

VERITAS Foundation Suite™ 2.0 for Linux

FOUNDATION SUITE、EXT3、REISERFS の パフォーマンスの比較 — 概要

Linux Kernel 2.4.9-e3 enterprise

目次

はじめに.....	3
概要.....	3
テスト構成.....	3
ハードウェア テスト構成.....	3
ソフトウェア テスト構成.....	4
ソフトウェア リリース.....	4
ソフトウェア構成の詳細.....	4
結果と分析.....	5
詳細結果.....	5
まとめ.....	8
付録 A.....	9

はじめに

VERITAS Foundation Suite™ 2.0 for Linux には、SPECsfs3.0 NFS サーバー ベンチマーク テストで、標準 Linux Multiple Device ドライバ(MD)上で実行する他の 2 つのオープン ソース ジャーナリング ファイルシステム、EXT3 と ReiserFS を大幅に上回る最高スループットが達成しました。Foundation Suite は、EXT3 を 478%上回り、ReiserFS を 236%上回る最高スループットを達成しました。Foundation Suite は、このどちらのオープン ソース製品に比べて非常に高速かつスケラブルなクライアント要求応答時間を実現しました。NFSを実行しているエンタープライズ環境では、オープンソースの EXT3 や ReiserFS では得られない予測可能でスケラブルなパフォーマンスが Foundation Suite で得られます。

概要

この報告書の目的は、次の 3 つの環境で SPECsfs NFS サーバー ベンチマーク テストを実行してパフォーマンスを比較することです。

- VERITAS Volume Manager (VxVM) と VERITAS File System (VxFS) on Advanced Server 2.1 を含む VERITAS Foundation Suite 2.0 for Linux
- Red Hat Advanced Server 2.1 上の EXT3 ファイルシステム上で実行する MD
- Red Hat Advanced Server 2.1 上の ReiserFS 上で実行する MD

EXT3 は Red Hat のジャーナリング ファイルシステムです。ReiserFS は Namesys のジャーナリング ファイルシステムです。Linux の制約により、このベンチマーク テストは UDP を使用する NFS バージョン 3 に限定されました。ファイルサーバーのパフォーマンスを評価するために、SPECsfs97_R1 と呼ばれる Standard Performance Evaluation Corporation (SPEC) System File Server(SFS)ベンチマークを使用しました。これは SPECsfs 3.0 とも呼ばれています。

テスト構成

SPECsfs3.0 NFS サーバー ベンチマーク テストは、次のようなソフトウェア、ハードウェアおよびセットアップを使用して実施しました。

ハードウェア テスト構成

- 4 つの 700 MHz P-III Xeon CPU と 6 GB のメモリを備えた Dell 6450
- Dell PowerVault 200S ディスクアレイ

4 つの 700 MHz P-III Xeon CPU と 6 GB のメモリを備えた Dell 6450 は NFS version 3 サーバーとして使用しました。4 つの Dell PowerVault 200S ディスクアレイをすべてのテストに使用しました。各ディスクアレイは 8 つの 18 GB Seagate Cheetah ST318451LC 15K RPM ディスクドライブを備えており、合計 32 ドライブをテストに使用しました。各 PowerVault は、Adaptec 39160 SCSI Ultra-160 ホスト バス アダプタの 1 つのポートを介して 6450 に接続しました。2 つの 2 ポート Adaptec 39160 をベンチマーク テストに使用しました。6450 は、内蔵式 SCSI インターフェイスを使用して 18 GB の内臓ドライブにも接続しました。これらのドライブは、OS や VxVM シングル ディスク rootdg ディスクグループなど、様々なオーバーヘッド要件のために使用しました。最大 6 つのクライアントを使用して NFS 作業負荷を生成し、NFS クライアントには 1 つの 400 MHz UltraSparc III CPU と 128 MB の RAM を備えた Sun Microsystems Ultra5 ワークステーションを使用しました。1 つの Cisco 100/1000 BaseT ネットワーク スイッチ (Catalyst 3500XL) を使用してプライベート クライアント ネットワークを作成しました。NFS サーバーは 128 個の nfsd を使用するよう設定しました。

