

## 媒体操作機構を用いた大容量オンラインファイルシステムと そのコスト評価†

木ノ内 康夫<sup>††</sup> 櫻井 紀彦<sup>†††</sup> 小谷 尚也<sup>†††</sup>

可換媒体の機械的な操作により大容量のファイルシステムを構成するライブラリ方式について、オンライン環境下でのコスト評価方法を明らかにし、これを用いてサービス特性、ファイル特性に応じた各種方式の適用領域や構成法を考察した。大容量ファイル系の構成としては、(1)媒体をリード/ライト用のドライブに固定する常駐方式、(2)ドライブと媒体格納用の棚、媒体自動操作機構を持つライブラリ方式、(3)磁気ディスク装置上にファイルをステージングする機能を(2)に付加したライブラリ方式の3方式を対象とした。コスト評価は、ファイル系に対する主な要求条件としてスループットと容量に着目して行った。これを現状の装置に適用し、ファイルのオープン頻度、ファイル容量をはじめファイルオープン後の端末におけるユーザの考慮時間等オンラインサービスに特徴的な要因と適用領域の関係を考察した。またライブラリ装置の容量やステージング方式について検討し、光ディスクライブラリ方式でのドライブの台数不足等にステージング方式の採用が有効な領域を示した。

### 1. はじめに

磁気テープや書換え型光ディスク等の可換媒体を機械的に操作し、大容量のファイルシステムを経済的に構成する、いわゆるライブラリ方式が広く使われてきている。本論文はこれら大容量ファイルシステムをオンライン環境で用いる場合についてコスト評価方法を明らかにし、これを用いてサービス特性、ファイル特性に応じた各種方式の適用領域や構成法を考察したものである。大容量ファイル系の構成としては、(1)媒体をリード/ライト用の記憶装置(ドライブ)上に固定する常駐方式、(2)ドライブと媒体格納用の棚、媒体自動操作機構(ハンドラ)を持つライブラリ方式、(3)磁気ディスク装置(DK)上にファイルをバッファリング(ステージング)する機能を(2)に付加したライブラリ方式を対象とした。

コスト評価にあたっては、ファイルシステムに対する主な要求条件としてスループットに着目した。従来 Lum<sup>1)</sup>、藤井<sup>2)</sup>はファイル装置のコスト算出を容量によっているが、これに木ノ内<sup>3)</sup>が DK で用いた、スループットに基づく使用時間の条件を加えた。木ノ内<sup>4)-6)</sup>が光 DK を中心に概略を示しているが、ここでは装置構成、サービス特性とコストとの関係をさらに詳細化している。この評価方法を現状の装置に適

用し、ファイルのオープン頻度、ファイル容量をはじめファイルオープン後の端末におけるユーザの考慮時間やファイルオープン中のファイルへのアクセス回数等オンラインサービスに特徴的な要因と適用領域の関係を考察した。またライブラリ装置容量の大小、光ディスク(光 DK)の性能等を変化させた場合や光 DK に対してステージング方式を採用した場合について検討を加え、ステージング方式の有効な領域を示した。

### 2. 大容量ファイル系のモデル化

オンライン環境下で、多数のユーザ端末から同時に使用される数十 GB 以上の大容量のファイルシステムを対象とした。ファイルのオープン中はランダムにアクセスされるとし、磁気テープ装置(MT)使用時は DK へのステージングが必須とした。また光 DK は書換え型を対象とした。

#### 2.1 大容量ファイルシステムの構成と制御法

図 1 に示すように 3 種類にモデル化した。

##### (1) ファイル常駐(媒体固定)方式(方式 1)

記録媒体はドライブに固定され、媒体交換は行わない方式である。DK と光 DK を対象とした。光 DK の場合は媒体の片面のみを使用することになるため、片面媒体の使用を想定した。

##### (2) ライブラリ方式(方式 2)

機械操作により可換媒体を置き換える方式で、媒体を操作するハンドラ、格納用の棚からなるライブラリ装置とドライブから構成される。本論文ではハンドラと棚は装置として一体化されているが、ドライブは別

† Evaluation of Very Large Capacity Online Storage Systems Using a Media Handling Mechanism by YASUO KINOCHI (NTT Communication Switching Laboratories), NORIHIKO SAKURAI and NAOKA KOTANI (NTT Communications and Information Processing Laboratories).

†† NTT 交換システム研究所

††† NTT 情報通信処理研究所

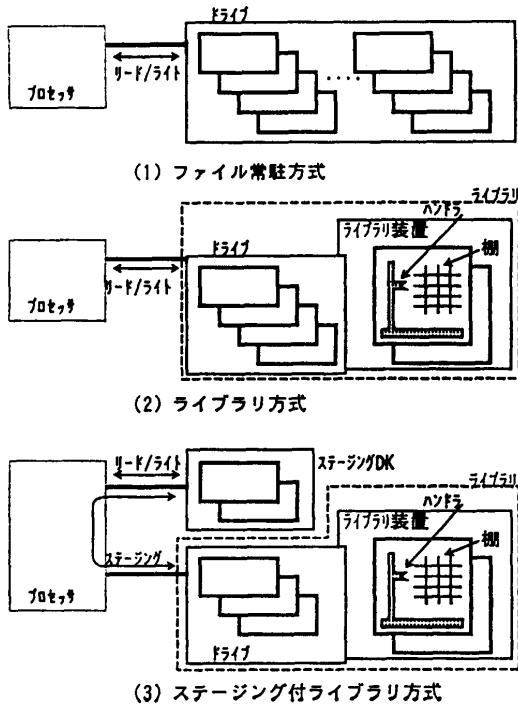


図 1 大容量ファイル系のモデル

Fig. 1 Models of very large online file systems.

