

## 階層型ディスクキャッシュ・サブシステムの 方式と性能評価

宮地泰造 三石彰純 溝口徹夫

三菱電機 (株) 情報電子研究所

### 1. はじめに

CPUの処理能力の飛躍的な向上により入出力処理がシステムのボトル・ネックになりつつある。とくにディスク装置へのアクセスの高速化は重要な問題となっている。ディスク・アクセスの高速化には、ディスク装置やチャンネルの増設、ブロッキング・ファクタの変更、ディスク・キャッシュ<sup>3)4)</sup>の導入、主記憶増設、CCD・磁気バブルなどの高速デバイスとの置換などの手段が考えられる。また、その他の手段としてバッファ内蔵型ディスク装置<sup>5)</sup>(以下B-DISKと略記する)があり、B-DISKは次の3点を特長としている。

- 1) RPS<sup>1)</sup>ミス時の待ち時間を短縮する(A)。
- 2) 書き込み時、シーク動作とデータ転送とをオーバーラップさせていることによって見掛け上シーク時間と回転待ち時間がなくなる(B)。
- 3) 読み出し時、先読みしたデータがバッファに存在すれば回転待ち時間がなくなる(C)。

我々は、CPUの処理能力の著しい向上に対応するための高速入出力処理のためにB-DISKを提案した。今回の報告ではB-DISKの構成と動作およびシミュレーションによる性能評価の詳細を示す。さらに、より高速な入出力処理の実現のために、B-DISKとディスク・キャッシュ(以下DCと略記する)とを階層的に結合することによりその相乗効果を実現する階層型ディスクキャッシュ・サブシステム(以下MESSIAH: Memory Subsystem of Hierarchical Disk-Cacheと略記する)を提案する。また、実サイトで稼働中のシステムに対してシミュレーションによるMESSIAHの性能評価を行ったので報告する。

### 2. 階層型ディスクキャッシュ・サブシステム

#### 2.1 入出力処理時間の短縮

入出力処理時間の短縮の直接的な手段としてディスク・キャッシュ(DC)がある。ここで、DCが入出力処理時間を短縮する基本的な理由とその制御方式を確認する。

まず、DCが入出力処理時間を短縮する基本的な理由は次の2つである。

- 1) ディスク装置のアームの移動回数を削減することによりシーク時間を削減する(D)。
- 2) ディスクの回転待ちの回数を削減することにより回転待ち時間を削減する(E)。

しかし、一方では

- 3) ディスク、DC間のデータ転送量の増大(すなわち、データ転送時間の増大)が起きる(F)。

よって、DCの効果を十分に引き出すためには、a式をより十分に満足させることがポイントになる。

$$(Dの効果 + Eの効果) \gg Fの効果 \dots (a)$$

そのためのDCの制御方式としては、ライト・スルー方式、ライト・アフター方式をはじめバイパス・モード、順次アクセスのモード<sup>2)</sup>、DCのみを使用するモード<sup>2)</sup>等がある。

つぎに、B-DISKの入出力処理時間の短縮の原理について示す。B-DISKが入出力処理時間を短縮できる基本的な理由は、第1章のA~Cで示した3点である。この中でCはDCの原理と重複していると考えられる。また、B-DISKでは、その特長が発揮できない状況であっても、一般の磁気ディスク装置に比較して能力が低下しないことは注意しておく必要がある。

#### 2.2 バッファ内蔵型ディスク装置の構成と動作

B-DISKは階層型ディスクキャッシュ・サブシステムの重要な要素である。以下に、その構成と動作について述べる。

ディスク装置のRPS機能は、ディスク・コントローラ(DKC)の利用効率を高め、回転待ち時間を短縮する効果があるが、反面DKCのビジー率が高くなるとRPSミスが頻発し余分な回転待ち時間が増加する。稼働中の計算機システムにおけるRPSミス率は予想外に大きく、それによって生じる待ち時間は無視できない。この待ち時間を最小限に抑える手段としてディスク装置内にバッファを内蔵させることを検討した。これは従来ディスク(円板)の回転と同期して行っていたDKC(ディスク・コントローラ)とディスク装置間のデータ転送をバッファを用いて非同期に行うもので、待ち時間の短縮を目的としたものである。また更にバッファを持つて

とによって可能となる読み出しデータの先読みやデータ転送とシーク動作のオーバーラップによる性能向上も図っている。その具体的な構造は図2.1に示す様に従来のディスク装置に読み出し用バッファ (RDB), 書き込み用バッファ (WDB), バッファ・コントローラ (BC) を付加したものである。

RDBは読み出しデータのバッファリングを行い、シーク命令で指定されたトラック (目的トラックと称する) とその前後 (-1, +1トラックと称する) 計3トラック分のデータを格納する。ディスクからのデータ読み出しは3トラック分同時に行い、各々直並列変換回路で並列データに変換してからRAMに書き込む。並列データのデータ幅は任意であるが、以下では8ビットの並列データに変換するものとして説明する。ディスクからRDBへのデータ転送は、シーク動作が終了して最初に遭遇したセクタから順次行い、次のシーク動作が始まるまで続行する。RDBからDKCへのデータ転送はRDBに目的のデータが到着した時に開始するが、読み出し要求が発生した時点で既に目的のデータがRDB内に格納されていた場合 (同一トラックに連続してアクセスした場合に発生する) は要求の発生と同時に開始する。RAMから読み出したデータは直並列変換回路で再び直列データに復元してDKCへ転送される。

