

マルチメディアサーバ「スマートストリーマ」 のアーキテクチャ

1 F-4

浅野 滋博 鈴木 真樹 相川 健 金井 達徳 内堀 郁夫
(株) 東芝

1 はじめに

コンシューマコンピューティング [1] には不可欠な動画を含むマルチメディア情報のネットワークへの送出しは様々な技術課題を含んでいる。特に動画は、転送する情報が膨大で、しかも途切れなく供給する必要があるため、現在の通信、およびコンピュータの技術では量および質の面で十分とはいえない。21世紀の初頭にはATMを始めとする技術が実用化を迎え、通信基盤の整備が進むと考えられる。

本論文では動画をネットワークに途切れなく送出する技術いわゆるビデオサーバについて考察し、筆者らが新しく開発したビデオサーバアーキテクチャであるスマートストリーマについて提案する。

2 ビデオサーバの技術的課題

ビデオサーバはディスク装置などに蓄積された動画をネットワークに送り出すサーバである。動画を蓄積する媒体としては磁気テープ、磁気ディスク、CDROMなど様々な媒体が考えられるが、コストとともにバンド幅および応答速度が重要である。一人のユーザーが一つの媒体を占有するテープなどの装置では、人気のあるプログラムのためには複数のコピーを用意する必要がある。一方、磁気ディスクなどの媒体ではバンド幅が大きく、応答速度も速いので、一つの媒体で複数のユーザーにサービスを行なうことができる。例えば、ディスクの平均バンド幅を16MBpsとして、一人のユーザーが4MBps使うとすると、1台の磁気ディスク装置で4人にサービスすることができる。現在のところ、磁気ディスク装置がコストおよび性能の面で最適であるといえる。

さて、例としてあげたように、一台のディスクでサービスするユーザーの数には限りがあるので、ディスク装置の台数を増やさずに同じプログラムをサービスするユーザーの数を大きくするにはストライピングという技術を使用する。ストライピングは映像を細切れにして複数のディスク装置に分散して記憶する手法で、これにより同時に多数のアクセスが同じプログラムに集中しても複数のディスク装置に負荷が分散することによりサービスを継続することが可能となる。

ところで、ディスク装置からのデータ転送はヘッドの移動や回転待ちなど機械的動作による遅延時間のゆらぎがあるため、ディスク装置からのデータを直接ユーザーに送ることができない。また、一回にディスクからアクセスするデータのサイズは大きいほど高いスループットが得られる傾向があるので、このデータを蓄えて、一定のレートで出力するバッファメモリがビデオサーバに必要である。

ビデオサーバに必要なもう一つの機能は、ストライピングされ複数のディスク装置に分散され記憶された映像を連続した一本の映像として複数のネットワークに送り出すスイッチの機能である。

現在、発表されているビデオサーバのアーキテクチャではこのバッファメモリとスイッチの部分に既存のアーキテクチャが適用されている。例えば、超並列マシンを使用した場合はプロセッシングノード内のメモリをバッファメモリとして使用し、ノード間のネットワークをスイッチとして使用している。この方法では、スイッチとプロセッサが必要以上に含まれているのでコスト高になってしまう。

一方、汎用のワークステーションサーバをビデオサーバとして用いることもできるが、この場合は、メモリのバンド幅やバスのバンド幅がボトルネックとなってしまう。

3 スマートストリーマのアーキテクチャ

筆者らは、安価にビデオサーバを実現するビデオサーバ専用のアーキテクチャであるスマートストリーマ（以下SSと略す）を開発した。以下では図1を参照しながらSSのアーキテクチャについて説明する。

図1に示すようにディスク装置に蓄積された動画はディスク制御装置により読み出され、ストリームスイッチアレー上のバッファメモリ $M(i,j)$ に転送される。さらにネットワーク制御装置により読み出され、ネットワークに送り出される。このようにSSではディスクからの読みだしとネットワークへの送り出しをパイプライン並列で処理している。

