &\quad \left. + \int_{C_1}^{\infty} \int_{y_3}^{\infty} x \cdot f(x, y) dy dx \right\} \end{aligned}$$

$$\begin{aligned} \text{総参照回数} &= N \left\{ \int_0^{C_1} \int_{y_1}^{\infty} y \cdot f(x, y) dy dx \right. \\ &\quad \left. + \int_{C_1}^{\infty} \int_{y_3}^{\infty} y \cdot f(x, y) dy dx \right\} \end{aligned}$$

(2) MSS 需要量

$$\begin{aligned} \text{ファイル数} &= N \left\{ \int_0^{C_1} \int_0^{y_1} f(x, y) dy dx \right. \\ &\quad \left. + \int_{C_1}^{C_2} \int_{y_2}^{y_1} f(x, y) dy dx \right\} \end{aligned}$$

$$\begin{aligned} \text{総ファイル容量} &= N \left\{ \int_0^{C_1} \int_0^{y_1} x \cdot f(x, y) dy dx \right. \\ &\quad \left. + \int_{C_1}^{C_2} \int_{y_2}^{y_1} x \cdot f(x, y) dy dx \right\} \end{aligned}$$

$$\begin{aligned} \text{総参照回数} &= N \left\{ \int_0^{C_1} \int_0^{y_1} y \cdot f(x, y) dy dx \right. \\ &\quad \left. + \int_{C_1}^{C_2} \int_{y_2}^{y_1} y \cdot f(x, y) dy dx \right\} \end{aligned}$$

(3) テープ需要量

$$\begin{aligned} \text{ファイル数} &= N \left\{ \int_{C_1}^{C_2} \int_0^{y_1} f(x, y) dy dx \right. \\ &\quad \left. + \int_{C_2}^{\infty} \int_0^{y_3} f(x, y) dy dx \right\} \end{aligned}$$

$$\text{総ファイル容量} = N \left\{ \int_{C_1}^{C_2} \int_0^{y_1} x \cdot f(x, y) dy dx + \int_{C_2}^{\infty} \int_0^{y_1} x \cdot f(x, y) dy dx \right\}$$

$$\text{総参照回数} = N \left\{ \int_{C_1}^{C_2} \int_0^{y_1} y \cdot f(x, y) dy dx + \int_{C_2}^{\infty} \int_0^{y_1} y \cdot f(x, y) dy dx \right\}$$

5.2 算出例

筆者らの計算センタにおける現在のディスク・ファイルに関するデータをケース1とし、ケース1に対してファイル数を4倍、平均サイズを2倍、1ファイルあたりの平均参照頻度を約1/2倍とした予測例をケース2とする。それぞれの振分け結果を以下に示す。

(1) ケース1 (現在例: $N=10,000$, $1/\lambda_1=400$ kB, $\bar{y}=3+7 \log_{10} x$)

	ディスク	MSS	テープ	合計
ファイル数	3,481	6,476	43	10,000
総容量 (MB)	559	3,116	325	4,000
総参照 (回/月)	75,447	28,474	125	104,046

(2) ケース2 (予測例: $N=40,000$, $1/\lambda_1=800$ kB, $\bar{y}=1.5+3.5 \log_{10} x$)

	ディスク	MSS	テープ	合計
ファイル数	6,093	33,064	843	40,000
総容量 (MB)	1,146	24,763	6,091	32,000
総参照 (回/月)	116,538	132,692	1,344	250,574

6. ファイル管理システムの最適装置構成

5で述べたファイル需要の振分けによって、ディスク、MSS、テープの各記憶媒体に保存すべきファイルの数、総容量、総参照頻度が得られる。この結果から、ファイル需要に対する最適なディスク装置台数、MSSの総容量、テープ装置台数が決定でき、最適装置構成にはMSSのステージング・ディスク (以後SDISKと略称する) 台数の決定が残る。

ここでは、損益分岐ラインによる振分けによっては評価できなかったMSSのSDISKの最適台数の決定法について検討する。

MSSの代表的な制御方式には仮想ディスク方式と仮想ファイル方式がある。前者の方式はIBM、富士通、日立のMSSに採用され、SDISKと一般のディスクを区別し、ページ (2MB) 単位で領域の割付けを行う。後者の方式は、CDC、日電のMSSに採用され、SDISKと一般のディスクを区別せず、ファイルの大きさに応じてSDISK上に領域をとる。

