

国立国語研究所学術情報リポジトリ

光ディスクを使用した大量日本語データの蓄積

メタデータ	言語: Japanese 出版者: 公開日: 2017-03-31 キーワード (Ja): キーワード (En): information storage and retrieval system, Write-Once-Read-Many optical disk, large-scale Japanese language data, distributed processing 作成者: 斎藤, 秀紀, SAITO, Hidenori メールアドレス: 所属:
URL	https://doi.org/10.15084/00001106

光ディスクを使用した大量日本語データの蓄積

齋藤秀紀

要旨：本稿は、追記 (WORM) 型光ディスクに大量日本語データを記録する場合の問題を把握するためにおこなった実験報告である。問題を確認する目的で、次の2点に焦点をおいて実験をおこなった。

- 1) 光ディスクの処理速度は、書き込みデータ長によってどう変化するか。
- 2) データ検索用の索引構成は、検索時間にどの程度影響を与えるか。

分散処理端末上で実験の結果、日本語データ (10.21Mバイト) の書き込み時間を325分 (4kバイト) から167分 (16kバイト) に短縮できた。また、書き込みデータ長を光ディスクのトラック長 (32kバイト) と一致させることによって、さらに50%高速化できる見通しを得た。

検索処理は、データ管理表と検索用索引の統合で、読み取り時間を900分から180分に短縮できた。そのほか、磁気ディスク装置と同程度の処理速度を得るためには、Read After Write 方式の改善、コンパイラのデータ転送単位の拡張、通信回線上のデータ長の変更が必要であることを示した。

キーワード：情報検索システム、追記型光ディスク、大量日本語データ、分散処理

On Large-Scale Japanese Language Storage on Optical Disks

Hidenori Saito

Abstract: This paper is a report of our experiment on large-scale language data (Japanese) storage on Write-Once-Read-Many (WORM) type optical disk. The experiment focused on the following two questions:

- 1) What effect does data length have on the processing speed of the optical disk?
- 2) To what degree does index organization affect data retrieval time?

We succeeded in reducing the time necessary for writing Japanese language data (10.21M bytes) on a distributed processor from 325 minutes (4k bytes) to 167 minutes (16k bytes). The results of the experiment suggest we could further reduce the writing time by 50% by making the length of written data accord with the track length of the optical disk.

The reading time was also reduced from 900 minutes to 180 minutes by integrating the data management table and data retrieval index. We suggested that in order to make the optical disk work as fast as the magnetic disk device we need to improve the Read-After-Write method, enlarge the data transfer unit on the compiler, and modify the data length on the communication line.

Key words: information storage and retrieval system, Write-Once-Read-Many optical disk, large-scale Japanese language data, distributed processing

1. はじめに

国立国語研究所にコンピュータが導入されて21年が経過した。コンピュータは、用語用字調査を始め各種の調査・研究に使用されてきた。この間、作成されたデータは、磁気テープ (Magnetic Tape: MT) 500本にもなる。しかし、MTに蓄積されたデータを長期間保存するためには、保存環境を安定させ一定期間ごとに複写が必要になる。MTによるデータ保存は、媒体の増加とともに保管場所、コンピュータ処理時間、要員の確保の点で限界がある。これらの問題を解決するために、新システムの切り替え(昭和60年2月)とともに光ディスクの利用を計画した。

光ディスクをコンピュータの記憶装置として利用する実験は、1982年に石島らがおこなっている [文献1]。検索用キー・コードをイメージ情報に付加した図形検索システムの試作である。また、最近では読み取り専用形式で電話帳、辞書などが提供され始めた。これに対して、利用者がデータを追記できる型は、イメージ処理が主体でありコード対応で実用化されているものはない。光ディスクは、大量データを長期間、安定した状態で蓄積できるといわれながらも実体は未知の部分が多い。

これらの点から、光ディスクにコード情報を記録する問題点と、実用化の可能性を探るため KWIC 用例検索システムを試作し実験をおこなった。使用したデータは、新聞 KWIC 46500 (255バイト) 件である (最終的には190万用例を記録する)。ハードウェア構成は、ACOS-S550 をホストとする N6300-55N インテリジェント端末と端末付属の光ディスク装置である。以下、本稿で述べる事項を示す。

2. 磁生媒体の問題点と利点。
3. 利用形態から見た光ディスクの効果。
4. 磁気ディスク (Direct Access Device: DASD)・光ディスクの階層性と索引。
5. 実験システムの概要。
6. 光ディスクの処理時間の測定。

2. 磁性媒体の問題点と利点

2.1. 磁気テープの特徴

大量データを長期間保存するためにMTを使用してきた。その理由として次の項目が考えられる。

- 1) 記録方式・媒体が標準化されている。
- 2) 記録容量が高く価格が安い。また、取り扱いが容易である。
- 3) 可搬性がある。
- 4) 順序処理向きであり不定長データの記録に適している。

第一の記録方式と媒体規格の標準化は、記録するデータに一定の枠をはめる。枠組の設定は、データ操作とプログラムの入出力処理を標準化させ、結果としてデータの利用期間・対象者を広げる。標準化は、装置・プログラム双方に理論的な基準を与えることになる。さらに、データ管理・保守などの間接作業を軽減させる。

しかし、装置・プログラムは、機器更新とともに一定期間ごとに変更をとめない、プロジェクトの途中で発生する人事異動、組織の変更もシステムの継続運用に影響する。装置・プログラム・運用組織は、それぞれ相補性が要求されながらも独立していることが必要である。

第二は、これまでデータ保存用媒体としてMTが利用されてきた理由である。次に述べるMTの「可搬性」、「順序性」と対の関係にある。取り扱いの容易性は、MTの構造が簡単であり軽量かつ小型であることによる。これらの特性は、長期保存を目的とした媒体では重要な条件である。

第三のMTの可搬性は、異なった処理で限られた装置を共有できること。大量のデータを記録保存する上で物理的な限界がないこと。保管は、棚などで管理できるため、少量データでは費用・運用面で省力化できる利点がある。しかし、MTが大量になる場合、保管場所、媒体の複製、履歴管理など自動化できにくいことが障害になる。

第四のMTのデータ処理の順序性は、一括入出力と分類処理ではDASD

に比べ実行速度が早い。反面、試行処理ではデータ探索が直列になるため応答時間が遅くなる。データの順序配列特性は、コンピュータの共同利用、ファイルの共有化がすすむにつれて利用上の制限が多くなる。

可搬性による問題

- 1) 使用データは、利用者による事前選択が必要である。また、選択した MT を装置へ装填しなければならない。
- 2) ファイルは、それぞれ独立しているため、管理情報の統合と標準化が困難である。
- 3) 大量データを保存する場合、保管のための空間が膨大になる。また、保存管理のために空調設備が必要である。
- 4) ファイルを長期間保存するためには、定期的な複写が必要である。

順序性による問題

- 1) 同一ファイルへの並行利用が困難。
- 2) 応答時間が遅く、会話処理が困難。
- 3) データの更新処理が複雑。
- 4) 階層化されたデータの直接処理が困難。

