

VERITAS Volume Manager™ 3.1

Administrator's Guide (日本語版)

Solaris

2001年2月
P/N 30-000102-011


VERITAS

免責事項

本書に記載されている情報は、予告なしに変更される場合があります。VERITAS Software Corporation は、本書に関して、商品性や特定目的に対する適合性の黙示保証などの一切の保証を行いません。本書に含まれる誤りや本書の提供、遂行、または使用に伴う付随的または間接的な損害に対して一切の責任を負わないものとします。

著作権

Copyright © 1999-2000 VERITAS Software Corporation. All rights reserved. VERITAS は、米国およびその他の国における VERITAS Software Corporation の登録商標です。VERITAS のロゴおよび VERITAS Volume Manager は、VERITAS Software Corporation の商標です。その他、本書に記載されている会社名、製品名は、各社の商標または登録商標です。

Printed in the USA, February, 2001.

VERITAS Software Corporation
1600 Plymouth St.
Mountain View, CA 94043
電話 (1) 650-335-8000
ファックス (1) 650-335-8050
www.veritas.com



目次

まえがき	xiii
対象読者	xiii
内容	xiii
本書の構成	xiii
本書の用途	xiv
関連マニュアル	xiv
表記規則	xv
第 1 章 Volume Manager とは	1
はじめに	1
データの格納方法	2
Volume Manager の概要	2
物理オブジェクト	3
物理ディスクおよびディスク名の命名	3
パーティション	3
ボリュームと仮想オブジェクト	4
Volume Manager ディスク	4
ディスク グループ	5
サブディスク	5
プレックス	6
ボリューム	7
Volume Manager の仮想オブジェクト間の関係	9
ボリュームのレイアウト	10
コンカチネーション	11



ストライピング (RAID-0)	14
RAID-5	17
ミラーリング (RAID-1)	18
ミラーリング アンド ストライピング (RAID-1 + RAID-0)	18
ストライピング アンド ミラーリング (RAID-0 + RAID-1)	19
Volume Manager および RAID-5	20
従来の RAID-5 アレイ	20
Volume Manager RAID-5 アレイ	21
ロギング	23
階層化ボリューム	25
Volume Manager およびオペレーティング システム	26
Volume Manager のレイアウト	26
Volume Manager のユーザ インタフェース	28
ユーザ インタフェースの概要	28
Volume Manager を使用する目的	28
第 2 章 設定と起動	31
はじめに	31
Volume Manager の初期化	32
Volume Manager のデーモン	33
設定デーモン (vxconfigd)	33
ボリューム入出力デーモン (vxioid)	34
システムの設定	35
システム設定手順の例	35
システム設定ガイドライン	36
ホットリロケーションのガイドライン	36
ストライピングのガイドライン	37
ミラーリングのガイドライン	38
ダーティ リージョン ロギング (DRL) のガイドライン	39
ミラーリング アンド ストライピングのガイドライン	40

ストライピング アンド ミラーリングのガイドライン	40
RAID-5 のガイドライン	41
システムの保護	42
第 3 章 Volume Manager の機能	45
はじめに	45
オンライン再レイアウト	46
ストレージレイアウト	46
オンライン再レイアウトの使用方法	47
レイアウト変更の種類	48
レイアウト変更の特性	50
レイアウト変更およびボリュームの大きさ	50
ホットリロケーション	51
ホットリロケーションの仕組み	51
リロケート用領域の選択方法	52
リロケート処理の解除ユーティリティ	53
ボリュームの再同期化	53
ダーティリージョン ロギング	55
高速ミラー再同期 (FMR)	56
FMR コンポーネント	57
VxVM のスナップショット機能に対する FMR の拡張機能	58
Volume Manager のルート機能	61
ルート ボリュームを使用したブート	62
ブート時のボリュームの制限	62
動的マルチパス (DMP)	64
パス フェイルオーバー機構	65
負荷の平準化	65
DMP デバイスからのブート	65
コントローラの有効化と無効化	66
DMP データベース情報の表示	66



VxSmartSync Recovery Accelerator	67
データ ボリュームの設定	67
REDO ログ ボリュームの設定	68
Volume Manager タスク モニタ	68
Volume Manager のクラスタ機能	68
第 4 章 ディスク関連タスク	69
はじめに	69
標準ディスク デバイス	70
ディスク グループ	72
ディスク コマンドおよびディスク グループ コマンド	73
ディスクの初期化と追加	74
Volume Manager へのディスクの追加	77
ディスク グループへのディスクの追加	80
ディスク グループからのディスクの削除	83
ディスクの移動	84
ディスク名の変更	85
ディスクの予約	86
ディスクのオフライン	86
ディスクのミラーリング	87
ディスクの削除	88
ディスク情報の表示	89
障害のあるディスクの検出と交換	91
ホットリロケーション	91
vxunrelocate の使用	96
障害のあるディスクの検出	99
ディスクの交換	101
ディスク グループの作成	103
ディスク グループのアップグレード	105
ディスク グループの削除	107



システム間でのディスク グループの移動	108
ディスク グループの破棄	112
特殊デバイスの使用	112
特殊なカプセル化用の vxdisk の使用	112
RAM ディスクでの vxdisk の使用	114
vxdisk を使用したマルチパス情報の表示	115
vxdiskadm メニュー インタフェース処理	116
vxdiskadm の起動	116
vxdiskadm メイン メニュー	117
Volume Manager 管理下への物理ディスクの配置	121
今後の使用に備えたディスクの追加	134
Volume Manager で使用するためのディスクの再初期化	134
VM ディスクのホットリロケーション プールへの追加	137
VM ディスクのホットリロケーション プールからの削除	138
ホットリロケート済みサブディスクのディスクへの移動	139
ディスクのホットリロケーション利用対象からの除外	140
ディスクのホットリロケーション設定	140
物理ディスク上の情報の表示	141
物理ディスクの削除	142
物理ディスクの無効化（物理ディスクのオフライン化）	143
物理ディスクの有効化	144
物理ディスクの交換	145
障害ディスクまたは削除済みディスクの交換	146
ディスク グループへのディスクの追加	147
ディスク グループの作成	147
ディスク グループのデポート	151
ディスク グループのインポート	152
vxdiskadm の終了	153



第 5 章 ボリューム関連タスク	155
はじめに	155
ボリューム、プレックス、およびサブディスク関連タスクとオンライン バックアップ	155
ボリュームの作成	157
コンカチネイテッド ボリュームの作成	159
ストライプ ボリュームの作成	160
RAID-5 ボリュームの作成	161
ミラー ボリュームの作成	161
ボリュームのサイズ変更	162
vxassist を使用したボリュームのサイズ変更	162
vxresize を使用したボリュームのサイズ変更	163
ボリュームの削除	164
ボリュームのミラーリング	164
ミラー ボリュームの作成	164
既存ボリュームのミラーリング	164
ダーティ リージョン ロギングを使用したボリュームの作成	164
すべてのボリュームのミラーリング	165
ミラーの削除	165
RAID-5 ログの追加	166
DRL ログの追加	166
DRL ログの削除	166
RAID-5 ログの削除	167
ボリュームの停止	167
ボリュームの起動	167
起動できないボリュームの一覧表示	168
ボリュームの読み取りポリシーの変更	168
バックアップからリストアするためのボリュームの準備	169
ボリュームのリカバリ	170
ボリューム情報の表示	171



プレックス関連タスク	172
プレックスの作成	172
プレックスの関連付け	173
プレックスの関連付けの解除とプレックスの削除	174
ミラーを使用したバックアップ	174
プレックス情報の表示	175
プレックス属性の変更	176
プレックスの状態の変更:プレックスの切り離しと接続	177
プレックスの移動	179
プレックスのコピー	179
サブディスク関連タスク	180
サブディスクの作成	180
サブディスクの削除	180
サブディスク情報の表示	181
サブディスクの関連付け	181
サブディスクの関連付けの解除	183
サブディスク属性の変更	184
サブディスクの移動	184
サブディスクの分割	185
サブディスクの結合	185
オンライン バックアップの実行	186
高速ミラー再同期	187
VM ディスク上でのボリュームのミラーリング	189
VM ディスクからのボリュームの移動	191
第 6 章 Volume Manager のクラスタ機能	193
はじめに	193
クラスタ機能の概要	194
Volume Manager におけるオブジェクトの共有	194
VxVM クラスタ内のディスク	203



ディスクの切り離しポリシー	203
ディスクグループのアクティブ化	203
データリージョン ロギングとクラスタ環境	204
ログのフォーマットとサイズ	205
互換性	205
クラスタ環境における DRL の仕組み	206
Volume Manager クラスタ機能のアップグレード	207
クラスタ関係の Volume Manager のユーティリティとデーモン	208
vxclust	208
vxclustadm	209
vxconfigd	209
vxconfigd のリカバリ	210
vxdg	211
vxdisk	214
vxrecover	215
vxdctl	215
vxstat	217
クラスタ処理と高速ミラー再同期 (FMR)	217
第 7 章 リカバリ	219
はじめに	219
UNIX のブート プロセス	220
障害発生後のブート	221
ルート (/)、swap、および usr 構成の設定	221
ボリューム上のルート (/) ファイルシステムまたは /usr ファイルシステム の修復	222
ルート ファイルシステムのバックアップとリストア	225
障害とリカバリ手順	226
UNIX パーティションでの障害	226
ブート デバイスへのアクセス時の障害	226
/etc/vfstab 中の不正なエントリによる障害	227

/etc/system の欠損または破損による障害	228
使用できない陳腐化したプレックスからブートしたことによる障害	230
ホットリロケーションおよびブート ディスク障害	232
ブート ディスクの再追加と交換	233
障害を起こしたブート ディスクの再追加	233
カプセル化されたサブディスクの新しいディスクへのリロケート処理の解除 ..	235
障害を起こしたブート ディスクの交換	236
ディスクの再接続	237
再インストールによるリカバリ	238
一般的な再インストール情報	238
再インストールおよび構成の再設定の手順	239
プレックスとボリュームの状態	249
プレックスの状態	249
プレックスの状態のサイクル	252
プレックスのカーネル状態	252
ボリュームの状態	252
ボリュームのカーネル状態	254
RAID-5 ボリュームのリカバリ	254
RAID-5 ボリュームのレイアウト	254
RAID-5 ボリュームの作成	256
RAID-5 ボリュームの初期化	256
障害と RAID-5 ボリューム	257
RAID-5 のリカバリ	259
RAID-5 のさまざまな処理	262
RAID-5 ログの操作	262
RAID-5 サブディスクの操作	263
RAID-5 ボリュームの起動	264
RAID-5 ボリュームを起動できない場合	265
RAID-5 ボリュームを強制的に起動する場合	266
RAID-5 ボリューム起動時のリカバリ	267



RAID-5 ボリュームの属性の変更	268
RAID-5 ディスク アレイへの書き込み	269
第 8 章 VxVM パフォーマンスの監視	275
はじめに	275
パフォーマンスのガイドライン	275
データの配置	276
ストライピング	276
ミラーリング	277
ミラーリング アンド ストライピング	278
ストライピング アンド ミラーリング	279
RAID-5	279
パフォーマンスの監視	280
パフォーマンスの優先事項	280
パフォーマンス データの収集	280
パフォーマンス データの使用	282
Volume Manager のチューニング	285
一般的なチューニングのガイドライン	285
チューニング可能パラメータ	285
大型システムのチューニング	291
用語集	293
索引	305

まえがき

本書では、VERITAS Volume Manager™ の使用方法について説明します。

対象読者

本書では、VERITAS Volume Manager の管理下でシステムのインストール、設定、および管理を行うシステム管理者を対象としています。

本書では、読者が以下の知識と経験を有していることを前提とします。

- ◆ UNIX オペレーティング システムに関する実務知識
- ◆ システム管理の基本的な理解
- ◆ ボリューム管理の基本的な理解

内容

本書では、システム管理者を対象に、Volume Manager を使用したボリューム管理およびシステム管理に関する概念と手順を詳細に説明します。Volume Manager のさまざまな機能の活用方法、Volume Manager コマンドを使用したオブジェクトの作成および操作方法、ディスク障害からのリカバリ方法なども説明します。

本書の構成

本書の構成は以下の通りです。

- ◆ Volume Manager とは
- ◆ 設定と起動
- ◆ Volume Manager の機能
- ◆ ディスク関連タスク
- ◆ ボリューム関連タスク
- ◆ Volume Manager のクラスタ機能
- ◆ リカバリ
- ◆ VxVM パフォーマンスの監視



本書の用途

本書では、Volume Manager のシステム管理に関するタスクを実行する手順について説明します。Volume Manager の管理機能を実行するには、次の1つまたは複数のインタフェースを使用します。

- ◆ 一連の複雑なコマンドの集まり
- ◆ 単一の自動化コマンド (vxassist)
- ◆ メニュー方式のインタフェース (vxdiskadm)
- ◆ Storage Administrator (グラフィカル ユーザ インタフェース)

本書では、さまざまなVolume Managerコマンド ライン インタフェースを使用した、Volume Manager による管理方法を説明します。Storage Administrator のグラフィカル ユーザ インタフェースの使用方法については、『VERITAS Volume Manager Storage Administrator Administrator's Guide』を参照してください。Volume Manager ユーティリティ、各ユーティリティのオプション、およびそれらの使用方法の詳細については、Volume Manager マニュアルページを参照してください。

注 Volume Manager コマンドを実行するには、多くの場合スーパーユーザ権限または適切なアクセス権限が必要です。

関連マニュアル

関連情報については、次のマニュアルを参照してください。

- ◆ 『VERITAS Volume Manager Installation Guide』
- ◆ 『VERITAS Volume Manager Release Notes』
- ◆ 『VERITAS Volume Manager Hardware Notes』
- ◆ 『VERITAS Volume Manager Reference Guide』
- ◆ 『VERITAS Volume Manager Storage Administrator Administrator's Guide』
- ◆ VERITAS Volume Manager マニュアル ページ

表記規則

次の表は、本書で使用している表記規則を示したものです。

表記	用途	例
固定スペース	コンピュータによる出力、ファイル、ディレクトリ、コマンド オプションなどのソフトウェア要素、関数名、およびパラメータ	/etc/vx/tunefstab ファイルからチューニング可能パラメータを読み取ります。 詳細については、ls (1) マニュアルページを参照してください。
固定スペース(太字)	ユーザによる入力	# mount -F vxfs /h/filesys
斜体	実際の名前または値を代入する変数	変数 <i>ncsize</i> は、... の値を決定します。

記号	用途	例
%	C シェル プロンプト	
\$	Bourne/Korn シェル プロンプト	
#	スーパーユーザ プロンプト (全シェル)	
\	次の行に入力が続きます。ユーザがこの文字を入力する必要はありません。	# mount -F vxfs \ /h/filesys
[]	コマンド内で使用される角かっこは、オプション引数を示します。	ls [-a]
	コマンド内で使用される垂直バーは、相互に排他的な引数を区切ります。	mount [suid nosuid]





はじめに

本書では、VERITAS Volume Manager (VxVM[®]) に関する機能、仕組み、ユーザ インタフェースを介した操作方法、および概念について説明します。Volume Manager の詳細情報が記載されている関連マニュアルについては、「まえがき」を参照してください。

VERITAS Volume Manager を使用すると、コンピューティング環境において使いやすいオンライン ディスク ストレージ管理を実現することができます。従来のディスク ストレージ管理では、マシンを頻繁にオフラインにする必要があり、非常に不便でした。分散型クライアント / サーバ環境においては、データベースを始めとするリソースは、高い可用性の維持、容易なアクセス、およびハードウェアの故障による障害からの保護が要求されます。

VERITAS Volume Manager では、パフォーマンスを向上させ、データの可用性と整合性を保証するツールを提供します。Volume Manager では、システムの稼働中でも動的にディスク ストレージの構成を設定することができます。

この章では、VERITAS Volume Manager の概念を紹介し、ストレージ管理を行うために Volume Manager で使用するツールについて説明します。

この章では、次のトピックについて説明します。

- ◆ データの格納方法
- ◆ Volume Manager の概要
- ◆ 物理オブジェクト
- ◆ ボリュームと仮想オブジェクト
- ◆ ボリュームのレイアウト
- ◆ Volume Manager および RAID-5
- ◆ 階層化ボリューム
- ◆ Volume Manager のユーザ インタフェース



データの格納方法

データを物理ディスクに格納する方法は何通りかあります。いずれの方法でも、データを格納して効率よく検索できるように、ディスク上にデータを編成します。このようなディスクの基本的な編成方法は、フォーマットと呼ばれる。フォーマットとは、あらかじめ設定されたストレージを設定する基本的なパターンに従って、データをディスクに書き込んだり、ディスクから検索したりできるように、ハードディスクを準備することです。

ハードディスクをフォーマットして情報を格納する方式には、物理ストレージレイアウトと論理ストレージレイアウトの2通りがあります。Volume Manager では、論理ストレージレイアウト方式を採用しています。この章では、Volume Manager がサポートしているストレージレイアウトのタイプについて説明します。

Volume Manager の概要

Volume Manager ではオブジェクトを使用してストレージ管理を行います。Volume Manager で使用するオブジェクトには、物理オブジェクトと仮想オブジェクトの2種類があります。

- ◆ 物理オブジェクト

Volume Manager では、物理ディスクを使用してデータを格納します。

- ◆ 仮想オブジェクト

Volume Manager では、物理ディスク上にボリュームと呼ばれる仮想オブジェクトが作成されます。各ボリュームでは、1つ以上の物理ディスクにわたるデータを記録および検索することができます。ボリュームには、ファイルシステム、データベース、またはほかのアプリケーションから、物理ディスクにアクセスする場合と同じ方法でアクセスします。また、ボリュームはさらにほかの仮想オブジェクトからも構成されます。それらの仮想構成要素を使用してボリュームの構成を変更します。ここでは、ボリュームおよびその仮想コンポーネントを、仮想オブジェクトまたは Volume Manager オブジェクトと呼びます。Volume Manager オブジェクトを使用すると、アプリケーションやユーザの作業を妨げることなく、ディスク上で管理タスクを実行できます。

物理オブジェクト

この節では、Volume Manager で使用する物理オブジェクト（物理ディスク）について説明します。

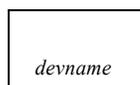
物理ディスクおよびディスク名の命名

物理ディスクとは、最終的にデータが格納される基本ストレージデバイス（メディア）です。物理ディスク上のデータにアクセスするには、デバイス名を使用して対象のディスクを特定します。物理ディスク デバイス名は、使用するコンピュータ システムによって異なります。すべてのシステムですべてのパラメータが使用されるわけではありません。一般的なデバイス名には、`c#t#d#` が含まれます。各記号の意味は、以下の通りです。

- ◆ `c#`: コントローラ
- ◆ `t#`: ターゲット ID
- ◆ `d#`: ディスク番号

本書における物理ディスクとデバイス名 (*devname*) の表記を図 1「物理ディスクの例」に示します。たとえば、デバイス名 `c0t0d0` は、0 番のコントローラに接続されており、そのターゲット ID が 0、物理ディスク番号が 0 であることを表わします。

図 1. 物理ディスクの例



パーティション

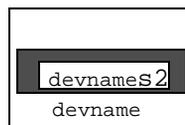
コンピュータ システムによっては、物理ディスクを 1 つ以上のパーティションに分割することができます。パーティション番号 (`s#`) は、`devname` の末尾に追加します。図 2「パーティションの例」に示すように、物理ディスク全体を 1 つのパーティションとすることもできます。

図 2. パーティションの例

パーティション



1 つのパーティションで構成される物理ディスク



ボリュームと仮想オブジェクト

物理オブジェクトと Volume Manager オブジェクトの関係は、物理ディスクを Volume Manager の管理下に置いたときに確立されます。

Volume Manager では、仮想オブジェクト (Volume Manager オブジェクト) を作成し、物理オブジェクトと仮想オブジェクトの間の論理関係を確立します。それ以降は、Volume Manager では仮想オブジェクトを使用して、ストレージ管理タスクを行います。Volume Manager オブジェクトには、以下のものがあります。

- ◆ ディスク グループ
- ◆ VM ディスク
- ◆ ボリューム
- ◆ プレックス (ミラー)
- ◆ サブディスク

これらのオブジェクトについては、後の節で説明します。

ボリュームは仮想ディスク デバイスであり、アプリケーション、データベース、およびファイルシステムからは物理ディスクとして見えます。ただし、ボリュームには物理ディスクの制限は適用されません。Volume Manager を使用する場合、アプリケーションは物理ディスクではなく、Volume Manager ディスク (VM ディスク) 上に作成されたボリュームにアクセスすることになります。

ボリュームにはほかの仮想オブジェクトが含まれます。それらのオブジェクトを使用してボリューム内のデータを操作できます。ボリュームに含まれる仮想オブジェクトには、サブディスクとプレックスがあります。仮想オブジェクトの詳細については、以降の節で説明します。仮想オブジェクトの組み合わせや、ボリュームによる仮想オブジェクトの取り扱いについては、「ボリューム (7 ページ)」で説明します。

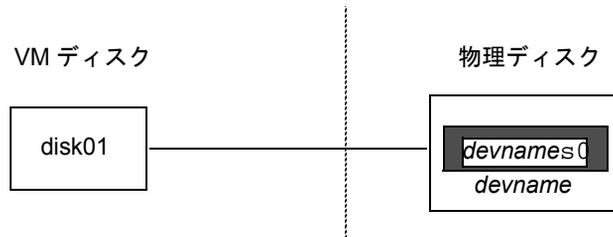
Volume Manager ディスク

物理ディスクを Volume Manager の管理下に組み入れると、その物理ディスクに 1 つの Volume Manager ディスク (VM ディスク) が割り当てられます。Volume Manager の管理下にある VM ディスクは、通常、ディスク グループに属します。各 VM ディスクは、1 つ以上の物理ディスクに対応しています。

通常、VM ディスクには、共有領域 (割り当てられた記憶領域) と専有領域 (Volume Manager の内部設定情報が格納されている領域) があります。

各 VM ディスクには、一意のディスク メディア名 (仮想ディスク名) を指定します。この場合、ディスク名を指定することも、Volume Manager によってデフォルト名 (通常は `disk##`) が割り当てられるようにすることもできます。図 3 「VM ディスクの例」に、物理ディスク `devname` に割り当てられている VM ディスク (メディア名 `disk01`) を示します。

図 3. VM ディスクの例



ディスク グループ

ディスク グループは、構成が共通である VM ディスクの集まりです。ディスク グループ構成は、関連する Volume Manager オブジェクト、それらの属性、およびオブジェクト間の関係についての詳細情報を記録したレコードの集まりです。デフォルトのディスク グループは、rootdg (ルート ディスク グループ) です。

必要に応じて、ディスク グループを追加作成することができます。ディスク グループを使用すると、ディスクを論理的なグループに分けることができます。ディスク グループとその構成要素は、あるホスト マシンから別のホスト マシンにひとまとまりの単位として移動することができます。

ボリュームはディスク グループ内に作成されます。ボリュームを構成するディスクは、同じディスク グループに属している必要があります。

サブディスク

サブディスクは、連続するディスク ブロックの集まりです。ブロックは、ディスク上の領域の単位です。Volume Manager は、サブディスクに基づいてディスク領域を割り当てます。VM ディスクは 1 つ以上のサブディスクに分割できます。各サブディスクは VM ディスクの特定の部分を表し、その部分は物理ディスクの特定の領域にマッピングされています。

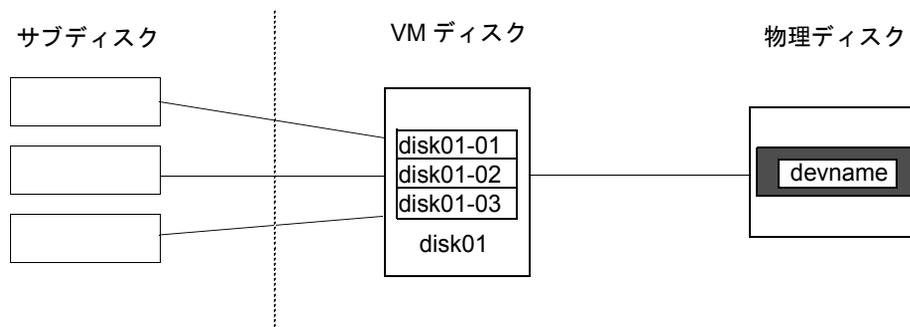
VM ディスクのデフォルト名は disk## (disk01 など) であり、サブディスクのデフォルト名は disk##-## です。図 4 「サブディスクの例」では、VM ディスク名が disk01 であり、その最初のサブディスクに disk01-01 という名前が付けられています。

図 4. サブディスクの例



VM ディスクは複数のサブディスクを含むことができますが、VM ディスク内のサブディスクは重複したり、VM ディスクの同じ部分を共有することはできません。図 5「1 つの VM ディスクに割り当てられた 3 つのサブディスクの例」に、3 つのサブディスクから成る VM ディスクを示します。この VM ディスクは 1 つの物理ディスクに割り当てられています。

図 5. 1 つの VM ディスクに割り当てられた 3 つのサブディスクの例



サブディスクに割り当てられていない VM ディスク領域は空き領域です。空き領域を使用して新しいサブディスクを作成できます。

Volume Manager リリース 3.0 以降では、サブディスクにボリュームを含めることができるようになりました。それ以前のバージョンでは、サブディスクにボリュームを含めることはできませんでした。詳細については、「階層化ボリューム (25 ページ)」を参照してください。

プレックス

Volume Manager では、サブディスクを使用してプレックス (ミラー) と呼ばれる仮想オブジェクトを構成します。プレックスは、1 つ以上の物理ディスク上にある 1 つ以上のサブディスクから構成されます。サブディスク上のデータを編成してプレックスを構成するには、次の方法を実行します。

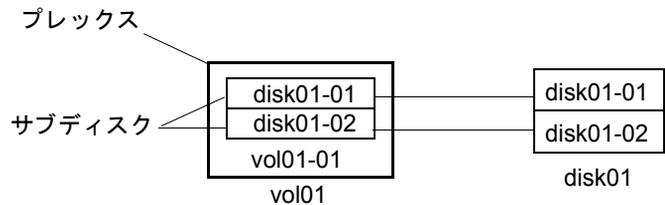
- ◆ コンカチネーション
- ◆ ストライピング (RAID-0)
- ◆ パリティ付きストライピング (RAID-5)
- ◆ ミラーリング (RAID-1)

注 RAID (Redundant Array of Independent Disks) とは、記憶容量の一部を使用してディスク アレイ内のデータに関する複製情報を格納しておくタイプのディスク アレイを指します。この複製データを使用すれば、ディスクに障害が発生したときにデータを復元することができます。

図 6 「2 つのサブディスクから成るプレックスの例」に、2 つのサブディスクから成るプレックスを示します。

コンカチネーション、ストライピング (RAID-0)、RAID-5、およびミラーリング (RAID-1) については、「ボリュームのレイアウト (10 ページ)」を参照してください。

図 6. 2 つのサブディスクから成るプレックスの例



ボリューム

ボリュームは 1 つ以上のプレックスで構成されます。各プレックスにはボリューム内のデータの一部が保持されます。仮想的な性質上、ボリュームは特定のディスクまたはディスク上の特定の領域に限定されることはありません。ボリューム構成は、**Volume Manager** のユーザ インタフェースを使用して変更することができます。構成を変更する場合は、ボリュームを使用しているアプリケーションまたはファイルシステムを中断させる必要はありません。たとえば、1 つのボリュームを複数のディスク上にミラーリングしたり、ボリュームを別のディスク ストレージに移動したりすることができます。

ボリュームには最大 32 個のプレックスを含めることが可能です。各プレックスには 1 つ以上のサブディスクを含めることができます。ボリュームは、少なくとも 1 つ以上のサブディスクから成る 1 つ以上のプレックスで構成され、そこに全データが格納されている必要があります。ボリューム内のすべてのサブディスクは、同じディスク グループに属している必要があります。

複数のプレックスを持つボリュームは「ミラーリング」され、データのミラー イメージを含みます。各プレックスには、ボリューム データの完全なコピーが含まれます。ミラーボリュームの詳細については、「ミラーリング (RAID-1) (18 ページ)」を参照してください。

Volume Manager では、ボリューム名として vol##、ボリューム内のプレックス名として vol##-## というデフォルトの命名規則を使用します。ボリュームには、識別しやすい名前をつけてください。図 7「1つのプレックスから成るボリュームの例」に、1つのプレックスから成るボリュームを示します。

図 7. 1つのプレックスから成るボリュームの例

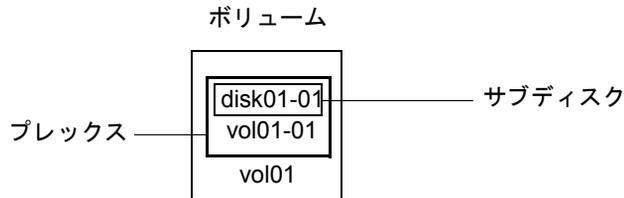


図 7 のボリューム vol01 には、次のような特徴があります。

- ◆ プレックス vol01-01 を 1 つ含んでいる。
- ◆ プレックスにはサブディスク disk01-01 を 1 つ含んでいる。
- ◆ サブディスク disk01-01 は VM ディスク disk01 から割り当てられている。

図 8「2つのプレックスから成るボリュームの例」に、2つのプレックスから成るミラーボリュームを示します。

図 8. 2つのプレックスから成るボリュームの例

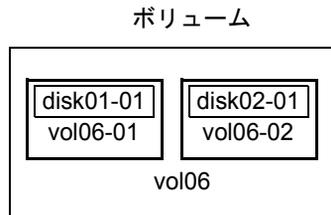


図 8 のボリューム vol06 には、次のような特徴があります。

- ◆ 2つのプレックス vol06-01 および vol06-02 を含んでいる。
- ◆ 各プレックスにサブディスクが 1 つずつ含まれている。
- ◆ 各サブディスクは別々の VM ディスク (disk01 および disk02) から割り当てられている。

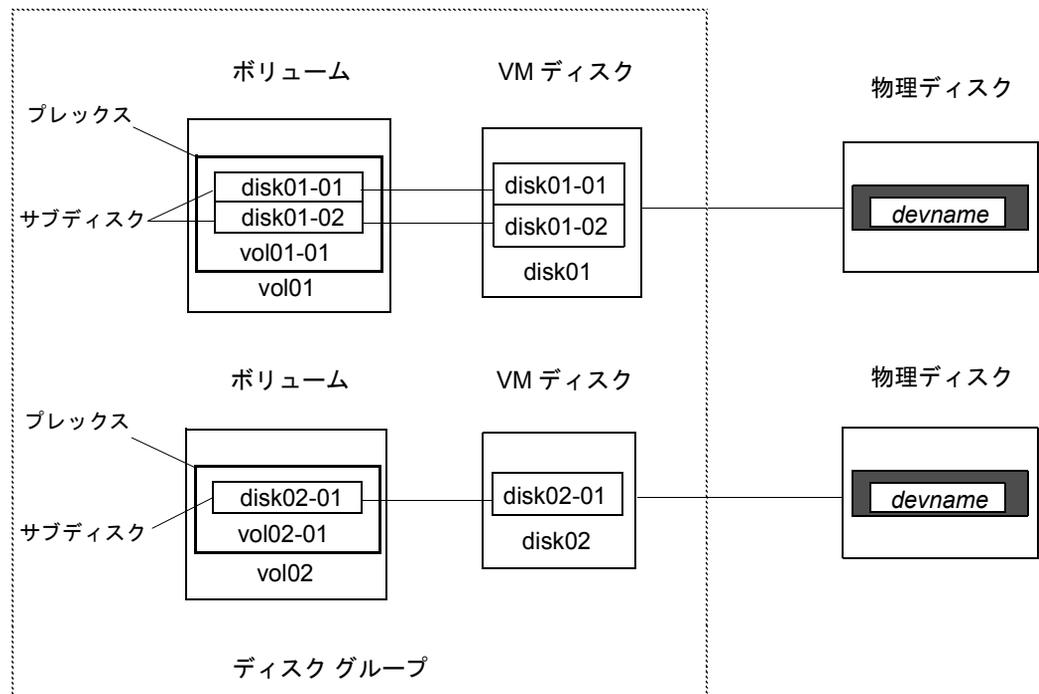
Volume Manager の仮想オブジェクト間の関係

Volume Manager では複数の仮想オブジェクトを組み合わせ、ボリュームを構築します。ボリュームに関連する仮想オブジェクトには、VM ディスク、ディスク グループ、サブディスク、プレックスなどがあります。Volume Manager のそれらのオブジェクトの間には次のような関係があります。