装置として組み合わせる構成を想定した。ライブラリ装置とドライブを含めた構成はライブラリと呼称する。

ファイルオープンを契機に媒体が棚からドライブにセットされ、ファイルクローズでドライブから棚に戻されるものとした。光 DK をこの方式の対象とし、媒体は両面媒体を想定した。なお方式 3 も同様である。

### (3) ステージング付ライブラリ方式 (方式 3)

方式 2 にステージング用 DK を組み合わせた方式で、ステージング DK はデータチャネルを介してライブラリに接続される。ファイルオープンを契機に媒体がドライブにセットされ、オープンされたファイル内のデータはすべてステージング DK にロードされる。ロードが終了すると媒体は棚に戻され、ステージング DK に対してアクセスが行われる。処理が終了し当該ファイルをクローズした時点でファイルの内容が更新されていれば、当該媒体が再びドライブにセットされファイル全体が書き戻される。更新されていなければ書き戻しは行わない。対象装置は MT と光 DK とした。

## 2.2 サービス特性、ファイル特性のモデル化

サービス種別としては、TSS、DB、リモートバッチを想定した。これらサービスの特性をシステム内の総

ファイル容量、ファイルの容量 (ファイルサイズ)、ファイル操作頻度 (オープン/クローズ頻度) およびファイルオープン中のアクセス回数等に着眼してモデル化した。また端末でのユーザの考慮時間や通信に要する時間等オンラインファイル特有の時間的な条件を反映させるため、オープン中のファイルに対するアクセスの平均的な時間間隔を、サービス種別に応じて設定可能とした。

## 3. コスト性能算出の方法

本論文ではコスト評価の方法として、要求される容量とスループット (ファイルオープン頻度とリード/ライト頻度) を満足する範囲で近似的に最小コストを求める方法を採用した。異なる方式間でより厳密にコスト比較を行うためにはスループット条件、容量条件だけではなく、レスポンスタイム等応答時間に対する条件を一致させる必要がある。しかしここでは通常大容量ファイル系に対してはスループット条件、容量条件に比べて応答時間条件は二次的な要求となることから、応答時間条件を厳密にはコストに反映しない Lum、藤井らの方法を採用した。

算出方法の骨子は以下のとおりである。

(1) 大容量ファイルシステムは種々のリソース (装置、媒体) で構成されるが、独立に必要な量を加減できるリソースをユニットと定義した。ドライブ制御装置、ドライブ、記録媒体、ライブラリ装置、ステージング DK 制御装置、ステージング DK をユニットとした。

(2) 要求されるスループット条件、容量条件から各ユニットの延べ使用時間、総使用ファイル量を算出し、ユニットごとの必要量 (台数、枚数) を求める。ユニットごとに使用時間に起因する量と使用ファイル量に起因する量の両方を算出し、大きい値を選択する。

(3) 各ユニットについて必要量とユニット単体コストとの積を求め全ユニットについて総和をとる。

(2)でのユニット量の大きい値をとる操作は、Lum、藤井の方法を拡張した部分である。アクセス頻度がユニット 1 台の能力を越えた場合、容量面では装置に余裕があったとしても増設が必要となる。この影響を算定可能とした。これにより例えばライブラリ装置について、容量面で余裕があるが、ハンドラがボトルネックとなる場合についても台数算出が容易となる。これは木ノ内らが DK のコスト評価に用いた方

法をライブラリ方式に拡張したものである。

具体的なコスト算出式は以下のように表せる。

$$C_i = \sum_j (c_{i,j} \cdot N_{i,j}) \quad (1)$$

$C_i$  : 方式  $i$  での大容量ファイル系のコスト

$c_{i,j}$  : 方式  $i$  におけるユニット  $j$  1 台の価格

$N_{i,j}$  : 方式  $i$  におけるユニット  $j$  の必要量

ここでドライブ制御装置、ドライブ、記録媒体、ライブラリ装置、ステージング DK 制御装置、ステージング DK の順に  $j=1\sim 6$  とした。なお簡単化のためにチャンネル、プロセッサはコスト評価から除外した<sup>7)</sup>。 $N_{i,j}$  は時間、容量それぞれの条件から次式で示される。

$$N_{i,j} = \max \{A_{i,j}/\rho_{i,j}, B_{i,j}/(\beta_{i,j} \cdot D_{i,j})\} \quad (2)$$

$A_{i,j}$  : ユニット  $j$  の延べ使用時間

$\rho_{i,j}$  : ユニット  $j$  の平均使用率 (ユニット 1 台、単位時間あたりのユニット使用時間の平均)

$B_{i,j}$  : ユニット  $j$  に格納されるファイル容量

$D_{i,j}$  : ユニット  $j$  1 台もしくは記録媒体 1 枚当たりの物理容量

$\beta_{i,j}$  : ユニット  $j$  の平均ファイル収容率 (ユニット 1 台または媒体当たりの平均ファイル格納容量/ $D_{i,j}$ )

$\rho_{i,j}$  や  $\beta_{i,j}$  はスループットやファイルサイズの平均値からの変動に対する余裕を示す値である。これらは単なる余裕としてだけでなく、高信頼化のための冗長構成や、レスポンスタイム等の応答時間条件の影響を考慮して決定される項である。本論文では実システムでの使用状況に基づいて  $\rho_{i,j}$  や  $\beta_{i,j}$  を設定している。

