

JAERI-M

8626

磁気ディスク，大容量記憶装置(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

1979年12月

藤井 実・浅井 清

この報告書は、日本原子力研究所が JAERI-M レポートとして、不定期に刊行している研究報告書です。入手、複製などのお問い合わせは、日本原子力研究所技術情報部（茨城県那珂郡東海村）あて、お申しこしください。

JAERI-M reports, issued irregularly, describe the results of research works carried out in JAERI. Inquiries about the availability of reports and their reproduction should be addressed to Division of Technical Information, Japan Atomic Energy Research Institute, Tokai-mura, Naka-gun, Ibaraki-ken, Japan.

磁気ディスク，大容量記憶装置（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

日本原子力研究所東海研究所計算センター
藤井 実・浅井 清

（1979年11月29日受理）

磁気ディスク，大容量記憶装置（MSS）および磁気テープの3種の記憶媒体を用いた，簡素で効果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。

ユーザ・ファイルについて，(1)ファイル分布，ファイル連続参照特性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。(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等，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等を採用した，簡素で実用的，効果的なファイル自動管理の一方式を提案した。

Multi-level, Automatic File Management System
using Magnetic Disk, Mass Storage System and
Magnetic Tape

Minoru FUJII and Kiyoshi ASAI
Computing Center, Tokai Research Establishment,
JAERI

(Received November 29, 1979)

A simple, effective file management system using magnetic disk, mass storage system(MSS) and magnetic tape is described.

Following are the concepts and techniques introduced in this file management system.

- (1) File distribution and continuity character of file references are closely approximated by memory retention function. A density function using the memory retention function is thus defined.
- (2) A method of computing the cost/benefit lines for magnetic disk, MSS and magnetic tape is presented.
- (3) A decision process of an optimal organization of file facilities incorporating file demands distribution to respective file devices, is presented.
- (4) A method of simple, practical, effective, automatic file management, incorporating multi-level file management, space management and file migration control, is proposed.

Keywords:

Computer, Mass Storage System, File Management, Memory Retention, Cost/Benefit Line, Optimization, Magnetic Disk, Magnetic Tape

目 次

1. はじめに	1
2. システム設計の概要	3
2.1 背景	3
2.2 システム作成過程	3
3. 記憶保存関数について	6
3.1 Wickelgren の記憶保存関数	6
3.2 記憶保存分布密度関数	7
3.3 記憶保存関数の特徴	9
3.4 記憶保存関数との出会い	9
4. ファイル分布と記憶保存関数	11
4.1 ファイル・サイズ分布と記憶保存関数	11
4.2 ファイル参照頻度分布と記憶保存関数	12
4.3 ファイル・サイズとファイル参照頻度の関係	12
5. 磁気ディスク, MSS, 磁気テープの損益分岐ライン	16
5.1 評価方法と前提条件	16
5.2 評価項目	16
5.3 算出結果	17
6. ファイル需要の各記憶媒体への振分け	18
6.1 振分け方法	18
6.2 算出例	19
7. ファイル連続参照特性と記憶保存関数	21
8. ファイル管理システムの最適装置構成	23
8.1 MSS のSDISKの最適台数	23
8.2 算出例	25
9. ファイル生成・消滅データと空き領域減少データ	27
10. ファイルの階層的管理	30
10.1 2レベル・ファイル管理の概要	30
10.2 S-s (補充点-消去点) 空き領域管理方式	31
10.3 最適消去量	31
10.4 損益分岐ライン制御と参照特点の管理	33
11. おわりに	34
謝 辞	34
参考文献	35
付録1. 損益分岐ラインの算出	36
付録2. 総ステージング時間(TST)の算出	40

Contents

1. Introduction	1
2. Outline of system design	3
2.1 Background	3
2.2 Construction process of file management system	3
3. Memory retention function	6
3.1 Memory retention function of Wickelgren	6
3.2 Density function by memory retention function	7
3.3 Characteristics of memory retention function	9
3.4 Encounter with memory retention function	9
4. File distribution vs. memory retention function	11
4.1 File size distribution vs. memory retention function	11
4.2 File reference distribution vs. memory retention function ..	12
4.3 Relationship between file size and file reference	12
5. Cost/benefit lines for magnetic disk, MSS, magnetic tape	16
5.1 Evaluation method and premises	16
5.2 Evaluation items	16
5.3 Computed result	17
6. Distribution of file demands to respective storage devices ...	18
6.1 Allocation method	18
6.2 Example of distribution	19
7. Characteristic of continuous file references vs. memory retention function	21
8. Optimal configuration of devices for file management system ...	23
8.1 Optimal staging capacity of MSS	23
8.2 Computed amounts of staging volume	25
9. Data of file creation, deletion and free space decreasement ..	27
10. Hierarchical file management	30
10.1 Outline of two-level file management	30
10.2 S-s space management	31
10.3 Optimal volumes of deletion	31
10.4 File migration control by cost/benefit line, process of reference data	33
11. Concluding remarks	34
Acknowledgment	34
References	35
Appendix 1 Calculation of cost/benefit line	36
Appendix 2 Calculation of total staging time	40

1. はじめに

大容量記憶システム（MSS：Mass Storage System）の出現により、情報処理システムにおける記憶階層は図1（Fig. 1、以後文中ではFig. nを図nで表わす）のように表わされる^[1]。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上位の記憶媒体に保存され、参照頻度が低くなればより低コストの下位の記憶媒体に移されるのが望ましい。

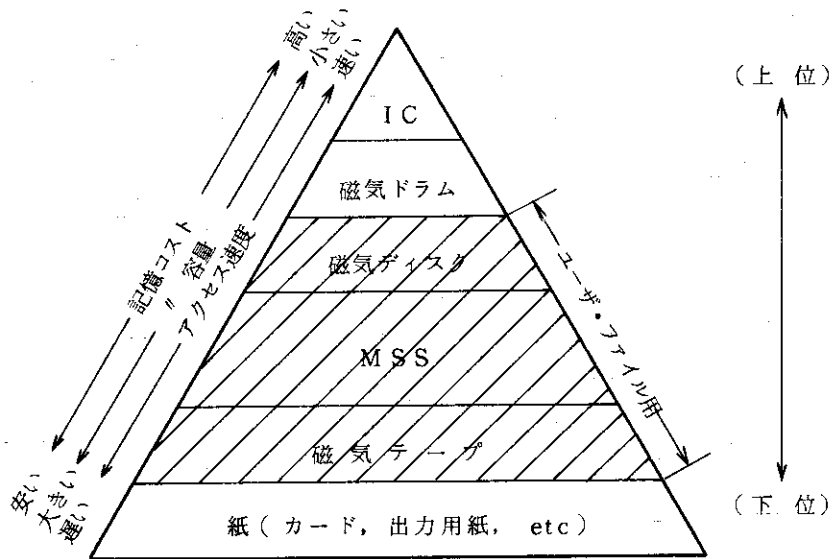


Fig. 1 Hierarchical Structure of Storages.

従来こうした記憶階層の効率的利用は、ファイル管理においては、磁気ディスク、磁気テープ間のファイル移動で図られ、実行されてきているが^[2]、人手による介入を必要としてきた。この自動化は、IBMのHSM^[3]（Hierarchical System Manager）で磁気ディスク、MSS上のファイルを対象に最近実現されたが、わが国では昭和54年7月現在未稼働である。

従来のファイル管理やHSMは、LRU（Least Recently Used：一番長い間使用されていないものから取出す）方式でファイルを下位の記憶媒体に移動させる。しかし、この方式は効率的なディスク運用を圧迫する。なぜなら、費用効果比からみると本来テープに入れて使用した方が望ましいような大きなファイルでさえも、LRU方式では1回参照するだけでディスクに居残るからである。ディスク装置費はファイル・サイズに比例してかかり、実際の運用においてはファイル・サイズに応じて参照頻度の高いものをディスクに入れておいた方がよい。

この報告は、磁気ディスク、MSS、磁気テープを用いたユーザー・ファイル管理システムとその設計法を考察したものである。筆者らが追求した望ましいファイル管理システムの要件は、

- (i) ファイルをその大きさ、参照頻度、参照形態などに応じて最も経済的な記憶媒体に移動できること、

(ii) ファイル需要に応じて最適なファイル装置構成をとれること、
 (iii) ファイル管理は自動化でき、簡素で実用的であること
 である。

(i)の要件には評価基準、例えばディスク、MSS、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〔1〕、〔4〕に概説されている。損益分岐ラインの作成は、Lum, et al.〔5〕、藤井〔6〕等で試みられている。(ii)は従来経験的に決められている。(iii)の自動化の一例はHSM〔3〕である。

この報告では、上記の要件を満すファイル管理システムの設計を現実のデータの上に乗って追求し、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、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た。後者についてはさらに、(イ)記憶保存分布密度関数を新しく定義した。(ロ)ディスク媒体全体が1つの記憶主体と考えられる解釈を示した。筆者らは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) サイズ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等、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等を採用した、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。

この報告では、主にディスク・ファイルのデータを使っ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、実システムの作成においては両方のデータを考慮しなければならない。この報告で使用しているデータは、特に断らないかぎり、原研計算センタの磁気ディスク上のユーザ・ファイルから収集したものである。

2. システム設計の概要

2.1 背景

原研計算センタでは課金制を採用せず、約700人のユーザに必要なに応じて磁気ディスクを使用させている。このため、ユーザのディスク利用は必ずしも効率的に行われていない。

現在ディスクが一杯になると参照頻度の少いファイルから磁気テープに退避して空き領域を作成しているが、図2、図14に見られるように消去後半月もたてばディスクは再び一杯となる。この状態に至ると、計算センタがファイル圧縮、未使用域の解放等の介入を行わなければ、小容量のファイルでさえも作成できなくなり、ジョブ処理に支障をきたす。

この磁気テープへの退避、空き領域作成等の作業量は、人手によるファイル管理に行詰りを感じさせるほどに月々増加している。このため、筆者らの計算センタでもMSSの導入、効率的なファイル自動管理手法等を考えなければならなくなった。

2.2 システム作成過程

図3にファイル管理システムの作成過程を示す。図のブロックの左肩の番号は以下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を使うと、ファイル需要を⑤で示す損益分岐ラインによって磁気ディスク、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、⑦の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、⑥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。

原研計算センタでは、ユーザの仕事の性格上、記憶階層間のファイル移動は半月に1回程度行えばよく、IBMのHSMのような常駐タスクの制御を必要としない。⑨のファイル生成・消滅データ、空き領域減少データを使用すると、在庫管理を応用したS-s^[7](補充点-消去点)空き領域管理方式が設定できる。通常の記憶階層間のファイル移動はユーザにまかせ、ディスク、MSSの一方の空き領域が下限s_i以下になった時に全体調整(記憶階層間のファイル移動)を行う2レベルのファイル管理を考える。こうすると計算センタの行う全体調整は半月に1回程度行えばよく、損益分岐ラインを使った効果的なファイル移動制御が可能となる。この制御は、指数平滑法を使うと各ファイルについて2個の収集データの追加で可能であり、容易に実用化できる。