改善方法は、(1) MT 管理およびファイル内容をコンピュータによって自動管理できること、(2) MT にかわる長期安定型媒体を採用すること、(3) ファイル・アクセスのオンライン化をすすめることが必要である。これらの三点は、後述する DASD を使用した大量データ保存の一形式となり光ディスクを導入する条件の指標となる。

2.2. 磁気ディスクの特徴

DASD の利用法は、即時応答処理を中心にした、オンライン型処理が基本にある。また、媒体は大容量化の方向にある。オンライン処理は、大容量と高速ランダム機能を中心にデータの共有化、応答時間の短縮、ファイル管理情報の把握の容易性、複数の構造化されたデータの利用、多様なファイル・アクセス手段に対応できる。これらの特徴は、MT にない機能である。

そのほか、DASD の有効な利用法にデータの共有化がある。共有化は、デ

ータの集中管理と、データベース (Data Base : DB) による標準的利用を明確にする。DB 化は、データ中心のシステム開発をすすめるために効果がある。また、端末側の簡易ファイル言語とホスト側 DB との結合は、双方の機能を疑似的に統合しファイル操作と内容を一元化させる。

ファイルの疑似的結合の考え方は、コンピュータ本体を大容量ファイル処理装置とした、ブラックボックス化と仮想化の方向づけである。利用者は、階層化された装置を意識することなく、ファイル処理をおこなうことができる。ファイルは、一元化されたホスト・端末間の関係を通して、端末側の OA パッケージを操作すればよい。

磁気ディスクの第二の特徴は、高速ランダム処理機能である。ランダムと高速データ処理機能は、データへの並列利用、アクセス手段の多様性、データの階層化の手段を提供する。これは、DB を構築する上で利用できるデータの範囲を広げる。また、データの集中化は、データ管理・運用情報の統一的な扱いを可能にする。

しかし、MT に比べ現行の固定型 DASD は、記憶容量の増設に迅速に対応できないことが多い。装置価格が高くビット当たりの費用が高価になることが理由である。さらに、媒体の大容量化とデータの DB 化は、移行処理においてファイルの再編成を必要とする。DB システムの変更は、データ構造の再編成、理論構造の整合性、運用面で支援プログラムを複雑にする。これらは、DASD が長期的なデータ保存用媒体に不向きであることを示している。

一方、処理方式は、データの多様化とともに分散化がすすんでいる。分散化は、処理とデータが主なものである。

分散したデータの処理は、二つの方式がある。一つは、マイクロフレームリンクによる仮想ファイルの考え方であり、ほかの一つは LAN (Local Area Network) 上のファイルサーバとしての利用法である。マイクロフレームリンクは、データと階層的な装置のあいだを論理的に統合し、ファイルの仮想化による統一処理をはかる方式である。

LAN は、端末の高性能化、OA などの不定型処理の増加とともに、今後

の主流になると考えられる。しかし、データの分散化はデータを統一的に把握し、運用するためには問題が多い。この問題の対応には、集中的なファイルの運用・管理が必要である。実験システムでは、インテリジェント型端末装置を分散化し、ファイルの集中化は、光ディスクの大容量を利用する方法をとった。

以下、「4. DASD・光ディスクの階層性と索引」で分散されたファイルについて、DB との疑似的結合の有効性を述べる。

3. 利用形態から見た光ディスクの効果

3.1. 光ディスクの特徴

この節では、光ディスクの特性を明らかにするため使用対象者の特徴を示す。情報利用には、過去に蓄積された膨大なデータを試行的に検索する処理が基本にある。この面から見た光ディスクの利用法は二つである。

第一は、MT にかわる長期間のデータ保存用として、第二は、大量データのオンライン利用である [文献 3, 4, 5]。処理方式からは、ホスト・コンピュータ接続型、分散処理型の二つである [文献 1, 2]。追記 (Write Once Read Many: WORM) 型光ディスクの特性は次のような特徴がある [文献 6]。

- 1) 記憶容量が大きいため、ほかの媒体に比べビット当たりの情報蓄積価格が安い。
- 2) 物理的な記録方式を採用しているため、長期保存用媒体として安定している。
- 3) データの更新ができない特性を利用して、保存情報の改ざんを防ぐことができる。
- 4) コード、イメージの混在した情報の蓄積ができる。
- 5) 可搬性を利用したオフライン形式の蓄積が可能であり、媒体が小型であるため保管場所の確保が容易である。
- 6) 傷・ほこりに強く、保存環境は磁性媒体より簡単である。

7)同一情報の大量複製が容易である。

3.2. 光ディスクの問題点

実用化されている光ディスク・システムの多くは、イメージ処理が主体であり大量コードを記録している例は少ない（本システムもイメージ処理が基本にある）。メーカーが提供できる光ディスク装置も限られ、基礎実験に必要なソフトウェアも十分に整備されていない。光ディスクのシステム開発に対する環境は、未整備状態であるといえる。そのため、光ディスクをコード対応で実用化するためには、基本的な実験を通して技術上の問題を解決していかなければならない。

光ディスクの問題点

- 1) 技術的に創成期にあり、装置・プログラムともに利用法・規格が定まっていない。書き換え可能な光ディスクの研究もすすんでいる。
- 2) WORM 型光ディスクは、更新用ファイルに利用できない。
- 3) 入出力処理速度が遅い。
- 4) 索引および検索用キーを作るさい、柔軟性を維持するために書き換え可能な媒体との併用が必要になる。

第一の書き換えできる消去型は、まだ実験段階にある。WORM 型の標準化は、5 インチ型が先行しているが、8, 12 インチ型は対象になっていない。さらに、メーカーから提供される光ディスクは、それぞれ記録方式が異なり同一メーカー内でも互換性がない場合が多い。

一方、長期保存用の媒体は、記憶容量が大きいことが必要である[文献 6]。大容量化は、記録されているデータの統一的な把握を容易にし、管理費用、保存面積を少なくする。また、多様なデータを併記できることも特徴になる。そのほか、データ交換の問題がある。光ディスクなどの大容量化している媒体は、必要データの抽出が中心になり、それ自身で「閉じた」状態で利用できる。処理は、標準媒体への変換の過程に、プログラムを介することで対応でき、媒体の直接交換は少ないことになる。この点で、大容量の光ディスクの記録方式の標準化は、実務的には決定的な問題にならないと考えてよ

い。データ交換は、各装置間の接続点のみ留意すればよい。

第二の、WORM 型と消去型は、それぞれ特徴に応じて使い分けることが必要である。WORM 型は、書き換えができないが、データの改ざんを防止できる。データの変更ができないことは、DB 内のデータをファイル保守のたびに保存処理をおこなう必要がなくなる。大量データの管理には有効な特性である。

第三は、分散型でホスト・コンピュータとの間でデータ交換をおこなう場合の問題である。分散処理では、大量データの交換にも回線が使われる。現行の回線は、少量のコード処理を中心に設定(1 Mbps: mega bit per second)されており、大量データの長時間転送やイメージ処理では応答に長時間を要する。