$A_{i,j}$ 、 $B_{i,j}$  は方式別に次のように算出される。以下、式の右辺の添字については表記が煩雑になるため省略した。なお自明のことであるが、制御装置での容量条件、記録媒体での時間条件は該当する  $A_{i,j}$ 、 $B_{i,j}$  を 0 とした。

#### (1) ファイル常駐方式 (方式 1)

ドライブの延べ使用時間を次式で示す。

$$A_{1,2} = W \cdot Q \cdot R \quad (3)$$

$W$  : 単位時間のファイルオープン回数

$Q$  : ファイルオープン中のファイルアクセス回数

$R$  : 1 アクセス当たりのドライブの保留時間

ここで  $W$  は次式で示される。

$$W = Z \cdot X / Y \quad (4)$$

$Z$  : 大容量ファイル系内の総ファイル容量

$X$  : 単位時間に 1 ファイルがオープンされる回数

$Y$  : 平均ファイルサイズ (個々のファイル容量の平均)

本論文では  $Q$  を  $Y$  とファイルアクセス 1 回当たりのデータサイズ  $L$  で表される関数として次式で表した。

$$Q = q \cdot Y / L \quad (5)$$

$q$  はアクセスされる容量の割合を示している。例えば  $q \ll 1$  のときはファイルの一部のエリアがアクセスされることに対応し、 $q=1$  のときは全面またはそれに相当する容量がアクセスされることに対応する。 $R$  については、

$$R = T_a + L / V \quad (6)$$

で、 $T_a$  はシークタイムと回転待ち時間を含むドライブアクセスタイムを、 $V$  はデータ転送速度を示す。

また常駐方式のドライブについて  $B_{1,2} = Z$ 、媒体について  $B_{1,3} = Z$  である。ドライブ制御装置の  $A_{1,1}$  は式 (3)~(6) で  $T_a = 0$  とすることで得られる。

#### (2) ライブラリ方式 (方式 2)

ドライブの延べ使用時間は次式で示される。

$$A_{2,2} = W \{Q(T_{in} + R) + 2T_h \cdot M\} \cdot M \quad (7)$$

$T_h$  はハンドラによる媒体操作時間であり、 $M$  は 1 つのファイルが複数の記録媒体にまたがる場合も考慮した、ファイル当りに必要な媒体数の平均を示している (付録)。なお、方式 2 では光 DK の表裏両面の使用を想定したが、ファイル内全域へのランダムアクセスを可能とするため、1 つのファイルを表裏に続けて格納せず、別媒体に格納するとした。すなわち同一媒体の表裏には異なるファイルが格納される。簡単化のためこの前提を方式 3 にも適用している。また  $T_{in}$  はファイルオープン中のアクセスの平均時間間隔を示し、端末でのユーザ考慮時間、通信時間、プロセッサ処理時間の和の平均で与えられる。 $T_{in}$  をドライブの保留時間に加算したのは、媒体がドライブ上にある間、オープンされたファイルを含む媒体はドライブと共に、他の端末から使用不可としたためである。一方、方式 1 の場合は複数ユーザからドライブを共用可能としたため  $T_{in}$  をドライブの占有時間に加算していない。また 1 つのファイルが複数の媒体にまたがる場合、ファイルを格納した媒体すべてがドライブを同時に占有するとした。式 (7) の右端の  $M$  はこの効果を示している。

ライブラリ方式の場合、容量条件からのドライブ台数算出には意味がないので、 $B_{2,2} = 0$  とした。ドライブ制御装置については  $A_{2,1} = A_{1,1}$  である。

次にライブラリ装置については、装置 1 台当たりハ

ンドラは1台とした。延べ使用時間とファイル容量を次式で示す。

$$A_{2,4} = W \cdot M \cdot 2T_h \quad (8)$$

$$B_{2,4} = Z \quad (9)$$

媒体については同様に  $B_{2,3} = Z$  となる。

(3) ステージング付きライブラリ方式 (方式3)

ライブラリとステージング DK 間の一括転送や、ステージング DK を設置することによるコストが方式2におけるコストに加算される。また MT は部分更新が困難なため、複数のファイルを格納すると更新後の書き戻し処理が複雑となる。このためオンラインファイルとして実用的でないと考え、MT 使用時は記録媒体当たり1ファイルの格納を前提とした。

ドライブの延べ使用時間は次式で示される。

$$A_{3,2} = W \cdot (1+P) \{T_s + 2T_h \cdot M + T_r\} \quad (10)$$

ここで、 $P$  は内容更新のあるファイル数とステージングされるファイル数の比であり、ファイル内容の更新により処理終了後、元の記録媒体に書き戻されるファイルの割合を示す。 $T_s$  はステージングに要する時間で次のように表される。

$$T_s = \max \{ (T_{s1} + L_s/V_s) \cdot Y/L_s, (T_s + L_s/V) \cdot Y/L_s \} \quad (11)$$

$T_{s1}$ : ステージング DK のアクセスタイム

$L_s$ : ステージング時のリード/ライトデータサイズ

$V_s$ : ステージング DK のデータ転送速度

中括弧内の第1項はステージング DK 側、第2項はドライブ側の占有時間を示す。

ドライブ制御装置の延べ使用時間は、

$$A_{3,1} = W \cdot (1+P) \cdot Y/V \quad (12)$$

またライブラリ装置の延べ使用時間は、

$$A_{3,4} = W \cdot (1+P) \cdot M \cdot 2T_h \quad (13)$$

となり、さらに光 DK については、

$$B_{3,4} = B_{3,3} = Z \quad (14)$$

であり、一方 MT の場合には、次式となる。

$$B_{3,4} = B_{3,3} = Z/Y \cdot M \cdot D_{3,3} \quad (15)$$

また、ステージング DK の延べ使用時間は、

$$A_{3,6} = W \{ (1+P)(T_{s1} + L_s/V_s) \cdot Y/L_s + Q(T_{s1} + L_s/V_s) \} \quad (16)$$

