

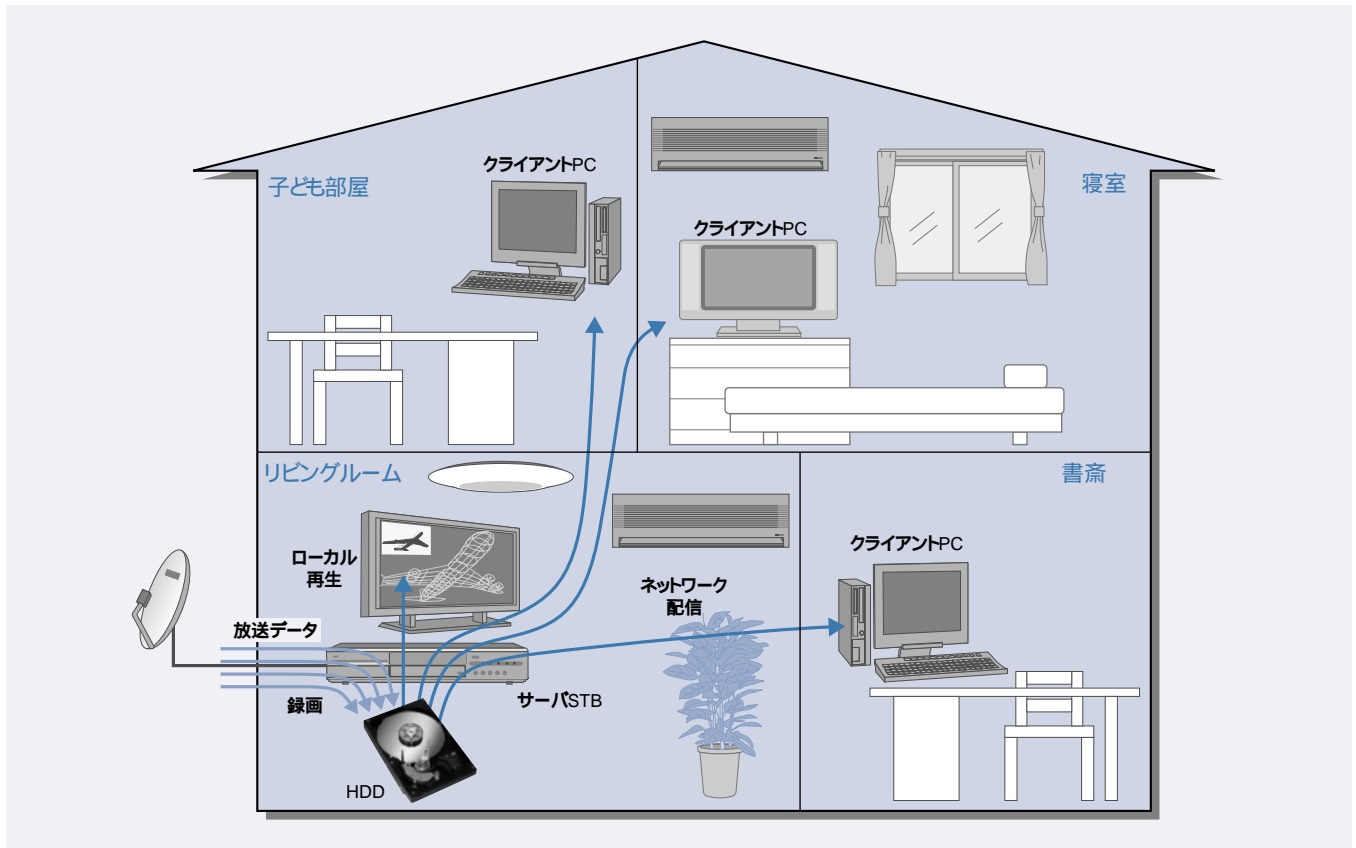
マルチハイビジョン時代に最適な AV用ファイルシステム

A Suitable AV File System for Digital Hi-Vision Contents

小日向宣昭 Nobuaki Kohinata

レモアル ダミアン Damien Le Moal

水谷美加 Mika Mizutani



注:略語説明 HDD(Hard Disk Drive), STB(Set Top Box)

