

# テープアーカイブシステムにおける ホットレプリケーションの性能評価

根本 利弘 喜連川 優

東京大学生産技術研究所

〒106-8558 東京都港区六本木7-22-1  
PHONE:03-3402-6231 FAX:03-3423-2834  
{nemoto,kitsure}@tkl.iis.u-tokyo.ac.jp

あらまし 高アクセス頻度データの複製(レプリカ)をテープ上の予め確保しておいた領域に作成するホットレプリケーションと呼ばれる手法について述べる。ホットレプリケーションでは、高アクセス頻度データのクラスタリングによるシーク時間の短縮、および複製によるアクセシビリティの向上により応答時間が短縮される。本稿では、まずホットレプリケーションについて、そのアルゴリズムを説明し、その後、シミュレーションによりその基本性能を示す。加えて、当研究所で公開している衛星画像データベースのアクセスログを用いたシミュレーション結果を示し、実システムにおける有効性を示す。

キーワード ホットレプリケーション, テープアーカイブ, 性能評価

## Performance Evaluation of the Hot Replication on a Tape Archiving System

Toshihiro NEMOTO Masaru KITSUREGAWA

Institute of Industrial Science, University of Tokyo

7-22-1, Roppongi, Minato-ku, Tokyo, JAPAN, 106-8558  
PHONE:03-3402-6231 FAX:03-3423-2834  
{nemoto,kitsure}@tkl.iis.u-tokyo.ac.jp

**Abstract** In this paper, we explain the hot replication method. In hot replication, frequently accessed (hot) files are replicated on the reserved area of the tape. Reduction of seek length by clustering of hot files and improvement of accessibility to hot data by replication can reduce the response time. After the description of the algorithm of hot replication, its basic performance is shown through simulations. In addition, we clarify its effectiveness for the real applications through the simulation using the access logs of our satellite image database system.

**Key words** hot replication, tape archiver, performance evaluation

## 1 はじめに

近年、マルチメディアデータベースや地球環境情報システムなどの膨大なデータを扱うシステムが盛んに構築されるようになり [1], 大規模 3 次記憶システムの需要が高まってきた。大規模 3 次記憶システムとしては、容量、コストの点で優れた磁気テープが用いられることが多いが、一般にテープ上のデータへアクセスする場合にはシークが必要であり、そのために長時間を要することも多い。データを記録する際に、アクセス頻度の高いデータをシークが短くなる位置に配置することで応答時間を短縮することが可能であるが [2], データが生成された時点においてそのデータのアクセス頻度を予測することは一般に困難である。また、一度テープ上へ記録されたデータの再配置のためには時間的、空間的に大きなコストを要するため、各データのアクセス頻度が判明した後に再配列を行うことは必ずしも得策とは言えない。

本論文では、テープ上に予め確保した領域へ高アクセス頻度データを複製し、クラスタリングをすることで応答性能の向上を図るホットレプリケーションと呼ばれる手法について述べる。ホットレプリケーションでは、複製されたデータへのアクセスが不利にならないようにテープ途中でロード/イジェクト可能なテープドライブ装置を用いる。ホットレプリケーションでは、データの多重化によるアクセシビリティの向上、および高アクセス頻度データのクラスタリングによるシーク時間の短縮により応答時間の短縮を図る。

以降 2 節においてレプリカの作成手法について説明を行い、3 節においてホットレプリケーションの基本性能について述べる。さらに、4 節では衛星画像データベースに対するアクセス履歴を用いた性能評価結果を示し、実アプリケーションに対するホットレプリケーションの有効性を明かにする。

## 2 ホットレプリケーション

### 2.1 ホットレプリケーション手法

ホットレプリケーションとは、予めテープ上に確保しておいた領域へ高アクセス頻度データ(ホットデータ)のレプリカを作成してクラスタリングを行い、応答時間の短縮を図る手法である。データがレプリカを持たない場合、要求されたデータが存在するテープが使用されていたり、あるいはそのテープをアクセスするために必要なドライブ装置、テープハンドラロボットなどの資源が使用されている場合にはリクエストはブロックされることとなるが、レプリカを作成することで、オリジナルデータにアクセスできない場合においてもレプリカをアクセスすることができるために応答時間が短縮される。また、高アクセス頻度データのレプリカがクラスタリングされることにより、高アクセス頻度データが連続してアクセスされる場合のシーク長が短縮され、応答時間が短縮さ

れることとなる。すなわち、ホットレプリケーションは、レプリカの作成によるアクセシビリティの向上、および高アクセス頻度データのクラスタリングによるシーク長の短縮により応答時間の短縮を図るものである。