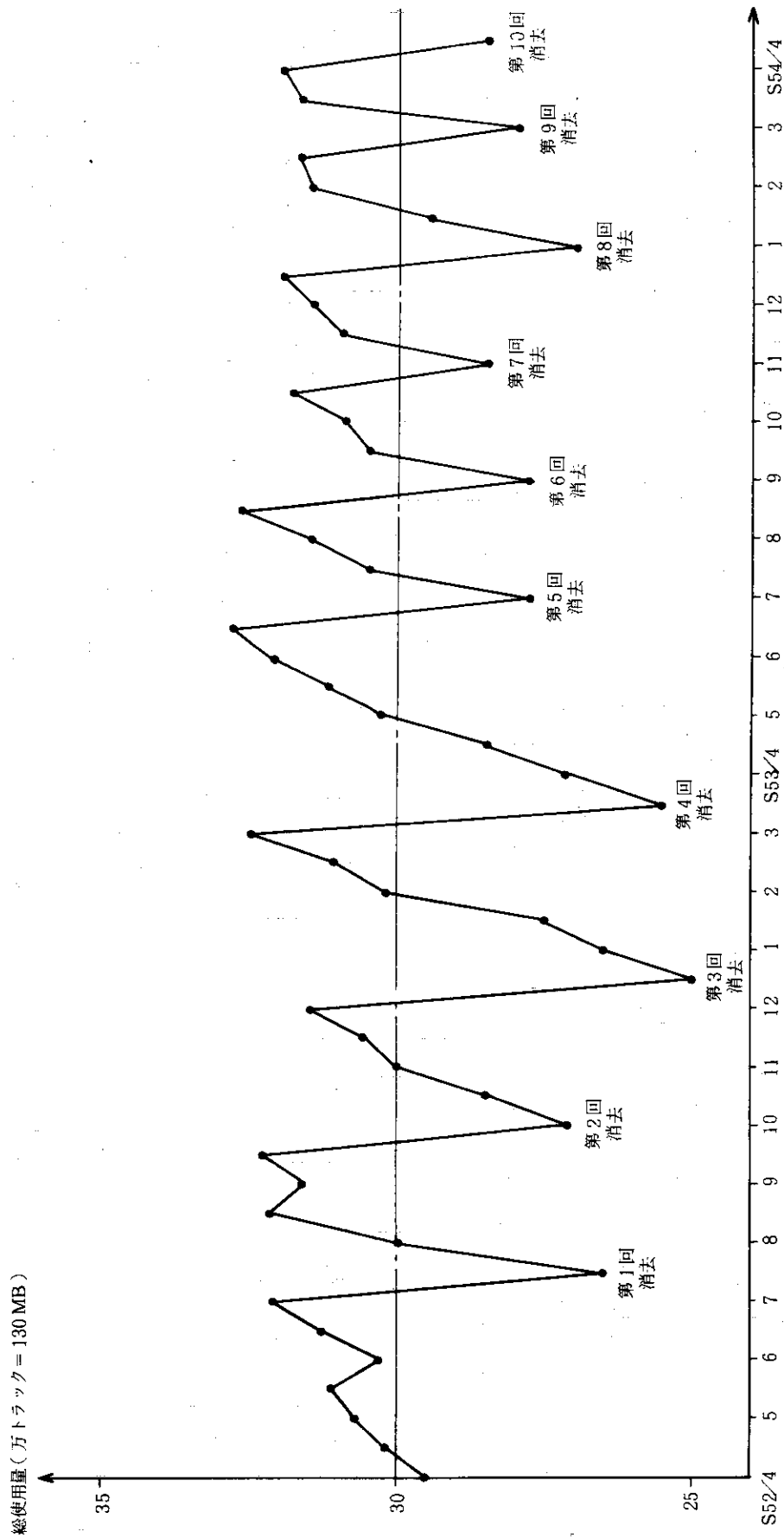


Fig. 2 Total Volume of user files.

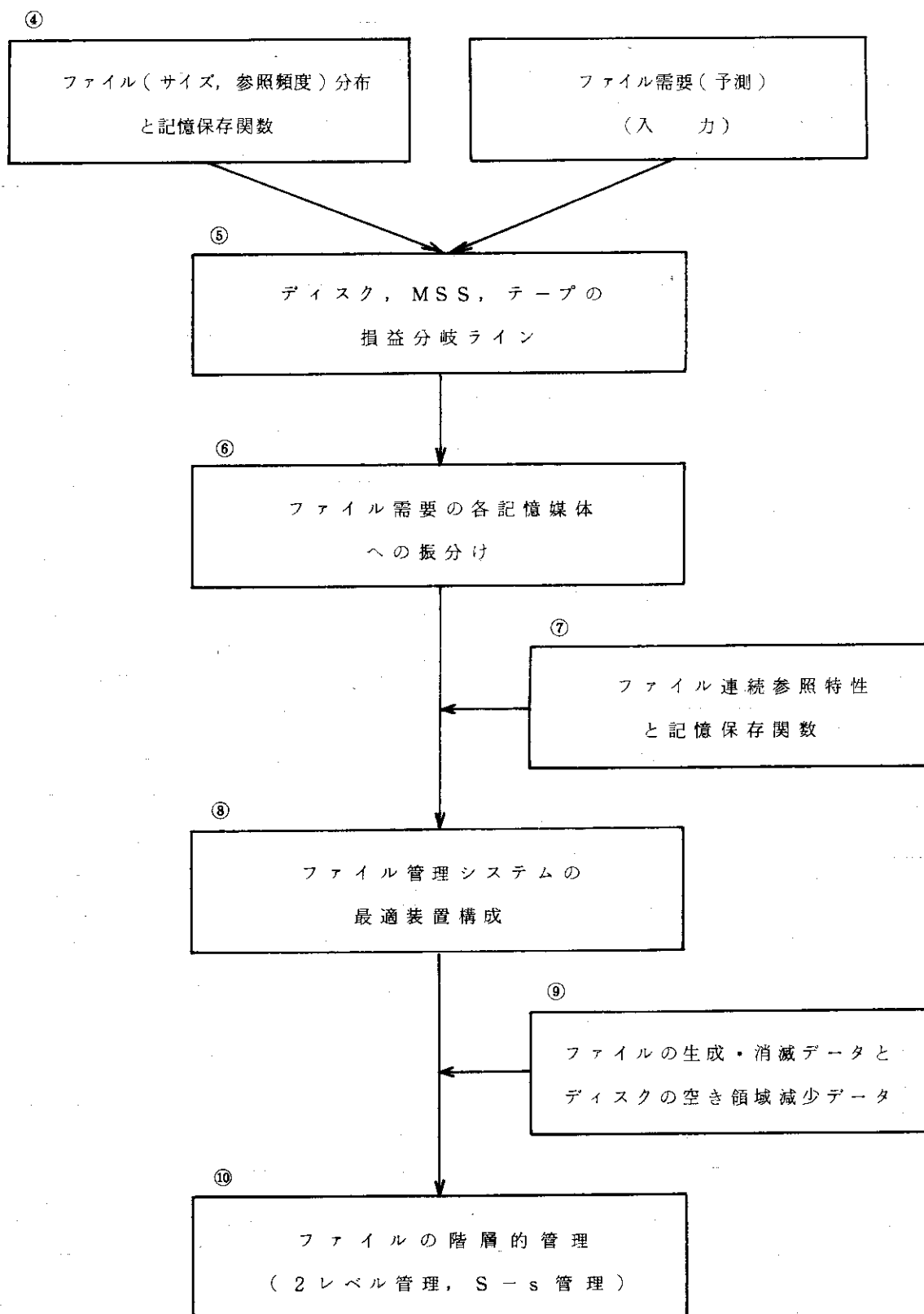


Fig. 3 A Construction Process of the File Management System.

3. 記憶保存関数について

この章では Wickelgren の示した記憶保存関数と、これから作り出した記憶保存分布密度関数について述べる。

4章のファイル・サイズ分布、ファイル参照頻度分布はこの記憶保存分布密度関数の3次の式、7章のファイル連続参照特性はこの2次の式で非常に良く近似できる。

3.1 Wickelgren の記憶保存関数

記憶モデルの研究では、1950年代に Miller や Broadbent によって人間を1つの情報処理系とみなす方向づけがなされ、60年代には図4に示す Atkinson & Shiffrin のモデル^[10]のように記憶を短期記憶と長期記憶の2つの成分に分離するモデルが主に考えられてきたようである^[9]。その後70年代に入ると、記憶を情報処理のレベルの違いとして連続的に取扱う Craik & Lockhart の研究、記憶の保持量（あるいは忘却率）と時間経過の関係を実験的に調べ、これを数理表現した Wickelgren^[8]の研究等が行われてきているようである^[9]。

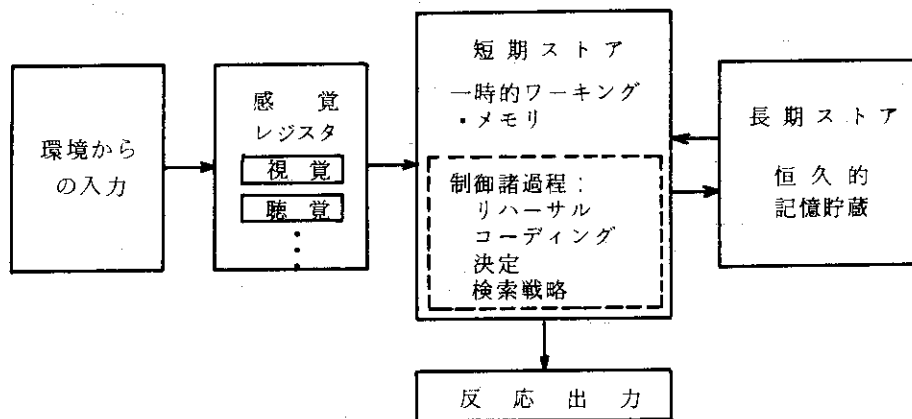


Fig. 4 The Model of Memory System by Atkinson & Shiffrin^{[9],[10]}

ここでは、Wickelgren の示した記憶保存関数に着目する。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、次の2つの微分方程式が記憶現象をよく近似すると述べている。

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

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

ここで、

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

である。(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、この式が数秒から数年にわたる記憶の痕跡の減衰データによく適合すると述べている。(3.3)の $e^{-\pi t}$ の項が10~20秒以下の記憶減衰を説明し、 $(1 + \beta t)^{-\phi}$ の項が10~20秒から少くとも2年位までにわたる記憶減衰を説明すると述べている。

3.2 記憶保存分布密度関数

ここでは、(3.3)において干渉項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る。

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

[密度関数として定義できる条件]

(3.4)が密度関数として定義できるのは、

$$\int_0^{\infty} m(x) dx = 1$$

なる正規化条件を満たすとき、即ち

$$\phi > 1 \text{ かつ } \frac{\lambda}{(\phi - 1) \cdot \beta} = 1 \quad (3.5)$$

を満たすときである。これは、

$$\lambda = (\phi - 1) \cdot \beta = (\eta - \mu) \cdot \rho$$

と変形され、

$$\eta - \mu = \frac{\lambda}{\rho}$$

つまり,

$$\begin{aligned} & \left(\begin{array}{c} \text{記憶の強さの時間に} \\ \text{よる減衰パラメータ} \end{array} \right) - \left(\begin{array}{c} \text{記憶のもろさの時間に} \\ \text{よる減衰パラメータ} \end{array} \right) \\ &= \frac{\text{最初の記憶の強さ}}{\text{最初の記憶のもろさ}} \end{aligned} \quad (3.6)$$

の場合である。

[分布関数]

(3.4) の密度関数をもつ分布関数 $F(x)$ は,

$$\begin{aligned} F(x) &= \int_0^x m(t) dt = \int_0^x \lambda (1 + \beta t)^{-\phi} dt \\ &= \frac{\lambda}{(\phi-1) \cdot \beta} \left\{ 1 - (1 + \beta x)^{-(\phi-1)} \right\} \end{aligned}$$

となり, (3.5) の条件を代入すると,

$$F(x) = 1 - (1 + \beta x)^{-(\phi-1)}, \quad \phi > 1 \quad (3.7)$$

となる。

[分布の平均値]

(3.4) の密度関数をもつ分布の平均 \bar{x} は,

$$\begin{aligned} \bar{x} &= \int_0^{\infty} x \cdot m(x) dx = \int_0^{\infty} \lambda x (1 + \beta x)^{-\phi} dx \\ &= \left[\lambda x \cdot \frac{(1 + \beta x)^{-(\phi-1)}}{(\phi-1) \cdot \beta} \right]_0^{\infty} - \int_0^{\infty} \lambda \cdot \frac{(1 + \beta x)^{-(\phi-1)}}{(\phi-1) \cdot \beta} dx \end{aligned}$$

となり, (3.5) の条件を代入すると,

$$\bar{x} = \frac{\infty}{(1 + \beta \cdot \infty)^{(\phi-1)}} + \frac{1}{(\phi-2) \cdot \beta} \quad (3.8)$$

となる。故に,

$$\begin{cases} \phi > 2 \text{ の時, } \bar{x} = \frac{1}{(\phi-2) \cdot \beta} \\ \phi \leq 2 \text{ の時, } \bar{x} = \infty \end{cases} \quad (3.9)$$

[n 次の記憶保存分布密度関数]

(3.5) の条件を(3.4)に代入し, ϕ を n , β を λ と表現すると,

$$f(x) = (n-1) \lambda (1 + \lambda x)^{-n} \quad (3.10)$$

となる。ここで、 $n > 1$ 。 $n > 2$ の時、平均 \bar{x} は

$$\bar{x} = \frac{1}{(n-2) \cdot \lambda}$$

となる。(3.10)の λ は先の記憶の強さとは関係のないパラメータとなっている。

筆者らは(3.10)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： $f(x)$ という用語で使用する。

特に $n = 3$ の時は、

$$f(x) = 2 \lambda (1 + \lambda x)^{-3} \quad (3.11)$$

で、 $\bar{x} = 1/\lambda$ となる。(3.11)を次章で使用する。

3.3 記憶保存関数の特徴

ここでは、記憶保存関数(3.12)と指数分布密度関数(3.13)の比較を行う。

$$f(x) = (n-1) \lambda (1 + \lambda x)^{-n} \quad (3.12)$$

$$g(x) = \lambda e^{-\lambda x} \quad (3.13)$$

指数分布密度関数 $g(x)$ の特徴は、

$$\frac{g(x+1)}{g(x)} = \frac{\lambda e^{-\lambda(x+1)}}{\lambda e^{-\lambda x}} = \frac{1}{e^\lambda} = \text{一定} \quad (3.14)$$