次は、端末の処理装置の問題である。現行の端末装置に使用されているマイクロコンピュータは、16ビット型であり開発指向では処理能力が不足する。光ディスクのシステム効率を高めるためにも、記録容量、処理性能、回線速度の整合をはかることが重要である。これらが改善されれば、光ディスクの容量を十分に利用できる性能が得られる。

そのほか、OA システムを中心にした運用は、操作性の点でホスト側プログラムより優位にある。操作は、ファイルが端末で閉じている場合、データはパーソナル用検索が主体になり、大容量ファイルに対する端末での問題は、処理速度と記憶容量の二つに絞られる。

内部記憶装置は、光ディスクの低速性を改善するキャッシュメモリ (Cache Memory) に利用できる [文献 7]。キャッシュメモリは、検索対象になるデータを事前に予測し、必要情報を内部メモリ内に先き読みする方法である。データ要求時の逐次処理に比べ DASD・光ディスクの入出力回数を減らし、実行処理速度を高速化する。光ディスクの高速化に対する有力な手段の一つである。

第五の問題は、索引の分離問題として次の節で述べる。

3. 3. 光ディスク利用対象者

光ディスクの特徴と問題点を明らかにした。次に、光ディスクを使用して恩恵を受ける業務および対象者について述べ、運用面から生じる問題を検討する。対象者は、次の五つが考えらる。

- 1) 大量データの管理者。
- 2) 過去の膨大な時系列データを必要とする研究者。
- 3) 法的にデータ保存を義務づけられている情報管理者。
- 4) 多種の情報を併用する CAI 利用者。
- 5) 辞書や百科事典的な情報の利用者。

第一は、データ管理者の立場で利用する場合である。分散型処理では、ファイルは端末ともに分散され、管理に必要なデータ内容や利用履歴の把握が困難になる。一方、光ディスクは、ファイルの階層を記憶の大容量化によって物理的に一元化でき、端末側での DB 構築に道を開く。分散化されたファイルの端末側の集中管理である。つまり、階層化ファイルの統合と光ディスクを階層の最深部におくことによって、見せかけの大容量化をはかるファイルの仮想化による対応である。

仮想化の利点は、索引、検索用キー、データを利用目的に応じて分割できることにある。例えば、索引を磁性媒体へ、データを光ディスクへ分割配分し、両データを理論的に一体化する方法である。これは、ホスト・コンピュータ接続型の集団型光ディスクのデータ管理にも拡張できる。

利用法の第二は、長期保存の対象になるデータと、データの再利用の問題である。調査は、一回の調査が長期にわたる場合と、短期調査が繰り返しおこなわれるものがある。しかし、いずれも調査データ・資料の管理が中心になる。資料の管理には、調査目的・方法、データに対する取り扱い規定、過去の調査結果を常に参照できることが必要である。

光ディスクは、データの長期安定保存とともに調査表・地図・回答に対する記述内容を視覚できる形で保存できる。調査で発生する多様な資料を柔軟に処理できることは資料の利用を容易にする。

第三の問題は二番と同様である。データが改ざんできないことが特徴にな

る。一定期間、安定した状態で情報を保存する場合に有効である。

第四は、大量データを時系列的に見る場合である。図形・文字コード・混在情報の三種が対象になる。CAI で使用する各種の情報の一元化と、コンピュータとの接続を想定した構成である。光ディスクは、多様な情報を統一的に記録できるため、CAI に要求される多様な情報の検索・加工手段をDB化する指針を示す例となる。

第五は、光ディスクの一形態である読み取り専用 (Read Only Memory : ROM) 型の利用法である。辞書、百科事典的な情報をマイクロフィルム、写真、書籍にかわるマルチ情報の伝達用に利用する。この特性は、光ディスクが大量に複製できる特性を生かし、大量データの配布とともに外部情報の受け取りに利用できる。

以上、光ディスクの利用対象者を述べた。特に重要になるのは、保存データの取り扱いである。データは、文献、用例、調査表などの第一次資料となるもの、統計データの第二次資料となるもの、さらに分析後の第三次資料となるものがある。資料は、コンピュータ処理されたデータと、それ以外の音声・図形データも含まれる。現在、国語研究所に保存されているデータに、調査結果、文献、カード資料、用例がある。調査結果は、地図、調査表、録音テープ、語彙表などであり、カード、用例は、過去の用語調査で得られた用語の使用法を記録したものである。この中には、コンピュータ処理されたKWIC も含まれる。

新たに調査をおこなう場合、過去に調査されたデータを知ることは、調査をすすめる上で不可欠である。蓄積されたデータは、当時の調査方法の目的、精度を確認するための基礎資料であり、蓄積されている用例は、新調査での用語の確定に基準を与える。その点で、調査終了時に保存すべきデータの決定は、調査担当者に重大な責任を課すことになる。保存データの選定は、データの履歴・検索手段・階層性を規定する索引の在り方などを留意すべきである。また、文献類も調査を間接的にささえる資料として重要な位置にある。

4. DASD・光ディスクの階層性と索引

4.1. データの利用形態と索引

光ディスクの利用は、大量データの蓄積と検索手段の柔軟性にある。大量データの保存は、標準的なデータ格納方法とともに検索効率・容易性から見た索引のありかたが重要である。格納方法には、データに対する更新処理への対応も含まれるが、一部DBで使用しているものを利用できる。

DBのデータは、網構造、リスト構造が使われているが、データと索引構造は一致していた。DBの格納構造は検索主体で決められ、検索主体のデータ構造を抽象したものに索引があるとする考え方があった。DB利用者のデータに対する見方は、格納されたデータにあるのではなく、データ構造を写した索引にある。索引は、利用者がデータをどう認識するか（または「したか」）を示す指標になる。

光ディスクを利用する上で、索引についての問題は二つある。第一は、記録されたデータの追記にともなう索引の再編成である。WORM型光ディスクでは、索引とデータの関係が固定され、記録後の索引・データ部ともに更新はできない。WORM型光ディスクは、索引部およびデータ部が安定するまで最終的な記録ができないことになる。これは実務上問題である。

第二は、データの利用形態と索引のもちかたである。データの検索のさい、検索効率を上げるために索引が作られる。

表1 索引とデータの対応関係

No.	索引数	データ数
1)	1	1
2)	1	N
3)	N	1
4)	N	N

このとき、索引とデータの関係は、(1) 索引・データは1対1に対応、(2) 1対Nに対応、(3) N対1に対応、(4) N対Nに対応、の四種にわけられる。(表1)。

1)の形態は、単一データに対して利用目的が一種の場合の利用法である。索引・データの関係は、1対1に固定されるため索引・データともに光ディスクへの記録も固定される。同一索引の拡張・再編成処理では、新規作成またはメタ索引の設定が必要になる。

2) は、複数データを一個の索引で統轄する型である。この型の索引は、データ検索の機能から、データ間の関係を疑似的に統合する管理処理を含むことになる。データが安定している場合のみ索引も安定する。

3) は、単一データを共有する使用例である。利用者が一人のときは、複数の見方への対応であり、複数で利用する場合、辞書データなどの共有データに対する異なった見方を示す設定になる。ただし、索引の追加は、1) と同じ追記処理で対応が可能である。