RAMは3トラック分のデータを格納するために必要な容量を持ち、A, B, Cの3領域に分割して使用する。RAM内のデータの有効性はデータ有効性フラッグ (DAF)、A, B, C各領域と目的、+1, -1各トラックとの対応はDAF内のDAF管理回路が管理する。RAMはディスクが1バイト分回転する時間を1サイクルとして動作し、1サイクルはW1, W2, W3, Rの4ステージから構成される。W1, W2, W3各ステージではそれぞれ目的、+1, -1トラックのデータをRAMに書き込み、RステージではRAMからデータの読み出しを行う (図2.2参照)。

WDBはディスクへの書き込みデータをバッファリングするFIFO型のメモリである。DKCからWDBへのデータ転送はシーク動作開始直後に行い、シーク動作とデータ転送をオーバーラップさせることで応答特性の向上を図っている。WDBからディスクへのデータ転送はアクセスヘッドが目的のセクタに達した時、BCが転送開始信号を出すことによって開始され、FIFOメモリが空になるまで続行される。FIFOメモリが空になった時、DAF管理回路はデータ転送を終了させると共にその旨をBCに通知する。ディスクへの書き込みはRDBのW1ス

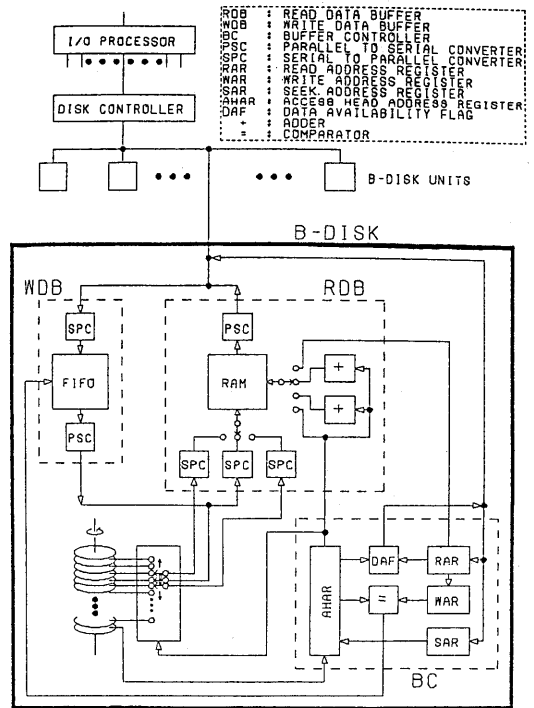


図2.1 B-DISKの構成

テージに同期して行い、ディスクとRDB双方に同時に書き込むため両者間にデータの不一致は生じない。

BCは4種のレジスタ (RAR, WAR, SAR, AHAR), 比較器, DAF等から成り、RDB, WDBの管理やDKCとの制御信号の交換を行う。

比較器はAHARとWARのトラック・セクタ・アドレスを比較して両者が一致した時、WDBに対してWDB→ディスクのデータ転送開始を指示する信号を供給する。

DAFはRDB内のデータが有効であるか否かを示すフラッグで3トラック分のラッチ群とDAF管理回路から成る。ラッチは各セクタの先頭で発生するセクタ・パルスによって対応するビット (AHARでアドレッシングされる) がonにされ、シーク命令又はRARの桁上りによって全部又は一部がoffにされる。RARによってアドレッシングされる目的セクタのビットがonになった時RDB→DKCのデータ転送が可能になる。一方DAF管理回路はA, B, C各領域と目的、+1, -1各トラックとの対応付けを管理し、必要に応じてラッチ群をクリアする。

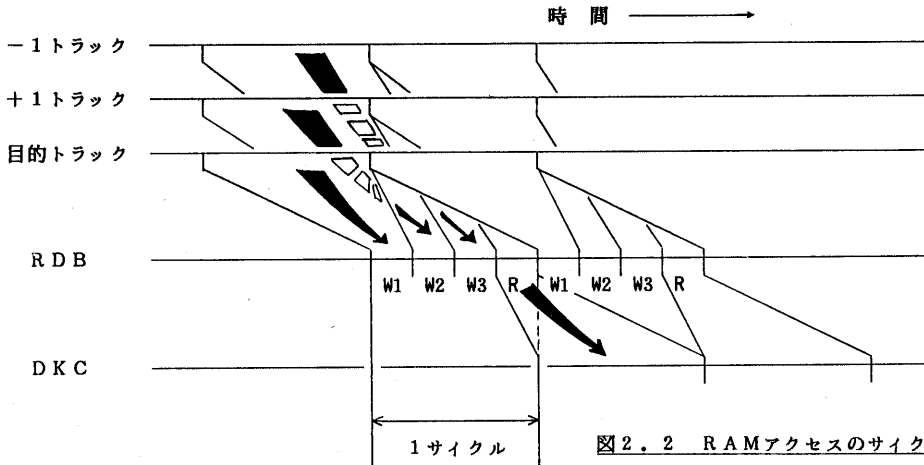


図 2. 2 RAMアクセスのサイクル

### 2. 3 階層型ディスクキャッシュ・サブシステムの構成と動作

以上に、2. 1節、2. 2節で、DCとB-DISKとが入出力処理時間を短縮できる原理を示し、その相違を明確にした。また、DCの制御方式、B-DISKの構造と動作の詳細についても述べた。