本来は損益分岐ラインによってMSSに振分けたファイル需要に対して、最適なSDISKの台数を求めるべきである。しかし、筆者らの計算センタはMSSを導入していないため、MSS上のファイルに関する連続参照データを持っていない。このため、3で示したディスク・ファイルの連続参照を使って、後者のMSS方式を採用する場合について、最適SDISK台数を求める方法を次節に示す。前者の方式におけるSDISKの最適台数は、MSSファイルの連続参照特性が得られれば同様の方法で決定できよう。

6.1 MSSのSDISKの最適台数

MSSのSDISK領域をうまく利用すると、MSSに保存されているファイルを参照するとき、毎回MSSからステージングを行う必要はなくなる。これは参照するファイルがすでにSDISK上に存在すれば、ステージングする必要がなくなるからである。しかし、ステージングの回数を減らし、計算機システムとユーザの待ち時間の減少をはかるとSDISKの装置費が増加する。この損益分岐点、すなわちSDISK装置費とステージングによって生じる計算機システム、ユーザの仕事遅れ損失費の総和を最小とするSDISKの台数を求める。

ここでは、一般のディスクとMSSのSDISKを区別しないため、5でディスクとMSSに振分けたファイル需要をすべてMSSに保存した場合について、SDISKの最適台数 (容量) の算出法を示す。

x : SDISKの総容量 (kB), s : MSSのファイルの平均サイズ (kB), k : MSSのファイルへの総参照回数 (回/日), n : あるファイルが1度参照されたときその日につづけて参照される平均回数 (回/日) とすると、1日平均参照ファイル数 d は、

$$d = k/n \quad (\text{ファイル/日}) \quad (6.1)$$

となる。以下において、 i 日前に参照したファイルまでがSDISK上に残っていると仮定する。このときのSDISKの必要容量 x は、図9より

	α	$1-\alpha$				
今日参照したファイル	NEW	前日参照分 β_1	β_2	β_3	β_4	β_5
1日前に "	NEW	" β_1	β_2	β_3	β_4	β_5
2日前に "	NEW	" β_1	β_2	β_3	β_4	β_5
3日前に "	NEW	" β_1	β_2	β_3	β_4	β_5

図9 ステージング・ディスク必要容量 ($i=3$ の場合、斜線部)

Fig. 9 A necessary capacity of staging disks. (ex. the shaded part is for $i=3$ case)

$$x = s \cdot d \{ (i+1) - \sum_{t=1}^i (i+1-t) \cdot \beta(t) \} \quad (6.2)$$

である。ここで、 $\beta(t) = (1-\alpha) \cdot m(t)$ 、 α は新規作成ファイルの割合、 $m(t)$ は(3.8)の記憶保存関数である。

SDISK 上のファイルに関して LRU 制御を行うと、ファイル参照においてステージングの行われる確率は、 $\left(\sum_{t=i+1}^{\infty} \beta_t \right) / n$ である。このときの総費用 TC (円/日) は、

$$\begin{aligned} TC &= K_1'(x) + K_2(i) + K_3(i) \\ &= K_1(i) + K_2(i) + K_3(i) \end{aligned} \quad (6.3)$$

と表われる。ここで、 $K_1'(x)$ 、 $K_1(i)$ はSDISKの装置費、 $K_2(i)$ 、 $K_3(i)$ はそれぞれステージング時間による計算機システムの使用機会損失費、ユーザの仕事遅れ損失費である。 $K_1(i)$ 、 $K_2(i)$ 、 $K_3(i)$ は、

$$\begin{aligned} K_1(i) &= K_1'(x) = a_1 x \\ &= a_1 [s \cdot d \{ (i+1) - \sum_{t=1}^i (i+1-t) \cdot \beta(t) \}] \end{aligned}$$

$$K_2(i) = a_2 \cdot TST, K_3(i) = a_3 \cdot TST \quad (6.4)$$

で求められる。ここで、 TST は総ステージング時間(秒/日)、 a_1 はSDISKの単位領域あたりの装置費(円/日・kB)、 a_2 、 a_3 はそれぞれ単位時間あたりの計算機システムの使用機会損失費(円/秒)、ユーザの仕事遅れ損失費(円/秒)である。 TST の算出方法については文献13)に示されている。