4) は、複数の利用者が、複数のデータを専用または共有する場合である。この形態は、データ検索の役割と、データ管理の両方の機能が要求され、索引は検索のための機能からデータ管理への機能に拡張される。この中には、索引を統括するメタ索引も含まれる。

複数データを多数が共用する場合は、ファイルを総合的に管理するため、より上位の索引が必要になる。複数の索引の統合は、逐次処理におけるデータ追加と索引部の再編成に対応できる支援プログラムが必要である。

以上、二つの問題の解決案は、(1) 索引作成用の支援プログラムに OA 用として提供されている簡易ファイル言語を使用する、(2) 支援プログラムおよび索引は、端末の制御装置に付属する DASD におくの二点である。

支援プログラムの利用は、光ディスクの物理特性から索引部を独立させる。また、データ追記のさいに生じる、複数索引の総合管理と再編成、データが安定性するまでの一時記録に対応できる。これによって、光ディスクのデータ管理機能は、一部であるが端末側の索引処理に吸収できる。この方式は、端末でも集団型と同じ処理が支援プログラムで対応できることを示している。

光ディスクは、大量データ処理を指向した場合、ホストでの対応が強化されつつある。ホスト対応型は集中型である。しかし、ファイルの統一性、開発ツールとしての端末機能の二点が満足できれば分散をとるべきである [文献 8]。

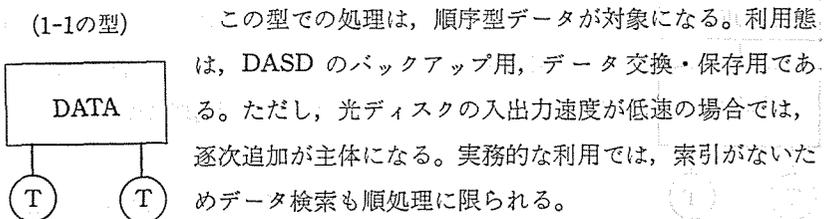
4.2. 索引とデータの分割配置

データの検索効率を上げるためには、索引が重要な役割をはたしている。索引は、データに対する利用者の見かたを示したものであり、利用者が目的に応じて設けることが多い。このことは、データに対して利用者の見方が異なれば、複数の索引が併用されることになる。

一方、索引を使った検索は、データの記録構造と媒体の特性に影響をうける。また、WORM 型ディスクは、記録したデータの変更ができないことが DASD 型媒体と異なる。このような索引への要求は、互いに矛盾した条件を満足させなければならない。特に、WORM 型媒体に対する追記処理と索引の更新問題は、光ディスクを利用する上で解決しなければならない最初の課題である。索引は、データ管理の指標であり、データの追記・削除にともなう管理・制御にも拡張できる。

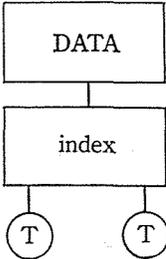
本節では、これらの条件について、プログラムで対応できる範囲と試案を述べる。試案は、索引・データの更新部分と固定部分を分離し、利用目的に応じて DASD・光ディスクに分割配置する方法である。DASD と光ディスクは、端末側に付属している装置を使用する。端末は、ホスト・コンピュータに対して分散処理形態をとる。索引処理用のソフトウェアは、簡易表操作言語を想定している [文献 9]。

データおよび索引分割形態は、光ディスク側では、(1) データのみ記録、(2) データと索引の併記、の二種である。端末側は、(1) ノンインテリジェント型端末を使用した場合、(2) 索引のみ分散、(3) データと索引の分散の三種である。それぞれの基本機能は、次の組み合わせに分類できる。

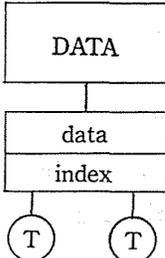


分離処理の基本型である。索引とデータの関係は、1対1 大文字： 光ディスクへ記録 に対応する場合と、N対1 の二種がある。しかし、いずれ

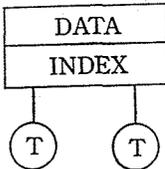
小文字：
 ① 端末DASDへ記録
 ② 端末装置
 (1-2の型)



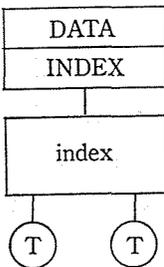
(1-3の型)



(2-1の型)



(2-2の型)



も端末側で独立して利用できるプログラムが必要である。端末側のプログラムは、簡易表操作言語またはDBでの対応がある。これらのソフトウェアは、索引の作成・登録・更新・検索処理のプログラム作成から利用者の労力を軽減させる。また、索引の論理構造をDASDの格納構造から独立させ、データ管理の統一を可能にする。更新処理は、索引部が肥大しないかぎりこの部分で吸収できるため、光ディスクのデータ更新問題は、一部であるが解決する。

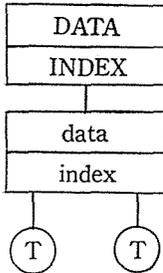
データが安定するまでの、暫定的なデータを端末側DASDへ分割する型である。索引は、光ディスクへのデータ追記、端末側データの再加工、それにとまう索引の再編成に対応できる。共通利用データは、光ディスク側におき、個別データはDASD上に分割配分する利用法も可能になる。

データ・索引ともに、光ディスクに固定する形式をとる。索引・データは、記録後の更新ができないため、安定が確認されたデータの最終処理である。複数の索引を必要とする利用法、辞書類など事前にキー内容を把握できるときに、それぞれ異なった設定が可能である。

光ディスク側には、固定できる共用データと索引をもたせ、端末側には共用データに対する個別索引をあてる型である。端末側の索引は、光ディスク側にある複数個の索引の統合用メタ索引として、また記追データに対して変更の可能性のある索引に利用できる。試行処理についても同様である。

光ディスク・端末間の索引とデータは、直列型では階層的になり、並列型では目的別の分離使用になる。例えば、光ディスクに共用データと索引を、端末側には個別データ

(2-3の型)



・索引をおく場合である。また、端末側で検索した情報を条件に、光ディスク側のデータを検索する処理にも利用できる。この条件は、索引部の多重構造の設定であり、シーラスなどのネットワーク表現やツリー構造の混在利用に拡張できる。

六種の構造について述べた。1-1, 2-1型以外は、すべて仮想ファイルの考え方が基本にある。実用面から見ると、1-2, 1-3または2-2, 2-3は、検索目的ごとに複数の索引を許しているため、利用目的に応じた最適構造をとることができる。DASDの索引機能と光ディスク上のディレクトリの併用は、修正データ処理を可能にする。データ追記と検索効率を上げるためには、光ディスク側にも索引が必要になるため、運用面では2-2, 2-3が最適型と考えられる。

索引の分割使用の型は、2-2, 2-3の二種である。このような索引の分割は二つのデータの見方を可能にする。一つは、利用者が光ディスクから必要部分を抽出し、個人用索引を作成する場合である。共有データの目的別再編成も含まれるが、索引の使いかたは並列的である。