- ◆ Volume Manager ディスクはディスク グループにグループ化されている。
- ◆ プレックスは1つ以上のサブディスクを連結して形成されている（各サブディスクはディスクの特定の領域を表す）。
- ◆ ボリュームは1つ以上のプレックスで構成されている。

図9「Volume Manager オブジェクト間の関係」に、Volume Manager の仮想オブジェクト間の関係およびそれらの物理ディスクとの関連性について例を示します。この図は、2つのVM ディスク（disk01 および disk02）から成るディスク グループを示しています。disk01 には、1つのプレックスと2つのサブディスクを含むボリュームがあり、disk02 には1つのプレックスと1つのサブディスクを含むボリュームがあります。

図 9. Volume Manager オブジェクト間の関係



ボリュームのレイアウト

仮想オブジェクト内のデータを編成してボリュームを作成するには、下記のレイアウト方式があります。

- ◆ コンカチネーション
- ◆ ストライピング (RAID-0)
- ◆ RAID-5 (パリティ付きストライピング)
- ◆ ミラーリング (RAID-1)
- ◆ ミラーリング アンド ストライピング
- ◆ ストライピング アンド ミラーリング

上記の各レイアウト方式について、以降の節で説明します。

Volume Manager の仮想デバイスは、ボリュームに基づいて定義されます。ボリュームのレイアウトは、ボリュームと1つ以上のプレックスとの関係に基づいて定義されます。各プレックスはサブディスクと関係付けられます。ボリュームは仮想デバイス インタフェースを備え、そのインタフェースを介して **Volume Manager** クライアントはデータにアクセスすることができます。この論理構造に基づいてボリュームのアドレス空間が再構成され、それに基づいて実行時に入出力がリダイレクトされます。

ボリューム レイアウトに応じて、提供されるストレージ サービスのレベルは異なります。ボリューム レイアウトは、必要なストレージ サービスのレベルに合わせて、構成および再構成することができます。

Volume Manager の以前のリリースでは、サブディスクは VM ディスクに直接関係付けることしかできませんでした。これにより、サブディスクで、VM ディスクの共有領域に裏付けのある一連のストレージの領域を定義することができました。VM ディスクは、アクティブな場合、下位のディスクに関連付けられます。**Volume Manager** は、このようにして、論理オブジェクトを物理オブジェクトに対応付け、データを安定的なストレージに格納します。

ボリューム レイアウトと実際のデータ格納場所である物理ディスクの組み合わせによって、所定の仮想デバイスから利用できるデータ ストレージ サービスが決まります。

Volume Manager の 3.0 以降のリリースでは、「階層化ボリューム」を構成する際には、サブディスクを従来通りに VM ディスクに関係付けることも、ストレージ ボリュームと呼ばれる新しい論理オブジェクトに関係付けることもできるようになりました。ストレージ ボリュームは、最上位ボリュームに類似したレイアウトを使用して繰り返し下層へとマッピングしていきます。最終的に、「最下位」のレベルで、VM ディスクならびに接続されている物理ストレージへの関連付けが必要となります。

階層化ボリュームでは、論理構成要素を様々な組み合わせることが可能です。その中には仮想デバイスの設定に望ましいものもあります。しかし、階層化ボリュームをコマンドレベルから自由に使用できるようにすると、結果的に管理が難しくなる恐れがあります。そこで、Volume Manager 3.0 リリースでは、あらかじめ構成が定義された階層化ボリュームがいくつか用意されています。

これらの定義済み構成は、組み込まれているルールに基づいて、指定された制約条件の範囲内で望ましいサービスレベルに自動的に適合します。この自動的構成の設定は、現在の設定に対して処理を行っている現在のコマンド呼び出しの「最大効果」に基づいて実現されます。

一連の仮想デバイスを使用して望ましいデータストレージサービスを実現するには、適切な一連の VM ディスクを含むディスクグループを作成し、複数の設定コマンドを実行する必要があります。

Volume Manager 3.0 リリースでは、一連のレイアウトと管理インターフェースを使用して初期設定とオンライン再設定を最大限に行い、この作業をより容易にし、システムをより決定的なものにします。

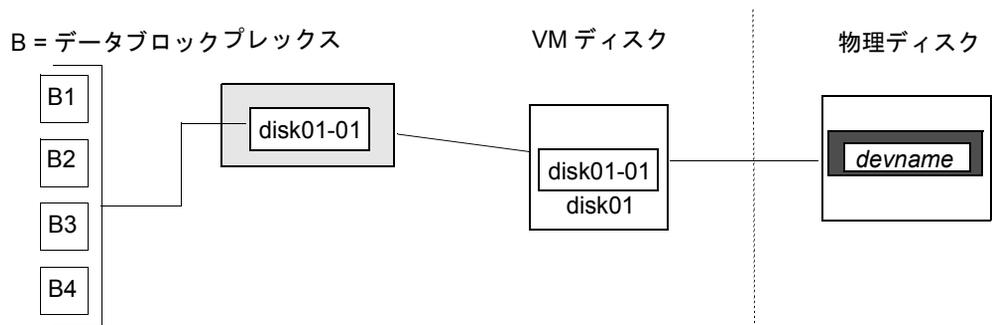
コンカチネーション

コンカチネーションとは、プレックス内の1つ以上のサブディスクにデータを直線的に対応付けることです。コンカチネイテッドプレックスのすべてのデータに順次アクセスするには、まず1番目のサブディスクのデータが先頭から末尾までアクセスされます。残りのサブディスクのデータも順々に先頭から末尾までアクセスされ、最後のサブディスクの末尾で終了します。

コンカチネイテッドプレックス内のサブディスクは、物理的に隣接している必要はなく、複数の VM ディスクに属していてもかまいません。複数の VM ディスク上に散在するサブディスクを連結することを、スパンニングと呼びます。

図 10「コンカチネーションの例」に、1つのサブディスクでのコンカチネーションを示します。

図 10. コンカチネーションの例



1つのディスク上に、プレックスに必要な隣接した領域が不足している場合は、複数のサブディスクでコンカチネーションを使用することができます。このコンカチネーション形式は、ディスク間で負荷を平準化したり、個々のディスク上でのヘッドの動作を最適化したりする場合に使用します。

図 11 「コンカチネーション構成内のボリュームの例」は、コンカチネーション構成内のボリュームを示したものです。

図 11. コンカチネーション構成内のボリュームの例

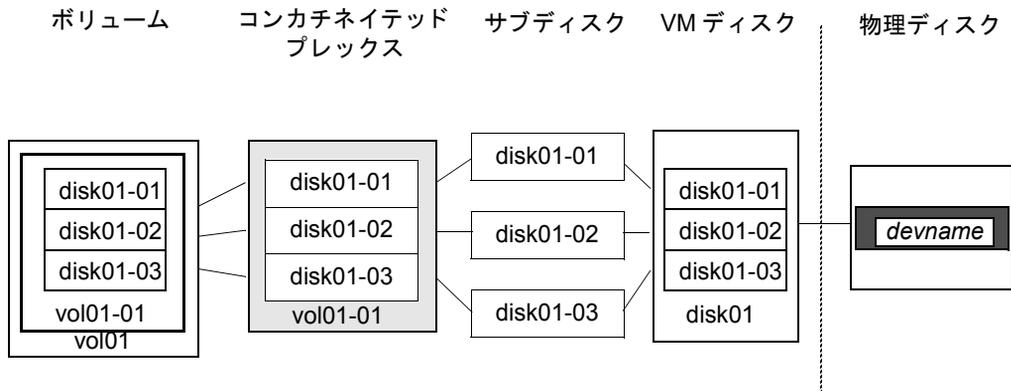
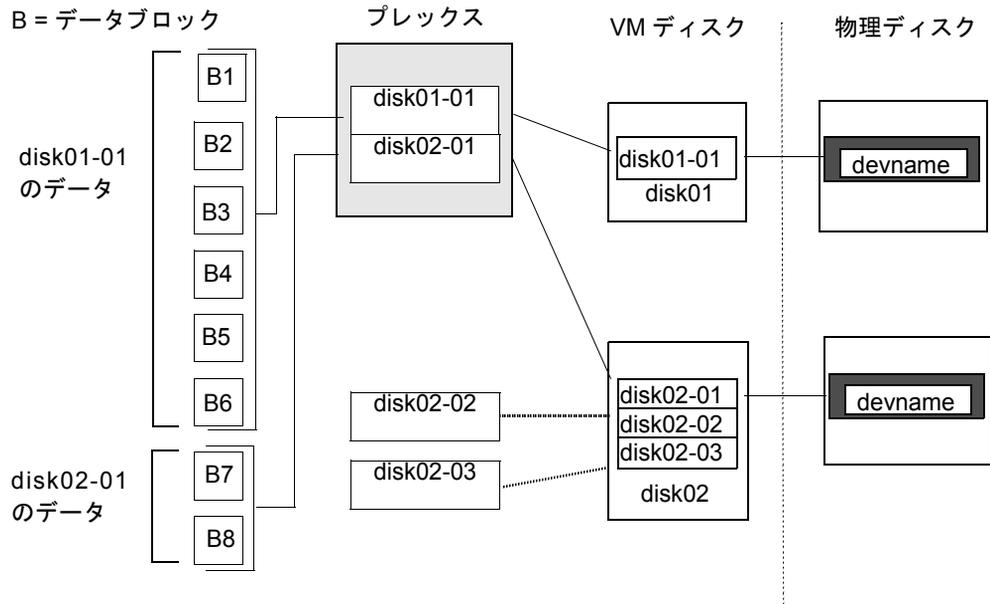


図 12 の例では、最初の 6 つのデータブロック (B1 から B6) が、VM ディスク `disk01` が割り当てられているディスク上のほとんどの領域を使用します。これには、VM ディスク `disk01` 上のサブディスク `disk01-01` 上の領域のみが必要となります。ただし、最後の 2 つのデータブロック (B7 および B8) は、VM ディスク `disk02` が割り当てられているディスク上の領域の一部しか使用しません。

VM ディスク `disk02` 上の残りの空き領域は、ほかの用途に使用できます。この例では、サブディスク `disk02-02` および `disk02-03` は、ほかのディスク管理タスクで使用可能になっています。

図 12「スパンニングの例」に、スパンプレックス内の2つのサブディスクに分散したデータを示します。

図 12. スパンニングの例



注意 1つのプレックスを複数ディスクにわたってスパンニングすると、1件のディスク障害が発生した結果として、ボリューム全体が障害に陥る可能性が高くなります。1件のディスクの障害がボリューム全体の障害を引き起こす可能性を低くするには、ミラーリングまたは RAID-5 (後の説明を参照) を適用してください。



ストライピング (RAID-0)

ストライピング (RAID-0) とは、複数の物理ディスク間を順番につなぐように、データをマッピングすることです。ストライプ ブレックスは、複数の物理ディスクに分散された複数のサブディスクから構成されます。データはストライプ ブレックスのサブディスクの間で順番に均等に割り当てられます。

サブディスクは「カラム」にグループ化されます。その際、各物理ディスクは1つのカラムに限定されます。各カラムには1つ以上のサブディスクが含まれます。それらのサブディスクの出所は、複数の物理ディスクにわたっていてもかまいません。カラムごとのサブディスクの数およびサイズは異なります。必要に応じて、カラムにサブディスクを追加することができます。

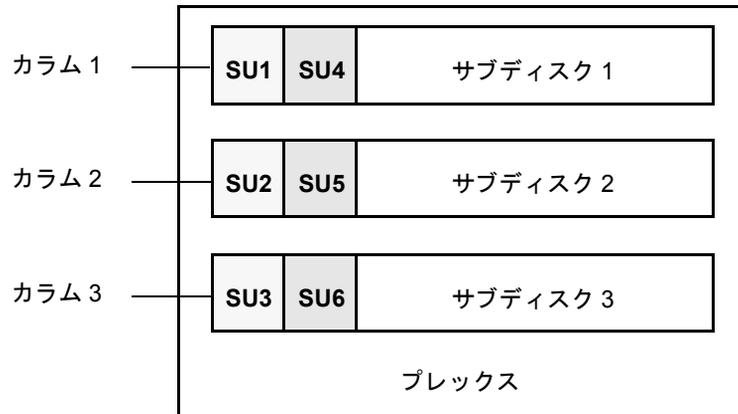
注意 1つのボリュームをストライピング、つまり、複数のディスクにわたってボリュームを分割すると、1件のディスクの障害がボリューム全体の障害につながる確率が高くなります。たとえば、5つのボリュームを5つのディスクにわたってストライピングしている場合に、いずれか1つのディスクに障害が発生したときは、5つのボリュームすべてをバックアップからリストアする必要が生じます。各ボリュームが別々のディスク上にある場合は、1つのボリュームをリストアするだけで済みます。1つのディスクの障害が多数のボリュームの障害を引き起こす可能性を低くするには、ミラーリングまたは RAID-5 を使用してください。

データは、均等なサイズ単位 (ストライプ ユニットと呼ばれるストライプの単位) に分けられ、各カラムに順番に割り当てられます。各ストライプ ユニットは、ディスク上の連続するブロックの集まりです。デフォルトのストライプ ユニットは64キロバイトです。

たとえば、ストライプ ブレックス内に3つのカラムと6つのストライプ ユニットがある場合、データは図 13 「3つのディスク (カラム) にわたるストライピング」に示すように、3つの物理ディスクにわたって次のようにストライピングされます。

- ◆ 1番目と4番目のストライプ ユニットがカラム1に割り当てられる。
- ◆ 2番目と5番目のストライプ ユニットがカラム2に割り当てられる。
- ◆ 3番目と6番目のストライプ ユニットがカラム3に割り当てられる。

図 13. 3つのディスク（カラム）にわたるストライピング



SU = ストライプ ユニット

ストライプは、すべてのカラムにわたって同じ位置にあるストライプ ユニットの集まりで構成されます。図 13 では、ストライプ ユニット 1、2、および 3 が 1 つのストライプを構成しています。

順に見ていくと、1 番目のストライプは次の要素で構成されています。

- ◆ カラム 1 のストライプ ユニット 1
- ◆ カラム 2 のストライプ ユニット 2
- ◆ カラム 3 のストライプ ユニット 3

2 番目のストライプは次の要素で構成されています。

- ◆ カラム 1 のストライプ ユニット 4
- ◆ カラム 2 のストライプ ユニット 5
- ◆ カラム 3 のストライプ ユニット 6

ストライピングはカラムの長さ分（すべてのカラムが同じ長さの場合）、または最も短いカラムの末尾に達するまで続行されます。それよりもカラムが長いサブディスクの末尾に残った領域は、未使用領域となります。

ストライピングは、物理ディスクに対する大量データの読み書きを実行する必要がある場合に、複数のディスクへのパラレルデータ転送を行うことによって、効果を発揮します。また、複数のディスクにわたるマルチユーザ アプリケーションによる入出力の負荷を分散する場合にも有効です。



図 14「カラムごとに1つのサブディスクを含むストライププレックスの例」に、サブディスクを1つずつ含む同じサイズのカラムが3つあるストライププレックスを示します。各物理ディスクにはカラムが1つずつ含まれます。

図 14. カラムごとに1つのサブディスクを含むストライププレックスの例

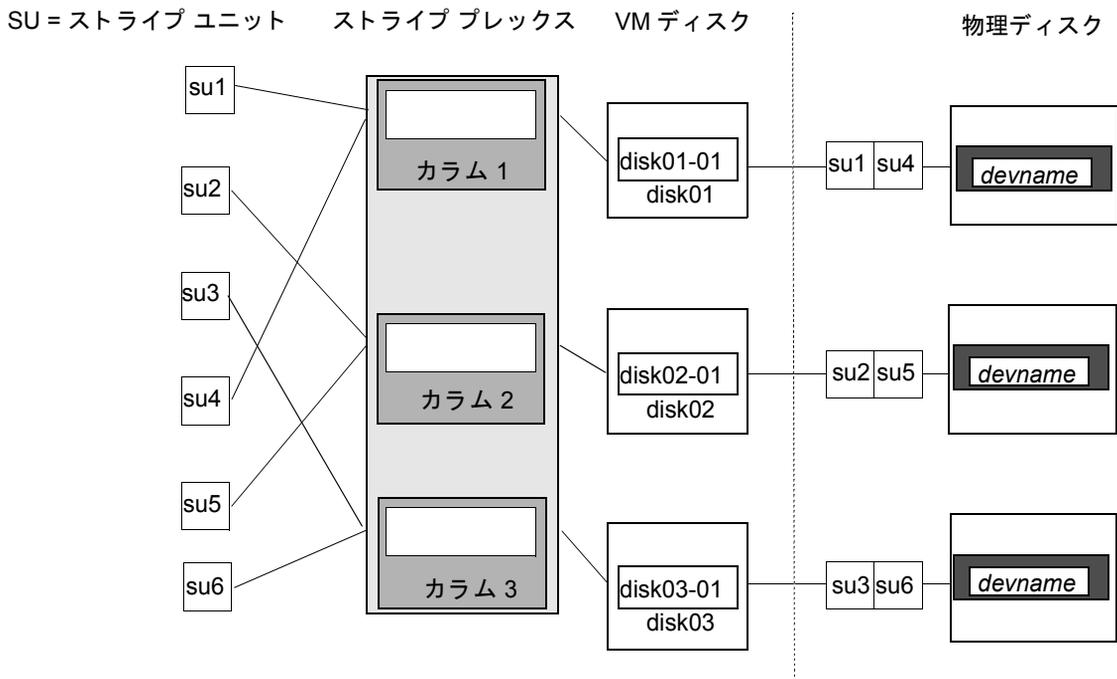
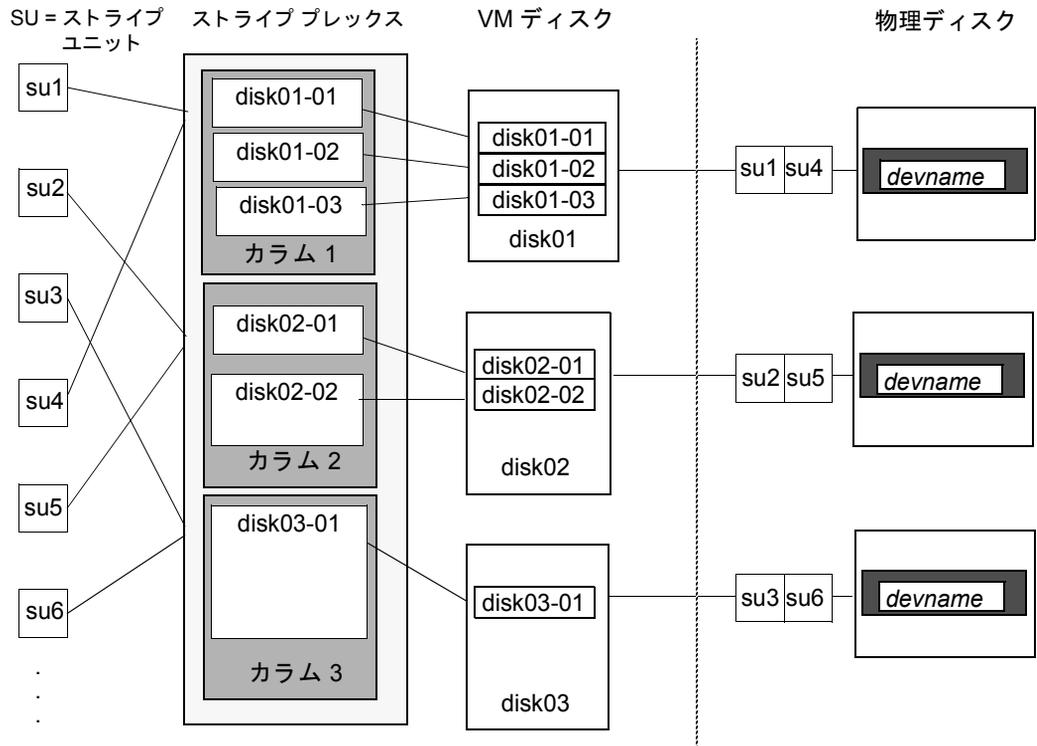


図 14 の例は、VM ディスク上の全領域を占めている 3 つのサブディスクを示しています。ストライププレックスの各サブディスクに VM ディスクの一部分のみを割り当てることも可能です。その場合、空き領域をほかのディスク管理タスク用に残しておくことができます。

図 15「カラムごとにコンカチネイテッド サブディスクを含むストライププレックスの例」に、サイズの異なるサブディスクを含むカラムが 3 つあるストライププレックスを示します。各カラムに含まれるサブディスクの数が異なります。各物理ディスクにはカラムが 1 つずつ含まれます。ストライププレックスは、ストライピングする各 VM ディスクの 1 つのサブディスクを使用して作成されます。同じディスクの異なる領域または別のディスクから領域を割り当てることもできます (プレックスのサイズが大きくなった場合など)。カラムには、異なる VM ディスクのサブディスクを含めることもできます。

図 15. カラムごとにコンカチネイテッド サブディスクを含むストライプ プレックスの例



RAID-5

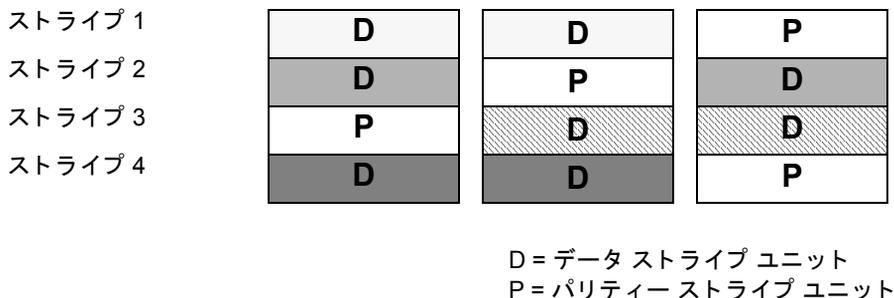
RAID-5 では、パリティを使用して、データに冗長性を持たせます。パリティとは、障害発生後にデータを復元するときに使用する計算値です。RAID-5 ボリュームへのデータ書き込み中に、データの排他的論理和 (XOR) 演算を実行することによって、パリティを計算します。計算されたパリティは、ボリュームに書き込まれます。RAID-5 のいずれかの部分に障害が発生した場合は、残りのデータとパリティを使って該当する部分のデータが復元されます。

RAID-5 ボリュームでは、ボリューム内でデータの冗長性が維持されます。RAID-5 ボリュームではデータのコピーと計算されたパリティが、複数のディスクにわたって「ストライピングされた」プレックス内に保持されます。ディスクに障害が発生した場合は、RAID-5 ボリュームでは、パリティを使用してデータを復元します。このレイアウトに、コンカチネーションとストライピングを組み合わせることもできます。

RAID-5 ボリュームでは、ログを記録することによって、リカバリ時間を最小に抑えることができます。RAID-5 ボリュームでは、RAID-5 ログを使用して、書き込み中のデータとパリティのコピーを保持します。RAID-5 のログ機能はオプションであり、RAID-5 ボリュームに作成時に設定することも、後から追加することもできます。

図 16「RAID-5 モデルでのパリティの配置」に、RAID-5 アレイ構成におけるパティの配置を示します。各ストライプは、パティ ストライプ ユニットの含むカラムとデータを含むカラムで構成されます。パティはアレイ内のすべてのディスクに分散されます。それによって、1つのパティ ディスクでデータの受け入れが可能になるのを待たずに書き込むことができるため、大量の独立した書き込みに必要な時間が短縮されます。

図 16. RAID-5 モデルでのパティの配置



詳細については、「Volume Manager および RAID-5 (20 ページ)」を参照してください。

ミラーリング (RAID-1)

ミラーリングとは、複数のミラー (プレックス) を使用して、ボリュームに保持されている情報を複製することです。物理ディスクに障害が発生した場合、そのディスク上のプレックスは使用できなくなります。しかし、影響を受けていないミラーを使用することによって、システムの稼働は続行されます。ボリュームの中のプレックスは1つでも問題ありませんが、データの冗長性を確保するには少なくとも 2 つ以上のプレックスが必要です。各プレックスには、冗長性を確保するために、異なるディスクの領域を含める必要があります。

多くのディスクにわたってストライピングまたはスパンニングする場合、それらのディスクのいずれかで障害が発生すると、プレックス全体が使用できなくなる可能性があります。したがって、ボリュームの信頼性 (および可用性) を向上させるために、ストライピングまたはスパンニングしたミラーリングを適用することをお勧めします。

ミラーリング アンド ストライピング (RAID-1 + RAID-0)

Volume Manager では、ミラーリングにストライピングを組み合わせることができます。この組み合わせは、ミラー ストライプ レイアウトと呼ばれます。ミラーリングしたボリュームをストライピングすると、データを複数のディスクに分散 (ストライピング) させながら、データの冗長性 (ミラー) を確保することができます。

ミラーリング アンド ストライピングを効率よく行うには、ミラーとストライプ プレックスを別々のディスクから割り当てる必要があります。ミラーのレイアウト タイプには、コンカチネーションまたはストライピングを使用できます。

ストライピング アンド ミラーリング (RAID-0 + RAID-1)

Volume Manager では、ストライピングにミラーリングを組み合わせたことができます。この組み合わせは、ストライプミラーレイアウトと呼ばれます。以前のリリースでは、ミラーリングを主として、ストライピングを従にする必要がありました。今回のリリースからは、どちらを主とすることもできるようになりました。

ストライピングを主としミラーリングを従とすると、ストライプの各カラムがミラーリングされます。カラムの中に複数のサブディスクがあるサイズの大きいストライプの場合は、各サブディスクが個別にミラーリングされます。このレイアウトでは、冗長性が高まり、障害発生時のリカバリ時間が短縮されます。

ミラー ストライプ レイアウトでディスク障害が発生した場合は、プレックス全体が切り離されるので、ボリューム全体の冗長性が損なわれます。ディスクを入れ替えた後で、プレックス全体をリカバリする必要があります。それには、かなりの時間を要します。ストライピングにミラーリングを準用した場合、ディスクに障害が発生したときに切り離されるのは、障害が発生したサブディスクだけです。したがって、ボリュームの該当個所の冗長性が失われるだけです。ディスクを入れ替えた後でリカバリが必要なのはボリュームの該当部分のみです。

ミラーリング アンド ストライピングに比べると、ストライピング アンド ミラーリングの方が、ディスクの障害に対するボリュームの耐性が強くなります。ディスクの障害発生時のリカバリ時間も、ストライピング アンド ミラーリングの方が短くなります。詳細については、「階層化ボリューム (25 ページ)」を参照してください。



Volume Manager および RAID-5

この節では、Volume Manager を使用した RAID-5 の実現方法について説明します。

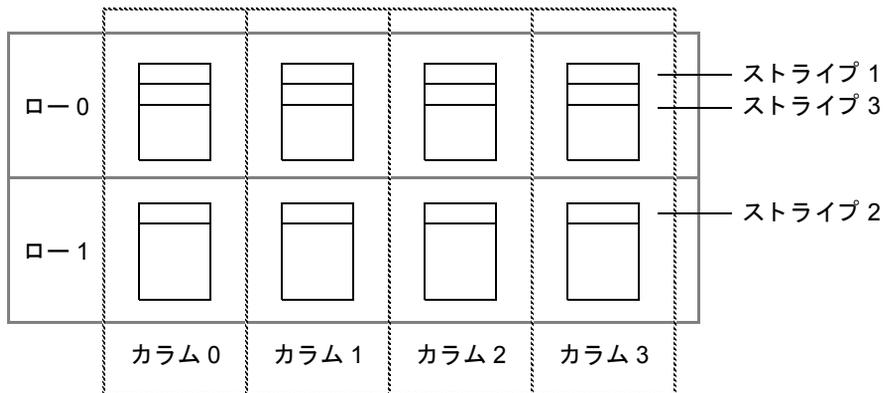
ミラーリング (RAID-1) と RAID-5 はどちらもデータに冗長性を持たせますが、その方法が異なります。ミラーリングでは、1つのボリューム内のデータの完全なコピーを複数作成することによって冗長性を確保します。ミラーリングされているボリュームに書き込まれたデータは、すべてのコピーに反映されます。ミラーリングされているボリュームの一部に障害が発生した場合、データの別のコピーが使用されます。

RAID-5 では、パリティを使用して、データに冗長性を持たせます。パリティとは、障害発生後にデータを復元するとき使用する計算値です。RAID-5 ボリュームへのデータ書き込み中に、データの排他的論理和 (XOR) 演算を実行することによって、パリティを計算します。計算されたパリティは、ボリュームに書き込まれます。RAID-5 のいずれかの部分に障害が発生した場合は、残りのデータとパリティを使って該当する部分のデータが復元されます。

従来の RAID-5 アレイ

従来の RAID-5 アレイは、ローとカラムに編成された複数のディスクで構成されます。カラムは、アレイ内で同じ序数の位置にあるディスクの集まりです。ローは、パリティストライプの全幅をサポートするために必要な最低数のディスクの集まりです。図 17「従来の RAID-5 アレイ」は、従来の RAID-5 アレイのローとカラムの配列を示しています。

図 17. 従来の RAID-5 アレイ

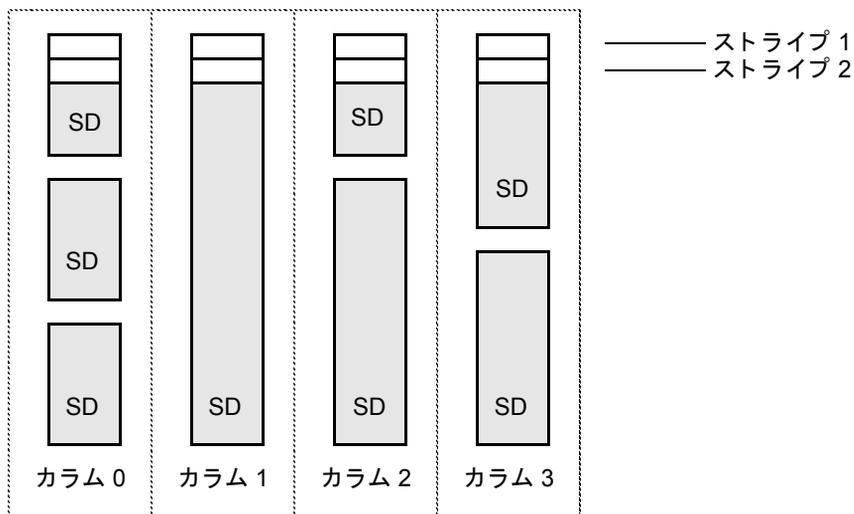


こうした従来のアレイ構造は、ローを追加することによってデータの増加に対処しています。ストライピングを実現するには、1番目のストライプをロー0のディスクに適用し、2番目のストライプをロー1のディスクに適用し、3番目のストライプをロー0のディスクに適用します。このタイプのアレイでは、すべてのディスク、カラム、およびローのサイズを等しくする必要があります。

Volume Manager RAID-5 アレイ

Volume Manager の RAID-5 アレイの構造は従来のアレイ構造とは異なります。ディスクをはじめとするオブジェクトが仮想であるという性質上、Volume Manager ではローを使用せず、代わりに可変長のサブディスクで構成されるカラムを使用します（図 18「Volume Manager RAID-5 アレイ」を参照）。各サブディスクは、ディスクの特定の領域を示します。

