

## 分散 R A I D 方式ビデオサーバー

中村俊一郎 峯村治実 山口智久 (三菱電機・情報技術総合研究所)  
清水洋 渡辺尚 水野忠則 (静岡大学・工学部)

ビデオサーバーの性能(同時ビデオ配信数)向上を目指した分散 R A I D 方式ビデオサーバーの提案を行う。本方式は通常ディスク群に対して適用される R A I D 技術をサーバー群に対して適用するものであり、ディスク R A I D 技術と同様に性能及び信頼性の向上を図ることを狙ったものである。これが実現すると普及型の安価なパソコンサーバーを追加していくことにより、ビデオサーバーをスケーラブルに性能向上出来るという非常に大きな利点が得られる。ここでは R A I D 0 型の分散 R A I D 方式ビデオサーバーの試作を行ったので、その方式と評価結果について述べる。

### Distributed RAID Style Video Server

Shunichiro Nakamura, Harumi Minemura, Tomohisa Yamaguchi  
(Mitsubishi Electric Corp.)  
Hiroshi Shimizu, Takashi Watanabe, Tadanori Mizuno  
(Shizuoka Univ.)

In this paper, we present a distributed RAID style video server that addresses the problem of increasing video stream supplying capability in VOD systems. "Distributed RAID Style" means extending the RAID method usually applied to disks to the servers, so as to achieve improvements in performance and reliability. The great advantage of this architecture is that the linear performance improvement can be achieved by adding inexpensive servers. Here we describe the implementation of RAID0 Style video server that we conducted and its evaluation result.

## 1. まえがき

近年のマルチメディアブームの中で VOD (Video on Demand) の研究が盛んである。我々は以前、通常使われている LAN 上の UNIX サーバに、複数のクライアントから NFS を用いてビデオデータの同時実時間アクセスを行うことにより、ビデオサーバとして動作させる実験を行った。<sup>(1)</sup> 又 Windows NT 環境でも同様の実験を行った。<sup>(1)</sup> この結果 LAN 性能さえ満足されれば、これらのサーバ 1 台当たり 1.2Mbps の画面良好なビデオストリームをそれぞれ 35、20 程度供給できそうとの予測結果を得た。

本論では上記の方式をさらに進めて、サーバのビデオ供給能力及び信頼性の向上を狙った分散 RAID 方式のビデオサーバの提案を行い、その試作評価結果およびシミュレーションによる評価結果について述べる。

## 2. 分散 RAID 方式とは

分散 RAID 方式とは通常ディスクに対して使われる RAID の手法を、サーバにまで拡張して適用してデータの分散配置 (ストライピング) により性能向上を図るものである。ユーザ数が増えて 1 台のサーバのビデオ供給能力を超えた時、サーバ台数を増やしてビデオプログラム単位でいずれかのサーバに分散配置させて能力向上を図る方法があるが、この方法では特定のビデオプログラムに要求が集中した場合に結局 1 台のサーバだけが対応することになり複数台にした効果が得られない。分散 RAID 方式では図 1 に示されるように、1 つのビデオプログラムが複数台のサーバにストライピングされて分散配置されるので、上記の場合でも複数台のサーバが対応することになり、サーバ台数に比例したスケラブルな性能向上が実現出来る。特にビデオデータの場合には各ストリーム毎に連続的にデータがアクセスされるため、各サーバへのアクセスの均等化が確実になる。このように複数台のサーバをまとめて 1 台のビデオサーバを構成することになるので全体としての故障率

が上昇するという問題があるが、分散 RAID 方式においても既存の RAID ディスクシステムと同様にデータにパリティを持たせて RAID4 型、RAID 5 型を構成し、1 台のサーバが故障してもデータ無喪失で処理を続行できる高信頼化が可能である。

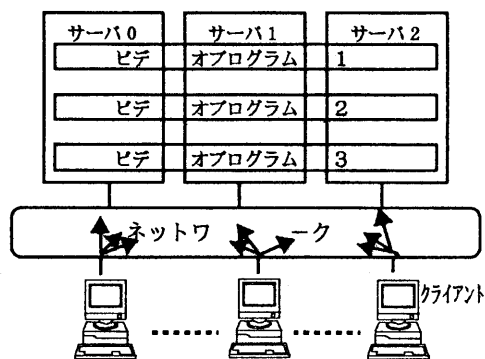


図 1 分散 RAID 方式ビデオサーバの概念図

マルチプロセッサシステムにおいてプロセッサ数に比例した性能向上はなかなか難しいが、分散 RAID 方式ではほぼサーバ数に比例した性能向上が図られる。特に普及型の廉価なサーバを追加していくことによりリニアに性能向上が図れる点は大きな魅力である。

