マルチメディアサーバ“スマートストリーマ”

SmartStreamer : Multimedia Server for Consumer Computing Era

天野　真家
S. Amano

相川　　健
T. Aikawa

金井　達徳
T. Kanai

21世紀に向けて高速広域ネットワークの基盤整備が進み，それを利用したマルチメディアを用いた情報サービスシステムの構築が現実のものになろうとしている。このような情報システムでは，映像に代表される連続データを途切れることなくしかも同時に多数のユーザに対して配信するマルチメディアサーバの役割が重要になる。

今回開発した“スマートストリーマ”は，数百人のユーザに異なる映像を同時に配信することを可能にするマルチメディアサーバである。連続データの処理に適したデータパスを採用することにより，コストパフォーマンスの良いマルチメディアサーバを構築することができる。

The development of a high-speed wide-area network infrastructure has made it possible to realize networked multimedia information systems. The multimedia server is a key component of such a system, capable of continuously and simultaneously distributing hundreds of video streams to customers.

This paper describes a new architecture for a multimedia server called SmartStreamer. SmartStreamer has a dedicated data path structure for continuous data processing and can be implemented in a highly cost-effective way.

1 まえがき

ARPA (Advanced Research Project Agency) 計画の一環として始まったインターネットが非常な勢いで発展している。インターネットに接続されている計算機は昨年300万台と言われていたが，現在ではすでに600万台を超え，6,000万人が使用しているであろうと推定されている。インターネットの用途としては，メール，WWW (World-Wide Web) での使用が著しく，WWWを用いたオンラインショッピングに出店する店舗の数も増加しつつある。

一方，VOD (Video On Demand)，NOD (News On Demand) のような動画を中心とする計算機応用がマルチメディアと呼ばれる分野として急速に台頭しつつある。動画はたとえディジタル化しても，その情報量が膨大になるため，従来の計算機の処理速度，特にバス，周辺機器の速度となじまず，また，通信路容量ともなじまなかった。しかし90年代に入り，MPEG (Moving Picture Experts Group) と呼ばれる圧縮方式が国際標準化されるとともに，計算機処理に適する処理量にまで低下してきた。特に注目されているVODの試みに，タイムワーナーケーブル(社)がフロリダのオーランドで試行を始めたFSN (Full Service Network) がある。これはMPEG2方式で圧縮した映画を契約している各家庭からの要求に応じて即時にケーブルにより配信するもので，従来の放送形態によるケーブルテレビとは決定的に異なるものである。

マルチメディアは文字どおりに解すれば“複数のメディア”ということに過ぎない。そのなかでも，MPEG技術の開発により動画をリーズナブルな計算量として扱うことができるようになり始めたことがマルチメディアの本質である。しかし，マルチメディアには現時点ではまだ制約が多い。インターネットでは一般的にはまだ，MPEG2による高品質な動画を送ることはできない。オンデマンド方式による通信は，放送に比べて非常に多くの負荷をバスや周辺機器に課すため，計算機システムが高価格になる，などである。

しかし，これらの問題も解決の方向に向かいつつある。前者では，米国のGlobal Information Infrastructure計画，わが国の電気通信審議会の活動などに見られるように，光ファイバが国際的に敷設されるのはそれほど遠いことではないだろう。後者に関しては，現在のところ大半は汎(はん)用計算機，超並列計算機などを用いて対処している。しかし，これらの従来の計算機はマルチメディアの処理――すなわち，連続メディアをとぎれることなく多数の端末からの要請に応じて配信する処理――のために設計されたものではないため，連続メディアの大量配信機構を実現すると，それに応じて，CPU周辺まで比例して大きくなり，高価格のものとなっている。

ここではこれらを考慮に入れ，マルチメディアサーバの新しいアーキテクチャ，特にその中核となるストリーマについ

て述べる。

2 システム構成

マルチメディアサーバは，21 世紀に実現されるマルチメディアを用いた情報サービスシステム FSN のキーコンポーネントである。**図 1** に FSN の全体システム構成を示す。21 世紀にはワールドワイドに接続されたネットワークを介しマルチメディアを用いて，これまで人間が出向いて受けてきたサービスが，いながらにして受けたいとき（オンデマンド）に受けられることを目ざしている。