図 18. Volume Manager RAID-5 アレイ



SD = サブディスク

Volume Manager の RAID-5 アレイの構造では、カラムごとにサブディスクの数が異なってもかまいません。あるカラムに属する複数のサブディスクが別々の物理ディスクに属していてもかまいません。また必要に応じて、カラムにサブディスクを追加することもできます。ストライピング（「ストライピング (RAID-0) (14 ページ)」を参照）を実現するには、1 番目のストライプを各カラムの先頭にあるサブディスクにわたって適用し、その下に次のストライプを適用し、カラムの末尾に達するまで同様に適用していきます。各ストライプでは、等しいサイズのストライプユニットが各カラム内に配置されます。RAID-5 のデフォルトのストライプユニットサイズは 16 キロバイトです。

注 RAID-5 ボリュームのミラーリングは、現在のバージョンではサポートされていません。



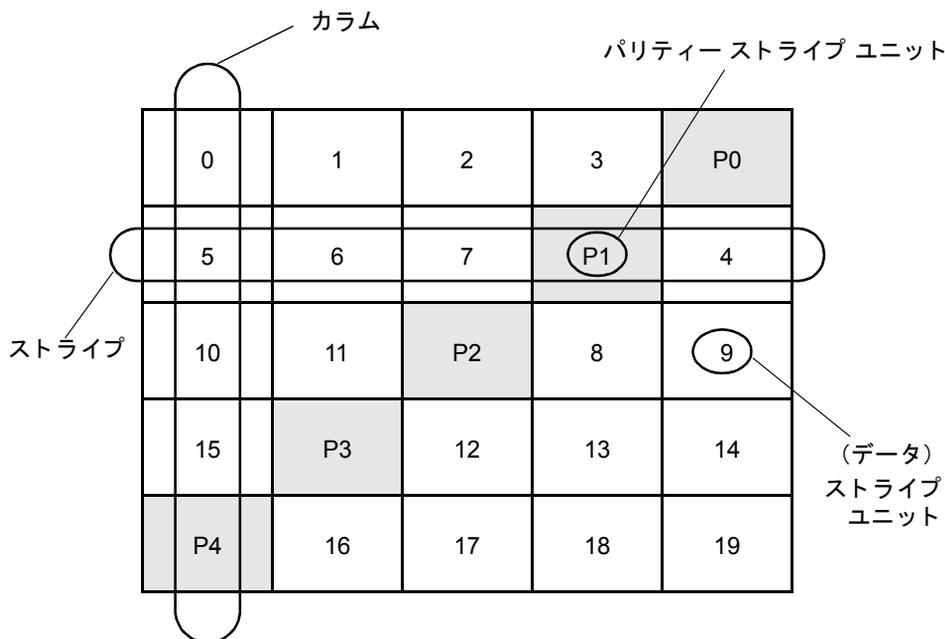
左対称レイアウト

RAID-5 アレイの設定で利用できるデータとパリティのレイアウトには、何種類かあります。Volume Manager の RAID-5 では、左対称レイアウトを採用しています。左対称パリティレイアウトは、ランダムな入出力操作および大量の順次入出力操作の場合に、最適なパフォーマンスを実現します。パフォーマンスの面から考えると、レイアウトの選択は、カラムの数やストライプユニット サイズの選定ほど重要ではありません。

左対称レイアウトでは、データとパリティの両方が複数のカラムにわたってストライピングされ、パリティは各ストライプごとに別々のカラムに配置されます。1 番目のパリティ ストライプユニットは、ストライプの右端のカラムに配置されます。以降の各パリティ ストライプユニットは、次のストライプ内で、直前のパリティ ストライプユニットより 1 つ左のカラムに配置されます。カラムよりもストライプが多い場合、パリティ ストライプユニットは、右端のカラムに再び配置されます。

図 19 「左対称レイアウト」は、5 つのディスク（カラムごとに 1 つ）から成る左対称パリティレイアウトを示しています。

図 19. 左対称レイアウト



各ストライプでは、データはパリティ ストライプユニットから右方向へ編成されます。図 19 では、最初のストライプのデータ編成は P0 から始まり、ストライプユニットの 0 から 3 の順に続きます。2 番目のストライプのデータ編成は P1 から始まり、ストライプユニットの 4、続いてストライプユニットの 5 から 7 の順に続きます。残りのストライプも同様に編成されます。

各パリティーストライプユニットには、同じストライプ内のデータストライプユニットデータに適用された排他的論理和（XOR）演算の結果の値が収められます。ハードウェアまたはソフトウェアの障害によって、あるカラムに対応するディスク上のデータにアクセスできなくなった場合、データを復元することができます。そのためには、同じカラム内の残りのデータストライプユニットの内容を、それぞれのパリティーストライプユニットで（ストライプごとに）XOR 演算します。

たとえば、図 19 の左端のカラムに対応するディスクで障害が発生した場合、そのボリュームは縮退モードに入ります。縮退モードでは、ストライプユニット 1 から 3 をパリティーストライプユニット P0 と XOR 演算してストライプユニット 0 を復元し、続いてストライプユニット 4、6、および 7 をパリティーストライプユニット P1 と XOR 演算してストライプユニット 5 を復元し、以下同様の処理を続けてデータを復元します。

注 RAID-5 レイアウトでプレックス内の複数のカラムに障害が発生した場合、そのボリュームは切り離されます。切り離されたボリュームは、読み取り要求や書き込み要求に応じることができなくなります。障害からリカバリした後で、そのカラムのデータをバックアップから復元する必要があります。

ロギング

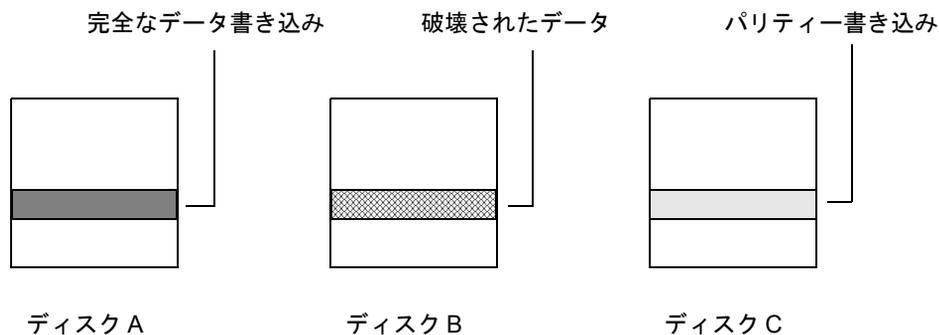
ロギング（記録）は、リカバリデータの破損を防止するために行います。新しいデータやパリティのログが永続性のあるデバイス（ディスク常駐ボリューム、または非揮発性 RAM）上に作成されます。その後、新しいデータやパリティがディスクに書き込まれます。

ログを記録していないと、ディスクとシステムの両方に障害が発生した場合、アクティブな書き込み処理に含まれていないデータが損失したり、知らないうちに破損する可能性があります。このように二重に障害が発生した場合は、ディスクのデータ部に書き込まれるデータ、またはパリティ部に書き込まれるパリティが実際に書き込まれたかどうかを確認する方法はありません。そのため、壊れたディスクのリカバリを実行しても結果は壊れている可能性があります。



図 20 「不完全な書き込み」は、ディスク B のリカバリの成否が、ディスク A のデータおよびディスク C のパリティの両方の書き込みが完了しているかどうかによって左右されることを示しています。この図では、ディスク A へのデータの書き込みは完了しているが、パリティの書き込みが不完全なため、ディスク B 上のデータを正しく復元できないことを示しています。

図 20. 不完全な書き込み



このような障害は、データをアレイに書き込む前に、すべての書き込みのログを記録しておくことによって防止できます。この方法で、ログを適用することによってデータとパリティを修正してから、障害が発生したドライブの内容を復元することができます。

ログは、ログ プレックスとして割り当てることによって、RAID-5 ボリュームと関連付けることができます。各 RAID-5 ボリュームに複数のログ プレックスを割り当てて、ログ領域をミラーリングすることもできます。

階層化ボリューム

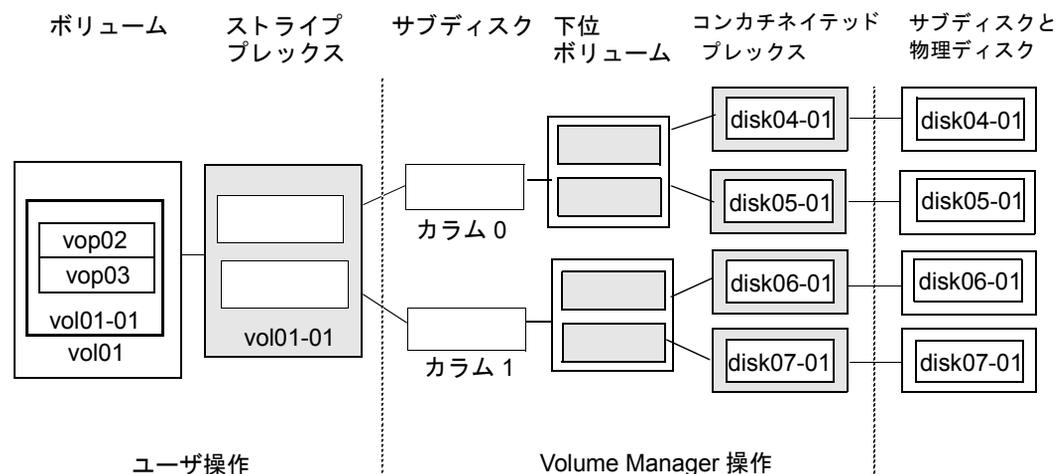
階層化ボリュームとは、ボリュームの上部に構築する Volume Manager の仮想オブジェクト データのことです。階層化ボリューム構造は標準のボリューム構造に比べて、障害に対する耐性や冗長性に優れています。たとえば、ストライピングおよびミラーリングした階層化ボリュームでは、各ミラー（プレックス）が扱うストレージの容量が少ないため、標準ミラー ボリュームよりすばやくリカバリできます。図 21 「ストライピングおよびミラーリングした階層化ボリュームの例」に階層化ボリュームの例を示します。

注 階層化ボリュームは、Volume Manager 3.0 以降のバージョンでサポートされており、それ以前のバージョンではサポートされていません。また、Volume Manager 3.0 以降のバージョンではボリューム（ストレージ ボリューム）上にサブディスクを構築することができますが、それ以前のバージョンではサブディスクはサポートされていません。

ユーザ処理は、階層化ボリュームの最上位ボリュームに対してのみ実行できます。階層化ボリュームを切り離したり、内部構造を操作して下位のボリューム上で別の操作を実行することはできません。ボリュームおよびストライプ プレックスを含むユーザ操作領域から、ユーザは必要な操作をすべて実行できます。図 21 に示すように、「ユーザ操作」領域内のボリュームおよびストライプ プレックスに対して、Volume Manager の通常の処理を実行することができます。

図 21 の「Volume Manager 操作」領域には、各ボリュームが内部的にミラーリングされている下位のボリューム上に構築された、2つのカラムを含むサブディスクが示されています。階層化ボリュームは Volume Manager 内の基礎構造です。この構造に基づいて、一定の機能を Volume Manager に追加することができます。下位のボリュームは Volume Manager によって独占的に使用されるものであり、ユーザが操作するようには設計されてはいません。ここでは、階層化ボリュームの仕組みと Volume Manager での役割について理解するために、下位ボリューム構造について説明します。

図 21. ストライピングおよびミラーリングした階層化ボリュームの例



システム管理者は、トラブルシューティングやその他の処理（データの特定ディスクへの配置など）を実行するために、階層化ボリューム構造を操作する必要がある場合があります。Volume Manager は階層化ボリュームを使用して、次の処理や操作を実行します。

- ◆ ストライプ ミラー (vxassist マニュアル ページを参照)
- ◆ コンカチネイテッド ミラー (vxassist マニュアル ページを参照)
- ◆ オンライン再レイアウト (vxrelayout および vxassist マニュアル ページを参照)
- ◆ RAID-5 サブディスクの移動 (vxsd マニュアル ページを参照)
- ◆ RAID-5 のスナップショット (vxassist マニュアル ページを参照)

Volume Manager およびオペレーティング システム

Volume Manager は、オペレーティング システムとデータ管理システム（ファイル システム、データベース管理システムなど）の間で、サブシステムとして機能します。

ディスクを Volume Manager の管理下に入れるためには、オペレーティング システム デバイス インタフェースを介してディスクにアクセスできる必要があります。Volume Manager は、オペレーティング システム インタフェース サービスの上位に置かれたサブシステムです。したがって、オペレーティング システムの物理ディスクへのアクセス方法に左右されます。

Volume Manager は、下記の点に関してオペレーティング システムに依存します。

- ◆ オペレーティング システム（ディスク） デバイス
- ◆ デバイス ハンドル
- ◆ VM ディスク
- ◆ Volume Manager 動的マルチパス (DMP) メタデバイス

Volume Manager のレイアウト

Volume Manager の仮想デバイスは、ボリュームに基づいて定義されます。ボリュームのレイアウトは、ボリュームと1つ以上のプレックスとの関係に基づいて定義されます。各プレックスはサブディスクと関係付けられます。ボリュームは仮想デバイス インタフェースを備え、そのインタフェースを介して Volume Manager クライアントはデータにアクセスすることができます。この論理構造に基づいてボリュームのアドレス空間が再構成され、それに基づいて実行時に入出力がリダイレクトされます。

ボリューム レイアウトに応じて、提供されるストレージ サービスのレベルは異なります。ボリューム レイアウトは、必要なストレージ サービスのレベルに合わせて、設定および再設定することができます。

Volume Manager の以前のリリースでは、サブディスクは VM ディスクに直接関係付けることしかできませんでした。これにより、サブディスクで、VM ディスクの共有領域に裏付けのある一連のストレージの領域を定義することができました。VM ディスクは、アク

ティブな場合、下位のディスクに関連付けられます。Volume Manager は、このようにして、論理オブジェクトを物理オブジェクトに対応付け、データを安定的なストレージに格納します。

ボリュームレイアウトと実際のデータ格納場所である物理ディスクの組み合わせによって、所定の仮想デバイスから利用できるデータストレージサービスが決まります。

Volume Manager の 3.0 以降のリリースでは、「階層化ボリューム」を構成する際には、サブディスクを従来通りに VM ディスクに関係付けることも、ストレージボリュームと呼ばれる新しい論理オブジェクトに関係付けることもできるようになりました。ストレージボリュームは、最上位ボリュームに類似したレイアウトを使用して繰り返し下層へとマッピングしていきます。最終的に、「最下位」のレベルで、VM ディスクならびに接続されている物理ストレージへの関連付けが必要となります。

階層化ボリュームでは、論理構成要素を様々に組み合わせることが可能です。その中には仮想デバイスの設定に望ましいものもあります。しかし、階層化ボリュームをコマンドレベルから自由に使用できるようにすると、結果的に管理が難しくなる恐れがあります。そこで、Volume Manager 3.0 リリースでは、あらかじめ構成が定義された階層化ボリュームがいくつか用意されています。

これらの定義済み構成は、組み込まれているルールに基づいて、指定された制約条件の範囲内で望ましいサービスレベルに自動的に適合します。この自動的構成の設定は、現在の設定に対して処理を行っている現在のコマンド呼び出しの「最大効果」に基づいて実現されます。

一連の仮想デバイスを使用して望ましいデータストレージサービスを実現するには、適切な一連の VM ディスクを含むディスクグループを作成し、複数の設定コマンドを実行する必要があります。

Volume Manager 3.0 リリースでは、一連のレイアウトと管理インタフェースを使用して初期設定とオンライン再設定を最大限に行い、この作業をより容易にし、システムをより決定的なものにします。



Volume Manager のユーザ インタフェース

この節では、VERITAS Volume Manager のユーザ インタフェースについて簡単に説明します。

ユーザ インタフェースの概要

Volume Manager では、次のユーザ インタフェースをサポートしています。

- ◆ Volume Manager Storage Administrator (VMSA)
- ◆ Storage Administrator は、Volume Manager へのグラフィカル ユーザ インタフェースです。Storage Administrator では、アイコン、メニュー、ダイアログ ボックスなどのビジュアル要素を使って Volume Manager のオブジェクトを操作します。また Storage Administrator は、一般的なファイル システム操作に関するインタフェースとしても機能します。Storage Administrator については、『VERITAS Volume Manager Storage Administrator Administrator's Guide』を参照してください。
- ◆ コマンド ライン インタフェース
Volume Manager コマンドには、単純なものから詳細なユーザ入力を必要とする複雑なものまで、さまざまな種類があります。通常、Volume Manager コマンドを使用するには、Volume Manager の概念を理解している必要があります。Volume Manager の概念については、この章で説明します。ほとんどの Volume Manager コマンドを実行する場合に、スーパーユーザなどの適切な権限が必要です。コマンド ライン インタフェースについては、『VERITAS Volume Manager Command Line Interface Administrator's Guide』を参照してください。
- ◆ Volume Man コマンド ログビューア `ager Support Operations`
[Volume Manager Support Operations] インタフェース (`vxdiskadm`) は、ディスクおよびボリュームを管理するためのメニュー方式のインタフェースです。`vxdiskadm` によって提示されるメイン メニューから、実行するストレージ管理タスクを選択できます。`vxdiskadm` については、『VERITAS Volume Manager Command Line Interface Administrator's Guide』を参照してください。

あるインタフェースで作成された Volume Manager オブジェクトには、別のインタフェースで作成されたオブジェクトとの互換性があります。

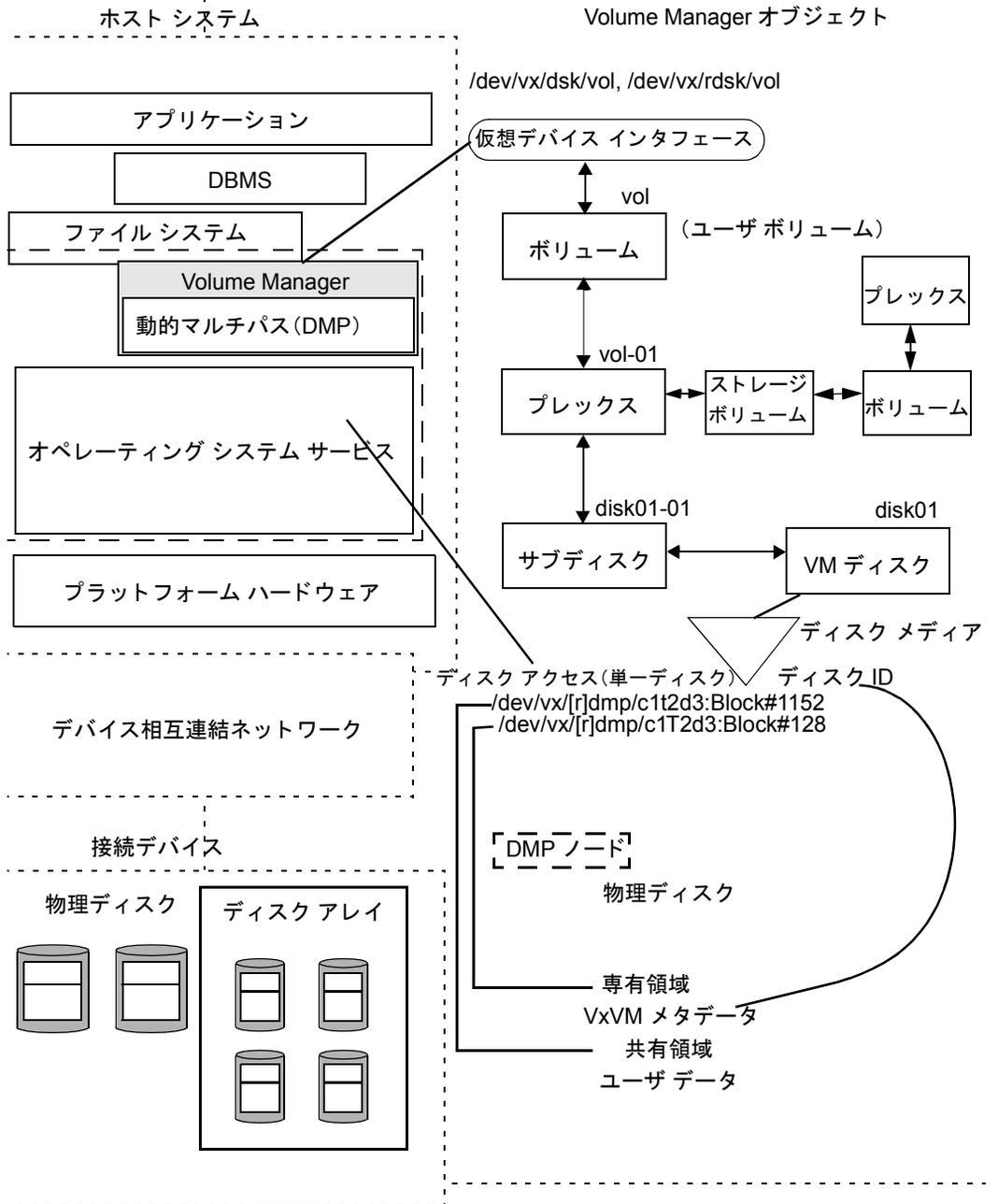
Volume Manager を使用する目的

Volume Manager では、データ管理の物理的な側面と論理的な側面を分けることによって、高度なデータ ストレージ サービスを提供します。Volume Manager は、ストレージに関する次の側面を制御することにより、データ ストレージ機能を向上させます。

- ◆ 領域 — 割り当ておよび使用
- ◆ パフォーマンス — データ伝達機能の向上
- ◆ データの可用性 — 継続操作および複数システムへのアクセス
- ◆ デバイスのインストール — サポートの集中化および最適化
- ◆ システム — 複数システムのサポートおよび専用 / 共有システムの監視

29 ページの図 22 「Volume Manager システムの概念」を参照してください。

図 22. Volume Manager システムの概念



ホスト システム上に **Volume Manager** をインストールしたら、**Volume Manager** オブジェクトの設定および使用を開始する前に、下記の作業を実行しておく必要があります。

- ◆ 物理ディスクの内容を **Volume Manager** の管理下に組み入れる。
- ◆ **Volume Manager** 下のディスクを集めてディスク グループを編成する。
- ◆ ディスク グループ領域を割り当てて論理ボリュームを作成する。

次の場合にのみ、物理ディスクの内容を **Volume Manager** の管理下に組み入れることができます。

- ◆ **Volume Manager** で物理ディスクを管理できる。
- ◆ 物理ディスクが別のストレージ マネージャの管理下でない。

Volume Manager は、その管理下にある物理ディスク（管理下ディスク）に識別情報を書き込みます。管理下ディスクは、物理ディスクを切り離した後やシステムに障害が発生した後でも識別できます。**Volume Manager** はディスク グループと論理オブジェクトを再編成して、障害を検出し、システムのリカバリを早めることができます。

はじめに

この章では、Volume Manager の設定に必要な作業および Volume Manager が動作するために実行する必要があるデーモンについて簡単に説明します。また、ストレージ管理を行うシステムを設定するために役立つガイドラインについても説明します。

Volume Manager および Storage Administrator をインストールおよび設定する方法の詳細については、『VERITAS Volume Manager Installation Guide』を参照してください。

この章では、次のトピックについて説明します。

- ◆ Volume Manager の初期化
- ◆ Volume Manager のデーモン
- ◆ システムの設定
- ◆ システム設定ガイドライン
- ◆ システムの保護



Volume Manager の初期化

Volume Manager を初期化するには、`vxinstall` プログラムを使用します。`vxinstall` は、指定されたディスクを Volume Manager の管理下に組み入れます。デフォルトでは、これらのディスクは `rootdg` ディスク グループ内に含まれています。最初に、`vxinstall` を使用し、1 つ以上のディスクを初期化して `rootdg` に含める必要があります。次に、`vxdiskadm` または `Storage Administrator` を使用し、追加のディスクを初期化またはカプセル化し、別のディスク グループ内に配置します。

パッケージのインストールが完了したら、次の手順で Volume Manager を初期化します。

1. スーパーユーザとしてログインします。
2. Volume Manager で管理しないディスクがある場合は、`disks.exclude` ファイルを作成します。`vxinstall` は、このファイルに含まれているディスクについては処理対象外とします。このファイルは `/etc/vx/disks.exclude` に置きます。
3. コントローラ上のすべてのディスクを Volume Manager で管理しない場合は、`cntrls.exclude` ファイルを作成します。このファイルは `/etc/vx/cntrls.exclude` に置きます。
4. コマンド `vxinstall` を入力して `vxinstall` を起動します。

`vxinstall` は次の処理を実行します。

- ◆ ライセンス情報の表示と確認のためのキーの要求
- ◆ システムに接続されているすべてのコントローラの確認および一覧表示
- ◆ 初期化プロセスの選択（クイック インストールまたはカスタム インストール）

クイック インストールでは、すべてのディスクの初期化またはカプセル化を選択できます。指定したコントローラ上の一部のディスクをカプセル化し、その他のディスクを初期化する場合は、カスタム インストールを実行してください。

カスタム インストールでは、Volume Manager の管理下に追加するディスクとその追加方法を制御できます。コントローラ上のすべてのディスクを初期化またはカプセル化するか、またはコントローラ上の一部のディスクを初期化して、その他のディスクをカプセル化することができます。

クイック インストールおよびカスタム インストール手順の詳細については、『*VERITAS Volume Manager Installation Guide*』を参照してください。

選択したシステムの設定方法によって、シャットダウンとリブートが必要であるかどうかが決まります。任意のディスクをカプセル化する場合は、リブートが必要になります。`vxinstall` の実行後にリブートが必要な場合は、メッセージが表示されます。

VM のインストールおよび初期化後、次のコマンドを実行すると、Volume Manager の主要プロセス (`vxconfigd`、`vxnotify`、および `vxrelocd`) が稼働しているかどうかを確認することができます。

```
# ps -ef | grep vx
```

Volume Manager のデーモン

Volume Manager が正しく動作するには、次の 2 つのデーモンが稼働している必要があります。

- ◆ vxconfigd
- ◆ vxiod

設定デーモン (vxconfigd)

Volume Manager 設定デーモン (vxconfigd) は、Volume Manager ディスクおよびディスク グループの設定を管理します。vxconfigd は、設定の変更をカーネルに伝え、ディスクに格納されている設定情報を変更します。

Volume Manager 設定デーモンの起動

vxconfigd は、ブート処理中に起動スクリプトによって呼び出されます。

このデーモンが有効であることを確認するには、次のコマンドを入力します。

```
# vxdctl mode
```

vxconfigd が既に稼働しており、有効な場合は、次のメッセージが表示されます。

```
mode:enabled
```

vxconfigd が既に稼働しているのが、有効でない場合は、次のメッセージが表示されます。

```
mode:disabled
```

このデーモンを有効にするには、次のコマンドを入力します。

```
# vxdctl enable
```

vxconfigd が稼働していない場合は、次のメッセージが表示されます。

```
mode:not-running
```

vxconfigd を起動するには、次のコマンドを入力します。

```
# vxconfigd
```

vxconfigd は、起動後は自動的にバックグラウンド プロセスになります。

デフォルトでは、vxconfigd はコンソールにエラー メッセージを出力しますが、ログ ファイルに出力するように設定することもできます。

vxconfigd デーモンの詳細については、vxconfigd (1M) および vxdctl (1M) マニュアル ページを参照してください。



ボリューム入出力デーモン (vxiod)

ボリューム拡張入出力デーモン (vxiod) を使用すると、呼び出しプロセスを中断せずに拡張入出力操作を実行できます。

vxiod の詳細については、vxiod (1M) マニュアルページを参照してください。

ボリューム入出力デーモンの起動

vxiod デーモンは、システムブート時に起動します。通常、複数の vxiod デーモンが常に稼働しています。初回インストール後にリブートすると、vxiod が起動します。

vxiod デーモンが稼働していることを確認するには、次のコマンドを入力します。

```
# vxiod
```

vxiod はカーネル スレッドであり、ps コマンドを使用して参照できないため、前述の方法でしか vxiod デーモンが稼働していることを確認できません。

vxiod デーモンが稼働している場合は、次のメッセージが表示されます。

```
10 volume I/O daemons running
```

10 は稼働している vxiod デーモンの数を示します。

vxiod デーモンが 1 つも稼働していない場合は、次のコマンドを入力し、任意の数のデーモンを起動してください。

```
# vxiod set 10
```

10 は、起動する vxiod デーモンの任意の数に変更できます。システムの各 CPU につき 1 つ以上の vxiod デーモンを起動しておくことをお勧めします。

システムの設定

この節では、効率良くストレージ管理を行うためのシステム設定に役立つ情報を紹介します。特定の設定タスクについては、『VERITAS Volume Manager Reference Guide』や『VERITAS Volume Manager Storage Administrator Administrator's Guide』を参照してください。

以下に述べるシステム設定手順は一般的なものなので、例として参照してください。ご使用のシステムによっては必要事項が異なる場合があります。特定の要件に合わせた設定については、後述の「システム設定ガイドライン」を参照してください。

システム設定手順の例

ストレージ管理システムの設定時に実行する一般的な作業手順について以下に説明します。

初期設定

- ◆ Volume Manager の管理下にディスクを組み入れる。
- ◆ 新しいディスク グループを作成する (rootdg を使用しない場合や、別のディスク グループを使用する場合)。
- ◆ ボリュームを作成する。
- ◆ ファイル システムをボリューム上に配置する。

選択オプション

- ◆ boot/root ディスクをカプセル化してミラーリングし、予備のブート ディスクを作成する。
- ◆ ホットリロケーションのスペア ディスクを指定する。
- ◆ ボリュームにミラーを追加する。

保守管理

- ◆ ボリュームおよびファイル システムのサイズを変更する。
- ◆ ディスク / ディスク グループをさらに追加する。
- ◆ スナップショットを作成する。



システム設定ガイドライン

以下の一般的なガイドラインは、効率的なストレージ管理システムを理解し計画する際に役立ちます。各節に記載されている参照先では、それぞれのガイドラインの詳細情報を得ることができます。

ホットリロケーションのガイドライン

以下の一般的なガイドラインに従って、ホットリロケーションを使用できます。詳細については、「ホットリロケーション (51 ページ)」を参照してください。

- ◆ ホットリロケーション機能は、デフォルトで有効になっています。無効にすることもできますが、そのまま有効にしておくことをお勧めします。
- ◆ ホットリロケーション機能を使用する場合は、特にスペア ディスクを指定する必要はありませんが、各ディスク グループ内に少なくとも 1 つ以上のスペア ディスクを指定できます。スペア ディスクを指定すると、リロケート用のディスクについても管理することができます。スペア ディスクを指定しないと、ディスク グループ内の使用可能な空き領域が使用されます。リロケートに空き領域を使用すると、リロケート後のパフォーマンスが低下する場合があります。
- ◆ ホットリロケーションの開始後に、1 つ以上の追加ディスクをスペアとして指定し、スペア領域を増やすことができます (元のスペア領域の一部がリロケートされたサブディスクで使用される場合があります)。
- ◆ ディスク グループが複数のコントローラにわたっており、複数のスペア ディスクがある場合は、スペア ディスクを別々のコントローラ上で設定できます (いずれかのコントローラに障害が発生した場合の対策)。
- ◆ ミラーリングされているボリュームの場合は、ディスク グループに、そのボリュームのミラーが含まれていないディスクが少なくとも 1 つ必要です。このディスクは、使用可能な領域があるスペア ディスクか、または空き領域がある通常のディスクのいずれかである必要があり、ホットリロケーションで使用されます。
- ◆ ミラー アンド ストライプ ボリュームの場合は、ディスク グループに、そのボリュームのいずれかのミラー、またはストライプ ブレックス内の別のサブディスクが含まれていないディスクが少なくとも 1 つ必要です。このディスクは、使用可能な領域があるスペア ディスクか、または空き領域がある通常のディスクのいずれかである必要があり、ホットリロケーションで使用されます。
- ◆ RAID-5 ボリュームの場合、ディスク グループに、そのボリュームの RAID-5 ブレックス (またはそのログ ブレックス) が含まれていないディスクが少なくとも 1 つ必要です。このディスクは、使用可能な領域があるスペア ディスクか、または空き領域がある通常のディスクのいずれかである必要があり、ホットリロケーションで使用されます。
- ◆ ミラーリングされているボリュームのデータプレックスの一部に DRL ログ サブディスクが含まれる場合、そのプレックスをリロケートすることはできません。ログ サブディスクは、データが含まれていないプレックス (ログ ブレックス) 内に配置することができます。

