

衛星画像データを対象とした階層ファイルシステムの実装

2F-9

迫 和彦[†] 高橋 一夫[‡]喜連川 優[‡] 高木 幹雄[†][†]: 東京大学生産技術研究所[‡]: 株式会社 ウェザーニュース

1 はじめに

莫大な衛星画像データを対象とした地球環境データを扱うことを目的とし、磁気テープジュークボックスなどの三次記憶装置を用いた超大容量アーカイブファイルシステムを構築している。現在、本システムでは、莫大なデータ量を持つ NOAA 気象衛星からのデータを柔軟に利用すべく効率の良い階層記憶制御方式の確立を目指している。

本稿では、気象衛星 NOAA の画像データを効率良く利用するための『部分マイグレーション』を用いた階層ファイルシステムについて報告する。

2 システム設計

NOAA の画像データは原画像 1 シーンで約 100MB と非常に大きく、ユーザが実際に処理に利用するのは、この莫大なデータのうちの一部分であることが多い。また、ユーザが利用するのはある程度連続した部分であることが多い。

従来の階層記憶ファイルシステムでは、ファイル単位のマigreション機能しか支援されていない。これに対し、NOAA 気象衛星画像の利用形態において効率的に階層記憶管理を行うために、本システムではユーザが必要とする部分のみをアーカイブから二次記憶装置に取り出すブロック単位の『部分マイグレーション』を実現することを試みる。これにより、目的の画像データが二次記憶装置にない場合のマigreション時間の短縮、及びワーキングセットの縮小にともなう二次記憶装置の効率的な利用が期待できる。また、NOAA 気象衛星画像は連続した画像ブロックにアクセスされる可能性が高いため、ブロックの先行マイグレーション機能を導入し、更に高性能化を図ることとした。

3 試作アーカイブファイルシステムの構築

アーカイブファイルシステムは、デバイスによるアクセス手段の違いを吸収しマイグレータに対して共通のアクセス手段を提供するデバイスドライバ、階層記憶管理を行うマイグレータ、ユーザにファイルへのアクセス手段を提供する I/O ライブラリから構成される。

マイグレータは、UFS (Unix File System) のファイルとして存在し、ユーザにデバイスを意識させない

Implementation of Hierarchical File System for Satellite Imagery Data

Kazuhiko Sako, Masaru Kitsuregawa and Mikio Takagi

Institute of Industrial Science, University of Tokyo

Kazuo Takahashi

Weathernews, Inc.

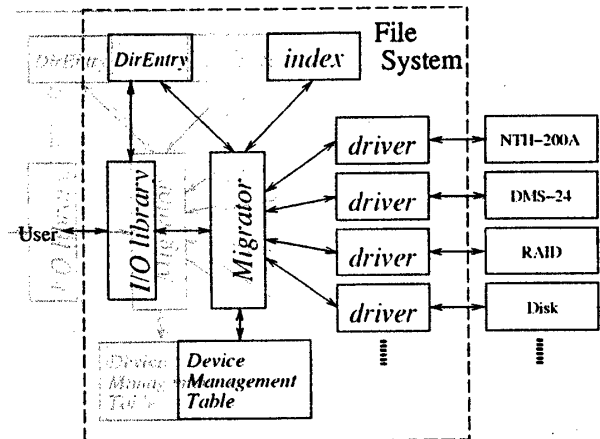


図1: 試作アーカイブファイルシステムのソフトウェア構成

ランスペアレントな機能を提供するためのディレクトリエントリ、ファイルシステム上のファイルを幾つかのブロックに分割して管理するためのインデックス、そして各デバイスの利用状況を管理するためのデバイス管理テーブルを利用して階層記憶管理を行う (図1)。

ファイルをリードする際のマイグレータの動作は次のようになる。

1. I/O ライブラリからファイルのリードリクエストを受けとる
2. ディレクトリエントリの内容からインデックスを参照し、目的の画像ブロックの格納デバイスデバイス位置を検索
3. 目的の画像ブロックが二次記憶装置上にあるならば I/O ライブラリにそのブロックアドレス情報を返す
4. ない場合は、その画像ブロックのマigreションを開始し、マigreションが終了すると画像ブロックの格納情報を返す
5. ユーザは I/O ライブラリを通じてデータを読み込む。この間マイグレータは次の画像ブロックのマigreションを行う

4 ファイルシステムの実装

SPARC station 10/41 (SunOS 4.1.3) 上で、二次記憶装置として SCSI Disk を使い、アーカイブとして Exabyte8500 ドライブが 1 ユニットに 2 台搭載されている NTH-200A 8mm テープジュークボックスを用い、試作ファイルシステムを実装した。