(6.3)の総費用 TC を最小とする i を求め、(6.2)に代入することにより、SDISKの最適容量を求めることができる。

6.2 算出例

5.2のケース1で示した筆者らの計算センタのディスク・ファイル参照データ ($s=400$ kB, $k=4,000$ 回/日, $n=4$ 回/日) の場合、前節で示した方法で最適SDISK容量を算出すると $i=0.6$ で 513 MB となる。

5.2のケース2(予測例)でディスクとMSSに振分けたファイル需要 ($s=660$ kB, $k=10,000$ 回/日, $n=2.5$ 回/日) の場合、最適SDISKの容量は $i=0.8$ で 3,591 MB となった。

これらは、CDC、日電のMSS(仮想方式)の場合の算出で、ステージングにおける中央演算処理装置のオーバーヘッド費を考慮していないため、最適SDISKの量がかなり少なく出た。しかし、5.2で示した結果と合わせて考えると、参照頻度の低いMSSに振分けられたファイルは、現在のところまだ毎日ステージングした方が経済的に有利であると解釈できる。またこれは、IBM型のMSSにおけるSDISKの台数もその

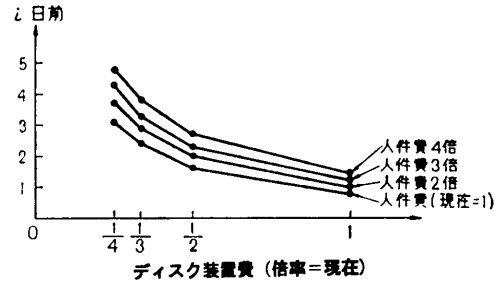


図10 ディスク装置費、人件費の変化とステージング・ディスクの最適容量

Fig. 10 Relationship between optimal staging capacity, human cost and disk cost.

日に参照されるファイル容量を満たす程度で経済的に有利であることを意味する。

図10に前述のケース2において、ディスク装置の単位領域あたりのコストを現在より安くし、ユーザの仕事遅れ損失費の算出に使う人件費を現在より高くしたときの i の値の変化を示す。たとえば、ディスク装置が1/3倍と安くなり、人件費が2倍になった場合、最適SDISK容量は $i=2.9$ で 5,153 MB となる。これは近い将来、ステージングにかかる待ち時間もコスト的に無視できなくなり、SDISKの容量を増やして数日前に参照したファイルもSDISK上に残しておいた方が経済的に有利となるかもしれないことを示唆している。

7. ファイルの階層的管理

7.1 2レベル・ファイル管理の概要

筆者らの提案するファイル管理の概要を図11に示す。通常のファイル管理は各記憶媒体ごとに行われ、ファイル作成媒体の選択や記憶媒体間のファイル移動はユーザに任せ、計算センタは一定の運用基準を提示し、ファイル移動用のTSSコマンド等を用意するだけである。

ディスクまたはMSSのどちらか一方の空き領域が、ジョブ処理に影響を及ぼすと予想される管理点以下に減少したときに、ファイル管理統合部がファイル移動の全体調整を行う。

7.2 S-s(補充点-消去点) 空き領域管理方式

在庫管理のS-s方式⁸⁾は、一定時間間隔で在庫量を調べ、そのときの在庫量 u が事前に定めた在庫の下限 s より少ない場合に $S-u$ だけ発注するように最大最小の在庫 S と s を決める発注方式である。

本論文のファイル管理統合部が行うS-s空き領域管理は、ディスクの場合、1日のジョブ処理終了時に

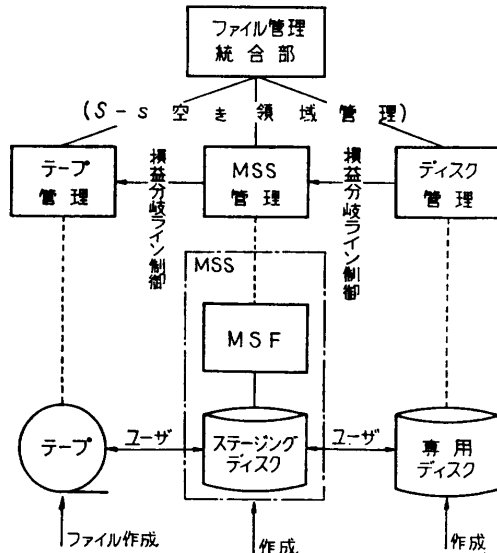


図 11 ファイルの階層的管理
Fig. 11 Hierarchical file management system.