ソフトウェア テスト構成

このパフォーマンス テストは RAID-0(ストライプ)ボリューム構成に対して実行しました。VxFS SPECsfs の実行には VxVM を使用し、EXT3 と ReiserFS の実行には MD を使用しました。どちらの場合も、使用するボリューム マネージャに関係なくボリューム レイアウトが同じになるように構成しました。テストでは、各ボリューム マネージャを使用して 16 個の 4GB ボリューム(1 ボリュームに 2 つのディスク)を構成し、ボリュームの 2 つのディスクのそれぞれを(同じコントローラの異なるポートではなく)別々なコントローラに接続しました。

VxFS ボリュームは vxassist コマンドを使用して作成しました。EXT3 と ReiserFS のボリュームには、VxVM で作成したボリュームに匹敵するサイズのパーティションを使用して、ブート時に raidtab エントリをロードしました。どの構成にも同じディスクセットを使用しました。たとえば、VxVM 構成の最初のボリュームとして使用したディスク セットは、MD 構成の最初のボリュームを作成するときにも使用しました。

ソフトウェア リリース

次のオペレーティングシステムとベリタスソフトウェアのリリースをテストに使用しました。

- Red Hat Advanced Server 2.1 (kernel:2.4.9-e3enterprise)
 - この製品に同梱のEXT3とReiserFS
- VxFS 3.4 Update 2(Linux Foundation Suite 2.0 より)
- VxVM 3.2 Update 1(Linux Foundation Suite 2.0 より)

ソフトウェア構成の詳細

Foundation Suite テストでは、次のようなコマンド ラインを使用してボリュームとファイルを作成しました。(< > 内のプレースホルダは各ボリューム専用の値に置き換えました。)

VxVM RAID-0 ボリュームの作成方法:

```
vxassist -o ordered -g <disk group> make <volume> 4g layout=stripe < disk names >
```

すべてのテストで使用した一般的な VxFS ファイルシステム作成コマンド ライン:

```
/sbin/mkfs.vxfs <volume name>
```

Foundation Suite スタックで使用したマウント コマンドの例:

```
/sbin/mount.vxfs -o log <volume name> <mount point>
```

オープン ソース テストでは、次のようなエントリを使用してボリュームとファイルを作成しました。

次の/etc/raidtab エントリは、MD RAID-0 ボリュームを作成するために使用した raidtab の例です。

```
raiddev                /dev/md1
raid-level              0
nr-raid-disks          2
chunk-size              64
nr-spare-disks         0
persistent-superblock  1
device                 /dev/sdc1
raid-disk               0
device                 /dev/sds1
raid-disk               1
```

すべてのテストで EXT3 と ReiserFS のファイルシステムの作成に使用したコマンドライン:

```
/sbin/mkfs.ext2 -j <volume name>
```

```
/sbin/mkreiser -q <volume name>
```

EXT3 と ReiserFS に使用したマウント コマンドの例:

```
/bin/mount <volume name> <mount point>
```

結果と分析

Foundation Suite が EXT3 上で実行する MD と ReiserFS 上で実行する MD を上回った最高スループットを表 1 に示します。Foundation Suite スタックは、EXT3 スタックを 478% 上回り、ReiserFS スタックを 236% 上回る最高スループットを達成しました。

構成	Foundation Suite スタックと EXT3 & MD の比較	Foundation Suite スタックと ReiserFS & MD の比較
RAID-0 UDPV3	478%	236%

表 1: Foundation Suite スタックが達成した SPECsfs97_R1 最高スループット

UDP v.3		
	最高スループット (NFS 処理数/秒)	総応答時間
Foundation Suite	4480	3.72
EXT3 & MD	818	10.66
ReiserFS & MD	1335	7.39

表 2: UDP V3 最高スループットと総応答時間

表 2 の NFS UDPv.3 の結果は、Foundation Suite スタックが競合スタックのスループットを上回っていることを示しています。EXT3 スタックの総応答時間は Foundation Suite スタックより 187% 長く、ReiserFS スタックは Foundation Suite より 99% 長い。

詳細結果

異なるファイルシステムとボリューム マネージャに対して実行した各ベンチマーク テストの詳細結果を表 3 に示します。この結果は、Foundation Suite は EXT3 や ReiserFS 上で実行している MD よりスループットが著しく高いことを示しています。