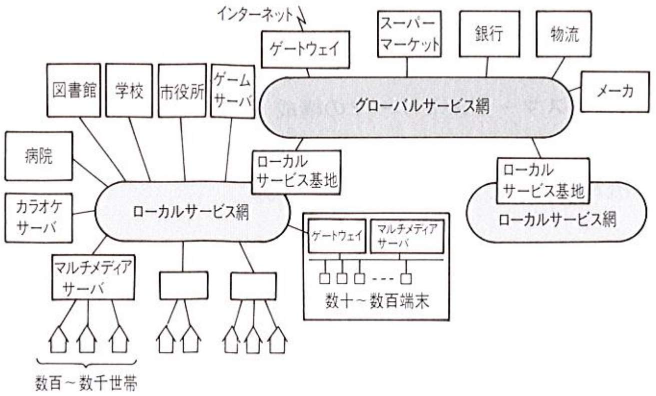

図 1．フルサービス ネットワークシステム　ネットワークを介してさまざまなサービスを家庭や職場からオンデマンドで受けることができる。
Full-service network system

一方，映像を用いてサービスを行うためには，映像の配信技術だけを取りあげてみても，通信回線の伝送能力，通信回線に映像を送り出すサーバの能力をはじめとして多くの問題が存在するが，現在の最高の技術を用いれば映像を用いたサービスが実現可能であることは，各所で行われている FSN システム実験の例からもわかる。後はサービスが事業的に成り立つ価格でシステムを実現する技術が要求されているわけである。われわれはサーバ部分の低価格化を実現する技術として“スマートストリーマ”と呼ぶマルチメディアサーバを開発した。ここでストリーマとは，マルチメディアサーバの中でも特に映像のような連続データの蓄積と配信に特化したものを指す。

3 マルチメディアサーバの技術的課題

映像を扱うことによるマルチメディアサーバの技術的課題は以下のとおりである。

(1)　映像データは大量のデータである。
(2)　単位時間に送るべきデータ量が多い。
(3)　映像は連続データである。
(4)　大量の映像蓄積メディアの耐故障性が必要。

これらの技術的課題を解決して，同時に多数の映像データをネットワークを介し配信するメカニズムについて，**図 2** に示すマルチメディアサーバの抽象化したアーキテクチャを用いて説明する。

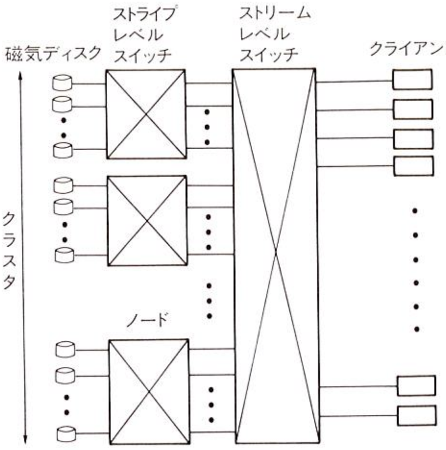

図 2．抽象化したマルチメディアサーバ　ストライプレベルの交換を行うスイッチと，ストリームの交換を行うスイッチの 2 段から成る。
Abstract structure of multimedia server

映像データは大量のデータであるため，蓄積メディアとしては磁気ディスクが使われる。4 G バイトの磁気ディスクには MPEG2 で圧縮された映像 2 時間分が蓄積可能である。また，映像データは単位時間に送るべきデータ量が多いので，複数のディスクに分割して格納されている。これを一般的にストライピングと呼び，同時に一つの映像コンテンツに対しアクセスが集中したときにも連続的に映像データの読出しを可能にする。

複数のディスクから読み出された映像コンテンツは，ストライプレベルスイッチで 1 本の映像ストリームにまとめ上げる。一本にまとめ上げられた映像ストリームは，ストリームレベルスイッチでスイッチングされ映像データを要求したクライアントに送られる。磁気ディスクとストライプレベルスイッチのペアをマルチメディアサーバと呼ぶ。磁気ディスクの読出し速度は映像の再生に必要な速度の数倍高速であるので，速度変換のためのバッファメモリが磁気ディスクからクライアントまでのどこかに最低一つ必要である。ストライプレベルスイッチとバッファメモリの実現法には，汎用計算機を用いる方法，超並列計算機を用いる方法，専用ハードウェアを用いる方法がある。同時にクライアントに供給可能なストリーム数がストライプレベルの出力ポート数を超える場合にはマルチメディアサーバを複数個並べる。これをクラスタリングと呼ぶ。システムの要求ストリーム数の実現は，一つ

のストライプレベルスイッチの規模と，クラスタリングの台数で多くのやりかたがある。ストリームレベルスイッチとメモリの実現法が同一ならば，ストライプレベルスイッチの規模を大きくしてクラスタリングの数を減らせばスイッチのコストが上がる。逆の場合にはストライプレベルスイッチのコストは下がるが，アクセス頻度の高いコンテンツは各サーバにコピーするため磁気ディスクの個数が増える。各種方式の特長を**表1**に示す。