となる。中括弧内の第1項がステージング動作による負荷を、第2項がステージング DK 上へのアクセスによる負荷を示す。ステージング DK 制御装置の  $A_{3,5}$  は式(16)で  $T_{s1} = 0$  としたものと同一である。

次にステージング DK 上で確保しなければならないエリアの容量  $B_{3,6}$  は、ステージング DK 上でのファイル平均滞留時間を  $T_k$  として、次式で近似した。

$$B_{3,6} = W \cdot Y \cdot T_k \quad (17)$$

$$T_k = Q(T_{in} + R) + (1+P) \cdot (T_s + 2T_h \cdot M + T_r) \quad (18)$$

(18)式の第1項がステージング DK 上へのファイルアクセスによる占有時間を示し、第2項はステージングのためのデータ転送による占有時間を示す。ここで、 $B_{3,6}$  は平均的なステージング DK の使用容量を示しており、実際には平均値からの変動を考慮した余裕を見込む必要がある。この余裕は(2)式に含まれる  $\beta_{3,6}$  を1未満で設定することで算入可能とした。

#### 4. 評 価

前章までに示した方法を用いて各種の大容量ファイル系のコスト算出例を示し、これに基づいて適用領域や構成を考察する。本評価例では装置の性能、価格や使用特性は、現状での大規模システムでの使用例をベースに表1に示すよう設定した<sup>8)-11)</sup>。総ファイル容量  $Z$  は各算出例とも同一とした。また、 $M$  と  $Y$  の

表1 装置諸元および価格の設定値  
Table 1 A list of parameters.

	パラメータ名	MT ライブラリ	光DK ライブラリ	DK
性 能	装置・媒体容量 $D_{i,1}$	0.2 GB	0.6 GB	1 GB
	ライブラリ容量 $D_{i,2}$	500 GB	500 GB	—
	アクセスタイム $T_s$	~0	130 ms	20 ms
	転送速度 $V$	3 MB/s	0.7 MB/s	4.5 MB/s
	媒体操作時間 $T_h$	10 s	10 s	—
	リワインド時間 $T_r$	50 s	—	—
	価 格	$C_{i,1}$ (ドライブ制御装置)	0.4	0.4
$C_{i,2}$ (ドライブ)		0.1	0.4	0.2
$C_{i,3}$ (媒体/1 TB 当たり)		0.75	5.0	—
$C_{i,4}$ (ライブラリ)		1.0	1.0	—
$C_{i,5}$ (ステージング DK 制御装置)		0.4	0.4	—
$C_{i,6}$ (ステージング DK ドライブ)		0.2	0.2	—

その他の設定値  
 総ファイル容量  $Z = 10$  TB  
 転送データサイズ  $L = 4$  KB  
 ステージングデータサイズ  $L_s = 64$  KB  
 デステージングファイル数の割合  $P = 1/3$   
 平均ファイル収容率 MT:  $\beta_{i,j} = 1.0$   
 ステージング DK, 光DK:  $\beta_{i,j} = 0.8$   
 平均ユニット使用率  $\rho_{i,j} = 0.4$

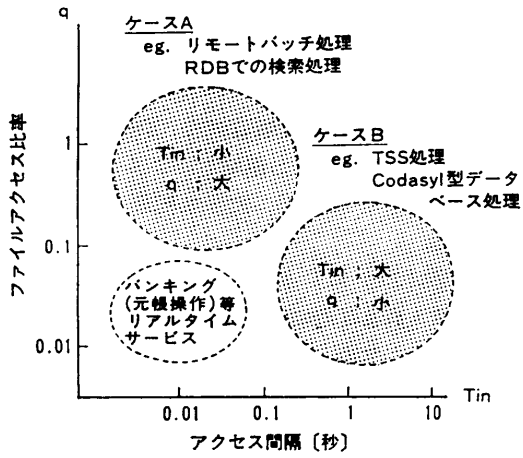


図2 サービス種別とアクセス特性  
Fig. 2 Models of access patterns in online services.

関係は付録に示す近似式を適用した。

考察にあたってファイルへのアクセス特性とサービス内容の関係を示すものとして  $T_{in}$  と  $q$  の関係に着目した。代表的な例として、図2に示すように  $T_{in}$  が小さく  $q$  が大きいケースAと、 $T_{in}$  が数秒以上と大きく  $q$  が小さいケースBに大別した。ケースAはファイルアクセスごとに端末に回答せず、 $T_{in}$  としてプロセッサ処理時間だけ加算される場合で、かつファイルのほとんど全面がアクセスされる場合を想定した。具体的なサービスとしてはリモートバッチ処理や、RDBでの検索処理等が対応すると考えた。ケースBはファイルの一部が参照・更新され、アクセスごとに端末に回答が返る場合である。具体的には TSS や Codasyl タイプの DB アクセスが対応すると考えた。

$T_{in}$  と  $q$  が共に小さいケースはバンキングでの元帳操作等リアルタイム処理が考えられるが、従来から DK が使用され、ライブラリを使用する場合は少ないと考えて算出例から除外した。

