3. VOD システムにおけるビデオサーバと セットトップの基礎技術

3-1 ビデオサーバ

正会員 小 西 智 夫[†]

1. ま え が き

ビデオオンデマンドとかビデオサーバという言葉が一般にも聞かれるようになって、そろそろ2年目になるようにしている。当初の誤解も徐々に解け、ビデオサーバといえども、単一の映像ファイルから無限のストリームを配信できるものではないことが、共通の認識となってきたようである。

ところで、いまだに耳に馴染まないこのストリームという言葉は、ビデオサーバの独立した映像チャンネルを表し、1000 ストリームのビデオサーバといえ、1000 本の独立した (個々に独立してコントロール可能な) 映像を供給できることを意味する。そうであれば、1000 チャンネルとか、1000 クライアント等といったもよさそうであるが、本来、間欠的なデータの処理を第1目的としてきた通常のトランザクションサーバとの違いを、ストリームという言葉を使うことによって、途切れることのない継続性を暗示的に表すことができ、まことに都合が良いようである。

さて、現在アメリカで計画されているインタラクティブテレビで、サービス世帯数と同じ数のストリームを用意する所はない。例えば、カリフォルニア州サンノゼ周辺で行われるパシフィック・テレシスの場合、発表によれば、契約総世帯数を10万と見積もっているが、ビデオサーバの総ストリーム数は1万である。

ここに、あるビデオサーバがあって、1000 ストリームの能力があるとしよう。しかし、この仕様だけでは、単純にビデオサーバの最終段における出力能力を示しているだけで、はたしてすべてのストリームが単一のプログラムを要求できるように設計されているの

か、多くの違ったプログラムを比較的ランダムに視聴することを前提にしているのかは、わからない。最初にも触れたように、単一のファイルから無限にストリームを得ることはできないからである。単一ファイルから複数のストリームを取り出す原理は単純で、タイムシェアリング以外の何物でもない。すなわち、100 MB/s の連続読出し能力の記憶媒体があれば、0.5 MB/s のストリームを200本確保できる。その後、続くビデオサーバの内部バスの連続転送速度は、それを上回らなくてはならないことはいうまでもない。

ここでは、このようなビデオサーバの基礎技術について紹介する。

2. ビデオサーバの構成

2.1 アーキテクチャ

ビデオサーバの実現手法には種々あるが、いずれの場合も主記憶装置の映像データを多くの独立したストリームに分けて、リアルタイムで出力することには変わりはない。しかし、ビデオサーバをコンピュータとして見た場合、今までの設計の目標からは根本的に違っているものがある。近年のコンピュータの高速化には、CPU の高速化が最重要課題であって、次々に最新の技術が投入された。CPU 自身が非常に高速になったため、I/O に係わる時間が問題でキャッシュが導入された。このような設計は、主に演算処理など比較的限られた記憶領域に頻繁にアクセスを繰り返す場合に特に有効であるが、ディスクと通信のために入出力を繰り返すことが最も優先度の高い仕事となるビデオサーバにとっては、効率の良い方法ではない。

図1に、ディスクとI/O間の転送能力のみを優先させた専用ハードウェアを使用したビデオサーバのアーキテクチャを示す。機能としては大きく3つに分かれており、映像入出力専用のビデオトランスファエンジン (VTE)、VTE の制御を行うリアルタイムコント

[†] 横河・ヒューレット・パッカード株式会社
“Fundamental Technologies of Video Server and Settop for VOD; Video Server” by Tomoo Konishi (Yokogawa-Hewlett-Packard, Ltd., Tokyo)

図 1 専用ハードウェアによるビデオサーバ

図 2 ディスクからの多重読出しの模式図

ロール、課金やその他の顧客情報を扱うトランザクションマネージャである。

家庭にあるセットトップボックスから送られてくるリクエストは、リアルタイム OS を搭載した PA-RISC ベースのマスターストリームコントローラ (MSC) で、一時停止や早送り等の即時実行コマンドか、新しい映画のリクエストやショッピング等の顧客情報へのアクセスの必要なコマンドかに分けられる。即時実行コマンドの場合は、すぐさま現在アクティブなストリームを持つローカルストリームコントローラ (LSC) に対して適切な指令を行う。このような構成によって、即時実行コマンドの実行遅延を違和感の目立たない 500 ms 程度におさえることができる。

LSC は CPU、ディスクアレイコントローラ、アウトプットストリームコンディショナから成り立ち、ディスクアレイの高速入出力をつかさどる VTE の心臓部分である。適切な設計をすれば、特別な部品を使わずとも 1 つの LSC で OC-3 (150 Mb/s) の出力を 2 系統ドライブすることができる。