空き領域量を調べ、空き量の残量が下限 s より少なくなっている場合にのみ次のファイル移動を行うユーティリティ・プログラムを起動する。起動されたプログラムは、ディスク上の領域管理テーブルから必要なデータを収集し、7.4 で述べる参照頻度得点を計算する。そして、参照頻度得点の小さいファイルから順に図 12 の $C1$ から $A2$ の点までのファイル量をディスクから MSS に移動させる。以上の操作は、最近採用されはじめている計算機無人運転の機能を利用すると容易に自動化できる。

$B1$ から $C1$ の点までの時間は、下限を越えてからファイル移動が実施されるまでの時間でリード・タイムと呼ばれる。ディスクの場合のリード・タイムは、空き領域量の監視を 1 日 1 回行う場合は最大 1 日となる。 s はこのリード・タイム間の空き領域の減少量も考慮して決定されなければならない。 S は次節で述べる最適問題の解から決定される。

7.3 最適消去量

図 13 の 1 回の消去量 v (MB) を増やすとディスクの有効利用度が落ちる。逆に 1 回の消去量を減らすと消去回数が増える。図 13 の斜線部で示すディスク使用機会損失費用とファイル消去作業費用の和がある単位期間において最小となる 1 回の最適消去量を求める。

ファイル増加量 (生成量 - 消滅量) の平均を d (MB/日) とすると、ファイル消去作業間隔 T (日) は、

$$T = v/d \quad (7.1)$$

となる。ディスクの使用機会損失費を c (円/MB・日)

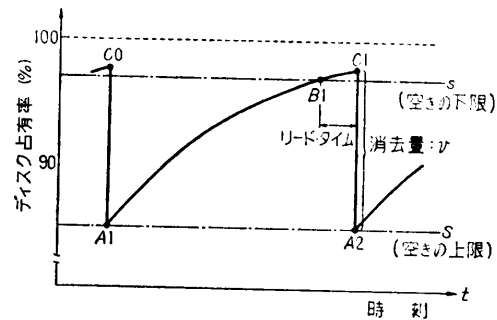


図 12 S-s 空き領域量管理方式 (ディスクの場合)
Fig. 12 S-s method for free space management.

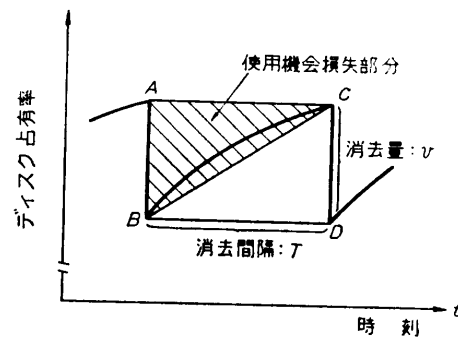


図 13 最適消去量
Fig. 13 Optimal amount for removal.

とし、ディスク占有率の増加曲線 $\widehat{BC}$ を直線 $\overline{BC}$ で近似すると、図の斜線部で示されるディスクの使用機会損失の 1 日あたりの費用 L (円/日) は、

$$L = c \cdot v / 2 \quad (7.2)$$

となる。1 回のファイル消去作業にかかる費用を $a + bv$ (円, a : 固定費, bv : 変動費) とすると、1 日あたりの総費用 G は、

$$G = (a + bv) / T + cv / 2 \\ = (a + bv) \cdot (d/v) + cv / 2 \quad (7.3)$$

となる。 G を最小にする最適消去量 v^* は、(7.3) より

$$v^* = \sqrt{\frac{2ad}{c}} \quad (7.4)$$

と計算される。ゆえに $S-s$ 空き領域管理の空き領域量の上限 S (補充点) は、

$$S = s - \Delta d + v^* \quad (7.5)$$

で求められる。ここで、 Δd はリード・タイム間のファイル増加量である。

7.4 損益分岐ライン制御と参照得点の管理

ファイル管理統合部の行う記憶媒体間のファイル移動は、損益分岐ラインを使って以下に示す方法で算出

される参照頻度得点 K_i の小さいファイルから順に行われる。

今消去期*(t)の経過時間 T_i (月)
損益分岐ラインの形状 $y=f(x)$

とする。各ファイルについて、

ファイル・サイズ x (kB)
今期内の参照回数 y_i (回)
前期までの参照頻度得点 K_{i-1}

のデータから

今期の参照得点 k_i
今期の参照頻度得点 K_i

を求める。

サイズ x のファイルがアクセスの速い上位の記憶媒体に格納され、そこから参照された方が経済的に有利であるのは、期間 T_i に次の基準参照回数 y_i^* 以上参照される場合である。