本項では、B-DISKとDCとを階層的に持つことにより、B-DISK, DC 両方の特長を有することができる階層型ディスクキャッシュ・サブシステム (MESSIAH) を提案する。

図 2. 3は、MESSIAH の構成を表している。MESSIAH は通常の磁気ディスク装置に代えてB-DISKを持ち、DCを入出力処理装置に持っている。ここで、DCを入出力処理装置に置いた理由は以下の2つである。

- 1) CPUに負荷を追加しない。
- 2) より多くのB-DISKにDCを共有させて、実質のDCのメモリ・サイズの増大を図る。

そして、MESSIAH は図 2. 3の構成により、従来の計算機システムの磁気ディスク装置に小容量のバッファ・メモリを、入出力処理装置に大容量のバッファ・メモリを有する形をとることにより、高速な入出力処理処理することができる。

ところで、MESSIAH が高速な入出力処理を行うことができる理由は、B-DISKとDCとの特長をともに実現できること以外に、Fに示したDCの欠点を改善するという特長をも有していることが上げられる。これはB-DISKが、DCのヒット率を向上させるためにB-DISK, DC間のデータ転送のブロック・サイズを大きくすることによりおきるRPSミスの頻発に対して、性能が著しく低下しないことを活用しているからである。すなわち、B-DISKとDCとの

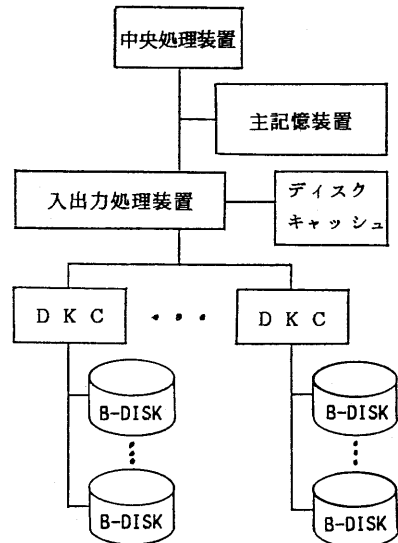


図 2. 3 MESSIAH の構成

相乗効果を実現したのがMESSIAHである。ここでMESSIAHの特長をまとめると次の6つになる。

- 1) ディスク装置 (B-DISK) のアームの移動回数を削減することによりシーク時間を削減する。
- 2) ディスクの回転待ちの回数を削減することにより回転待ち時間を削減する。
- 3) RPSミス時の待ち時間を短縮する。
- 4) 書き込み時、シーク動作とデータ転送とをオーバーラップさせていることによってディスク装置 (B-DISK) は見掛け上シーク時間と回転待ち時間がなくなる。

- 5) 読み出し時、当該データがB-DISKのバッファに存在すれば回転待ち時間がなくなる。
- 6) B-DISK~DC間のデータ転送のブロック・サイズの増大に伴うR P Sミスの頻発に対応できる。

### 3. 性能評価の方法

性能評価は、現在稼働中の計算機システム（仮にDB16と呼ぶ）を対象に行った。DB16は最適なファイル・アロケーションが施してある300MBディスク16スピンドル（システム・ディスク：2スピンドル、プライベート・ディスク：14スピンドル）のシステムである。また、DB16はデータベース処理およびファイル処理のバッチ・ジョブが8多重で実行されているシステムである。ここで、データベースはCODASYL型である。

#### 3.1 性能評価の流れ

性能評価は、以下に示す手順で行った（図3.1）。

- 1) 実システムの統計データ（I/O発生頻度など）から対象とすべき時間帯を選択する。そして、ディスク・アクセスのトレース・データのアクセス・パターンを分析し、DCの性能評価のためのトレ

ース・データを抽出・編集する。

- 2) DCの性能評価を行う。
- 3) DCの性能評価結果からB-DISKおよびMESSIAHの性能評価のためのトレース・データを決定する。
- 4) B-DISKおよびMESSIAHの性能評価用トレース・データを編集し、I/O要求発生率を算出する。
- 5) B-DISKおよびMESSIAHの性能評価を行う。

#### 3.2 ディスク・キャッシュのシミュレータ

DCのシミュレータは、ディスク・アクセスのトレース・データからキャッシュ・ヒット率、アクセス時間の改善度、ヒット深度のパターンを求めてDCの性能を評価する。ここで、アクセス時間はディスク装置のシーク長を考慮して算出している。また、DCは入出力処理装置に置かれている。つぎに、シミュレータの特徴として、DC~ディスク装置間のデータ転送のブロック・サイズを4種類（1ページ、1トラック、2トラック、1シリンダ）選択できることがあげられる。ここで、DC内でのデータ管理単位は、転送ブロック・サイズが2トラックの場合に1トラックである以外は転送ブロック・サイズと同一である。これによりディスク上のアドレスによる物理的なアドレスのプリフェッチの効果を調べることができる。