2.2 ハードディスクストレージ

ビデオサーバの複数読出しの基本はタイムシェアリングである。大型のキャッシュを用いたり、RAM を

直接映像のソースとして使用しない限り、その転送速度、アクセス時間、コスト等から、ハードディスクを主記憶装置として使用するのが現実的である。

映像ファイルは、たとえ MPEG 等の圧縮を使用した後でも巨大である。3 Mb/s のデータレートを使用したとすると、2 時間の映像ソースを 100 本持つには最低 270 G バイト必要である。光磁気ディスクの転送速度も最近上昇してきているが、2 次記憶装置としてハードディスクと組合せて階層的に使用する方が实际的である。ディスク装置をビデオサーバの主記憶装置として見た場合には、ピーク転送速度と同時に平均アクセス時間も重要なファクタになるからである。図 2 にディスクからの様子を示す。

平均アクセス時間は、平均シーク時間と平均回転遅延時間が主要なファクタである。平均シーク時間は、ヘッドが読出しトラックまで移動する時間であるが、近年非常に軽量化されたヘッド回りとアクチュエータの改良によって、10 ms 内外になった。平均回転遅延時間は、ディスクの回転による先頭データの読出し遅延であるが、5400 rpm のディスクの場合、約 5.6 ms である。

このような平均アクセス時間も考慮に入れた連続読

出し速度の上昇は、そのままサポートできるストリーム数の上昇になるが、ディスク自身の改良以外にもデータをストライプ（分割）して複数のディスクに同時に記録することによって、同様の効果をあげることができる。この手法については、RAID（Redundant Array of Inexpensive Disks）として知られているが、この最初の報告は5つのディスクを使ったモデルとしてUCB（University of California at Berkeley）の3人の研究者によってまとめられた²⁾。

RAIDの形態については、0から5の基本形と、その組合せの6, 10, 53の9種類が知られている³⁾⁴⁾。これらのどれもがビデオサーバに適しているとはいえないが、簡単に紹介しておく。

(1) RAID レベル 0

エラー訂正情報なしに、セクタ長で複数ディスクに分割。

(2) RAID レベル 1

通常のディスクにミラーディスク（全く同じものをコピーしたもの）を追加。読出し速度は2倍。

(3) RAID レベル 10

レベル0/1とも呼ばれ、セクタ長で複数のディスクに分割した後、レベル1のようにすべてにミラーを持つ。

(4) RAID レベル 2, 3

エラー訂正情報は冗長ディスクに、データは通常バイト単位で複数ディスクに、それぞれのスピンドルを同期させて同時読み書き。レベル2は、より確実なエラー訂正のために冗長ディスクの数がレベル3に比べて多い仕様だが、実際はアルゴリズムの改良により冗長ディスクを1つにしたレベル3で十分な場合が多い。大容量ファイルの転送に最も高い転送レートが得られる。

(5) RAID レベル 4

レベル3の分割単位をセクタ長にし、スピンドルの同期運転をやめた。データがセクタ長以下の場合には、アレイを構成するドライブに独立に書込むようにしてI/Oレートの改善を行った。

(6) RAID レベル 5

レベル4のエラー訂正情報をデータディスクに順番に割り振っていく方法で、最もよく使われている。

(7) RAID レベル 6

レベル5に2重のエラー訂正情報を付加して、もう1つの冗長ディスクに書込む。より確実なエラー訂正ができるが、書込み速度が低下。

(8) RAID レベル 53

レベル0とレベル3を組合せたもの。データはセ

クタ長で分割された後、再びレベル3に合わせてデータの置換えとエラー訂正情報の記録が行われる。

2.3 クラスタ型システム

特殊な構造のビデオサーバ以外においては、通常1000ストリームもの出力を得ようとするばクラスタ型になる。図1のビデオサーバにおいても、1つのLSCで100ストリームが得られるとすれば、10台のLSCをMSCが制御することになる。このような場合、リソースのダイナミックな制御が不可欠になる。

例えば、仮りにこのビデオサーバに100本の映画が蓄積されているものとする。当然それらの映画は、各LSCに10本ずつ位に分けられている。通常リクエストが均等に発生している場合には問題ないが、1つのプログラムにリクエストが集中したような場合に、それぞれのLSCのリソースをうまく配分しなければ、すぐにサービスが得られなくなってしまう。

具体的には、プログラムAにリクエストが集中し始めたら、そのプログラムを持っているLSCのリソースが使い切られない間にプログラムAのコピーを別のLSCに作り、複数のLSCでそれらのリクエストに答えるようにすべきである。実際、全体の50%にもものぼるストリームが単一のプログラムに集中したような場合には、そのプログラムはすべてのLSCでシェアできるように管理した方が安全である。入出力バンド幅を十分に持ったLSCが相互に接続されている場合には、そのLSCのリソースがまだ残っている適当な時点でコピーを行えば、120分の映像ソースを数分で完了することができる。