Foundation Suite の最高スループットは、RAID-0 ボリューム構成の EXT3 と MD を 478% 上回り、ReiserFS と MD を 236% 上回っています。サーバーの負荷を大きくしても、EXT3 と MD の総応答時間 (ORT) は Foundation Suite を 187% 上回り、ReiserFS と MD の場合は Foundation Suite を 99% 上回っています。(ORT は、テストした負荷の全体にわたってシステムがどのように応答したかを計る尺度です)。

UDP V3					
Foundation Suite		EXT3		ReiserFS	
処理数/秒	応答時間 (ミリ秒/処理)	処理数/秒	応答時間 (ミリ秒/処理)	処理数/秒	応答時間 (ミリ秒/処理)
448	1.6	80		139	2.3
906	1.7	160	3.3	283	2.5
1372	1.8	240	8.4	425	15.7
1825	1.9	325	6.2	566	3.5
2263	2.4	397	9.7	711	11.4
2728	3.5	506	12.9	839	8.4
3193	4.7	574	14.4	996	4.8
3656	6.4	652	16.5	1139	9.7
4097	8.5	566	25.0	1289	13.0
4480	11.5	818	16.4	1335	12.7

表 3: RAID-0 ボリューム構成に対する SPECsfs97_R1 統計値。スループット (処理数/秒) の欄では、UDP V3 の結果で Foundation Suite は EXT3 を 478% 上回り、ReiserFS を 236% 上回る最高スループットを達成しました。

RAID-0 構成のファイルシステムのスループットと応答時間を図 1 に示します。

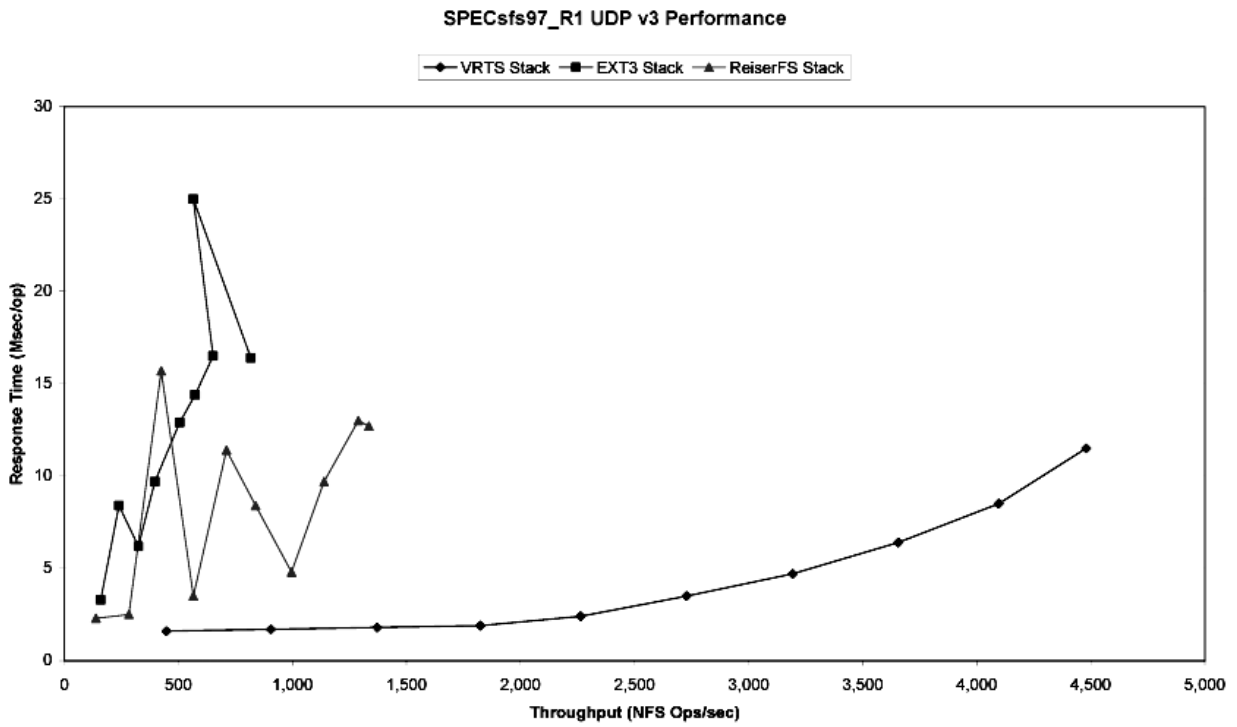


図 1: Foundation Suite、EXT3、ReiserFS のスタックのスループットと応答時間。このグラフの下に行くほど、また右に行くほどパフォーマンスが優れています。スループット(処理数/秒)の欄では、UDP V3 の結果で Foundation Suite は EXT3 を 478%上回り、ReiserFS を 236 %上回る最高スループットを達成しました。