となることである。

同様に、記憶保存関数 $f(x)$ を調べると、

$$\begin{aligned} \frac{f(x+1)}{f(x)} &= \frac{(n-1) \lambda \{1 + \lambda(x+1)\}^{-n}}{(n-1) \lambda (1 + \lambda x)^{-n}} \\ &= \left(\frac{1 + \lambda x}{1 + \lambda + \lambda x} \right)^n = \left(1 - \frac{\lambda}{1 + \lambda + \lambda x} \right)^n \end{aligned} \quad (3.15)$$

となる。これは、 x が大きくなるに従って $f(x+1)/f(x)$ が大きくなり、徐々に1に近づく。だから、 x が大きくなっても $f(x)$ はなかなか0にならない性格をもつ。

3.4 記憶保存関数との出会い

筆者らと Wickelgren の示した記憶保存関数との出会いは、7章で述べるファイル連続参照

分布の近似関数を模索していた時に始まる。この連続参照分布の近似を従来定義されている種々の分布関数で行ってみたが、どれも実用的に使えるだけの近似精度は得られなかった。そこで、従来まだ分布関数としては考えられていない関数を作成したり、適用して近似関数を探した。この過程で、文献〔9〕の記憶モデルに目がとまり、Wickelgren の記憶保存関数を適用してみて、非常に良い結果が得られることがわかった。

その後、ファイル分布の近似においても、その形状から記憶保存関数を適用して、3次の式で良く近似できることがわかった。

4. ファイル分布と記憶保存関数

ここでは、ファイル・サイズ分布、ファイル参照頻度分布がともに記憶保存関数の3次の式で良く近似できることを示す。又、ファイル・サイズとファイル参照頻度の関係も示す。

4.1 ファイル・サイズ分布と記憶保存関数

図5に原研計算センタのファイル・サイズ分布とこれを指数分布密度関数、記憶保存関数を使って近似した曲線を示す。図6は図5の分布を対数目盛で表現したものである。

$$\text{指数分布密度関数} : f_1(x) = \lambda_1 e^{-\lambda_1 x} \quad (4.1)$$

$$\text{記憶保存関数} : f_2(x) = 2 \lambda_1 (1 + \lambda_1 x)^{-3} \quad (4.2)$$

$$\text{超指数分布密度関数} : f_3(x) = \sum_{i=1}^n \alpha_i^2 \lambda_i e^{-\alpha_i \lambda_i x}, \quad \sum_{i=1}^n \alpha_i = 1 \quad (4.3)$$

ここで、ファイル・サイズの平均は、 $1/\lambda_1 = 31 \times 13 \text{ KB}$ (キロバイト) である。

図6に見られるように(4.1)の指数関数で近似すると特にファイル・サイズの大きい所では全く適合しない。又、n種類のユーザ・グループがいると仮定して、(4.3)の超指数関数でも近似を試みたがあまり良い結果は得られなかった。ところが、(4.2)の記憶保存関数を適用すると、 $10^{-1} \sim 10^{-6}$ の範囲にわたって非常に良く近似できることがわかった。そこで、この因果関係に対する1つの解釈を以下に試みる。

一般に大きいファイルを完成するには時間がかかる。今、サイズx(KB)のファイルを完成するには、その大きさに比例して時間tがかかると考える。

$$x = at, \quad a : \text{比例定数} \quad (4.4)$$

この時、ファイル・サイズの平均を $1/\lambda_1$ 、ファイル完成時間の平均を $1/\lambda_1^*$ とすると、

$$\frac{1}{\lambda_1} = \bar{x} = at = a \cdot \frac{1}{\lambda_1^*} \quad (4.5)$$

となる。(4.4)、(4.5)を(4.2)に代入すると、

$$f_2(t) = \frac{1}{a} \{ 2 \lambda_1^* (1 + \lambda_1^* t)^{-3} \} \quad (4.6)$$

となる。

ここで、任意のファイルについての記憶時間は、そのファイル完成までに要した時間に比例すると考える。すると、(4.2)、(4.6)より、ディスク媒体全体を1人の人間のよう1つの記憶

主体と考えることができ、任意の時点における大きなファイルと小さなファイルの存在割合は、長い間覚えていたような事象と短い間しか覚えていないような事象の存在割合と良く似ていると解釈できる。

4.2 ファイル参照頻度分布と記憶保存関数

図7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。

$$\text{記憶保存関数: } f_4(y) = 2 \lambda_2 (1 + \lambda_2 y)^{-3} \quad (4.7)$$

ここで、 $1/\lambda_2$ は月平均参照回数であり、 $1/\lambda_2 = 10.86$ (回/月)である。

(4.7)の解釈を以下に試みる。参照頻度 y (回/月)の度数(記憶の強さと考える)に比例して、ファイル保存時間(記憶保存時間)が長くなるとする。

$$y = bt, \quad b: \text{比例定数} \quad (4.8)$$

この時、ファイル参照頻度の平均を $1/\lambda_2$ 、ファイル保存期間の平均を $1/\lambda_2^*$ とすると

$$\frac{1}{\lambda_2} = \bar{y} = b\bar{t} = b \cdot \frac{1}{\lambda_2^*} \quad (4.9)$$

となる。(4.8)、(4.9)を(4.7)に代入すると、

$$f_4(t) = \frac{1}{b} \{ 2 \lambda_2^* (1 + \lambda_2^* t)^{-3} \} \quad (4.10)$$

となる。これは、ディスク媒体全体を1つの記憶主体と考えることができ、参照頻度の高いファイルと低いファイルの存在割合は、記憶の強い事象と弱い事象の存在割合に良く似ていると解釈できる。

4.3 ファイル・サイズとファイル参照頻度の関係

図8にファイル・サイズと各ファイル・サイズ毎に算出した月平均ファイル参照回数との関係を示す。この関係を次の式で近似した。

$$\begin{aligned} \frac{1}{\lambda_2(x)} = \bar{y}(x) &= a + b \log_{10} x \\ &= 3 + 7 \log_{10} x \end{aligned} \quad (4.11)$$

ここで、 $1/\lambda_2(x)$ 、 $\bar{y}(x)$ はファイル・サイズ x (KB) のファイルの月平均ファイル参照回数である。

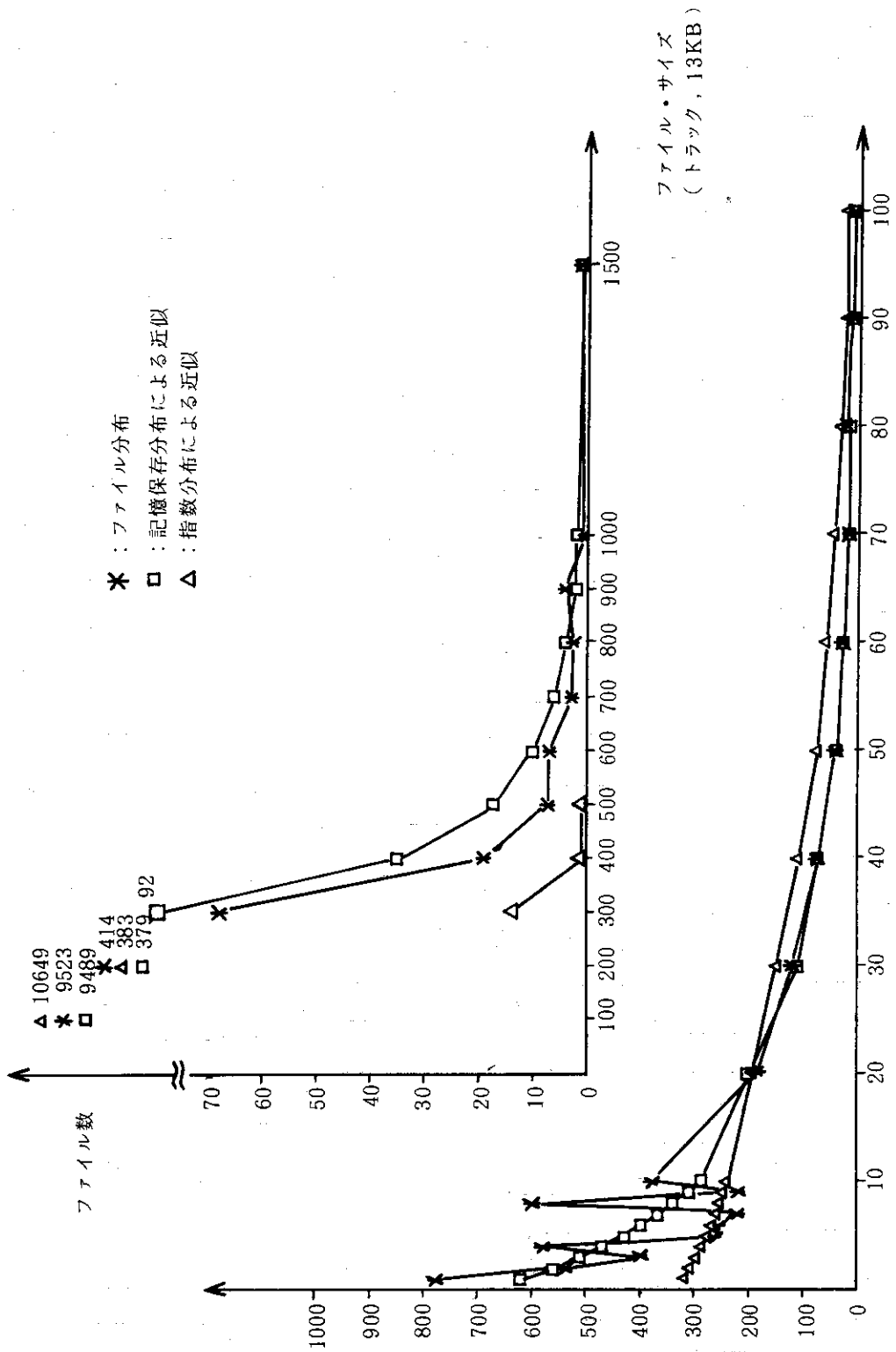


Fig. 5 File Size Distribution, Exponential Distribution and Memory Retention Function.

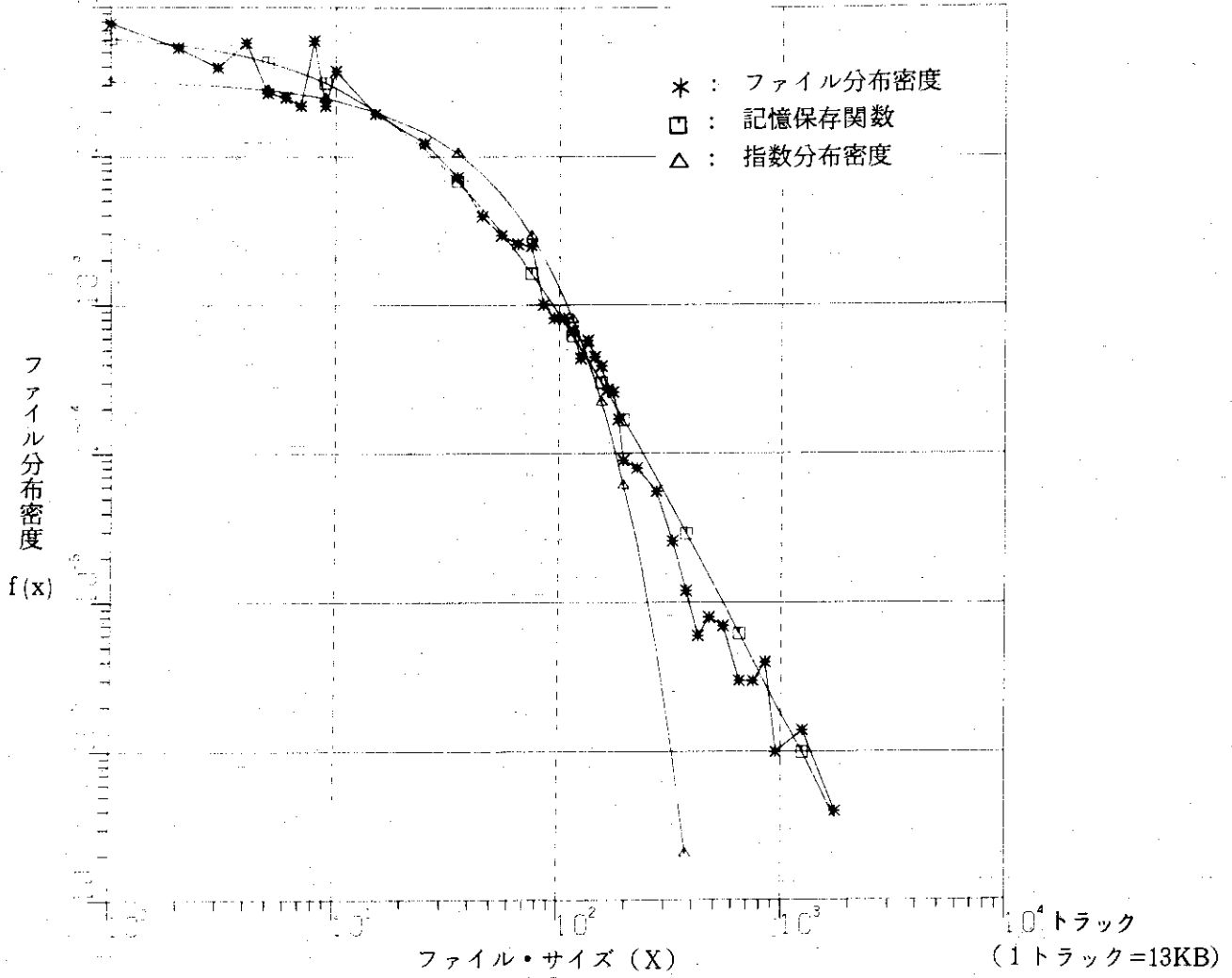


Fig. 6 File Size Distribution and Memory Retention Function.

