連続転送方式に基づくカートリッジ型 MT の先読み／まとめ書きスケジューリング・アルゴリズムとその性能解析†

山本 彰†† 坪井俊明††† 北嶋弘行††

連続転送方式を基本にした，カートリッジ型 MT の先読み／まとめ書きスケジューリング・アルゴリズムの提案と，その性能解析を行う．カートリッジ型 MT においては，MT の高速スタート／ストップ動作には不可欠である真空カラムを，装置の小型化を目的として除去する．このため，スタート／ストップ回数の削減を目的として，制御装置内にバッファを設け，バッファと MT の間は複数ブロックの先読み／まとめ書きを実行する．したがって，先読み／まとめ書きスケジューリング方式がカートリッジ型 MT の性能に大きい影響を与えることになる．連続転送方式とは，バッファと MT の間のデータ転送路を間断なく使用し，MT の転送速度に等しいスループットを保証するというものである．解析モデルにより，真空カラム除去により生ずる性能劣化を防止するために，十分条件とはならないが，必要条件となるバッファ容量を導いた．さらに，本論文で評価対象としたシミュレーションでは，性能劣化の防止が困難な領域においても，解析モデルを用いて導出したバッファ容量より 20% 大きいバッファ容量によって，真空カラム除去に起因して生ずる性能劣化の防止が達成されるという結果を得た．

1. はじめに

省スペース化，掛け換え数の少なさなどに対するニーズから，小型化，軽量化，および，ライブラリ化の容易なカートリッジ型の MT の開発が進められている[1),2)]．

カートリッジ型 MT においては，MT 装置自体の小型化，軽量化のために，MT の高速スタート／ストップ動作には必須である真空カラムを除去することが必要となる．(MT におけるスタート動作とは，MT の走行速度を停止した状態からデータの読み書きが可能となる一定の速度に引き上げる操作を意味する．) 真空カラムが存在する従来の MT の場合には，MT のスタート／ストップ動作は，MT 上の記録単位であるブロックとブロックの間に存在するギャップ（間隔）の範囲内で実行可能であった[3)]．CPU と MT の間のデータ転送単位も通常ブロック単位であるため，従来 MT の制御装置は，チャネルから入出力要求を受け取ったとき，MT をスタートさせ，データ転送後，MT を停止させるという単純な制御で十分な性能を確保できた．

真空カラムを除去した場合，スタート／ストップ動作に要する時間が従来の数十倍となる．さらに，MT を停止させる際，停止までの間に，MT の読み書きヘッドが次のブロックを通過してしまうため，再び，次のブロックから先読み／まとめ書きを開始する場合，再位置付け処理と呼ばれる位置付け処理が必要となる．したがって，カートリッジ型 MT においては，従来の制御方式をそのまま踏襲すると大幅な性能劣化が生ずるため，ディスク・キャッシュ[4)]と同様に，半導体メモリから構成されるバッファを制御装置内に設け，以下に示す制御方式をとる[1),2)]．

制御装置内の半導体メモリを中間バッファとして，チャネルとバッファ，バッファと MT，それぞれのデータ転送処理を MT の転送速度で並列に実行させる．さらに，スタート／ストップ動作の実行回数を削減することを目的として，MT とバッファの間は複数ブロックの先読み／まとめ書き処理を実行する．

以上に加え，データ転送路を効率的に利用するため，スタート／ストップ動作を制御装置とは切り離した状態（オフライン）で動作させる．これは，ディスク装置におけるオフライン・シーク／サーチ制御と等価な制御である．ただし，ディスク装置で用いているオフライン・サーチ制御[5)]は，サーチ要求を受け付けると直ちに要求をディスクに対して発行するため，要求完了時に，他のディスクが転送路を占有している可能性があった．この場合，ディスクが 1 回転してから，データ転送に入る必要が生ずる[5)]．

以上より，カートリッジ型 MT においては，先読み／まとめ書きの実現方式が，その性能に大きく影響

† A Continuous Data Transfer Scheduling Algorithm for Preload/Batch-Write of Cartridge Type MTs, and Its Performance Analyses by AKIRA YAMAMOTO (Systems Development Laboratory, Hitachi, Ltd.), TOSHIAKI TSUBOI (Hitachi Microcomputer Engineering, Ltd.) and HIROYUKI KITAJIMA (Systems Development Laboratory, Hitachi, Ltd.).

†† (株)日立製作所システム開発研究所
††† 日立マイクロコンピュータエンジニアリング(株)

することになる．本論文では，真空カラム除去により生ずる性能劣化を防止し，従来 MT と同等の性能を得ることを目標とする連続データ転送方式に基づく先読み／まとめ書きアルゴリズムを提案する．従来のディスク装置で用いられているオフライン・サーチ制御では，データ転送路の空き時間をまったくなくすという保証が困難であった．連続データ転送方式とは，MT のスタート時間が一定分布であることに着目し，バッファと MT の間のデータ転送路を間断なく利用し，MT の転送速度と同等なスループットを得ることを可能とした制御方式である．MT とバッファ間の転送路のスループットが MT の速度と同等になると，バッファを介して動作しているチャネルとバッファ間の転送路のスループットも MT の速度と同等にすることが可能なため，上位システムから見て，従来 MT の性能を達成するという目標を実現できることになる．