ほかの一つは、索引の階層化である。検索効率を上げるための、メタ索引の設定に利用できる。検索用キーは、上位概念を端末側に下位概念を光ディスク側へ分割配分する型である。データ利用の支援や、データ管理情報の記録用にも対応する。広義には、データディクショナリ/ディレクトリシステム(Data Dictionary/Directory Systems: DD/DS)への発展であり、調査・文献資料の管理情報への拡張である [文献10]。

5. 実験システムの概要

実験は、ホスト・コンピュータ (ACOS-S550) と、インテリジェント端末 (N6300-55N) で構成する分散処理システム上でおこなった (図1)。光ディスクは、外部記憶装置として DASD (120M バイト) とともに端末に接続されている。端末に接続されている機器は、フロッピーディスク2台、低速

の漢字プリンタ、漢字ディスプレイの各装置である。なお、ホスト・端末間は、1 Mbps の宅内専用通信回線で結合している。

ホストと端末の処理分担は、ホスト側でデータを通信路に乗せる前処理を、端末側は光ディスクへデータを記録するための編集処理をおこなう。ホスト側の前処理は、MT に保存されているデータを通信回線の規定に調整する形式変換と、伝送路に低触するコードの変換である。データ転送単位は、通信規格にあわせ 255 バイトに固定した。

端末側は、光ディスクへデータを記録する物理データ長への変換と、管理ファイルの作成である。光ディスクへの書き込み単位は 16k バイトである。データ管理ファイルは、リラソ用、データのディレクトリ用、記憶管理用の三種である。

通信路は、端末・ホストに付属している DASD 上にデータ転送用ファイルを設け、回線に対する入出力バッファ機能にも利用している。DASD、伝送路、光ディスクのデータ長は、シミュレーションによる不良セクタ発生の最少値と、実効データの転送速度から最適値を求めた(表2)。プログラムは、ホスト側がシステム開発用言語を使い端末側は COBOL で作成した。プ

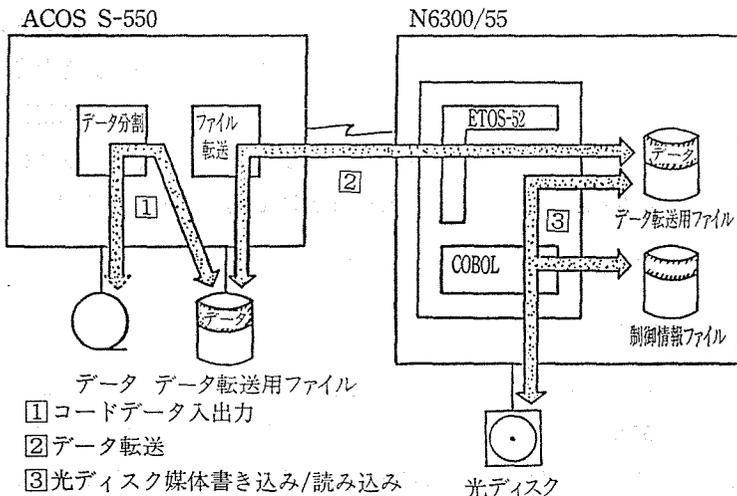


図1 実験システム概要

プログラムは、ホスト側のデータ入出力処理、端末側では光ディスクの編集・書き込みの二本である。作業工数は、COBOL 言語で 4000 行である。実験処理に使用した OS は、ホスト・コンピュータが ACOS4/MVP、端末は ETOS-52 である。

5.1. ファイル機能

実験システムで使用する作業用ファイルは五種である。入出力用が二種、データ管理用が二種、保存用ファイルが一種である。入出力用は、(1) ホスト・端末間のデータ転送用ファイル、(2) リラン・ファイルである。管理用は、(1) ディレクトリ・ファイル、(2) 記憶管理ファイルである。そのほかは、データ格納用の光ディスク・ファイルである。

処理の中心になるファイルは、制御・索引などのデータ管理情報を記録するためのものである。これらのファイルは、作業経過によって内容が変化する。ファイルは、データの変化に対応できるように、ホスト・端末ともに DASD 上に割り付けた。一部の管理情報は、変更がないものとして、光ディスクに保存対象データとともに併記した。この型は、「4.2. 索引とデータの分割配置」で述べた 2-2 の形式になる。

1) データ転送用ファイル

実験システムで使用した通信回線は、データ転送単位 255 バイトに固定されている。ホストから端末、逆方向のデータ伝送は、すべてこの長さデータ長を変更しなければならない。データ転送用ファイルは、光ディスクへの書き込み、検索結果のホスト・コンピュータへのデータ転送に共用する。ファイルは、端末側・ホスト側にそれぞれ 2 個、作業用ファイルとして設定した。

2) 記憶管理ファイル

記憶管理ファイルは、光ディスクの制御プログラム (Office Electronic Filling System: OEF) の一部として提供されている。管理ファイルは、OEF の起動時に自動的に生成される。入出力コマンドは、OEF の制御下におかれている。ファイル内容は、入力データにつけられている媒体のボリューム名、データ作成日付、コマンド実行時に使用するファイル名、そのほか、

データ追記処理用の光ディスク・データの記録位置情報である。データの記録位置情報は、光ディスクの更新処理に利用できる（ただし、更新機能は第二期のシステム拡張で対応予定）。

データ記録位置情報は、利用者から見た論理データと、光ディスクに格納する物理的データとの接続関係を管理する。光ディスクに記録されているデータは、媒体上で直接のデータ修正ができない。しかし、汎用的に利用するには、疑似的にせよデータの修正ができることが望ましいことになる。

対応方法は、(1) 修正済みデータを光ディスクに追記する、(2) DASD 上の記憶管理ファイルに記録されている旧データの接続関係を追記データに変更する二つである。この二つの処理は、記録されている光ディスクのレコード情報を索引から切りはなし、追記する新データに制御を変更する。旧データは、光ディスク内にそのまま残ることになるが、データの追加・削除・置換にも対応でき、光ディスクの書き替え問題は実務上問題ないことになる。

3) ディレクトリ・ファイル

情報管理の主体となるファイルである。DASD のデータ管理に使用するディレクトリが記録され、記憶管理ファイルと対で使用される。管理の対象となる情報は、転送されてきた論理レコードの長さ、ブロック長、レコード形式、コードの種類などである。ファイルにあるデータの先頭には、キーの位置、長さ、ファイルの先頭と最終データの内容が制御コードとして記録されている。記録するデータが、複数の光ディスク媒体を使用する場合の識別情報も保持している。

4) リラン・ファイル

長時間連続した処理をおこなうためには、作業を任意の位置で中断できなければならない。この処理をリラン処理という。リラン処理では、作業を中断するとき、その時点のコンピュータ処理内容を保持しておく必要がある。リラン・ファイルは、光ディスクへデータを読み・書きする処理の中断時点の内容を保存するために使用する。処理の再開は、DASD 上に待避させたリラン・ファイル内容をコンピュータに読み込ませ、中断した命令の次の位