- ◆ ホットリロケーションでは、元のパフォーマンス特性やデータレイアウトが必ず維持されるとは限りません。リロケート後のサブディスクの新しい位置を確認し、より適切なディスクに配置し直して、元のパフォーマンス特性を回復する必要があるかどうかを判断します。
- ◆ ミラーリングされているルート ディスクに障害が発生した場合、ホットリロケーションでは、ルート ディスクのミラーを新しく作成できます。ルート ディスクの新しいミラーを作成するためには、rootdg ディスク グループに、ルート ディスク上のボリュームを格納するために十分な連続したスペア領域か空き領域が必要です (rootvol および swapvol では連続したディスク領域が必要)。
- ◆ VxVMオブジェクトをスペア ディスク上に作成することもできますが (vxmake1または Storage Administrator インタフェースを使用)、用意したスペア ディスクはホットリロケーション専用として使用することをお勧めします。

ストライピングのガイドライン

以下の一般的なガイドラインに従って、ストライピングを使用できます。詳細については、「ストライピング (RAID-0) (14 ページ)」を参照してください。

- ◆ 同じ物理ディスク上に、ストライプ ブレックスのカラムを複数配置しないでください。
- ◆ ストライプ ユニット サイズは慎重に計算してください。通常は、中規模のストライプ ユニット サイズ (vxassist で使用するデフォルトサイズの 64 キロバイトなど) を使用することをお勧めします。ストライプ ユニット サイズをトラックサイズに合わせて設定できず、アプリケーションの入出力パターンが不明な場合、ストライプ ユニット サイズは 64 キロバイトにしてください。

注 最近のディスク ドライブは「可変ジオメトリ」であることが多く、トラック サイズがシリンダ間で異なります (つまり外側のトラックの方が内側のトラックより多くのセクタを含みます)。そのため、トラック サイズをストライプ ユニット サイズとして使用することが常に適切であるとは限りません。このようなドライブで、アプリケーションの入出力パターンが不明な場合は、中規模のストライプ ユニット サイズ (64 キロバイトなど) を使用してください。

- ◆ ストライプ ユニット サイズが小さいボリュームは、スピンドルが同期していない場合に、順次入出力の処理時間が長くなる場合があります。通常、スピンドルが同期していない複数のディスクにわたってストライピングを実行する場合は、ストライプ ユニット サイズが大きくマルチスレッドであるか、または非同期のランダム入出力ストリームである方が効率よく実行できます。
- ◆ 通常、ストライプを構成する物理ディスク数が多いほど入出力のパフォーマンスは向上しますが、ボリュームの障害発生頻度は高くなります。この点が問題となる場合は、ストライピングとミラーリングを併用することによって、ボリュームのパフォーマンスおよび信頼性を向上させることができます。



- ◆ ミラーリングされているボリュームの1つのプレックスだけがストライピングされている場合は、ストライププレックスに対して、ボリュームのポリシーを `prefer` に設定します (デフォルトの読み取りポリシー `select` では、これは自動的に実行されます)。
- ◆ ミラーリングされているボリュームの複数のプレックスがストライピングされている場合は、各ストライププレックスのストライプユニットサイズを等しくしてください。
- ◆ 可能な場合は、ストライピングボリュームのサブディスクを、別のコントローラやバスに接続されているドライブにわたって分散してください。
- ◆ 重複シークをサポートしていないコントローラは使用しないでください (ほとんどのコントローラではサポートされています)。

`vxassist` コマンドは、ボリューム内のストライププレックスに領域を割り当てる際に、上述の規則の多くを自動的に適用し遵守します。

ミラーリングのガイドライン

以下の一般的なガイドラインに従って、ミラーリングを使用できます。詳細については、「ミラーリング (RAID-1) (18 ページ)」を参照してください。

- ◆ ミラーボリュームの異なるプレックスのサブディスクを、同じ物理ディスク上に配置しないでください。同じディスク上に配置すると、ミラーリングの利点である可用性が損なわれ、パフォーマンスが低下する危険があります。このような問題を避けるには、`vxassist` を使用してください。
- ◆ ミラーリングを使用して最適なパフォーマンスを実現するには、物理入出力処理の 70 パーセント以上が読み取りである必要があります。読み取りの割合が高いほど、パフォーマンスが高くなります。書き込みの方が多い場合は、ミラーリングするとパフォーマンスが向上しないか、結果的に低下する可能性があります。

注 UNIX オペレーティングシステムは、ファイルシステム キャッシュを搭載しており、多くの場合、読み取り要求にはこのキャッシュで対応できます。このため、ファイルシステムを介した物理入出力処理の読み取り / 書き込み比率は、書き込みの方が高くなる傾向があります (アプリケーションレベルの読み取り / 書き込み比率と比較した場合)。

- ◆ ミラーリングまたはストライピングする場合は、可能な限り、異なるコントローラに接続されているディスクを使用してください。多くのディスクコントローラでは、2つのディスクで同時にシークを開始する重複シークがサポートされています。重複シークをサポートしていないコントローラに接続されているディスク上には、同じボリュームの2つのプレックスを設定しないでください。これは、ドライブ上にキャッシュを持たない古い種類のコントローラや SCSI ディスクの場合に特に重要です。最近のワークステーションやサーバマシンで使用されている新しい SCSI ディスクやコントローラでは、ほとんどの場合問題ありません。複数のコントローラにわたってミラーリングすると、1つのコントローラに障害が発生してもシステムの稼働を続行で

きるという利点があります。障害が発生していないコントローラは、ほかのミラーのデータを提供し続けることができます。

- ◆ 複数のディスクにわたってストライピングまたはコンカチネーションを実行したり、より高速なデバイス上に配置すると、プレックスのパフォーマンスが向上する場合があります。読み取りポリシーは、その「高速」プレックスを優先して設定することができます。デフォルトでは、ストライププレックスからなるボリュームでは、ストライププレックスの読み取りを優先にして設定されます。

ダーティ リージョン ロギング (DRL) のガイドライン

以下の一般的なガイドラインに従って、ダーティ リージョン ロギングを使用できます。詳細については、「ダーティ リージョン ロギング (55 ページ)」を参照してください。

ダーティ リージョン ロギング (DRL) では、システム障害発生後のミラー ボリュームのリカバリを高速化できます。DRL が有効になっている場合、Volume Manager はプレックスへの書き込みによって変更されたボリューム内の領域を監視します。Volume Manager はビットマップを維持し、この情報をログ サブディスクに保存します。ログ サブディスクはボリュームに対して定義および追加され、DRL を実現します。ログ サブディスクはプレックスとは独立しており、プレックスのポリシーは適用されず、DRL 情報を保持する目的でのみ使用されます。

注 ダーティ リージョン ロギングを使用すると、書き込み処理では、システムのパフォーマンスに影響のある場合があります。

DRL を使用するには、以下のガイドラインに従ってください。

- ◆ DRL の効果を発揮させるには、ボリュームをミラーリングしておく必要があります。
- ◆ DRL が機能するためには、ボリューム上に1つ以上のログ サブディスクが必要です。ただし、各プレックスに含むことのできるログ サブディスクは1つです。
- ◆ ログ サブディスクとして使用するサブディスクは、必要なデータを含まないようにする必要があります。
- ◆ ボリュームに複数のログ サブディスク (プレックスごとに1つ) を含めることによって、ログ サブディスクを「ミラーリング」することができます。このようにしておくと、ディスクに障害が発生して1つのログ サブディスクがアクセス不能になった場合でも、ログの記録を続行できます。
- ◆ ログ サブディスクは複数のセクタで構成する必要があります (セクタ数が奇数の場合は、ログ サブディスク中の最後のセクタが使用されないため、できる限り偶数にしてください)。ログ サブディスクのサイズは通常、ボリューム サイズに比例します。ボリューム サイズが2 ギガバイト未満の場合は、ログ サブディスクのセクタは2つで十分です。ボリューム サイズが2 ギガバイト増えるごとに、サブディスク サイズも2 セクタずつ増やす必要があります。ただし、vxassist ではデフォルトで適切なサイズが選択されるため、通常は、このデフォルトのログ サブディスク長を使用してください。



- ◆ ログサブディスクは、可能な限り、使用頻度の低いディスク上に配置するようにしてください。
- ◆ ログサブディスクには、持続性のある（非揮発性）ストレージディスクを使用する必要があります。

ミラーリング アンド ストライピングのガイドライン

以下の一般的なガイドラインに従って、ミラーリング アンド ストライピングを使用できます。詳細については、「ミラーリング アンド ストライピング (RAID-1 + RAID-0) (18 ページ)」を参照してください。

- ◆ ストライピング編成およびミラーリング編成を取る場合、使用可能なディスクを十分に確保します。ストライププレックスには2つ以上のディスク、ミラーリングにはそれとは別に1つ以上のディスクが必要です。
- ◆ あるプレックスのサブディスクを、別のプレックスのサブディスクと同じ物理ディスク上に配置しないでください。「ストライピングのガイドライン (37 ページ)」の具体的な指示に従ってください。
- ◆ 「ミラーリングのガイドライン (38 ページ)」の具体的な指示に従ってください。

ストライピング アンド ミラーリングのガイドライン

以下の一般的なガイドラインに従って、ストライピングとミラーリングを併用できます。詳細については、「ストライピング アンド ミラーリング (RAID-0 + RAID-1) (19 ページ)」を参照してください。

- ◆ ストライピング編成およびミラーリング編成を取る場合、使用可能なディスクを十分に確保します。ストライププレックスには2つ以上のディスク、ミラーリングにはそれとは別に1つ以上のディスクが必要です。
- ◆ あるプレックスのサブディスクを、別のプレックスのサブディスクと同じ物理ディスク上に配置しないでください。「ストライピングのガイドライン (37 ページ)」の具体的な指示に従ってください。
- ◆ 「ミラーリングのガイドライン (38 ページ)」の具体的な指示に従ってください。

RAID-5 のガイドライン

以下の一般的なガイドラインに従って、RAID-5 を使用できます。詳細については「RAID-5 (17 ページ)」を参照してください。

通常、RAID-5 についても、ミラーリングおよびストライピングの両方のガイドラインを適用できます。RAID-5 では、さらに以下のガイドラインも考慮する必要があります。

- ◆ 1つのRAID-5 ボリュームには、1つのRAID-5 プレックスのみを含めることができます (ただしログ プレックスは複数含めることも可能)。
- ◆ RAID-5 プレックスは、複数の物理ディスク上の2つ以上のサブディスクから構成する必要があります。ログ プレックスがある場合、それらのログ プレックスは RAID-5 プレックスで使用されているディスクとは別のディスクに属している必要があります。
- ◆ RAID-5 ログはミラーリングおよびストライピングすることができます。
- ◆ ボリューム長が明示的に指定されていない場合で、ボリュームに関連付けられている RAID-5 プレックスがあるときは、その長さに設定されます。それ以外の場合は、ゼロに設定されます。ボリューム長が明示的に指定されている場合で、関連付けられている RAID-5 プレックスがあるときは、そのストライプ ユニット サイズの倍数にする必要があります。
- ◆ ログ長が明示的に指定されていない場合で、関連付けられている RAID-5 ログ プレックスがあるときは、そのうちの最も小さい値に設定されます。関連付けられた RAID-5 ログ プレックスがない場合は、ゼロに設定されます。
- ◆ 空の RAID-5 ログ プレックスは無効です。



システムの保護

ディスクに障害が発生すると、障害を起こしたディスク上のデータが失われ、システムにアクセスできなくなるという、2種類の問題が発生する可能性があります。システムにアクセスできなくなった原因が、主ディスク（システムの運用に使用されるディスク）の障害であることがあります。VERITAS Volume Manager を使用すると、これらの問題からシステムを保護することができます。

システムの可用性を維持するには、システムの稼働や起動に必要なデータをミラーリングしておく必要があります。また、ミラーリングしたデータは、障害発生時に備えて保持しておきます。

システムおよびデータを保護するための手順については、以下の説明を参考にしてください。

- ◆ ルートファイルシステムを含むディスク（root または boot ディスク）をカプセル化して Volume Manager の管理下に組み入れます。この処理によって、root デバイスおよび swap デバイスがボリュームに変換されます（rootvol および swapvol）。次にルートディスクをミラーリングして、ブート専用の予備ルートディスクを作成します。ブートするのに必要なディスクをミラーリングしておくこと、1つのディスクに障害が発生しても、システムがブートできなくなったり、使用できなくなったりすることはありません。

システムの可用性を最大限に高めるには、rootvol、swapvol、usr、および var の各ボリュームについてミラーを作成します。詳細については、『VERITAS Volume Manager Administrator's Guide』の「リカバリ」に関する章を参照してください。

- ◆ ミラーリングによってデータを保護します。データをミラーリングすることにより、ディスク障害発生時のデータの損失を避けることができます。そのために、2つ以上のデータプレックスを含むミラーボリュームを作成し使用します。これらのプレックスは別々のディスク上にある必要があります。ディスク障害の発生によっていずれかのプレックスが使用できなくなった場合でも、ほかのディスク上のミラーリングされているボリュームのデータを使用できます。

vxassist mirror を使用してミラーを作成する場合、どのディスクが損傷した場合でもデータが失われないようにミラーが配置されます。デフォルトでは、vxassist を実行してもミラーボリュームは作成されません。/etc/default/vxassist ファイルを編集して、デフォルトのレイアウトをミラーリングするように設定できます。

- ◆ Volume Manager のホットリロケーション機能を有効にしておくと、障害の検出、障害内容の通知、障害による影響を受けた冗長性のあるサブディスクのリロケートの試行、およびリカバリ手順の開始などが自動的に行われます。1つのディスクグループにつき1つ以上のホットリロケーション用のスペアディスクを用意して、障害発生時のリロケートに必要な十分な領域を確保します。

root ディスクをミラーリングすると、元の root ディスクに障害が発生した場合に、ホットリロケーションによって root ディスクのミラーがもう1つ自動的に作成されます。rootdg ディスクグループには、ルート ディスク上のボリュームを格納するための連続したスペア領域か空き領域が必要です (rootvol ボリュームおよび swapvol ボリュームでは、連続したディスク領域が必要)。

- ◆ ミラーリングされたボリュームに関して、ダーティ リージョン ログ機能を活用して、システム障害後のミラー ボリュームのリカバリ処理を高速化できます。ミラー ボリュームに1つ以上のログ サブディスクがあることを確認します (rootvol ボリューム、swapvol ボリューム、および usr ボリュームは、DRL ボリュームにすることができません)。
- ◆ RAID-5 ボリュームの場合、ログを記録することによって、リカバリ データの破損を避けることができます。各 RAID-5 ボリュームに1つ以上のログ プレックス確保してください。

定期的にバックアップを実行して、データを保護します。バックアップは、ボリュームのすべてのコピーが何らかの理由で失われるか壊れた場合に必要になります。たとえば、電圧の急増によって、システム上のいくつかの (またはすべての) ディスクが破損する場合があります。また、コマンドを間違えて使用して、個別に重要なファイルを削除したり、ファイルシステムを破壊したりしてしまう場合も考えられます。



はじめに

この章では、VERITAS Volume Manager の機能について詳しく説明します。

この章では、次のトピックについて説明します。

- ◆ オンライン再レイアウト
- ◆ ホットリロケーション
- ◆ ボリュームの再同期化
- ◆ ダーティ リージョン ロギング
- ◆ 高速ミラー再同期 (FMR)
- ◆ Volume Manager のルート機能
- ◆ 動的マルチパス (DMP)
- ◆ VxSmartSync Recovery Accelerator
- ◆ Volume Manager タスク モニタ
- ◆ Volume Manager のクラスタ機能



オンライン再レイアウト

オンライン再レイアウトを行うと、データ アクセスを中断せずに、Volume Manager でサポートされているストレージレイアウトを別のレイアウトに変換することができます。通常、Volume Manager でストレージレイアウトを変更するのは、ストレージの冗長性やパフォーマンスなどの特性を変更するためです。Volume Manager は、アドレス空間を複製（ミラーリング）するか、パリティを追加（RAID-5）することによって、ストレージの冗長性を高めます。Volume Manager でストレージのパフォーマンス特性を変更するには、カラム数やストライプ幅などのストライピングに関するパラメータを変更します。

レイアウトの変更には次の種類があります。

- ◆ RAID-5 からミラーリング
- ◆ ミラーリングから RAID-5
- ◆ パリティの追加または削除
- ◆ カラムの追加または削除
- ◆ ストライプ幅の変更

ストレージ レイアウト

現在のリリースでは、オンライン再レイアウトは、次のストレージレイアウトをサポートします。

- ◆ コンカチネイテッド
- ◆ ストライピング
- ◆ RAID-5
- ◆ ミラーリング
- ◆ ストライプ ミラー
- ◆ コンカチネイテッド ミラー

注 VERITAS Volume Manager Storage Administrator では、ストライププロがストライプミラーの GUI 用語として使用されています。また、コンカチネイテッドプロはコンカチネイテッド ミラーの GUI 用語です。

オンライン再レイアウトの使用方法

VERITAS オンライン再レイアウト機能を使用すると、データアクセスを中断せずに、既に設定されているストレージレイアウトを変更することができます。必要に応じて、特定のレイアウトのパフォーマンス特性を変更できます。コマンドを1つ実行するだけで、レイアウトを変更できます。

ストライプ ミラー プレックスは、ミラーリングされているボリュームにあるストライプ プレックスを重ね合わせたものです。このプレックスはミラーリングとストライピングの両方を備えています。この組み合わせによって形成されるプレックスは、ストライプ ミラー プレックスと呼ばれます。コンカチネイテッド プレックスも同じようにミラーリングされます。オンライン再レイアウトでは、ストライプ ミラー プレックスからコンカチネイテッド ミラー プレックスへの変更と、コンカチネイテッド ミラー プレックスからストライプ ミラー プレックスへの変更の両方をサポートしています。

注 現在のリリースでは、レイアウト変更時にミラー数を変更することはできません。

たとえば、ストライプユニットサイズが128キロバイトのストライプレイアウトがあり、そのパフォーマンスが十分でないとは想定します。この場合、再レイアウト機能を使用して、レイアウトのストライプユニットサイズを変更することができます。

ボリュームにマウントされたファイルシステムは、オンラインでの縮小および拡大処理に対応していれば、レイアウト変更時にマウント解除する必要はありません。VFS にはこの機能があります。

オンライン再レイアウトでは、既存のストレージ領域を再利用して、新しいレイアウトでの必要に対処するための割り当てポリシーを採用しています。レイアウト変更処理では、最小限の一時領域を使用して、指定されたボリュームをデスティネーションレイアウトに変換します。

レイアウト変更は、ソースレイアウト内のデータの一部をデスティネーションレイアウトに移動することによって実行されます。データはソース ボリュームから一時領域へコピーされます。データはソース ボリューム ストレージ領域からは削除されます。その後、ソース ボリューム ストレージ領域が新しいレイアウトに変更され、一時領域にコピーされていたデータが新しいレイアウトに書き込まれます。この処理は、ソース ボリューム内のすべてのストレージとデータが新しいレイアウトに変更されるまで、繰り返し実行されます。

オンライン再レイアウトを使用すると、カラム数やストライプ幅の変更、パリティの削除または追加、および RAID-5 のミラーリングへの変更を実行することができます。



レイアウト変更の種類

オンライン再レイアウト処理を有効に実行するには、次の条件の1つ以上を満たす必要があります。次の1つ以上の処理を実行します。

- ◆ RAID-5 からミラーリングへの変更
- ◆ ミラーリングから RAID-5 への変更
- ◆ カラム数の変更
- ◆ ストライプ幅の変更
- ◆ パリティの削除または追加

レイアウト変更を実行するには、ミラー ボリュームのプレックスのレイアウトにおいてストライプ幅とカラム数がすべて同じである必要があります。表1「サポートされているレイアウト変更」を参照してください。

表 1. サポートされているレイアウト変更

変更元 / 変更先	ストライプ ミラー	コンカチネイ テッド ミラー	通常の ミラー	RAID-5	コンカチネイ テッド	ストライプ
ストライプ ミラー	可 1	可 2	不可 3	可 4	可 5	可 6
コンカチネイ テッド ミラー	可 7	不可 8	不可 9	可 10	不可 11	可 12
通常の ミラー	可 13	可 14	不可 15	可 16	不可 17	不可 18
RAID-5	可 4	可 10	不可 19	可 20	可 21	可 22
コンカチネイ テッド	可 5	不可 11	不可 17	可 21	不可 23	可 24
ストライプ	可 6	可 12	不可 18	可 22	可 24	可 25

表 1 のエント리는以下の通りです。

- ◆ 「可」は、オンライン再レイアウト処理を実行できることを示します。
- ◆ 「不可」は、処理を実行できる場合もありますが、再レイアウトは実行できないことを示します。
- ◆ 数値は、特定のレイアウト変更の変更可能な処理についての簡単な説明に対する参照番号を示します。「各数値の説明 (49 ページ)」を参照してください。
- ◆ 処理は、双方向で実行可能です。

各数値の説明

表 1 内の数値は、次のレイアウト処理を示します。

1. ストライブ幅またはカラム数を変更します。
2. すべてのカラムを削除します。
3. 再レイアウト処理でなく、変換処理を行います。
4. ミラーリングから RAID-5 への変更とストライブ幅/カラムの変更のいずれかまたは両方を行います。
5. ミラーリングから RAID-5 への変更とストライブ幅/カラムの変更のいずれかまたは両方を行います。
6. ストライブ幅/カラムを変更し、ミラーを削除します。
7. カラムを追加します。
8. 再レイアウト処理ではありません。
9. 変換処理です。
10. ミラーリングから RAID-5 へ変更します。vxconvert 処理を参照してください。
11. ミラーを削除します。再レイアウト処理ではありません。
12. ミラーを削除してストライピングを追加します。
13. ミラー ボリュームをストライブ ミラーに変更します。これは、カラムまたはストライブ幅に変更があった場合にのみ再レイアウト処理となり、それ以外の場合は変換処理になります。vxconvert 処理を参照してください。
14. 古いミラー ボリュームをコンカチネイテッド ミラーに変更します。これは、カラムまたはストライブ幅に変更があった場合にのみ再レイアウト処理となり、それ以外の場合は変換処理になります。
15. 変更はありません。再レイアウト処理ではありません。
16. 古いミラー ボリュームを RAID-5 に変更します。再レイアウトに使用するプレックスを古いミラー ボリュームから選択します。ほかのプレックスは、再レイアウト処理の最後に削除されます。
17. ミラー ボリュームのプレックスを選択し、カラム/ストライブ幅を変更した場合のみ、これは再レイアウト処理になります。
18. ミラー ボリュームのプレックスを選択し、カラム/ストライブ幅を変更した場合のみ、これは再レイアウト処理になります。
19. 再レイアウト処理ではありません。
20. ストライブ幅/カラムを変更します。
21. パリティとすべてのカラムを削除します。



- 22. パリティ
- 23. 変更はありません。再レイアウト処理ではありません。
- 24. カラムを削除します。
- 25. ストライブ幅 / カラム数を変更します。

レイアウト変更の特性

データのレイアウトを変更すると、既存レイアウトを取るデータが新しいレイアウトに再編成されます。レイアウトの変更中、オンライン再レイアウトでは、使用されている一時領域をミラーリングすることによって、データの冗長性を維持します。データの読み取り / 書き込みアクセスは、レイアウトの変更中でも中断されません。

レイアウトの変更中にシステムに障害が発生した場合でも、データが損傷を受けることはありません。レイアウト変更処理は、システムのリストア後に続行され、読み取り / 書き込みアクセスも維持されます。

レイアウト変更処理は、いつでも元に戻すことができます。ただし、データが以前のストレージ位置に正確に戻されない場合があります。元に戻す前に、ボリューム内で実施中のレイアウト変更を中止する必要があります。

`vxrelayout status` コマンドを使用すると、レイアウト変更の方向を指定できます。

これらの変更では、データを移動するために十分な冗長性を維持していれば、入出力処理の障害の影響を受けずに済みます。

レイアウト変更およびボリュームの大きさ

レイアウト変更によっては、実行するとボリュームの大きさが変わることがあります。レイアウト変更でボリュームの大きさが変更された場合、オンライン再レイアウトでは、`vxresize` を使用してファイルシステムを縮小または拡大します。

オンライン再レイアウトを実行しても、空白プレックスは変更されません。また、プレックスを空白にすることもありません。

注 オンライン再レイアウトは、`vxassist` コマンドまたは **Storage Administrator** で作成したボリュームにのみ使用することができます。

次のレイアウト変更はサポートされていません。

- ◆ ログプレックスの変更
- ◆ ボリューム上でオンライン再レイアウト処理が実行されているときのボリュームのスナップショット

ホットリロケーション

ホットリロケーション機能により、冗長化（ミラーリングまたは RAID5）されている Volume Manager オブジェクトに入出力障害が発生してもシステムが自動的に対処してこれらのオブジェクトへのアクセスと冗長性をリストアすることができます。Volume Manager は、オブジェクトでの入出力障害を検出し、影響のあったサブディスクをリロケートします。サブディスクは、スペア ディスクやディスク グループ内の空き領域として指定されているディスクにリロケートされます。その後、Volume Manager は障害直前のオブジェクトを復元して、再び冗長性を持たせアクセスを可能にします。

ディスクの部分的な障害（ディスク上の一部のサブディスクにのみ影響を与える障害）が発生した場合は、ディスクの該当部分の冗長性のあるデータがリロケートされます。ディスク上で影響のなかった部分にある既存ボリュームは、引き続きアクセスすることができます。

注 ホットリロケーションは、障害が発生したディスク上の冗長性のある（ミラーリングまたは RAID-5）サブディスクに関してのみ実行されます。障害が発生したディスク上の冗長性のないサブディスクはリロケートされませんが、システム管理者には障害発生が通知されます。

ホットリロケーションの仕組み

ホットリロケーション機能は、デフォルトで有効になっています。障害発生時にホットリロケーションを起動するために、システム管理者が実行する処理はありません。

ホットリロケーション デーモン `vxrelocd` は、Volume Manager で冗長性に影響するイベントを監視し、ホットリロケーションを実行して冗長性をリストアします。また、`vxrelocd` によって、システム管理者に障害の発生、リロケート操作、およびリカバリ操作が通知されます。`vxrelocd` の詳細については、`vxrelocd (1M)` マニュアル ページを参照してください。

`vxrelocd` デーモンは、システムの起動中に開始され、Volume Manager のもとでディスク、プレックス、または RAID-5 サブディスクに障害が発生しないかどうかを監視します。障害が発生すると、`vxrelocd` デーモンはホットリロケーションの起動を試みます。

ホットリロケーションが正常に実行されると、次のような処理が実行されます。

1. ディスク、プレックス、または RAID-5 サブディスクの障害による Volume Manager イベントが検出されます。
2. システム管理者（およびその他の指定されたユーザ）に、障害の発生が通知され、影響のあった Volume Manager オブジェクトが特定されます。この通知は、電子メールで送信されます。
3. リロケート可能なサブディスクの判定、ディスク グループ内のそれらのサブディスク用の空き領域の検出、およびサブディスクのリロケートを実行します。これらの操作とその成否をシステム管理者に通知します。



4. ボリュームおよびデータのリストアに必要なリカバリ処理を開始します。リカバリ処理の結果をシステム管理者に通知します。

注 ホットリロケーションでは、実行前のデータのレイアウトやパフォーマンスが実行後も同じであるとは限りません。システム管理者は、ホットリロケーションの実行後に、設定の変更を行う必要があります。

リロケート用領域の選択方法

スペア ディスクは交換処理で使用する前に、初期化し、スペアとしてディスク グループに配置しておく必要があります。障害発生時に、スペアとしてディスクが指定されていない場合、**Volume Manager** は自動的に障害が発生したディスク グループ内の使用可能な空き領域を使用します。スペア ディスク領域が不足している場合は、スペア領域と空き領域を組み合わせ使用します。

ホットリロケーションで指定した空き領域は、ホットリロケーションで常に利用できます。ホットリロケーションの利用対象からディスクを除外するには、**Storage Administrator** インタフェース (`vxdiskadm` または `vxedit`) を使用します。

システム管理者は、各ディスク グループ内で 1 つ以上のディスクをホットリロケーション スペアとして指定できます。**Storage Administrator** インタフェース (`vxdiskadm` または `vxedit`) 使用すると、ディスクをスペアとして指定できます。スペアとして指定されたディスクは空き領域モデルに配置することはできません。また、ストレージ領域を割り当てることもできません。

リロケート用の領域を選択する際、ホットリロケーションではリロケートされるサブディスクが属する **Volume Manager** オブジェクトの冗長性特性を保持します。たとえば、ホットリロケーションでは、障害を起こしたプレックスのサブディスクが障害を起こしたプレックスのミラーを含むディスクにリロケートされないようにします。使用可能なスペアディスクや空き領域を使用して冗長性が保持されない場合は、ホットリロケーションは実行されません。リロケーションが実行できない場合はシステム管理者に通知され、それ以上の処理は実行されません。

ホットリロケーションが使用するディスクには、使用可能なディスクの中から障害が発生したディスクに「最も近いもの」が選択されます。近さは、障害が発生したディスクのコントローラ、ターゲット、およびディスク番号によって決まります。障害が発生したディスクと同じコントローラ上にあるディスクは、別のコントローラ上のディスクよりも近くにあります。また、障害が発生したディスクと同じ ターゲット上にあるディスクは、別のターゲット上のものよりも近くにあります。

可能な場合、ホットリロケーションは、障害の発生しているドライブから同一のデスティネーションディスクにすべてのサブディスクを移動します。

ルート ディスクの障害時は、同一のディスク上にすべてのファイル システムをリロケートできる場合に限り、ホットリロケーションは動作します。そのようなディスクが見つからなかった場合は、システム管理者に電子メールで通知されます。

ホットリロケーションが行われると、障害を起こしたサブディスクは構成データベースから削除されます。Volume Manager では障害を起こしたサブディスクが使用していたディスク領域が空き領域として再利用されないようにします。

ホットリロケーションを無効にする方法については、『VERITAS Volume Manager Installation Guide』を参照してください。

リロケート処理の解除ユーティリティ

VxVM のホットリロケーションを使用すると、システムは冗長性のある VxVM オブジェクト上の入出力障害にサブディスクレベルで自動対応し、オブジェクトを使用可能な状態にするための処理を実行します。この機構によって、サブディスク上の入出力障害の検出、サブディスクのリロケート、サブディスクに関連付けられたプレックスのリカバリなどが行われます。ディスクの交換後は、ディスク障害発生前と同じ設定にシステムをリストアするための `vxunreloc` ユーティリティを使用できます。`vxunreloc` は、ホットリロケーションで作成されたサブディスクを、ディスク障害によって交換されたディスクに戻します。

ボリュームの再同期化

ミラー ボリュームまたは RAID-5 ボリュームを使用して、冗長性を保持しながらデータを保存する場合、Volume Manager では、データのすべてのコピーを正確に一致させることができます。ただし、一定の条件下（通常は完全なシステム障害が原因）では、ボリューム上のいくつかの冗長データの整合性が崩れたり、同期しないことがあります。ミラーデータは、元のデータとまったく同一ではありません。このような状態は、通常の設定変更（プレックスの切り離しや再組み込みなど）を除くと、データのボリュームへの書き込み中にシステムに障害が発生した場合にのみ発生します。

データおよびパリティが RAID-5 ボリュームに書き込まれる場合と同様に、データはボリュームのミラーに並行して書き込まれます。個々の書き込みがすべて完了する前にシステムに障害が発生した場合は、一部の書き込みが失敗し、一部の書き込みは完了している可能性があります。このような状態によって、データが同期しなくなることがあります。ミラー ボリュームでは、ボリュームの同じ領域から 2 つの読み取りが行われ、読み取り要求を満たすために 2 つのミラーが使われた場合は、異なる結果が返されることがあります。RAID ボリュームの場合は、パリティの破損や不正なデータの復元の原因となることがあります。