SSのもう一つの並列性は、複数のディスク制御装置と複数のネットワーク制御装置を並列に動作させることである。これにより、制御装置の負荷を分散させている。ストリームスイッチアレーはクロスバススイッチの交点部分にバッファメモリを搭載した構造になっている。バッファメモリには、ストライピングの結果分割された動画の一部分が入っている。ネットワーク制御装置からバッファメモリを所定の順序で読み出すことで連続した動画のイメージを得ることができる。

SSの特徴として、複数のバスおよびメモリを並列にアクセスすることで負荷の分散が行なわれ、ボトルネックになりにくいことあげられる。ストリーム制御装置はディスク制御装置およびネットワーク制御装置にディスクのどのブロックをどのバッファメモリに読み出すのか、バッファメモリのどの部分をネットワークに読み出すのかを一定時間毎に指令している。この指令はストリーム制御装置と各制御装置をつなぐ回線を経由して行なわれる。

ストリーム制御装置では、ディスクの読みだしと応答時間の関係が最適になるような要求をディスク装置に対して発行するためのスケジューリングアルゴリズム [2] を実行している。

4 スマートストリーマの RAID 機構

SSの特徴の一つとしてハードウェアの RAID 機構 [3] をストリームスイッチアレーに内蔵していることがあげられる。図1の場合の m 台のディスク制御装置に各々接続される m 個のディスク装置の組のうち $m-1$ 個にデータが格納され、1つがパリティ用のディスクになる。この m 台のディスク装置の組のうち1台が故障した時は、残り $m-1$ 台のディスク装置の内容から故障した1台の内容が復元できる。

ストリームスイッチアレーの各交点に内蔵された EXOR 回路がバッファメモリから動画データを読み出す際に演算しながらデータを転送することでネットワーク制御装置が正しいデータを受けとれる。

新しい動画データをディスクに格納する場合は、この逆のバスをたどってパリティを生成しながらバッファメモリにデータを蓄え、さらにディスク装置に書き込みを行なう。

ディスクが故障して交換した場合、交換したディスクにもともと入っていたデータを復元しなければならない。これも RAID 機構を使用することでパフォーマンスを落さずにデータの復元をすることができる。

5 まとめ

以上のようにSSではメモリとバスを分散して配置することでボトルネックをなくし、高いコストパフォーマンスを実現できる。また、ハードウェアによる RAID 機構で性能を落さずに高信頼性を得ることができる。

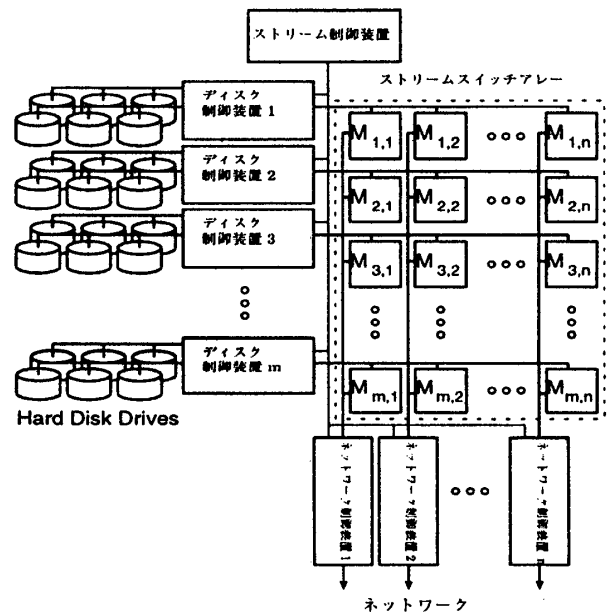


図1: スマートストリーマの構成

参考文献

- [1] 金井ら, 'マルチメディアサーバ「スマートストリーマ」の概要' 本予稿集
- [2] 矢尾ら, 'マルチメディアサーバ「スマートストリーマ」の制御ソフトウェア' 本予稿集
- [3] P.Chen, G.Gibson, R.H.Katz, D.A.Patterson and M.Schulze, 'Two Papers on RAIDs,' USB/CSD 88/479, University of California Berkeley, 1988