ホットレプリケーションでは少数の高アクセス頻度データのみを複製の対象とする。全データのレプリカを作成すればデータのアクセシビリティは更に向上すると考えられるが、全データの複製のためには、オリジナルデータの読み込み、書き込みに要する時間的なコスト、およびレプリカを格納するための空間的なコストを要することとなり、また、アクセス頻度の低いデータのレプリカを作成してもそのデータがアクセスされることは少ないため、効果は小さいことが予想される。さらに、テープドライブにおいてはテープのイジェクト/ロード、シークに要する時間的なコストが大きいため、レプリカを作成するには通常のリクエストへの影響を最小限に抑えるよう以下の方針を採用する。

1. キャッシュディスク上に存在する高アクセス頻度データを複製する。
2. 使用されていないテープドライブ装置を用い、そのドライブ内に存在するテープにレプリカを作成する。

キャッシュディスク上の高アクセス頻度データのみをレプリカ作成の対象とすることにより、新たにレプリカ作成対象となるデータを読み込む必要がなくなる。また、既にテープドライブ内に存在するテープを用いることで、新たにテープをロードする必要がなくなる。これらの方針を採用し、レプリカ作成に要するコストを最小限に抑える。上記方針によってレプリカを作成することができない場合には、新たなデータの読み込みやテープのマウントは行わず、レプリカ作成そのものを行わない。

ホットレプリケーションでは、各データに対してある程度のアクセスが行われ、各データのアクセス頻度が判明した時点でレプリカを作成することとなるが、多くの商用テープドライブ装置では、テープ上の既存のデータを破壊せずに新たなデータを書き込むためには、新たに書き込むデータを既存データの後に追記する以外の方法がない。従って、このようなテープドライブ装置においてはレプリカのための領域はテープの終端部に確保せざるをえないが、テープのイジェクト時に先頭まで巻戻さねばならず、テープ後方に存在するデータへのアクセス時には長いシークが必要となる。すなわち、テープ終端部に確保された領域の高アクセス頻度データのレプリカへのアクセスは時間的なコストの面で不利となってしまう。ホットレプリケーションではこの問題を回避し、レプリカへのアクセスが不利にならないようにするため、テープを巻戻すことなく、テープの途中でロード/イジェクトすることが可能なテープドライブ装置を用い

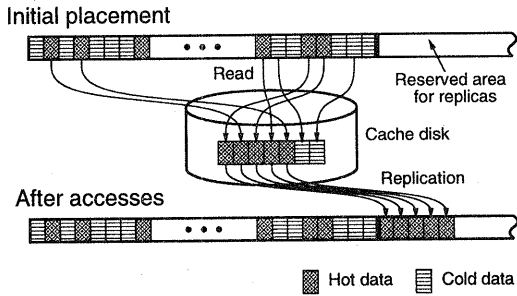


図1: テープ途中でロード/イジェクト可能なテープドライブ装置によるホットレプリケーション

る。データ記録時には、テープ全てにオリジナルデータは記録せず、テープ終端部を高アクセス頻度データのレプリカのための領域として残す。この領域に、高アクセス頻度データのレプリカを作成し、クラスタリングする(図1)。テープ途中でロード/イジェクト可能なドライブ装置を用いるため、テープの入れ替えを繰り返しても、高アクセス頻度データがクラスタリングされているテープ終端部にヘッドが位置することが多く、高アクセス頻度データへのアクセスのシーク長が短縮され、応答性能が向上することとなる。

## 2.2 テープ途中でロード/イジェクト可能なテープドライブ装置

多くの商用テープドライブ装置では、テープメディアの先頭にディレクトリ情報やメディアの状態に関する情報などを記録しており、テープロード時にはこの情報を読み込むようになっている。このため、イジェクト時にはテープを巻戻すようになっている。テープアーカイブシステムでは多数のテープの交換を繰り返すため、ロード/イジェクト時のシークに多くの時間が費やされる。このロード/イジェクト時のシークを削減するために、テープ先頭のみではなく、テープ上の複数の位置あるいは任意の位置にもディレクトリ情報やメディア情報などを記録できるようにしたり、あるいはテープカセット内に不揮発性メモリ素子を設け、これにディレクトリ情報やメディア情報などを記録することで、テープを巻戻すことなく、ヘッドがテープの途中にあるときでもロード/イジェクトが可能であるドライブが開発され、商用化されている。例えば、AMPEX社製DST312やSONY社製GY-2120では、ヘッドがテープの先頭以外の位置に存在するときでもロード/イジェクトが可能である。テープを巻戻すことなくロード/イジェクト可能なテープドライブ装置はホットレプリケーションのためばかりでなく、一般的なアーカイブシステムでの利用においても性能面で大きな利点を持つため、今後の普及が期待される。