マルチルーム対応サーバSTB

地上デジタル放送の開始やホームネットワークの普及に伴い、サーバSTBのHDDは複数同時の読み書きに対応しなければならない。サーバSTBの部品コストを上昇させることなく性能を上げるには、ハイビジョンコンテンツやHDDの特徴を考慮したファイルシステムを使用することが必要である。

大容量HDDと高速ネットワークI/Fを搭載したAV機器の普及に伴い、今後は全録画したコンテンツを宅内のどこでも自由に楽しみたい、という新しいコンテンツ視聴スタイルが主流になると考えられる。このスタイルに対応するファイルシステムには、コンテンツを格納するHDDへの書き込みと、HDDからの読み出しにおけるアクセスレートを保証しながら高効率に処理することが求められる。

日立製作所は、この課題を解決するミドルウェアソリューションとして、複数のハイビジョンコンテンツへの同時アクセスを可能とするAV用ファイルシステムを開発し、製品化を進めており、HDDの付加価値を向上するミドルウェアの開発を中心に、新しいデジタルライフを支える基盤技術の研究開発を強化、推進している。

1 はじめに

大容量HDD(Hard Disk Drive)とネットワークI/F (Interface) を搭載したAV(Audio-Visual) 機器の普及と地上デジタル放送の開始に伴い、今後ユーザーのコンテンツ視聴スタイルが変化していくと考えられる。例えば、

全チャンネルの番組をHDDに録画しながら、すでに録画したハイビジョンコンテンツを宅内で自由に視聴するといった形態である。これを実現するには、複数の読み書きをHDDへ同時に高効率で処理するとともに、各アクセスのレートを保証する仕組みが必要である。

日立製作所は、このような課題に対応し、HDDの付

加価値を向上させるためのモデルウェアソリューションの提案に取り組んでいる。

ここでは、部品コストの上昇を抑えながら複数の映像コンテンツデータの同時処理を可能とする高性能なAV用のファイルシステムと、その適用事例と効果について述べる。

2 ファイルシステムへのニーズ

ファイルシステムは、対象とするシステムによって多様に解釈することができることから、ここではAV機器やSTB (Set Top Box) といった組み込み機器が備えるローカルなHDDにデータを記録したり、記録したデータを読み書きしたりするためのデータ管理方法、もしくはそれに関する一連の概念やソフトウェアプログラムを指すこととする。

このようなファイルシステムは、一般にOS (Operating System) の1機能として提供されており、現在、各種OS用にさまざまなファイルシステムが存在している(表1参照)。

本来、パソコンやサーバが備えているファイルシステムの機能には、高速にアクセスできる揮発性メモリに比べて低速ではあるが、HDDのように大容量かつ不揮発な記録デバイスに確実にデータを記録し、管理させるという基本的ニーズがある¹⁾。

そして、パソコンやサーバ以外のさまざまな組み込み機器にもHDDが利用されるようになり、ファイルシステムへのニーズが多様化してきている。例えば、不測の電源障害などでシステムが正常終了できなかった時でも、データの管理情報(メタデータ)の一貫性が保証できることや、そのメタデータが高速に検査、修復できることなど、運用の容易性が求められている。

表1 各種OS用ファイルシステム

下記以外にも光ディスクやフラッシュメモリといった記録メディア用ファイルシステムがある。

OS	ファイルシステム	開発元
MS-DOS ^{*1}	FAT16	Microsoft
MS-Windows ^{*1}	FAT32	Microsoft
	NTFS	Microsoft
Mac OS ^{*2}	HFS+	Apple
Linux ^{*3}	Ext2FS	Remy Card
	Ext3FS ²⁾	Stephen Tweedie
	ReiserFS ³⁾	Namesys
	JFS ⁴⁾	IBM
	XFS ⁵⁾	SGI

注:略語説明ほか FAT(File Allocation Table), NTFS(NT File System)
HFS+(Hierarchical File System Plus)
Ext2FS(Second Extended File System)
Ext3FS(Third Extended File System), JFS(Journaled File System)

