

マルチプラットフォーム対応ディスクシステムにおけるデータ交換方式

2Q-1

北村 学, 大枝 高

(株)日立製作所システム開発研究所

1. はじめに

メインフレームと、UNIX などのいわゆるオープンシステムとが混在する複合構成によるシステムが増加しており、メインフレームのレガシーデータをオープンシステムで利用したいという需要は多い。今回、メインフレームとオープンシステムの両方で使用出来る「マルチプラットフォームディスク装置」と、それによるデータ交換機能を開発した。

2. ディスクシステムの構成

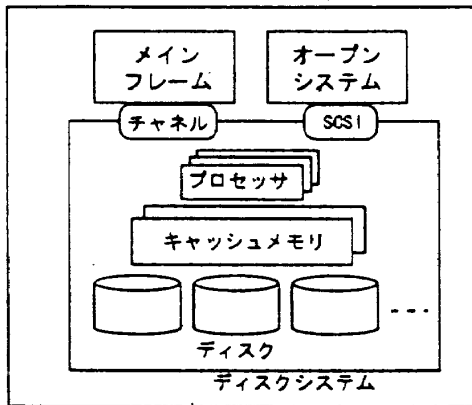


図1ディスクシステムの構成

図1にディスクシステムの構成を示す。ディスクにはSCSIディスクを使用し、複数のディスクを組にしてRAID構成を取る。ホストとのインタフェースにはチャンネルとSCSIが用意され、チャンネルからのCKDデータ<sup>[1]</sup>は、ディスクシステム内でCKD形式データを独自の方式でFBA形式に変換して格納し、従来からあるメインフレームのディスクをエミュレートする機能を有する。各論理ボリュームは、チャンネルからはCKD形式のディスクが見え、SCSIからはSCSIディスクがあるように見える。また、チャンネル、SCSIのいずれからアクセスできるボリュームを定義す

ることも出来、これを共有ボリュームと呼ぶ。その場合、チャンネルから見える1つのボリュームはSCSIからは1つのLUに見える。

3. データ交換機能

本機能では、メインフレームファイル(データセット)をオープンシステムから読み書き可能にする機能を提供する。この機能をHMDE(Hitachi Multiplatform Data Exchange)と呼ぶ。

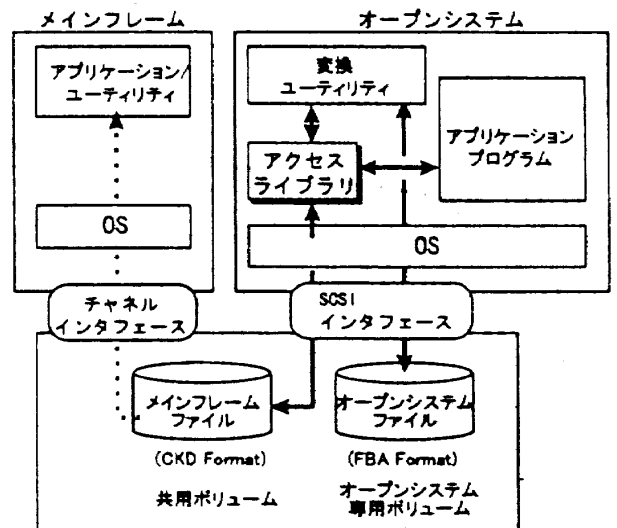


図2ファイル交換機能

オープンシステムからメインフレームファイルにアクセスする場合、ファイルシステムの解釈のほかに、CKD形式とFBA形式の相互変換が必須である。HMDEではこれらの処理を主にオープンシステム上で行わせる方式を取る。

図2はデータ交換方式の構成図である。オープンシステム上のソフトウェアとしてアクセスライブラリと呼ぶC言語のライブラリ関数と、ファイル操作をGUIベースで行う変換ユーティリティと呼ぶアプリケーションを提供する。アクセスライブラリでは共有ボリューム上のファイル管理情報(VTOC)の内容を解釈して、ファイルに対するアクセスを行う。この時同時にCKD,FBAの形

式変換も行う。本方式では、オープンシステムからメインフレームファイルにアクセスする際にメインフレームが介在する必要が無いので、データマイニングなど、メインフレーム側の大量のデータをオフラインでオープンシステム側に移行するといった用途に有効である。

#### 4. データ転送方式

データ交換機能を利用すると、メインフレームとオープンシステム間のデータ転送を行うことも可能である。

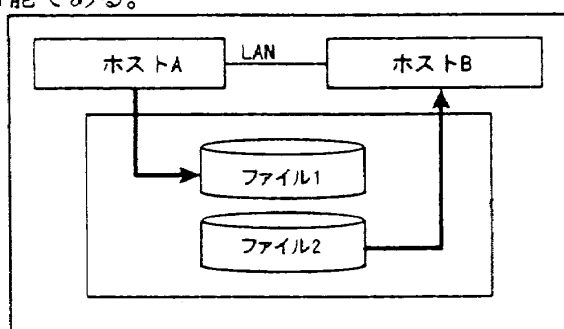


図3 データ転送の構成図

共有ボリュームにファイル1、ファイル2を用意し、中間バッファとして使用する。ホストAがファイル1に書き込んでいる間は、ホストBではファイル2を読み込む。逆にホストAがファイル2に書き込んでいる間にホストBはファイル1を読み込む、という処理を交互に繰り返し行えば、ホストAからの書き込みとホストBの読み出しが並列に行える。ただし、2つのホストが同時に同じファイルにアクセスすることがないように、ネットワークでホスト間の同期を取りながらアクセスする。これにより、ディスク装置のスループットに近い転送速度が得られると考えられる。

検証のため、図3の構成で転送速度を測定した。ホストA,BにはSun Ultra1を使用し、共有ボリュームには、あらかじめ一定サイズのメインフレームファイルを2つ作成しておく。中間ファイルのリード・ライトはアクセスライブラリを使用した。ホストA,ホストB間の同期はネットワーク

で(10Mbps Ethernet)行った。中間ファイルを繰り返し使用して、各ホストが並列に動作することの効果が観測できるように、中間ファイルよりも大きい、600MBのデータ転送を行った。

表1ディスク装置経由のデータ転送速度

中間ファイル サイズ	リード 速度 (MB/s)	ライト 速度 (MB/s)	データ 転送速度 (MB/s)
720KB	9.49	7.02	5.20
7200KB	8.08	9.50	8.77
72000KB	9.06	7.71	6.28

表1に測定結果を示す。リードのみ、ライトのみを行った時の転送速度も測定し、ホスト間データ転送との速度を比較した。ディスク装置のキャッシュ制御の影響のため、中間ファイルサイズによっては性能にばらつきがでたが、中間ファイルサイズを適切に選べば、ディスクアクセス速度と同じ速度でのデータ転送が可能になることがわかる。

#### 5. まとめ

マルチプラットフォームディスク装置のデータ交換方式を紹介し、これを利用したデータ転送方式の提案、性能測定を行った。一般にEthernetなどネットワークでは大量データ転送ではスループットがあまり上がらない[2]。大量データ転送をディスク装置経由で行う本方式は、ネットワークの欠点を補う方式として、今後様々な応用が期待出来ると考える。

#### 参考文献

- [1] C. P. Grossman, "Evolution of the DASD storage control", IBM SYSTEMS JOURNAL, VOL28, NO.2.
- [2] Donald J. Becker et al, "BEOWULF: A PARALLEL WORKSTATION FOR SCIENTIFIC COMPUTATION.", Proceedings, International Conference on Parallel Processing, 95