

4H-4

VODサービスにおける ディスク入出力制御方式

杉本 欽一

NEC 機能エレクトロニクス研究所

1. はじめに

近年の、メディアの多様化に伴い双方向テレビ技術が注目されている。特にコンピュータ技術を活用した、デジタル動画のサービスとしてビデオオンデマンド（以下VOD）サービスが期待されている[1]。このようなサービスを実現するファイルサーバーに要求される条件は、次の3点である。

- より多くのビデオストリーム（多重度）の実現
- データのとぎれないサービスの保証
- 再生要求に対するレスポンスの保証

本論文では、サーバーシステムにキューを使用したディスクスケジューリングを適用することにより、レスポンスを保証したままディスク装置の稼働率を向上し、高多重度が実現可能であることを示す。

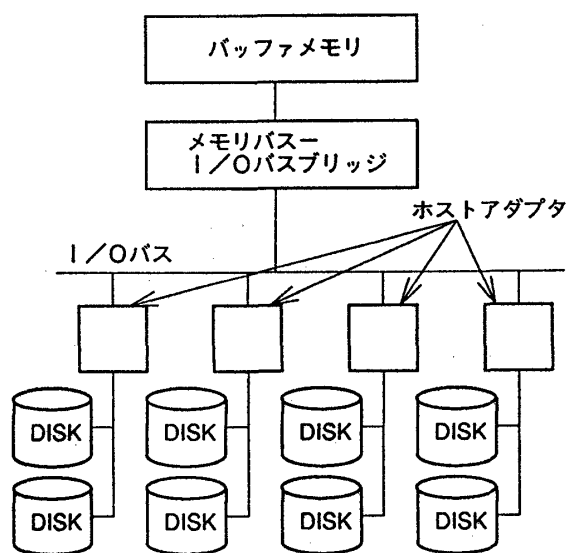


図1 ディスクの接続形態

2. VODアプリケーションの特性

2.1. ディスク装置の接続形態

VODサーバーではディスクアクセスは主に読み出し動作であるため、本論文では、図1のような構成の複数のディスク装置からの読みだし動作に要する時間に関して検討を行う。

2.2. ディスク制御方式

ディスク制御を行うソフトウェアは、図2に示すようにビデオ再生制御プロセスからのデータ読みだし要求を、複数のディスク装置に対する読みだし要求に変換してキューに入れる動作を繰り返す。ソフトウェアアレイ制御ドライバは、各ディスクに対する読みだし要求を単位時間当たりの処理数に上限（キューの長さの制限）を設け、負荷バランスが崩れるのを防いでいる。

3. ディスク入出力動作モデル

3.1. ディスクのシーク特性

ディスク装置のシーク動作は最大速度以内では、バンバン制御に近い最大加速・最大減速制御が行われ、最大速度に達すると一定の速度で動作する。この場合ディスクのシーク距離（横切るシリンダ数）とアクセス時間の関係は、(1)式で与えられる[2]。

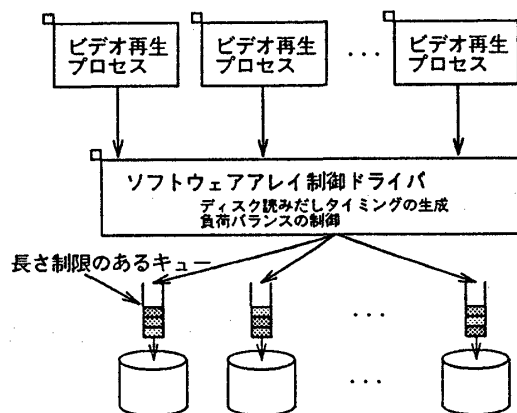


図2 ディスク周りの制御ソフトウェアの動作

ここで、 n はヘッドが横切るシリンダ数であり、 t_{seek} はシーク時間である。 n_{max} は、最大速度に達するまでに横切るシリンダ数を表わす。

$$\begin{aligned} t_{seek}(n) &= a + bn \quad (n \geq 2n_{vmax}) \\ t_{seek}(n) &= c + d\sqrt{n} \quad (n \leq 2n_{vmax}) \end{aligned} \quad (1)$$