*1 MS-DOS, Windowsは、米国およびその他の国における米国Microsoft Corp.の登録商標である。

*2 Mac OSは、米国Apple Computer, Inc.の登録商標あるいは商標である。

*3 Linuxは、Linus Torvaldsの米国およびその他の国における登録商標あるいは商標である。

表2 代表的な映像コンテンツのデータサイズ

120分のHDコンテンツでは約18 GBのディスク容量を消費する。

コンテンツの種類	平均ビットレート	動画時間	データサイズ
SD	8 Mbps	120分	7.2 GB
HD	20 Mbps	120分	18 GB

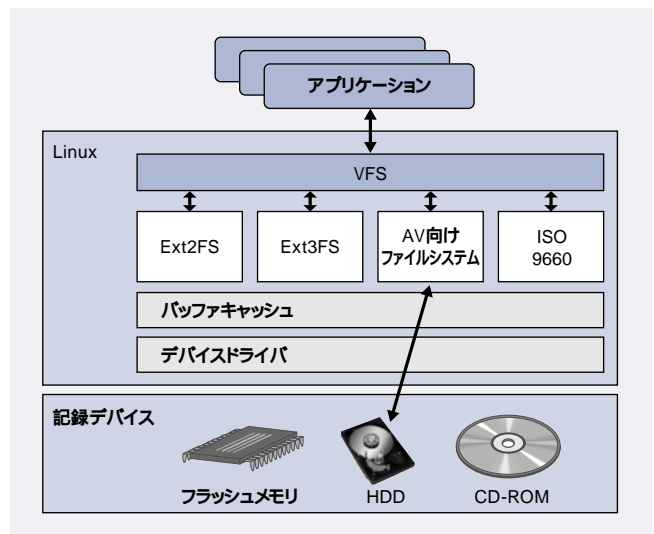
注:略語説明 SD(Standard Definition), HD(High Definition)

今後は数百Gバイトクラスの大容量HDDと高速ネットワーク/Fを搭載したAV機器の普及によって、いったん、全番組を録画しておき、あとはユーザーが好きな時刻に好きな場所でコンテンツを視聴するというスタイルが普及すると考えられる。このスタイルに対応し、AV機器のコストを抑えつつ、HDDが一定のアクセスレートで複数同時の読み書きを効率よく処理するには、アプリケーションに指向したファイルシステムが必要になってくる(表2参照)。

3 AV用ファイルシステムの開発

3.1 実装方針

日立製作所は、ファイルシステムへの新たなニーズに応えるため、AV用のファイルシステムを開発し、製品化を進めている。AV用ファイルシステムは、AV機器やSTB用のOSとして広く使用されているLinuxをターゲットとし、VFS(Virtual File System)の下にExt3FS (Third Extended File System) などと同じ階層に実装する(図1参照)。これにより、例えば、1台のHDDにおいて、番組表や録画したコンテンツに関する付加情報はExt3FSのボリュームを、コンテンツ本体はAV用ファイル



注:略語説明 VFS(Virtual File System)
IS(International Organization for Standardization)

図1 AV用ファイルシステムの位置づけ

AV用ファイルシステムは、アプリケーションからファイルシステム関連のシステムコールを呼び出すことによって、他のファイルシステムとまったく同様に利用することができる。

システムのボリュームを利用するといった使い方ができる。また2.5型や3.5型といった形態やディスク容量，メーカーを問わずさまざまなHDDに適用することができる。

3.2 開発目標

(1) 性能

映像コンテンツに恒常的に高いアクセス性能を得るには、それが細かく分割(フラグメンテーション)して記録されないようにする必要がある。これはHDDの特性において、ファイルのフラグメンテーションがHDDへのI/O(Input-Output)フラグメンテーションを引き起こし、その結果として、ディスクヘッドのシーク動作を増大させてしまうからである。

また、映像コンテンツが、ディスク外周部より先アクセス速度が遅い内周部だけに偏って記録されないようにする必要もある。

(2) 運用性

不測の電源遮断といった異常終了に際して、高速なファイルシステムの検査、修復を実施し、ユーザーにその一連の作業時間を意識させないようにするには、メタデータに対するジャーナリング(ordered data mode)をサポートする必要がある。