$$y_i^*(x) = T_i \cdot f(x) \quad (7.6)$$

今期の参照得点 k_i は、各ファイルの今期の参照回数がこの基準参照回数の何倍であるかを示すもので、

$$k_i = y_i / y_i^* \quad (7.7)$$

で求められる。

ファイル移動の判断基準として使う参照頻度得点 K_i の算出には、過去の参照データをできるだけ多く考慮するのが望ましい。しかし、保存データが多くなると実用的でなくなる。そこで保存データの集約を指数平滑法⁸⁾を用いて次のように行う。

$$\begin{aligned} K_i &= \alpha \cdot k_i + (1-\alpha) \alpha \cdot k_{i-1} + (1-\alpha)^2 \alpha \cdot k_{i-2} + \dots \\ &= \alpha \cdot y_i / y_i^* + (1-\alpha) \cdot K_{i-1} \end{aligned} \quad (7.8)$$

ここで、 α は指数平滑法の重み定数 ($0 \leq \alpha \leq 1$) である。

ゆえに K_i を使ったファイル移動制御 (これを損益分岐ライン制御と呼ぶ) は、各ファイルについて前期の参照頻度得点 K_{i-1} と今期内の参照回数 y_i の2つのデータを追加収集すれば実行できる。

なお、全体のファイル参照が増え、この参照頻度得点が1以上であるファイルが下位の記憶媒体に移動されるようになれば、上位の記憶媒体の増量が経済的に正当化できる。

8. 本管理方式と他管理方式との比較

本論文は、筆者らの計算センタで必要に迫られ、数

年前から設計してきたものであるが、最近同じ目的で作成された IBM の HSM¹⁾、富士通の ADF^{2),3)} 等の階層的ファイル自動管理システムが発表されている。ここでは、本論文の管理方式とこれらの管理方式との長短を ADF (HSM の機能、処理方式と同様) と対比させて述べる。なお HSM, ADF の管理方式を A 方式と略述する。

[本方式の長所]

① ファイル移動の基準にファイルの大きさと参照頻度の両方を考慮している点

A方式は、ファイルの大きさに関係なく LRU 方式でファイルを MSS に移動させる。このため費用効果比からみると本来テープで使用した方が望ましいような大きなファイルでさえも1回の参照でディスク上に居残る。

② 簡素でオーバーヘッドの少ない点

A方式では MSS に移動 (migration) されたファイルは、MSS から直接参照できず、MSS のステージング・ディスクから一度実ディスク**にもどして (recall), 実ディスクから参照しなければならない。本方式では MSS 上のファイルは MSS のステージング・ディスクから直接参照でき、オーバーヘッドが少ない。

③ 独自の考慮を示している点

最適装置構成、パラメータ設定法等についても示しており、A方式を採用する場合にも有効と思われる。

[本方式の短所]

① MSS から実ディスクへのファイル移動は自動化せず、ユーザに任せている点

A方式では MSS から直接ファイルを参照できないので、実ディスクへのファイル移動は自動化されている。

本方式は、MSS に移されたファイルは MSS から直接参照でき、MSS から実ディスクへのファイル移動は必要ならばユーザが自分で行う。したがって MSS から実ディスクへのファイル移動については、ファイル移動専用の TSS コマンド等を用意するにとどめている。

② ファイルの自動バック・アップ、世代別管理等の機能を含まない点

本方式では、これらの機能は別のユーティリティ・プログラムに任せる考えである。

9. む す び

簡素で効率的な新しいファイル管理方式とその設計

* 前回のファイル移動の行われたときから今回のファイル移動の行われるときまでの期間。

** ステージング・ディスクに対比させた用語で、普通の磁気ディスクのことである。

に必要な一連の手法を提案した。これらはファイル参照に関する現実のデータが記憶保存関数を使うとよく説明できるという事実を利用しているため、実用的にも意味あるものとなっている。

3.3~3.5 で述べた記憶保存関数によく適合したデータは、ユーザ・ファイル作成、参照に対して比較的規制のゆるやかな筆者らの計算センタで得られたデータで、特殊なケースかもしれない。しかし、見方によっては規制のゆるいのかえって自然現象的法則性のようなものがあらわれたのかもしれない。筆者らはこれらのデータを合理的に説明するために次の仮説をたて、ひきつづき調査を行っている。

