

分散ビデオサーバにおける多重アクセス制御

3R-7

深澤 勝彦 鈴木 偉元 佐々木 主税

NTTヒューマンインタフェース研究所

1 まえがき

近年、高多重度の実現とともに、コスト性能比に優れ、システム構築の拡張性を有するビデオオンデマンドシステムが求められ、PCをプラットフォームとして、複数のサーバ群を高速ネットワークで接続した分散ビデオサーバの開発が活発化している [1][2].
本稿では、分散ビデオサーバにおいて、サーバ間通信オーバーヘッドが少く、応答時間に優れた多重アクセス制御方式を提案し、シミュレーションによりその有効性を評価する。

2 分散ビデオサーバの課題

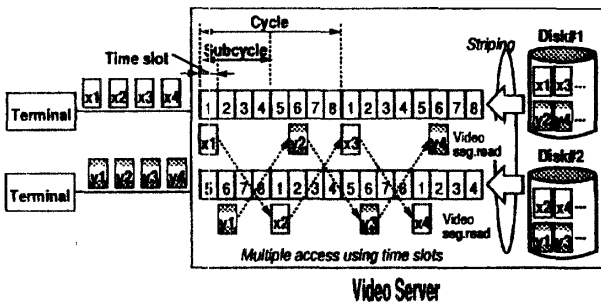


図 1: ストライピング/タイムスロット方式の動作原理

従来の単一サーバでは、多重度を向上し、途切れることなく連続読出しを保証するために、ビデオデータを固定長のセグメントに分割し、全ディスクに分散配置する「ストライピング方式」と、周期的なタイムスロット列に基づいてディスクへのアクセス処理を時分割多重化する「タイムスロット方式」を組み合わせた方式が提案されている [3].

図 1 に上記方式の動作原理を示す。1セグメントの再生時間に相当する周期（小周期 (Subcycle) と呼ぶ）毎にセグメントの蓄積順にディスクを切替えて読出しを行う。多重度 m のディスク n 台の並行動作により、全体として $m \times n$ の多重度を実現できる。

上記方式を適用するために、従来の分散ビデオサーバでは、ある一つのサーバ（以下、マスタサーバ）がタイムスロットを一元管理し、他のサーバは、タイムスロット管理情報をマスタサーバからの通知により共有化している。そのため、以下に示す課題がある。

(1) サーバ間通信オーバーヘッドの削減

端末からのビデオデータの新規再生や特殊再生の要求を受信するたびに、マスタサーバは、全サーバに対して選択/解放されたタイムスロット情報を通知する必要が

ある。したがって、端末からのインタラクティブな要求が頻繁に生じる場合やシステム規模の拡大によってサーバ数が増える場合には、サーバ間通信によるオーバーヘッドが増大する。

(2) 高速応答性

多重度を向上するにはタイムスロット周期 (Cycle) を長くする必要はあるが、一度割り当てたタイムスロットは終了時や特殊再生時を除いて解放されない。したがって、同時接続数が増えるにつれて、新規再生要求に対する応答時間が長くなる。

3 提案方式

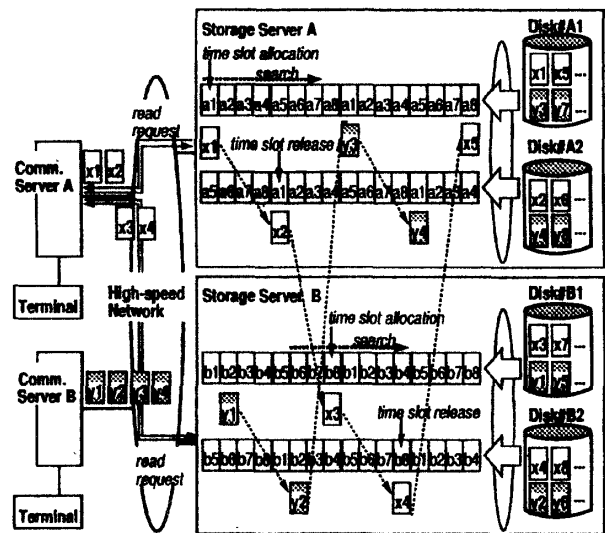


図 2: タイムスロット分散管理方式の動作原理

図 2 に提案方式の動作原理を示す。提案する分散ビデオサーバは、ビデオデータを蓄積し要求に応じて読み出すサーバ（以下、蓄積サーバ）と、読み出したデータを端末に配信するサーバ（以下、通信サーバ）と各サーバを接続する高速ネットワークから構成される。

ビデオデータはセグメントに分割され、全蓄積サーバの全ディスクにストライピングして配置する。端末からビデオデータの再生要求を受信した通信サーバは、番組先頭のセグメントを蓄積している蓄積サーバへその読出しを指示する。通信サーバは、ユーザ端末毎にセグメントの読出し状況を管理し、ビデオデータが途切れないように蓄積サーバに後続するセグメントの読出しを指示する。読出しを指示された蓄積サーバは、未使用のタイムスロットを選択し、セグメントを読み出し、指示した通信サーバに転送する。ディスク数分の読出し/転送処理後、そのタイムスロットを解放する。

提案方式では、通信サーバがタイミングをとって、各蓄積サーバに対して読出しを指示する。このため、蓄積サーバはサーバ単位に自サーバ内のディスク数分のタイ

ムスロット列を管理するだけでなく、タイムスロット選択時にタイムスロット情報を他蓄積サーバに通知する必要がない。また、全サーバの全ディスク数分のタイムスロットを一括管理する場合と比べ、タイムスロット周期はサーバ数に依らず一定（サーバ内ディスク数×ディスクの多重数）であり、タイムスロットの選択/解放の周期が短いため、新規要求の応答時間を短縮できる。