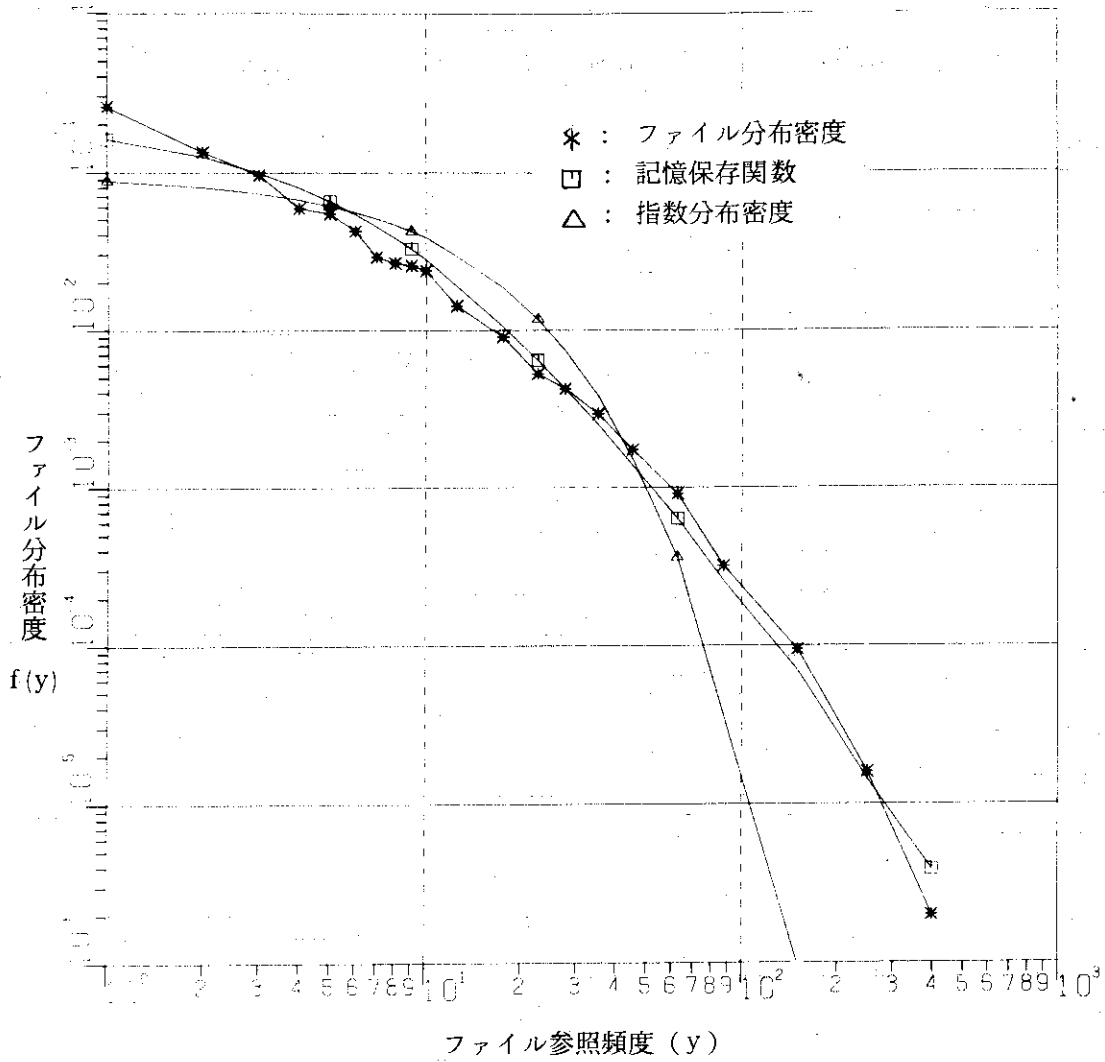


Fig. 7 File Reference Distribution and Memory Retention Function.

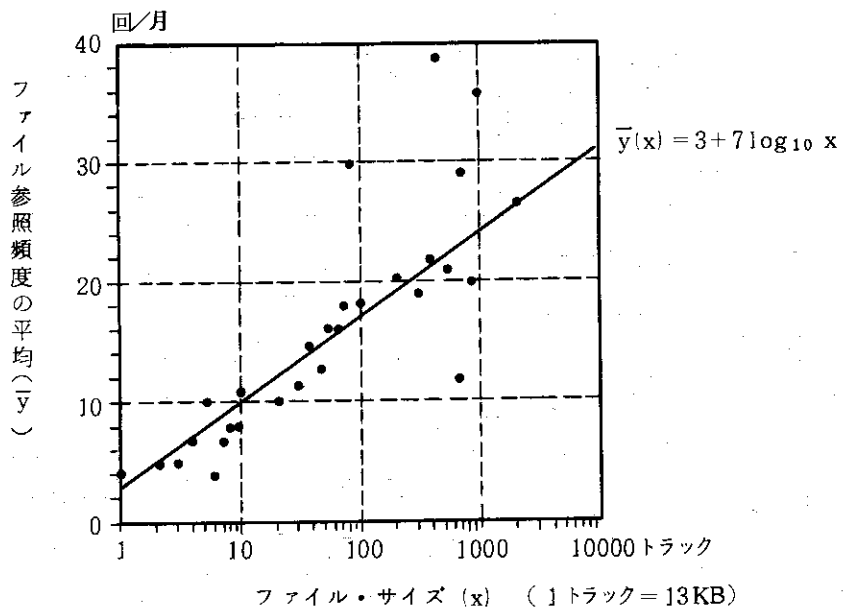


Fig. 8 File Size and File References.

5. 磁気ディスク, MSS, 磁気テープの損益分岐ライン

5.1 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費用を各記憶媒体毎に計算し, 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), ファイル参照形態 (オンライン, バッチ) 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保存する場合, 1ファイルが1本のテープを占有するものとする。
- (3) テープに保存されているランダム・アクセス・ファイルを使用する場合は, 一度作業用ディスクに入れてから使用し, 1参照あたり1日ディスクにとどまるものとする。
- (4) MSS は IBMタイプの MSS を対象とする。

5.2 評価項目

各記憶媒体の総費用を算出するの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1, 類似の算出は文献〔6〕に詳述している。

(1) 装置費

ディスク, MSS の場合はファイル保存費, テープの場合はテープ装置使用費である。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) に各装置の平均占有率, 平均稼働率を加味して算出される。

(2) 計算機使用費 (チャネル費, システム費1)

各記憶媒体からファイルを Read/Write する時, データ転送時間に対してチャネル費を負担させる。

ファイル参照における計算機の正味使用時間 (Read/Write のデータ転送時間, 実行時間) に対し, 演算装置費, 主記憶装置費等について評価係数を乗じて負担させる費用をシステム費1とする。

(3) 計算機使用機会損失費 (システム費2)

ファイル・アクセス時間の計算機使用機会損失に対する費用を評価し, システム費2とする。

(4) ユーザの仕事遅れ損失費 (ユーザ費1, ユーザ費2)

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。ユーザがファイルをオンラインで参照している時は待ち時間のほとんどをユーザの仕事遅れ損失とし, バッチで参照している場合はその何割か (入力パラメータ) をユーザの仕事遅れ損失とする。オンライン参照の場合をユーザ費1, バッチ参照の場合をユーザ費2とした。

(5) テープ関係費 (一時ディスク費, テープ費, オペレータ費)

テープ・ファイルの保存・参照にのみかかる費用で, ランダム・アクセス・ファイルをテープに保存した時に1参照あたり1日作業用ディスクを利用するためにかかる一時ディスク費, テープ媒体費, テープ・ハンドリング等にかかるオペレータ費。

5.3 算出結果

図9に原研計算センタにおける損益分岐ラインの算出結果を示す。図の損益分岐ラインは,

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

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

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

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, バッチ参照) の場合である。

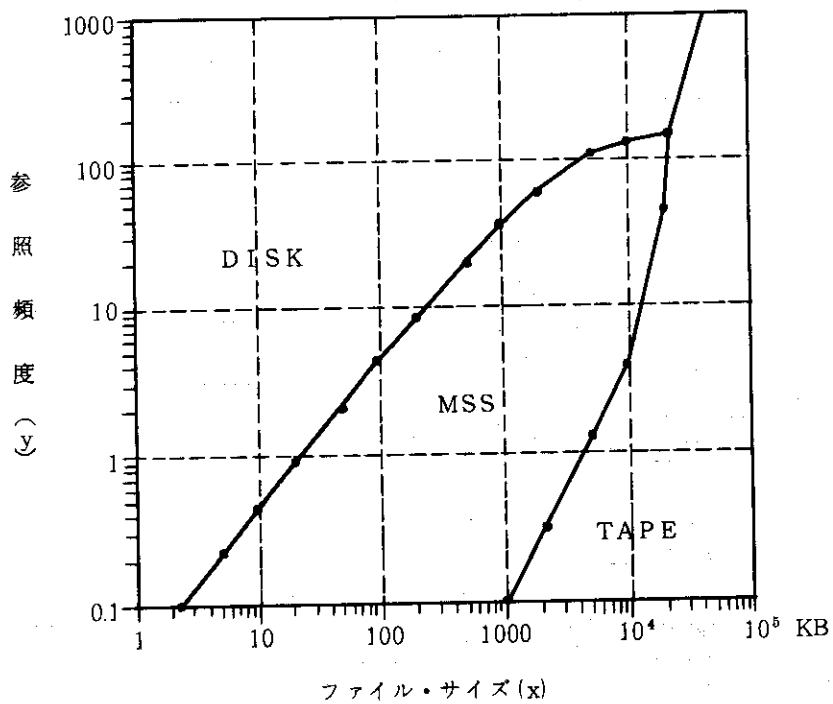


Fig. 9 Break-Even Lines for Disk, MSS, Tape.

6. 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、テープ装置がどの程度必要かという最適装置構成が大きな問題である。ここでは、サイズ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、5章で作成した損益分岐ラインを使って各記憶媒体に振分ける方法を示す。

6.1 振分け方法

4章で述べたように、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

$$z_1 = f_1(x) = 2 \lambda_1 (1 + \lambda_1 x)^{-3} \quad (6.1)$$

$$z_2 = f_2(y) = 2 \lambda_2 (1 + \lambda_2 y)^{-3} \quad (6.2)$$

で表わされた。又、ファイル・サイズと月平均ファイル参照回数の関係は、

$$\frac{1}{\lambda_2(x)} = y(x) = 3 + 7 \log_{10} x \quad (6.3)$$

で表わされた。

4章では、ファイル・サイズと参照頻度の関係は無視して、それぞれ独立に(6.1)、(6.2)の式を求めた。しかし、原研計算センタのファイル・データでは、任意のファイル・サイズをもつファイル群のファイル参照頻度分布も(6.2)で近似できる形状を示している。したがって、ファイル・サイズ(x)と参照頻度(y)に関する同時分布の確率密度は、

$$z = f(x, y) = f_1(x) \cdot f_2(y | x) \\ = 4 \lambda_1 \cdot \lambda_2(x) (1 + \lambda_1 x)^{-3} (1 + \lambda_2(x) \cdot y)^{-3} \quad (6.4)$$

で近似できる。

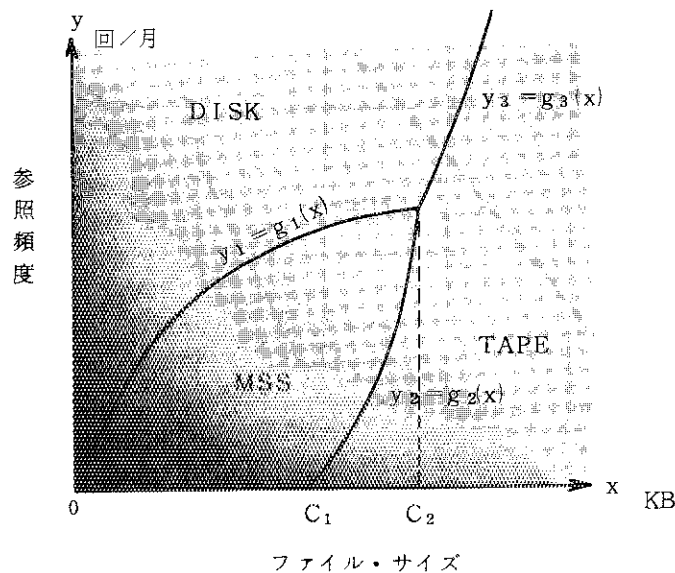


Fig. 10 File Demands and Break-Even Lines.