RAID-0 構成のファイルシステムのスループットと CPU 使用率を図 2 に示します。この図では、ベンチマーク テストで EXT3 と ReiserFS のスタックはどちらも CPU 使用率が大きく変動していますが、Foundation Suite の CPU 使用率は非常に均等であることがわかります。

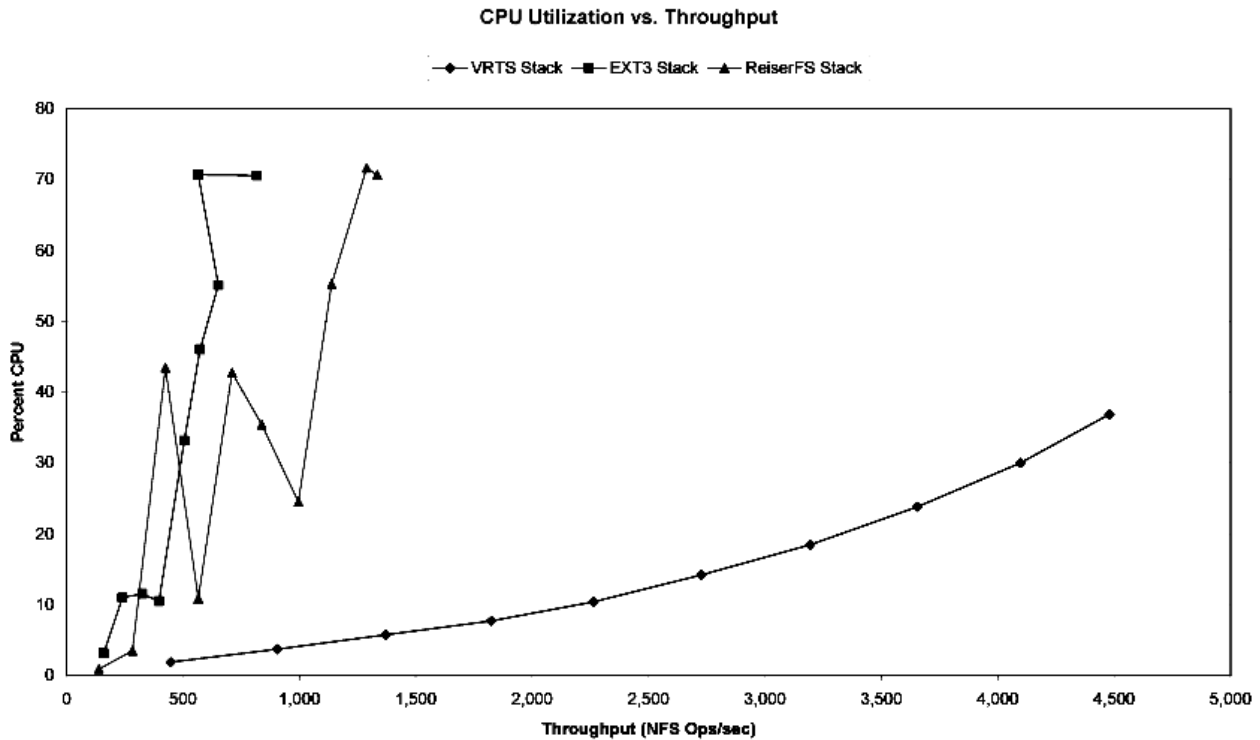


図 2: Foundation Suite、EXT3、ReiserFS のスタックのスループットと CPU 使用率。このグラフの下に行くほど、また右に行くほどパフォーマンスが優れています。ベンチマーク テスト全体にわたって EXT3 と ReiserFS のスタックの CPU 使用率が変動している点に注目してください。