Volume Manager では、すべてのミラーにまったく同じデータが含まれること、および RAID-5 ボリュームのデータとパリティが整合することを確保する必要があります。この処理はボリュームの再同期化と呼ばれます。ブート時に自動的にインポートされたディスクグループの一部であるボリューム (`rootdg` など) の場合、システムのリポート時に再同期化処理が実行されます。



すべてのボリュームがシステム障害発生後に再同期化を必要とするわけではありません。システムの障害発生時に一度も書き込みされていないボリューム、または休止中の（アクティブな入出力がない）ボリュームは、書き込み途中であった可能性はなく、再同期化は必要ありません。

Volume Manager では、ボリュームに最初に書き込まれたときに記録し、ボリュームにダーティフラグを設定します。ボリュームがすべての処理によって閉じられるか、管理者によってきちんと停止されたときに、すべての書き込みが完了し、ボリュームからダーティフラグが削除されます。システムのレポート時にダーティフラグが立っているボリュームのみ、再同期化を実行する必要があります。

再同期化処理は、ボリュームの種類によって異なります。**RAID-5** ログを含む **RAID-5** ボリュームの場合は、それらのログを「再生」することができます。使用可能なログがない場合は、ボリュームは復元 - リカバリ モードになり、すべてのパリティが再生成されます。ミラー ボリュームの場合は、ボリュームをリカバリ モード（読み込み - 書き戻しリカバリ モード）にすることによって、再同期化が実行されます。ボリューム内のデータの再同期化は、バックグラウンドで処理されます。これにより、リカバリ処理中にボリュームの使用が可能となります。

再同期化処理は、コストがかかり、システムのパフォーマンスに影響を与えるものです。リカバリ処理は、リカバリを分散することで特定のディスクやコントローラへの負荷の集中を防ぎ、再同期化に付随する影響をいくらか低減します。

サイズの大きいボリュームや数の多いボリュームの場合は、再同期化処理に時間がかかることがあります。これらの影響に対しては、ミラー ボリュームではダーティリージョンロギングを適用し、**RAID-5** ボリュームでは有効な **RAID-5** ログがあることを確認することによって対応できます。データベースアプリケーションで使用されるボリュームの場合は、**VxSmartSync™ Recovery Accelerator** を使用することができます（「**VxSmartSync Recovery Accelerator**（67 ページ）」を参照）。

ダーティ リージョン ロギング

ダーティ リージョン ロギング (DRL) は、任意のボリューム属性で、システム障害発生後にミラー ボリュームを迅速にリカバリするために使用されます。DRL では、入出力のミラー ボリュームへの書き込みによって変更されたリージョンの記録を取ります。DRL はこの情報を使用して、ボリュームのうちのリカバリが必要な部分のみをリカバリします。

DRL を使用していないときに、システムに障害が発生した場合は、ボリュームのすべてのミラーを整合性のとれた状態にリストアする必要があります。リストア処理では、ボリュームのすべての内容をミラー間でコピーします。この処理は、入出力が多く時間がかかります。また、既に整合性のとれたボリュームの領域もリカバリする必要があります。

DRL は、ボリュームを連続した一連のリージョンに論理的に分割し、書き込みのあったボリュームのリージョンを記録します。ダーティ リージョン ロギングには、ボリュームの各リージョンの状態を表わすビットが必ず含まれます。ボリュームへの書き込み操作では、データの書き込み前に、ログ内の該当リージョンにダーティ フラグが設定されます。書き込みによって、以前はクリーンだったログ内のリージョンがダーティになる場合は、書き込み操作の開始前に、ログが同期的にディスクに書き込まれます。システムの再起動時に、Volume Manager はダーティ リージョン ロギングにダーティ フラグが設定されているボリュームのリージョンのみをリカバリします。

ログ サブディスクは、DRL が有効になっているボリュームのダーティ リージョン ロギングを保存するために使用されます。DRL が有効なボリュームには、1 つ以上のログ サブディスクがあり、複数のログ サブディスクを使用して、ダーティ リージョン ロギングをミラーリングすることもできます。各ログ サブディスクは、ボリュームの 1 つのプレックスと関連付けられます。1 つのプレックスにつき 1 つのログ サブディスクのみ含むことができます。ログ サブディスクのみを含み、データ サブディスクを含まないプレックスは、ログ プレックスと呼ばれます。

ログ サブディスクは、データ サブディスクを含む通常のプレックスに関連付けることもできます。この場合、ログ サブディスクは、データ サブディスクの 1 つが破損してプレックスが切り離された場合に、使用できなくなることがあります。

vxassist コマンドを使用してダーティ リージョン ロギングを作成すると、デフォルトでは、1 つのログ サブディスクを含むログ プレックスが作成されます。ダーティ リージョン ロギングは、ログ サブディスクを作成し、プレックスと関連付けることによって、手動で作成することもできます。このようにして、プレックスにログ サブディスクとデータ サブディスクを含めることができます。

ログ内では、限られた数のビットのみにダーティ フラグが常に設定されます。リージョンに設定されたダーティ ビットは、そのリージョンへのデータの書き込みが完了しても、すぐには解除されません。対応するボリューム リージョンが使用されたリージョンの中で最も古いものになるまで、ダーティ フラグは設定されたままです。既にダーティ フラグが設定されているリージョンに対して別の書き込みが発生した場合は、書き込み操作の開始前にログをディスクに書き込む必要はありません。



Oracle のログの再適用に使用されるボリュームなどでは、順次書き込みが行われるため、このような時間のかかる DRL ビットのクリーニングは効果がありません。これらのボリュームには、シーケンシャル DRL を使用して、ダーティビット数を制限し、リカバリ時間を短縮することができます。シーケンシャル DRL で処理できるダーティビット数を制限するには、チューニング可能パラメータ `voldr1_max_dirty` を使用します。順次書き込みのボリューム上でシーケンシャル DRL を使用すると、入出力のスループットに重大な影響を与える可能性があります。

注 DRL では、ほとんどの書き込みアクセスパターンにおいて、多少の入出力オーバーヘッドが追加されます。

高速ミラー再同期 (FMR)

高速ミラー再同期 (FMR) 機能は、バックアップや意志決定支援などの支援処理に対する VxVM のスナップショット機構の効果を高めることにより、陳腐化したミラーに対して高速で効果的な再同期化処理を実行します。通常、これらの処理には、データストアボリュームが静止状態であること、またはデータストアへのセカンダリアクセスがプライマリアクセスに影響を与えていないことが必要となります (スループット、更新、整合性など)。

これを実現するために、VxVM にはスナップショット機構があります。スナップショットは、ある時点のプライマリ ボリュームの正確なコピーを作成する機能です。スナップショットを作成すると、元のボリュームから独立してスナップショットにアクセスできるようになります。共有またはクラスタの VxVM 環境では、別のマシンからスナップショットにアクセスすることによって、リソースの競合やオーバーヘッドを排除することができます。

高速ミラー再同期では、以前のバージョンのスナップショット機構にあった問題が解決されました。以前のバージョンには以下のような問題がありました。

- ◆ スナップショットの作成後は、プライマリ ボリュームとスナップショットが切り離されます。これらの整合性は確保されません。その結果、スナップショットの内容が使用できなくなった時点で、そのスナップショットを破棄し、新たにスナップショットを作成して最新のプライマリ データのコピーを得る必要があります。
- ◆ スナップショットの作成に時間がかかるため、スナップショット機能の実用性に制限があります。これは、スナップショットの作成時間がボリューム サイズに正比例しているためです。企業などの大規模なボリュームでは、この作成時間の影響を受け、オフラインでの使用を余儀なくされる場合があります。

FMR コンポーネント

FMR は、VxVM に対して 2 つの基本的な拡張機能を提供します。1 つは、ミラー再同期化処理の最適化 (高速ミラー再同期) です。もう 1 つは、スナップショットを破棄せずに更新して再使用する機能を提供するためのスナップショット モデルの拡張 (高速ミラー再接続) です。

高速ミラー再同期コンポーネント

FMR では、データストアの更新を適用する際にミラーが使用できなかったために失われた更新を記録しておきます。ミラーが使用できるようになったら、ミラーによって失われた更新だけを再度適用します。FMR を使用すると、復元中のミラーにデータストア全体をコピーする現在の方法に比べ、リストアするデータ量が大幅に少なくなります。

使用できないミラーとは、ボリュームから切り離されたミラーのことです。切り離しは、エラーのために VxVM によって自動的に、または管理者によって直接 (VxVM ユーティリティ `vxplex` または `vxassist` など) 実行されます。復元中のミラーとは、一度切り離されてから元のボリュームに再度組み込むために、`vxrecover` または `vxplex att` で処理中のミラーのことです。

FMR では、現在のミラーに関する障害は修正されていません。また、管理モデルは修復されていません。ユーザが確認できる効果は、一般的なミラーの修復処理が高速になったことだけです。

拡張された再同期化機能では、管理者はボリュームの単位で FMR の有効 / 無効を設定できます。また、FMR の状態も確認できます。

高速ミラー再接続コンポーネント

高速ミラー再接続により、既存のスナップショット使用モデルが増えました。高速ミラー再同期化を使用しないと、ボリュームの独立したコピーはスナップショット機構を通じて作成されます。元のボリュームとレプリカボリュームは、互いに完全に独立しており、これらのデータは内容が異なる可能性があります。

高速ミラー再接続のスナップショット拡張機能を使用すると、スナップショットボリュームを元のボリュームに再度関連付けることができます。これは、VxVM のスナップショット機能に深く依存する周期的処理による作業負荷を低減することが目的です。

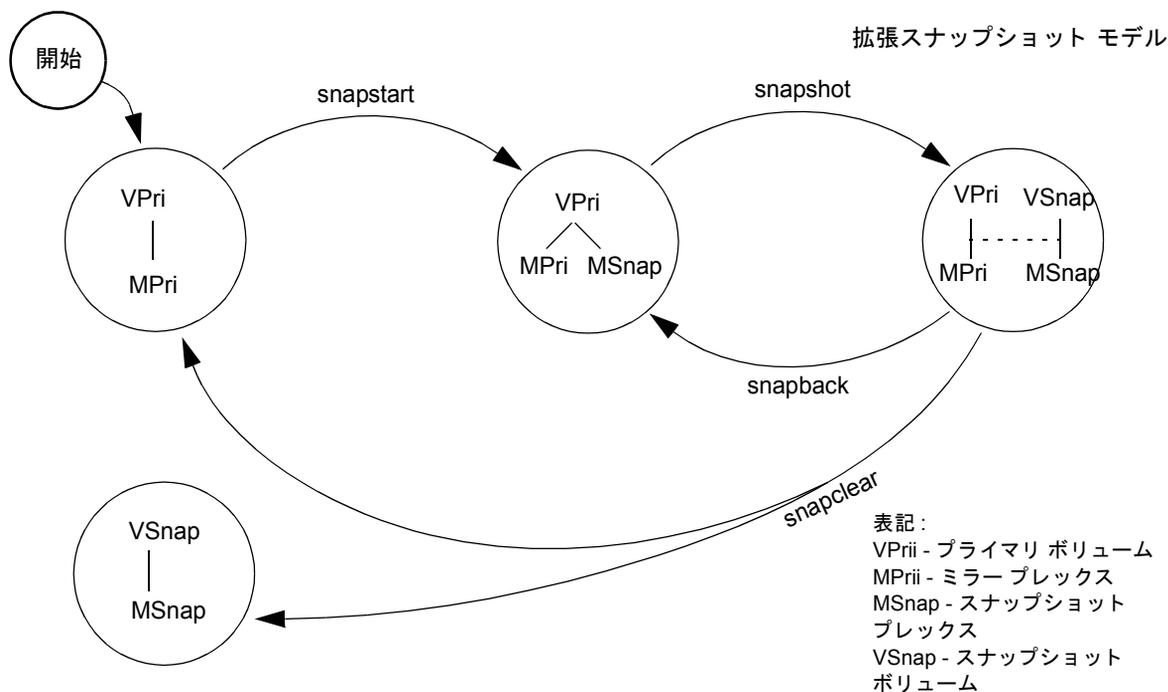


VxVM のスナップショット機能に対する FMR の拡張機能

FMR スナップショットのリリース 3.1 に対する拡張機能では、58 ページの図 23 「FMR 拡張スナップショット」に示すようにスナップショット モデルが拡張されています。リリース 3.1 では、`snapshot` コマンドは、元のボリュームとスナップショット ボリュームを関連付ける点を除けば、以前と同じ動作をします。新しい `vxassist snapback` コマンドは、この関連付けを使用して、スナップショット プレックス (M_{Snap}) をスナップされたボリューム (V_{Pri}) に戻します。

図 23 「FMR 拡張スナップショット」は、`snapback` コマンドおよび `snapclear` コマンドの導入により拡張されたスナップショット モデルの移行を示します。

図 23. FMR 拡張スナップショット



さらに、新しい `vxassist snapclear` コマンドは、`snapshot` コマンドが作成した関連付けを完全に削除します。これにより、スナップショットの記録によるボリュームの管理上のオーバーヘッドを軽減します。この機能は、スナップショットを作成元のボリュームに戻すことがない環境では有用です。

動作原理

FMR の基本は変更の記録です。ミラーがオフラインまたは切り離されている時、およびミラーのスナップショットが作成されている時には、更新は記録されません。ミラーを戻すときは、これらの更新だけを適用することによって、ボリュームの再同期化にかかる時間を大幅に削減できます。

この変更記録の基礎となるのは、ビットマップの利用です。ビットマップ内の各ビットは、ボリュームのアドレス空間の連続した領域（範囲）を表します。この連続した領域を領域サイズと呼びます。通常、この領域サイズは1ブロックです。ボリュームの各ブロックは、ビットマップ内で1ビットとして示されます。ただし、チューニング可能パラメータ `vol_fmr_logsz` によって、FMR マップの最大サイズ（ブロック単位）を制限することができます。マップサイズの計算アルゴリズムは、領域サイズが1の場合から開始されます。計算されたマップサイズが `vol_fmr_logsz` よりも小さい場合は、その値がマップサイズとなります。 `vol_fmr_logsz` よりも大きくなった場合は、 `vol_fmr_logsz` に適応するように領域サイズを2にして計算されます（以降マップサイズが `vol_fmr_logsz` よりも小さくなるまで領域サイズの値を増加して計算が繰り返されます）。

たとえば、次のように入力します。

```
ボリューム サイズ=1G
vol_fmr_logsz=4
```

ここでは、ブロックサイズが512バイトのシステムで、 $4 \times 512 = 2048$ バイト、すなわち $2048 \times 8 = 16384$ ビットとなります。

したがって、1Gのボリュームでは、領域サイズ1個が2048ビットになります。これは4ブロックよりも小さく、その結果マップサイズは2048ビット、すなわち256バイトとなります。

ボリュームサイズが増加した場合は、マップサイズが `vol_fmr_logsz` を超えないように再計算する必要があることに注意してください。

持続性のあるトラッキングと持続性のないトラッキング

VxVM 3.1 では、FMR マップはメモリに割り当てられます。DRL とは異なり、これらはディスクや固定ストアには配置されません。システムに障害が発生すると、この情報が失われるため、ボリューム全体を同期させる必要があります。この方法の利点は、ディスクに対する更新がないため、FMR の更新（このマップへの更新）がパフォーマンスに影響しないことです。ただし、システムに障害が発生した場合は、情報が失われるため、ミラーの完全な再同期化が再度必要となります。



スナップショットと FMR

スナップショット使用時に FMR の差分トラッキングを有効に活用するには、新しいスナップショット オプションを使用します。スナップショットの作成後、スナップショット オプションを使用してスナップショット プレックスを再度組み込みます。スナップショットを作成する前に FMR が有効になっていて、スナップショットの作成が完了する前に FMR が無効になった場合は、FMR ビットマップに反映される FMR の差分が、スナップバック中にボリュームの再同期化に使用されます。複数ボリュームのスナップショットを同時に簡単に作成するには、複数ボリュームに対応できるようにスナップショット オプションを拡張し、命名体系を追加しておきます。デフォルトでは、各レプリカ ボリュームは SNAP-**<元のボリューム>** という名前になります。デフォルト値は、コマンドラインでオプションを設定すると変更できます。複数ボリュームのスナップショットを同時に作成できるのは、ボリュームが同一ディスク グループ内にある場合だけです。

1 つのディスク グループ内の全ボリュームのスナップショットを簡単に作成するために、vxassist に新しいオプションが追加されていますが、ディスク グループ内に完全なスナップショット プレックスを持たないボリュームがあるとエラーになります。同一ボリュームで複数のスナップショットを作成することが可能です。新しい FMR ビットマップが各スナップショットごとに作成されるため、各スナップショットの再同期化にかかる時間は最小限に抑えられます。

スナップバックの実行時に、スナップショット プレックスを優先データ セットとして選択できます。スナップバック オプションに `-o resyncfromreplica` を追加すると、スナップショット (レプリカ) プレックス上のデータが元のボリュームに接続されているすべてのミラー上にコピーされます。デフォルトでは、元のボリューム上のデータが優先設定になっていて、スナップショット プレックス上にコピーされます。

レプリカ ボリュームまたは元のボリュームを拡大して、FMR の使用を続けることが可能です。ボリュームを拡大すると、FMR が差分を記録するために使用するビットマップが拡張されます。これにより、ビットマップのサイズまたは領域サイズが変更されることがあります。いずれの場合も、ビットマップのうち、ボリュームの拡大された領域に対応する部分には、「ダーティ」フラグが設定され、再同期化されます。スナップショットが不完全なスナップショット プレックスを作成しようとする、スナップショット処理に障害が発生します。このような場合は、スナップバック オプションの実行前に、レプリカ ボリュームまたは元のボリュームを拡大する必要があります。2 つのボリュームを別々に拡大すると、ボリューム内のほかのミラーと物理ディスクを共有するスナップショットが作成されることがあります。これを防止するには、`snapback` コマンドが完了してからボリュームの拡大を行ってください。

レプリカ ボリュームのレイアウトを変更する処理では、スナップショットの FMR マップに「ダーティ」フラグが設定されるため、スナップバック中に完全な再同期化が必要となります。この処理には、サブディスクの分割、サブディスクの移動、レプリカのオンライン再レイアウトなどがあります。スナップショットが完了してから、これらの処理を実行すると安全です。詳細については、`vxvol (1M)`、`vxassist (1M)`、および `vxplex (1M)` の各マニュアル ページを参照してください。

FMRと書き込み可能なスナップショット

書き込み可能なスナップショットへの変更を記録するには、次に示す2つのオプションのいずれかを使用します。

- ◆ スナップショット ボリュームへの変更を記録する個別のマップを作成する。
- ◆ スナップショット ボリュームの親のマップを更新する。バックアップおよび DSS (意志決定支援システム) アプリケーションのように、スナップショット ボリュームの更新がほとんどない場合は、この簡略方式を使用してください。

VxVM 3.1 では、後者の方式が実装されています。すなわち、スナップショットへの書き込みがあると、スナップショット ボリュームの親のマップが更新されます。

警告および制限事項

FMR は、RAID-5 ボリュームではサポートされていません。

サブディスクがリロケートされると、プレックス全体に「ダーティ」フラグが設定されるため、完全な再同期化が必要となります。

Volume Manager のルート機能

Volume Manager では、種々のシステムのさまざまなファイル (ルート ファイル システム、swap デバイス、usr ファイル システム、stand ファイル システムなど) を Volume Manager の管理下に組み入れることができます。これをルート機能と呼びます。ルート ディスク (ルート ファイル システムを含むディスク) は、カプセル化処理を実行して Volume Manager の管理下に置くことができます。

カプセル化とは、ディスク上の既存のパーティションをボリュームに変換することです。root および swap デバイスは、Volume Manager の管理下に置かれると、ボリュームとして見えるようになり、ほかの Volume Manager ボリュームと同じ特性を持つようになります。スワップ領域として使用するよう設定されたボリュームはスワップ ボリューム、ルート ファイル システムを含むボリュームはルート ボリューム、スタンド ファイル システムを含むボリュームはスタンド ボリュームとそれぞれ呼ばれます。

rootvol ボリューム、swapvol ボリューム、および standvol ボリュームは、システムを正常にブートするために必要なルート ディスクのほかの部分 (/usr など) と同じように、ミラーリングすることができます。これにより、ディスクの障害発生したときに備えた、完全な冗長性およびリカバリ可能性が得られます。Volume Manager のルート機能を使用しないと、root、swap、_usr ルート、またはスタンド パーティションが損傷を受けた場合に、生き残ったディスクからシステムをリブートすることができなくなります。



ブートに不可欠なディスクドライブをミラーリングしておく、たった 1 つのディスクに障害が発生しただけでシステムが使用できなくなるという事態を防ぐことができます。重要なディスクを別の使用可能なディスクにミラーリングするように設定してください (vxdiskadm コマンドを使用)。root、stand、および swap パーティションを含むディスクに障害が発生した場合、ルートのミラーを含むディスクからシステムをリブートできます。ブート (ルート) ディスクのミラーリングおよびシステム リカバリ処理については、第 7 章「リカバリ」を参照してください。

ルート ボリュームを使用したブート

通常、オペレーティング システムをブートすると、root ファイル システム、stand ファイル システム、および swap 領域がブート処理の初期の段階で使用できるようになります。その後、ユーザ処理を実行して Volume Manager の設定を読み込み、ボリュームを起動させることができます。root、stand、および swap の各デバイスの設定は、Volume Manager を起動する前に完了している必要があります。Volume Manager vxconfigd デモンを init 処理の一部として開始した後では、ボリュームを root または swap デバイスとして使用するように設定することができなくなります。

この制限を回避するために、rootvol ボリューム、standvol ボリューム、および swapvol ボリュームのミラーは、起動時にシステムからアクセスされます。起動時には、システムは rootvol、standvol、および swapvol の各ボリュームを通常のパーティションとして認識し、標準的なパーティション番号設定方式を使用してアクセスします。rootvol、standvol、および swapvol ボリュームはそれぞれ 1 つのパーティションで対応付けられた連続するディスク領域から作成されます。この制限により、rootvol、standvol、または swapvol ボリュームのプライマリ プレックス (ブートに使用されるプレックス) をストライピングまたはスパン化することはできません。ブートに必要なこれらのボリュームのミラーは、どれもストライピングまたはスパン化できません。

ブート時のボリュームの制限

rootvol、standvol、swapvol、および usr ボリュームは、ボリュームの設定に関して非常に特殊な制限がある点で、ほかのボリュームと異なります。

- ◆ ルート ボリューム (rootvol) は、デフォルトのディスク グループ rootdg に含まれている必要があります。rootdg 以外のディスク グループに rootvol という名前のボリュームを作成することはできますが、システムのブート時に使用されるボリュームは rootdg 内の rootvol のみです。
- ◆ rootvol ボリュームは、特定のマイナー デバイス番号 (マイナー デバイス 0) を持ちます。また、swapvol はマイナー デバイス番号 1 を持ちます。usr ボリュームは特定のマイナー デバイス番号を持ちません。第 7 章「リカバリ」を参照してください。

- ◆ rootvol、var、usrrootvol、standvol、および swapvol デバイスの制限されたミラーには、これらのデバイス用に作成された「オーバーレイ」パーティションがあります。オーバーレイパーティションとは、制限付きのミラーで使用するディスク領域と同じ大きさのパーティションのことです。ブート時は、rootvol、var、usrrootvol、standvol、および swapvol ボリュームが完全に設定される前に、デフォルトのボリューム設定により、オーバーレイパーティションを使用してディスク上のデータにアクセスします。
- ◆ パフォーマンスを向上させるために、ストライプ ミラーを rootvol デバイスに追加することはできますが、プライマリプレックスや、rootvol のミラー（プライマリプレックスが破損した場合に、システムのリカバリやブートに必要な場合がある）をストライピングすることはできません。
- ◆ rootvol、standvol、および swapvol はスパン化できません。つまり、プライマリプレックスを複数の連続しないサブディスクに分散して配置することはできません。カプセル化されたブート ディスクに関連付けられたボリューム（rootvol、usr、var、opt、swapvol など）は、ディスク上の基礎となる物理パーティションにマッピングされていて、連続している必要があるため、拡大または縮小できません。対処としては、ブート ディスクのカプセル化を解除してから、必要に応じてブート ディスクを再パーティション化し（パーティションの拡大または縮小）、再びカプセル化を行う方法があります。
- ◆ ブート ディスクの一部をミラーリングする場合、ミラーリングする元のディスクに元のプレックス上のデータを保存するための十分な領域がないと、ミラーリングできないことがあります。
- ◆ rootvol、standvol、swapvol、および usr は、ダーティ リージョン ロギング ボリュームにできません。

これらの要件に加えて、root、usr、var、opt、varadm、usrkvm、および swap の各ボリューム用に1つ以上の連続する（必要に応じて、シリンダ整列の）ミラーを作成することをお勧めします。これによって、これらのボリュームから通常のディスクパーティションへの変換が容易になります（オペレーティングシステムの更新時など）。



動的マルチパス (DMP)

一部のシステムでは、Volume Managerはマルチポートに対応したディスク アレイをサポートします。そのような Volume Manager はディスク アレイ内の特定ディスク デバイスへの複数の入出力パスを自動認識します。Volume Manager の動的マルチパス機能は、パスフェイルオーバー機構を提供することによって、さらに高い信頼性を実現します。ディスクへの1つの接続が失われた場合、システムはそのディスクへの別の正常な接続経路を使用して、重要なデータへのアクセスを続行します。DMP は、入出力による負荷をディスク デバイスへの複数の入出力パスに均等に分散することによって、入出力スループットも大幅に向上させます。

Volume Manager では、システムに接続されているすべての物理ディスクは、1つ以上の物理アクセス パスを持つメタデバイスとして表されます。システムに接続されている単一の物理ディスクは、1つのパスを持つメタデバイスとして表されます。ディスク アレイに含まれるディスクは、2つの物理アクセス パスを持つメタデバイスとして表されます。Volume Manager 管理ユーティリティ (vxdisk など) を使用すると、メタデバイスのすべてのパスと、さまざまなパスの状態を参照できます。

オペレーティング システムの複数のデバイス ハンドルを通じて物理ディスクにアクセス可能な場合、マルチパス条件が存在しえます。それらのデバイス ハンドルを使用すると、ホストとデバイス間で、別のパスを介して、データにアクセスしたり、データを制御したりすることが可能です。

Volume Manager に DMP システムを組み込んで、マルチパス デバイスにアクセスできるようことができます。Volume Manager は、ユニバーサル ワールドワイドデバイス識別子 (WWD ID) を使用して、マルチパス デバイスを検出します。DMP でマルチパス デバイスにアクセスするには、物理ディスクに、明確な WWD ID 識別子を付ける必要があります。

WWD ID によって DMP で物理ディスクを識別できない場合、識別処理は Volume Manager のデバイス検出機能に委ねられます。デバイスの検出は、Volume Manager でディスク上のメタデータ識別子を認識できるかどうかには依存します。

Volume Manager DMP は、検出した各マルチパス ターゲットのメタデバイスを示すメタノードを作成します。各メタノードは、オペレーティング システムの一連のデバイス ハンドルに対応付けられ、適切なマルチパス ポリシーが設定されます。Volume Manager DMP は、オペレーティング システムのデバイス ハンドルを使用してアクセス可能な、接続されているすべての物理ディスクに関するメタノードを作成します。

Volume Manager DMP は、ディスク アレイのような複数のパスの使い方に関するポリシーを定義しているマルチパス ターゲットを管理します。一部のディスク アレイでは、複数のパスを同時にアクティブにすることができます (アクティブ/アクティブ)。1つのパスだけをアクティブにし、別のパスは既存のパスに障害が発生した場合のスペアとして設定できるディスク アレイもあります (アクティブ/パッシブ)。さらに、もっと精巧なポリシーを備えたディスク アレイもあります。

一般に、Volume Manager は、VM ディスクが 1 つの Volume Manager DMP メタノードにマッピングされるように設計されています。VxVM 論理操作を単純化するために、各 VM ディスクは一意的 Volume Manager DMP メタノードに対応付けられます。これは、物理ディスク デバイスがマルチパス構成を取っているかどうかにかかわらず行われます。

パス フェイルオーバー機構

DMP をマルチポート対応のディスク アレイに適用すると、システムの信頼性が向上します。ディスク アレイへの 1 つの接続が失われた場合、DMP は入出力要求に関して次の入出力パスを動的に選択します。この処理は自動的に実行されるため、管理者による操作は不要です。

DMP を使用すると、管理者は Volume Manager の DMP サブシステムに、失われた接続が修理またはリストアされたことを通知できます。この機能を DMP 再構成と呼びます。この再構成手順では、新しく追加されたデバイスや、システムが完全に起動した後で削除されたデバイスも検出することができます (オペレーティング システムにそれらのデバイスが正しく認識されている場合のみ)。

負荷の平準化

複数のパスにわたって負荷を平準化するため、DMP ではアクティブ / アクティブ状態のディスク アレイの場合はパス負荷平準化機構を使用します。負荷を平準化すると、すべてのパスの全帯域幅を最大限に活用することによって、入出力スループットを確実に向上させることができます。ただし、ディスクへの順次入出力は、入出力スループットを最適化するために、同じパスに送出されます。これは、ディスクトラック キャッシュを効果的に活用するためです。

アクティブ / パッシブ タイプのディスク アレイの場合、入出力は障害が発生するまでプライマリ パスに送られます。プライマリ パスに障害が発生すると、入出力は、使用可能なほかのプライマリ パスまたはセカンダリ パスに引き継がれます。コントローラ間で LUN の所有権の移転が連続して行われると、入出力の速度が大幅に低下するため、アクティブ / パッシブ ディスク アレイでは、複数のパスにわたる負荷の平準化は行われません。

DMP デバイスからのブート

ルート ディスクが Volume Manager の管理下に置かれている場合、そのディスクは、1 つのディスクである場合は 1 つのパス、ディスク アレイに含まれている場合は複数のパスを持つ DMP デバイスとして、自動的に認識されます。ルート ディスクをカプセル化することにより、ディスクへの 1 つ以上の既存の物理パスが失われた場合のシステムの信頼性が向上します。



コントローラの有効化と無効化

DMP 機能を使用すると、管理者は管理処理を実行するために、ホスト入出力コントローラへの入出力を無効にすることができます。この機能は、ホストに接続されているコントローラや、Volume Manager でサポートされているディスク アレイを管理する場合に使用します。ホスト入出力コントローラへの入出力処理は、管理処理の完了後に有効にすることができます。これには、Volume Manager で提供される `vxdmpadm` コマンドを使用します。

たとえば、システムに StorEdge A5000[™] アレイがあり、このディスク アレイに接続されている A5000 インタフェース ボードをユーザが変更する必要がある場合は、`vxdmpadm` コマンドを使用してこの A5000 インタフェース ボードに接続されているホスト入出力コントローラの一覧を取得します。一覧のすべてのコントローラを無効にする必要があります。この処理を行うと、それ以降はこれらのコントローラを通じて実行されるディスクへの入出力操作は停止されます。

次に、ディスク アレイ中のディスクに対して行われている入出力操作を中断せずに、このインタフェース ボードを入れ替えます。通常、この操作はアクティブ / アクティブ タイプのディスク アレイ (この例も同様) で必要な操作です。Volume Manager はパス負荷平準化機構を使用して、複数のパスを持つディスクへの入出力のスケジュールを設定します。その結果、いかなる時点でどのパスにも入出力処理が流れる可能性があるからです。

アクティブ / パッシブ タイプのディスク アレイの場合、Volume Manager は、障害が発生するまで入出力操作をプライマリ パスに送るように設定します。したがって、ディスク アレイ上のインタフェース カード、またはディスク アレイに接続されているホスト上のカード (可能な場合) を変更するには、ホスト コントローラへの入出力操作を無効にする必要があります。この処理によって、すべての入出力操作をアクティブなセカンダリパスまたは別の入出力コントローラ上のアクティブなプライマリ パスに引き継いでからハードウェアを変更することができます。