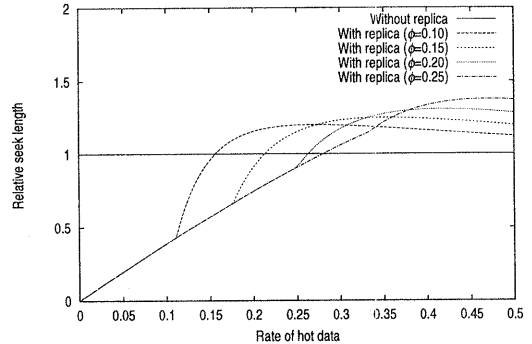


図2: 高アクセス頻度データをテープ終端部に複製した場合のシーク長短縮効果

## 2.3 ホットレプリケーションによるシーク長の短縮

本節では、全体の $p$ の割合を占める高アクセス頻度データと $1-p$ の低アクセス頻度データの2種類のデータで構成され、それらがランダムに配置されている場合について、解析的に求めたホットレプリケーションによるシーク長の短縮の効果を示す。レプリカが存在するデータに関しては必ずレプリカをアクセスするものとしている。図2は、レプリカ用の領域を除くオリジナルデータ用の領域上にランダムにデータを配置した場合の平均シーク長を1としたときの、レプリカが存在する場合の相対シーク長を示している。 $\phi$ は全テープ長に対するレプリカ用領域の割合である。オリジナル領域が全体に占める割合 $1-\phi$ は、単位データあたりに要するコストと反比例関係にあるため、図2は、コストで正規化したシーク長の短縮効果と言うこともできる。

図2によると、レプリカ用の領域が小さく、その領域以上に高アクセス頻度データが存在する場合、すなわち全高アクセス頻度データのレプリカを作成できない場合には、平均シーク長は急激に長くなる。クラスタリングの効果を保つためには、レプリカの量よりもリザーブ用の領域を大きくとる必要がある。一方で、アクセス頻度が70/30則よりも緩やかになると、たとえレプリカ用領域が十分でも、ホットデータのクラスタリングの効果は薄れ、低アクセス頻度データをアクセスするためにオリジナルデータ領域へシークする回数が増えるために平均シーク長は延びる。すなわち、シーク短縮の観点のみから言えば、70/30則よりも緩い分布をなすデータに対しては効果がなく、リザーブ領域をテープ全体の20~25%以上としても意味はない。

表 1: シミュレーションパラメータ

エレメントアーカイバ	
全エレメントアーカイバ数	16 台
最大テープ数	200 本/台
テープドライブ数	2 台/台
テープドライブ	
ロード時間	35 秒
シーク速度	25MB/秒
リード/ライト速度	0.5MB/秒
イジェクト時間	20 秒
テープハンドラロボット	
移動時間 (テープの操作なし)	2 秒
移動時間 (テープの操作あり)	14 秒
テープマイグレーション装置	
ワゴンの移動時間	9 秒

### 3 基本性能評価

#### 3.1 シミュレーション条件

本節では、ホットレプリケーションの基本性能をシミュレーションにより評価する。16 台のエレメントアーカイバにより構成されるスケラブルテープアーカイバにおいて、使用されている Exabyte 8mm テープドライブ装置がテープ上の任意の位置でテープをロード/イジェクト可能であると仮定し、シミュレーションを行う。ホットデクラスタリング [3] を導入した場合、しない場合について、それぞれ高アクセス頻度データのレプリカがテープ終端部に存在する場合、しない場合の平均応答時間を測定する。応答時間とはリクエストが発行されてからデータの読み込みが完了するまでの時間である。スケラブルテープアーカイバのパラメータを表 1 に示す。各テープの容量は 7GB とし、先頭より 5.5GB までをオリジナルデータ用、終端部の 1.5GB の領域を高アクセス頻度データのレプリカ用領域とする。リクエスト到着間隔は負の指数分布に従う。リクエストのスケジューリングに関しては、高アクセス頻度データのクラスタリングによるシーク長の短縮の効果とレプリカによるアクセシビリティの向上による効果の両者が得られるよう、次のスケジューリングを採用する。まず、レプリカが存在するデータはクラスタリングされたレプリカのみが存在すると仮定し、リクエストキュー内のサービス可能なものの中で最も先に発行されたリクエストを優先し、サービスを行うテープを選択する。ここで、サービス可能なリクエストが存在しない場合には、改めてレプリカ、オリジナルデータの両者がともに存在するものとし、サービスを行うテープを選択する。その後、そのテープ上のデータに対する全てリクエストをテープの先頭方向から順にまとめてサー