置から自動的に復帰させることができる。

5) 光ディスク・ファイル

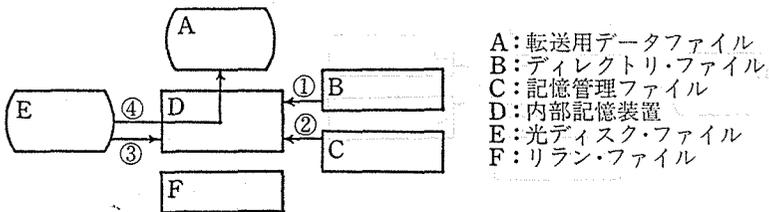
永久保存の対象になるデータを記録するためのファイルである。記録対象となるデータは、各レコードの先頭位置情報と保存データの二種である。記録順序は、順配列であるが、光ディスク上に各データの位置を索引化することによって、疑似的に索引順ファイルとして利用できる。検索処理のさい処理時間の短縮に効果的である。媒体単位のデータ交換も容易になる。そのほか、DASDの記憶管理ファイルとの併用で、データ修正も追記処理で代行できる。

なお、光ディスク1枚に記録できるコード・データ容量は、255バイト長のデータ46500件を1ページとしたとき84ページ分である。ただし、複数枚にわたるデータの記録は現在許していない。

5.2. データの入出力と管理ファイル

1) 読み込み処理

データの読み出し処理は、ディレクトリ・ファイル、記憶管理ファイルの二つの管理ファイルを使用する。処理は、端末からパラメータで指示を与える。入力するパラメータは、読み込み処理の指示、データの識別情報（光ディスク・ホストファイルのボリューム名、ファイル名、装置番号）、キー情報（キー位置、長さ、先頭と最終キーの内容）である。入力したパラメータをもとに、ディレクトリ・ファイルからはレコード長、ブロック長、レコード形式、コード識別情報を、記憶管理ファイルからは、光ディスクにあるデータの先頭位置情報を読む（1, 2）。



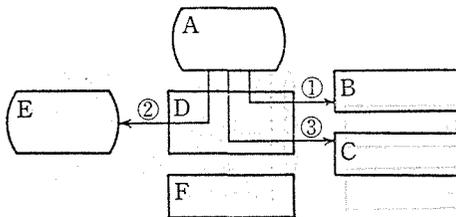
ファイルの処理は、入力した管理情報をもとに、ディレクトリ・ファイルにあるデータの先頭および最終内容を見て、光ディスク・ファイルに記録されているデータ位置を確認する。次に、該当するデータを光ディスクから読み取り、データの先頭から逐次探索をおこなう (3)。キー情報とデータが一致したレコードを転送ファイルに送る。

この後、利用者は DASD 上の転送ファイルから任意にデータを読み出すことができる (4)。また、光ディスクからデータを読み出すとき、キー情報は DASD におかれているため、検索ごとに光ディスクを探索する必要はない。

2) 書き込み処理

光ディスクへのデータの書き込み処理は、次の手続きをおこなう。パラメータは、書き込み指示、光ディスクのボリューム名、装置番号、キー情報および、追記すべきデータ位置の各情報である。指定されたパラメータ値を参照して、転送用ファイルの先頭にある書き込みデータの制御情報を読む (1)。制御情報には、データ識別情報 (MT のボリューム名、ファイル名、レコード長、ブロック長) と、キー情報が記録されている。制御情報をディレクトリ・ファイルに複写した後、転送用ファイルにある 255 バイト長のデータを論理データ長に再編成し、光ディスクへ書き込みを開始する (2)。

光ディスクへデータを書き込むさい、各論理レコードの先頭にデータの制御情報を挿入する。キー情報は、記憶管理ファイルへ登録する (3)。以下、データが終了するまでこの処理を繰り返す。最終データ処理は、最後のデータにあるキー情報をディレクトリ・ファイルに、またデータの識別情報を記憶管理ファイルに登録する。最終処理は、処理を再開するさい、新しく読まれるデータの識別と照合をおこなうための処置である。

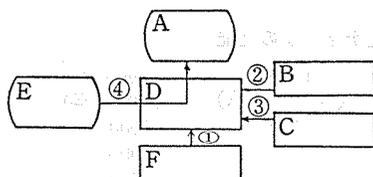


3) リラン処理

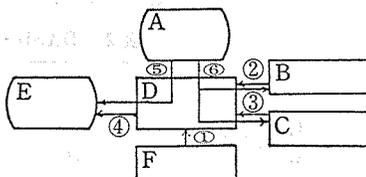
中断した処理を継続させるには、端末のキーボードから、データの追記または読み出しの指示と、対象ファイルの識別情報、キー情報、リラン処理の指示をおこなう。リラン・ファイルの起動後 (1)、情報をもとに、ディレクトリ、記憶管理情報の二つのファイルを読み込み、中断時点のデータ内容を処理装置上に再現する (2, 3)。

これらの作業を終了した後、光ディスクからデータの読み出しを再開する (4)。書き込みの継続処理では、追記するデータが必要とする領域を光ディスク上に割り当て (4)、中断された次のファイル位置から処理を開始する (5)。ディレクトリ、記憶管理ファイルの更新・作成は、書き込み処理中に、同時におこなう (6)。

(リラン処理・読み込み)



(リラン処理・書き込み)



5.3. DASD・光ディスクの実効データ転送時間の比較

DASD のデータ処理に遅れを生じさせる要因は、DASD のヘッドのトラック間の移動時間 (シーク時間)、データを探すための媒体の回転による平均待ち時間 (1回転に必要な時間の2分の1を平均アクセス時間とする)、通信回線のデータ転送能力の三つである。三条件のうち、データの入出力処理時間に直接影響を与えるのが、シーク時間、平均アクセス時間である。また、プログラムと関連するものにデータ長がある。DASD の処理速度は、上記の三種の処理の遅れを少なくすることによって、実効効率を上げることができる。入出力処理の効率化には、装置固有の物理的特性から処理速度の最適値を求め、データ長に反映させる必要がある。同様に、光ディスクについても実務に耐えるシステム構築には、DASD、データ伝送路、データの最

適単位を決める必要がある。

以下、DASD のハードウェア条件から、光ディスクを利用する上での処理速度を推定し、DASD・光ディスク間で調整できる最適データ長を求める。なお、計算値の基礎となる平均アクセス時間は、シーク時間と平均回転待ち時間の和で表される [文献7]。

$$\text{データ転送速度} = \text{データ長} / \text{平均アクセス時間}$$

例えば、表2からデータ転送速度625kバイト/秒、平均アクセス時間38.5ミリ秒 (millisecond: ms) の DASD を使用した場合、データ長と実効データ転送速度の関係は表2の項目8のように推定できる。1トラックの最長データ24062バイトは理論値である。計算値は、DASD の物理的なデータ長、24062バイトのときに、装置の最大データ転送速度を確保できることを示している。物理レコード長は、データの最大長を決める目やすのほか、入出力用バッファ領域の決定にも利用できる。