第2章では，先読み／まとめ書きアルゴリズムの内容を示し，第3章では，各 MT に対するアクセス比率が異なっていてよいという条件の下で，提案アルゴリズムを適用したとき，この値未満では，真空カラム除去により生ずる性能劣化を防止できないというバッファ容量を解析する．最後に，第4章では，3章の解析値に基づき，シミュレーション・モデルにより，提案アルゴリズムの評価を行う．ただし，第3章，第4章の解析は，動作中の MT の集合は変化しないという前提に基づいている．動作中の MT の集合の変化は，MT の入出力処理の開始／終了動作に人手の介入が通常必要なことから，数十秒単位であることが一般的である．これに対し，各 MT の先読み／まとめ書き処理のサイクルが数百 msec のオーダ（スタート時間＋再位置付け時間のオーダは高々 100 msec のオーダである．本論文で提案する先読み／まとめ書きのサイクルは，第2章で示すように，これと同じオーダとなる．）であるため，ある一定の MT の集合に対して先読み／まとめ書き処理は定常状態になると考えてよい．したがって，第3章，第4章の解析は十分有用であると考える．

2. 連続転送方式に基づく先読み／まとめ書きアルゴリズム

図1にカートリッジ型 MT を含む計算機システムの構成，**表1**に本論文で用いる記号の定義を示す．以下，本論文では，ライト動作中の MT をライト MT，

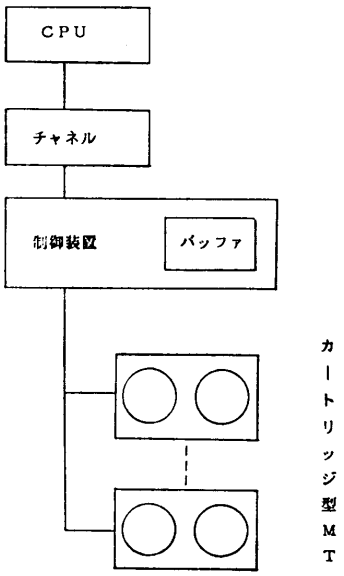

図1 カートリッジ型MTを含む計算機システムの構成
Fig. 1 A computer system with cartridge type MTs.

表1 記号の定義
Table 1 Definition of symbols.

s	MTのスタート時間
r	MTの再位置付け時間
t	MTの転送速度（転送バイト数／単位時間）
βi	MTi のサイクル内先読み／まとめ書き時間
β_{MAX}	それぞれのMTのサイクル内先読み／まとめ書き時間の最大値
αi	MTi のアクセス比率
α_{MAX}	最もアクセス比率の高いMTのアクセス比率
E_W	動作中のMTがライトMTのみの場合，各MTのバッファ占有量の合計
E_R	動作中のMTがリードMTのみの場合，各MTのバッファ占有量の合計
$E_{R/W}$	リード MT／ライト MT が混在する場合の各 MT のバッファ占有量の合計
$E_{R/W,R}$	リード MT／ライト MT が混在する場合，リード MT のバッファ占有量の合計
$E_{R/W,W}$	リード MT／ライト MT が混在する場合，ライト MT のバッファ占有量の合計

リード動作中の MT をリード MT と呼ぶ．制御装置内には，バッファを設け，第1章で述べたように，バッファと MT の間で複数ブロックの先読み／まとめ書き処理を実行する．また，チャネルから受け付けた要求は，ライト MT の場合，バッファが満杯のとき，リード MT の場合，その MT の先読みデータがないとき，待ち状態に入ることになる．以下，連続転送方式の基本的な考え方を示し，次に具体的なアルゴリズムについて述べる．

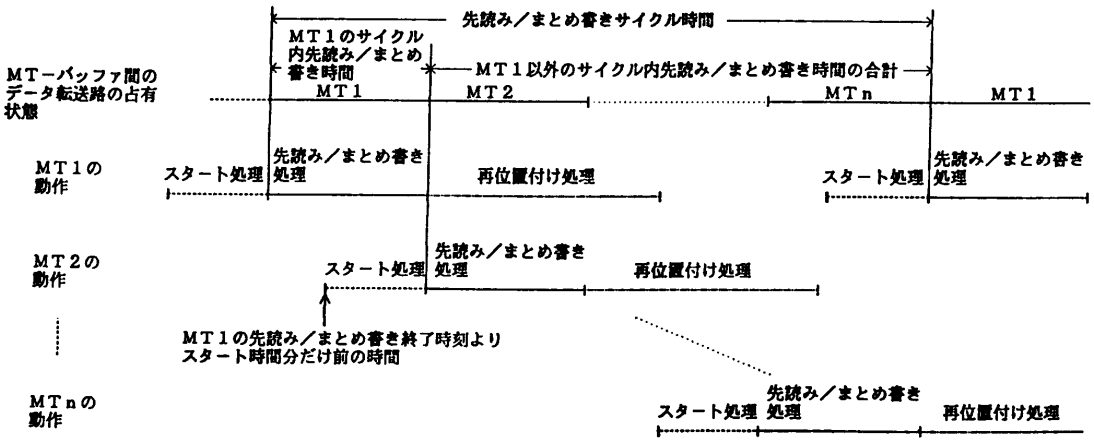

図 2 連続転送方式の基本的原理
Fig. 2 Basic principle of continuous transfer mechanism.

2.1 連続転送方式の考え方

図2に連続転送方式の基本的な考え方を示し，以下に，その制御目標を示す．図2では，先読み／まとめ書きを行う MT の集合を，MT1, MT2, …, MTn としている．

制御目標：MT とバッファの間のデータ転送路を間断なく利用可能とし，MT の転送速度と等価なスループットを引き出す．

上記制御目標を，以下の基本制御方式より実現する．

基本制御方式 1：先読み／まとめ書き対象 MT を切り換える際に，現在転送中の MT の転送終了時刻からスタート時間分だけ前に，次に先読み／まとめ書き処理を行う MT に対してスタート要求を発行する．カートリッジ型 MT の場合，MT のスタート処理に要する時間は，ほぼ一定時間である．したがって，図2に示したように，以上の制御により，MT1 から MT2 への先読み／まとめ書き処理の切り換えから，MTn−1 から MTn までの先読み／まとめ書き処理の切り換えの際に，データ転送路を間断なく利用することができる．