ビスする。

#### 3.2 シーク時間の短縮による効果

本シミュレーションでは、データのアクセス分布は、データ全体が  $z = 1.5$  の Zipf 分布<sup>1</sup>に従うとし、それらをテープ上のオリジナル領域にランダムに配置する。このとき、データ全体の約 10% が全リクエストの約 90% を受けることとなるため、アクセス頻度の高い順に 10% のデータをホットデータとみなし、これらのレプリカをテープ終端部の領域にランダムに配置する。1 つのファイルのサイズは 10MB としている。各データをランダムに配置しているため、各テープのヒートはほぼ均一となり、従って、複数のリクエストが 1 本のテープに集中する可能性は低く、レプリカによるアクセシビリティの向上による性能向上は期待できず、シーク長の短縮による応答性能の向上の効果が明確になる。

図 3 は初期状態から 50000 アクセスまでの平均応答時間である。図 3 には、高アクセス頻度データのレプリカが存在しない場合、テープ終端部に高アクセス頻度データのレプリカが存在する場合、およびテープ終端部に高アクセス頻度データのレプリカが存在し、さらに高アクセス頻度データにのみに対してアクセスリクエストが発行される場合の、3 つの場合の平均応答時間を示している。高アクセス頻度データにのみに対してリクエストが発行される場合は、ホットレプリケーションによる応答性能向上の限界値と考えられる。図 3 より、ホットでクラスタリングの有無に関わらず、テープ終端部に高アクセス頻度データのレプリカを作成することで、平均応答時間が大幅に短縮されていることがわかる。データをランダムに配置しているため、各エレメントアーカイバのヒートが平衡に達しており、ホットデクラスタリングはほとんど応答性能の向上に影響を与えない。一方、ホットデクラスタリングは、既に各エレメントアーカイバのヒートが平衡に達している状態においても、テープ内のホットデータがクラスタリングされているためにシーク時間が短縮され、応答性能がさらに向上する。

#### 3.3 レプリカによるアクセシビリティの向上による効果

図 4 は各テープに対するアクセス頻度が  $z = 1.2$  の Zipf 分布に従い、さらに各テープ上のデータそれぞれに対するアクセス頻度もまた、 $z = 1.2$  の Zipf 分布に従うとした場合の初期状態から 50000 アクセスまでの応答時間である。また、図 5 は、データ全体としては図 4 のデータと同じ分布を持つが、それらを全テープ上にランダムに

<sup>1</sup> 全要素数を  $N$  としたとき、 $i$  番目の要素の確率  $p_i$  が

$$p_i = \frac{1}{N} \cdot \frac{1}{i^z} = \frac{1}{N \sum_{n=1}^{\infty} \frac{1}{n^z}}$$

により表される分布。

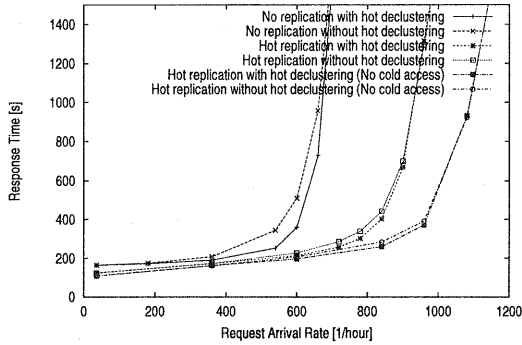


図 3: ホットレプリケーションによる応答時間

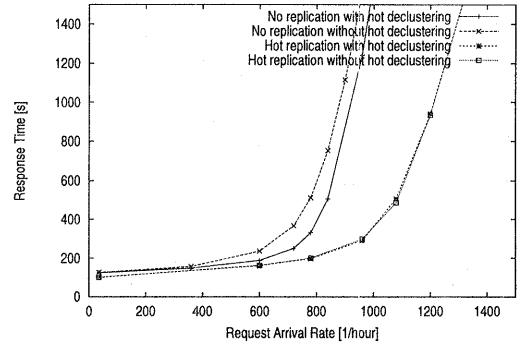


図 5: ホットレプリケーションによる応答時間 (テープ間のアクセス頻度の偏りなし)

