

DVD ライブラリアレイ装置の開発（1）

～開発コンセプトと信頼性に対する考察～

4H-8

兼田 泰典[†], 大枝 高[†], 寺岡 忠浩[†], 三部 良太[†], 伴野 洋二[‡]
[†](株)日立製作所システム開発研究所, [‡](株)日立製作所情報・通信グループ

1. はじめに

従来主に磁気ディスク装置(HDD)に適用されていたディスクアレイ(RAID: Redundant Arrays of Inexpensive Disks)技術[1]を、リムーバブルドライブを搭載したライブラリ(媒体交換機構)に適用した場合、ドライブ、ライブラリ、媒体の組み合わせにより提供できる信頼性が異なってくる。そこで本稿では、ドライブとライブラリの組み合わせによる可用性と、予備媒体と予備ドライブによるデータの信頼性について考察する。

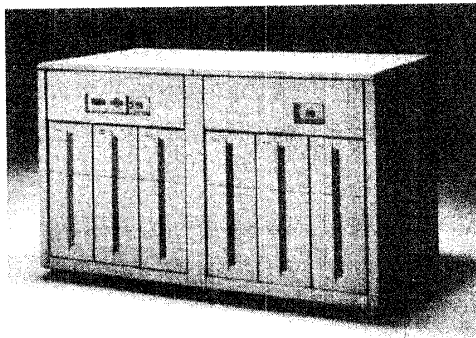


図1 DVD ライブラリアレイ装置

2. 開発のコンセプト

低ビットコストと可搬性(リムーバブル性)を特徴とするDVD-RAMをライブラリと組み合わせ、膨大なデジタルコンテンツを管理するメディアサーバやアーカイブ・バックアップサーバに最適なストレージとして提供する。さらに、エンタープライズ規模のシステムに対応する高信頼かつ高速(ハイスループット・高速アクセス)なストレージを低価格で提供するため、従来主に磁気ディスク装置に適用されていたディスクアレイ技術を、DVD-RAMドライブを搭載したライブラリに適用することにした。

Development of DVD Library Array System (1) -Product concept and study of system reliability-

Yasunori Kaneda[†], Takashi Oeda[†], Tadahiro Teraoka[†], Ryota Mibe[†], Youji Tomono[‡]

[†]Systems Development Laboratory, HITACHI, Ltd.

[‡]Information & Telecommunication Systems, HITACHI, Ltd.

3. 課題

HDDはデータを記録保持する媒体とドライブ(駆動機構と回路)が一体であったため、どちらか一方が壊れれば故障となり使用できなくなる。

しかしDVD-RAMは、媒体をドライブから取り外し可能なリムーバブルドライブであるため、媒体とドライブの障害を独立事象として処理できる。よって、ドライブが壊れた場合には媒体を他のドライブに装着することで読み書きでき、媒体が壊れた場合には新しい媒体を装着しRAIDアルゴリズムによるデータ復元を行うことができる。さらにこれら媒体の搬送を、ライブラリを用いて自動化した場合には、①ドライブ、②ライブラリ、③媒体の組み合わせにより提供できる信頼性が異なってくる。

4. 信頼性に対する考察

4.1 ドライブとライブラリの組み合わせ

表1にドライブとライブラリの組み合わせを示す。mはライブラリ数、nはドライブ数、rはRAIDを構成するドライブの台数、sは予備に用意するドライブまたはライブラリの数を示す。r台のドライブで1つのRAIDグループ(論理ドライブ)を構築する。

表1 ドライブとライブラリの組み合わせ

ドライブ数 (n)	ライブラリ数(m)			
	m=1	1<m<r	m=r	m=r+s
n=r	(1)	(3)	(4)	-
n=r+s	(2)	(3)	-	(6)
n=q·r	(1)	(3)	(5)	-
n=q·(r+s)	(2)	(3)	-	(6)

ライブラリは一つの媒体搬送機構を有すと仮定
 q: 2以上の整数

(1) m=1, n=r/n=q·r の場合

1台のライブラリ内のr台のドライブを用いてRAIDを構成した場合に相当する。この場合、ドライブが1台故障してもRAIDのパリティを使用し処理を続行できるが、ライブラリが壊れた場合には運用が停止する。システ