基本制御方式 2：各 MT の先読み／まとめ書きはサイクリックに行うものとし，このサイクルを先読み／まとめ書きサイクル，サイクル時間を先読み／まとめ書きサイクル時間と呼ぶ．また，サイクル内の各 MT の先読み／まとめ書き時間をサイクル内先読み／まとめ書き時間と呼ぶ．このとき，すべての MT に関し，先読み／まとめ書きサイクル時間からその MT のサイクル内先読み／まとめ書き時間を引いた時間が，MT のスタート時間＋再位置付け時間以上になるようにする．以上の制御により，図2に示すように，MTn の先読み／まとめ書き処理が終了する時刻のスタート時間前には，MT1 の再位置付け処理がすでに完了しているため，MT1 のスタート処理の開始に入ることが可能となる．

したがって，最もサイクル内先読み／まとめ書き時間の長い MT に対し，基本制御方式2を実現させることにより，すべての MT の先読み/まとめ書きの切り換えの際に，連続的な転送路の利用が可能となる．これより，基本制御方式2を満足する先読み／まとめ書きサイクル時間の最小値は以下の式で表すことができる．

$$T=(s+r)+\beta_{MAX} \qquad (1)$$

以上により，連続転送方式の制御目標を実現できることを示すことができた．次節では，本方式に基づく具体的な先読み／まとめ書きアルゴリズムを示す．

2.2 先読み／まとめ書きアルゴリズム

本節では，前節で示した連続転送方式に基づく先読み／まとめ書きスケジューリング・アルゴリズムを示す．まず，スケジュール処理を実行するそれぞれのスケジュール時点で決定する具体的項目を，以下に示す．

項目1：次のスケジュール時点（実行時刻）の設定．

項目2：MT へのスタート要求の発行．

項目3：先読み／まとめ書き対象（スケジュール対象）MT の切り換え．

項目4：設定，あるいは，追加する先読み／まとめ書き時間．

なお，ここで提案するスケジューリング・アルゴリ

ズムにおいては，一度あるMTのスケジューリングを開始すると，ある条件を満足するまで，このMTのスケジューリングを行う．この一連の動作を，連続スケジュール処理と呼ぶ．さらに，連続スケジュール処理における最初のスケジュール時点を，連続スケジュール初期時点，最後のスケジュール時点を，連続スケジュール最終時点と呼ぶ．以上より，各MTの先読み／まとめ書き時間は，連続スケジュール初期時点において，新たに設定し，以降のスケジュール時点で，追加していく形をとるようにした．したがって，各MTの連続スケジュール初期時点から連続スケジュール最終時点までの間に，設定，追加した先読み／まとめ書き時間の合計が，それぞれのMTのサイクル内先読み／まとめ書き時間になることになる．

以下，スケジュール対象MTをMTiとして，具体的な方式を示す．

[方式1＝基本制御方式1の実現：項目1，項目2の決定]

次のスケジュール時点を，現スケジュール時点において，MTiに対して設定，あるいは，追加した先読み／まとめ書き時間分だけ後に設定する．

さらに，現スケジュール時点が，MTiの連続スケジュール初期時点であるときには，MTiに対するスタート要求を発行する．

[方式2＝基本制御方式2の実現：項目3の決定]

次式が成立したとき，現スケジュール時点を，連続スケジュール最終時点とする．すなわち，次のスケジュール時点において，スケジュール対象MTを，MTiからMTi+1に切り換える．

(現スケジュール時点－前回の連続スケジュール最終時点)－(前回の連続スケジュール最終時点以降のMTiのチャネル転送時間)$\geqq(s+r)$ (2)

前回の連続スケジュール最終時点とは，MTiに対する前回の連続スケジュール処理における，連続スケジュール最終時点を表す．

[方式3：項目4の決定]

設定，追加する先読み／まとめ書き時間を，この前MTiのスケジュールを行った時点以降の，MTiのチャネル転送時間とする．ただし，リードMTの場合，最初に要求を受け付けたとき，チャネルとの転送は行っていないため，適当な値，たとえば，バッファ容量をMT台数で割った値等を用いてスケジュールを行う．この場合，次の先読み／まとめ書きの順番までに，先読みデータがなくなったときなど，最初のスケジュール量を調節する．(これにより，定常状態においては十分な先読みデータを確保できると考えられる．)

以下，図3を用いて，方式1-方式2により，前節で示した基本制御方式1，基本制御方式2が実現できること，および，方式3の考え方について説明する．

図3に示したように，スケジュール時点$k-1$では，スケジュール時点kを，スケジュール時点$k-1$で設定した先読み／まとめ書き時間$k-1$だけ後に設定する．同様の関係は，スケジュール時点$k+1$とスケジュール時点kの間でも成立する．さらに，スケジュール時点$k-1$は，連続スケジュール初期時点であるため，MTiにスタート要求を発行する．以上により，図3に示したように，各スケジュール時点を，それぞれのスケジュール時点で設定，あるいは，追加する先読み／まとめ書き処理の開始時刻に比較して，常に，MTのスタート時間分だけ，先行させることができる．このため，図3においては，スケジュール時点$k+1$で，MTi+1にスタート要求を発行することにより，転送対象MTをMTiからMTi+1に切り換える際の転送路の連続利用を可能にしている．以上により，方式1により，基本制御方式1を実現できることは明らかである．