図10に示すように、(6.4)の同時分布をもつファイル需要に5章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ァイルの数、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

$$\text{ファイル数} = N \left\{ \int_0^{c_2} \int_{y_1}^{\infty} f(x, y) dy dx + \int_{c_2}^{\infty} \int_{y_3}^{\infty} f(x, y) dy dx \right\}$$

$$\text{総ファイル容量} = N \left\{ \int_0^{c_2} \int_{y_1}^{\infty} x \cdot f(x, y) dy dx + \int_{c_2}^{\infty} \int_{y_3}^{\infty} x \cdot f(x, y) dy dx \right\}$$

$$\text{総参照回数} = N \left\{ \int_0^{c_2} \int_{y_1}^{\infty} y \cdot f(x, y) dy dx + \int_{c_2}^{\infty} \int_{y_3}^{\infty} y \cdot f(x, y) dy dx \right\}$$

(2) MSS 需要量

$$\text{ファイル数} = N \left\{ \int_0^{c_1} \int_0^{y_1} f(x, y) dy dx + \int_{c_1}^{c_2} \int_{y_2}^{y_1} f(x, y) dy dx \right\}$$

$$\text{総ファイル容量} = N \left\{ \int_0^{c_1} \int_0^{y_1} x \cdot f(x, y) dy dx + \int_{c_1}^{c_2} \int_{y_2}^{y_1} x \cdot f(x, y) dy dx \right\}$$

$$\text{総参照回数} = N \left\{ \int_0^{c_1} \int_0^{y_1} y \cdot f(x, y) dy dx + \int_{c_1}^{c_2} \int_{y_2}^{y_1} y \cdot f(x, y) dy dx \right\}$$

(3) テープ需要量

$$\text{ファイル数} = N \left\{ \int_{c_1}^{c_2} \int_0^{y_2} f(x, y) dy dx + \int_{c_2}^{\infty} \int_0^{y_3} f(x, y) dy dx \right\}$$

$$\text{総ファイル容量} = N \left\{ \int_{c_1}^{c_2} \int_0^{y_2} x \cdot f(x, y) dy dx + \int_{c_2}^{\infty} \int_0^{y_3} x \cdot f(x, y) dy dx \right\}$$

$$\text{総参照回数} = N \left\{ \int_{c_1}^{c_2} \int_0^{y_2} y \cdot f(x, y) dy dx + \int_{c_2}^{\infty} \int_0^{y_3} y \cdot f(x, y) dy dx \right\}$$

6.2 算出例

原研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、ファイル数を4倍、平均ファイル・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として、それぞれの振分け結果を以下に示す。

(1) ケース1 (現在例: $N = 10000$, $1/\lambda_1 = 400\text{KB}$, $\bar{y} = 3 + 7 \log_{10} x$ 回/月)

	ディスク	MSS	テープ	合計
ファイル数	3481	6475	43	10000
総容量(MB)	559	3114	325	3999
総参照回数(回/月)	75447	28474	125	104046

(2) ケース2 (予測例: $N = 40000$, $1/\lambda_1 = 800\text{KB}$, $\bar{y} = 1.5 + 3.5 \log_{10} x$ 回/月)

	ディスク	MSS	テープ	合計
ファイル数	6093	33065	843	40000
総容量(MB)	1146	24744	6091	31980
総参照回数(回/月)	116538	132692	1344	250575

7. ファイル連続参照特性と記憶保存関数

大容量記憶システム(MSS)は、カートリッジ収納部を含む大容量記憶装置(MSF)とジョブが使うファイルをステージングするためのステージング・ディスク・サブシステムからなる。

ここでは、MSSのステージング・ディスク(以下、SDISKと略称する)の最適台数を算出するために、ファイルの連続参照特性を調査した。即ち、今日参照されたファイルのその直前の参照は何日前であったかを調べた。原研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、

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

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 $\beta(t)$ とこれを次の記憶保存関数で近似した値 $m(t)$ を図11に示す。

$$\begin{cases} m(t) = \int_{t-0.5}^{t+0.5} (1+t)^{-2} dt & (t \geq 2) \\ m(1) = \int_0^{1.5} (1+t)^{-2} dt \end{cases} \quad (7.1)$$

ここで、 $m(t)$ は参照間隔日数が t であるファイルの既存ファイルに占める割合である。

(7.1) 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 $\int_{0.5}^{1.5} (1+t)^{-2} dt$ では説明できず、 $\int_0^{1.5} (1+t)^{-2} dt$ で近似できる値と

なった。

これを図4の記憶モデルに対比させると、

$$\begin{cases} \text{前日参照} & \text{短期記憶 (連続参照)} \\ \text{2日以上前の参照} & \text{長期記憶 (非連続参照)} \end{cases}$$

と考えられ、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る。2日以上前に参照されたファイルの分布特性は、一度長期記憶として記憶された事象の記憶保存割合に等しいと考えられる。

解釈上、記憶モデル的には短期と長期の2つの概念で解釈したが、このファイル参照間隔の分布特性は(7.1)の1つの記憶保存関数で表現できることがわかった。これを8.1節で使う。

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: $\beta(t)$	記憶保存関数: $m(t)$
新規作成	14.4 %	0.0 %	} 60.0 %
1 日前	48.5	56.7	
2 "	10.5	12.2	11.4
3 "	5.3	6.2	6.4
4 "	3.4	4.0	4.0
5 "	2.4	2.8	2.8
6 "	1.6	1.9	2.1
7 "	1.5	1.7	1.5
8 "	1.0	1.2	1.3
9 "	1.0	1.2	1.0
10 "	0.7	0.8	0.8
11 ~ 20 日前	5.1	5.9	4.0
21 日以上前	4.6	5.4	4.7

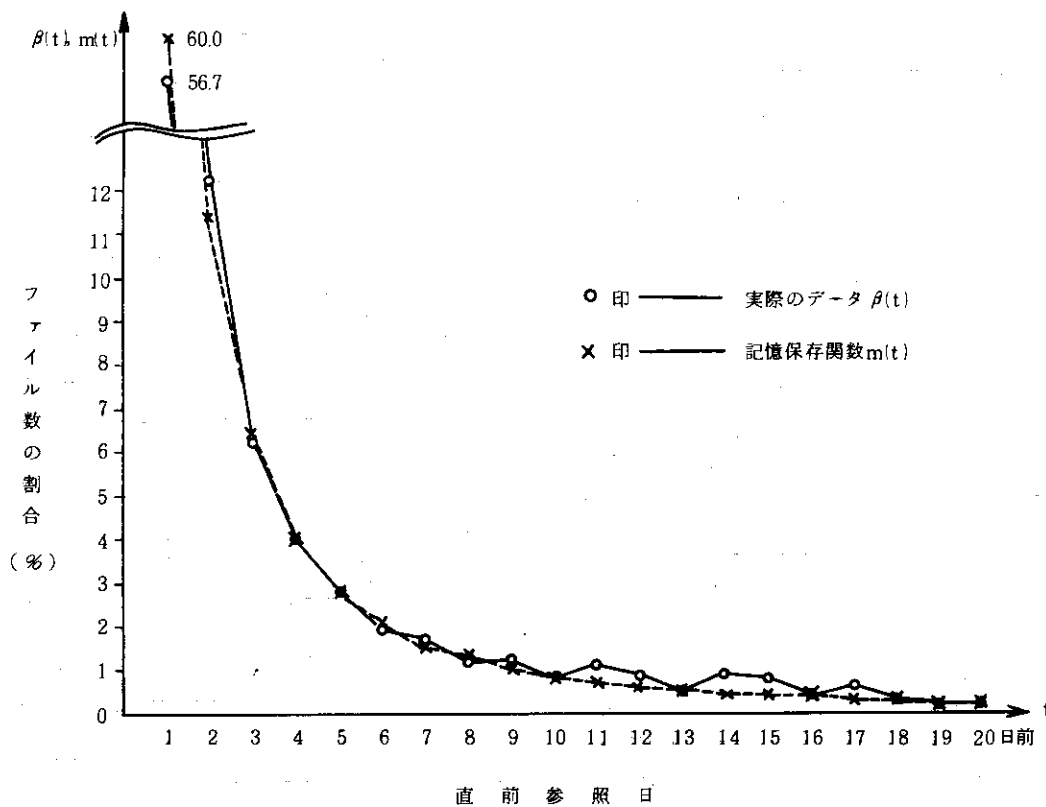


Fig. 11 Distribution of Continuous File References and Memory Retention Function.

8. ファイル管理システムの最適装置構成

6章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、最適装置構成の決定にはMSSのSDISK台数の決定が残る。

この章では、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、原研計算センタではこの報告書作成時点でまだ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。そこで7章で示したディスク・ファイルの連続参照特性を使って、後者のMSS方式を採用する場合、つまりディスクとMSSに振分けたファイル需要をすべてMSSに入れた場合の最適SDISK台数を求める方法を考え、これを次節に示す。前者のMSS方式を採用した場合、つまりMSSに振分けたファイル需要をMSSに入れた場合の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。

8.1 MSSのSDISK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時、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既に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。SDISKの台数を増すと参照するファイルがSDISK上に存在する確率は高くなり、ステージングの回数が減り、計算機システムとユーザのステージングによる待ち時間が減少する。しかし、SDISKの装置費が増加する。この損益分岐点、即ちSDISK装置費とステージングによって生ずる計算機システム、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。

以下にSDISKの最適容量の算出方法を示す。

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 : ひとつのファイルが1度参照された時、その日につづけて参照される平均回数 (回/日)

とおく。この時、1日平均参照ファイル数は、

$$d = \frac{k}{n} \quad (\text{ファイル/日}) \quad (8.1)$$

となる。ここで、

i : i 日前に参照したファイルまでがSDISK上にLRU制御でデステージされずに残っている

とする。この時、SDISKの必要容量 x は、図12より

$$x = s \cdot d \left\{ (i+1) - \sum_{t=1}^i (i+1-t) \cdot \beta(t) \right\} \quad (8.2)$$

で求まる。ここで、 $\beta(t)$ は7章で示した確率密度で、

$$\beta(t) = (1-\alpha) \cdot m(t) \quad (8.3)$$

である。 α は新規作成ファイルの割合、 $m(t)$ は(7.1)の記憶保存関数である。

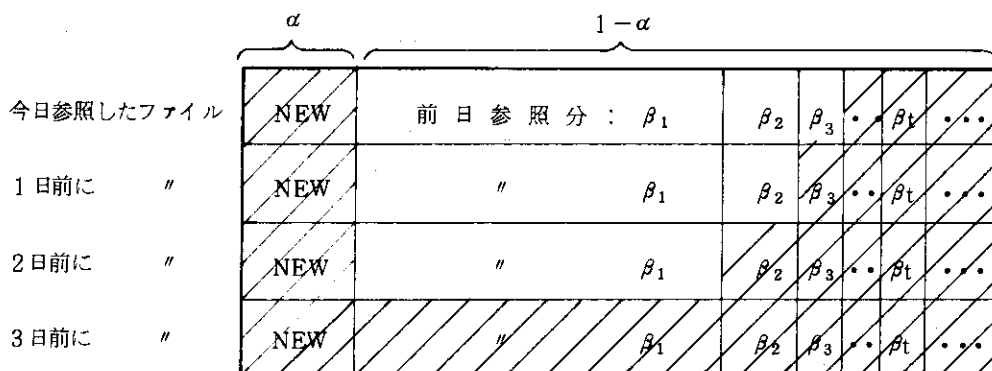


Fig. 12 A Necessary Capacity of Staging Disks.