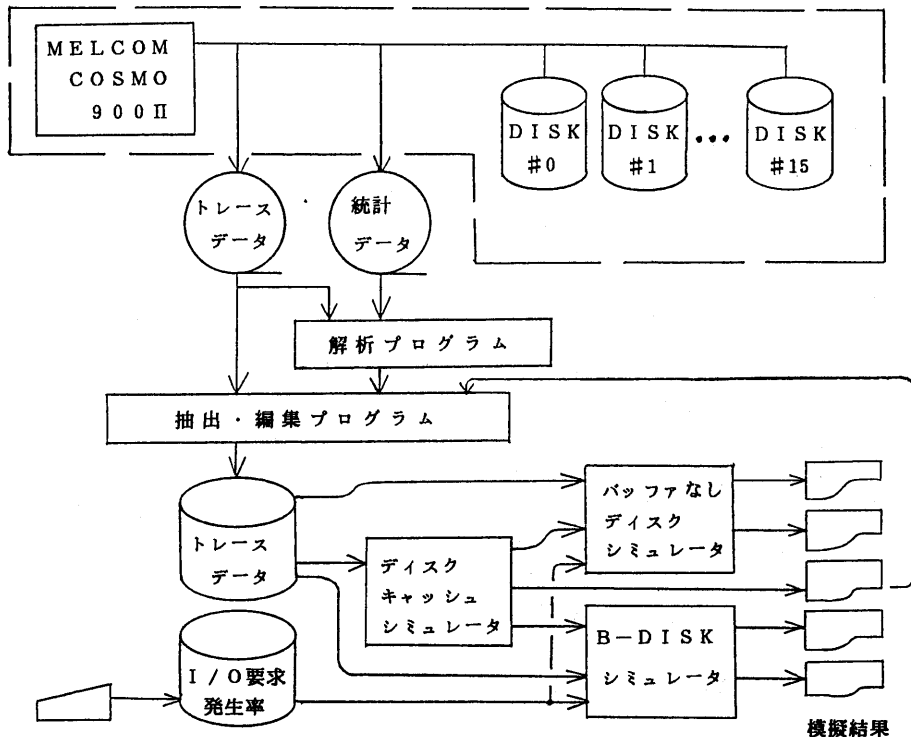


図3.1 性能評価の流れ

また、DCの制御方式はファイルの更新時に信頼性の高いライト・スルー方式を用いている。ここで用いたライト・スルー方式は、Read時: i) 目標のデータがDC上にある(ヒット)の場合DCから主記憶にデータ転送。ii) 目標のデータがDC上がない場合ディスクからDC, 主記憶の両方にデータ転送。Write時: i) ヒットの場合DC上の当該レコード, ディスクを更新。ii) ヒットしない場合ディスクだけ更新。である。そして、キャッシュ上での置換はLRU方式である。

ヒット率は、対象としたシステムのI/O要求が2KB以下であったことから「I/O要求の対象であるページがDCに存在する場合にヒットしたという」と定義した。

入出力サービス時間(Tc)は、ヒット時:  $Tc = \text{平均DCアクセス時間}$ , ヒットしない時:  $Tc = (\text{シーク時間}) + (\text{平均回転待ち時間}) + (\text{データ転送時間})$ とした。

### 3.3 バッファ内蔵型ディスク装置のシミュレータ

B-DISKの性能を評価するため、バッファ有り・なし2種のシミュレータを作成し、種々の負荷における性能を調べた。バッファなしディスクのシミュレータは当社300MBディスクパック M2838-Fをモデルとし、B-DISKシミュレータは M2838-F にバッファを付加した状態を想定している。モデルとした M2838-Fの仕様を表3.1に示す。

この2つのシミュレータはCPU, DKC各1台とディスク8台を対象としてシミュレーションを行う。IOP(I/Oプロセッサ)は性能にあまり関与しないことからモデルに組み込んでいない。

#### ●モデル化における仮定

モデル化を行うに当たって以下の仮定を設けた。

- (全体) ソフトウェア/ハードウェアのエラーは発生しない。
- (CPU) I/O要求発生時間間隔は指数分布とする。
- (CPU) I/O要求の内容はシーク, リード, ライトの3種のみとする。
- (DKC) データ転送を行っている期間のみビジー状態とする。
- (DISK) トラック内の17セクタは物理的に均等に配置されているものとし、1セクタのデータを転送するのに必要な時間は  $1 / (3600 \times 17)$  分とする。
- (DISK) WDBはオーバーフローしない。

表3.1 M2838-Fの仕様

記憶容量	300MB	
データ容量	267.2MB	
記録形式	byte / sector	1,024
	sector / track	17
	track / cyl	19
	cyl / unit	808+7代替
平均回転待ち時間	8.3ms	
位置決め時間	min.	10ms
	ave.	30ms
	max.	55ms