図3においては，スケジュール時点kが，連続スケジュール最終時点となっている．また，スケジュール時点hが，前回の連続スケジュール最終時点となっている．すでに述べたように，(2)式が成立したスケジュール時点は連続スケジュール最終時点となる．したがって，(2)式が成立したとき，(2)式中の(現スケジュール時点－前回の連続最終スケジュール時点)は，図3に示したように，明らかに，先読み／まとめ書きサイクル時間となる．さらに，方式3で示したように各MTの先読み／まとめ書き時間を決定した場合，図3より明らかに，(前回の連続最終スケジュール時点以降のMTiのチャネル転送時間)が，MTiのサイクル内先読み／まとめ書き時間となる．これより，(2)式の左辺は，先読み／まとめ書きサイクル時間から，MTiのサイクル内先読み／まとめ書き時間を引いた値になる．したがって，(2)式が成立した時点で，各MTのスケジュール処理を切り換えることにより，基本制御方式2の実現が可能となる．

方式3は，各MTに対する先読み／まとめ書き量の配分に関する．方式3は，図3に示したように，1回の先読み／まとめ書きサイクルの間に，各MTの先読み／まとめ書き量とチャネルとの転送量を等しく

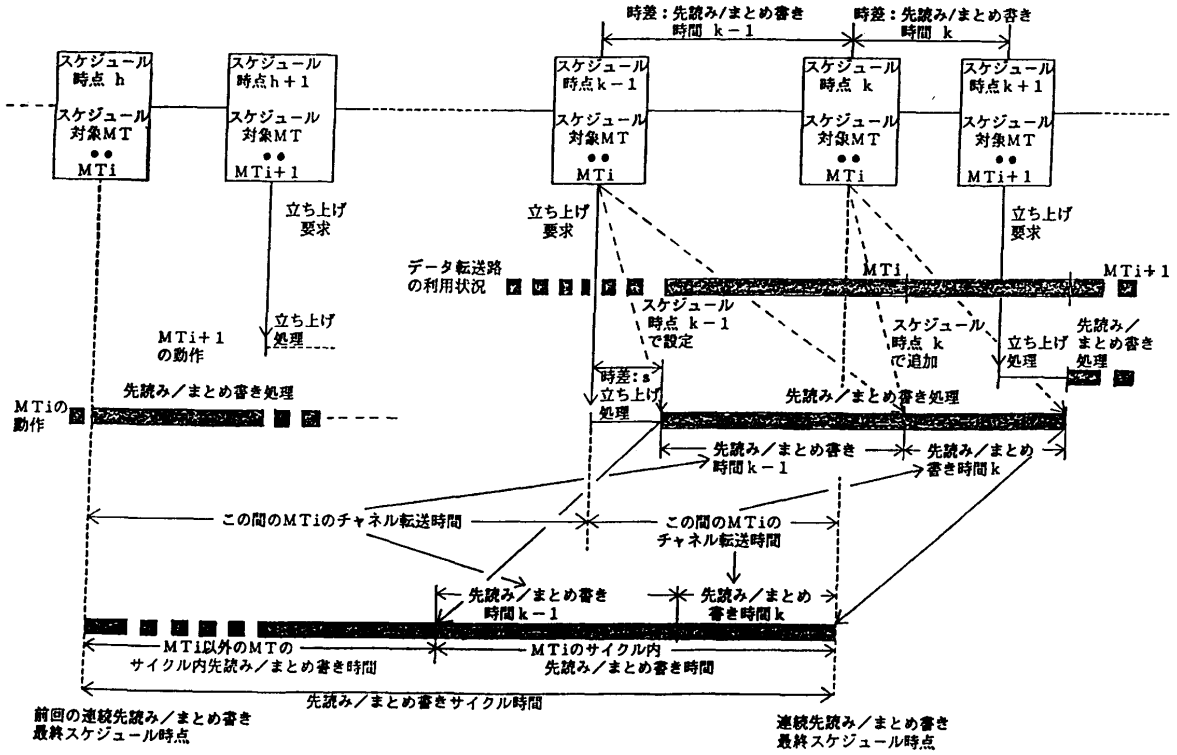

図 3 スケジューリング・アルゴリズムの概要
Fig. 3 Outline of the presented scheduling algorithm.

することを目標としている．先読み／まとめ書き量とチャネルとの転送量は，定常的には等しくする必要があるため，本方式は，この保証を1回の先読み／まとめ書きサイクル内で実現しようとしていることになる．

以上が，本論文で提案する先読み／まとめ書きスケジューリング・アルゴリズムの内容である．次章以降，提案アルゴリズムの性能解析を行う．

3. 先読み／まとめ書きに必要なバッファ・サイズの解析

本章では，各 MT へのアクセス比率（利用率）の差異を許し，前章で示した先読み／まとめ書きアルゴリズムを適用したとき，バッファ容量がこの値未満では，従来の真空カラムを有する MT と同等の性能を得ることができないという値を解析モデルにより見積もる．ここで見積もるバッファ容量は，制御目標を満たすための十分条件ではなく，必要条件となるため，以下，必要条件バッファ容量と呼ぶ．したがって，実際の装置では，解析値より数十%の余裕をもつのが，実用的である．ただし，以下の解析においては，各 MT のアクセス比率の総和は1とした．各 MT のアクセス比率の総和が1に達しない場合は，総和が1になる場合に比較して必要条件バッファ容量は小さくなる．これは，転送路が空いている時間をある仮想的な MT に対して入出力処理を実行していると考えると，この仮想的な MT に割り当てるバッファが実際には必要なくなるためである．まず，解析にあたり設けた仮定を以下に示す．

仮定1：各 MT のバッファの割り当て単位は微小な単位とし，バッファを確保，解放するタイミングは以下のタイミングとする．

確保時点：割り当て単位にデータを入力した時点

解放時点：割り当て単位からデータを出力した時点

仮定 2：各 MT のアクセス比率は，時間によらず一定とする．

仮定 3：リード MT のバッファ占有量は，この MT の先読み開始時に0となるものとする．

仮定 4：アクセス比率が最大の MT がライト MT である場合，この MT の連続スケジュール最終時点で設定，追加する先読み／まとめ書き時間はほぼ0に等しい時間であるとする．