(ex. the shaded part is for $i=3$ case)

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 P_s は、

$$P_s = \frac{\sum_{t=i+1}^{\infty} \beta(t)}{n} \quad (8.4)$$

である。この時の総費用 TC (円/日) は、

$$\begin{aligned} TC &= K_1'(x) + K_2(i) + K_3(i) \\ &= K_1(i) + K_2(i) + K_3(i) \end{aligned} \quad (8.5)$$

と表わされる。ここで $K_1'(x)$, $K_1(i)$ はSDISKの装置費をそれぞれ x , i の関数で表わしたものである。 $K_2(i)$, $K_3(i)$ 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, 次の式で求められる。

$$\begin{aligned} K_1(i) &= K_1'(x) = a_1 x \\ &= a_1 \left[s \cdot d \left\{ (i+1) - \sum_{t=1}^i (i+1-t) \cdot \beta(t) \right\} \right] \end{aligned} \quad (8.6)$$

$$K_2(i) = a_2 * TST \quad (8.7)$$

$$K_3(i) = a_3 * TST \quad (8.8)$$

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付録2に示す。

(8.5)の総費用TCを最小とする i を求め, (8.2)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。

この算出は, 各入力ケースに対し, i を0から0.1きざみで大きくして行ってTCが最小となる i を求めた。

8.2 算出例

原研計算センタの現在のディスク・ファイルの参照データ($s = 400$ KB, $k = 4000$ 回/日, $n = 4$ 回/日)のケース1の場合, 最適SDISK容量は $i = 0.6$ で513MBとなった。6.2節のケース2でディスクとMSSに振分けたファイル需要($s = 660$ KB, $k = 10000$ 回/日, $n = 2.5$ 回/日)の場合, 最適SDISKの容量は $i = 0.8$ で3591MBとなった。

ケース1は, CDC, 日電の仮想ファイル方式の場合で, ステージングにおける中央演算処理装置(cpu)のオーバーヘッド費等を考慮していないため, 最適SDISKの容量がかなり少く出た。しかし, 6.2節の結果と合わせて考えると, MSSに振分けられたファイルは, 毎日ステージングした方がコスト最適であると解釈できる。又, これはIBMタイプの仮想ディスク方式の場合のSDISKの台数もその日に参照されるファイル容量を満す程度あれば, 現在のハードウェア価格においてはコスト最適であることを意味している。

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, 人件費を現在より高くした場合の i の値の変化を図13に示す。例えば, ディスク装置費が1/3倍になり, 人件費が2倍になった場合, 最適SDISK容量は $i = 2.9$ で5151MBとなる。これは近い将来, ステージングにかかる時間(わずか10~15秒)もコスト的に無視できなることを示唆している。

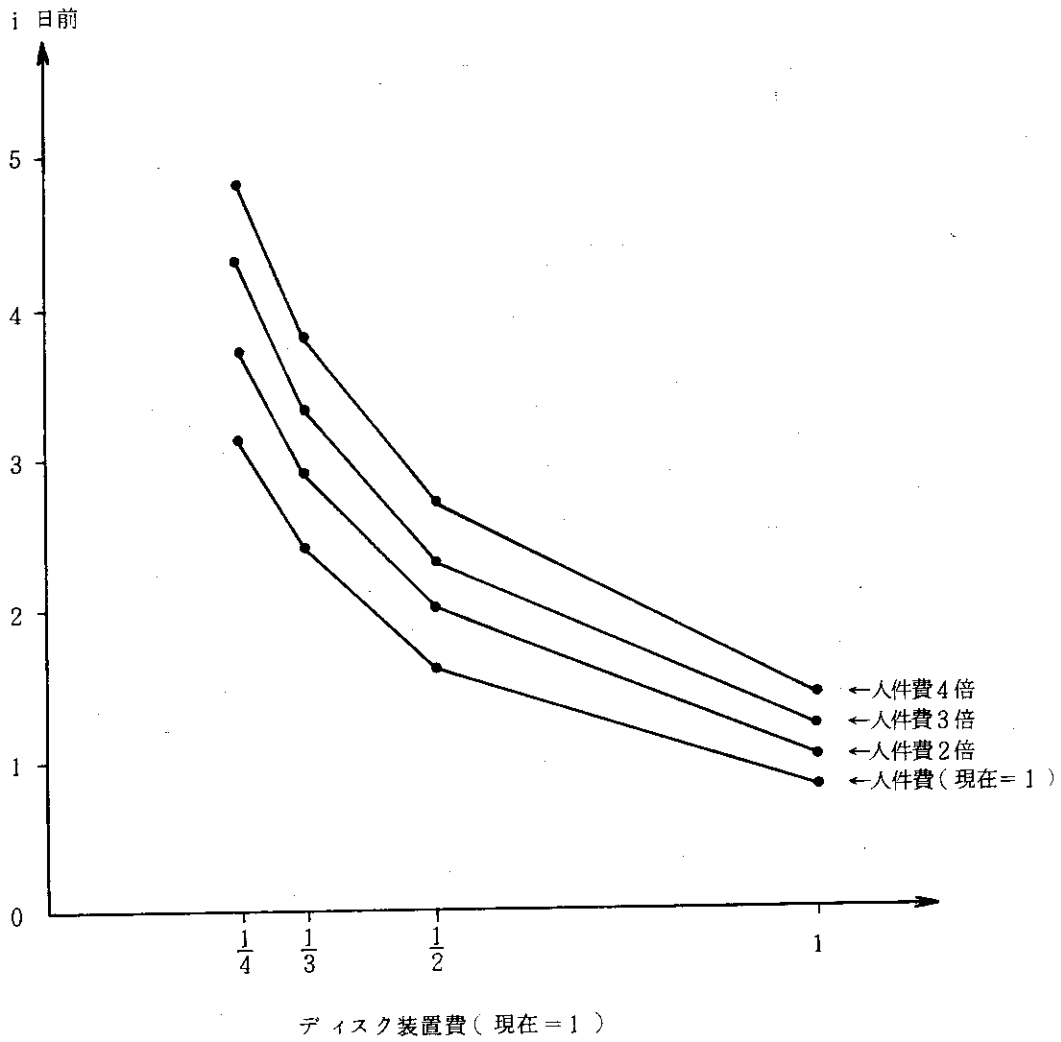


Fig. 13 Relationship between Staging Capacity, Human Cost and Disk Cost.

9. ファイル生成・消滅データと空き領域減少データ

図14に原研計算センタにおけるユーザ用ディスクの占有率とファイル生成・消滅量の変化を示す。ファイル生成量は、計算センタによる未使用ファイルの強制消去後に非常に多い。これはディスクが空いたため、大きなファイルを必要として取れなかったユーザがファイルを作成するためである。ディスク占有率が高くなるとファイル作成量が少くなる一般的傾向も見られる。顕著な傾向は曜日によるファイル生成量の違いの大きいことである。以下に昭和53年度下半期の各曜日の平均ファイル作成量を示す。

月曜日	$3344 \times 13 \text{ KB}$
火曜日	$4338 \times 13 \text{ KB}$
水曜日	$4723 \times 13 \text{ KB}$
木曜日	$5348 \times 13 \text{ KB}$
金曜日（土が出勤日）	$5137 \times 13 \text{ KB}$
金曜日（土が休日）	$3451 \times 13 \text{ KB}$
土曜日	$2294 \times 13 \text{ KB}$

以上より、ファイル生成量は、ディスク占有率、曜日、強制消去直後、季節変動等を因子とする回帰式で近似することは可能であると思われる。しかし、まだ十分なデータがそろっていないのでここでは以上のデータを示すにとどめる。

次に磁気ディスクにおけるディスク占有率と空き領域減少データを図15に示す。

ファイル作成の要求があった時、オペレーティング・システムの領域管理は指定された割当要求量を満す空き領域を1ボリューム内に見つけて割当ててゐる。この場合、連続する領域を取ることができないと数個の増分を使って要求を満す。例えば筆者らの計算センタでは増分が5個までだから、800KBのファイル作成を行うには、図15からディスク占有率が98%以下でないとなかなか難しいことがわかる。

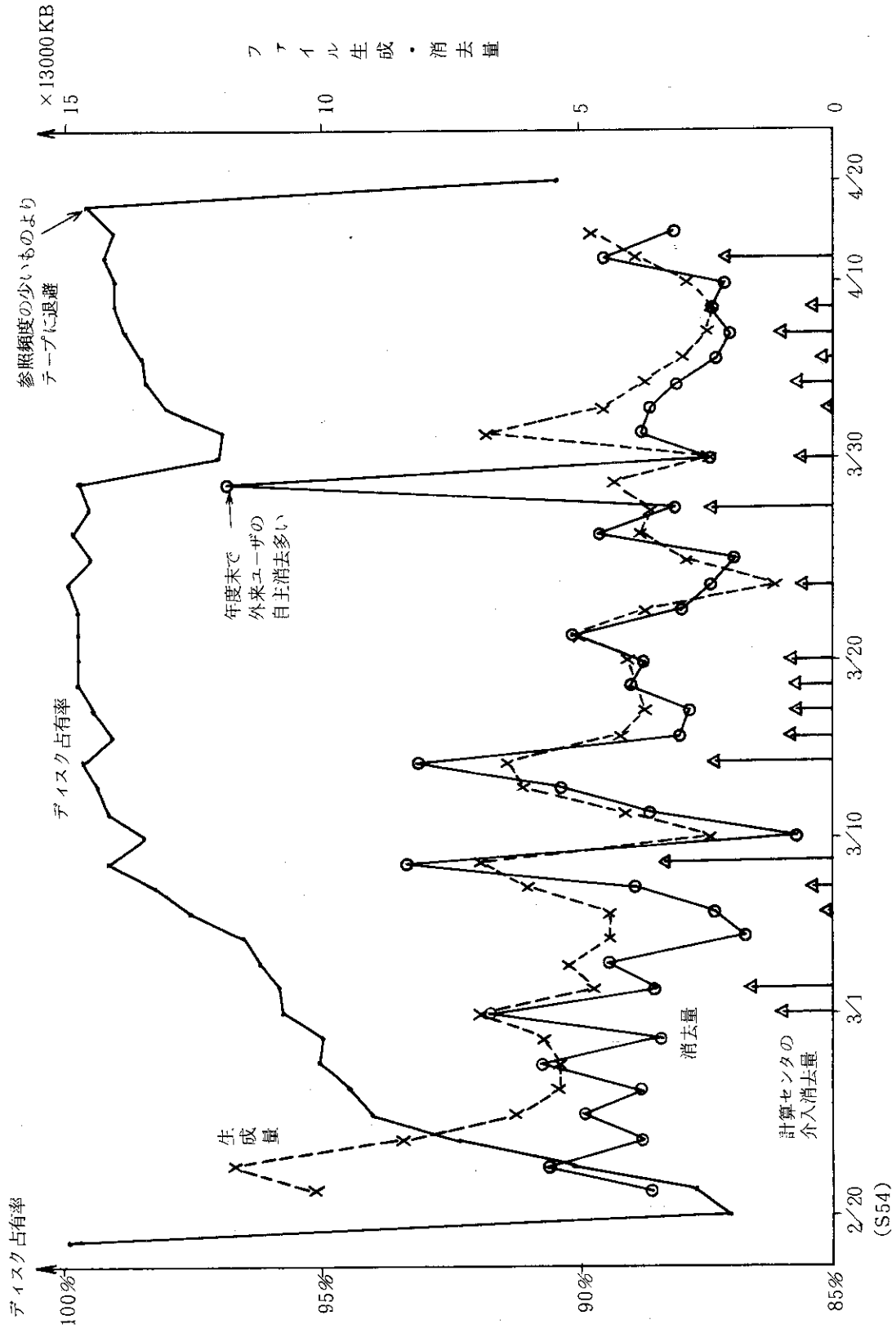


Fig. 14 Disk Utilization, Creations and Deletions of Files.