4.1 現状装置におけるコストと適用領域

現在主に使用されている方式として、ステージング付 MT ライブラリ方式、光 DK ライブラリ方式、DK と光 DK の常駐方式を選び、算出例を図3~7に示す。

(1) 各方式のコスト

図3, 4はケースA, Bにおけるファイルオープン頻度  $X$  とコスト  $C_i$  の関係を示し、図5には光 DK ライブラリ方式につ

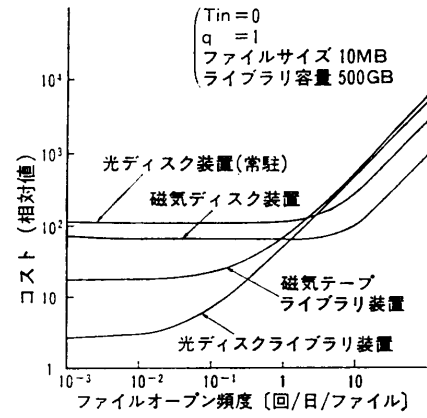


図3 ファイルオープン頻度とコストの関係 (ケースA)  
Fig. 3 Dependence of cost on frequency of file opening (case A).

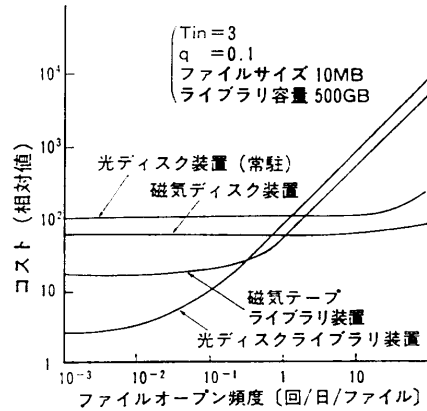


図4 ファイルオープン頻度とコストの関係 (ケースB)  
Fig. 4 Dependence of cost on frequency of file opening (case B).

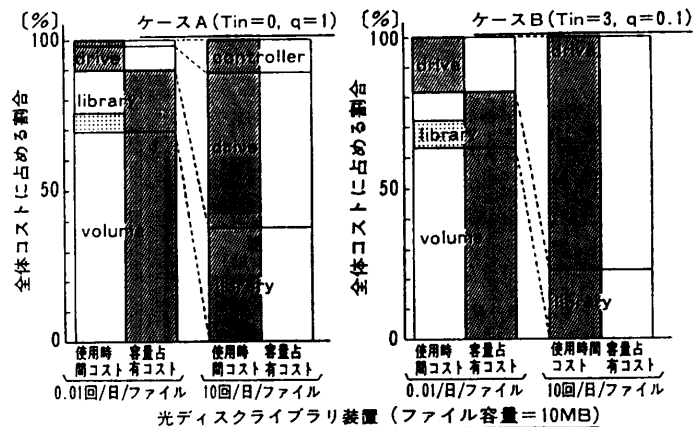


図5 各構成要因ごとのコスト寄与分の比較  
Fig. 5 Contribution of each component to subsystem cost.

いて、 $X$ を 0.01 と 10 [回/日/ファイル] とした場合、各ユニットごとのコスト寄与分を使用時間、容量のいずれがコスト決定に影響しているかを含めて示した。図 3~5 共にファイルサイズ  $Y$  は 10 MB 一定とした。これらの図から分かるように  $X$  が 1 [回/日/ファイル] 程度を境として、それより低頻度の領域でライブラリを用いた方式が、高頻度の領域で常駐方式が低コストとなっている。  $X$  が低頻度のときは媒体やライブラリ装置の容量要因がコストへ寄与しているが、高頻度のときはドライブ、ライブラリ装置の使用時間要因がコストのほとんどを占める。

次に、 $X$ を一定としたときの  $Y$  とコストの関係を図 6 に示す。ハンドラの使用頻度 [回/日/ファイル] は  $Z \cdot X/Y$  に比例することから、ライブラリを使用した方式のコストは、媒体 1 枚の容量程度まで  $Y$  の増加につれて減少し、それ以上になるとドライブ数の増加によりコストが上昇する。

(2) 各方式の適用領域

図 7 にファイルオープン頻度  $X$  と、ファイルサイズ  $Y$  を軸として、ケース B についての算出例を示す。図

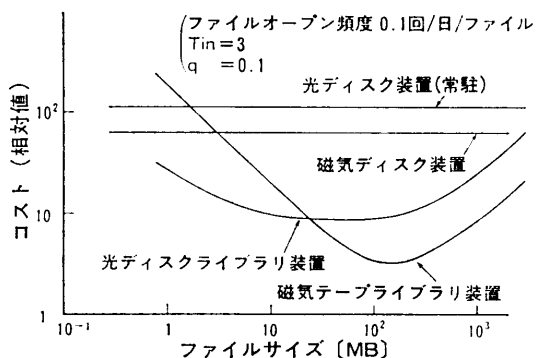


図 6 ファイルサイズのコストに与える影響 (ケース B)  
Fig. 6 Dependence of cost on file size (case B).

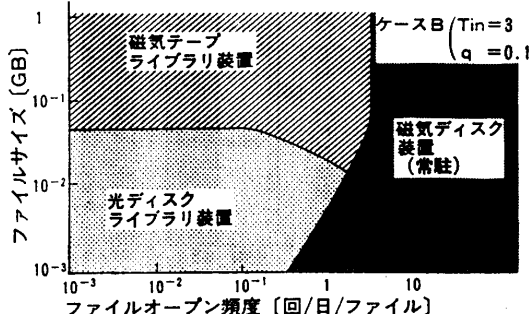


図 7 大容量記憶装置の適用領域 (ケース B)  
Fig. 7 Application range of very large file systems (case B).