操作の完了後、無効にしたコントローラを使用するパスは、`vxdmpadm` コマンドを実行して有効にすることにより、元どおり使用できるようになります。

Volume Manager では、最後に残っているルート ディスクのアクティブ パスを無効にすることはできません。

DMP データベース情報の表示

`vxdmpadm` コマンドを使用して、DMP データベース情報を一覧表示し、ほかの管理処理を実行することができます。このコマンドを使用すると、ユーザはディスクに接続されているシステム上のすべてのコントローラと DMP データベースに格納されているほかの関連情報の一覧を参照することができます。この情報を活用して、システム ハードウェアの所在を把握したり、有効 / 無効にするコントローラを決定することができます。

また、`vxdmpadm` コマンドを使用すると、ディスク アレイのシリアル番号、ディスク アレイ コンピュータ システムに接続されている DMP デバイス (ディスク) の一覧、特定のコントローラへのパスの一覧など、ほかにも役立つ情報を参照できます。

VxSmartSync Recovery Accelerator

いくつかのシステムでは、VxSmartSync_Recovery Accelerator を使用できます。

VxSmartSync for Mirrored Oracle Databases は、ボリュームの再同期化処理 (*resilvering*) を高速化する機能を集めたもので、Oracle Universal Database と共に使用されます。これらの機能では、Volume Manager ボリュームとデータベース ソフトウェア間の拡張インターフェースを使用するため、ミラー再同期化の不必要な作業を省くことができます。これらの拡張機能により、ボリュームのリカバリ時間を大幅に短縮できます。

Oracle は、SmartSync が有効になっている場合は自動的に SmartSync を利用します。

システム管理者はボリュームを設定して、VxSmartSync を正しく使用する必要があります。Volume Manager では、データベースで使用できるボリュームは次の 2 種類になります。

- ◆ REDO ログ ボリューム。データベースの REDO ログが含まれます。
- ◆ データ ボリューム。データベースで使用するほかのすべてのボリューム (制御ファイルおよびテーブル領域ファイル)。

VxSmartSync はこの 2 種類のボリュームで異なった動作をします。また、拡張インターフェースを最大限に活用するために、これらのボリュームを正しく設定する必要があります。2 つのボリュームの違いは、REDO ログ ボリュームにはダーティ リージョン ロギングが含まれ、データ ボリュームには含まれないことだけです。

データ ボリュームの設定

データ ボリュームのリカバリ時間の短縮は、リカバリが必要なボリューム部分をデータベース ソフトウェアに判断させることで達成されます。データベースは、データに加えられた変更のログをデータベースに保持しており、ボリュームのどの部分をリカバリすべきか判断できます。リカバリが必要な領域を小さくし、データベースでリカバリ処理を制御できるようにすると、リカバリ時間全体を短縮できます。

また、リカバリ処理はシステム起動時ではなく、データベース ソフトウェアの起動時に行われます。これにより、システムのレポート時にリカバリの影響全体を抑えることができます。リカバリ処理はデータベースにより制御されるため、ボリュームのリカバリ時間はデータベースの再同期 (*resilvering*) 時間 (REDO ログの再生に必要な時間) になります。

データベースが独自のログを保持するため、Volume Manager でログを記録する必要はありません。そのため、データ ボリュームは、ダーティ リージョン ロギングを使用しないミラー ボリュームとして設定されます。リカバリ時間の短縮に加えて、DRL による実行時の入出力オーバーヘッドも避けられるようになり、通常のデータベースへの書き込みアクセスが速くなります。



REDO ログ ボリュームの設定

REDO ログは、データベースのデータに加えられた変更のログです。REDO ログへの変更のログはデータベースに保持されないため、データベース自体は再同期 (resilvering) が必要な部分についての情報を提供できません。また REDO ログは順次書き込みされます。従来のダーティ リージョン ロギングはデータがランダムに書き込まれた場合にも有効なため、REDO ログのリカバリ時間を短縮するには、ダーティ リージョン ロギングはほんの少ししか役に立ちません。ただし、Volume Manager ではダーティ リージョン ロギング機能の動作を変更することによって、ダーティ リージョン数を減らし、順次アクセスパターンを有効活用することができます。これにより、リカバリに必要なデータ量を減らし、リカバリ時間によるシステムへの影響を抑えることができます。

REDO ログの拡張インタフェースを使うと、ボリュームを REDO ログとして使用すべき時間にデータベース ソフトウェアから Volume Manager に通知できます。これにより、Volume Manager はボリュームの DRL 動作を変更し、アクセスパターンを有効に活用できます。リカバリ時間の短縮はダーティ リージョン ロギングに依存するため、REDO ログ ボリュームはダーティ リージョン ロギングを使用するミラー ボリュームとして設定する必要があります。

Volume Manager タスク モニタ

Volume Manager タスク モニタは、処理の作成、管理、および完了を監視することにより、システムのリカバリの進行状況を把握します。タスク モニタにより、処理の進行状況を監視し、一時停止やリカバリ率などの処理特性を変更できます (システムのパフォーマンスへの影響を抑える場合など)。オンライン再レイアウト機能の進行状況も監視および変更できます。詳細については「オンライン再レイアウトを参照してください」。

Volume Manager のクラスタ機能

Volume Manager には、クラスタ環境において VxVM を使用するためのオプションのクラスタ機能があります。Volume Manager のクラスタ機能を利用すると、Volume Manager の管理下にある一連のディスク (VM ディスク) に、複数のホストから同時にアクセスして管理することができます。

クラスタは、一連のディスクを共有するホストの集まりです。各ホストは、クラスタ内ではノードと呼ばれます。ノードはネットワークを通じて接続されます。1 台のノードが故障しても、ほかのノードは引き続きディスクにアクセスすることができます。Volume Manager のクラスタ機能では、すべてのノードにおいて、ディスクの構成の設定 (変更を含む) を同一の論理ビューとして見ることができます。このクラスタ機能を有効にすると、クラスタ内のすべてのノードによって、Volume Manager オブジェクトを共有することができます。

Volume Manager のクラスタ機能の詳細については、本書のクラスタ機能に関する章を参照してください。

注 Volume Manager のクラスタ機能は別途ライセンス提供されます。

はじめに

この章では、Volume Manager を使用したディスクの管理操作について説明します。ディスク グループの操作についても説明します。

注 ほとんどの Volume Manager コマンドを実行する場合に、スーパーユーザまたはその他の適切な権限が必要です。

この章では、次のトピックについて説明します。

- ◆ 標準ディスク デバイス
- ◆ ディスク グループ
- ◆ ディスク コマンドおよびディスク グループ コマンド
- ◆ ディスクの初期化と追加
- ◆ Volume Manager へのディスクの追加
- ◆ ディスク グループへのディスクの追加
- ◆ ディスク グループからのディスクの削除
- ◆ ディスクの移動
- ◆ ディスク名の変更
- ◆ ディスクの予約
- ◆ ディスクのオフライン
- ◆ ディスクのミラーリング
- ◆ ディスクの削除
- ◆ ディスク情報の表示
- ◆ 障害のあるディスクの検出と交換
- ◆ ディスク グループの作成
- ◆ ディスク グループのアップグレード



- ◆ ディスク グループの削除
- ◆ システム間でのディスク グループの移動
- ◆ ディスク グループの破棄
- ◆ 特殊デバイスの使用
- ◆ vxdiskadm メニュー インタフェース処理

標準ディスク デバイス

Volume Manager では、標準デバイスと特殊デバイスの 2 種類のディスク デバイスを使用できます。特殊デバイスについては、この章の後半で説明します。

Volume Manager は、物理ディスク上で最大 8 パーティション (スライス) をサポートします。これらのパーティションには、順に 0 から 7 までの名前が付けられます。パーティション 2 は、ディスク全体を指定するために予約されています。

注 システムによっては、Volume Manager で最大 16 パーティションがサポートされます。これらのシステムでは、パーティションには 0 から 15 までの名前が付けられ、パーティション 0 がディスク全体を指定するために予約されています。

パーティションが Volume Manager の管理下に置かれると、VM ディスクがそのパーティションに割り当てられます。VM ディスクには、disk0 などのわかりやすい名前 (ディスク名またはディスク メディア名) を使用できます。

注 システムによっては、例とは異なるデバイス名を使用している場合があります。

パーティションは、c#b#t#d#s# という形式の物理アドレス (通常、デバイス名またはディスク アクセス名) を使用してアドレス付けされます。物理アドレスは次の要素からなります。

- ◆ c# - ディスク ドライブが接続されているコントローラの番号
- ◆ b# - 対応するバス番号 (システムで使用している場合)
- ◆ t# および d# - コントローラ上のディスク ドライブのアドレスを構成するターゲット ID およびデバイス番号
- ◆ s# - ディスク ドライブ上のパーティション番号

たとえば、デバイス名は c0t0d0s2 のようになります。表記上、s2 は Volume Manager で使用される標準パーティション方式を示します。システムによっては、Volume Manager で s0 が標準パーティション方式として使用されます。物理ディスクは、Volume Manager では c#b#t#d#s# (b# はバスを使用するシステム用) として識別されます。

表示コマンドを実行するとデバイスに `s#` が付いた形で報告されますが、ほとんどのコマンドではサフィックス `s#` はオプションです。Volume Manager の `vxdiskadm` および `vxdiskadd` コーティリティでは、サフィックス `s2` (または `s0`) を付けずにデバイス名を使用します。たとえば、1番目のコントローラに接続された2番目のディスクを `vxdiskadd` に指定するには、`c0t1d0` という名前を使用します。

ブートディスク (ルート ファイルシステムを含み、システムのブート時に使用) は、通常、Volume Manager ではデバイス名 `c0t0d0` で識別されます。

Volume Manager ディスクには、次の2つの領域があります。

- ◆ 専有領域 – 設定情報が格納されている小さい領域。ディスク ラベルと設定レコードが格納されています。
- ◆ 共有領域 – ディスクの残りの領域。サブディスクの格納 (および記憶域の割り当て) に使用します。

Volume Manager では、次の3つの基本的なディスク タイプが使用されます。

- ◆ `sliced` – 共有領域および専有領域が異なるディスク パーティション上にあります。
- ◆ `simple` – 共有領域および専有領域が同じディスク パーティション上にあります (専有領域の次に共有領域が続く)。
- ◆ `nopriv` – 占有領域がありません (サブディスクの割り当てに使用する共有領域のみ)。

Volume Manager は、使用パーティション数をできるだけ少なくして新しいディスクを初期化します (通常は物理ディスクごとに2パーティション)。 `s2` (または `s0`) で終わっているディスク アクセス名では、デフォルトのディスク タイプは `sliced` です。



ディスクグループ

ディスクは、**Volume Manager** によってディスクグループへ編成されます。ディスクグループは、共通の構成を共有するディスクの集まりに名前を付けたものです。ボリュームはディスクグループ内に作成され、そのディスクグループ内のディスクのみを使用するように制限されます。

Volume Manager がインストールされているシステムには、デフォルトのディスクグループ `rootdg` が存在します。デフォルトでは、操作は `rootdg` ディスクグループに対して実行されます。システム管理者は、必要に応じて別のディスクグループを作成できます。多くのシステムでは、多数のディスクがない限り、複数のディスクグループを使用することはありません。**Volume Manager** オブジェクトを作成するためにディスクが必要になるまでは、ディスクグループにディスクは追加されません。後からディスクを初期化および予約して、ディスクグループへ追加できます。ただし、**Volume Manager** をインストールするには、1つ以上のディスク(パーティション)を `rootdg` に追加する必要があります。

ディスクグループに追加したディスクには名前が付けられます (`disk02` など)。この名前は、ボリュームを操作 (ボリュームの作成やミラーリング) する際にディスクを識別するために使用します。この名前は、物理ディスクに直接関連付けられます。物理ディスクを別のターゲット アドレスまたは別のコントローラに移動した場合でも、`disk02` という名前で引き続きこの物理ディスクを識別できます。ディスクを交換するには、まず交換するディスクの名前を別の物理ディスクと関連付け、次に元のディスクに格納されているボリューム データを (ミラーやバックアップ コピーから) リカバリします。

ディスクグループのサイズが大きくなると、専有領域がいっぱいになる場合があります。大きなディスクグループでは、ディスクの設定でログインする専有領域を大きく指定する必要があります。専有領域の大部分は、ディスクグループ設定データベース (このディスクグループの各 **Volume Manager** オブジェクトのレコードを含む) が使用します。各設定レコードで 256 バイト (半ブロック) が使用されるため、ディスクグループ内に作成できるレコード数は、設定データベースのコピーサイズの 2 倍になります。コピーサイズは、コマンド `vx dg list` ディスクグループ名を出力して確認できます。

ディスク コマンドおよびディスク グループ コマンド

Volume Manager には、ディスク管理用に次のインタフェースが用意されています。

- ◆ グラフィカル ユーザ インタフェース
- ◆ 一連のコマンド ラインのコマンド
- ◆ vxdiskadm メニュー方式のインタフェース
- ◆ リロケート処理の解除コマンド

この章では、次のコマンドについて説明します。

- ◆ vxdiskadm - Volume Manager Support Operations メニュー インタフェース。このコマンドにより、ディスク操作のメニューが表示されます。メイン メニューの各項目を使用すると、表示される情報やプロンプトに従って、特定の処理を実行することができます。多くの質問に対してデフォルトの回答が用意されているので、一般的な設定を簡単に選択できます。vxdiskadm の使用方法については、vxdiskadd (1M) マニュアル ページを参照してください。
- ◆ vxdiskadd - Volume Manager に標準ディスクを追加するために使用されるコマンド。vxdiskadd により、情報やプロンプトが表示され、新しいディスクを初期化するプロセスが実行できます。vxdiskadd の使用方法については、vxdiskadd (1M) マニュアル ページを参照してください。
- ◆ vxdisk - ディスクを VxVM の管理下に配置するコマンド。vxdisk により、特殊ディスク デバイスの定義、ディスク上に格納されている情報 (Volume Manager がディスクを識別および管理するとき使用) の初期化、およびその他の特殊な操作の実行が可能になります。vxdisk の使用方法については、vxdisk (1M) マニュアル ページを参照してください。
- ◆ vxdbg - ディスク グループを操作するコマンド。vxdbg により、新しいディスク グループの作成および既存のディスク グループの管理が可能になります。vxdbg の使用方法については、vxdbg (1M) マニュアル ページを参照してください。
- ◆ vxunrelocate - ホットリロケーション機能によってリロケートされたサブディスクを元のディスクに戻すコマンド。

vxdiskaddユーティリティおよび多くのvxdiskadm操作は、標準ディスク デバイスでのみ使用できます。

ほとんどの Volume Manager コマンドでは、-g オプションを使用してディスク グループを指定できます。たとえば、ディスク グループ mktldg にボリュームを作成するには、次のコマンドを使用します。

```
# vxassist -g mktldg make mktvol 50m
```

このボリュームの (ブロック) ボリューム デバイスは次のようになります。

```
/dev/vx/dsk/mktldg/mktvol
```



オブジェクト名が一意的な場合は、ディスク グループを指定する必要はありません。ほとんどの Volume Manager コマンドは、コマンド ラインで指定されたオブジェクト名を使用して、対象のディスク グループを特定します。たとえば、ディスク グループ名を指定しないでディスク mkt dg01 上にボリュームを作成するには、次のように入力します。

```
# vxassist make mktvol 50m mkt dg01
```

2つのディスク グループに同じ名前のオブジェクトが存在しない限り、多くのコマンドはこのような働きをします。たとえば、Volume Manager では、root dg と mkt dg の両方のディスク グループ上にボリューム mktvol を作成できます。この場合、mkt dg ディスク グループ上でボリュームを操作するすべてのコマンドに -g mkt dg を追加する必要があります。

ディスクの初期化と追加

ディスクを VxVM に追加すると、ディスクは初期化またはカプセル化されます。カプセル化ではディスク上の既存データが保持されますが、初期化では既存データが削除されます。

Volume Manager でのディスクの初期化には、次の2つのレベルがあります。

1. ディスク メディア自体をフォーマットする。これは、Volume Manager の外部で実行する必要があります。
2. Volume Manager で使用する識別情報と設定情報をディスクに格納する。このレベルのディスク初期化は、Volume Manager インタフェースを使用して実行できます。

完全に初期化されたディスクはディスク グループに追加され、障害を起したディスクの交換や新しいディスク グループの作成用に使用できます。これらについては、この章で後述します。

ディスク メディアのフォーマット

初期化の第一段階として、対話型の format (一部のシステムでは diskadd) コマンドを使用して、ディスクのメディア フォーマットを実行します。

注 SCSI ディスクは通常フォーマット済みです。一般に、format (または diskadd) コマンドは、フォーマットが著しく破損している場合のみ必要となります。

Volume Manager ディスクのインストール

vxdiskadm メニューまたは vxdiskadd のいずれかを使用して、ディスクの初期化を行うことができます。ここでは、vxdiskadd の使用方法について説明します。コントローラ上にある1つのディスクまたはすべてのディスクを初期化するための vxdiskadm の使用方法については、本章の「vxdiskadm メニュー インタフェース処理 (116 ページ)」を参照してください。

`vxdiskadd` を使用して指定したディスクを初期化できます。たとえば、1 番目のコントローラ上にある 2 番目のディスクを初期化するには、次のコマンドを入力します。

```
# vxdiskadd c0t1d2
```

`vxdiskadd` は、ディスクが既に初期化されているかどうかを確認し、その結果に応じてプロンプトを表示します。また、カプセル化できるディスク（「特殊なカプセル化用の `vxdisk` の使用（112 ページ）」を参照）、**Volume Manager** に既に追加されているディスク、およびその他の条件についても確認します。

注 初期化されていないディスクを追加すると、`vxdiskadd` の実行中に警告およびエラー メッセージがコンソールに表示されます。これらのメッセージは無視してください。ディスクが完全に初期化されると、これらのメッセージは表示されなくなります。初期化が完了すると、処理が正常に完了したことを示すメッセージが表示されます。

次に示すプロンプトに対して、「**y**」を入力するか、または **Return** キーを押して処理を続行します。

ディスクの追加または初期化

Menu:VolumeManager/Disk/AddDisks

選択したディスクの一覧を以下に表示します。出力形式:[Device_Name]

c0t1d0

操作を続行しますか?[y,n,q,?] (デフォルト:y) **y**

ディスクが初期化されていない場合や、ディスクを再初期化することを選択した場合は、上記のプロンプトが表示されます。

ディスクを既存のディスク グループに追加するか、新しいディスク グループに追加するかを選択します。また、後の作業で追加や交換操作を行うためにディスクを確保しておくこともできます。新しいディスク グループを作成するには、使用していないディスク グループ名を指定します。ディスクを後の作業で使用するために確保しておくには、ディスク グループ名に「none」を指定します。

ディスク グループの指定 [<group>,none,list,q,?] (デフォルト:rootdg)

このディスクをデフォルトのグループ `rootdg` に追加するには、**Return** キーを押します。交換用ディスクとして空けておく（ディスク グループに追加しない）場合は、「**none**」と入力します。この後で、ディスク グループに追加するディスク名の指定を促すプロンプトが表示されます。

このディスクにデフォルトのディスク名を使用しますか?[y,n,q,?] (デフォルト:y) **y**



通常はデフォルトのディスク名を使用します（特殊ディスクの名前を入力する場合を除く）。

ディスクをホットリロケーション スペア ディスクとして使用しない場合は、次に示すプロンプトに対して「**n**」を入力します。

```
ディスクを rootdg のスペア ディスクとして追加しますか?[y,n,q,?] (デフォルト :n) n
```

ディスクをホットリロケーション利用対象から除外しない場合は、次のプロンプトで「**n**」を入力します。

```
他のディスクをホットリロケーションへの利用対象から除外しますか?[y,n,q,?]  
[デフォルト :n] n
```

「**y**」を入力するか、または **Return** キーを押して操作を続行します。

選択したディスクは、デフォルトのディスク名でディスク グループ rootdg に追加されます。

```
c0t1d0
```

```
操作を続行しますか?[y,n,q,?] (デフォルト :y) y
```

保存する必要のあるデータがディスク上にないことが確実な場合は、次に示すプロンプトに対して「**n**」を入力します。

次のディスク デバイスには有効な VTOC がありますが、Volume Manager 用に初期化されていません。ディスク上に重要なデータがある場合は、ディスクを新しいディスクとして追加するのではなく、既存のディスクパーティションをボリュームとしてカプセル化します。出力形式:[Device_Name]

```
c0t1d0
```

```
Encapsulate this device?[y,n,q,?] (デフォルト :y)
```

ディスクの初期化を促すプロンプトが表示されたら、「**y**」を入力します。

```
Instead of encapsulating, initialize?[y,n,q,?] (デフォルト :n) y
```

ディスク `c1t0d1` が Volume Manager の管理下に組み入れられることを示す次のようなメッセージが表示されます。システムによっては、表面分析を実行するかどうかを選択するオプションが表示される場合もあります。

デバイス `c0t1d0` を初期化しています。

表面分析の実行 (推奨)

`[y,n,q,?]` (デフォルト :`y`) **n**

ディスク デバイス `c0t1d0` を、`disk33` というディスク名でディスク グループ `rootdg` に追加します。

Volume Manager へのディスクの追加

ボリュームのディスク領域を使用するには、あらかじめディスクを Volume Manager の管理下に置いて、ディスク グループに追加しておく必要があります。Volume Manager の管理下に配置せずに使用していたディスクは、ディスク上の既存データを維持したまま Volume Manager の管理下に置くことができます。この処理には、Volume Manager のカプセル化機能を使用します。カプセル化ではディスク上の既存データが保持されます。ディスクが新しい場合は、初期化する必要があります。初期化すると、ディスク上の既存データは削除されます。Volume Manager の管理下に配置せずに使用していたディスクも、既存データを維持する必要がない場合は初期化する必要があります。

ディスクを追加するには、次のコマンドを使用します。

```
# vxdiskadd デバイス名
```

デバイス名には追加するディスクのデバイス名を指定します。

デバイス `c1t0d0` を Volume Manager の管理下に追加するには、次の処理を実行します。

1. 次のように入力して `vxdiskadd` を起動します。

```
# vxdiskadd c1t0d0
```

2. 処理を続行するには、次に示すプロンプトで、「**y**」を入力するか、または Return キーを押します。

ディスクの追加または初期化

```
Menu:VolumeManager/Disk/AddDisks
```

選択したディスクの一覧を以下に表示します。出力形式 :`[Device_Name]`

```
c1t0d0
```

操作を続行しますか ?`[y,n,q,?]` (デフォルト :`y`) **y**



- 次に示すプロンプトで、ディスクの追加先ディスク グループを指定するか、または **Return** キーを押して **rootdg** を使用します（詳細については、「ディスクのオフライン（86 ページ）」を参照してください）。

ディスクを既存のディスク グループに追加するか、新しいディスク グループに追加するかを選択します。また、後の作業で追加や交換操作を行うためにディスクを確保しておくこともできます。新しいディスク グループを作成するには、使用していないディスク グループ名を指定します。ディスクを後の作業で使用するために確保しておくには、ディスク グループ名に「none」を指定します。

ディスク グループの指定 [`<group>,none,list,q,?`] (デフォルト :`rootdg`)

- 次に示すプロンプトで、**Return** キーを押してデフォルト ディスク名を使用するか、または別のディスク名を入力します。

このディスクにデフォルトのディスク名を使用しますか？ [`y,n,q,?`] (デフォルト :`y`)

- ディスクをホットリロケーション スペア ディスクにするかどうかを確認するプロンプトが表示されたら、「**n**」を入力するか、または **Return** キーを押します。

ディスクを `rootdg` のスペア ディスクとして追加しますか？ [`y,n,q,?`] (デフォルト :`n`) **n**

詳細については、「`vxdiskadm` メニュー インタフェース処理（116 ページ）」を参照してください。

- ディスクをホットリロケーションの利用対象から除外しない場合は、次のプロンプトで「**n**」を入力します。

他のディスクをホットリロケーションへの利用対象から除外しますか？ [`y,n,q,?`]

[デフォルト :`n`] **n**

- 処理を続行するには、次に示すプロンプトで、「**y**」を入力するか、または **Return** キーを押します。

選択したディスクは、デフォルトのディスク名でディスク グループ `rootdg` に追加されます。

`c1t0d0`

操作を続行しますか？ [`y,n,q,?`] (デフォルト :`y`) **y**

8. ディスク上のデータを維持する必要がある場合は、「**y**」を入力してカプセル化を選択します。

次のディスク デバイスには有効な VTOC がありますが、Volume Manager 用に初期化されていません。ディスク上に重要なデータがある場合は、ディスクを新しいディスクとして追加するのではなく、既存のディスクパーティションをボリュームとしてカプセル化します。出力形式:[Device_Name]

```
c1t0d0
```

```
Encapsulate this device?[y,n,q,?] (デフォルト :y) y
```

9. 処理を続行するには、次に示すプロンプトで **Return** キーを押します。

次のディスクがカプセル化の対象として選択されています。

出力形式:[Device_Name c1t0d0]

カプセル化を続行しますか?[y,n,q,?] (デフォルト :y)

Volume Manager で使用するためにディスクがカプセル化されることを示す次のようなメッセージが表示されます。

ディスク デバイス c1t0d0 はカプセル化され、disk01 という名前でディスク グループ rootdg に追加されます。

ディスク c1t0d0 はカプセル化用に設定されています。

カプセル化の初期段階が問題なく完了しました。

できるだけ早い段階でシステムをリブートしてください。

カプセル化操作では数回リブートする必要があります。これからリブートすると、その後自動的に数回リブートされます。

リブートするには、次のコマンドを実行します：

```
shutdown -g0 -y -i6
```

このコマンドにより /etc/vfstab ファイルが更新され、ボリューム デバイスを使用してファイル システムをこのディスク デバイス上にマウントできるようになります。

バックアップ スクリプト、データベース、または手動作成したスワップ デバイスなどの参照項目を更新する必要があります。

終了します。

必要に応じて、シャットダウンおよび再起動してください。



ディスク グループへのディスクの追加

新しいディスクを作成済みのディスク グループに追加することができます。たとえば、アプリケーションやワーク グループの要件が変更されたため、現在のディスクに必要な領域が足りなくなった場合などにディスクを追加できます。

初期化したディスクをディスク グループに追加するには、次のコマンドを使用します。

```
# vdiskadd デバイス名
```

デバイス `c1t1d0` を `rootdg` に追加するには、次の処理を実行します。

1. 次のコマンドを入力して `vxdiskadd` を起動します。

```
# vxdiskadd c1t1d0
```

`vxdiskadd` が起動すると、次のようなメッセージが表示されます。

ディスクの追加または初期化

Menu:VolumeManager/Disk/AddDisks

選択したディスクの一覧を以下に表示します。出力形式:[Device_Name]

```
c1t1d0
```

操作を続行しますか?[y,n,q,?] (デフォルト:y) **y**

2. 次に示すプロンプトで、ディスクの追加先ディスク グループを指定するか、または **Return** キーを押してデフォルトの `rootdg` を使用します。

ディスクを既存のディスク グループに追加するか、新しいディスク グループに追加するかを選択します。また、後の作業で追加や交換操作を行うためにディスクを確保しておくこともできます。新しいディスク グループを作成するには、使用していないディスク グループ名を指定します。ディスクを後の作業で使用するために確保しておくには、ディスク グループ名に「none」を指定します。

ディスク グループの指定 [`<group>`,none,list,q,?] (デフォルト:rootdg)

3. 次に示すプロンプトで、**Return** キーを押してデフォルト ディスク名を使用するか、または別のディスク名を入力します。

このディスクにデフォルトのディスク名を使用しますか?[y,n,q,?] (デフォルト:y)

4. ディスクをホットリロケーション スペア ディスクにするかどうかを確認するメッセージが表示されたら、「**n**」を入力するか、または **Return** キーを押します。

ディスクを `rootdg` のスペア ディスクとして追加しますか?[y,n,q,?] (デフォルト:n)**n**

5. ディスクをホットリロケーションの利用対象から除外しない場合は、次のプロンプトで「**n**」を入力します。

他のディスクをホットリロケーションへの利用対象から除外しますか?[y,n,q,?]