今回のシミュレーションではアクチュエータの性能としては、 $n_{max}/3 = 2n_{vmax}$ となるように設計したもものとして評価を実施している。

3.2. シーク動作以外のオーバーヘッドモデル

通常のディスクアクセスに伴うオーバーヘッドとして、回転待ちが発生するが、これは回転角に対して一様なバラツキとして近似できる。

その他のオーバーヘッドとして、ディスクドライブのトラック間シーク、ヘッド切り替え、SCSI プロトコルオーバーヘッドなどを考慮するべきであるが、今回の評価においては固定値と一様なバラツキの和として近似する。

4. シミュレーション評価及び考察

図3に実験及びシミュレーションモデルによる入出力処理時間の分布を示す。ここで、横軸は読みだし要求を出してからレスポンスがかえってくるまでの処理時間を表わし、縦軸は頻度を表す。この例では、キューの長さを5、I/Oサイズを32KBとした場合のリード動作の処理時間の分布を示すが、実測とシミュレーションで同様の傾向を示しており、このモデルを使用してシステムの特性を推定することが出来る。

キューを使用した本制御方式では、一定時間 T 内に全 I/O 処理を保証する必要がある。ただし T は次式で与えられる。

$$T = \sum_{i=0}^q t_i \quad (2)$$

ディスクの入出力処理時間 t の分布が (平均、標準偏差) $= (\bar{t}, s)$ で与えられるとすれば、 T の分布は $(q\bar{t}, \sqrt{qs})$ で与えられる。すなわち、平均の I/O 処理時間は $(q\bar{t}, s/\sqrt{q})$ となるため、結果として平均の I/O 処理時間のばらつきは小さく見積もることが可能となる。特に、 q の値が大きい場合は、中心極限定理より正規分布に近づくため、統計的に処理性能を見積もることも可能となる。

各入出力処理時間の標準偏差が、長さ制限のあるキューを用いることで平均化され低く抑えられるため、ディスク装置の入出力処理のばらつきによる変動分を小さく評価することが可能となる。VOD システム設計で用いられる最悪ケースを用いた設計は不要となり、バスに大きな負荷をかけたような使用条件においてもレスポンスを保証したままディスク装置の稼働率を向上可能となる。ただし、十分なキューの長さを確保できない場合はバッファによる変動分の吸収が必要になると考えられる。

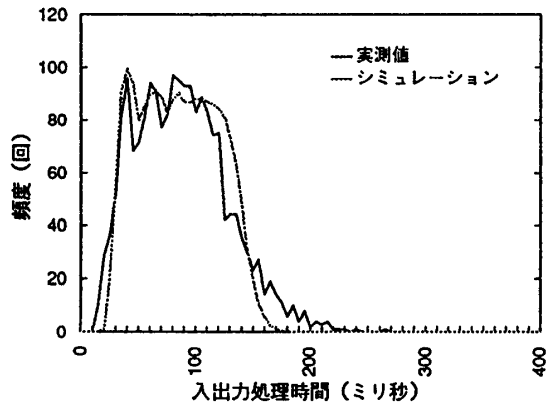


図3 ディスク入出力処理時間の分布

5. 終わりに

本報告では、キューを効果的に使用したVODサーバーのディスク入出力制御方式を提案し、その動作をシミュレーションで検証した。また、その結果より、本制御方式ではI/O処理時間の変動分を相対的に減少させシステムのスループットを向上できる事が明らかとなった。

今後、十分にモデル化出来なかった処理時間を含めたシステム全体の性能評価が課題である。

参考文献

- [1] F.A.Tobaji, J. Pang, R. Baird and M. Gang, "Streaming RAID: A disk Array Management System For Video Files", First ACM International Conference on Multimedia (1993).
- [2] 杉本ら, "マルチディスク装置 (MD-1) の性能評価", 信学会春季全国大会講演論文集, 6-139(1991).