信頼できる *iostat* がないので、Linux システムでディスク I/O 使用率を測定するのは非常に困難です。VxVM を使用したベンチマーク テストでは *vxstat* を使用して I/O 使用率を測定することができました。限られたディスク I/O 使用率の測定値から、Foundation Suite スタックはディスク I/O のボトルネックにはならないという結論に達することができました。

この報告書のベンチマーク テスト結果の捕捉説明と数字で部分的にしか示せないという事実が、EXT3 と ReiserFS のベンチマーク テストで遭遇した困難の大きさを示しています。この結果からわかるように、EXT3 と ReiserFS のベンチマーク テストの結果には非常にむらがあります。これは SPECsfs ベンチマーク テストの正常な結果ではなく、この不均一さが EXT3 や ReiserFS に固有のものであることを SPECsfs ベンチマーク テストで明らかにするには、非常に多くのテストを実施する必要があります。EXT3 と ReiserFS のベンチマーク テストの第二の問題は、必要なデータが得られる前に失敗した SPECsfs ベンチマーク テストの数です。EXT3 と ReiserFS に対して実施したベンチマーク テストの約半数が終了前に失敗しました。これに対して、Foundation Suite スタックに対して実施した 30 のベンチマーク テストはすべて成功しました。

さらに注意することは、このプロジェクトの初期段階では Red Hat 7.2(kernel 2.4.9-13)をテスト OS として使用した点です。この OS を使用している間は、VxVM を使用する EXT3 のベンチマーク テストで 2,600 処理数/秒の最高スループットを達成しました。同じハードウェアを Advanced Server 2.1 で使用したときには、VxVM を使用する EXT3 が達成した最高スループットは約 800 処理数/秒でした。どうしてこうなるのかという質問に対する Red Hat からの回答はすぐには得られませんでした。

まとめ

この報告書で説明したベンチマーク テストの結果は、Foundation Suite が EXT3 や ReiserFS のスタックよりも非常に高いスループットを達成したことを示しています。RAID-0 と UDP プロトコルを使用したベンチマーク テストで、Foundation Suite は約 4,500 処理数/秒の最高スループットを達成しました。EXT3 スタックの最高スループットは約 800 NFS 処理数/秒で、ReiserFS と MD は約 1,300 NFS 処理数/秒でした。このような結果から、Foundation Suite のスループットが ReiserFS より非常に高く、EXT3 より圧倒的に高いことがわかります。

ファイル サービスのために NFS を実行しているエンタープライズ環境では、VERITAS Foundation Suite を使用すると EXT3 や ReiserFS よりも高速で予測可能なパフォーマンスが得られます。Foundation Suite では CPU 使用率が非常に均一ですが、EXT3 や ReiserFS では作業負荷が増えると大きく変動します。

付録A

NFS version 3 を使用する SPECsfs 3.0 の作業負荷配分を表 4 に示します。

NFS Op	配分
getattr	11%
setattr	1%
lookup	27%
readlink	7%
read	18%
write	9%
create	1%
remove	1%
readdir	2%
fsstat	1%
access	7%
commit	5%
readdirplus	9%

表 4: SPECsfs NFS v3 の作業負荷配分

次に示すのは、今回ベンチマーク テストを実施した 3 つのスタックに関する組合せ SPECsfs の完全な情報開示用レポートです。

SPECsfs97_R1.v3 の結果

Foundation Suite RAID-0

SPECsfs97_R1.v3 = 4480 処理数/秒 (総応答時間 = 3.72)

EXT3/MD RAID-0

SPECsfs97_R1.v3 = 818 処理数/秒 (総応答時間 = 10.66)

ReiserFS/MD RAID-0

SPECsfs97_R1.v3 = 1335 処理数/秒 (総応答時間 = 7.39)

RAID-0

Foundation Suite		EXT3/MD		ReiserFS/MD	
処理数/秒	応答時間 (ミリ秒/処理)	処理数/秒	応答時間 (ミリ秒/処理)	処理数/秒	応答時間 (ミリ秒/処理)
448	1.6	80		139	2.3
906	1.7	160	3.3	283	2.5
1372	1.8	240	8.4	425	15.7
1825	1.9	325	6.2	566	3.5
2263	2.4	397	9.7	711	11.4
2728	3.5	506	12.9	839	8.4
3193	4.7	574	14.4	996	4.8
3656	6.4	652	16.5	1139	9.7
4097	8.5	566	25.0	1289	13.0
4480	11.5	818	16.4	1335	12.7