表1．マルチメディアサーバの方式比較
Comparison of multimedia server architectures

ノードの実現手段	超並列	汎用サーバ	PC	スマートストリーマ
バッファメモリの実現法	PU 内メモリ	メインメモリ	メインメモリ	ストリームスイッチアレイ（メモリ付専用スイッチ）
ストライプレベルスイッチの実現法	PU 間結合網	高速バス 数Gバイト/秒	中速バス 133Mバイト/秒	
実現可能ストリーム数	数千	数百	数十	数百
ノード当たりの価格	高	高	低	低
中規模システム実現時のコンテンツのコピー	不要	不要	必要	不要

PU：Processing Unit

スマートストリーマではストライプレベルスイッチに低価格化の可能な専用ハードウェアを用い，スイッチ一つで240ストリームまでのサポートを行い，それ以上の規模はクラスタ化でサポートする方式を採った。数百ストリームサポートを行う場合には，ストライプレベルスイッチのほうが低価格で実現でき，かつコンテンツのコピーが不要であるため，コストパフォーマンスの良いマルチメディアサーバの実現が可能である。

また，マルチメディアサーバは数十個から数百個の磁気ディスクを使用するが，磁気ディスクの故障時にもサービスを継続できることが重要である。スマートストリーマでは4：1のRAID (Redundant Arrays of Inexpensive Disks) 構成を採用し，かつストライプレベルスイッチの中でディスク故障時のエラー訂正演算を行う。そのためディスク故障時にも性能劣化なしにサービスを継続することができる。

4 スマートストリーマのアーキテクチャ

4.1 スマートストリーマの仕様

スマートストリーマの仕様を以下に示す。

(1) 蓄積可能データ量　256 Gバイト
(2) 最大同時サービス映像数　240 ストリーム MPEG2 (4 Mbps)
(3) コンテンツ蓄積媒体　4 Gバイト磁気ディスク 80台
(4) 高信頼化方式　4：1のRAID方式
(5) スイッチ実現方式　メモリ付きスイッチアレイ
(6) 送出開始遅延時間　1秒以内

スマートストリーマはアプリケーションサーバから指示されたビデオストリームをディスク装置から取り出し，ATMネットワークに一定速度で送り出す装置である。

ハードウェア構成としては，1秒間に960 Mビットという多量のデータを扱うために複数の処理ユニットに分散して処理を行っていることに特長がある。分散の方法としては，処理を直列に分配するパイプライン処理と並列処理の両方を行っている。また，装置につけられた80個を超える大量のディスク装置の故障に対応して，ディスク装置単位で冗長符合をとり，ディスク故障時にもシステムが停止することなく動作するRAID構成を採用している。

4.2 スマートストリーマの構成

図3にスマートストリーマの構成を示す。以下では，処理の流れに沿って各部の構成を述べる。

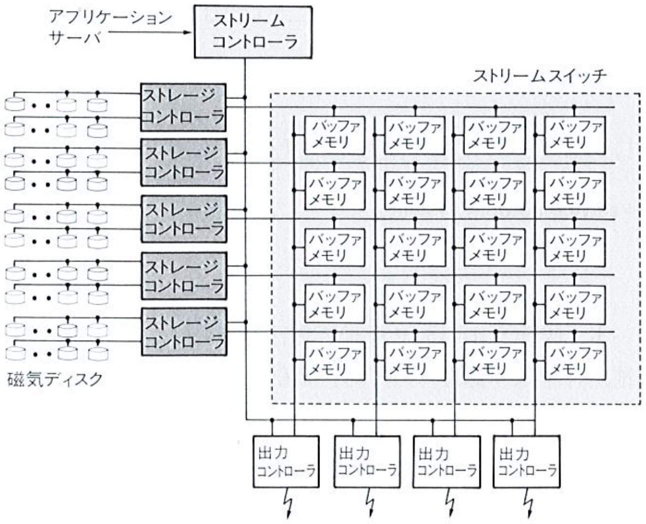

図3．スマートストリーマの構成　複数のストレージコントローラと出力コントローラの間を，バッファメモリのアレイで接続する。
Block diagram of SmartStreamer

4.2.1 ディスク装置　スマートストリーマには960 Mbpsを引き出す80個以上のディスク装置が接続されている。現在，接続にはWIDE SCSI (Small Computer System Interface) を使用している。

4.2.2 ストレージコントローラ　ディスク装置を制御する。ストリームコントローラ（後述）の制御に従ってディスクを読み出し，その内容をストリームスイッチに転送する。RAIDは四つのディスクに対して一つのパリティをもち，システム中には五つのストレージコントローラをもつ。