[デフォルト :n] **n**

6. 処理を続行するには、次に示すプロンプトで、「**y**」を入力するか、または Return キーを押します。

選択したディスクは、デフォルトのディスク名でディスクグループ rootdg に追加されます。

c1t1d0

操作を続行しますか?[y,n,q,?] (デフォルト :y) **y**

7. このディスクが既に Volume Manager での使用に備えて初期化済みであることを示す次のようなプロンプトが表示されます。「**y**」を入力して、このディスクの使用を確定します。

次のディスク デバイスは既に初期化されています。

ディスクは現在、交換ディスクとして使用できます。

出力形式 :[Device_Name]

c1t1d0

Use this device?[y,n,q,?] (デフォルト :y) **y**

8. ディスクを再初期化するには、次に示すプロンプトで、「**y**」を入力するか、または Return キーを押します。

選択した次のディスクは、既に Volume Manager 用に初期化されています。ディスクが既に Volume Manager 用に初期化されていることが確実な場合は、ディスクデバイスを再初期化する必要はありません。

出力形式 :[Device_Name]

c1t1d0

Reinitialize this device?[y,n,q,?] (デフォルト :y) **y**



Volume Manager で使用するためにディスクを再初期化していることを示す次のようなメッセージが表示されます。システムによっては、表面分析を実行するかどうかを選択するオプションが表示される場合もあります。

デバイス c1t1d0 を初期化しています。

表面分析の実行 (推奨)

[y,n,q,?] (デフォルト :y)n

ディスク デバイス c1t1d0 を、disk03 というディスク名で追加しています。

ディスクがディスク グループに追加されていることを確認するには、次のように入力します。

```
# vxdisk list
```

上記のコマンドを実行すると、次のような結果が表示されます。

DEVICE	TYPE	DISK	GROUP	STATUS
c0t0d0s2	sliced	disk04	rootdg	online
c1t0d0s2	sliced	disk01	rootdg	online
c1t1d0s2	sliced	disk03	rootdg	online

ディスク グループからのディスクの削除

サブディスクを含まないディスクは、ディスク グループから削除できます。次のコマンドを使用します。

```
# vxdg [-g グループ名] rmdisk ディスク名
```

グループ名には、デフォルトの rootdg 以外のディスク グループのみを指定できます。

たとえば、rootdg から disk02 を削除するには、次のように入力します。

```
# vxdg rmdisk disk02
```

ディスク上にサブディスクがある場合は、削除しようとする、次のエラー メッセージが表示されます。

```
vxdg: ディスクディスク名はサブディスクで使用されています。
```

-k オプションを指定して vxdg を実行し、デバイスの割り当てを削除します。-k オプションを使用すると、サブディスクがある場合でもディスクを削除できます。詳細については、vxdg (IM) マニュアルページを参照してください。

注 vxdg に -k オプションを指定すると、データが失われる恐れがあります。

ディスクをディスク グループから削除した後で、そのディスクを Volume Manager の管理下から完全に削除できません (オプション)。次コマンドを使用します。

```
# vxdisk rm デバイス名
```

たとえば、c1t0d0 (または c1b0t0d0) を Volume Manager の管理下から削除するには、次のように入力します。

```
# vxdisk rm c1t0d0s2
```

バスを使用しているシステムの場合は、次のように入力します。

```
# vxdisk rm c1b0t0d0s0
```

複数のサブディスクが定義されているディスクを削除することができます。たとえば、すべてのボリュームを 1 つのディスク上にまとめることができます。vxdiskadm を使用してディスクを削除する場合は、そのディスクからボリュームを移動するように選択できます。この処理を行うには、vxdiskadm を実行し、メイン メニューの項目 3[ディスクの削除] を選択します。

そのディスクにいくつかのサブディスクが含まれている場合、通常、次のようなメッセージが表示されます。



次のボリュームは、ディスク `disk02` の一部を現在使用しています：

```
home usrvol
```

ボリュームを削除するには、その前に `disk02` から移動する必要があります。

ボリュームを他のディスクに移動しますか？`[y,n,q,?]` (デフォルト :n)

y を選択すると、すべてのサブディスクがディスクから移動されます (可能な場合)。移動できないサブディスクもあります。サブディスクを移動できない場合、一般に次のことが原因と考えられます。

- ◆ 残りのディスクに十分な領域がない。
- ◆ ボリューム内にある既存のプレックスやストライプ サブディスクを別のディスクに割り当てることができない。

`vxdiskadm` によって移動できないサブディスクがある場合、ディスクの削除を続行する前に、一部のディスクからプレックスの一部を削除し、空き領域を確保する必要があります。

ディスクの移動

ディスク グループ間でディスクを移動するには、移動するディスクをディスク グループから削除し、別のディスク グループに追加します。たとえば、ディスク `disk04` という名前が接続されている物理ディスク `c0t3d0` をディスク グループ `rootdg` から移動して、ディスク グループ `mktgdg` に追加するには、次のコマンドを使用します。

```
# vxdbg rmdisk disk04  
# vxdbg -g mktgdg adddisk mktgdg02=c0t3d0
```

注 この方法では、ディスク上の設定やデータは保存されません。

`vxdiskadm` を使用してディスクを移動することもできます。メイン メニューの項目 3[ディスクの削除]を選択してから、項目 1[ディスクの追加または初期化]を選択します。

ディスク名の変更

ディスクに Volume Manager 名を指定していない場合は、ディスクを Volume Manager の管理下に配置すると、デフォルトの名前が設定されます。Volume Manager 名は、ディスクの位置やディスクタイプを識別するために使用されます。ディスク名を変更して所有権や使用権に反映するには、次のように入力します。

```
# vxedit rename old_diskname new_diskname
```

disk01 を disk03 に変更するには、次のように入力します。

```
# vxedit rename disk01 disk03
```

名前が変更されたことを確認するには、次のように入力します。

```
# vxdisk list
```

上記のコマンドを実行すると、次のような結果が表示されます。

DEVICE	TYPE	DISK	GROUP	STATUS
c0t0d0s2	sliced	disk04	rootdg	online
c1t0d0s2	sliced	disk03	rootdg	online
c1t1d0s2	sliced	-	-	online

注 デフォルトでは、サブディスクオブジェクト名は、それらのオブジェクトが属する VM ディスクの名前に基づいて設定されます。ただし、VM ディスク名を変更しても、そのディスク上のサブディスク名は自動的に変更されません。



ディスクの予約

デフォルトでは、`vxassist` は空き領域のある任意のディスクから領域を割り当てます。処理速度が特に遅いディスクまたは速いディスクを通常の場合には使用しないようにする場合など、特別な目的のためにディスクを予約することができます。

ディスクを特別な目的のために予約するには、次のコマンドを使用します。

```
# vxedit set reserve=on ディスク名
```

このコマンドを入力しても、`vxassist` は、コマンド ラインで特にディスクを指定しない限り、選択されたディスク領域を割り当てません。たとえば、ディスク `disk03` が予約されている場合は、次のコマンドを使用します。

```
# vxassist make vol03 20m disk03
```

このコマンドを実行すると、予約が無効になり、`disk03` に 20 メガバイトのボリュームが作成されます。ただし、次のコマンドを入力した場合は、結果が異なります。

```
# vxassist make vol04 20m
```

このコマンドを実行すると、ほかのディスクに空き領域がない場合でも `disk03` は使用されません。

ディスクの予約を無効にするには、次のように入力します。

```
# vxedit set reserve=no ディスク名
```

ディスクのオフライン

必要に応じて、物理ディスクをオフライン状態にします。ディスクが破損している場合は、ディスクを無効にして削除する必要があります。物理ディスク デバイスを移動し、別のシステムに接続する場合も、事前にディスクを無効にする必要があります。

物理ディスクをオフライン状態にするには、まずディスク グループからディスクを削除します（詳細については、「[ディスクの削除 \(88 ページ\)](#)」を参照）。次のコマンドを使用して、ディスクを「オフライン」状態に設定します。

```
# vxdisk offline デバイス名
```

注 ディスクがディスク グループから削除され、管理名が使用できないため、ここではデバイス名が使用されます。

ディスクのミラーリング

ルート ディスクをミラーリングすると、起動に必要なルート ボリュームやその他の領域が別のディスクにミラー化されます。

使用中のブート（ルート）ディスクをほかのディスクにミラー化するには、ここで説明する手順または `vxdiskadm` を実行してください。ブート ディスクをミラー化しておく、ブート ディスクの障害時にブート ディスクをそのミラーと入れ替えることによって、障害をリカバリすることができます。

ディスクの交換が必要な場合に、Volume Manager ブート ディスク エイリアスを利用してブート ディスクのミラーを識別するには、ブート ディスクをミラー化する前に、EEPROM 変数 `use-nvramrc?` を `true` に設定しておく必要があります。この変数が `false` に設定されている場合、どれがブート可能なディスクかをユーザ自身で判断する必要があります。この変数を `true` に設定するには、次のように入力します。

```
eeeprom use-nvramrc?=true
```

ブート ディスクをミラー化するには、次の処理を実行します。

1. ブート ディスク以上のサイズを持つディスクを選択します。
2. `vxdiskadd` コマンドを使用して、選択したディスクを新しいディスクとして追加します（まだ追加していない場合）。
3. 次のコマンドを実行します。

```
# /etc/vx/bin/vxrootmir 代替ディスク
```

代替ディスクには、ミラー化によって作成されたディスクに割り当てる名前を指定します。

`vxrootmir` を実行すると、`rootvol`（代替ディスク上のルート ファイルシステムのボリューム）のミラーが作成されます。

プライマリ ブート ディスクに障害が発生した場合に、代替ブート ディスクから起動するように設定できます。

ブート ディスク上には、`/home` または `/tmp` ファイル システムなどのボリュームがある場合があります。これらのボリュームは、`vxassist` ユーティリティを使用して個別にミラー化できます。たとえば、`homevol` ボリューム上に `/home` ファイル システムがある場合は、次のコマンドを使用して代替ディスクにミラー化することができます。

```
# vxassist mirror homevol 代替ディスク
```

代替ブート ディスク上にファイル システムのコピー用の領域がない場合は、別のディスクにミラー化できます。また、これらのボリュームを、システムに接続されているほかのディスクにわたってスパンまたはストライプすることもできます。

プライマリ ブート ディスク上のすべてのボリュームを一覧表示するには、次のコマンドを使用します。

```
# vxprint -t -v -e\'aslist.aslist.sd_disk="ブート ディスク"
```



このディスク上の連続するすべてのボリュームを代替ブート ディスクにミラー化するには、次のコマンドを使用します。

```
# /etc/vx/bin/vxmirror ブート ディスク 代替ディスク
```

ディスクの削除

ディスクに障害が発生した場合は、ディスクをシステムから削除して別のシステムに移動することができます。現在のシステムからディスクを削除する前に、次の処理を実行する必要があります。

1. ボリューム上のすべてのファイル システムをマウント解除します。
2. ディスク上のボリュームを停止します。
3. ボリュームをほかのディスクに移動するか、バックアップを作成します。ボリュームを移動するには、1つ以上のほかのディスク上にボリュームをミラー化してから、元のボリュームを削除します。ボリュームが不要になった場合は、移動せずに削除することもできます。

ディスクを削除するには、次の処理を実行します。

4. 次のコマンドを使用して、ディスク グループからディスクを削除します。

```
# vxdbg [-g グループ名] rmdisk ディスク名
```

グループ名にはディスクの属するグループの名前、**ディスク名**には削除するディスクの名前をそれぞれ指定します。

たとえば、rootdg から disk01 を削除するには、次のように入力します。

```
vxdbg rmdisk disk01
```

rootdg はデフォルトのディスク グループであるため、ここでは特に指定する必要はありません。

5. 次のコマンドを使用して、ディスクを Volume Manager の管理下から削除します。

```
# vxdisk rm デバイス名
```

たとえば、c1t0d0 を Volume Manager の管理下から削除するには、次のように入力します。

```
# vxdisk rm c1t0d0s2
```

ディスク情報の表示

ディスクを使用する前に、そのディスクが初期化され、Volume Manager の管理下にあることを確認する必要があります。また、ディスクグループに含まれないディスク上にボリュームを作成することはできないため、ディスクがディスクグループに属しているかどうかも確認しておく必要があります。vxdisk list コマンドを実行すると、認識されたすべてのディスクのデバイス名、ディスク名、各ディスクに関連付けられているディスクグループ名、および各ディスクの状態が表示されます。

次のコマンドを使用すると、Volume Manager に定義されているすべてのディスク上の情報を表示することができます。

```
# vxdisk list
```

上記のコマンドを実行すると、次のような結果が表示されます。

DEVICE	TYPE	DISK	GROUP	STATUS
c0t0d0s2	sliced	disk04	rootdg	online
c1t0d0s2	sliced	disk01	rootdg	online
c1t1d0s2	sliced	-	-	online

Volume Manager に定義された特定のディスクの詳細を表示するには、次のように入力します。

```
# vxdisk list disk01
```

vxdisk コーティリティによって、特定のメタデバイスの動的マルチパス情報が表示されます。メタデバイスとは、システムの I/O コントローラからの複数の物理パスを持つ特定の物理ディスクのことです。Volume Manger では、システムのすべての物理ディスクを 1 つ以上の物理パスを持つメタデバイスとして表します。

次のコマンドを使用すると、特定のメタデバイスのマルチパス情報を表示することができます。

```
# vxdisk list デバイス
```

デバイスは、Volume Manager の DMP サブシステムによって構成されたメタデバイスです。

上記の出力では、メタデバイス c2t0d0s2 として示される物理デバイスへの 2 つのパスが表示されています。パス c2t0d0s2 はアクティブ状態 (state=enabledÅj で、もう一方のパス c1t0d0s2 は非アクティブ状態 (state=disabledÅj です。



上記のコマンドを実行すると、次のような結果が表示されます。

```
Device      device
devicetag   c2t0d0
type        sliced
hostid      aparajita
disk        name=disk01 id=861086917.1052.aparajita
group       name=rootdg id=861086912.1025.aparajita
flags       online ready autoconfig autoimport imported
pubpaths    block=/dev/vx/dmp/c2t0d0s4 char=/dev/vx/rdmp/c2t0d0s4
privpaths   block=/dev/vx/dmp/c2t0d0s3 char=/dev/vx/rdmp/c2t0d0s3
version     2.1
iosize      min=512 (bytes) max=2048 (blocks)
public      slice=4 offset=0 len=1043840
private     slice=3 offset=1 len=1119
update      time=861801175 seqno=0.48
headers     0 248
configs     count=1 len=795
logs        count=1 len=120
Defined regions
config      priv 000017-000247[000231]:copy=01 offset=000000 enabled
config      priv 000249-000812[000564]:copy=01 offset=000231 enabled
log         priv 000813-000932[000120]:copy=01 offset=000000 enabled
Multipathing information:
numpaths:   2
c2t0d0s2    state=enabled      type=primary
c1t0d0s2    state=disabled     type=secondary
```

アクティブ/パッシブ タイプ ディスク アレイのディスクには、**type** に追加情報が表示されます。ここには、ディスクへのプライマリ パスおよびセカンダリ パスが表示されます。たとえば、DG Clariion、Hitachi DF350 などのディスク アレイの場合です。

この **type** 情報はアクティブ/アクティブ タイプ ディスク アレイ上のディスクには表示されません。これらのディスク アレイ上のディスクにはプライマリ パスとセカンダリ パスの概念がないためです。たとえば、StorEdge A5000、Sparc Storage Array (SSA) などのディスク アレイの場合です。

障害のあるディスクの検出と交換

ここでは、ディスクの障害を検出して障害のあるディスクを交換する方法について説明します。まず、障害の発生時に、冗長性のある **Volume Manager** のオブジェクトを自動的にリストアするホットリロケーション機能について説明します。

ホットリロケーション

ホットリロケーションは、冗長性のある（ミラーリングまたは RAID-5）**Volume Manager** オブジェクトで発生した入出力障害に自動的に対処し、これらのオブジェクトに対する冗長性およびアクセスをリストアします。**Volume Manager** は、オブジェクト上の入出力障害を検出し、影響を受けたサブディスクを、指定されているスペア ディスクまたはディスク グループ内の空き領域にリロケートします。次に、オブジェクトを障害発生前の状態にリストアし、冗長性とアクセスを再び確保します。ホットリロケーションについては、第 3 章の「ホットリロケーション（51 ページ）」を参照してください。

注 ホットリロケーションは、障害が発生したディスク上の冗長性のある（ミラーリングまたは RAID-5）サブディスクに関してのみ実行されます。障害ディスク上の冗長性を持たないサブディスクはリロケートされませんが、システム管理者に障害が発生したことが通知されます。

ホットリロケーションはデフォルトで有効になっており、システム管理者による介入を必要とすることなく障害発生時に機能します。ホットリロケーション デモン `vxrelocd` は、次の種類の障害を示す **Volume Manager** のイベントの検出および対処を行います。

- ◆ ディスク障害 — 通常、**Volume Manager** オブジェクトの入出力に障害が発生すると検出されます。**Volume Manager** はエラーの修復を試みます。エラーが修復できない場合、**Volume Manager** はディスクの専有領域にある設定情報へのアクセスを試みます。専有領域にアクセスできない場合、ディスクが破損していると判断されます。
- ◆ プレックス障害 — 通常、プレックス内に修復できない入出力エラーが発生すると検出されます（プレックス内のサブディスクに影響します）。ミラー ボリュームの場合、プレックスは切り離されます。
- ◆ RAID-5 サブディスク障害 — 通常、修正できない入出力エラーが発生すると検出されます。サブディスクは切り離されます。

このような障害が検出された場合、`vxrelocd` は、障害の内容と影響を受けた **Volume Manager** オブジェクトを通知する電子メールをシステム管理者に送信します。リロケートできるサブディスクがある場合は、そのサブディスクが特定されます。リロケートできる場合、適切なリロケート用の領域が探し出され、サブディスクがリロケートされます。

ホットリロケーション用の領域は、障害が発生したディスク グループ内でホットリロケーション用に予約されたディスクから選択されます。有効なスペア ディスクが存在せず、領域の追加が必要な場合は、同一のディスク グループ内の空き領域が使用されます。



ただし、`nohotuse` フラグの設定されたディスク (ホットリロケーションの利用対象から除外されたディスク) は使用されません。サブディスクのリロケート後、リロケートされた各サブディスクがプレックスに再接続されます。

最後に `vxrelocd` は、適切なリカバリ処理を開始します。たとえば、ミラー ボリュームのミラーの再同期化や RAID-5 ボリュームのデータのリカバリなどが実行されます。システム管理者には、ホットリロケーションとリカバリ処理が実行されたことが通知されます。

リロケート処理を実行できない場合はシステム管理者に通知され、それ以上の処理は実行されません。次の場合には、リロケート処理を実行できません。

- ◆ サブディスクが冗長性を持たない場合 (ミラー ボリュームまたは RAID-5 ボリュームに属さない場合)、そのサブディスクはリロケートできません。
- ◆ ディスク グループ内に十分な領域 (スペア ディスクまたは空き領域のうち `nohotuse` フラグが設定されホットリロケーションの利用対象から除外されているものを除く) がない場合、障害が発生しているサブディスクはリロケートできません。
- ◆ 唯一使用できる領域が、障害が発生しているプレックスのミラーが含まれるディスク上にある場合、そのプレックス内のサブディスクはリロケートできません。
- ◆ 唯一使用できる領域が、RAID-5 ボリュームのログ プレックスまたは正常なサブディスクの 1 つが含まれるディスク上にある場合、RAID-5 プレックス内の障害サブディスクはリロケートできません。
- ◆ ミラー ボリュームにデータ プレックスの一部としてダーティ リージョン ロギング サブディスクがある場合、そのプレックスに属するサブディスクはリロケートできません。
- ◆ RAID-5 ボリュームのログ プレックスまたはミラー ボリュームの DRL ログ プレックスに障害が発生した場合、新しいログ プレックスが別の場所に作成されます (ログ プレックスは実際にはリロケートされません)。

ホットリロケーション用のスペア ディスクをディスク グループごとに1つ以上指定して、ホットリロケーションに備えておきます。ディスクをスペアとして設定する方法については、本章の「VM ディスクのホットリロケーションプールへの追加 (137 ページ)」を参照してください。障害発生時に有効なスペア ディスクがない場合、またはスペア ディスク上に十分な領域がない場合は、ディスクに `nohotuse` フラグが設定されていてホットリロケーションから除外されていない限り、空き領域が自動的に使用されます。

スペア ディスクを指定しておく、障害時にリロケートに使用される領域を制御できません。空き領域とスペア ディスク上の領域を合計しても十分ではない場合や、冗長性の制約を満たしていない場合、サブディスクはリロケートされません

ディスクに `nohotuse` フラグを設定しておく、交換用のディスクとして使用しないディスクを制御できます。

ディスクをホットリロケーションの利用対象から除外するには、2通りの方法があります。たとえば、次のように入力します。

```
# vxedit -g rootdg set nohotuse=on disk01
# vxdiskadm
```

vxdiskadm メイン メニューの項目 15[ディスクのホットリロケーション利用対象からの除外]を選択します。

ディスクをホットリロケーションの利用対象にするには、2通りの方法があります。たとえば、次のように入力します。

```
# vxedit -g rootdg set nohotuse=off disk01
# vxdiskadm
```

vxdiskadm メイン メニューの項目 16[ディスクのホットリロケーション設定]を選択します。

スペア ディスクまたはホットリロケーションの利用対象から除外されているディスクを確認するには、2通りの方法があります。たとえば、次のように入力します。

```
# vxdisl list
# vxprint
```

正常にリロケートが実行されたら、障害を起したディスクを削除して交換する必要があります（「ディスクの交換（101 ページ）」を参照）。リロケートされたサブディスクの場所によっては、ホットリロケーションの実行後、リロケートされたサブディスクを別の場所に移動することもできます（「リロケートされたサブディスクの移動（95 ページ）」を参照）。

vxrelocd の変更

ホットリロケーションは、vxrelocd の稼働中は有効になっています。ホットリロケーションを有効な状態にしておくと、障害発生時にこの機能を活用できます。ただし、この機能を無効にする場合（ディスク上の空き領域をリロケートに使用しない場合）は、システムの起動時に vxrelocd を起動しないようにしてください。

システム起動時にホットリロケーションを無効にする方法については、『VERITAS Volume Manager Installation Guide』を参照してください。vxrelocd プロセスを強制終了して、いつでもホットリロケーションを停止できます（ホットリロケーションが進行中の場合は、強制終了しないでください）。

vxrelocd を呼び出す起動ファイル（/etc/rc2.d/S95vxvm-recover）の vxrelocd 行を編集するか、実行中の vxrelocd プロセスを強制終了し別のオプションを指定して再び起動すると、vxrelocd の動作を多少変更することができます。起動ファイル内で vxrelocd を呼び出す方法を変更した場合は、システムをリブートして変更を有効にする必要があります。デーモンを強制終了して再起動する場合は、vxrelocd プロセスを強制終了するときに、ホットリロケーションが進行中でないことを確認してください。この場合も、障害発生時にホットリロケーションを活用できるように、デーモンをすぐに再起動する必要があります。



vxrelocd の動作を変更するには、次の操作を実行します。

- ◆ デフォルトでは、障害が検出されてリロケートが行われると、vxrelocd は電子メールをルート ユーザに送信します。ほかのユーザにも通知するように vxrelocd を設定するには、次のように適切なユーザ名を追加して vxrelocd を呼び出します。

```
# vxrelocd root ユーザ名1 ユーザ名2 &
```

- ◆ リカバリ処理によるシステム パフォーマンスへの影響を軽減するには、ボリュームの各領域をリカバリする際に時間間隔を空けるように指定できます。次のように入力します。

```
# vxrelocd -o slow[= 入出力遅延時間 ] root &
```

入出力遅延時間には、必要な遅延時間（ミリ秒単位）を指定します。デフォルト値は、250 ミリ秒です。詳細については、vxrelocd (1M) マニュアル ページを参照してください。

-O: 旧バージョンに戻すオプションです。VxVM_version -O と指定すると、vxrelocd は指定したバージョンのリロケート処理を実行します。

-s: vxrelocd によるリロケートの開始前に、/etc/vx/saveconfig.d に現在の設定のスナップショットが作成されます。-s オプションには、各ディスク グループに保持される設定の最大数を指定します。デフォルト値は 32 です。

スペア ディスク情報の表示

コマンド vxdbg spare を使用して、リロケートに使用できるすべてのスペア ディスクの情報を表示できます。次のような結果が表示されます。

GROUP FLAGS	DISK	DEVICE	TAG	OFFSET	LENGTH
rootdgs	disk02	c0t2d0s2	c0t2d0	0	658007

この例では、disk02 が唯一指定されているスペア ディスクです。LENGTH フィールドには、このディスク上で現在使用できるリロケート用のスペア領域のサイズが示されています。

また、次のコマンドを使用して、現在指定されているスペア ディスクの情報を表示できます。

- ◆ vxdisklist - ディスク情報の一覧と spare フラグが設定されているスペア ディスクを表示します。
- ◆ vxprint - ディスクおよびその他の情報の一覧と spare フラグが設定されているスペア ディスクを表示します。

リロケートされたサブディスクの移動

ホットリロケーションを実行すると、サブディスクはスペアディスクやディスクグループ内の使用可能な空き領域にリロケートされます。サブディスクが移された場所では、ホットリロケーションの実行前と同じパフォーマンスやデータレイアウトを実現できない場合があります。ホットリロケーションの完了後、リロケートされたサブディスクを移動して、パフォーマンスを向上させることができます。

リロケートされたサブディスクをスペアディスクから移動して、次のホットリロケーションの必要時に使用できるようにスペアディスクを空けておくこともできます。また、ホットリロケーションが実行される前の構成設定に戻すためにサブディスクを移動する場合があります。

ホットリロケーション中、次のような電子メールメッセージがルートユーザに送信されます。

To: ルート

Subject: ホスト teal 上の Volume Manager エラー

プレックス home-02 からサブディスク disk02-03 をリロケートしようとしています。

Dev_offset 0 length 1164 dm_name disk02 da_name c0t5d0s2.

使用可能なプレックス home-01 を使用してデータをリカバリします。

このメッセージには、リロケート前のサブディスクの情報が含まれており、リロケート後にサブディスクの移動先を決定するために使用できます。

次のメッセージは、リロケートされたサブディスクの新しい場所を示しています。

To: ルート

Subject: VxVM を teal 上にリロケートします。

ボリューム home サブディスク disk02-03 を disk05-01 へリロケートしましたが、まだ、リカバリは実行していません。

リロケートされたサブディスクを移動する前に、前節の説明に従って、障害が発生したディスクを修復または交換します。この処理を終了すると、リロケートされたサブディスクを元のディスクに移動できるようになります。たとえば、リロケートされたサブディスク disk05-01 を disk02 に移動するには、次のコマンドを実行します。

```
# vxassist -g rootdg move home !disk05 disk02
```

注 サブディスクの移動処理中は、RAID-5 ボリュームは冗長性を持ちません。



vxunrelocate の使用

VxVM ホットリロケーションを使用すると、システムは冗長性のある VxVM オブジェクトの入出力障害に対してサブディスクレベルで自動対応し、オブジェクトを使用可能な状態にするための処理を実行できます。この機構によって、サブディスク上の入出力障害の検出、サブディスクのリロケート、サブディスクに関連付けられたプレックスのリカバリなどが行われます。ディスクの交換後は、`vxunreloc` を使用すると、ディスク障害が発生する前の設定にシステムを戻すことができます。`vxunreloc` は、障害によりホットリロケート済みのサブディスクを元のディスクに戻すことができます。

`vxunreloc` を起動したら、ホットリロケート済みサブディスクの元の位置を示すディスクメディア名を指定する必要があります。`vxunreloc` がサブディスクを移動する場合は、元のオフセットに移動します。障害が発生した元のディスクよりも小さいディスクのリロケート処理を解除しようとする、`vxunreloc` はエラーを返して処理を行いません。

`vxunreloc` には、サブディスクを元のディスクとは異なるディスクに移動するオプションがあります。また、サブディスクのリロケート処理を解除して異なるオフセットに移動するオプションもあります。ただし、これはデスティネーションディスクがすべてのサブディスクを収容できる大きさを持つ場合に限り指定できます。

`vxunreloc` がサブディスクを元のオフセットに戻すことができない場合は、強制オプションを使用すると、元のオフセットを使用しないでサブディスクを指定したディスクに移動できます。詳細については、`vxunreloc(1M)` マニュアル ページを参照してください。

次に示す例で、`vxunreloc` の使用方法を説明します。

例 1:

`disk01` に障害が発生し、すべてのサブディスクがリロケートされたと仮定します。`disk01` の交換後、`vxunreloc` を使用してホットリロケート済みのサブディスクをすべて `disk01` に移動するには、次のコマンドを実行します。

```
vxunreloc -g newdg disk01
```

例 2:

`vxunreloc` ユーティリティには、リロケート前の元のディスクとは異なるディスクにサブディスクを移動するための `-n` オプションがあります。

`disk01` に障害が発生し、ディスク上のすべてのサブディスクがホットリロケーション処理によってほかのディスクに移動したと仮定します。`vxunreloc` には、リロケート前の元のディスクとは異なるディスクにサブディスクを移動するためのオプションがあります。修復したディスクは、`disk05` など、別の名前を持つディスクグループに追加して戻すことができます。ホットリロケート済みのサブディスクをすべてこの新しいディスクに戻すには、次のコマンドを使用します。

```
vxunreloc -g newdg -n disk05 disk01
```

例 3:

disk01 に障害が発生し、サブディスクがリロケートされ、ホットリロケート済みのサブディスクを disk05 に移動すると仮定します。disk05 には、ほかのサブディスクが既に存在しています。次の強制オプションを使用すると、ホットリロケート済みのサブディスクを disk05 に移動できますが、正確なオフセットには移動できません。

```
vxunreloc -g newdg -f -n disk05 disk01
```

例 4:

ディスク障害が複数発生してサブディスクに複数回ホットリロケーション処理が実行されても、サブシステムのリロケート処理を解除して元の位置に戻すことができます。たとえば、disk01 に障害が発生し、サブディスク disk01-01 が disk02 に移動した後、disk02 にディスク障害が発生した場合、disk02 上のサブディスクはホットリロケーション処理されたサブディスクも含めてすべて再度移動できます。disk02 が交換された場合、disk02 に対して vxunreloc を実行しても、ホットリロケート済みのサブディスク disk01-01 には何も処理されません。ただし、disk01 交換の直後に vxunreloc を実行すると、disk01-01 は disk01 に戻されます。

障害の発生したディスクの修復または交換後は、vxunreloc により、ホットリロケート済みのサブディスクをすべて元のディスクに移動できます。サブディスクのホットリロケーション処理中、元のディスク メディア名およびディスクのオフセットは、設定データベースに保存されます。これらの情報は、vxunreloc がサブディスクを元のディスクまたは新しいディスクに移動する時点で消去されます。元の dm 名および元のオフセットは、サブディスク レコードに保存されています。rootdg ディスク グループの disk01 からホットリロケーション処理されたサブディスクをすべて印刷するには、次のコマンドを使用します。

```
# vxprint -g rootdg -se 'sd_orig_dmname="disk01"'
```

disk01 からホットリロケーション処理されたサブディスクをすべて元のディスクに戻すには、次のコマンドを使用します。

```
# vxunreloc -g rootdg disk01
```

vxunreloc ユーティリティには、リロケート前の元のディスクとは異なるディスクにサブディスクを移動するための -n オプションがあります。たとえば、disk01 に障害が発生し、disk01 上のすべてのサブディスクがホットリロケーション処理されて、別のディスクに移動したと仮定します。修復したディスクは、disk05 など、別の名前を持つディスク グループに追加して戻すことができます。ホットリロケート済みのサブディスクをすべてこの新しいディスクに戻すには、次のコマンドを使用します。

```
# vxunreloc -g rootdg -n disk05 disk01
```

デスティネーション ディスクには、少なくとも元のディスクと同じ記憶容量が必要です。十分な領域がない場合は、リロケート処理の解除中にエラーが発生し、サブディスクの移動は実行されません。



vxunreloc がホットリロケート済みのサブディスクを移動する場合は、元のオフセットに移動します。ただし、デスティネーションディスク上の一部またはすべての領域を使用しているサブディスクがあると、vxunreloc にエラーが発生します。基本的には、次の2つの選択肢があります。(1) 既存のサブディスクを別の場所に移動し、再度vxunreloc を実行する。(2) vxunreloc に -f オプションを指定して、サブディスクをデスティネーションディスクに移動する。ただし、ディスク上の領域は vxunreloc が検出します。デスティネーションディスクに十分な領域があり、すべてのサブディスクが保存可能な場合は、元のオフセットを使用することなく、ホットリロケート済みのサブディスクのリロケート処理がすべて解除されます。

エラー後の vxunreloc の再起動

内部的には、vxunreloc によるサブディスクの移動は3段階で実行されます。第1段階では、リロケート処理が解除されたサブディスクと同数のサブディスクが、指定されたデスティネーションディスクに作成されます。サブディスクの作成時に、vxunreloc はサブディスクレコードのコメントフィールドに文字列「UNRELOC」を識別情報として書き込みます。第2段階では、実際のデータの移動が実行されます。すべてのサブディスクの移動が成功すると、第3段階で、サブディスクレコードのコメントフィールドが補整されます。

サブディスクの作成は、部分的に処理されることはありません。すべてのサブディスクを完全に作成できない場合、サブディスクは1つも作成されず、vxunreloc は何も処理を行いません。サブディスクの移動は、部分的にも処理されます。サブディスクの移動は1つの処理ごとに独立しているため、各サブディスクの移動処理にエラーが発生した時点で vxunreloc ユーティリティはエラーメッセージを出力して終了します。ただし、処理されなかったサブディスクはすべて、ホットリロケーション処理によって移動されたディスク上に残り、戻されることはありません。このとき戻されたサブディスクでは、第3段階の消去処理が実行されていないため、サブディスクレコードのコメントフィールドが「UNRELOC」のまま残ります。

デスティネーションディスクに新しいサブディスクを作成後、サブディスクが戻される前にシステムに障害が発生した場合は、システムの回復後に再度 vxunreloc ユーティリティを実行できます。上記のように、新しいサブディスクが作成されると、vxunreloc はサブディスクのコメントフィールドに「UNRELOC」を設定します。vxunreloc は、再実行時に、デスティネーションディスク上にある既存のサブディスクの offset、len、および comment フィールドをチェックして、前回の vxunreloc 処理時にディスク上に設定された「UNRELOC」を検出し、そのディスクを使用します。

コメントフィールドの文字列「UNRELOC」は手動で変更しないでください。

一連のサブディスクの移動で1つの処理にエラーが発生すると、vxunreloc は処理を中止します。この場合、サブディスクの移動処理が失敗した原因となるエラーを確認し、リロケート処理の解除を続行できるかどうか判断する必要があります。サブディスクの移動処理を再開するために vxunreloc を再起動すると、前回の実行時に作成されたサブディスクが使用されます。

消去段階は、単一のトランザクションで処理されます。vxunreloc は、デスティネーションディスク上で「UNRELOC」が設定されているすべてのサブディスクについて、コメントフィールドを NULL 値でリセットします。消去処理は、以前に完了しなかった vxunreloc 処理でリロケート処理が解除されたサブディスクも対象となります。

障害のあるディスクの検出

注 Volume Manager のホットリロケーション機能は、自動的にディスク障害を検出し、システム管理者に電子メールで障害を通知します。ホットリロケーションが無効になっている場合や電子メールを紛失した場合は、vxprint コマンドの出力またはグラフィカルユーザインタフェースからディスク障害を確認できます。また、ドライバのエラーメッセージをコンソールまたはシステムメッセージファイル内で参照することもできます。

