

映像情報サーバにおける連続データ転送方式の評価

1V-9

高倉 健 櫻井 紀彦 石川 篤

NTT 情報通信研究所

1 はじめに

複数端末からの多重アクセスを可能とする映像情報サーバの記憶系に磁気ディスク装置を用いる場合、ディスクのアクセスオーバーヘッドが基幹バス上の転送空き時間として現れるためシステムの実効転送速度に大きく影響する。その解決方法として以前に PreSS 転送 (Pre-Seek & Successive Data Transfer) というバスの使用効率を高める転送方式を提案したが [1]、本稿ではその性能評価について報告する。PreSS 転送の試作システムとその制御方式について述べ、シミュレーションを用いて性能解析した。クライアントからのサーバへのアクセスは読み出し専用とし、通常の動画再生に小容量ランダムアクセスが混在する場合も検討した。PreSS 転送によりバス使用効率が上昇し、システムの実効転送速度を向上させることができた。

2 PreSS 転送方式

サーバシステムを構築する上で、基幹バスの性能は避けることのできないボトルネックである。PreSS 転送方式の狙いは、転送性能低下の原因となるディスクのシーク・回転待ちの間にも、同一バスの別のディスクで可能な処理は予め進めておき、バスの限界性能を引き出そうというものである。

今回の PreSS 転送評価に用いる映像情報サーバの試作システムでは、I/O バスに SCSI バスを、システムバスに VME バスを使用している。SCSI バスの性能がボトルネックにならないように Disk Controller には 4 本の SCSI

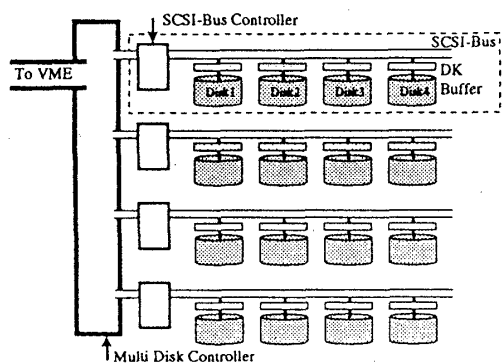


図 1: Multi Disk Controller

Evaluation of PreSS-Transfer for Multi-media Server.
 Takeshi Takakura, Norihiko Sakurai, and Atsushi Ishikawa.
 NTT Information and Communication Systems Laboratories
 1-2356 Take, Yokosuka, Kanagawa 238-03, Japan

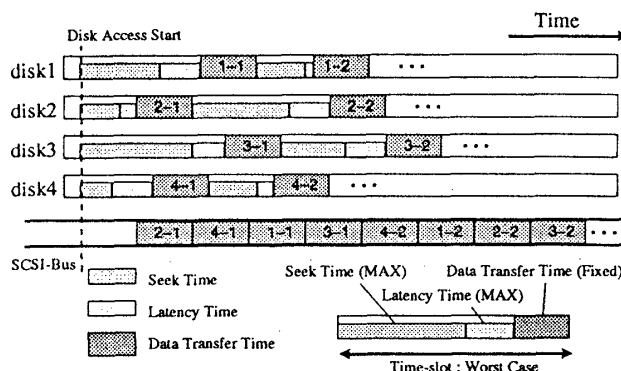


図 2: PreSS 転送の動作例

バスが接続されており、これを Multi Disk Controller と呼んでいる。図 1 の点線部が PreSS 転送を行うユニットであり、現在はディスク 4 台を用いた構成である。

PreSS 転送ユニットでのディスク動作と SCSI バス上のデータ転送の様子を図 2 に示す。SCSI バスコントローラ間では、RAID3,4 と類似した制御、すなわち 4 本のバスから得られるデータをバイト単位やセクタ単位でマルチプレクスすることが可能である。

システム全体の制御を見ると、main CPU は可変長のサイクル時間 T_{cyc} 内に複数クライアントから受け取ったリクエストをグループとして扱い、この中でアクセス時間が最小になるように並べ換えを行い、コマンドリストとして発行する。コマンドリストを受け取ったユニットでは、処理ディスク毎に Ordered Que Tag を割り付けて SCSI コマンドとして発行する。ディスク群では転送準備の整ったものから順次 SCSI バスをつかみデータ転送を始め、全てのコマンドの完了を確認して次のコマンドリストを要求する。という流れである。

3 シミュレーションモデル

性能評価のためのシミュレーションモデルを考える。4 台のディスクからなる PreSS 転送ユニットを 4 セット接続して、1 つのシステムを構成するものとした。

アクセスタイムは $T_{acc} = T_{cmd} + T_{seek} + T_{rot}$ と表される。 T_{cmd} はコマンド処理時間の実測値の平均から約 2.5ms と見積もった。 T_{seek} はディスクのシーク時間で平均 17.5ms, 最大 27.5ms の関数で、 T_{rot} はディスクの回転待ち時間で平均 5ms, 最大 10ms の関数でそれぞれ近似した。ディスクの転送速度とバスの転送速度は $V_{dk} = V_{bus} = 40Mbps$ として評価している。また映像表示系のデータの消費速度は $V_{play} = 4Mbps$ とする。

バス上でデータ転送時間の占める割合について考える。多重度を M_{ac} 、クライアント i の要求するデータ長を L_i とすると、バス使用効率 P_{DT} は、

$$P_{DT} = \frac{\sum_{i=1}^{M_{ac}} (L_i/V_{bus})}{\sum_{i=1}^{M_{ac}} (T_{acc}^i + L_i/V_{dk})}$$