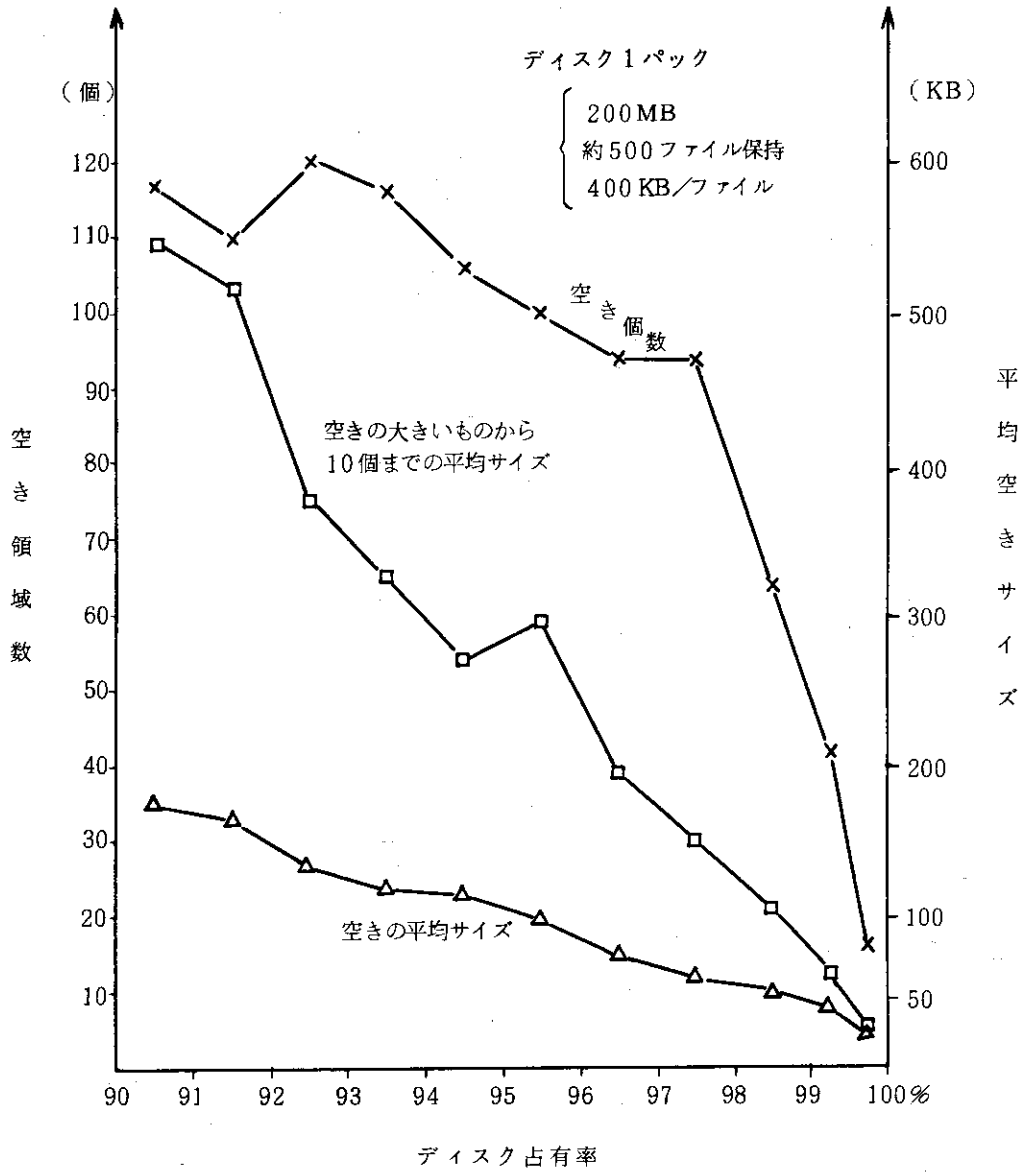


Fig. 15 Disk Utilization and its Free Spaces.

10. ファイルの階層的管理

10.1 2レベル・ファイル管理の概要

図16に示すように、筆者らの提案するファイル管理では、通常のファイル管理は各記憶媒体毎に行う。ファイル作成媒体の選択、媒体間のファイル移動等はユーザにまかせ、計算センタは一定の運用基準を提示するのみである。MSSとSDISKの間のファイル移動はオーバーヘッドの小さいLRU制御で行う。

ディスク、又はMSSの空き領域が、ジョブ処理に影響が出ると予想される管理点以下に減少すると、ファイル管理の統合部がファイル移動の全体調整に乗出す。この管理点は前章の空き領域減少データからあらかじめ決定することができる。空き領域残量の監視は1日1回程度行えば十分である。この方法を原研計算センタに適用すると、ディスクからMSSへのファイル移動は半月に1回、MSSから必要なファイルをテープに退避する作業は半年に1回程度と予想される。以下に管理方式の詳細について記す。

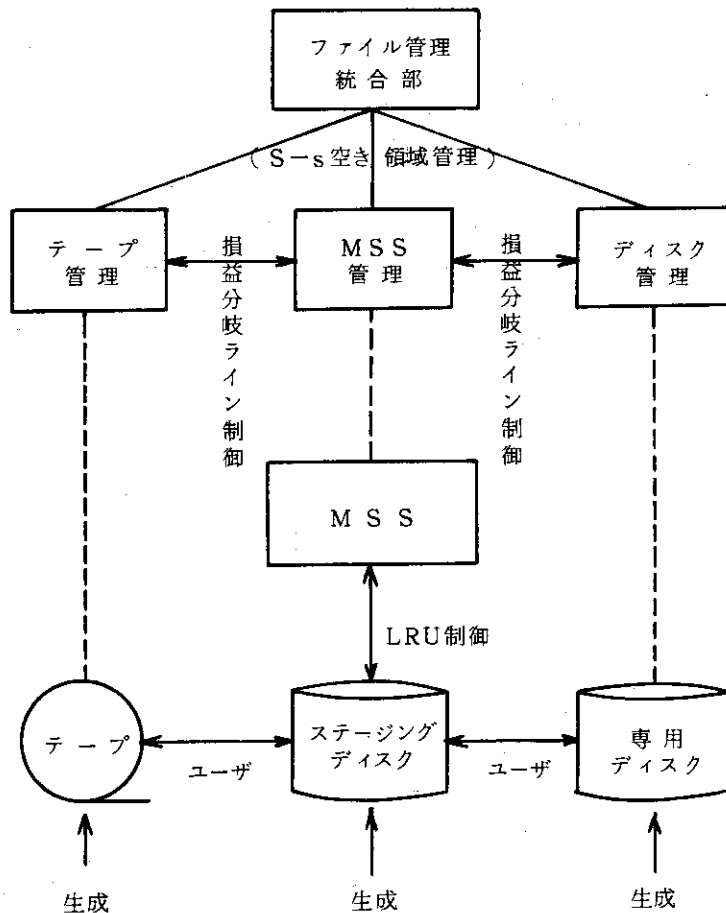


Fig. 16 Hierarchical File Management System.

10.2 S-s (補充点-消去点) 空き領域管理方式

在庫管理のS-s方式は、一定時間間隔で在庫量を調べ、その時の在庫量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量Sとsを決める発注方式である。

ディスクの空き領域管理では、図17に示すように空きの残りが下限sより少くなるとファイル管理統合部に連絡される。この時、統合部は必要なデータを収集し、10.4節で述べる参照頻度得点を計算し、その大小によって図のC1からA2の点までのファイル量をディスク、MSS間で移動させる。以上の操作は、現在採用されはじめた計算機無人運転の機能を利用すると1日のジョブ終了後自動的に実行できる。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・タイムと呼ぶ。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。MSSの場合のリード・タイムは約半月と予想される。これは、MSSから1回に消去するファイル量が非常に多いため、消去するファイルをユーザに通知し、ユーザ自身で必要なファイルをテープに退避してもらうためである。Sは次節で述べる最適消去量問題の解から決定される。sは9章のデータとリード・タイム間の空き領域減少量を考慮して決定される。

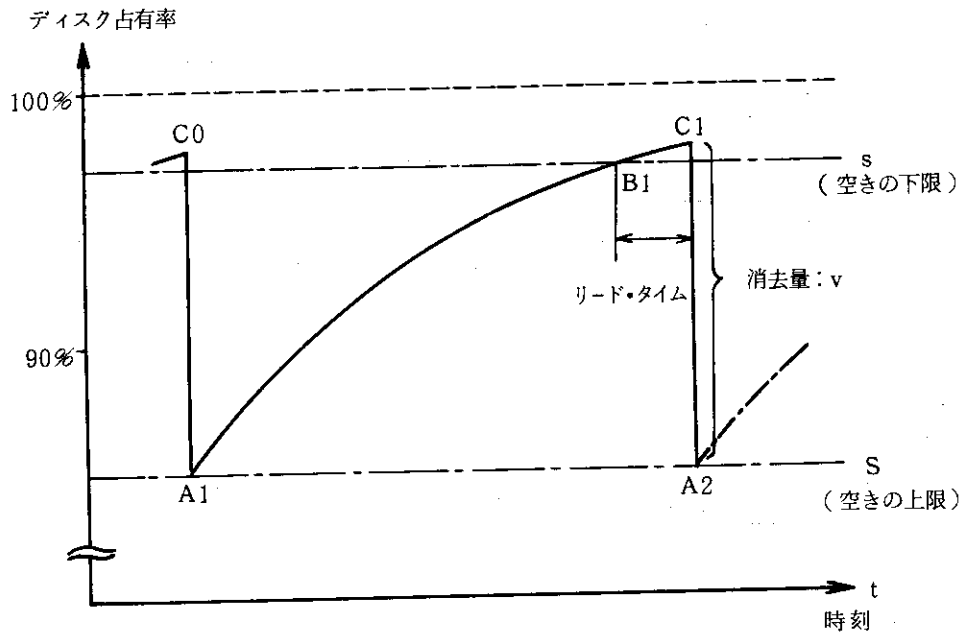


Fig. 17 S-s Method for Free Space Management.

10.3 最適消去量

図18に示す1回のファイル消去量: v (MB)を増やすとディスクの有効利用度(平均占有率)が落ちる。逆に消去量を少くすると消去回数増える。このディスク使用の機会損失費用と消去作業費用の総和が最小となるような1回の最適消去量を求める。

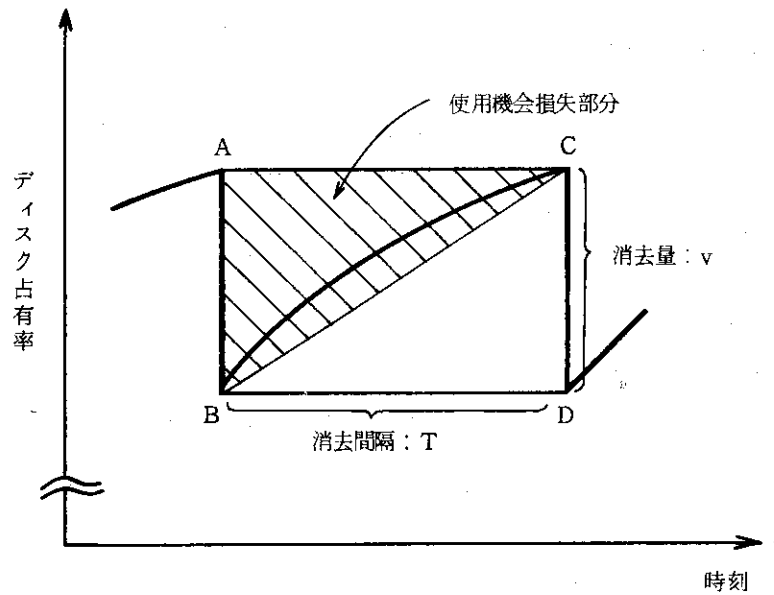


Fig. 18 Optimal Amount for Removal.

c : ディスクの使用機会損失費 (円/MB・日)

$a+bv$: 1回の消去作業費用 (円)

a が固定費, b が変動費

d : ファイル増加量 (生成量-消滅量) の平均 (MB/日)

とすると, ファイル消去作業間隔 T は,

$$T = \frac{v}{d} \quad (\text{日}) \quad (10.1)$$

となる。次にディスク占有率の増加曲線 \widehat{AB} を直線 \overline{AB} で近似すると, 図の斜線部で示されるディスク使用機会損失費 L は,

$$L = c \cdot \frac{v}{2} \quad (\text{円/日}) \quad (10.2)$$

となる。故に, 1日あたりの総費用 G は,

$$\begin{aligned} G &= \frac{1}{T} (a+bv) + \frac{cv}{2} \\ &= \frac{d}{v} (a+bv) + \frac{cv}{2} \end{aligned} \quad (10.3)$$

となる。 G を最小とする最適消去量 v^* は, (10.3) より

$$v^* = \sqrt{\frac{2ad}{c}}$$

と計算される。

$S - s$ 空き領域管理の空き領域量の上限 (補充点) S は、

$$S = s - \Delta d + v^* \quad (10.4)$$

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。

10.4 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が行う記憶媒体間のファイル移動は、以下で算出する参照頻度得点 K_t の大小によって行われる。

今消去期(t)の経過時間	T_t (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_t (回)
今期の参照得点	k_t
前期の参照頻度得点	K_{t-1}
今期の参照頻度得点	K_t

とする。この時、各ファイルの今期内の基準参照回数 y_t^* と今期の参照得点 k_t は、

$$y_t^* = T_t \cdot f(x)$$

$$k_t = \frac{y_t}{y_t^*}$$

で求められる。

ファイル移動の判断基準とする参照頻度得点の算出には、過去のファイル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を用いて次のように行う。

$$\begin{aligned} K_t &= a \cdot k_t + (1-a) a \cdot k_{t-1} + (1-a)^2 a \cdot k_{t-2} + \dots \\ &= a \cdot \frac{y_t}{y_t^*} + (1-a) K_{t-1} \end{aligned} \quad (10.5)$$

ここで、 a は指数平滑法における重み定数 ($0 \leq a \leq 1$) である。(10.5) より、今期の参照頻度得点 K_t は、各ファイルについての前期の参照得点 K_{t-1} と今期内の参照回数 y_t のわずか2つのデータを余分に記録することによって算出できる。

11. お わ り に

簡素で効率的な新しいファイル管理方式とその設計に必要な一連の手法を提案した。これらは、ファイル分布、ファイル連続参照特性が記憶保存関数で現実のデータによく適合するという事実を基盤としているため、実用的に意味あるものとなった。もし、ファイル分布等を情報処理の分野でよく使われる指数関数で安易に近似したりするならば、できあがる手法や方式は現実と遊離したものになる。この記憶保存関数はその形状からみて、情報処理分野での適用可能性を今後さらに検討すべきと思われる。