また、1つのLSCにリソースのインタフェースの役割を与えた設計も考えられる。例えば、光磁気ディスク、磁気テープ等の2次記憶装置やネットワーク上に接続された別のビデオサーバのために1つのLSCを割り振る。ハードディスク上に存在しない映像ソースがリクエストされた場合、ネットワークや光磁気ディスクから読出してくるように、記憶装置を階層的に使用することができる。仮りにそのLSCのハードディスクに、2次記憶にあるソースの最初の一部を記憶しておくような構成にすれば、ユーザに頭出しの問題を感じさせるような心配はない。

このように、クラスタ型のビデオサーバは、ストリーム数の設計に自由度が大きいばかりでなく、事業者の要求に応じて柔軟に対応できる優れたアーキテクチャであるが、実際の運用時には、いかにダイナミックで、きめ細やかな制御が行えるかは、そのマネジメントソフトウェアに負うところが大きい。上述の事例は、1つのLSCの能力が大きくなればなるほど対処

しやすい。実際 10 ストリーム程度の能力の LSC を 100 個並べて 1000 ストリームのビデオサーバを作るというのは考えにくい。この点については、ストリーム当たり高ビットレートを与えた場合にも同様である。15 Mb/s 以上のビットレートの場合、300 Mb/s 程度の LSC を、クラスタ構造によってストリーム数を稼ごうとしても、全体の管理は結構複雑になると考えられる。

3. む す び

日常、我々のコンピュータ環境はクライアント/サーバによるネットワーク化が進み、実に快適に仕事が行われるようになった。ビデオサーバという言葉は、このような快適性をビデオの世界にも今すぐ現実のものとする響きを持っている。すでに一言では説明できないほど、多様なアーキテクチャのビデオサーバが紹

介されていることは、今後の発展を予期する嬉しい状況である。今後、実際の使用報告が期待されるところである。
(1994 年 12 月 15 日受付)

〔参 考 文 献〕

- 1) A. Kovalick: "The Fundamental Concepts of Media Servers", Hewlett-Packard internal document (Sep.1993)
- 2) David A. Patterson, Garth Gibson, Randy H. Katz: "A Case for Redundant Array of Inexpensive Disks" (1987)
- 3) W. L. Rosch: "Hardware Bible, Third Edition", Brady Publishing, pp.1199-1223 (1994)
- 4) "Disk Arrays", Hewlett-Packard (1993)

こにし ともお
小西 智夫 昭和 54 年、関西大学工学部電子工学科卒業。ソニーテクノロクス(株)、米国テクノロクス(株)TV 事業部にて、TV 信号自動測定器の開発に従事。平成 5 年、横河・ヒューレット・パッカード(株)に入社。現在、ビデオサーバ関連の開発を手掛ける。正会員。

3-2 セットトップ

阪 田 史 郎[†]

1. ま え が き

セットトップは、センタのビデオサーバやマルチメディアサーバから送られてくる番組を受信し、視聴者が家庭のテレビ受像機で見ることができる映像に変換する装置である。テレビ受像機の上に置くことが多いため、セットトップボックス(STB)やセットトップユニット(STU)、セットトップコンバータなどと呼ばれる。

従来のアナログ CATV 用に使われてきたセットトップは、多チャンネルの番組の中から視聴者が番組を選択した時に、その番組に対応した受信チャンネルに周波数を変換するチューナの役目を果たす。近年注目を集めているデジタルネットワーク(デジタル CATV や光ファイバによる ATM ネットワークなど)による VOD やマルチメディアオンデマンドにおいては、デジタル情報として符号化(エンコード)された MPEG の映像を復号化(デコード)してテレビ受像機側に出力する機能に加え、ソフトウェアによって実現

図 1 セットトップの全体構成
(アナログではハードウェア部分のみ)

されるビットマップ制御や画面メニュー操作、サーバ側とのインタラクションなどを制御する機能が要求される。ネットワークと家庭のテレビ受像機との間のインタフェース全般を司る装置(コンピュータ)と呼ばれるようになっている。また、オフィスで利用する場合には LAN を通した情報転送を行うことになり、端末(パソコン)にセットトップの機能が内蔵され、端末と一体となった形態となる。

図 1 にセットトップの全体構成の概要を示す。

[†] 日本電気株式会社 C&C 研究所
"Settop Box" by Shiro Sakata (C & C Research Laboratories, NEC Corporation, Kawasaki)