と表せる。またシステムの実効転送速度は P_{DT} を用いると $V_{eff} = P_{DT}V_{bus}$ と表される。

4 結果と解析

4.1 転送性能

インタラクティブ性の向上には単位アクセス当たりの転送データ量を小さくする必要がある [1]。単位アクセスデータ長 L を 250, 150, 100KB と変化させてシミュレーションを行った。映像再生時間では各々 500, 300, 200ms に対応し、この値がそのまま応答性能に反映される。

映像データに対するランダムアクセスを想定し、1 サイクル時間とその間の処理所要時間との差 ΔT_{cyc} をシミュレーションで調べた。システムで処理されるコマンドは 4 つの PreSS 転送ユニットに均等に分散されると仮定している。各クライアントの持つバッファ容量が十分で、 ΔT_{cyc} の平均値が 0 以下であれば、コマンド処理所要時間がサイクル時間内に完了しないことがあっても次以降のサイクルで吸収されることが示されているので [2]、映像情報サーバとして途切れることなく映像が再生できる。

このシミュレーションの結果からサイクル時間を 500ms, 300ms としたときに、多重度 36 (実効転送速度 144Mbps) が実現することがわかった。また 200ms の場合は、36 クライアントからのアクセス処理の平均所要時間がサイクル時間を越えているので不可能であり、実現できる多重度は 32 (実効転送速度 128Mbps) となる。1 ディスクアクセス処理には平均 25ms かかるが、先に述べたスケジューリング制御 [2] により 1 サイクルでの平均アクセス時間は 15ms 程度に抑えられた。

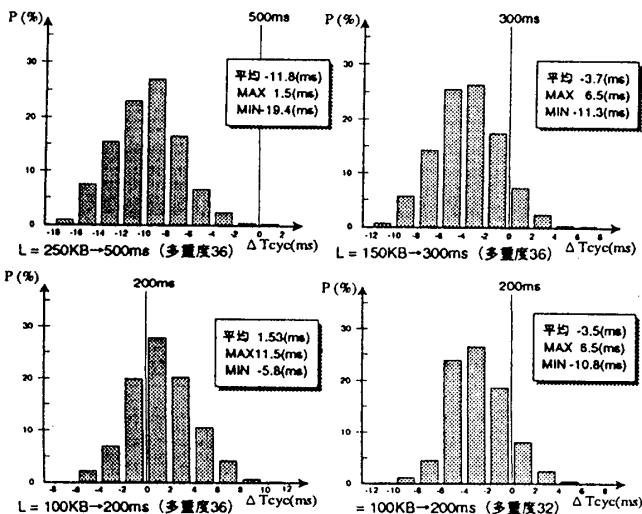


図 3: シミュレーション結果

4.2 小容量ランダムアクセスの許容性

映像情報サーバでは、動画像情報の提供と同時にプログラムやメニュー等の静止画やテキストファイルへのアクセスも頻繁に行われる。従来このような小容量ランダムアクセスの存在は転送性能を低下させていた。これに対し PreSS 転送では、先行処理された転送データがディスクバッファ上で順番待ちの状態にあれば、小容量アクセスは単に短い転送時間のアクセスになる。

今回、小容量ランダムアクセスが 10% ~ 30% 混在するという想定の下でシミュレーションしたところ、平均処理時間は混在率にほぼ比例して単純減少した。PreSS 転送は小容量ランダムアクセスに対し性能低下をもたらさない転送方式と言える。

この結果は、映像情報のような実時間性を考慮すべきファイルと、テキストデータのような小容量アクセスファイルとを意識的に区別して格納する必要が無いことを意味しており、様々な形態のファイルが混在するマルチメディアサーバとして使用する上で大きな利点である。

4.3 バス使用効率

今回のシミュレーションでは各クライアントのアクセスデータ長を一定にしたので、従来の転送方式におけるバス使用効率は次式のように表せる。

$$P_{DT} = \frac{L/V_{bus}}{T_{acc} + L/V_{dk}}$$

PreSS 転送を行った場合のバス使用効率もシミュレーションから計算できる。両者を下表に示した。PreSS 転送を行うと、バス上の空き時間が減少しバスをより有効に活用していることが解る。また使用効率は実効転送速度に比例しているため、この結果は PreSS 転送による転送速度の向上も意味している。

単位データ長 (KB)	250	150	100
RAID3 型	76.9	66.7	57.1
PreSS 転送	90.0	90.0	80.0

表: バス使用効率 (%)

5 まとめ

PreSS 転送のシステム構成と制御方法について述べ、実システムに基づくシミュレーションで性能を評価した。単位データ長を小さくした場合の転送速度の低下を抑えることができ、実効転送速度の大きな向上が見られた。また小容量ランダムアクセスが混在しても、転送性能に影響を及ぼすことなく対応できた。バスの使用効率に関する解析から、PreSS 転送はシステムリソースを有効利用する上で優れた方式であることがわかった。

参考文献

- 1 高倉 健, 櫻井 紀彦 映像情報サーバ MAMI における連続データ転送方式の検討 第 47 回情報処理学会全国大会 4-1 Oct.1993.
- 2 H. Fujii, A. Ishikawa, N. Kotani and N. Sakurai, "Multimedia server for on-demand services", SPIE Conference, San Jose, Vol.2188/179-188 Feb.1994.
- 3 A. Ishikawa, J. Kishigami, N. Sakurai and N. Kotani, "Multiple-access moving picture information system (MAMI)", IEEE GLOBECOM'92, 1992.