以下，各仮定が解析結果に与える影響について述べる．仮定1に関するバッファの確保，解放の単位は，

通常，チャネル，MTとの間のデータ転送単位，すなわち，ブロックであるため，実際の確保時点は，そのブロックのバッファへの書き込み転送開始時点であり，解放時点は，ブロックのバッファからの読み出し転送終了時点である．解析対象システムでは，チャネルとの転送，MTとの転送を並列に実行可能としたため，転送中のブロックは最大2個である．したがって，仮定1を設けたことにより，解析結果は，最大2ブロック・サイズだけ，制御目標を満たすバッファ容量を過少評価することになる．しかし，ブロック・サイズは通常数 kbyte から数十 kbyte であると考えられるため，上記の過小評価による影響は小さいと考えられる．

アクセス比率の時間によるばらつきを考慮すると先読みデータ量等に余裕量が必要となるため，仮定2は，制御目標を満たすバッファ容量を過小評価する傾向をもつ．しかし，本章の冒頭で述べた必要バッファ容量解析という目的から考えると本仮定は妥当である．

仮定2が成立したとき，前章の方式 1-方式3に示した先読み／まとめ書きスケジュール方式を適用すると，定常状態においては，先読み／まとめ書きサイクル時間は常に一定となる．また，各 MT のサイクル内先読み／まとめ書き時間，先読み／まとめ書き開始・終了時のバッファ占有量等が，各先読み／まとめ書きサイクルごとに，等しくなることは明らかである．

各 MT のバッファ占有量が最小となるのは，リード MT の場合，先読みの開始時点，ライト MT の場合，まとめ書きの終了時点である．リード MT の場合，バッファ占有量は，このときの値を最小値として，先読み中は，$(1-\alpha_i)$ の割合で増大し，先読みが終了すると α_i の割合で減少していき，再び先読み開始時には最小値をとることになる．以上より明らかに，仮定3が成立するとき，リード MT に関しては，バッファ占有量が最小となる．仮定3を前提とするのは，必要条件バッファ容量の解析という目的から，妥当であると考えられる．

ライト MT に関しても，まとめ書き中は $(1-\alpha_i)$ でバッファ占有量が減少し，まとめ書きを行っていないときには，α_i の割合で増大することになり，まとめ書き終了時点におけるバッファ内データ占有量が0のとき，バッファ占有量が最小となる．しかし，各スケジュール時点で，前章の方式3に示したように，まとめ書き時間を決定した場合，まとめ書き対象データは，前回のスケジュール時点以降チャネルとの間で転送したデータとなる．したがって，連続スケジュール最終時点以降バッファ内に転送されたデータのまとめ書き処理は，次の連続スケジュール処理において実行せざるをえない．このため，まとめ書き終了時点においては，連続スケジュール最終時点以降バッファ内に転送されたデータがバッファを占有していることになる．

本論文で提案する先読み／まとめ書き方式においては，前章の方式2に示したように，(2)式が成立したスケジュール時点が連続スケジュール最終時点となる．仮定2を前提にすると，定常的には，アクセス比率が最大の MT 以外の MT は，連続スケジュール初期時点で(2)式が成立し，連続スケジュール初期時点が連続スケジュール最終時点となる．この証明は付録で行う．

以上より，アクセス比率が最大の MT のみが複数回のスケジュール時点を必要とすることになる．連続スケジュール最終時点からまとめ書き終了時点までの時間は，前章で示した方式1，方式3から，(連続スケジュール最終時点において追加，設定した先読み／まとめ書き時間＋MT のスタート時間) となる．このため，アクセス比率が最大であるライト MT のバッファ占有量が最小となるのは，連続スケジュール最終時点において追加，設定した先読み／まとめ書き時間がほぼ0となった場合，すなわち，仮定4が成立した場合となる．同一の MT のスケジュール時点の時間間隔，および，先読み／まとめ書き時間は，前章の方式1，方式3で示したように決定するため，仮定2を前提にすると，この MT のアクセス比率の等比級数となり，徐々に減少して，0に収束していくことになる．以上より，制御目標を実現する必要条件バッファ容量を見積もるという本解析の目的より，仮定4の成立を前提とするのは妥当であると考えられる．

次に，必要条件バッファ容量を見積もる．必要条件バッファ容量は，仮定 1-仮定4を前提としたとき，各 MT のアクセス比率の任意の組合せに対し，MT の転送速度と同等のスループットを保証する最小のバッファ容量ということになる．ここでは，MT の集合を MT1, MT2, …, MTn とし，アクセス比率最大の MT を MTn とする．前章の方式3に示したように，各スケジュール時点の先読み／まとめ書き時間を決めた場合，各 MT のサイクル内先読み／まとめ書き時

間は $T\alpha_i$ となるため，(1)式は以下のようになる．

$$T=(s+r)/(1-\alpha_{MAX}) \quad (3)$$

以下，3つのケースに分類して，必要条件バッファ容量の見積もりを行う．

［ケース1］ すべての MT がライト MT であるとき

転送路の利用率が1であれば，チャネル側からバッファにデータが書き込まれる速度と同じ速度でバッファからデータが読み出され MT に書き込まれていることになるため，すべての MT のバッファ占有量の総和は時間によらず変化しない．したがって，ここでは，アクセス比率最大の MTn のまとめ書きが完了した時点の解析を行う．まず，各 MT のバッファ占有量を示す．

MTn のバッファ占有量

$\alpha_n st$

MTi $(i=1,\cdots,n-1)$ のバッファ占有量

$$\left(\sum_{j=i}^{n}\alpha_j\right)\alpha_i Tt+\alpha_i st$$

したがって，バッファ占有量の合計 E_W は次式で表せる．

$$E_W=\sum_{i=1}^{n-1}\left\{\left(\sum_{j=i}^{n}\alpha_j\right)\alpha_i Tt+\alpha_i st\right\}+\alpha_n st \quad (4)$$