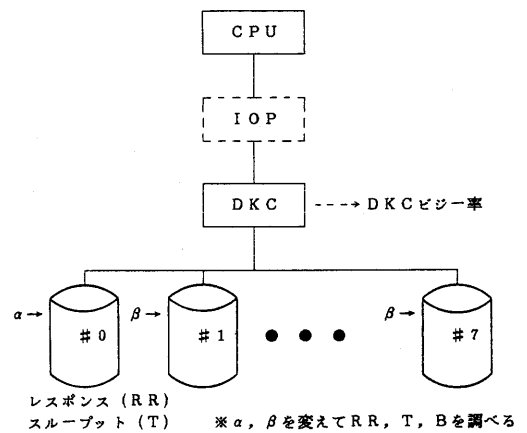


図3.2 シミュレーションの条件設定

#### ●シミュレーションの方式

本シミュレータはI/O要求発生時刻を乱数によって決定する自己駆動型シミュレータであると同時に個々のI/O要求の内容は稼働中の実システムにおいて収集したトレースデータを用いておりトレース駆動型でもある。従って外部から乱数発生に必要なパラメータとI/O要求のトレースデータを与える必要がある(図3.1)。シミュレータはモデル上の時間で1分間図3.2に示すシステムの動作を模擬し、その間に収集した統計量を結果として出力する。シミュレート直後の過渡状態の影響を除去するために統計量を収集する前に15秒間の空運転を行っている。

●シミュレーションの条件設定

今回作成した2種のシミュレータに同一のトレースデータを与えて種々のI/O要求発生率のもとでシミュレーションを行った結果をもとに性能評価を試みた。シミュレーションの条件設定は図3.2に示すように特定のディスク(#0)のI/O要求発生率 $\alpha$ と他のディスク(#1~#7)のI/O要求発生率 $\beta$ を変えながら、#0の応答時間が同ディスクのスループット(T)とDKCビジー率(B)にどの様に依存するかを調べた。8台のディスクには全て同一のトレース・データを与えた。

●評価の指標

評価には以下に定義する指標を用いている。

○ディスクのレスポンス  $RR = BDST / IOST$

○ディスクのスループット  $T = IORATE \times BDST$

○DKCのスループット  $B = DKC$ のビジー率

ここで、

$BDST$  = シーク時間 + 平均回転待ち時間 + データ転送時間

$IOST$  = I/O要求が発生してからその処理が完了するまでの時間

$IORATE$  = 単位時間当たりに実行完了したI/O要求数

である。

3.4 階層型ディスクキャッシュ・サブシステムのシミュレーション

MESSIAHのシミュレーションは、DCのシミュレータ、B-DISKのシミュレータ、そしてDCの動作を模擬して入出力処理からB-DISKに出すI/O要求を編集するプログラムを用いて行う。

4. 性能評価の結果

4.1 ディスク・キャッシュの性能評価結果

DCの性能評価は、

- 1) データベース・アクセス (A, B, L, 2, 4)
- 2) データベースとファイルのアクセス (C, T)
- 3) ファイルのアクセス (D)

の合計8種類のアクセス・パターンに対して行った。ここで、Bは全件検索に近いアクセス・パターンであり、Dは順次アクセス・パターンである。また、データベース(以下DBと略記する)は3種類ある。

図4.1は転送ブロック・サイズを2トラックとした

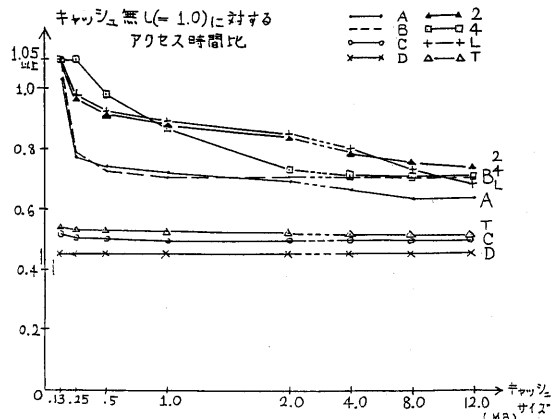


図4.1 入出力サービス時間の改善度 (転送ブロック・サイズ=2トラック)

表4.1 入出力処理時間の改善結果

アクセス・パターン	転送ブロック・サイズ			
	1ページ	1トラック	2トラック	1シリンダ
D	×	○	●	●
C	×	○	●	○
T	×	○	○	△
B	×	△	△	△
A	△	△	△	×
2	△	△	△	×
L	○	△	△	×
4	○	△	△	×

(注) ●: 50%以下, ○: 50~60%, △: 60~80%  
×: 80%以上を示す。

DC利用時の入出力サービス時間( $T_c$ )であり、キャッシュ・サイズは1スピンドル当たりの値を示している。また、 $T_c$ の有効度がある境界値に従って表すと、表4.1のようになる。

そして、DCの性能評価結果からアクセス・パターン: A, B, C, Dについて以下のことが言える。アクセス時間は、それぞれ、およそA: 62%, B: 70%, C: 49%, D: 42%に短縮される。これらの値は、それぞれ、転送ブロック・サイズがA: 1ページ, B: 2トラック, C: 2トラック, D: 1シリンダの場合である。そして、アクセス種類別にみると、A(DB)では転送ブロック・サイズが1シリンダ以外の場合に約30%~40%高速化でき、B(DB), C(DBとファイル), D(ファイル)では、転送ブロック・サイズが1ページ以外の場合