ムが故障し運用できなくなるまでの時間 $MTTF_{sys}$ (Mean Time To Failure)はライブラリの $MTTF_{lib}$ とドライブの $MTTF_{disk}$ により決まり、 $MTTF_{sys}$ は $MTTF_{lib}$ に等しくなる。 $m=1$ では媒体の搬送を r 回行わなければならない。比較的小規模なシステムの実現に適す。

(2) $m=1, n=r+s/n=q \cdot (r+s)$ の場合

(1)と同様であるが、ドライブが壊れた場合には予備のドライブを使用し、RAIDのパリティを使用することなく運用を続行できる。このため、(1)に比べデータを消失する確率を低減できる。

(3) $1 < m < r$ の場合

ライブラリ数が RAID を構成するドライブ数に満たない場合、壊れたライブラリに2台以上のドライブが搭載されていた場合には運用が停止する。このため、 $MTTF_{sys}$ は $MTTF_{lib}$ の $1/m$ に低下する。

(4) $m=r, n=r$ の場合

ライブラリ数が RAID を構成するドライブ数と同数な場合、ドライブ、ライブラリ、媒体のどれか一つが故障しても RAID のパリティを使用し運用を続行できる。また、媒体の搬送が複数のライブラリで同時に行えるため、媒体へのアクセスが高速になる。 $MTTF_{disk}$ と $MTTF_{lib}$ を同等と仮定すると、 $MTTF_{sys}$ は

$$MTTF_{sys} = \frac{1}{2} \cdot \frac{(MTTF_{lib})^2}{m \times (m-1) \times MTTR} \quad \dots (式1)$$

となる。ここで $MTTR$ (Mean Time To Repair)は故障箇所を修復するのに要する時間である。

(5) $m=r, n=q \cdot r$ の場合

m 台のライブラリに q 台ずつのドライブを搭載した場合に相当する。最大 q 個の RAID グループを構築できるので、並列運用により $n=r$ の場合に比べて q 倍の転送速度を実現できる。また、 q 台のドライブの内、幾つかのドライブを予備ドライブに割り当てることで(4)に比べドライブ障害時の信頼性を向上できる。 q 台のうち1台以上を予備ドライブとして使用した場合の $MTTF_{sys}$ は

$$MTTF_{sys} = \frac{(MTTF_{lib})^2}{m \times (m-1) \times MTTR} \quad \dots (式2)$$

となる。

(6) $m=r+s, n=r+s/n=q \cdot (r+s)$ の場合

予備のライブラリを用意した場合に相当する。ライブラリが故障した場合には、予備のライブラリを使用するが、まずライブラリに搭載された全ての媒体にデータを復元しなければならない。例えば1枚の媒体の復元に1時間必要な場合、100枚では約4日要することになる。媒体が多くなると、予備ライブラリによる自動復旧の $MTTR$ と、保守員による修理交換の $MTTR$ に大きな差は無く、予備ライブラリと媒体にかかるコストを考えた場合、そのメリットは小さくなる。

4.2 予備媒体の用意

媒体が壊れた場合、RAID のパリティ使用し処理を続行すると、データを消失する確率が高くなる。そこでライブラリ毎に予備の媒体を用意することで、媒体が壊れた場合には、予備の媒体にデータを復元し RAID グループを正常な状態に戻すことができる。

5. DVD ライブラリアレイ装置への適用

図1に示す DVD ライブラリアレイ装置では、ライブラリの媒体格納枚数(150枚)、故障率、コスト、および目標性能を鑑みて、構成を $r=6, m=6, n=12$ とした。予備ライブラリは設けていない。また予備媒体を各ライブラリ毎に3枚用意した。各ライブラリに搭載した2台のドライブの内、1台は予備ドライブとして使用する。装置容量は最大で $1.95TB (=2.6GB \times 150 \times (6-1))$ になる。

6. まとめ

ライブラリ数を RAID を構成するドライブ数と同数とし、ドライブ、ライブラリ、媒体のどれか一つが故障しても運用を続行することができる高可用性を提供する。さらに、予備のドライブと媒体を用意しデータの高信頼化を実現する。本考察に基づき「DVD ライブラリアレイ装置」(図1)を開発した。

参考文献

[1] David A. Patterson, Garth Gibson, and Randy H. Karz: "A Case for Redundant Arrays of Inexpensive Disks (RAID)", Computer Science Division Department of Electrical Engineering and Computer Science, University of California Berkeley.