分散 RAID 方式は我々が独自に考案したもののだが、調査の結果カリフォルニア大学で Swift<sup>(2)</sup>、<sup>(3)</sup> という類似の研究が行われていることが分かった。以下に述べる我々の研究では、Swift ではサーバ当たりのディスク台数が 1 台であったがこれを複数台としそのアレイ構成法を提案したこと、サーバの縮退 (故障隔離) 以外にディスクの縮退も出来るようにして可用性を高めたこと、ストライピングバッファを設けて縮退時の性能向上を図ったこと、ストライピング機能をサーバ側ではなくクライアント側に持たせる身軽な方式を実現したこと等の新たな試みを行った。さらに Swift では通常のデータの転送性能の測定だけで映像データ試験は行われていないが、我々の試作では実際のビデオデータを流して正常な映像が得られることを確認し、コマ落ち率測定プログラムによる

確認も行った。又サーバに疑似故障を発生させても、ビデオ映像には影響を与えずサーバの切り離しが瞬時に行われることも確認した。

### 3. RAID0型の試作評価

#### 3.1 実現方式

本章ではRAID0型の分散RAID方式ビデオサーバの試作結果について述べる。RAID0型ではビデオデータを単純に複数のサーバに渡ってストライピングしてスケラブルな性能向上を図る。図2はその実現手法を示した図である。

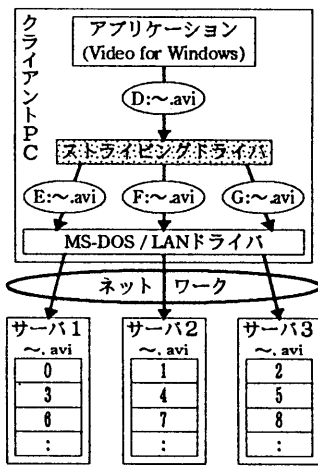


図2 RAID0型の実現方式

図2に示されるようにRAID0のストライピング機能はクライアント側で実現しており、サーバ側には手を加えていない。図中のストライピングドライバが今回作成した部分である。図2で、D:上の1つのファイルを16KBの固定長のブロックに分割してE:、F:、G:に順番に格納しておく。図2でサーバ上に0,1,2,3,4...で示されているのがそれであり、それぞれサーバのディスクに格納される。Video for Windows等のアプリケーションがD: ~.aviのブロック0,1,2へのREADを発行すると、ストライピングドライバがそれをE: ~.aviのブロック0、F: ~.aviのブロック1、G: ~.aviのブロック2へのREADに置き換えて

実行する。これを繰り返していくことにより、3台のサーバにストライピングされたビデオデータを順次読み出すこと出来る。ストライピングドライバはMS-DOSのTSR(Terminate and Stay Resident)として実現した。図3、図4はこの機構を示したものであり、図4にはアプリケーションからDOSへのシステムコールINT21をTSRがフックする様子が示されている。

なおVODではREAD性能が重要であるため今回はREAD機能のみ実現し、ビデオデータの格納は別プログラムを作成して実施した。

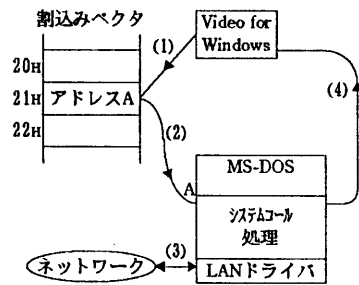


図3 MS-DOSのシステムコール

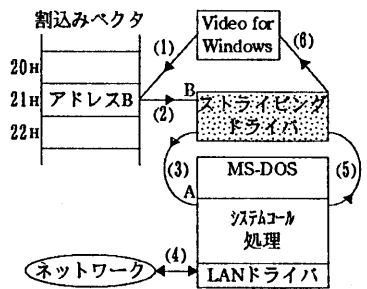


図4 ストライピングドライバの挿入

#### 3.2 システム構成

図5はこの試作システムの構成を示した図である。LANとしてイーサネットの10倍の性能を持つ100VG-AnyLANを用い、サーバにはWindowsNT3.5を搭載したパソコンサーバ3台をこれにつなげた。各サーバにはそれぞれディスクを3台接続し、ソフトウェアストライピングを施した。これは図