に、それぞれ約20%~30%、約40%~50%、約40%~60%高速化できる。この結果より、DCのデータベースのアクセスに対してよりもファイルのアクセスに対する場合の方が約10%~30%高いと言える。また、データベースのアクセスだけについてみると、Aは転送ブロック・サイズが1シリンダと大きい場合に、そして、Bは転送ブロック・サイズが1ページと小さい場合にそれぞれDCは有効ではない。この理由は、Aのアクセス・パターンがランダムであり、Bでは部分的にシーケンシャルであるからである。結局、データベース・アクセスにDCを使う場合にはアクセス・パターンを考慮することが必要であるといえる。

つぎに、転送ブロック・サイズが1トラックの場合と2トラックの場合とを比較すると約1%~12%アクセス時間が短縮されることがわかる。このことから、ディスク上のアドレスによるデータのプリフェッチがある程度有効であることがわかる。そして、このプリフェッチの効果が最もよく現れているのがDである。その理由は、Dがほぼシーケンシャルのアクセス・パターンをとるためである。

ところで、以上に示した性能評価は、磁気ディスク装置によるディスク制御装置(DKC)の競合から起きるRPSミスを考慮していない。そこで次にRPSミスを考慮したシミュレーションを行った。

表4.2は、RPSミス、DKCビジー率、IOに要するデバイス・サービス時間(DST)をアクセス・パターン:A、Dについてシミュレーションにより求めた結果である。この結果より、Aの処理では転送ブロック・サイズが2トラックの場合にDKCがシステムのボトル・ネックになることがわかる。その理由は、一般にDKCがシステムのボトル・ネックになる境界値はRPSミス、DKCビジー率ともに30%~40%であり、この境界値を越えているからである。また、転送ブロック・サイズが2トラック以上の場合も同様になることが予想できる。

そして、DB処理に対してDCを利用する場合は、この結果を参考にしてDKCがボトル・ネックにならないようにする必要がある。

#### 4.2 バッファ内蔵型ディスクの性能評価結果

シミュレーションはトレースデータとして順アクセス主体のものとデータベース・アクセスの2種について行った。ここでは順アクセスの結果を中心に述べ、データベース・アクセスについては【 】内に結論のみ記すことにする。図4.2は順アクセスのシミュレーション結果からTをパラメータとしてBとRRの関係をプロットしたものである。図からバッファによってレスポンスが

表4.2 RPSミスを考慮したDST

処理	DCの有・無	DCのサイズ (MB/シリンダ)	BS	RPSミス率 (%)	DKCビジー率 (%)	DST (MS)
A	無	———	最大1P	22.1~24.1	25.4	26.8~27.3
D	無	———	最大1P	28.7~30.9	25.8	29.0~29.7
A	有	0.25	1P	10.3~13.9	15.0	15.6~16.1
				11.0~12.3	13.8	14.8~14.8
		12.00	1T	38.4~39.7	47.7	22.8~23.2
				29.7~31.2	35.0	18.0~18.3
		0.25	2T	62.6~66.1	71.4	35.0~37.6
				47.2~51.8	52.3	22.7~24.1
D	有	1.00	1P	28.7~30.9	25.8	29.0~29.7
			1T	24.2~26.3	26.7	15.1~15.5
		1.00	2T	25.3~27.0	27.2	14.0~14.5
		1.00	1C	33.6~34.9	29.2	14.4~14.9

(注1) BS: 転送ブロック・サイズ, P: ページ, T: トラック, C: シリンダ を意味する。

(注2) 処理は、アクセス・パターンを示している。

大幅に改善されていることがわかる。また、各曲線の傾斜が緩やかになったためDKCビジー率がかなり高い所まで実用的になっている。

●レスポンスに関する評価

図4.3はバッファの有無による応答速度の比を示したものである。比較はバッファがない場合の実用的な範囲 ( $0.1 \leq B \leq 0.3$  かつ  $RR \geq 0.5$ ) に限定している。B及びTがそれぞれ等しい所で比較しているので、図は同一負荷における応答速度の比を示している。図から応答速度が2.5~3.8倍(応答時間は40~