表2 DASD・光ディスクの性能

	DASD (N6355-54A)	光ディスク (N6329-23)
1. 記憶容量 (Mb)	120	1000
2. 回転数 (rpm)	3510	900
3. 平均アクセス時間(ms)	38.5	303
4. 平均シーク時間 (ms)	30	270
5. 平均回転待ち時間(ms)	8.5	33
6. 隣接トラック間シーク時間(ms)	8	3
7. データ転送速度(kb/s)	625	785
8. 実効データ転送速度		
実効速度(kb/s): データ長(b)	13.3:512	1.7:512
	51.9:2000	6.6:2000
	103.9:4000	13.2:4000
	207.8:8000	26.4:8000
	415.6:16000	52.8:16000
	625.0:24062	105.6:32000

rpm: round per minute

そのほかの計算値は、表2の8で示した値になる。DASD では、データ長512バイトのとき実効データ転送速度は13.3kバイト/秒である。4kバイト単位の入出力処理では103.9kバイト/秒になる。この数値は、データ長を2倍にすると一定時間内のデータ転送数は13300バイトから103900バイトに

なることを示している。ここで、データの転送時間は、媒体への入出力速度と同じとした。したがって、単位時間内でのデータ転送の絶対数の増加と同意味になる。

同様に、光ディスクの場合も実効転送速度を算出した。512バイトで1.7kバイト/秒であるものが、4kバイト/単位では13.2kバイト/秒になり、16kバイトでは52.8kバイト/秒まで増加する。以上の数字からのもの、データ入出力単位を長くとることによって効率が上がることは明らかである。

6. 光ディスクの処理時間の測定

6.1. 処理速度と最適データ長

データ入出力速度は、対象となるデータの長さによって変化する。処理の効率化をはかるためには、データ処理単位を長くとり、入出力アクセス回数を少なくすることが必要である。しかし、実情は与えられる条件から、適当な値で妥協せざるをえない。

これに対して、光ディスクの実用化には、基礎的段階で未知の部分があり、DASDの技術をそのまま利用できないことがある。そのため、システム設計に当たっては、個々の問題を実験によって解決しなければならない。第一は、光ディスクへのデータ記録単位の問題、第二は、書き込み誤りへの対応である。

光ディスクにデータを書き込むさい、個々のデータが正しく処理されたかどうか確認が行われる。確認は、記録したデータを再度読み出し、双方のデータを比較・検証する Read After Write 方式である。比較の結果が一致しない場合、次の処理が必要になる。

- 1) 誤りが発生したセクタにフラグを立てフラグの検証をおこなう。
- 2) 不良セクタ部分のデータを代替セクタに書き込み代替セクタの検証をおこなう。
- 3) 誤りが発生した、残りのセクタにあるデータの検証を再開する。

データ検証では、不良セクタにフラグを立てる、代替セクタにデータを書

き込む、それぞれの処理を個別に検証するなど4回の処理が追加される。この操作は、書き込み処理中に誤りが発生した場合、4回の回転待ち時間が必要になることを意味する。さらに、トラックに残されたデータ検証が加わるため、全体では5回である。誤りが無い場合は、2回の入出力処理で対応できるのに対し、N個の誤りで5N回に増加する。このときの処理時間は、入出力と同様5Nである。

処理の効率化をはかるためには、誤りの発生を低くおさえることが重要である。実験システムで使用した光ディスクは、書き込み単位を4k, 8k, 12k, 16k バイトの4種が選択できる。書き込みデータ長は、誤り発生率が最低であり、かつ現行システムで許される最大長をとる二つの条件を満たす16k バイトを使用した。

実測の結果から、処理速度は約48.6%改善された。読み込み処理では、79.9%である。表3に示した処理時間は、それぞれのデータ長における計測値である。測定に使用したデータ数は、4k バイトが10.21Mb バイト、16k バイトでは11.86M バイトである。

なお、誤り発生時に使用するセクタは、各トラックの64から68セクタの4個を予備として用意している。トラックにある予備セクタをこえる修正は、次のトラック以降にあるセクタを順次使用する。

16k バイト長では、一部システムを改良して測定した。高速化が可能になった主な理由は、データ長を4倍にすることによって、光ディスクの回転待

表3 データと処理速度の関係

No.	処理内容	書き込み時間(分)	読み取り時間(分)
1.	MT-DASD (ホスト処理)	1) 5 (1.0%)	1) 5 (0.4%)
		3) 4 (1.2%)	3) 4 (1.4%)
		1) 160 (32.7%)	1) 100 (10.0%)
2.	データ転送	3) — (48.3%)	3) — (35.1%)
		1) 325 (66.3%)	1) 900 (89.6%)
		2) 194	2) 210
3.	DASD-MT (端末処理)	3) 167 (50.5%)	3) 180.8(63.5%)

1) 4kb (10.21Mb)

2) 16kb (11.86Mb)

3) 16kb (10.21Mb 換算値)

ち時間を4分の1近くに減少させたことにある。4k バイト長の記録単位では、1トラック分のデータ記録に8回の待ちが必要になる。照合処理は、同数回おこなわれるため全体では16回である。これに対し、16k バイト長では4回でよい。

しかし、現行の16k バイト長は、光ディスクのトラック長の2分の1の長さにすぎない。データ長は、可能な限り光ディスクのトラック長と同じに調整すべきである。32k バイト長では、回転待ちは2回でよく、16k バイト長の4回に対し処理時間の短縮が期待できる。表3で示した端末処理時間をさらに2分の1程度短縮することができる。

6.2. 通信回線の伝送効率の測定

光ディスクへのデータの書き込み処理と、検索したデータをホスト・コンピュータへ転送する処理では、常に回線を使用しなければならない。通信回線は、データ処理性能に十分に対応できる速度をもつことが必要になる。現行システムで使用できる回線は、少量のコードを伝送する目的で設定されており、光ディスクの大容量に対応できる伝送容量をもっていない。回線速度の効率化は、長時間の伝送処理を必要とするシステムの実用化にとって大きな問題となる。

回線能力は、各装置の処理速度を上げること、装置の実効データ速度に対し、アイドル時間が少ないこと、データ伝送単位の最適化をはかることで効率化できる。そのためには、回線能力の実態を知ることが必要である。

以上の点を踏まえ、装置の回線効率を測定した。測定は、回線を通して送られるデータの総伝送時間と、ホスト・端末で使用した DASD 処理時間の二つである。回線処理時間は、データ伝送に要した全処理時間から、ホスト・端末間の DASD 処理時間を引いた値で求められる。測定条件および計測値を以下に示す。

	1) データ転送単位	255 バイト
	2) 実験データ件数	40050 件
	3) 回線速度	1 Mbps
伝送効率は、単位時間に実際に転送できるデータ量を、理論値 (1 Mbps) で割った値の百分率で表される。データ総数	4) 全処理時間	100 分
	5) 回線効率	1.36 %