4 性能評価

(1) サーバ間通信オーバーヘッドの比較

通信サーバと蓄積サーバに機能分散した構成を前提とすると、従来方式の蓄積サーバおよびマスタサーバの小周期あたりのサーバ間通信オーバーヘッド（処理コスト） O_{cs} 、 O_{cm} は次式で表される。

$$O_{cs} = 2PC_m M_s \tag{1}$$

$$O_{cm} = 2PC_m M_s N_s \tag{2}$$

同様に提案方式の蓄積サーバの小周期あたりのサーバ間通信オーバーヘッド O_p は次式で与えられる。

$$O_p = \{(1 - P)/N_d + 2P\} C_m M_s / N_s \tag{3}$$

M_s : サーバ全体での同時端末数（提供ストリーム数）

N_d : サーバ内ディスク数

N_s : 蓄積サーバ数

P : 特殊再生あるいは番組終了の発生する確率

C_m : 1制御メッセージの通信に要する処理コスト

蓄積サーバでの小周期あたりの1ビデオセグメントに要する処理コスト（ディスクからの読出しや通信サーバへの転送処理コスト）を C_s とすると、蓄積サーバあたりの処理コストは $C_s M_s / N_s$ となる。蓄積サーバあるいはマスタサーバの処理能力の制約から、サーバ数が N のときの提案方式の許容多重度 $M_p(N)$ 、従来方式の許容多重度 $M_c(N)$ は次式で与えられる。

$$M_p(N) = M_{p1} N \tag{4}$$

$$M_c(N) = \min \{M_{cm}(N), M_{cs}(N)\} \tag{5}$$

$$M_{cm}(N) = \frac{(1 - P)/N_d + 2P + C_s/C_m}{2PN} M_{p1} \tag{6}$$

$$M_{cs}(N) = \frac{(1 - P)/N_d + 2P + C_s/C_m}{2PN + C_s/C_m} M_{p1} N \tag{7}$$

M_{p1} : サーバ数が1のときの提案方式の許容多重度

図3に M_{p1} で正規化した両方式の許容多重度を示す。サーバ内ディスク数 N_d は2とし、パラメータとして、

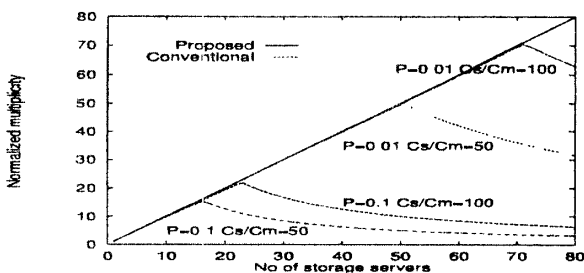


図3: 許容多重度

C_s / C_m を 50, 100, 特殊再生確率を 0.1, 0.01 と変化させている。サーバ数が大きくなるにつれ、従来方式はマスタサーバがボトルネックとなり多重度が向上しないが、提案方式はスケーラブルに向上している。

(2) シミュレーションによる性能評価

特殊再生を含むビデオサーバの動作をシミュレートし、多重数に対する読出し途切れ率（小周期以内に読み出せなかったビデオセグメントの割合）と、平均および最大応答時間を求めた。結果を図4および図5に示す。数値例として、蓄積サーバ数は10、サーバ内ディスク数は2、ディスクの多重度は5、特殊再生の発生頻度は0.01を用いている。結果より、読出し途切れ率と応答時間ともに提案手法が優れていることがわかる。

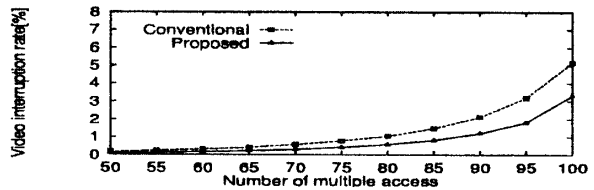


図4: 読出し途切れ率

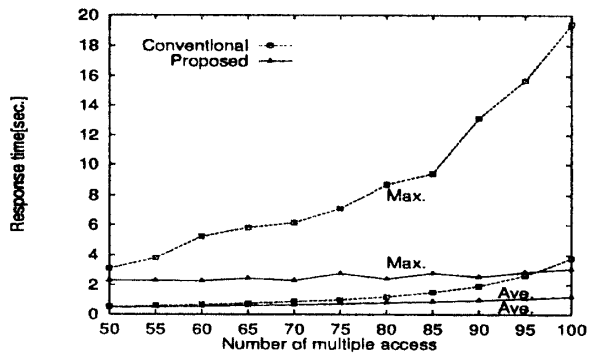


図5: 応答時間

5 むすび

サーバ間ストライピング配置に基づいて、サーバ間通信オーバーヘッドによる性能劣化の低減と応答時間の短縮をめざし、サーバ単位で独立にタイムスロットを管理する多重アクセス制御法を提案した。サーバ間通信オーバーヘッドに関する理論解析により、提案方式は規模に比例して多重度が向上できることを示した。シミュレーションを用いて従来方式と比較した結果、読出し途切れ率および応答時間ともに提案方式が優れていることを明らかにした。今後は、試作により実機検証を行なう予定である。

参考文献

- [1] 佐々木, 鈴木, 深澤: 分散ビデオサーバの構成法, 信学'97 春大, vol.D-11-109,1997.
- [2] 金田, 菊池: ATMネットワークを用いたクラスタ型ビデオサーバシステム, マルチメディア通信と分散処理ワークショップ, pp.321-326,1996.
- [3] 阪本, 西村, 中野: ビデオ情報の大規模多重アクセス方式, 信学論 (D-II), vol.J78-D-II, no.1, pp.76-85, Jan.1995.