<CPU、メモリ、処理能力>

モデル名	Dell 6450
プロセッサ	700 MHz P-III Xeon
プロセッサ数	4
一次キャッシュ	16KBI+16KBD オン チップ
二次キャッシュ	512KB(I+D)オン チップ
その他のキャッシュ	N/A
UPS	N/A
その他のハードウェア	N/A
メモリ サイズ	6 GB
NVRAM サイズ	N/A
NVRAM タイプ	N/A
NVRAM の詳細	N/A

サーバー ソフトウェア

OS 名とバージョン	Red Hat Advanced Server 2.1 (kernel 2.4.9-e3enterprise)
その他のソフトウェア	VxFS 3.4 Update 2、VxVM 3.2 Update 1
ファイルシステム	VxFS、EXT3、ReiserFS
NFS バージョン	3

サーバー チューニング

バッファ キャッシュ サイズ	デフォルト
NFS プロセッサ数	128
ファイルセット サイズ	43.5 GB (Foundation Suite) 7.7 GB (EXT3/MD) 13.5 GB (ReiserFS/MD)

ネットワーク サブシステム

ネットワーク タイプ	ギガビット イーサネット
ネットワーク コントローラの詳細	Intel PPro 1000
ネットワーク数	1
ネットワーク コントローラ数	1
プロトコル タイプ	UDP
スイッチ タイプ	Cisco 3500XL スイッチ
ブリッジ タイプ	N/A
ハブ タイプ	N/A
その他のネットワーク ハードウェア	N/A

<ディスク サブシステムとファイルシステム>

ディスク コントローラ数	3
ディスク数	34
ファイルシステム数	16 (F1 ~ F16)
ファイルシステム作成処理数	デフォルト
ファイルシステム構成	デフォルト

ディスク コントローラ	組込み SCSI コントローラ
コントローラ タイプ数	1
ディスク数	2
ディスク タイプ	Seagate Cheetah 18 GB 10 K RPM & Quantum Atlas 18 GB 10 K RPM
ディスク上ファイルシステム構成上の特記	OS、ルート ディスクグループ

ディスク コントローラ	Adaptec 39160
コントローラ タイプ数	2
ディスク数	32
ディスク タイプ	Seagate Cheetah 18 GB 15 K RPM
ディスク上ファイルシステム構成上の特記	F1 ~ F16

<ロード ジェネレータ(LG)構成>

ロード ジェネレータ数	2 (EXT3、ReiserFS)、6 (Foundation Suite)
1 LG あたりのプロセッサ数	8
Biod 最大読取り設定値	2
Biod 最大書込み設定値	2
LG タイプ	LG1
LG モデル	Sun Microsystems Ultra Enterprise 5
プロセッサ数とタイプ	1400 MHz UltraSPARC
メモリ サイズ	128 MB
オペレーティングシステム	Solaris 2.8
コンパイラ	SPEC 提供のプリコンパイル バイナリ
コンパイラ オプション	N/A
ネットワーク タイプ	オンボード 100 baseT

<テスト構成>

LG 数	LG タイプ	ネットワーク	ターゲット ファイルシステム	備考
-----	-----	-----	-----	-----
1 ~ 6	LG1	N1	16 (F1 ~ F16)	Foundation Suite
1 ~ 2	LG1	N1	16 (F1 ~ F16)	EXT3 と ReiserFS

<注記とチューニング>

なし。

※ベリタスソフトウェアの詳細な製品情報につきましては、弊社の Web サイト(<http://www.veritas.com/>)をご覧ください。

© Copyright 2003 VERITAS Software Corporation. All rights reserved. VERITAS、VERITAS SOFTWARE、VERITAS ロゴ、およびその他の VERITAS 製品は、米国および各国の VERITAS Software Corporation の商標または登録商標です。その他の会社名、製品名等は、それぞれ各社の商標または登録商標です。製品の仕様・性能等は予告なく変更する場合がありますので、ご了承ください。

ベリタスソフトウェア株式会社

〒100-0011 東京都千代田区内幸町 2 丁目 2 番 2 号 富国生命ビル
TEL.03-5532-8241 FAX.03-5532-0887
<http://www.veritas.com/jp>

お問い合わせ先