は、データ長とデータ件数の積である。なお、回線のデータ転送単位は、255バイト固定である。これは、ユーティリティ・プログラムの利用条件によるものである。

測定は、コンピュータの実際の利用状態の効率を把握するため、ホスト・端末双方の DASD を稼働させおこなった。測定結果は表3に示した。伝送時間の全システムに占める割合は48.3%である。測定値は、システムの改善後の回線の待ち時間、プログラム処理を含んだ所要時間である。正確には、回線そのものの伝送効率を示していないが、回線効率の1.36%は、この部分の改善が必要であることを示している。

7. おわりに

光ディスクの導入は、大量データを長時間保存する場合と、DASD のバックアップに利用する二つの目的があった。実験システムによって、大量データのバックアップ機能は、ほぼ確認できた。しかし、長期保存用媒体としての有効性は、作成したデータの保存内容と、確認するための再利用をまたなければならぬ（メーカは加速試験の結果から10年を保証している）。

小規模の実験の結果、光ディスクのデータ入出力速度は、DASD の5-10倍かかることが確認された。この数値は、大量データを光ディスクに蓄積する場合、対象業務へ与える影響は大きいといわねばならない。開発したシステムは、光ディスクの利点とともに問題も明らかにした。今後、解決すべき問題は、(1) 媒体と記録方式の標準化、(2) 光ディスクの処理の低速性への対応、(3) データ修正機能の導入とオペレーティング・システム (Operating System : OS) を含む、基本ソフトウェアの整備、(4) 光ディスクと DASD の最適ブロック化係数の設定、(5) 通信回線と装置間における最適データ長の選択機能などである。

1) 媒体および記録方式の標準化は、光ディスクの開発状態と利用法に依存する部分が多い。現在、光ディスクは、ROM 型、WORM 型が実用化され消去型は研究段階にある。速度制御方式も線速度一定型、角速度一定型の

二種が実用化されている。しかし、それぞれの特性に応じた利用方法はまだ定まっていない。

一方、システムは、複雑になるに従って多層構造をとる。多層化は、モジュール化をすすめる、モジュールはモジュールの独立性と、接続インタフェースの標準化を要求する。インタフェースは、データとデータおよびモジュールを制御する二つの機能をもつ必要がある。DB との併用を考えた場合この特性は重要である。DB との標準インタフェースの設定は、光ディスクをどのように利用するかにかかっている。

2) 低速性は、光ディスクの利用目的と応用の範囲を制限する。技術的な対応には、以下の方法が考えられる。(1) キャッシュ・メモリを使用したキー処理の先き読み予測の採用、(2) 端末装置の記憶容量の拡大と媒体の回転速度の高速化、(3) データ圧縮法、(4) プログラム言語に対する光ディスク処理機能の拡張などである。

さらに、総合的な解決をはかるためには、ホスト・コンピュータ、端末の双方の OS に光ディスクの処理機構を導入すべきである。特に、実験システムで作成の対象になった管理ファイルは、キャッシュ・メモリの制御用に拡張できる。これらの機能は、基本処理の一つでありシステムの中で実用化できる最も近い位置にある。

キャッシュ・メモリは、データの読み取り時間の推定に、平均シーク時間が適用される場合に有効である。これに対して、書き込み処理は複数トラックに連続して記録するため、隣接トラック間シーク時間が適用される。これは、光ディスクで期待できる最高値であり、読み取り速度の目標値になる。

3) 問題は二つある。一つは、入力したデータに対する利用者の修正機能であり、ほかの一つは入出力時の書き込み誤りの検出処理である。前者は、データ管理機能のシステム化で対応できる見通しがついた。問題は後者である。光ディスクへの書き込み照合は、DASD 同様 Read After Write 方式である。この処理は、一処理に 2 回の回転待ち時間を必要とする。書き込み処理中に読み取りができれば処理時間は半分になると予想される。光学ヘッド

の軽量化とともに改善すべき点である。また、誤り検出機能は、ソフトで検査しているが、これも標準化の一環としてハード化しやすい部分である。

4) 最適ブロック化係数の決定は、低速装置と高速装置の間に生じる、データ転送時のアイドル時間の問題を解決する手段に使用できる。表2で求めた実効データ転送速度をもとにした試算では、現行システムで約3対1である。DASDのデータ長を1とした場合、光ディスクに書き込む長さを3に設定することによって、速度上の整合をとることができる。

5) 回線速度は、現行の回線規定がデータ長を255バイトに固定しているため伝送回数が増加している。また、データ伝送途中で発生する誤りへの対応が処理効率を低下させている。早急に利用目的に応じたデータ長を確定すべきである。

光ディスクの入出力単位を16kバイトにすることによって、処理速度を改善できた。端末処理時間の全処理時間の割合は、66.3%から50.5%に減少した。しかし、逆にデータ伝送時間は、32.7%から48.3%に相対的に増加している。全処理時間の半分である。読み込み処理では、10.0%から35.1%である。この数字は、次に改善すべきプログラム位置を示している。

光ディスクの有効利用の形は、多様化した情報の統一的な処理であり、研究用DBも当然この要求が強い。しかし、光ディスクを利用した応用システムは、研究開発段階にあり、実用化の例も経験的な技術の積み重ねも少ない。その点で、本システムで得られた各種のデータは、これから光ディスクの利用と応用をすすめる上で貴重な情報となる。

最後に、この報告をまとめるにあたって、日本電気株式会社情報処理官庁システム事業部第三システム部、課長平野哲、千葉友夫、大生学、中島進の各氏の協力を得た。記して謝意を表する。

千葉友夫氏は、この8月4日に交通事故でお亡くなりになりました。御冥福をお祈り致します。(1986.9.30)

参考文献

- 1) 石島義弘・他 (1982) 「イメージ情報を結合した日本語情報検索システム」(情報処理学会第25回全国大会発表論文集) 699-700。
- 2) 『N6300 モデル55光ディスクシステム説明書』日本電気株式会社。
- 3) 「大量文書情報の保管・検索用に使われ始めたユーザ書き込み型光ディスク」『日経エレクトロニクス』(1983. 3. 28) 105-120。
- 4) 「広範な用途に進出する光ディスク」『日経エレクトロニクス』(1984. 3. 12) 159-174。
- 5) 市山義和 (1982) 「大容量ファイルとしての光ディスク」『情報処理』Vol. 23, No. 7, 634-640。
- 6) 笹森勝之助 (1985) 「初心者のための入門講座⑥資料の保管形態としての情報媒体」『ドクメンテーション研究』Vol. 35, No. 7, 405-414。
- 7) ロイ・チャネイ・他 「磁気ディスクの転送速度を5-15倍に上げるキャッシュの手法を探る」『日経エレクトロニクス』(1984. 8. 27) 249-273。
- 8) 斎藤秀紀 (1980) 「分散処理システムへの試み」『電子計算機による国語研究X』(国立国語研究所報告67) 73-88。
- 9) 『N6300 モデル 55LANFILE3 説明書』日本電気株式会社。
- 10) B. W. Leong-Hong・他 (穂鷹良介監訳, 成田光彰訳) (1986) 『データディクショナリ/ディレクトリシステム』オーム社。
- 11) 『光ディスク・システム設計書』日本電気株式会社。