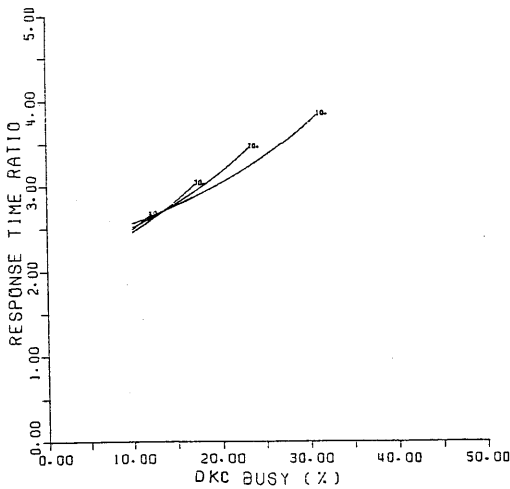


図4.3 レスポンスの向上率

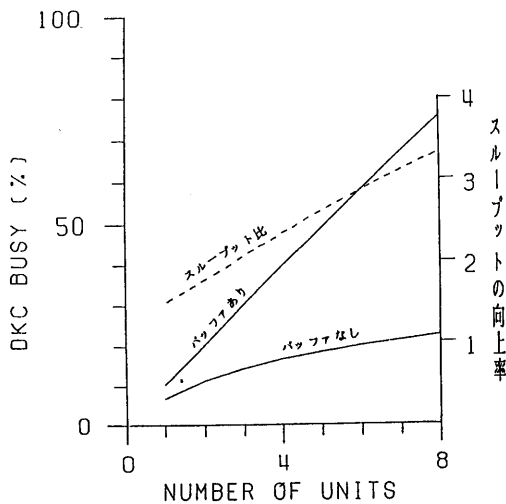
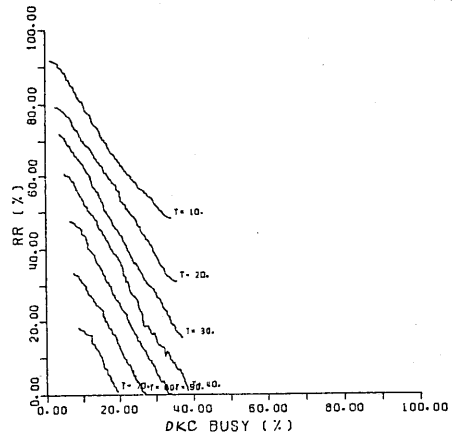
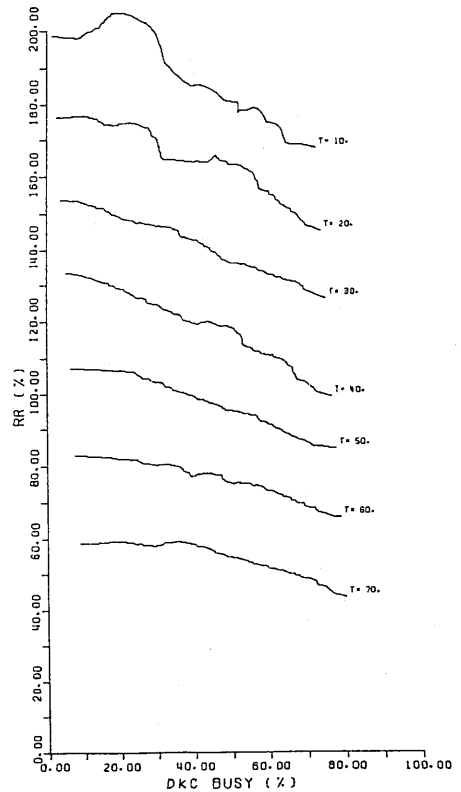


図4.4 スループットの向上率



(a) バッファなし



(b) バッファあり

図4.2 B-DISKの応答特性



26%に減少)になることが期待できる。【データベース・アクセスの場合、応答速度は1.7~2.7倍(応答時間は59~37%に減少)】

●スループットに関する評価

図4.4の実線はRR=0.5すなわち許容できる応答速度の限界におけるディスク台数とDKCビジー率の関係を示している。バッファがない場合DKCビジー率は30%前後で飽和してしまっているが、バッファがある場合は約80%までほぼ直線的に上昇している。両者のビジー率の比を図4.4に破線で示している。図からディスク8台の場合バッファによって3.3倍のI/O要求を処理できることがわかる。【データベース・アクセスの場合は2.0倍】

4.3 MESSIAHの性能評価結果

表4.3は、MESSIAH導入による入出力処理時間の改善率を示している。表4.3で、AはDBアクセスによく見られるランダムなアクセス・パターンであり、Dはほぼシーケンシャルなアクセス・パターンである。そして、表4.3から、MESSIAH利用により入出力処理をA:約1.1~3.6倍、D:約2.2~10.2倍高速化できることがわかる。尚、高速化は、DC使用時にA:約0.76~1.8倍、D:約1.0~2.1倍、B-DISK使用時にA:約1.6

倍、D:約2.2倍である。

ここで、アクセス・パターンAに対する平均デバイス・サービス時間(DST)の改善率を表すと図4.5のようになる。図4.5では、従来のディスク・サブシステムのDSTを1としたときのDC,B-DISK,MESSIAHのDSTを示している。

つぎに、評価結果において注目すべきことは、d式の関係が生じていることである。

$$(MESSIAHによる高速化) > (B-DISKによる高速化) \times (DCによる高速化) \dots (d)$$

d式は、B-DISKとDCとを階層的に結合したことにより、相乗効果が得られていることを示している。これにより、MESSIAHの構成が非常に有効であることを確認することができた。また、MESSIAHに必要なバッファ・メモリのサイズは、B-DISKでは数十KB/スピンドル、DCでは約1MB/スピンドルである。

5. まとめ

本報告では、B-DISKとDCとを階層的に持つ階層型ディスクキャッシュ・サブシステム(MESSIAH)を提案し、実験システムに対してシミュレーションによる性能評価を行った。また、MESSIAHの性能評価のためにB-DISK,DCの性能評価も行った。そしてその結果、DC部をライ