本稿のファイル管理方式と最近発表されたIBMのHSMとは類似点も多いが、本方式は、

- (1) ファイル管理システム全般にわたり、データと具体的な計算方式を示している、
 - (2) 記憶階層間のファイル移動は単なるLRU方式でなくファイル・サイズの大小も考慮した損益分岐ラインを使った制御方式によって行っている、
 - (3) 簡単なユーティリティ・プログラムで階層的自動ファイル管理が可能で、しかも実運用上の問題点をほとんど解決している
- こと等を特長としている。

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。1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。

謝 辞

本報告書を作成するにあたっての下記の方々よりいただいた御激励、御指導に対し深く感謝いたします。

計算センタ計算技術係の斉藤直之氏、山崎和彦氏には、ユーザ・ファイルの参照形態、参照データ等について御助言、御協力いただきました。システム係の中村康弘氏、オンライン係の石黒美佐子氏には、この報告書の御検討をいただきました。金沢コンピュータ・サービスの菊池克次氏、富士通CEの小笠原道生氏にはデータ収集に御協力いただきました。計算センタ室長の平川隆氏の御激励が本報告書作成の大きな力でありました。

11. お わ り に

簡素で効率的な新しいファイル管理方式とその設計に必要な一連の手法を提案した。これらは、ファイル分布、ファイル連続参照特性が記憶保存関数で現実のデータによく適合するという事実を基盤としているため、実用的に意味あるものとなった。もし、ファイル分布等を情報処理の分野でよく使われる指数関数で安易に近似したりするならば、できあがる手法や方式は現実と遊離したものになる。この記憶保存関数はその形状からみて、情報処理分野での適用可能性を今後さらに検討すべきと思われる。

本稿のファイル管理方式と最近発表された IBM の HSM とは類似点も多いが、本方式は、

- (1) ファイル管理システム全般にわたり、データと具体的な計算方式を示している、
 - (2) 記憶階層間のファイル移動は単なる LRU 方式でなくファイル・サイズの大小も考慮した損益分岐ラインを使った制御方式によって行っている、
 - (3) 簡単なユーティリティ・プログラムで階層的自動ファイル管理が可能で、しかも実運用上の問題点をほとんど解決している
- こと等を特長としている。

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。1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。

謝 辞

本報告書を作成するにあたっての下記の方々よりいただいた御激励、御指導に対し深く感謝いたします。

計算センタ計算技術係の齊藤直之氏、山崎和彦氏には、ユーザ・ファイルの参照形態、参照データ等について御助言、御協力いただきました。システム係の中村康弘氏、オンライン係の石黒美佐子氏には、この報告書の御検討をいただきました。金沢コンピュータ・サービスの菊池克次氏、富士通 CE の小笠原道生氏にはデータ収集に御協力いただきました。計算センタ室長の平川隆氏の御激励が本報告書作成の大きな力でありました。

参 考 文 献

- [1] 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).
- [2] 原研計算センタ Computer 情報, No. 30 (1978).
- [3] OS/VS 2 MVS Hierarchical Storage Manager : General Information, IBM JAPAN, GH35-0007-2 (Nov. 1978).
- [4] 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.
- [5] 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).
- [6] 藤井: 磁気ディスク対磁気テープの費用効果比, 日本原子力研究所, JAERI-M 7424 (1977).
- [7] 村松: 生産管理の基礎, 国元書房 (1973).
- [8] W. A. Wickelgren : Single-Trace Fragility of Memory Dynamics, Memory & Cognition, Vol. 2, pp. 775-780 (1974).
- [9] 阿部, 往住: 記憶のモデルについて, 数理科学, No. 179, pp. 22-27 (1978).
- [10] R.C. Atkinson and R.M. Shiffrin : The Control Processes of Short-Term Memory, Scientific American, pp. 82-90, (August 1971).
- [11] IBM 3850 大容量記憶システム (MSS) 入門, IBM JAPAN, GA32-0028-2-(1976).
- [12] 依田: 技術者の統計学, 宝文館出版 (1969).
- [13] 西田: 待ち行列の理論と応用, 朝倉書店 (1976).
- [14] 研野: 階層システム論, 共立出版 (1974).

付録1. 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

- ファイル属性, 参照形態比率 — 全ファイル参照に占める各ファイル属性, 参照形態の割合で, $\sum_i \alpha_i = 1$ である。

参照形態 \ 属性	ランダム	シーケンシャル
オンライン	α_1	α_2
バッチ	α_3	α_4

- ファイルR/W量 — 1回のファイル参照で実際にRead/Writeする量で, a_i はそのファイルのサイズの何倍R/Wするかを示すもので, 通常 $a_i \leq 1$ 。

参照形態 \ 属性	ランダム	シーケンシャル
オンライン	$a_1 x$	$a_2 x$
バッチ	$a_3 x$	$a_4 x$

- ファイル参照において書込み (Write) の行われる確率 δ
- 評価係数 (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 $\frac{\gamma_1}{N}$

アクセス時間 $\frac{\gamma_2}{N}$

ユーザの待ち時間に対する人件費負担率

{	オンライン参照の場合	γ_3
	バッチ参照の場合	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 — MSS 上のファイルを参照にいて、SDISK 上に
当該ファイルがない確率 β_m
- テープからワーク用ディスクへのステージ比率 — テープ上にあるランダム・アクセス
• ファイルを参照するとき、当該ファイルがワーク用ディスクにない確率 β_t
- 計算機システム費 (CPU + MEM) C_s (円/秒)
- チャンネル費 (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 (")
- テープ費 (近似式) $T_{a1} + T_{a2} \cdot x$ (円/月)
- 1 月の計算機稼動日数 D (日/月)
- ジョブ・セット時間 t_j (秒)
- テープ探し時間 t_s (")
- テープ・マウント時間 t_m (")
- テープからワーク用ディスクへのファイル読み込み実行時間 $\omega_1 + \omega_2 \cdot x$ (秒)
- テープ・ジョブの実行時間 $\omega_3 + \omega_4 \cdot x$ (秒)

(3) 費用項目

項目 \ 媒体	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	$U_{fd} \cdot xy$	$(U_{fm1} + U_{fm2} \cdot x) y$	$(U_{ft1} + U_{ft2} \cdot x) y$
一時ディスク費			$F_{dt} \cdot x$
テープ費			$T_{a1} + T_{a2} \cdot x$
オペレータ費			$T_{ao} \cdot y$
総費用	G_d	G_m	G_t

(4) 算出式

$$RAL = a_1 a_1 + a_2 a_2 + a_3 a_3 + a_4 a_4$$

$$RON = a_1 a_1 + a_2 a_2$$

$$ROF = a_3 a_3 + a_4 a_4$$

$$RPO = a_1 a_1 + a_3 a_3$$

$$RPS = a_2 a_2 + a_4 a_4$$

$$PON = a_1 + a_2$$

$$POF = a_3 + a_4$$

$$PPO = a_1 + a_3$$

$$PPS = a_2 + a_4$$

$$STM = \beta_m + \delta$$

$$STT = \beta_t + \delta$$

$$T_s = \max \{ 1/T_m, (A_d + B/T_d) / B \}$$

とおくと

$$S_{cd} = S_{cm} = RAL * C_c / T_d.$$

$$S_{1d} = S_{1m} = RAL * C_s * (\gamma_1 / N) / T_d.$$

$$S_{2d} = RAL * C_s * (\gamma_2 / N) * (A_d / b).$$

$$S_{2m1} = A_m * STM * C_s * (\gamma_2 / N).$$

$$S_{2m2} = T_s * STM * C_s * (\gamma_2 / N) + S_{2d}.$$

$$U_{nd} = (A_d + b/T_d) / b * RON * C_u * \gamma_3.$$

$$U_{fd} = (A_d + b/T_d) / b * ROF * C_u * \gamma_4.$$

$$U_{nm1} = A_m * STM * PON * C_u * \gamma_3.$$

$$U_{nm2} = T_s * STM * PON * C_u * \gamma_3 + U_{nd}.$$

$$U_{fm1} = A_m * STM * POF * C_u * \gamma_4.$$

$$U_{fm2} = T_s * STM * POF * C_u * \gamma_4 + U_{fd}.$$

$$F_{t1} = \{ PPO * (t_m + \omega_1) * STT + PPS (t_m + \omega_3) \} * C_t.$$

$$F_{t2} = (PPO * \omega_2 * STT + PPS * \omega_4) * C_t.$$

$$S_{ct} = \{ PPO * STT * (\frac{1}{T_t} + \frac{1}{T_d}) + \frac{RON}{T_d} + \frac{ROF}{T_t} \} * C_c.$$

$$S_{1t1} = PPO * STT * \omega_1 * C_s * (\gamma_1 / N).$$

$$S_{1t2} = (PPO * STT * \omega_2 + RPO / T_d + RPS / T_t) * C_s * (\gamma_1 / N).$$

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

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

$$U_{nt1} = \{ a_1 * STT * (t_j + t_s + t_m + \omega_1) + a_2 * (t_s + t_m) \} * C_u * \gamma_3.$$

$$U_{nt2} = (a_1 * STT * \omega_2 + a_1 * a_1 / T_d + a_2 * a_2 / T_t) * C_u * \gamma_3.$$

$$U_{ft1} = \{ a_3 * STT * (t_j + t_s + t_m + \omega_1) + a_4 * (t_s + t_m) \} * C_u * \gamma_4.$$

$$U_{ft2} = (a_3 * STT * \omega_2 + a_3 * a_3 / T_d + a_4 * a_4 / T_t) * C_u * \gamma_4.$$

$$F_{dt} = PPO * F_d / D$$

$$T_{ao} = (PPO * STT + PPS) * (t_s + t_m) * C_o.$$

(5) 損益分岐ライン

各記憶媒体における総費用は,

$$\begin{aligned}
 G_d &= D_1 * x + D_2 * xy \\
 &= F_d * x + (S_{cd} + S_{1d} + S_{2d} + U_{nd} + U_{fd}) * xy \\
 G_m &= M_1 * x + M_2 * xy + M_3 * y \\
 &= F_m * x + (S_{cm} + S_{1m} + S_{2m2} + U_{nm2} + U_{fm2}) * xy \\
 &\quad + (S_{2m1} + U_{nm1} + U_{fm1}) * y \\
 G_t &= T_1 + T_2 * x + T_3 * xy + T_4 * y \\
 &= T_{a1} + (F_{d1} + T_{a2}) * x + (F_{12} + S_{c1} + S_{112} + S_{212} + U_{nt2} + U_{ft2}) * xy \\
 &\quad + (F_{11} + S_{111} + S_{211} + U_{nt1} + U_{ft1} + T_{a0}) * y
 \end{aligned}$$

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]

$$G_d = G_m \text{ より}$$

$$y_1 = \frac{(D_1 - M_1) x}{(M_2 - D_2) x + M_3}$$

[MSS 対テープ]

$$G_m = G_t \text{ より}$$

$$y_2 = \frac{(M_1 - T_2) x - T_1}{(T_3 - M_2) x + T_4 - M_3}$$

[ディスク対テープ]

$$G_d = G_t \text{ より}$$

$$y_3 = \frac{(D_1 - T_2) x - T_1}{(T_3 - D_2) x + T_4}$$

付録 2. 総ステージング時間 (TST) の算出

$$TST = (PA + PD) * NS$$

$$PA = TACS / (1 - RA)$$

$$PD = (TDRD + PT) / (1 - RD)$$

$$RA = TA / (2 * TU)$$

$$RD = TD / (2 * TU)$$

$$TA = NS * TACS * 2 * (1 + W)$$

$$TD = NS * \{ TDRD * 2 * (1 + W) + PT \}$$

$$PT = \max \{ s / BUF, (AD + B / DK) * s / B \}$$

$$NS = d * \sum_{t=i+1}^{\infty} \beta(t) = d (1 - a) / (i + 1.5)$$

} ファイル参照間隔は指数分布を仮定

ここで,

PA : 平均アクセッサ (ACSSR) 処理時間 (秒)

PD : 平均データ記録機構 (DRD) 処理時間 (秒)

PT : 平均データ転送時間 (秒)

RA : 平均ACSSR使用率

RD : 平均DRD使用率

TA : 総ACSSR使用時間 (秒/日)

TD : 総DRD使用時間 (秒/日)

NS : 総ステージング回数

である。以下に入力データを示す。

TU : 計算機運転時間 (秒/日)

TACS : 待ちを除いたACSSR処理時間 (秒)

TDRD : " DRD処理時間 (秒)

W : ファイル参照において書込みの行われる割合

BUF : MSS からステージング用バッファへの転送速度 (KB/秒)

DK : バッファからSDISKへの転送速度 (KB/秒)

s : 平均ファイル・サイズ (KB)

B : ステージング用バッファ・サイズ (KB)

AD : ディスクのアクセス時間 (秒)

なお、ここでは2台のACSSR, 2台のDRDをもつMSSを仮定している。