

## RAID型ファイルシステムVAFS/HRの性能評価

6U-8

高橋 英男 山下 洋史 畠山 敦 裏谷 郁夫 山田 秀則\* 城田 浩二\* 高良 亜紀子\*  
 (株)日立製作所 \*日立コンピュータエンジニアリング(株)

### 1. はじめに

近年、ワークステーションの適用業務範囲をビジネス分野に拡大するために、信頼性を高める様々な試みがなされている。筆者らはUNIXのファイルシステム層でRAID (Redundant Array of Inexpensive Disks) を実現するVAFS/HR (Virtual Array File System/High Reliability) を開発し、ファイルシステムの高速度・高信頼性に取り組んでいる。VAFS/HRでは複数のディスクにファイルの分散格納とパリティデータの格納を行って高信頼化を図っている。本稿ではVAFS/HRのファイルアクセス性能を評価し、高速・高信頼ファイルシステムとしての有効性を検証する。

### 2. VAFS/HRの性能評価

#### 2.1 測定条件

VAFS/HRの読み出し性能と書き込み性能を、正常運転とディスク1台が故障した縮退運転について測定する。比較のためにUNIX標準ファイルシステムのUFSと、ストライピングは行いパリティデータは格納しないVAFS[1]の性能も測定する。表1に測定に用いる各ファイルシステムの構成を示す。測定は表2に示すようにシーケンシャルおよびランダムアクセスに対して行う。一般のアプリケーション・プログラムが用いるシーケンシャルアクセスと、データベースなどが用いるランダムアクセスについて性能を測定することにより典型的なファイルアクセスを網羅できる。シーケンシャルアクセスでは8MBのファイル

表1 ファイルシステム構成

項番	ファイルシステム	ディスク台数	運転モード	備考
1	VAFS/HR	5	正常	4データ+1パリティ
2	VAFS/HR	5	縮退	データディスク1台故障
3	VAFS	4	正常	4データ, パリティなし
4	UFS	1	正常	1データ

A performance evaluation of VAFS/HR - a software RAID file system

Hideo Takahashi, Hirofumi Yamashita, Atsushi Hatakeyama, Ikuo Uratani, Hidenori Yamada\*, Koji Shirota\* and Akiko Kora\*

Hitachi, Ltd.

\*Hitachi Computer Engineering Co., Ltd.

表2 アクセスパターン

	ファイルサイズ	アクセス総量	アクセス単位
シーケンシャル	8MB	8MB	32KB
ランダム		2MB	8KB

表3 使用システムコール

読み出し時	asread()+aswait()
書き込み時	aswrite()+aswait()

ルの読み出しと書き込みを行い、ランダムアクセスでは8MBのファイルに対して8KB単位で2MB分の読み出しと書き込みを行う。なお、書き込みは新規ファイルに対して行う。ファイルアクセスには表3に示すシステムコールasread(), aswrite(), aswait()を用いる。これらはVAFSで導入したシステムコールであり、asread()とaswrite()は複数のI/O要求を前のI/Oの完了を待たずに次々に発行することができ、aswait()はそれら複数のI/Oの完了待ちを一括して行う。

測定に用いるシステムの構成を図1に示し、各機器の仕様を表4に示す。

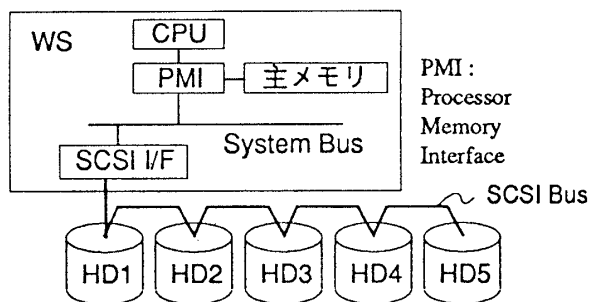


図1 測定機器構成

表4 測定機器の仕様

WS	HD
日立 3050RX	HP C2244
CPU : PA-RISC 80MHz, 主記憶容量 : 64MB, SCSI : 10MB/s	直径 : 3.5", 容量 : 566MB, 回転数 : 5400rpm