2で見た場合、図中の1つのサーバのディスクが実は3台のディスクがソフトウェアストライピングされて構成されていると考えることができる。クライアントとしては、Windows3.1搭載パソコンを6台接続した。

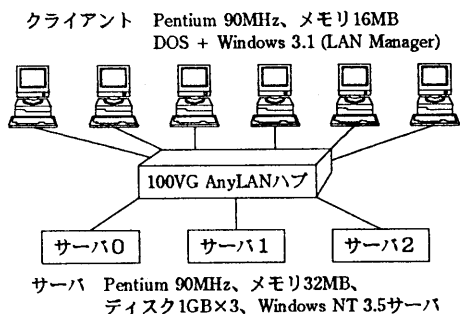


図5 試作機のシステム構成

### 3.3 性能評価

#### 3.1 概要

性能評価には Video for Windows に付いてきた AVI 形式の、1.2Mbps の約5分の長さのビデオデータを使用した。1台のクライアントではビデオデータをサーバ群から読み出して放映し、コマ落ち測定を行うが、他のクライアントでは疑似プログラムを走らせて、1つのクライアントからあたかも多数のクライアントがアクセスしているようにサーバ群に対して読み出しを行うようにし、クライアント数の不足を補った。性能のボトルネックになるものとして、サーバのCPU能力、サーバ内部バス、サーバのディスク系、LANが考えられるが、以前に行った実験<sup>(1)</sup>からサーバのディスクが3台のこの構成では、サーバのCPUネットワークが最もクリティカルであることが判っているため、サーバのCPU使用率を縦軸にして性能測定を行った。この結果が図6である。

図6はサーバ数がそれぞれ1、2、3台の3つのケースで、クライアント側に供給される総ストリーム数に対してサーバ側のCPU使用率を測定したものである。各ケースにおいて最初のストリーム数=1の所の点は、映像を放映する1台のク

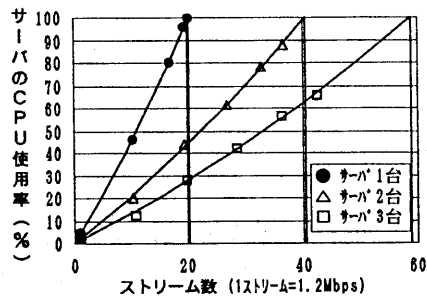


図6 RAID0型の性能評価

ライアントだけが動作した場合である。その次の点は疑似プログラムが動くもう1つのクライアントが動作した時であり、同様にしてクライアント数を増やしていったものである。実測点を結ぶ線は2次曲線で近似したものである。サーバ数=1のグラフにおいて、最初の疑似プログラム・クライアントの追加で9.25ストリーム分増加しており、これが疑似プログラム・クライアント1台の読み出し能力と言える。しかるに次は増加分が6.4ストリーム、2.84ストリーム、0.68ストリームと減少している。このようにリニアに増加しない原因は、サーバCPUのビジー率が増え、クライアント側の要求が待たされることが増えて、疑似プログラム・クライアントからの読み出しデータ量が減少するためと考えられる。

#### 3.2 コマ落ち率の測定

表1は図6の測定点に対応した、放映クライアントでのコマ落ち率を示したものである。コマ落ち率が1%以上になるとハッキリコマ落ちに気付くが、表1ではほとんど0%であり、多いところでも0.04%となっている。これから1~3サーバの全てに渡って非常に良好な映像が供給出来ることが確認できた。

なおサーバ数=1のケースでCPU使用率=100%の所を注目すると、ここでもコマ落ち率=0%となっている。即ちサーバのCPUビジー率が100%であっても、クライアント側にコマ落ち率0%の良好な映像が転送出来るという注目すべき結果も得られた。