(4)式に(3)式を代入し，整理すると以下の式が得られる．

$$E_W=(s+r)t-\left\{\sum_{i=2}^{n-1}\left(\sum_{j=1}^{i-1}\alpha_j\right)\alpha_i\right\}(s+r)t/(1-\alpha_n)+st \quad (5)$$

(5)式は α_n, すなわち，MTn のアクセス比率が1に収束すると上限値 $(2s+r)t$ をとる．したがって，ケース1の必要条件バッファ容量は $(2s+r)t$ となる．

［ケース2］ すべての MT がリード MT であるとき

ケース1と同様，すべての MT のバッファ占有量の総和は時間によらず変化しない．したがって，ここでも，アクセス比率最大の MTn の先読みが終了した時点の解析を行う．まず，各 MT のバッファ占有量を示す．

MTi $(i=1,\cdots,n)$ のバッファ占有量

$$\alpha_i Tt-\left(\sum_{j=i}^{n}\alpha_j\right)\alpha_i Tt$$

したがって，バッファ占有量の合計 E_R は次式で表せる．

$$E_R=\sum_{i=1}^{n}\left\{\alpha_i Tt-\left(\sum_{j=i}^{n}\alpha_j\right)\alpha_i Tt\right\} \quad (6)$$

ケース1と同様に展開すると，ケース2の必要条件バッファ容量は $(s+r)t$ となることがわかる．

［ケース3］ リード MT とライト MT が混在するとき

この場合，各 MT のバッファ占有量の総和は一定ではなく，先読みを実行しているときには増大傾向にあり，まとめ書きを実行しているときには減少傾向にある．この場合次式が成立する．

$$E_{R/W}=E_{R/W,W}+E_{R/W,R} \quad (7)$$

$$E_{R/W,W}<E_W \quad (8)$$

$$E_{R/W,R}<E_R \quad (9)$$

(8)式が成立する理由は，実際にはリード MT である MT をライト MT と仮定したとき，これらの MT が占有するバッファ容量と，$E_{R/W,W}$ をあわせた値が，E_W となるためである．同様の理由から(9)式が成立する．必要条件バッファ容量は，$E_{R/W}$ の上限値であるため，以上より，E_W, E_R それぞれの上限値の和である $(3s+2r)t$ より大きくなることはないことになる．以下，この値が，$E_{R/W}$ の上限値となることを示す．いま，MT 台数を2台とし，MT1 をリード MT, MT2 をライト MT とする．バッファ占有量の和は MT1 の先読みが完了したときが最大となるため，この時点における $E_{R/W}$ の値を解析する．ただし，以下では，簡単のため，$\alpha_1<\alpha_2$ の成立を前提とした．

$$\begin{aligned}E_{R/W}&=(1-\alpha_2)\alpha_2(s+r)t/(1-\alpha_2)\\&\quad+\alpha_1\alpha_2(s+r)t/(1-\alpha_2)+\alpha_2 st\\&=2\alpha_2(s+r)t+\alpha_2 st\end{aligned} \quad (10)$$

(10)式は明らかに，α_2, すなわち，MT2 のアクセス比率が1に収束するにつれ，$(3s+2r)t$ に収束する．したがって，この場合の必要条件バッファ容量は $(3s+2r)t$ となる．

通常の動作環境においては，リード MT とライト MT は混在するのが一般的であることから，制御装置内のバッファ容量は，リード MT/ライト MT 混在時の必要条件バッファ容量を基準として定めるべきである．次章では，本章の解析結果に基づき，バッファ容量と性能の関係を評価する．

4. シミュレーション・モデルによる性能評価

本章では，シミュレーション・モデル[6]により，前章で解析した必要条件バッファ容量を基準バッファ容量として，本論文で提案した先読み／まとめ書きアル

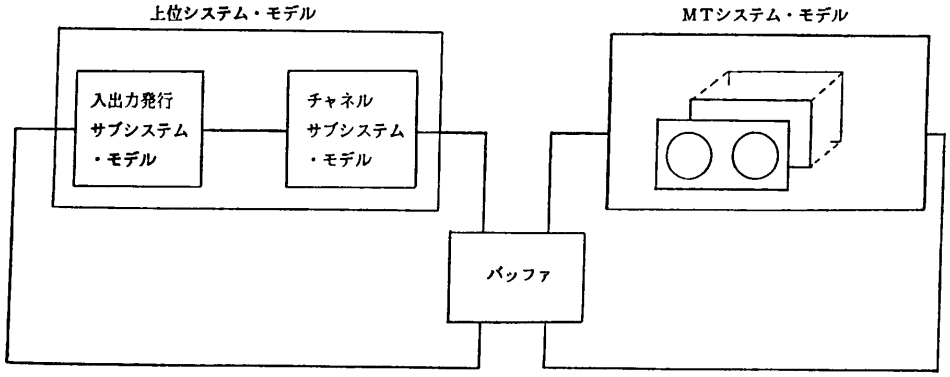

図4 シミュレーション・モデル
Fig. 4 Simulation model.