### 2.2 測定結果

シーケンシャルアクセス性能の測定結果を図2に、ランダムアクセス性能の測定結果を図3に示す。

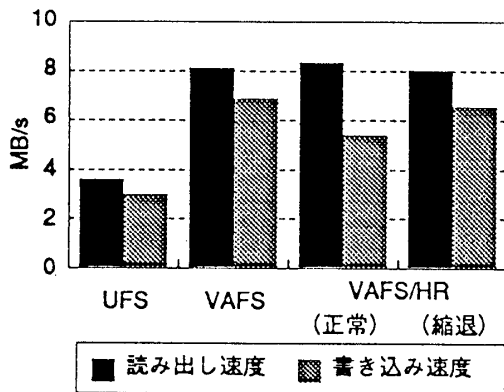


図2 シーケンシャルアクセス性能

シーケンシャルアクセスでは、VAFS/HRの読み出しで正常時に8.3 MB/s (UFSの2.4倍)、縮退時に8.0 MB/s (UFSの2.3倍)となった。これは正常、縮退共にVAFSと同等である。書き込みでは正常時に5.3MB/s (UFSの1.8倍)、縮退時に6.5MB/s (UFSの2.2倍)となった。VAFSと比較すると、正常時でVAFSの約8割、縮退時ではほぼ同等である。

ランダムアクセスでは、VAFS/HRの読み出しで正常時に1.7MB/s (UFSの1.7倍)、縮退時に1.0MB/s (UFSと同等)となった。書き込みでは正常時に0.75MB/s (UFSと同等)、縮退時に0.88MB/s (UFSの1.2倍)となった。これは、正常時の読み出しでVAFSと同等、それ以外ではVAFSの約5割の性能である。

### 2.3 考察

#### (1) シーケンシャル読み出し性能

正常時の読み出し性能は、SCSIバスの転送速度(10MB/s)がボトルネックとなるため、8.3MB/sであった。また、正常時の読み出し性能がVAFSと同等なのは、正常時にはパリティデータを読まないためVAFSとVAFS/HRで処理内容に差がないためである。

縮退時に僅かな性能低下が見られるのはデータ修復のためのxor演算処理のためである。読み出しデータ量については、各rowにつき正常時にはデータブロックを4つ読み出すのに対し、縮退時にはデータブロックを3つとパリティブロックを1つ読み出すため読み出しブロック数としては変わらず、性能に影響しない。

#### (2) シーケンシャル書き込み性能

正常時の性能がVAFSの約8割になっているのは、VAFS/HRではデータブロック4ブロックにつきパリティブロックを1ブロック多く書き込むためである。縮退時の性能が正常時に比べて向上するのは故障ディスクに対する書き込みを行わない分だけ正常

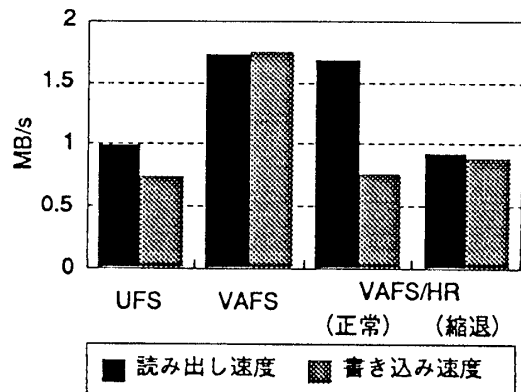


図3 ランダムアクセス性能

時よりも書き込みブロック数が減るからである。

#### (3) ランダム読み出し性能

縮退時の性能が正常時の約半分になるのは、故障ディスク上のデータブロックを復元する度に正常な全てのディスクから読み出しが生ずることにより、読み出しブロック数がほぼ2倍に増えるからである。

#### (4) ランダム書き込み性能

一般的なRAID4のランダム書き込みでは1ブロックの書き込み処理の度にパリティブロックの読み出しと書き込みが起こる。このため、パリティブロックを格納したディスクがボトルネックとなり、書き込み処理性能は単体ディスク性能の約半分になる。しかし、VAFS/HRでは、ファイルとして使用していないブロックは読み出さないため新規ファイルの書き込み処理では単体ディスク性能と同等の性能、すなわちUFSと同等の性能が得られる。

## 3. おわりに

VAFS/HRをUNIXのファイルシステムとして実装し、性能評価を行った。

シーケンシャルアクセスでは縮退時でもVAFSと比肩する性能が得られた。ランダムアクセスではUFSと同等か上回る性能が得られた。

これにより高速・高信頼ファイルシステムとしてのVAFS/HRの有効性を確認した。

## 参考文献

- [1] 鬼頭他6, 「高速UNIXファイルシステムの構想」他4件, 情報処理学会第47回全国大会講演論文集, 7B-1\*5, (平5-10)

注) UNIXオペレーティングシステムはUNIX System Laboratories, Inc.が開発し、ライセンスしています。