(i) 制約のない条件下における人間行動のある種のもの、人間の記憶保存特性により影響を受けている。

(ii) ファイル参照行動はランダムに行われるのではなく、人間の記憶保存特性に関連して行われている。

これらは大胆な仮説であるが、興味をもたれた読者の追証を期待する。

筆者らはこの数年間、ファイル関係データを分析してきた経験から、ここで提案したファイル管理方式は、ユーザの計算機利用が比較的長期間である大型科学技術計算センタにおいて実用上有効であると感じている。半年後に予定している MSS の導入を待ってこの方式の実用化を試みるつもりである。

この研究を暖かく見守っていただいた平川室長以下協力いただいた計算センタのみなさんに感謝します。また、本論文の記述方法等に関し適切な助言をいただいた査読者に感謝します。

参 考 文 献

- 1) IBM: OS/VS2 MVS Hierarchical Storage Manager: General Information, IBM マニュアル, GH 35-0007-2 (Nov. 1978).
- 2) 富士通: OS IV/F 4 ADF 運用手引書, 富士通マニュアル (Jan. 1980).
- 3) 富士通: OS IV/F 4 ADF 使用手引書, 富士通マニュアル (Jan. 1980).
- 4) 伊藤, 川田: 超大容量記憶装置の動向, 情報処理, Vol. 19, No. 5, pp. 465-471 (1978).
- 5) 岡本: 超大容量記憶の適用領域の検討, 昭 52 信学全大, 講演番号 1403.
- 6) Lum, V. Y. et al.: A Cost Oriented Algorithm for Data Set Allocation in Storage Hierarchies, C. ACM, Vol. 18, No. 6, pp. 318-322 (1975).
- 7) 藤井: 磁気ディスク対磁気テープの費用効果比, JAERI-M 7424 (1977).

- 8) 村松: 生産管理の基礎, 国元書房 (1973).
- 9) Wickelgren, W. A.: Single-Trace Fragility of Memory Dynamics, Memory & Cognition, Vol. 2, pp. 775-780 (1974).
- 10) 阿部, 往住: 記憶のモデルについて, 数理科学, No. 179, pp. 22-27 (1978).
- 11) Atkinson, R. C. and Shiffrin, R. M.: The Control Processes of Short-Term Memory, Scientific American, pp. 82-90 (Aug. 1971).
- 12) IBM: IBM 3850 大容量記憶システム (MSS) 入門, IBM マニュアル, GA 32-0028-2 (1976).
- 13) 藤井, 浅井: 磁気ディスク, MSS, 磁気テープを用いた階層的ファイル自動管理システム, JAERI-M 8626 (1979).

付 録 損益分岐ラインの算出

(1) 変 数

- ・ファイルサイズ x (kB)
- ・ファイル参照頻度 y (回/月)

(2) 入力データ

参 照 属 性	ランダム		シーケンシャル	
	オンライン	オフライン	オンライン	オフライン
・ファイル属性, 参照形態比率	a_1	a_2	a_3	a_4
参 照 属 性	ランダム		シーケンシャル	
	オンライン	オフライン	オンライン	オフライン
・ファイル R/W 量	a_1x	a_2x	a_3x	a_4x

- ・ファイル参照において Write の行われる確率 δ
- ・評価係数 (ジョブ実行多重度: N とする)
- R/W と実行時間に対するシステム費負担率 γ_1/N
- アクセス時間 " γ_2/N
- ユーザの待ち時間に対するユーザ人件費負担率
- オンライン参照: γ_3 , オフライン参照: γ_4

- ・ファイルのブロック・サイズ b (kB)
- ・ステージング用バッファ・サイズ B (kB)

・装置性能

装 置	性 能	
	アクセス時間(秒)	データ転送速度 (kB/秒)
ディスク	A_d	T_d
MSS	A_m	T_m
テ ー プ	A_t	T_t

- ・MSS からのステージング比率 β_m
- ・テープから作業用ディスクへのステージ比率 β_t
- ・計算機システム費 (CPU+MEM)= C_c (円/秒, 以下同じ), チャネル費 (BMC)= C_c , テープ装置費= C_t , ユーザ人件費= C_u , オペレータ人件費= C_o .
- ・ディスク装置費 F_d (円/kB・月)
- ・MSS 装置費 F_m (円/kB・月)
- ・テープ費 (近似式) $T_{a1} + T_{a2} \cdot x$ (円/月)

・ 1 か月の計算機稼動日数 D (日/月)
 ジョブ・セット時間 $=t_j$ (秒, 以下同じ), テープ探し時間 $=t_s$, テープ・セット時間 $=t_m$, テープから作業用ディスクへのファイル読み込み実行時間 $=\omega_1+\omega_2x$,
 テープ・ジョブの実行時間 $=\omega_3+\omega_4x$.