表1 コマ落ち率の測定

サーバ数	ストリーム数	コマ落ち数 (5128フレーム中)	コマ落ち率
1	1.0	0	0.00 %
	10.25	0	0.00 %
	16.65	0	0.00 %
	19.49	0	0.00 %
	20.17	0	0.00 %
2	1.0	0	0.00 %
	10.4	0	0.00 %
	19.4	0	0.00 %
	26.9	0	0.00 %
	32.9	0	0.00 %
	36.7	2	0.04 %
3	1.0	0	0.00 %
	10.9	1	0.02 %
	19.8	0	0.00 %
	28.6	0	0.00 %
	36.4	0	0.00 %
	42.8	1	0.02 %

### 3.3 LAN上の競合の問題

LAN を使ってデータ転送を行う主体（イニシエータ）は、回数、データ量共にクライアント側よりもサーバ側が圧倒的に多い。近似的にサーバだけがLANのイニシエータであるとする、1サーバの時にはLAN上でデータ転送がぶつかることはないが、2サーバ、3サーバの時はサーバ間でぶつかりが起き、片方に待ち時間が出来るため性能低下を来す可能性がある。この条件が最も厳しいのは3サーバで6クライアント（42.8ストリーム）の場合であるが、表1からそのコマ落ち率は0.02%と良好であり、少なくともこの時点までは上記による性能低下は起きていないことが確認出来た。

### 3.4 分散RAID方式の効果

最後に最も重要な結果について述べる。前述のようにサーバCPU ネット環境のため、図6においてサーバ側のCPU能力を100%使い切った時

のストリーム数が、そのビデオサーバの最大供給能力と見ることが出来る。図6からサーバ数1台、2台、3台の時のビデオストリーム数供給能力は、それぞれ20.2、40.7、59.0ストリームである。その比は1:2.01:2.92であるが、これはサーバ数の比1:2:3に近いものであり、RAID0型分散RAID方式ビデオサーバの目指すところの、サーバ数に比例した性能向上が達成されていることが確認出来た。

## 4. むすび

分散RAID方式ビデオサーバの提案を行い、この内のRAID0型について、試作及び評価を行った。性能評価の結果、サーバ数に比例したリニアな性能（供給可能ビデオストリーム数）向上が図られることが確認出来た。（Swiftの例ではこのようにきれいに比例していない。）また複数のサーバがイニシエータとして1つのLANを競合してデータ転送することによる性能低下は、我々が観測した範囲（42ストリーム）では（サーバ側において）起きていないことが確認出来た。又、Swiftでは分散RAIDにおいて、通常データの転送性能を測定したにとどまっているが、我々は実際に動画を表示しながら確認し、かつコマ落ち率測定プログラムにより画面の品質の測定も並行して行ったため、本方式のフィージビリティの確認としては、より説得力があるものと考えられる。また、この時サーバのCPUの使用率が100%になるまで良質の動画が供給出来るという興味深い評価結果も得られた。以上得られた評価結果は、分散RAID方式ビデオサーバの実現性を大きく確実にするものと考えられる。

## 参考文献

- [1] 中村, 山口, 峯村, 渡辺, 水野, “ビデオストリーム配信性能の一検証”, 情報処理学会研究報告 95-DPS-72, p.37 ~ 42, Sep.1995.
- [2] L.F. Cabrera, and D.D.E. Long “Swift: Using Distributed Disk Striping to Provide High I/O Data Rates,” Computing Systems, vol.4, no.4, Fall, 1991.

- [3] D.D.E. Long, and B.R. Montague, "Swift/RAID: A Distributed RAID system," *Computing Systems*, vol.7, no.3, Summer, 1994.
- [4] C. Federighi and L.A. Rowe, "A Distributed Hierarchical Storage Manager for a Video-on-Demand System," *IS & T/SPIE*, 1994.
- [5] J.H. Hartman, and J.K. Ousterhout, "The Zebra Striped Network File System," *Proceedings of the 14th ACM Symposium on Operating Systems Principles*, 1993.
- [6] F.A. Tobagi, J.Pang, R.Baird, and M.Gang, "A Disk Array Management System For Video Files," *ACM Multimedia 93*, Aug. 1993.
- [7] D.J.Gemmell, H.M. Vin, D.D. Kandlur, P.V Rangan, and L.A. Rowe, "Multimedia Storage Servers: A Tutorial," *Computer*, May. 1995.