ゴリズムを適用したときの性能と，従来 MT の性能を比較評価する．

図4は，評価に用いたシミュレーション・モデルである．シミュレーション・モデルは，カートリッジ型MTを含むモデルであり，制御装置内のバッファを中間バッファとして上位システム・モデルと MT システム・モデルより構成する．各システム・モデル内には，各 MT に1対1に対応する入出力要求が走行する．また，本シミュレーション・モデルにおいては，バッファ確保・解放の単位はブロックとした．

上位システム・モデルは，入出力発行サブシステム・モデルとチャネルサブシステム・モデルより構成する．入出力発行サブシステム・モデルは，各 MT に等しい窓口をもち，制御装置に対する入出力処理が完了してから，次に入出力処理を発行するまで各入出力要求が滞在する．本シミュレーションにおいては，各 MT に対するアクセス比率を時間的に変動させるために，その滞在時間分布を指数分布とした．チャネルサブシステム・モデルは，チャネルとバッファの間のデータ転送処理をモデル化しており，ブロック単位転送を行う．

MT システム・モデルは各 MT とバッファの間の先読み／まとめ書き処理のモデルであり，2章で示したスケジューリング・アルゴリズムに従って，この処理を実行する．

以下，評価結果をまとめる．**表2**は，アクティブ MT 台数が2台のとき，それぞれの MT のアクセス比率を1対1から8対1まで，および，入出力発行間隔比率を，アクセス比率の小さい MT に対し，1から 1/3 まで変化させたときの評価結果をまとめたものである．これに対し，**表3**はアクティブ MT 台数を増やした場合の評価である．まず，表2の評価について述べる．2台の MT のうち，(1)ライト MT が2台，(2)リード MT が2台，(3)ライト MT 1台，リード MT 1台，の3ケースについて評価した．本評価においては，バッファ容量を，前章で得たリード MT，ライト MT 混在時の必要条件バッファ容量であ

表2 シミュレーション結果-1
Table 2 Simulation results-1.

アクセス比率	I/O発行比率	(1)ライトMT2台 バッファサイズ 基準値×1.0～×1.2	(2)リードMT2台 バッファサイズ 基準値×1.0～×1.2	(3)ライトMT，リードMT1台ずつ バッファサイズ 基準値×1.2	基準値×1.1	基準値
24	3	○	○	○	0.2%	2.8%
‥	2	○	○	○	0.3%	2.3%
3	1	○	○	○	0.3%	1.4%
18	3	○	○	○	0.1%	0.4%
‥	2	○	○	○	0.2%	0.4%
3	1	○	○	○	0.3%	1.3%
12	3	○	○	○	○	○
‥	2	○	○	○	○	0.2%
3	1	○	○	○	○	0.1%
9	3	○	○	○	○	○
‥	2	○	○	○	○	○
3	1	○	○	○	○	0.1%
6	3	○	○	○	○	○
‥	2	○	○	○	○	○
3	1	○	○	○	○	○
3	3	○	○	○	○	○
‥	2	○	○	○	○	○
3	1	○	○	○	○	○

数値：従来MTとの利用率の差を表す．
○：従来MTとの利用率の差が 0.1% 未満．
バッファサイズの基準値：$(3s+2r)t$.

表 3 シミュレーション結果-2
Table 3 Simulation results-2.

アクセス比率	I/O 発行比率	バッファサイズ $(3s+2R)t\times 1.2$
24×1 台（ライト） 3×1 台（ライト） 3×1 台（リード）	1	○
24×1 台（ライト） 3×2 台（ライト） 3×2 台（リード）	1	○
24×1 台（ライト） 3×3 台（ライト） 3×3 台（リード）	1	○

○：従来MTとの利用率の差が 0.1% 未満.

る $(3s+2r)t$ の 1.0-1.2 倍にした．このとき，チャネルサブシステム・モデルの利用率と従来 MT の場合のチャネルサブシステム・モデルの利用率を比較対象とした．この値が等しければ，従来 MT と同等の性能が得られ，目標を達成していることになる．（従来 MT の場合，上位システム・モデルのみからなるモデルを評価した．）

以下，評価結果について述べる．表2の評価において○は，従来 MT の利用率との利用率の差が 0.1% 未満だった評価ケースであり，0.1% 以上だった評価ケースに関しては，従来 MT の利用率との利用率の差を示した．(1)ライト MT が2台，(2)リード MT が2台の場合は，前章で解析した必要条件バッファ容量が，ライト MT，リード MT が混在するときに比較して小さいため，すべての評価ケースにおいて○が得られている．一方，(3)ライト MT，リード MT が1台ずつの場合，○が得られていない評価ケースがある．バッファ容量が同一の場合，○が得られにくい傾向を以下に示す．

(a) アクセス比率にかたよりがある．

(b) 入出力発行間隔が小さい，すなわち，データ転送路の混雑度が大きい．

(a)の傾向は前章の解析と同様の傾向である．また，(b)の傾向が得られる原因については，3章の冒頭の部分で述べたとおりである．ただし，○が得られていない評価ケースにおいては，かえって，入出力発行間隔が大きい方が従来 MT との利用率の差が大きくなっている評価ケースが1つ存在する．具体的には，アクセス頻度が8対1，バッファ容量が $(3s+2r)t$ のときの評価ケースである．これは，バッファ不足が発生し，一方の MT にバッファが割り当てられないとき，入出力発行間隔が大きい方が，残りの MT が1台で動作しているときの利用率が低くなるため得られる傾向であると考えられる．しかし，アクセス頻度が8対1の場合でも，バッファ容量を 1.1 倍にした評価ケースでは，入出力発行間隔が小さくなるにつれ，利用率の差は，大きくなっているか，同じであるという，他の評価ケースと同様の傾向が得られている．以上は，入出力発行間隔が大きい方が，バッファ不足なしに，データ転送路を間断なく利用するために必要なバッファ容量は，小さくてすむが，さらに，バッファ容量を減らしたときの利用率の減少率が，場合によっては，大きくなることを示していると考えられる．

しかし，バッファ容量を 1.2 倍にした評価ケースに関してはすべて○が得られている．したがって，$(3s+2r)t$ の 1.2 倍のバッファ容量により，従来 MT と同等の性能を得るという目標は達成されていると考えてよい．