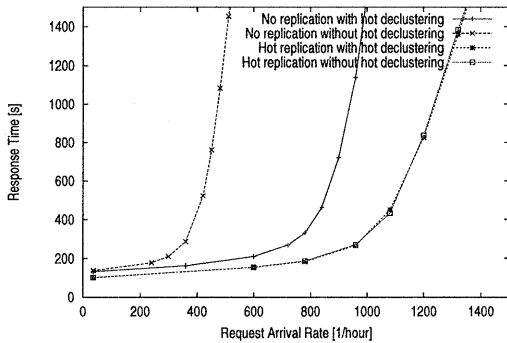


図 4: ホットレプリケーションによる応答時間 (テープ間のアクセス頻度の偏りあり)

配置した場合の初期状態から 50000 アクセスまでの応答時間である。図 4, 図 5 のいずれの場合も, エレメントアーカイバ内のテープ配置はランダムである。データ全体の約 10% が全リクエストの約 90% を受けることとなるため, アクセス頻度の高い順に 10% のデータをホットデータとし, これらのレプリカをテープ終端部の領域にランダムに配置している。両者の違いはテープ間のアクセス頻度の偏りの有無であり, テープ間にアクセス頻度の偏りがある場合には少数のテープにアクセスが集中することとなるため, これらの結果を比較することでレプリカによるアクセシビリティの向上による効果を見ることが出来る。

テープ間のアクセス頻度の偏りがある場合には, テープ間のアクセス頻度の偏りが無い場合に比べ, ホットデクラスタリング, ホットマイグレーションのいずれも用いないと大幅に応答性能は劣化する。これは, 少数のテープにリクエストが集中してしまうため, あるデータが記録されているテープが別のリクエストによって使用され

ているためにブロックされてしまったり, あるいはそのテープが存在するエレメントアーカイバ内のテープドライブが常に使用中となってしまうためにサービスを行うことができず, テープドライブを効率的に使用することができないためである。ホットデクラスタリングは, このような場合, 空いているテープドライブが存在するエレメントアーカイバへテープをマイグレートし, リクエストをサービスすることで応答時間の劣化を押さえるが, ホットレプリケーションでは, オリジナルデータとレプリカの 2 つが存在するため, どちらか一方の存在するテープが使用されていたり, そのテープが存在するエレメントアーカイバのテープドライブが使用されていても, 他の異なるエレメントアーカイバ内に存在するもう一方のデータへアクセスすることによりテープドライブを効率的に使用し, 応答性能の劣化を防ぐ。加えて, ホットレプリケーションでは, シーク時間の短縮の効果もあり, ホットデクラスタリングに対してより大幅な応答性能の向上が得られる。

#### 4 衛星画像データベースのアクセス履歴を用いた性能評価

本節では, 東京大学生産技術研究所において World Wide Web (WWW), gopher, ftp により公開している衛星画像データベースのアクセス履歴を用い, ホットレプリケーションの性能を評価する。

##### 4.1 アクセス履歴

図 6 は, 1996 年 4 月から 1998 年 10 月半ばまでの衛星画像データベースシステム上のクイックルック画像に対するアクセスの分布である。リクエスト数は ftp によるアクセス約 49000 件, gopher によるアクセス約 215000 件, WWW によるアクセス約 197000 件の計 461000 件である。横軸は 1996 年 4 月 1 日からの経過日数, 縦軸はデータ番号を表す。グラフの各点は当該データに対し当該日にア