中の各点に示す方式がその点でのコストの最も低い方式、すなわち適用領域を示す。全体として、ファイルサイズ  $Y$  に関わらず、ファイルオープン頻度  $X$  が 1 [回/日/ファイル] 以上の領域では、DK の常駐方式が有利である。これ以下の領域ではファイルサイズの小さい領域では光 DK が、大きな領域では MT が有利となる。

ステージング付 MT ライブラリ方式が  $Y$  の小さい領域でコスト高となるのは、MT の場合 1 巻当たり 1 ファイルとしたため、 $Y$  が小さいと多数の媒体やライブラリ装置が必要となるためである。光 DK では 1 枚の媒体上に複数のファイルを格納することが容易であり、ファイルサイズの小さい範囲で光 DK が低コストとなる傾向にある。ケース A についても、同様の算出結果が得られている。

4.2 装置特性、構成の変化がコスト、適用領域に及ぼす影響

ライブラリ装置容量の大小、光 DK の高性能化・ステージング適用の影響について算出例を示す。

(1) ライブラリ装置容量の影響

4.1 節で想定したライブラリ装置に比べて 1 台当たりの装置容量を 10 分の 1、コストを 5 分の 1 とした小型のライブラリ装置を使用したときの算出例を図 8 に示す。小型化により (単位容量×コスト) 当たりのハンドラ数が 5 倍になるが、コスト低減は  $X$  が 1 [回/日/ファイル] 前後の領域で 2~3 割程度となる。

(2) 光 DK の高性能化・ステージング方式適用の影響

光 DK について、現状の 3 倍程度まで高速化した場合とステージング方式を採用した場合の算出例を図 9、10 に示す。光 DK ライブラリを用いて端末ユー

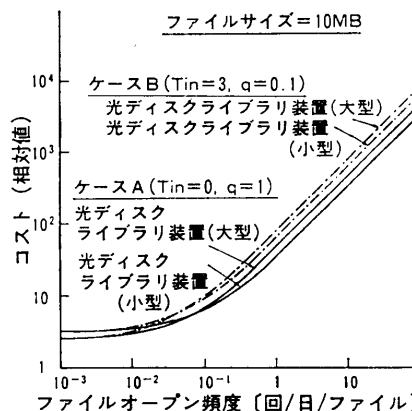


図 8 ライブラリ容量のコストに与える影響  
Fig. 8 Dependence of cost on library capacity.

ザに多数のディスク装置があるように見せた場合、端末側での考慮時間によりドライブの保留時間が増加し、必要なドライブ台数が急増するという問題が発生する。図9はドライブ数削減に対するドライブの高性能化とステージング適用の効果を示し、図10ではコストに対する影響を示した。ドライブの高性能化は $T_{in}$ が小さい場合に効果があるが、3秒前後より大きくなると効果が少ない。代わってステージング採用の効果が大きく現れてきている。MT使用時のステージング適用の狙いは、通常ランダムアクセスを可能とすることにある。しかし既に半沢が指摘しているように、DK上へのステージングにより、MTドライブが連

続してリード/ライトされることからMTドライブの保留時間が減少し、必要とするMTドライブ台数の削減が期待できる場合がある<sup>12)</sup>。ランダムアクセス可能な光DKの場合もステージングの適用により端末から保留されるドライブ台数の節減や経済化が可能

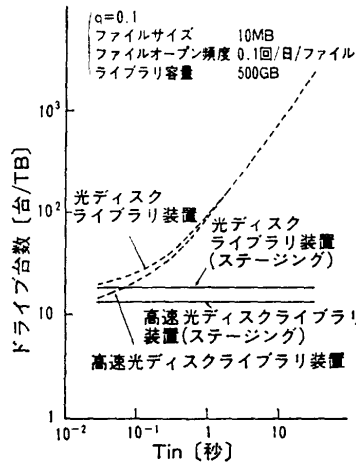


図9 ステージングによるドライブ数の削減効果  
Fig. 9 Reduction in the number of optical disk drive by staging.

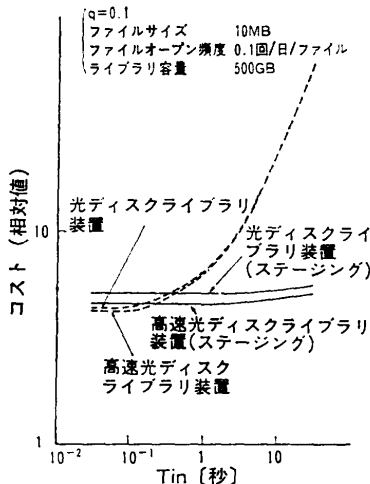


図10 ステージングによるコスト削減効果  
Fig. 10 Reduction in cost by staging.

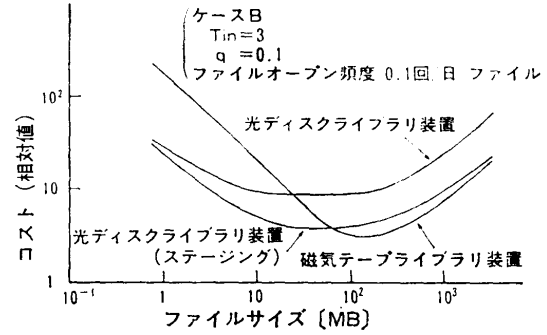


図11 ステージングによるコスト削減効果 (ケースB)  
Fig. 11 Reduction in cost by staging (case B).