以上の結果を踏まえ，表3には MT 台数を増やした評価ケースをまとめた．ここでは，バッファ容量は $(3s+2r)t$ の 1.2 倍とし，表2の評価ケースで，○が得られにくかった評価ケースについての評価を行った．すなわち，リード MT とライト MT が混在し，かつ，(a)，(b)の傾向として示したように，アクセス比率にかたよりがあり，転送路が混雑しているケースである．評価結果は，やはり，すべての評価ケースについて従来 MT の場合と利用率の差が 0.1% 未満であった．

したがって，以上の評価においては，従来 MT と同等の性能が得られにくい領域においても，前章で解析した必要条件バッファ容量 $(3s+2r)t$ の 1.2 倍程度のバッファ容量によって，真空カラム除去により生ずる性能劣化を防止し，従来 MT と同等の性能を達成するという目標を満たす結果を得た．

5. おわりに

連続転送方式を基本にした，カートリッジ型 MT の先読み／まとめ書きスケジューリング・アルゴリズムの提案とそのスケジューリング・アルゴリズムの評価を行った．カートリッジ型 MT においては，MT の高速スタート／ストップ動作には不可欠である真空カラムを装置の小型化を目的として除去する．このため，スタート／ストップ回数の削減を目的として，制御装置内にバッファを設け，バッファと MT の間は

複数ブロックの先読み／まとめ書きを実行する．連続転送方式とは，バッファとMTの間のデータ転送路を間断なく使用し，MTの転送速度に等しいスループットを保証するというものである．

提案したスケジューリング・アルゴリズムを適用した場合，各MTのアクセス比率が異なるとき，解析モデルを用いて，バッファ容量がこの値未満では，真空カラム除去により生ずる性能劣化を防止できないという必要条件バッファ容量を導いた．さらに，シミュレーション・モデルによって，導出した必要条件バッファ容量を基礎値とした評価を行った．表2，表3に示した評価においては，従来MTと同等の性能が得られにくい領域でも，導出した必要条件バッファ容量の1.2倍程度のバッファ容量により，真空カラム除去により生ずる性能劣化を防止し，従来MTと同等の性能を達成するという目標を満足する結果を得た．

謝辞 終わりに，本研究の機会を与えてくださった(株)日立製作所システム開発研究所堂面信義所長，同社小田原工場下矢吉孝部長，有益なご指導，ご助言を与えてくださった同所石原孝一郎主管研究員，久保隆重部長，同工場宮崎道生副技師長，影浦憲一主任技師，西村利文主任技師，尾形幹人技師，日立コンピュータ機器(株)土井隆部長に深く感謝いたします．

参考文献

1) 山本ほか：カートリッジ型MTにおける先読み・纏め書きスケジューリング方式の提案と解析，第35回情報処理学会全国大会論文集，pp. 219-220 (1987).
2) 山本ほか：カートリッジ型MTにおける先読み・纏め書きスケジューリング方式の評価，第35回情報処理学会全国大会論文集，pp. 221-222 (1987).
3) 日立マニュアル：H-8488-1形磁気テープ装置，H-8487-1形磁気テープ装置，H-8468-1形磁気テープ装置，H-8467-1形磁気テープ装置，H-8481-1形磁気テープ制御装置，8080-2-005.
4) Smith, **A.** J.: Disk Cache—Miss Ratio Analysis and Design Considerations, *ACM TOCS*, Vol. 3, No. 3, pp. 161-203 (1985).
5) Brandwajn, A.: Models of DASD Subsystems: Basic Model of Reconnection, *Performance Evaluation*, Vol. 1, No. 3, pp. 263-281 (1981).
6) Kobayashi, H.: *Modeling and Analysis*, Addison-Wesley (1978).

付　録

提案した先読み／まとめ書きアルゴリズムを適用したとき，仮定2を前提とすると，アクセス比率が最大でないMTは，連続スケジュール初期時点で，(2)式が成立するということを証明する．仮定2が成立したとき，先読み／まとめ書きサイクル時間，および，各MTのサイクル内先読み／まとめ書き時間は，各先読み／まとめ書きサイクルによらず一定となる．したがって，各MTの連続スケジュール初期時点も先読み／まとめ書きサイクル時間ごとに発生することになる．このとき，先読み／まとめ書きサイクル時間は(3)式で表される．一方，アクセス比率が最大でないMT, MTiの(2)式が成立してから次に(2)式が成立するまでの時間T_i'は次の式で表すことができる．

$$T_i'=(s+r)/(1-\alpha_i) \qquad (11)$$

(11)式は，明らかに先読み／まとめ書きサイクル時間より小さい値となる．したがって，連続スケジュール処理のある過渡的な時点において(2)式が成立したとしても，何回か先読み／まとめ書きサイクルを繰り返すうちに，連続スケジュール初期時点において(2)式が成立することになる．

(昭和63年11月10日受付)
(平成元年9月12日採録)

山本　彰（正会員）

昭和29年生．昭和52年京都大学工学部情報工学科卒業．昭和54年同大学院修士課程修了．同年(株)日立製作所入社．以来，同社システム開発研究所において，計算機システムの性能解析手法，計算機メモリ階層システム，ファイルシステム制御の研究に従事．第30回情報処理学会全国大会学術奨励賞受賞．ソフトウェア科学会，日本OR学会各会員．

坪井　俊明（正会員）

昭和33年生．昭和56年鳥取大学工学部卒業．昭和58年同大学院修士課程修了．同年日立マイクロコンピュータエンジニアリング(株)に入社．データベース管理システム，入出力装置の制御に関する研究に従事．

北嶋　弘行（正会員）

昭和44年東京大学工学部機械工学科卒業．昭和46年東京大学大学院工学系修士課程修了．同年(株)日立製作所入所．昭和48年以来，同所システム開発研究所にて研究．現在，同所主任研究員．昭和58～59年にUCLA修士課程修了．計算機システム・アーキテクチャ，特に，ファイルシステム，データベースマシンの研究，分散システムの計画・評価技法の研究に従事．