(3) 費用項目

	ディスク (d)	MSS (m)	テープ (t)
装置費	$F_d \cdot x$	$F_m \cdot x$	$(F_{t1}+F_{t2} \cdot x)y$
チャンネル費	$S_{cd} \cdot xy$	$S_{cm} \cdot xy$	$S_{ct} \cdot xy$
システム費 1	$S_{1d} \cdot xy$	$S_{1m} \cdot xy$	$(S_{1t1}+S_{1t2} \cdot x)y$
システム費 2	$S_{2d} \cdot xy$	$(S_{2m1}+S_{2m2} \cdot x)y$	$(S_{2t1}+S_{2t2} \cdot x)y$
ユーザ費 1	$U_{nd} \cdot xy$	$(U_{nm1}+U_{nm2} \cdot x)y$	$(U_{nt1}+U_{nt2} \cdot x)y$
ユーザ費 2	$U_{fd} \cdot xy$	$(U_{fm1}+U_{fm2} \cdot x)y$	$(U_{ft1}+U_{ft2} \cdot x)y$
作業用ディスク費			$F_{dt} \cdot x$
テープ費			$T_{a1}+T_{a2} \cdot x$
オペレータ費			$T_{so} \cdot y$
総費用	G_d	G_m	G_t

(4) 算出式

$$\begin{aligned} \text{RAL} &= \alpha_1 a_1 + \alpha_2 a_2 + \alpha_3 a_3 + \alpha_4 a_4, \text{RON} = \alpha_1 a_1 + \alpha_2 a_2, \\ \text{ROF} &= \alpha_3 a_3 + \alpha_4 a_4, \text{RPO} = \alpha_1 a_1 + \alpha_3 a_3, \text{RPS} = \alpha_2 a_2 \\ &+ \alpha_4 a_4, \text{PON} = \alpha_1 + \alpha_2, \text{POF} = \alpha_3 + \alpha_4, \text{PPO} = \alpha_1 + \\ &\alpha_3, \text{PPS} = \alpha_2 + \alpha_4, \text{STM} = \beta_m + \delta, \text{STT} = \beta_t + \delta, \\ T_s &= \max[1/T_m, (A_d + B/T_d)/B] \text{ とおくと,} \\ S_{cd} &= S_{cm} = \text{RAL} \cdot C_s / T_d, \\ S_{1d} &= S_{1m} = \text{RAL} \cdot C_s \cdot (\gamma_1/N) / T_d, \\ S_{2d} &= \text{RAL} \cdot C_s \cdot (\gamma_2/N) \cdot A_d / b, \\ S_{2m1} &= A_m \cdot \text{STM} \cdot C_s \cdot (\gamma_2/N), \\ S_{2m2} &= T_s \cdot \text{STM} \cdot C_s \cdot (\gamma_2/N) + S_{2d}, \\ U_{nd} &= (A_d + b/T_d) / b \cdot \text{RON} \cdot C_s \cdot \gamma_3, \\ U_{fd} &= (A_d + b/T_d) / b \cdot \text{ROF} \cdot C_s \cdot \gamma_4, \\ U_{nm1} &= A_m \cdot \text{STM} \cdot \text{PON} \cdot C_s \cdot \gamma_3, \\ U_{nm2} &= T_s \cdot \text{STM} \cdot \text{PON} \cdot C_s \cdot \gamma_3 + U_{nd}, \\ U_{fm1} &= A_m \cdot \text{STM} \cdot \text{POF} \cdot C_s \cdot \gamma_4, \\ U_{fm2} &= T_s \cdot \text{STM} \cdot \text{POF} \cdot C_s \cdot \gamma_4 + U_{fd}, \\ F_{t1} &= \{ \text{PPO} \cdot (t_m + \omega_1) \cdot \text{STT} + \text{PPS} \cdot (t_m + \omega_3) \} \cdot C_s, \\ F_{t2} &= (\text{PPO} \cdot \omega_2 \cdot \text{STT} + \text{PPS} \cdot \omega_4) \cdot C_s, \\ S_{ct} &= \left\{ \text{PPO} \cdot \text{STT} \cdot \left(\frac{1}{T_s} + \frac{1}{T_d} \right) + \frac{\text{RON}}{T_d} + \frac{\text{ROF}}{T_s} \right\} \cdot C_s, \\ S_{1t1} &= \text{PPO} \cdot \text{STT} \cdot \omega_1 \cdot C_s \cdot (\gamma_1/N), \\ S_{1t2} &= (\text{PPO} \cdot \text{STT} \cdot \omega_2 + \text{RPO}/T_d + \text{RPS}/T_s) \cdot C_s, \end{aligned}$$