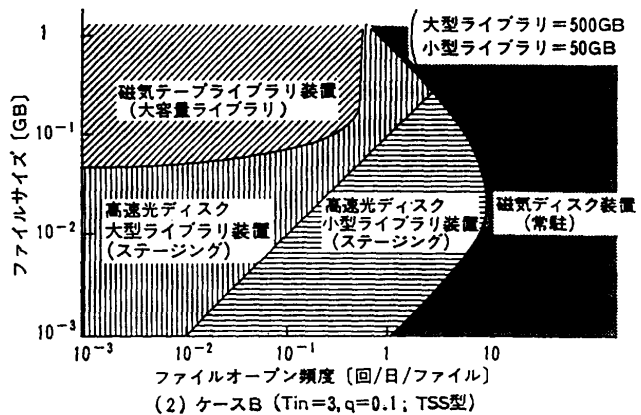
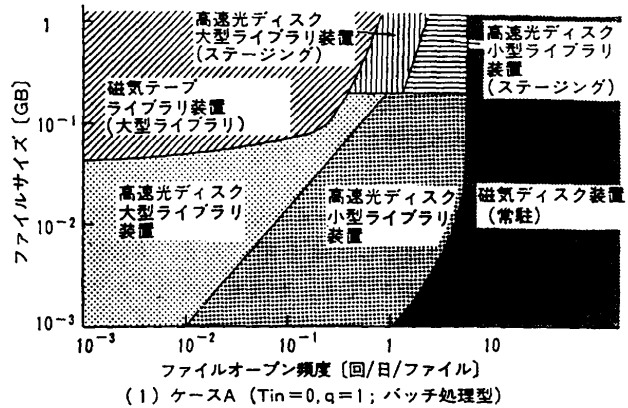


図12 ステージング方式およびドライブの高性能化の適用領域への影響

Fig. 12 Application range of very large file systems using several improved methods.

な場合があることを図 9, 10 は示している。また  $Y$  とステージング適用効果の関係を図 11 に示した。ステージングの適用によりファイルサイズの小さな領域で光 DK ライブラリ方式のコスト低減、適用領域の拡大が図られている。

### (3) 装置、構成の変化が適用領域に与える影響

以上の装置、構成の変化を加えたときの適用領域算出例を図 12 に示した。 $X$  が 1 [回/日/ファイル] 程度の領域でライブラリ装置の小型化が有効である。またバッチ指向のケース A ではステージング付光 DK ライブラリ方式のコスト低減効果は小さいが、TSS 指向のケース B では有用となっている。

## 5. おわりに

オンライン環境下でのライブラリ方式による大容量ファイルシステムについて、スループット、容量に着目したコスト算出の方法を明らかにし、これを用いてサービス特性やファイルアクセス特性、装置構成がコストに与える影響、適用領域を考察した。ここで示した方法を MT や光 DK のライブラリ構成を用いたファイル操作自動化方式や、大規模システムでのファイル構成方式の基本設計に用いて良好な結果を得ている。

本論文では装置使用率やファイル使用率等の装置使用特性はあらかじめ設定した値を用いたが、大容量ファイルシステムにおいても、レスポンスタイムやターンアラウンドタイム等の応答性能が明確な条件として規定されるような使用環境では、コストとのトレードオフとしてシステム設計に影響してくる場合がある。今後は応答性能条件を算入可能とすることで、より厳密な形で同一性能条件下でのコスト評価法を明確にしていく予定である。また並行して、半導体ディスク等高速系ファイルと統合したシステム全体のファイルコスト評価方式の明確化、サービス条件との関係の一層の詳細化等を図ってゆきたいと考えている。

**謝辞** 本研究の機会を与えて頂いた NTT 情報通信処理研究所石野福彌所長をはじめ、様々な面から長期にわたってご指導頂いた 同所松永俊雄研究企画部長、ソフトウェア研究所鶴保証城所長に深謝いたします。

## 参考文献

- 1) Lum, V. Y., Senko, M. E., Wang, C. P. and Ling, H.: A Cost Oriented Algorithm for Data Set Allocation in Storage Hierarchies,

*Comm. ACM*, Vol. 18, No. 6, pp. 318-322 (1975).

- 2) 藤井, 浅井: 階層的ファイル自動管理システムの設計, 情報処理学会論文誌, Vol. 21, No. 6, pp. 442-453 (1980).
- 3) 木ノ内, 山口, 櫻井: オンラインファイル記憶系装置のコスト評価に関する一考察, 情報処理学会論文誌, Vol. 27, No. 6, pp. 621-629 (1986).
- 4) Kinouchi, Y., Yamaguchi, H., Komatsu, T. and Sakurai, N.: A Survey of Very Large Capacity On-line File Systems Using Optical Disks, *IEEE COMPCON spring 87*, pp. 142-145 (1987).
- 5) 山口, 木ノ内, 花沢: MT 操作自動化方式に関する一検討, 昭和 58 年電子通信学会総合全国大会, 1602, pp. 6-87 (1983).
- 6) 小松, 櫻井, 木ノ内: 自動媒体交換機構付き光 DK の適用領域について, 第 34 回情報処理学会全国大会論文集, 2C-1, p. 441 (1987).
- 7) 小谷, 櫻井: 光ディスクを用いた操作自動化機構の制御方式に関する一考察, 第 39 回情報処理学会全国大会論文集, 7W-5, p. 1812 (1989).
- 8) 板生, 広野, 細川, 布谷: 磁気テープ記憶自動化システムの方式構成, 電気通信研究所研究実用化報告, Vol. 35, No. 7, pp. 697-704 (1986).
- 9) 6000 巻を収容するカートリッジテープライブラリ, 日経エレクトロニクス, No. 460, pp. 184-185 (1988).
- 10) 原, 中西, 吉澤, 山田: 高速 130 mm 径光ディスク装置のハードウェア構成, NTT 電気通信研究所研究実用化報告, Vol. 36, No. 10, pp. 1397-1404 (1987).
- 11) 三矢, 小口, 柿崎, 大久保: 大容量高速磁気ディスク装置の機構, NTT 電気通信研究所研究実用化報告, Vol. 36, No. 4, pp. 471-477 (1987).
- 12) 半沢: 磁気ディスク技法, 竹内書店新社 (1978).