4.2.3 ストリームスイッチ　ストレージコントローラからのデータと出力コントローラ（後述）へのデータを交換するストライプレベルスイッチおよびバッファメモリの機能をもつ専用ハードウェアである。クロスバスイッチ構造で各交点に16Mバイトのバッファメモリを備える。バッファメモリがストレージコントローラと出力コントローラ間のバッファリングを行っている。RAIDのための演算はストリームスイッチ内部で行われ，ディスク装置故障の場合でも正しいデータを出力し続ける。

4.2.4 出力コントローラ　ATMネットワークへの出力を制御する。ストリームスイッチの交点上のバッファメモリからデータを読み出し，ATMネットワークに出力する。

4.2.5 ストリームコントローラ　アプリケーションサーバからの指示に従い，ディスクの読出し動作のスケジューリングを行い，ストレージコントローラにディスク読出しの指示を行う。また，出力コントローラに対して出力の指示を与える。

以上のように，ストレージコントローラ，ストリームスイッチ，出力コントローラがパイプラインで動作し，処理を分担している。また，ストレージコントローラ，出力コントローラ，ストリームスイッチ上の複数のバッファメモリで並列に処理を行うことでデータのふくそうを防ぎ，多量のデータを流すことが可能になった。

5 スマートストリーマのスケジューリング方式

スマートストリーマでは，ストリームコントローラが数十msのタイムスロットごとに，ストレージコントローラへはどのディスク装置はどのクライアントの映像ストリームを読み出してどのバッファメモリに入れるか，出力コントローラへはどのバッファメモリの内容をどのクライアントへ送り出すかを指示することで，コンテンツの配信処理を進める。これらの各タイムスロットにおける指示は，各映像ストリームの連続性を満たすためにはどういう順序でディスク装置からの読出しを行い，どのようにバッファメモリを使用するかを考慮してスケジューリングする必要がある。

一般に映像ストリームを配信するコンテンツサーバのスケジューリング方式には，クライアントからの要求発生順にディスク読出し要求を発行するFIFO方式と，各クライアントのディスク読出し要求をディスク上の位置の順に並べ変えて発行するSCAN方式がよく用いられる。両者を比較した場合，応答時間に関しては要求発生順にディスク読出しを行うFIFO方式が優れるが，ディスク装置の能力をより有効に利用し，多数の映像ストリームを取り出せるという点ではSCAN方式が優れる。

マルチメディアを用いた情報サービスでは，利用者からの要求に対してより早く応答を返すことが必須(す)であるとの観点から，スマートストリーマのストリームコントローラではFIFO方式をベースにより応答性能を高めた方式を用いている。従来のFIFO方式では，新しい映像ストリーム要求が到着してからその配信のための空タイムスロットが確保できるまでの時間は，同時に配信している映像ストリーム数が多くなるほど確率的に長くなる。そこで，配信中のすべての映像ストリームの連続性を乱さない範囲で，各映像ストリームに対応するディスク読出し要求を発行するタイムスロットをずらして，空タイムスロットをつねに時間的にまばらに分散させておく。その結果，多数の映像ストリーム配信中でも，新しい映像ストリーム要求発生時に空タイムスロットを確保するまでの時間を短くすることが可能になっている。

6 あとがき

マルチメディアは，従来の情報処理，すなわち企業の勘定系，情報系のような数値計算と情報検索，そしてその後興ったワードプロセッシングという概念，すなわち，文書処理の次にくるコンピューティング技術と見ることができる。そして，この順に情報処理は，特殊な目的から家庭内への目的へと領域を広げてきた。マルチメディアは，これまで計算機が経験したことがないほど，生活に密着した形態のコンピューティングである。2章で挙げたように，VOD，NOD，各種の情報提供，オンラインショッピング，遠隔教育，遠隔医療，地域情報システム，ゲームのようなものが考えられている。このようなコンピューティング技術をその用いられかたから，われわれはコンシューマコンピューティングと名付けた。連続メディア用の低価格な機器とそれを実現する計算機技術，子供でも使える簡単なヒューマンインタフェース，公共システムとしての信頼性，ネットワーク上での課金／セキュリティ技術などが特長となる。計算機アーキテクチャはその領域をいっそう広げていくことになるだろう。

天野　真家　Shin'ya Amano

1973年入社。自然言語処理システムの研究・開発に従事。現在，研究開発センター 情報・通信システム研究所部長。
Communication & Information Systems Research Labs.

相川　　健　Takeshi Aikawa

1976年入社。計算機アーキテクチャの研究・開発に従事。現在，研究開発センター 情報・通信システム研究所主任研究員。
Communication & Information Systems Research Labs.

金井　達徳　Tatsunori Kanai

1989年入社。分散処理ミドルウェアの研究・開発に従事。現在，研究開発センター 情報・通信システム研究所研究主務。
Communication & Information Systems Research Labs.