例えば、ジャーナリング機能を持たないExt2FSでは、システムクラッシュ後の再起動時にfsckコマンドによるファイルシステムの検査と修復が実施されるが、HDDの大容量化に伴い、その時間が増大し、結果的にシステム起動までに多大な時間を要してしまうことになる。しかし、メタデータに対するジャーナリングのサポートにより、すばやくシステムを回復し、起動することができる。

(3) スケーラビリティ

ユーザーの利用環境に応じて性能や運用性を高めるために、さまざまな設定ができる必要がある。例えば、自由にメタデータを追加で定義できるファイル管理機能や、ワークロード(データの種類:テキスト、映像など)に最適なデータブロックサイズの値などを変更できることである。

3.3 AV用ファイルシステムの特長

開発目標に基づき開発したAV用ファイルシステムには、以下のような特長がある。

- (1) 最大ファイルシステム(パーティション)サイズは4,100 Pバイト(データブロックサイズが4 Mバイト、メタデータブロックサイズが4 kバイトの場合)
- (2) コンパクトなメタデータ構成により、メタデータへの高速アクセスとディスク消費の削減が可能
- (3) データブロックサイズは4 kバイトから4 Gバイトまで選択が可能
- (4) リアルタイム性を考慮したデータ配置が可能

- (5) メタデータジャーナリングをサポート
- (6) 容易に機能拡張可能なアーキテクチャ

3.4 モジュール構成

特長の一つである柔軟な拡張性により、コアオペレーションモジュールにさまざまな機能をプラグインすることができる(図2参照)。

(1) コアオペレーション

ファイルシステムとしてのメイン処理を行うモジュールであり、ジャーナリングや各種管理(メタデータブロック、ディスク、ボリューム、iノード、ファイル、I/Oなど)機能を提供する。

(2) システムラッパ

OSに依存する部分を吸収するところであり、HDDに発行されるすべてのI/O処理をサポートする。

(3) ボリュームオペレーション

メタデータの構成に付加的な情報を与えることができるプラグインモジュールである。例えば、コンテンツの著作権に基づいたアクセス権やアクセスの優先度を定義することができる。

(4) ブロックオペレーション

データブロックの割当アルゴリズムを提供することができるプラグインモジュールである。

(5) I/Oスケジューラ

例えば、ボリュームオペレーションによって定義されたアクセス優先度を用いて、ファイルI/Oの優先順位を決定するために使用する。

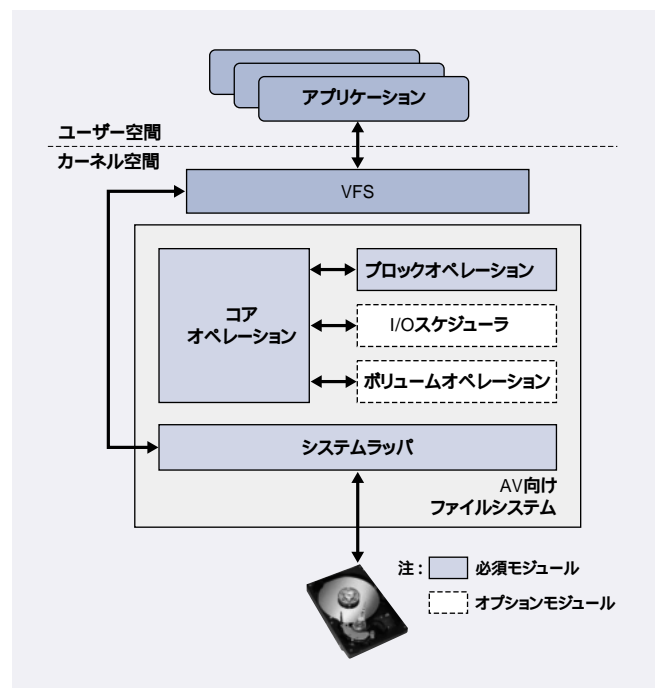


図2 AV用ファイルシステムのモジュール構成
OS依存部分はシステムラッパで吸収している。このモジュールを変更することにより、Linux以外のOSにも容易に移植することができる。