## 付録 ファイル当たりの平均必要媒体数 $M$ について

$M$  は格納方法の差異や、ファイルサイズの分布の影響を受ける。例えば、MT 媒体では、同一媒体内の複数ファイル格納を不可としたため、ファイルの先頭は媒体の端に一致する。しかし光 DK 媒体では同一媒体内に複数のファイルを格納可能としたため、ファイルが媒体内の中間的な位置から始まる場合があり、ファイルが小さなときでも複数媒体に跨る場合がある。また一般に個々のファイルサイズを  $y$  としたとき、平均ファイルサイズ  $Y$  が等しい場合でも  $y$  の分布の形状により  $M$  は異なる値をとる。

### (1) 媒体、格納方式と $M$ の関係

- ① 磁気テープ媒体



サイズ  $y$  のファイルの格納に必要な媒体数を  $m$  としたとき、 $m = \lceil y/D \rceil$  となる。(ここで  $\lceil \eta \rceil$  は、 $\eta$  より大きい、等しい最小の整数を示す。)  $M$  は  $y$  の分布を用いて、次式で示される。

$$M = \frac{\sum_0^{y_{\max}} (f(y) \lceil y/D \rceil)}{\sum_0^{y_{\max}} f(y)} \quad (A1)$$

$y_{\max}$ : ファイルサイズ  $y$  の最大値

$f(y)$ : ファイルサイズ  $y$  のファイルの個数分布

## ② 光 DK 媒体

媒体の端からファイルの先頭までの容量を  $s$  とすると、必要な媒体数は  $\lceil (s+y)/D \rceil$  となる。ここで  $s$  が媒体内で一様に分布すると仮定して期待値を求めると、 $m = 1 + y/D$  となる。さらに  $m$  が  $y$  の線型な関数となっていることから、 $f(y)$  の形に関わらず直ちに、 $M = 1 + Y/D$  と求められる。

ただし光 DK 媒体の場合、媒体内を隅から隅までファイルで埋めるのではなく、ある程度の空きスペースを用意して可能な限りファイルの先頭を媒体の端に合わせ、運用の簡単化を図るのが普通である。このため  $M$  は通常  $1 + Y/D$  より小となる傾向にある。

### (2) 算出例における $M$ の設定

算出例への適用にあたっては、 $M$  について MT、光 DK 共に次の関数により近似し簡単化した。

$$M = 1 \quad (Y/D < 0.5) \quad (A2)$$

$$M = 0.5 + Y/D \quad (Y/D \geq 0.5)$$

MT 媒体では  $M$  は  $f(y)$  の影響を受けるが、 $f(y)$  として  $y_{\max} = 2Y$  とし、 $y$  が 0 から  $2Y$  まで、一様に分布する関数を仮定 (通常のシステムにおいて  $y$  のばらつきは比較的大きいことによる) すると、ほぼ上式に近い結果を得る。また光 DK 媒体については  $y$  のばらつきは  $M$  に影響しないが、ファイル間の空きスペースの活用による  $M$  の減少を考慮する必要がある。

この値は個々のシステムにより大きく異なるが、ここでは全体として 0.5 程度の減少と想定し、簡単化を図った。

(平成 2 年 3 月 19 日受付)

(平成 2 年 10 月 9 日採録)



木ノ内 康夫 (正会員)

1969 年早稲田大学工学部応用物理学科卒業。1971 年同大学院修士課程修了。同年日本電信電話公社 (現 NTT) 入社。以来、ハードウェアシステムの研究実用化に従事。現在 NTT 交換システム研究所第二プロジェクト主幹研究員。電子情報通信学会会員。



櫻井 紀彦 (正会員)

1979 年早稲田大学工学部電気工学科卒業。同年日本電信電話公社 (現 NTT) 入社。以来、記憶階層構成の研究実用化に従事。現在 NTT 情報通信処理研究所基本アーキテクチャ研究部主任研究員。



小谷 尚也 (正会員)

1986 年東京理科大学工学部電気工学科卒業。1988 年同大学院工学研究科修士課程修了。同年日本電信電話 (株) 入社。現在 NTT 情報通信処理研究所基本アーキテクチャ研究部において、記憶系装置アーキテクチャの研究に従事。

**論 文 誌 編 集 委 員 会**

委員長	益田 隆司				
副委員長	名取 亮				
委員	石畑 清	魚田 勝臣	浮田 輝彦		
	大田 友一	小池 誠彦	小谷 善行		
	佐藤 興二	島津 明	戸川 隼人		
	徳田 雄洋	永田 守男	原田 紀夫		
	松田 晃一	三浦 孝夫	毛利 友治		
	吉澤 康文				