8mm テープアーカイブは、read/write 時に与えられる論理ブロック番号からテープ番号とその位置に変換し、RS232C 経由でロボティクスをコントロールして、目的

の位置にアクセスするドライバを作成することでファイルシステムに組み込む。データはあらかじめ2つのテープにストライピングしておき、同時に読み込むことにより、read/write速度を2倍の約1MB/secに高速化した。

	device driver	time(sec)
(a)	SCSI driver(mt command)	288
(b)	SCSI target driver	70

表 1: テープ頭出し性能

また、Exabyte8500をコントロールするのに、OS標準のSCSI テープドライバを用いず、新たにSCSIターゲットドライバを開発することにより、Exabyte8500の持つ75倍頭出し機能を十分に引き出すことを可能とした。

表1はテープの先頭からの頭出し性能を示す。測定は100MBのファイルが40シーン入れている8mmテープにおいて、21シーン目に頭出しを行う時間を測定した。ファイルマーク単位の読み出しを基本とする従来の方式に比べ、大幅な高速化が実現されていることがわかる。

5 ファイルシステムの性能評価

試作ファイルシステム上でアプリケーションプログラムを実行し、その実行時間を測定した。実行条件を表2に示す。ここでは頻繁に利用される処理として、放射能補正、幾何補正を行い、関東近郊の5°×5°の画像を抽出するプログラムを取り上げることにした。この処理は、必ずしもすべての画像を走査する必要はなく、部分マイグレーションが有効であることが予想される。すなわち、4297ラインの画像から621ラインを含むブロックのみを部分的にマイグレートすることを可能としている。

性能評価に於いて比較のため、通常用いられる方式としてtarで展開した後ディスク上のファイル进行处理する場合の時間を表3に示す。なお、8mmテープ上のデータをディスク上にダウンロードせずにアプリケーションを実行する手法も考えられるが、一般的な利用形態とは考えられないため、本稿では取り上げない。

画像データ

原画像サイズ	95307460bytes(4297lines)
ファイルシステム	
画像ブロックサイズ	1Mbytes
ディスクブロックサイズ	256Kbytes
8mm ブロックサイズ	256Kbytes
アプリケーション	
入力ライン	621 lines (from 1810th line)
入力サイズ	22180bytes×621
出力画像	5°×5°(512×512 pixels)
利用画像ブロック	14blocks(14MBytes)

表 2: アプリケーション実行条件

		time(sec)
(c)	down load	232.24
(d)	application	74.07

表 3: tar を用いたアプリケーションプログラム実行時間

		time(sec)
(e)	on disk	75.10
(f)	on 8mm (non-prefetch)	97.51 + 10
(g)	on 8mm (1 block-prefetch)	83.70 + 10

表 4: 試作アーカイブファイルシステムを用いたアプリケーションプログラム実行時間

表4に、アーカイブファイルシステムを用いた場合の実行時間を示す。目的の画像ブロックがディスク上に存在する場合の実行時間には約75秒、ブロックがミスヒットし、8mm ジュークボックス上のデータのマイグレーションを行った場合は約98秒、更に1画像ブロックだけプリフェッチをした場合には84秒という計測結果が得られた。ただし、ファイルの先頭から利用される画像ブロックへのテープのシーク時間としてさらに10秒程要する。本アーカイブファイルシステムはブロック単位でディスク上にマイグレーションを行い、当該ブロックのマイグレーション終了と同時にアプリケーションに返される。そのため、当該ブロックの処理と次ブロックのマイグレーションが並行に行われ、実行時間が短縮される。

ディスクにキャッシュされていない場合、ロボティクスによるテープの入れ換え時間は共通であることから考慮しないと、一般的な手法では(a)+(c)+(d)の時間が掛かり、アーカイブファイルシステムでは、(b)+(g)の時間で終了することとなり、大幅な高速化が達成された。

6 まとめ

アーカイブファイルシステムの実装し、NOAA 気象衛星画像を処理するアプリケーションプログラムを用いてその性能を測定し、有効性を確認できた。すなわち(1)専用のデバイスドライバの開発によるテープの頭出しの高速化(2)2-way テープストライピングの採用(3)ブロック単位のマイグレーションの導入(4)ブロックプリフェッチの採用により従来の階層記憶システムに比べ大きな性能向上を達成可能なことをNOAA アプリケーションプログラムを実際に走らせることにより明らかにすることが出来た。ディスク上にある場合と比べても大差のない処理時間を得ることが出来た。今後、より多くの応用を用いることにより、更に詳細な評価をすすめてゆく予定である。

参考文献

- [1] 高橋一夫, 喜連川優, 高木幹雄, “衛星画像の格納を目的とした超大容量アーカイブデータベースシステム” 情報処理学会第47回全国大会講演論文集