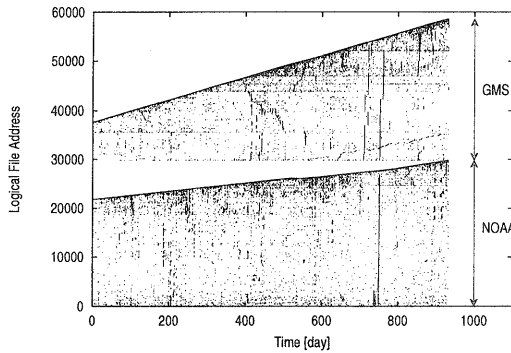


図 6: リクエスト分布

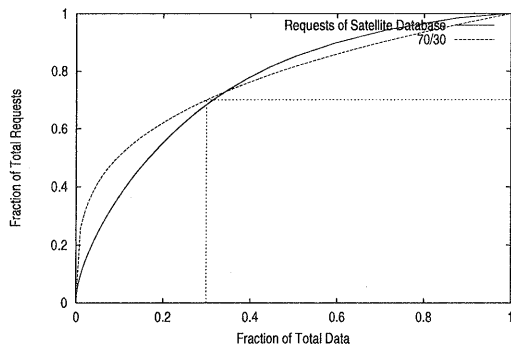


図 7: アクセスローカリティ

アクセスがあったことを示している。データ番号はデータ全体を NOAA, GMS に分けてそれぞれ観測日順に番号を付けており、0~29799 が NOAA データ、29800~58636 が GMS データである。図 6 において、NOAA, GMS とともに最上部に斜めの線状の分布がみられ、最新画像にアクセスが集中していることが示されている。また、縦方向の線状の分布が見られ、短期間に多くのデータがまとめてアクセスされたことがわかる。これらは、特定の利用者がある一定期間の画像データを一括して転送したことによる。

図 7 は、衛星画像データベース内の全データに対するリクエストに関して、アクセス頻度上位のデータに対するリクエストが全リクエストに対して占める割合を表したグラフである。比較のために論文 [4] において示された 70/30 則に従う曲線<sup>2</sup>も示している。衛星画像データベースに対するリクエストは全データに関しては全体の約 30% のデータが 70% のリクエストを受けているが、論

<sup>2</sup>Fraction of Total Requests = Fraction of Total Data  $\frac{\log 8}{\log 3}$  により表される分布。

文 [4] による分布とはやや異なり、アクセス頻度が極端に高いデータはなく、より緩やかな分布をしていることがわかる。

## 4.2 シミュレーション条件

シミュレーションには、28000 の新たに受信されたデータの書き込みリクエストを加えた 489000 のリクエストを用いる。読み込みリクエストはクイックルック画像に対するものであるが、シミュレーションではこれらのリクエストが対応する衛星原画像へのリクエストであると仮定する。クイックルック画像のアクセス分布と原画像へのアクセス分布とは必ずしも一致しないが、原画像とクイックルック画像は一対一に対応しており、また、最新画像に対するアクセスが多い、特定ユーザが短期間に一括してデータにアクセスしている、などの特徴は原画像に対するアクセスにも共通すると考えられる。アーカイブシステムはなるべく実システムに即した環境を想定するが、実システムでは各テープの全領域にオリジナルデータが記録されており、レプリカ用の領域は確保されていないため、シミュレーションにおいてはデータを再配置する。各テープの容量は 7GB (非圧縮時) であるとし、そのうち先頭より 5.5GB の領域をオリジナルデータ用の領域として、データ番号順に配置する。データは両者ともテープドライブ装置に備え付けられている圧縮機能を用いて圧縮されているとする。実システムにおいても、大部分のデータはテープドライブによって圧縮されているが、個々のデータに対する圧縮率を得ることは困難であるため、非圧縮時のテープ容量と圧縮機能を用いて実際に記録されたデータ量より求めた各データの圧縮率の平均値に基づき、NOAA データは一律に 67%、GMS データは一律 20% に圧縮されているものとする。例えば GMS データは、テープ上では見かけ上約 20MB のデータとなり、読み出し時間、書き込み時間とも 20% に短縮される。テープは NOAA データ用 328 本、GMS データ用 108 本の計 436 本で構成される。スケラブルテープアーカイブは 4 台のエレメントアーカイバ NTH-200B で構成され、初期状態では各エレメントアーカイバに均等に 109 本配置する。テープアーカイバ上のデータをキャッシュするためのディスクは 10MB/s のデータ転送速度をもち、LRU によりデータを管理する。ディスク容量は、十分な容量である 40GB、および平均的なファイルサイズの約 3 倍と極めて小さい容量である 300MB とする。アクセススケジューリングとしては、3 節と同様のスケジューリングを採用する。その他のパラメータは 3 節の表 1 に従う。本シミュレーションでは、シミュレーション開始時にはレプリカは存在せず、シミュレーション開始後に 10 度以上アクセスされたデータを高アクセス頻度データとみなし、そのレプリカをリザーブ領域に動的に作成する。レプリカ作成の対象となるテープは、レプリカ用の領域を除く

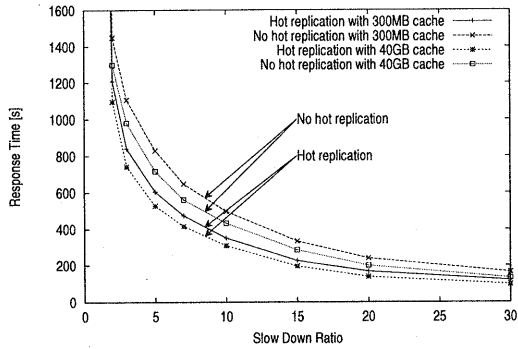


図 8: ホットレプリケーションによる平均応答時間 (ホットデクラスタリングあり)