4 AV用ファイルシステムの適用例と効果

4.1 マルチルームシステム

AV用ファイルシステムでフォーマットしたHDDに数種類のハイビジョンコンテンツを格納し、ローカルでの再生に加えてネットワークを介して同時に複数個所で視聴できるシステムである。

これにより、ローカルでの再生に加えて3本のネットワーク配信、さらにはローカル配信先での倍速再生といった要求にも対応することができる。

4.2 性能結果

マルチルームシステムの各クライアントにおいて、AV用ファイルシステムを適用したHDDから、同時に最大何本までハイビジョンコンテンツを読み出せるかを目視によって測定した(図3参照)。Ext2FSと比較し、HDD上の配置場所によらず、安定して最大10本のハイビジョンコンテンツが読み出せ、2倍以上の性能向上が得られることを確認した。

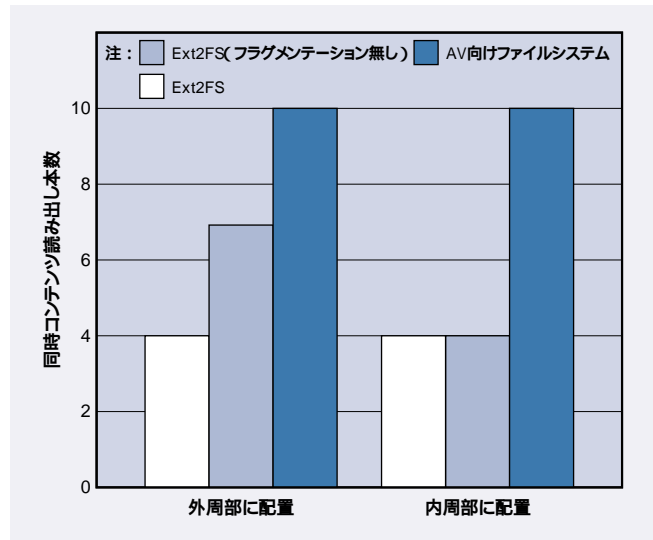


図3 AV用ファイルシステムの性能(読み出し)
25 Mビット/sのHDコンテンツをHDDの外周部と内周部にそれぞれ配置して測定した。

5 おわりに

ここでは、ハイビジョンのような高精細な映像コンテンツの特徴を考慮し、効率よくHDDを活用するためのAV用ファイルシステムについて述べた。このファイルシステムによって、複数同時にハイビジョンコンテンツを取り扱うことができる。

日立製作所は、今後も、HDDの付加価値を向上させるミドルウェアの開発を中心に、新しいデジタルライフを支える基盤技術の研究開発を推進していく考えである。

参考文献など

- 1) Dominic Giampaolo, Practical File System Design(1999)
- 2) Stephen Tweedie, Journaling the Linux ext2fs Filesystem, LinuxExpo '98(1998)
- 3) ReiserFSホームページ, <http://www.namesys.com>
- 4) JFSホームページ, <http://jfs.sourceforge.net>
- 5) Project XFSホームページ, <http://oss.sgi.com/projects/xfs/>

執筆者紹介



小日向宣昭

1996年日立製作所入社、システム開発研究所 情報サービス研究センター 第六部 所属
現在、情報家電のネットワークミドルウェア、ファイルシステムの研究開発に従事
E-mail: kohinata@sdl.hitachi.co.jp



レモアル ダミアン

2000年日立製作所入社、システム開発研究所 情報プラットフォームセンター 第三部 所属
現在、OS、ファイルシステム、およびストリーミング用ミドルウェアの研究開発に従事
E-mail: damien@sdl.hitachi.co.jp



水谷美加

1987年日立製作所入社、システム開発研究所 情報サービス研究センター 第六部 所属
現在、情報家電向けネットワーク・HDD向けミドルウェアソリューションの研究開発に従事
情報処理学会会員
E-mail: mizutani@sdl.hitachi.co.jp