* テープにあるファイルを磁気ディスクに入れるためにジョブ制御文をセットし、カード・リーダーで読み込ませるまでの時間。

$$\cdot (\gamma_1/N).$$

$$S_{2t1} = t_m \cdot (\text{PPO} \cdot \text{STT} + \text{PPS}) \cdot C_s \cdot (\gamma_2/N).$$

$$S_{2t2} = \{ \text{PPO} \cdot \text{STT} \cdot (A_d + A_s) + A_d \cdot \text{RPO} + A_s \cdot \text{RPS} \} / b \cdot C_s \cdot (\gamma_2/N).$$

$$U_{n1} = \{ \alpha_1 \cdot \text{STT} \cdot (t_j + t_s + t_m + \omega_1) + \alpha_2 \cdot (t_s + t_m) \} \cdot C_s \cdot \gamma_3.$$

$$U_{n2} = \{ \alpha_1 \cdot \text{STT} \cdot \omega_2 + \alpha_1 \cdot a_1 / T_d + \alpha_2 \cdot a_2 / T_s \} \cdot C_s \cdot \gamma_3.$$

$$U_{f1} = \{ \alpha_3 \cdot \text{STT} \cdot (t_j + t_s + t_m + \omega_1) + \alpha_4 \cdot (t_s + t_m) \} \cdot C_s \cdot \gamma_4.$$

$$U_{f2} = \{ \alpha_3 \cdot \text{STT} \cdot \omega_2 + \alpha_3 \cdot a_3 / T_d + \alpha_4 \cdot a_4 / T_s \} \cdot C_s \cdot \gamma_4.$$

$$F_{dt} = \text{PPO} \cdot F_d / D.$$

$$T_{so} = (\text{PPO} \cdot \text{STT} + \text{PPS}) \cdot (t_s + t_m) \cdot C_s.$$

(5) 損益分岐ライン

各記憶媒体における総費用は,

$$G_D = F_d \cdot x + (S_{cd} + S_{1d} + S_{2d} + U_{nd} + U_{fd}) \cdot xy \\ = D1 \cdot x + D2 \cdot xy,$$

$$G_m = F_m \cdot x + (S_{cm} + S_{1m} + S_{2m2} + U_{nm2} + U_{fm2}) \cdot xy \\ + (S_{2m1} + U_{nm1} + U_{fm1}) \cdot y \\ = M1 \cdot x + M2 \cdot xy + M3 \cdot y,$$

$$G_t = T_{a1} + (F_{dt} + T_{a2}) \cdot x + (F_{t2} + S_{ct} + S_{1t2} + S_{2t2} \\ + U_{f2}) \cdot xy + (F_{t1} + S_{1t1} + S_{2t1} + U_{n1} + U_{f1} \\ + T_{so}) \cdot y \\ = T1 + T2 \cdot x + T3 \cdot xy + T4 \cdot y$$

となる。これより各損益分岐ラインは次のように計算される。

[ディスク対 MSS]

$$G_d = G_m \text{ より}$$

$$y_1 = \frac{(D1 - M1)x}{(M2 - D2)x + M3}$$

[MSS 対 テープ]

$$G_m = G_t \text{ より}$$

$$y_2 = \frac{(M1 - T2)x - T1}{(T3 - M2)x + T4 - M3}$$

[ディスク対 テープ]

$$G_d = G_t \text{ より}$$

$$y_3 = \frac{(D1 - T2)x - T1}{(T3 - D2)x + T4}$$

(昭和 54 年 8 月 24 日 受付)

(昭和 55 年 6 月 19 日 採録)