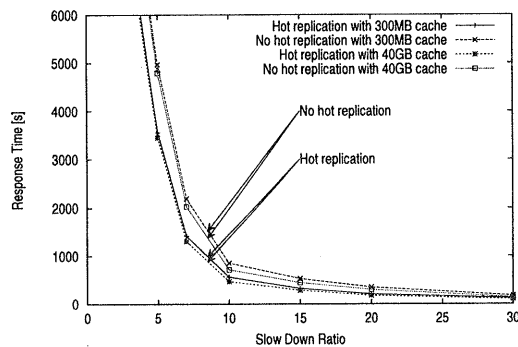


図 9: ホットレプリケーションによる平均応答時間 (ホットデクラスタリングなし)

オリジナル用の領域が全て記録されているテープである。シミュレーション開始後にデータが記録されるテープに関しては、リザーブ領域以外の全ての領域にオリジナルデータが記録された時点でレプリカが作成される対象のテープとなる。

### 4.3 シミュレーション結果

図 8 は、ホットデクラスタリングを用いたときの初期状態から 450000 アクセスまでの平均応答時間である。キャッシュディスクサイズが 300MB、および 40GB の場合それぞれに対し、ホットレプリケーションを用いた場合、およびホットレプリケーションを用いない場合についての結果を示している。横軸はリクエスト到着率を変化させるための各リクエスト間隔を延ばす際の倍率を表すリクエスト遅延率である。図 9 はホットデクラスタリングを用いない場合の初期状態から 450000 アクセスまでの平均応答時間である。図 8 と同様にキャッシュディスクサイズが 300MB、および 40GB の場合それぞれに対し、ホットレ

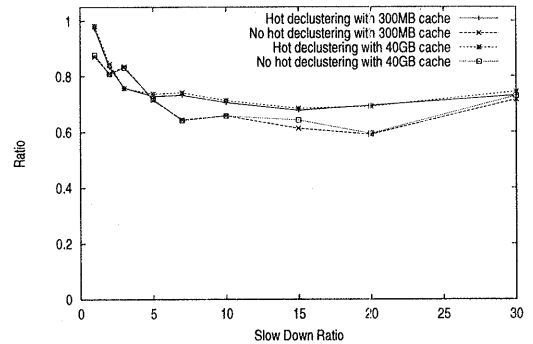


図 10: ホットレプリケーションによる平均応答時間の短縮率

プリケーションを用いた場合、およびホットレプリケーションを用いない場合についての結果を示している。また、図 10 は各シミュレーション条件におけるホットレプリケーションを用いない場合の応答時間を 1 としたときのホットレプリケーションを用いた場合の相対平均応答時間を表している。

ホットデクラスタリングの有無、ディスクによるキャッシュのサイズに関わらず、高アクセス頻度データのレプリカを作成することにより平均応答時間が短縮されることがわかる。ホットデクラスタリングを用いた場合、用いない場合とも、40GB のディスクによるキャッシュ以上にホットレプリケーションは平均応答時間を短縮している。また、ホットレプリケーションはキャッシュサイズによらず平均応答時間を短縮し、キャッシュディスクが小さくても十分にレプリカが作成され、平均応答時間が短縮されることがわかる。

一方、図 10 より、リクエスト遅延率が大きい場合には、ホットデクラスタリングを用いていないときは、用いたとき以上に平均応答時間を短縮する。これは、テープドライブが全て使用されているエレメントアーカイバ内のテープに新たにリクエストが発行された場合に、ホットデクラスタリングではそのテープの移動を行い、ホットレプリケーションでは別のエレメントアーカイバ内のレプリカを参照し、リクエストがブロックすることを避けるなど、ホットデクラスタリングとホットレプリケーションが効果を示す状況が一部重複しているためである。また、リクエスト遅延率が小さいときにはアクセスリクエストへの対応のためにテープドライブ装置の使用率が上昇し、レプリカの作成が効果的に行われなくなるためにホットレプリケーションの効果が低下する。ホットデクラスタリングはテープドライブ装置の使用を効率的にするため、リクエスト遅延率が小さいときにはホットデクラスタリングを用いた場合の方がホットレプリケーショ