表4.3 MESSIAHによるDSTの改善率

処理	MESSIAHの使用	DCのサイズ(MB/スピンドル)	BS	DST(MS)	平均速度比(倍)
A	不使用	————	最大1P	26.8~27.3	1.000
D	不使用	————	最大1P	29.0~29.7	1.000
A	使用	0.25	1P	8.5~8.7	3.011
		12.00		7.5~7.6	3.613
		0.25	1T	12.3~12.8	2.168
		12.00		8.8~9.5	2.978
		0.25	2T	23.9~29.2	1.059
		12.00		12.4~14.5	2.053
D	使用	1.00	1P	13.1~13.8	2.201
		1.00	1T	6.3~6.8	4.538
		1.00	2T	5.7~6.0	5.086
		1.00	1C	2.7~3.0	10.172

(注1) BS: 転送ブロック・サイズ, P: ページ, T: トラック, C: シリンダ を意味する。

(注2) 処理は、アクセス・パターンを示している。

ト・スルー方式で制御するMESSIAHを導入することにより、入出力処理をデータベース処理(A)で3.6倍、ファイル処理(D)で10.2倍程度まで高速化できる可能性があることを示した。

さらに、

(MESSIAHの効果) > (B-DISKの効果) × (DCの効果)の結果を得ることができた。これにより、MESSIAHがB-DISKとDCとの相乗効果を実現できることが確認できた。

以上に示したように、MESSIAHの方式が入出力処理の高速化に非常に有効であるといえる。また、以上の効果を得るために必要なバッファ・メモリのサイズは、B-DISK部: 数十KB/スピンドル、DC部: 約1MB/スピンドルであるが、これは十分商用化できる値である。

ところで、今回の報告ではMESSIAHのDC部をライト・スルー方式で制御する場合の性能を評価したが、さらに入出力処理を高速化する方法として、DC部をライト・アフター方式で制御することが考えられる。この場合の評価は今後の課題である。

〔謝 辞〕

本研究の機会を与えていただいた、当社情報電子研究所

情報処理部首藤勝部長、ならびに、熱心なご助言をいただいた情報電子研究所菅隆志氏に深謝致します。

〔参考文献〕

- 1) Brown, D.T. et al., "Channel and Direct Access Device Architecture," IBM Syst. J., Vol. 23, No. 11, pp. 186-199, 1972.
- 2) Tokunaga, T. et al., "Integrated Disk Cache System with File Adaptive Control," Proc. IEEE COMPCON Fall, pp. 412-416, 1980.
- 3) 菅他, 「ディスク・キャッシュ装置のシミュレーションによる効率測定」, 情報処理学会論文誌, Vol. 22, No. 1, pp. 22-28, 1981.
- 4) 宮地他, 「キャッシュ付きディスク装置の性能評価」, 情報処理学会 第24回全国大会論文集 6D-2, 1982.
- 5) 三石他, 「バッファ内蔵型ディスク装置とその性能評価」, 情報処理学会 第24回全国大会論集 6D-1, 1982.

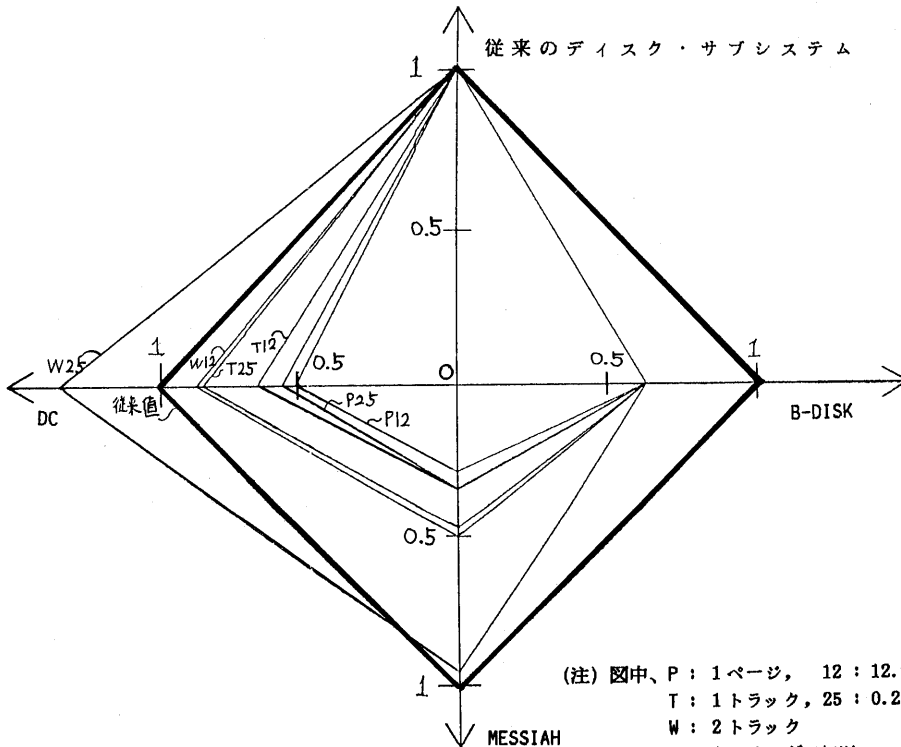


図4.5 平均デバイス・サービス時間の改善度 (アクセス・パターンAの場合)