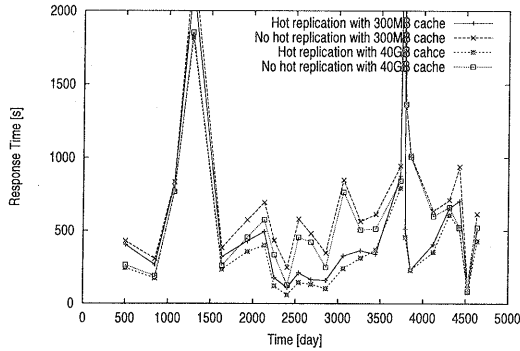


図 11: ホットレプリケーションによる平均応答時間の変化(ホットデクラスタリングあり)

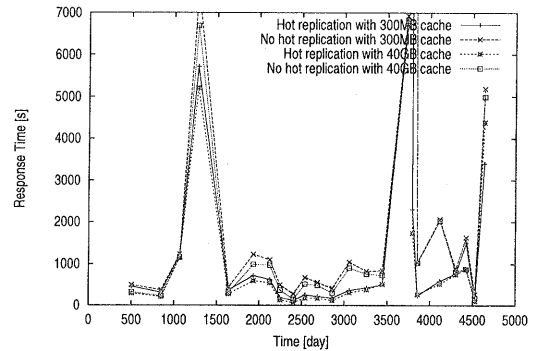


図 12: ホットレプリケーションによる平均応答時間の変化(ホットデクラスタリングなし)

ンによる平均応答時間の短縮率は向上する。

図 11, 図 12 はそれぞれホットデクラスタリングを用いた場合, およびホットデクラスタリングを用いない場合のリクエスト遅延率が 5 のときの 20000 アクセス毎の平均応答時間を表している。1600 日以前では, まだ十分にレプリカが作成されておらず, キャッシュがホットレプリケーション以上に効果を示しているが, 1600 日以降はホットレプリケーションはキャッシュ以上に応答時間を短縮している。また, 図 6 において垂直方向の線状の分布が見られ, 過去の連続する多数へのデータへアクセスが行われる期間においてもホットレプリケーションは効果を示している。これは, ホットレプリケーションを用いていない場合, 1 本のテープ上の連続するデータをシケンシャルに読み込まねばならず, 長時間を要し, また, アクセスされるデータは過去のデータであり, その直前にアクセスされていることはほとんどないためキャッシュは有効ではないが, ホットレプリケーションでは, このような場合, 異なるテープ上の高アクセス頻度データのレプリカにアクセスすることが可能であるためである。

## 5 おわりに

本節では, 高アクセス頻度データのレプリカを作成し, それらをテープ最後部にクラスタリングすることにより, データアクセス時のシーク長を削減するホットレプリケーションの提案を行い, シミュレーションによりその有効性を示した。さらに, 生産技術研究所において WWW, gopher, ftp により公開している衛星データのクイックルック画像に対するアクセス履歴を用い, これを衛星原画像へのアクセスであると仮定してシミュレーションを行うことで, 実システムに対しても, ホットレプリケーションは応答時間の短縮に有効であることを示した。また, キャッシュ, テープマイグレーションをともに導入した場合においても, さらにホットレプリケーションを導入

することで, 応答時間の短縮可能であることを示した。

## 参考文献

- [1] B. Kobler, J. Berbert, P. Caulk, and P. C. Hariharan. "Architecture and desing of storage and data management for the NASA earth observing system data and information system (EOSDIS)". In *Proceedings of Fourteenth IEEE Symposium on Mass Storage Systems*, pp. 65–76, Monterey, California, September 1995.
- [2] S. Christodoulakis, P. Triantafillou, and F. A. Zioga. "Principles of optimally placing data in tertiary storage libraries". In *Proceedings of the Twenty-third Very Large Database Conference*, pp. 236–245, Athenes, Greece, August 1997.
- [3] 根本利弘, 喜連川優. "スケーラブルテープアーカイバにおけるテープマイグレーションを用いた負荷分散手法とその性能評価". 電子情報通信学会論文誌, Vol. J82-D-I, No. 1, pp. 53–69, January 1999.
- [4] G. Copeland, W. Alexander, E. Boughter, and T. Keller. "Data placement in Bubba". In *Proceedings of the 1988 ACM SIGMOD International Conference on Management of Data*, pp. 99–109, Chicago, Illinois, June 1988.
- [5] T. Nemoto and M. Kitsuregawa. "Scalable tape archiver for satellite image database and its performance analysis with access logs – hot declustering and hot replication –". In *Proceedings of 16th IEEE Symposium on Mass Storage Systems in cooperation with the 7th NASA GSFC Conference on Mass Storage Systems and Technologies*, pp. 59–71, San Diego, California, March 1999.