ボリュームにディスクの入出力障害が発生した場合（ディスクに修正不能なエラーが発生した場合など）、Volume Manager は障害に関係するプレックスを切り離すことができます。

プレックスが切り離された場合、そのプレックスでの入出力は停止されますが、ボリュームの残りのプレックスでは継続して実行されます。ディスクが完全に破損した場合、Volume Manager はディスクグループからそのディスクを切り離します。

ディスクが切り離された場合、そのディスク上のすべてのプレックスは無効になります。切り離しが行われるときに、ディスク上にミラーリングされていないボリュームがある場合は、そのボリュームも無効になります。

部分的なディスク障害

障害によりプレックスまたはディスクが切り離されたときにホットリロケーションが有効な場合は、障害の発生したオブジェクトを通知するメールがルートユーザに送信されます。部分的なディスク障害が発生した場合、メールには障害を起したプレックスが示されます。たとえば、ミラーボリュームを含むディスクに障害が発生した場合、次のようなメールが送信されます。

```
To: ルート
Subject: ホスト teal 上の Volume Manager エラー

VERITAS Volume Manager がエラーを検出しました。

failed plexes:
    home-02
    src-02
```

ルートユーザ以外のユーザにメールを送信する方法については、「vxrelocd の変更 (93 ページ)」を参照してください。



上の例のようなメッセージを受信した場合は、次のコマンドを実行して、障害を引き起こしているディスクを特定できます。

```
vxstat -s -ff home-02 src-02
```

通常、次のような結果が表示されます。

```
                                FAILED
TYP NAME                        READS    WRITES
sd disk01-04                     0         0
sd disk01-06                     0         0
sd disk02-03                     1         0
sd disk02-04                     1         0
```

この出力は、障害が disk02 上で発生していることを示しています（サブディスク disk02-03 および disk02-04 が影響を受けています）。

ホットリロケーションは、影響を受けたサブディスクを自動的にリロケートし、必要なリカバリ処理を開始します。ただし、リロケートできない場合やホットリロケーション機能が無効の場合は、問題を特定し、プレックスのリカバリを試みる必要があります。これらのエラーは、ケーブルの障害によって発生する可能性があるため、ディスクとシステムを接続するケーブルをチェックしてください。明らかに問題があれば修正し、次のコマンドを使用してプレックスをリカバリします。

```
# vxrecover -b home src
```

このコマンドを実行すると、障害を起したプレックスのリカバリがバックグラウンドで開始されます（処理が終了する前にコマンドから制御が戻ります）。実行後にエラーメッセージが表示された場合や、プレックスが再び切り離された場合は、ケーブルに明らかな障害がなければ、ディスクを交換します（「ディスクの交換（101 ページ）」を参照）。

完全なディスク障害

ディスクが完全に破損し、ホットリロケーションが有効な場合は、障害が発生したディスクとそのディスク中のプレックスの一覧がメールで通知されます。たとえば、次のようなメールが送信されます。

```
To: ルート
Subject: ホスト teal 上の Volume Manager エラー

VERITAS Volume Manager がエラーを検出しました。

failed disks:

    disk02
```

```

failed plexes:
    home-02
    src-02
    mkting-01

failing disks:
    disk02

```

このメッセージは、障害が発生したために `disk02` が切り離されたことを示しています。ディスクが切り離されると、このディスクへの入出力は実行できなくなります。プレックス `home-02`、`src-02`、および `mkting-01` も同様に切り離されています（ディスク障害が原因と考えられます）。

ここでもケーブル障害が原因として考えられます。ケーブル障害が原因でない場合は、ディスクを交換してください（「ディスクの交換（101 ページ）」を参照）。

ディスクの交換

完全に破損したディスク（障害が発生したために切り離されたディスク）は、`vxdiskadm` を実行するか、メインメニューの項目 5[障害が発生したディスクまたは削除したディスクの交換] を選択して交換できます。初期化されているけれども追加されていないディスクがある場合は、それらのディスクの 1 つを交換用ディスクとして選択できます。

注 選択リストに古いディスクドライブが表示されている場合でも、そのディスクを交換用ディスクとして選択しないでください。適切な初期化済みディスクがない場合は、新しいディスクの初期化を選択できます。

ディスク障害によってボリュームが無効になった場合は、ディスクを交換した後にボリュームをバックアップからリストアする必要があります。ディスク障害が発生したために無効になったディスク上にボリューム全体が配置されているボリュームを識別するには、次のコマンドを使用します。

```
# vxinfo
```

`Unstartable` と表示されているボリュームは、バックアップからリストアする必要があります。次に `vxinfo` の出力画面の例を示します。

```

home          fsgen        Started
mkting        fsgen        Unstartable
src           fsgen        Started
standvol      gen          Started (used on some systems)
rootvol       root         Started
swapvol       swap         Started

```



バックアップからリストアできるように `mkting` を再起動するには、次のコマンドを使用します。

```
# vxvol -o bg -f start mkting
```

`-o bg` オプションの組み合わせにより、バックグラウンドでプレックスが再同期化されます。

障害の発生が認められるが、ディスクが完全に破損していない場合は、ディスクを交換します。次の2つの手順で実行します。

1. ディスクグループからディスクを切り離す。
2. 新しいディスクと交換する。

ディスクを切り離すには、`vxdiskadm` を実行してメインメニューの項目 4[交換用ディスクの削除] を選択します。交換用として使用できる初期化済みディスクがある場合は、そのディスクを処理の一部で指定できます。ここで指定しない場合は、メインメニューの項目 5[障害が発生したディスクまたは削除したディスクの交換] を選択し、後から交換用ディスクを指定する必要があります。

交換のために削除するディスクを選択する場合は、処理で影響を受ける可能性のあるすべてのボリュームが表示されます。次のような結果が表示されます。

この操作を行うと、次のボリュームのミラーが失われます：

```
home src
```

これらのボリュームのデータは失われません。

次のボリュームが使用中です。この操作の結果、これらのボリュームは無効になります。

```
mkting
```

これらのボリュームを使用しているすべてのアプリケーションが、ボリュームにアクセスできなくなります。ボリュームをバックアップからリストアする必要があります。

```
Are you sure you want do this?[y,n,q,?] (デフォルト :n)
```

ボリュームが無効になる可能性がある場合は、`vxdiskadm` を終了してボリュームを保存します。そのボリュームは、バックアップするかディスクから移動してください。ボリューム `mkting` を `disk02` 以外のディスクへ移動するには、次のコマンドを使用します。

```
# vxassist move mkting !disk02
```

ボリュームのバックアップまたは移動後に、もう一度 `vxdiskadm` を実行して、交換するディスクの削除を続行します。

交換するディスクが削除されたら、`vxdiskadm` メインメニューの項目 5[障害が発生したディスクまたは削除したディスクの交換] を選択し、ディスクの交換を実行します。

ディスクグループの作成

通常、ディスクグループは特定のユーザまたはアプリケーション用に作成します。ディスクを Volume Manager でボリューム用に使用するには、それらのディスクをディスクグループに入れておく必要があります。Volume Manager には、必ずデフォルトのディスクグループ rootdg がありますが、必要に応じて別のディスクグループを追加することもできます。

注 特に指定しない限り、すべてのボリュームは rootdg に作成されます。すべてのコマンドで、デフォルトとして rootdg が使用されます。

ディスクグループ newdg を作成するには、次の手順を実行します。

1. 次のコマンドを入力して vxdiskadd を起動します。

```
# vxdiskadd c1t1d0
```

2. 次に示すプロンプトで、Return キーを押して処理を続行します。

ディスクの追加または初期化

```
Menu:VolumeManager/Disk/AddDisks
```

選択したディスクの一覧を以下に表示します。出力形式:[Device_Name]

```
c1t1d0
```

操作を続行しますか?[y,n,q,?] (デフォルト:y)

3. 次に示すプロンプトで、ディスクの追加先ディスクグループを指定します (ここでは newdg)。

ディスクを既存のディスクグループに追加するか、新しいディスクグループに追加するかを選択します。また、後の作業で追加や交換操作を行うためにディスクを確保しておくこともできます。新しいディスクグループを作成するには、使用していないディスクグループ名を指定します。ディスクを後の作業で使用するために確保しておくには、ディスクグループ名に「none」を指定します。

```
ディスクグループの指定 [<group>,none,list,q,?] (デフォルト:rootdg) newdg
```

4. vxdiskadd を実行すると、同じ名前前のアクティブなディスクグループが存在しないことが確認され、この新しいディスクグループを作成するかどうかを確認するメッセージが表示されます。「y」を入力して処理を続行します。

newdg という名前前のアクティブなディスクグループはありません。

```
newdg という名前前で新しいグループを作成しますか?[y,n,q,?] (デフォルト:y) y
```



- 次に示すプロンプトで、**Return** キーを押してデフォルト ディスク名を使用するか、または別のディスク名を入力します。

このディスクにデフォルトのディスク名を使用しますか？[y,n,q,?] (デフォルト :y)

- ディスクをホットリロケーション スペア ディスクにするかどうかを確認するメッセージが表示されたら、「**n**」を入力するか、または **Return** キーを押します。

ディスクを rootdg (または newdg) のスペアディスクとして追加しますか？[y,n,q,?] (デフォルト :n) **n**

ディスクをホットリロケーションの利用対象から除外しない場合は、次のプロンプトで「**n**」を入力します。

他のディスクをホットリロケーションへの利用対象から除外しますか？[y,n,q,?]

[デフォルト :n] **n**

- 処理を続行するには、次に示すプロンプトで「**y**」を入力するか、または **Return** キーを押します。

新しいディスクグループ newdg が作成されます。選択したディスクはデフォルトのディスク名で新しいディスクグループに追加されます。

c1t1d0

操作を続行しますか？[y,n,q,?] (デフォルト :y) **y**

Volume Manager で使用するためにディスクを初期化していることを示す次のようなメッセージが表示されます。

デバイス c1t1d0 を初期化しています。

newdg01 という名前のディスク デバイス c1t1d0 が属する新しいディスク グループ newdg を作成しています。

- ディスクグループが作成されたことを確認するには、次のコマンドを使用します。

```
# vxdisk list
```

上記のコマンドを実行すると、次のような結果が表示されます。

DEVICE	TYPE	DISK	GROUP	STATUS
c0t0d0s2	sliced	disk04	rootdg	online
c1t0d0s2	sliced	disk03	rootdg	online
c1t1d0s2	sliced	newdg01	newdg	online

ディスク グループは、`vx dg init` を使用して作成することもできます。`vx dg` ユーティリティを使用してディスク グループを作成するには、次のコマンドを使用します。

```
# vx dg init ディスク グループ ディスク名 = デバイス名
```

たとえば、デバイス `c1t0d0s2` 上にディスク グループ `mktdg` を作成するには、次のように入力します。

```
# vx dg init mktdg mktdg01=c1t0d0
```

`vx dg` に指定するディスク デバイス名は、`vx diskadd` を使用して初期化済みである必要があります。ディスクはディスク グループに属している必要があります。

ディスク グループのアップグレード

Volume Manager 3.0 より前のリリースでは、ディスク グループがインポートされると、ディスク グループ バージョンは、必要に応じて、自動的にアップグレードされていました。

Volume Manager のディスク グループ アップグレード機能では、ディスク グループのインポートおよびアップグレードが、2つの処理に分割されています。低いレベルのバージョンのディスク グループをインポートし、アップグレードせずに使用することができます。

新しい機能を使用する場合は、ディスク グループをアップグレードできます。アップグレードは明示的に実行します。ディスク グループがアップグレードされると、新しいバージョンをサポートしていない VxVM の以前のリリースとの互換性はなくなります。

インポートされたディスク グループがアップグレードされる以前には、管理者が現在のリソースに明示的にアップグレードするまで、使用されないようにインポートされたディスク グループに対して変更は行われません。

アップグレードが完了するまで、ディスク グループは提供された「そのままの状態」で使用されます。最新バージョンの機能は使用できません。ディスク グループのインポート元のバージョンではなく、最新バージョンの機能を使用しようとすると、次のようなエラー メッセージが表示されます。

```
vxvm:vxedit:ERROR: ディスク グループ バージョンで機能がサポートされていません
```

新しい機能を使用するには、管理者がコマンドを実行して、新機能をサポートするバージョンにディスク グループを明示的にアップグレードする必要があります。

すべてのディスク グループには、それぞれバージョン番号が関連付けられています。Volume Manager のリリースによって、特定のディスク グループ バージョンがサポートされています。Volume Manager では、特定のバージョンのディスク グループに関して、インポートおよび処理の実行が可能です。ディスク グループのバージョンでサポートされている機能および処理だけを実行できます。



表 2 は、VxVM リリースで採用およびサポートするディスク グループ バージョンをまとめたものです。

表 2. ディスク グループ バージョンの割り当て

VxVM リリース	採用するバージョン	サポートするバージョン
1.2	10	10
1.3	15	15
2.0	20	20
2.2	30	30
2.3	40	40
3.0	60	20-60
3.1	70	20-70

注 VxVM 3.0 を除くその他すべての VxVM リリースでは、ディスク グループがインポートされたときに、サポートするディスク グループ バージョンのアップグレードが実行されます。

VxVM 3.0 システム上に以前のバージョンのディスク グループをインポートすると、そのバージョンのリリース以降に採用された機能は使用できません。表 3 は、特定のディスク グループ バージョンでサポートされていない機能をまとめたものです。

表 3. サポートされないディスク グループ機能

ディスク グループ バージョン	サポートされない機能
40	
30	ホットリロケーション
20	VxSmartSync Recovery Accelerator
10、15	RAID-5 ボリューム、新しいスタイルのストライプ、リカバリチェックポイント、ディスク グループ設定コピーの制限、ダーティバージョン ロギング、ミラー ボリューム ロギング

ディスク グループ バージョンを一覧表示するには、ディスク グループ名を指定して次のコマンドを実行します。

```
# vxdbg list ディスク グループ名
```

vxprint (1M) コマンドを -1 フォーマット オプションを指定して実行することによって、ディスク グループ バージョンを特定できます。

VxVMはディスクグループを、現在実行中のVxVMリリースでサポートされている最上位バージョンにアップグレードします。ディスクグループをアップグレードするには、次のコマンドを使用します。

```
# vxpdg upgrade ディスクグループ名
```

デフォルトでは、VxVMはVxVMリリースでサポートされている最上位バージョンのディスクグループを作成します。たとえば、VxVM 3.0ではバージョン60のディスクグループが作成されます。

下位バージョンのディスクグループを作成する必要がある場合もあります。VxVMリリース3.0を実行しているシステムで作成されるディスクグループは、デフォルトではバージョン60になります。このディスクグループは、バージョン40までをサポートするVxVMリリース2.5を実行しているシステムではインポートできません。そのため、VxVMリリース3.0を実行しているシステムで、VxVMリリース2.5を実行しているシステムにインポート可能なディスクグループを作成するには、バージョン50以下のディスクグループを作成する必要があります。

上の例のように、特定のバージョンのディスクグループを作成するには、`vxpdg init` コマンドで `-T` バージョンオプションを指定します。オプションバージョンには、ディスクグループのバージョンを指定します。たとえば、VxVM 2.5が稼働しているシステムでディスクグループを作成するには、次のコマンドを使用します。

```
# vxpdg -T 40 init newpdg newpdg01=c0t3d0s2
```

上記のコマンドを実行すると、バージョン40のディスクグループ `newpdg` が作成されます。このディスクグループは、VxVMリリース2.5にインポートできます。VxVM 2.5システムでディスクグループをインポートできても、VxVM 3.0機能で使用するとエラーになることに注意してください。

ディスクグループの削除

ディスクグループを削除するには、ディスクグループのすべてのボリュームをマウント解除および停止し、次のコマンドを使用します。

```
# vxpdg deport ディスクグループ
```

ディスクグループをデポートしても、ディスクグループが実際に削除されることはありません。ディスクグループがシステムで使用できなくなるだけです。デポートされたディスクグループ上のディスクは、再利用、再初期化、またはほかのディスクグループへ追加できます。



システム間でのディスク グループの移動

ディスク グループの重要な機能として、ディスク グループのシステム間の移動があります。ディスク グループ上のすべてのディスクをあるシステムから別のシステムに移動すると、そのディスク グループを 2 番目のシステムで使用できるようになります。構成の設定を指定し直す必要はありません。

次の手順に従って、システム間でディスク グループを移動します。

1. 1 番目のシステム上で、ディスク グループ上のすべてのボリュームを停止し、ディスク グループをデポート (ディスク グループへのローカル アクセスを無効化) します。次のコマンドを使用します。

```
# vxdg deport ディスクグループ
```

2. すべてのディスクを 2 番目のシステムに移動し、2 番目のシステムと Volume Manager で新しいディスクを認識するために必要な手順を実行します (システムによって異なります)。

システムのリポートが必要になる場合があります。システムをリポートすると、vxconfigd デーモンが再起動し、新しいディスクが認識されます。システムをリポートしない場合は、vxdctl enable コマンドを使用して vxconfigd を再起動すると、Volume Manager でディスクが認識されます。

3. ディスク グループを 2 番目のシステムにインポート (ディスク グループへのローカル アクセスを有効化) します。次のコマンドを使用します。

```
# vxdg import ディスクグループ
```

4. ディスク グループのインポート後、ディスク上のすべてのボリュームを起動します。次のコマンドを使用します。

```
# vxrecover -g ディスクグループ -sb
```

クラッシュしたシステムからディスクを移動できます。この場合、1 番目のシステムからディスク グループをデポートすることはできません。ディスク グループがシステム上に作成またはインポートされた場合、そのシステムはディスク グループのすべてのディスク上にロックを書き込みます。

注 ロックの目的は、デュアルポート ディスク (2 つのシステムから同時にアクセスできるディスク) が同時に両方のシステムから使用されないようにすることです。2 つのシステムが同時に同じディスクを制御しようとする、ディスクに格納されている設定情報が壊れ、ディスクおよびそのデータは使用できなくなります。

クラッシュしたシステムまたはグループを検出できなかったシステムからディスクを移動すると、ディスクに記録されたロックはそのまま残るので、解除する必要があります。システムは次のようなエラー メッセージを返します。

```
vxdg: ディスク グループ グループ名: インポートできません: ディスクは別のホストが使用中です。
```



特定のデバイスセット上のロックを解除するには、次のコマンドを使用します。

```
# vxdisk clearimport デバイス名 ...
```

次のコマンドを使用すると、インポート中にロックを解除できます。

```
# vxdg -C import ディスクグループ
```

注 vxdisk clearimport または vxdg -C import コマンドをデュアルポート ディスクのあるシステム上で使用するときには注意が必要です。ロックを解除すると、これらのディスクに複数のホストから同時にアクセスできるようになり、データが破損する可能性があります。

いくつかのディスクが使用できず、ディスク グループのインポートが必要な場合があります。ディスク グループのいくつかのディスクがシステムに接続されているディスク ドライブ上で検出できない場合、通常 import 処理はエラーとなります。import 処理がエラーとなると、次のようなエラー メッセージが表示されます。

```
vxdg: ディスク グループグループ名: インポートできません: ディスクグループに有効な設定コピーがありません。
```

このメッセージは、ハードウェアの修理や新しいディスク グループの作成を必要とする致命的なエラーを示しています。

```
vxdg: ディスク グループ グループ名: インポートできません: ディスク グループのディスクが見つかりません。
```

このメッセージは、リカバリ可能なエラーを示しています。

ディスク グループのディスクにエラーが発生した場合は、次のコマンドを使用して強制的にディスク グループをインポートできます。

```
# vxdg -f import ディスクグループ
```

注 -f オプションを使用するときには注意が必要です。同じディスク グループが別のディスク セットから 2 回インポートされると、ディスク グループに矛盾が生じる可能性があります。

これらの処理は vxdiskadm を使用して実行できます。vxdiskadm を使用してディスク グループをデポートするには、メニュー項目 9[ディスク グループへのアクセスの無効化 (デポート)] を選択します。ディスク グループをインポートするには、メニュー項目 8[ディスク グループへのアクセスの有効化 (インポート)] を選択します。vxdiskadm のインポート処理では、ホストのインポート ロックが確認され、ロックが検出された場合は、解除するかどうかを確認するプロンプトが表示されます。また、ディスク グループ内のボリュームも起動します。



ディスク グループ名の変更

1つのシステム内で複数のディスク グループに同じ名前を指定することはできません。同じ名前のディスク グループが移動先のシステムに存在する場合は、そのディスク グループをインポートまたはデポートすることはできません。この問題を避けるため、Volume Manager ではインポートまたはデポート処理中にディスク グループ名を変更できます。

たとえば、Volume Manager を実行しているすべてのシステムにはデフォルトの rootdg ディスク グループが1つあるため、システム間で rootdg をインポートまたはデポートする場合は問題です。同じシステム上に2つの rootdg ディスク グループは存在できません。この問題は、インポートまたはデポート処理中に rootdg ディスク グループの名前を変更することで回避できます。

インポート中にディスク グループに新しい名前を付けるには、次のコマンドを使用します。

```
# vxdbg [-t] -n 新しいディスク グループ名 import ディスク グループ
```

-t オプションを指定すると、インポートは一時的なものとなり、リブート後は元の状態に戻ります。この場合、保存されたディスク グループの名前は元のホスト通りに変更されない状態で残りますが、そのディスク グループはインポート先のホストには新しいディスク グループ名として認識されます。-t オプションを指定しないと、名前の変更は確定されます。

デポート中にディスク グループ名を変更するには、次のコマンドを使用します。

```
# vxdbg [-h hostname] -n 新しいディスク グループ名 deport ディスク グループ
```

デポート中に名前を変更する場合は、-h ホスト名オプションを指定すると、代替ホストへロックを割り当てることができます。これによって、代替ホストのリブート時に、ディスク グループは自動的にインポートされます。

次の手順に従って、rootdg ディスク グループをあるホストから別のホストへ一時的に移動し（ルート ボリューム上で修復作業を行う目的など）、次にそのディスク グループを元に戻すことができます。

1. 元のホスト上で、rootdg ディスク グループのディスク グループ ID を識別し、次のコマンドを使用して別のホストへインポートします。

```
# vxdisk -s list
```

ディスク グループ情報などが次のように表示されます。

```
dgname:rootdg
dgid:774226267.1025.tweety
```

2. インポート先ホストで、次のコマンドを使用して、ディスク グループ `rootdg` をインポートし、名前を変更します。

```
# vxpdg -tc -n 新しいディスク グループ名 import ディスク グループ
```

`-t` は一時的なインポート名を示します。`-c` はインポート ロックの解除を示します。`-n` はインポートする `rootdg` の一時的な名前を指定します(一時的に名前を付けることによって既存の `rootdg` との競合を回避します)。ディスク グループには、インポートするディスク グループのディスク グループ ID (774226267.1025.tweety など) を指定します。

この時点でリブートやクラッシュが発生した場合、一時的にインポートされたディスク グループはインポートされないため、再度インポートする必要があります。

3. インポートされた `rootdg` で必要な処理を実行した後は、次のコマンドを使用して、`rootdg` をデポートして元のホストに戻します。

```
# vxpdg -h ホスト名 deport ディスク グループ
```

ホスト名には、`rootdg` が戻されるシステムの名前を指定します(システム名は `uname -n` コマンドで確認できます)。

このコマンドを実行すると、インポート ホスト側のインポートされた `rootdg` が削除され、元のホストにロックが戻されます。次のリブート時に、元のホストは `rootdg` を自動的にインポートします。

ディスク グループ用のマイナー番号の予約

Volume Manager では、指定したディスク グループのマイナー番号の範囲を選択できます。ボリュームの作成中にこの番号範囲を使用します。これによって、リブートや再設定を行った後でも、各ボリュームのマイナー番号は同一に保たれます。2つのディスク グループの番号の範囲が重複している場合は、インポート時に競合が検出されるため、インポートを中止するか、または番号の範囲を指定し直します。

別の範囲にあるボリューム デバイス番号を各ディスク グループに割り当てると、マシンのグループ内にあるすべてのディスク グループを移動しても、デバイス番号の競合は発生しません。

ディスク グループに対するボリューム デバイスの基本マイナー番号を設定するには、次のコマンドを使用します。

```
# vxpdg init ディスク グループ minor= 基本マイナー番号 デバイス名
```

ディスク グループの各ボリューム デバイス番号のマイナー番号は、基本マイナー番号に指定された番号から始まるように設定されます。ほとんどのシステム上で、マイナー番号は 131071 までの範囲になります。一時的なデバイス番号の再割当用に、適当なサイズの番号範囲を最後の方に確保しておくことができます。



`vxvg init` コマンドラインに `minor` オペランドを指定しない場合、Volume Manager はランダムな番号を選択します。番号は 1000 または、1000 の倍数であり、1000 個のデバイス番号を取ることができます。選択されたデフォルト番号は、現在インポートされているディスクグループの 1000 個分の範囲と重複しません。また、現在割り当てられているボリューム デバイス番号とも重複しません。

注 デフォルトのポリシーを持っていると、少数のディスクグループはマシン間で正常にマージできます。ただし、フェイルオーバー機構を使用してディスクグループを自動的にマージする場合、重複を避けて範囲を選択する必要があります。

マイナー番号の予約の詳細については、`vxvg (1M)` マニュアル ページを参照してください。

ディスクグループの破棄

`vxvg` コマンドには破棄オプションが用意されています。このオプションを使用して、システムからディスクグループを削除し、そのディスクグループ内のディスクをほかのディスクグループで使用するために解放します。不要なディスクグループを削除して、ディスクをほかのディスクグループから使用できるようにするには、次のコマンドを実行します。

```
# vxvg destroy ディスクグループ
```

`vxvg deport` コマンドを使用して、ディスクにアクセスできないようにすることもできます。Volume Manager では、デポートしたディスクグループのディスクをほかのディスクグループで使用できないようにすることができます。

特殊デバイスの使用

ここでは、Volume Manager が管理タスクを実行するために使用する特殊デバイスについて説明します。

特殊なカプセル化用の `vxdisk` の使用

カプセル化とは、指定されたディスク上の既存のパーティションをボリュームに変換する処理のことです。ファイルシステムを含むパーティションの場合は、ファイルシステムがボリューム上にマウントされるように `/etc/vfstab` エントリが修正されます。

ディスクのカプセル化を行うには、Volume Manager の識別情報と設定情報を格納するためにディスク上に空き領域が必要です。この空き領域はほかのパーティションに含めることができません。詳細については、『*VERITAS Volume Manager Installation Guide*』および `vxencap (1M)` マニュアル ページを参照してください。

Volume Manager の専有領域パーティションに使用できる領域がないディスクをカプセル化できます。`vxdisk` ユーティリティは、使用可能な領域のないディスクをカプセル化します。この処理は、特殊ディスクデバイスである `nopriv` デバイスを使用して実行します。このデバイスには専有領域がありません。

`vxdisk` を使用するには、アクセスするディスクのすべてに対応するパーティションを作成します。次に、以下のコマンドを実行して、そのパーティションのパーティション デバイスを追加します。

```
# vxdisk define パーティション - デバイス type=nopriv
```

パーティションデバイスには、`/dev/dsk` ディレクトリの基本名を指定します。たとえば、ディスク デバイス `c0t4d0` のパーティション 3 を使用する場合は、次のように入力します。

```
# vxdisk define c0t4d0s3 type=nopriv
```

このディスク ドライブ上のほかのパーティション用のボリュームを次の方法で作成します。

- ◆ デバイスをディスク グループへ追加する。
- ◆ それらのパーティションのカプセル化パーティション内での位置を決定する。
- ◆ `vxassist` を使用して、該当のオフセットおよびサイズを指定してボリュームを作成する。

`vxassist` は、デフォルトで、作成したボリュームのデータ領域を再初期化します。パーティション上に維持すべきデータがある場合は、`vxassist` を使用しないでください。`vxmake` を使用してボリュームを作成し、`vxvol init active` を使用してそのボリュームを起動します。

`nopriv` デバイスを使用すると、**Volume Manager** によってディスクのアドレスやコントローラに加えられた変更を追跡できません。通常、**Volume Manager** は物理ディスク上の専有領域に格納されている識別情報を使用して、物理ディスクの位置の変更を追跡します。`nopriv` デバイスには専有領域および物理ディスクに格納されている識別情報がないため、追跡できません。

特殊なカプセル化パーティション デバイスを最大限に活用するには、ディスクをカプセル化する際に、**Volume Manager** で領域をディスクから移動できるようにします。ディスク上の領域が使用可能になると、特殊なパーティション デバイスを削除して、ディスクを標準ディスク デバイスとしてカプセル化できます。

ディスク グループを、`nopriv` デバイスで完全に形成することはできません。これは、`nopriv` デバイスにはディスク グループの設定情報を格納するための領域がないためです。設定情報は、ディスク グループ内の 1 つ以上のディスクに格納しておく必要があります。



RAM ディスクでの vxdisk の使用

注 この節の内容は、RAM ディスクを搭載したシステムにのみ適用されます。

システムの中には、RAM ディスクの作成をサポートしているものがあります。RAM ディスクは、システム RAM で作成されたデバイスであり、小さなディスク デバイスのようなものです。多くの場合、システムのリブート時に RAM ディスクの内容は消去されます。リブートにより RAM ディスクが消去されると、Volume Manager で物理ディスクを識別できなくなります。これは、物理ディスク上に格納された情報（リブート時に消去された情報）がディスクを識別するために必要なためです。

nopriv デバイスには、RAM ディスクをサポートするための特殊な機能である `volatile` オプションがあります。このオプションは Volume Manager にそのデバイスの内容がリブート時に失われることを示します。揮発性デバイスは、システムの起動時に特別に処理されます。ボリュームがミラーリングされている場合、不揮発性プレックスからデータをコピーすることによって、揮発性デバイスで構成されるプレックスは常にリカバリされます。

RAM ディスクを使用するには、ディレクトリ `/dev/dsk` および `/dev/rdsk` にディスク用のデバイス ノード (`/dev/dsk/ramd0`、`/dev/rdsk/ramd0` など) を作成する必要があります。RAM ディスク デバイスを Volume Manager に定義するには、次のコマンドを使用します。

```
# vxdisk define ramd0 type=nopriv volatile
```

通常、Volume Manager は揮発性サブディスクを使用したプレックスからすべてが構成されているボリュームを起動しません。これは、プレックスに最新のボリューム内容が含まれている保証がないためです。

いくつかの RAM ディスクは、リブート後にすべてのボリューム内容が再作成されてから使用されます。このような状況下では、RAM で構成されたボリュームをシステムのリブート時に強制的に起動することができます。次のコマンドを使用します。

```
# vxvol set startopts=norecov ボリューム名
```

このオプションは、gen タイプのボリュームでのみ使用できます。vxvol set の操作および norecov オプションの詳細については、vxvol (1M) を参照してください。

vxdisk を使用したマルチパス情報の表示

注 この節の内容は、DMP (Dynamic Multipathing: 動的マルチパス) 機能を持つシステムにのみ適用されます。

Volume Manager では、システムに接続されている物理ディスクは、1 つまたは複数の物理アクセスパスを持つメタデバイスとして表されます。このアクセスパスは、ディスクが単一ディスクであるか、システムに接続されているマルチポートに対応したディスクアレイの一部であるかによって異なります。vxdisk ユーティリティを使用すると、メタデバイスのパスを表示し、各パスの状態 (有効または無効など) を表示することができます。たとえば、ディスク disk01 の詳細を表示するには、次のように入力します。

```
# vxdisk list disk01
```

次のような結果が表示されます。

```
Device      c2t0d0s2
devicetag   c2t0d0
type        sliced
hostid      aparajita
disk        name=disk01 id=861086917.1052.aparajita
group       name=rootdg id=861086912.1025.aparajita
flags       online ready autoconfig autoimport imported
pubpaths    block=/dev/vx/dmp/c2t0d0s4 char=/dev/vx/rdmp/c2t0d0s4
privpaths   block=/dev/vx/dmp/c2t0d0s3 char=/dev/vx/rdmp/c2t0d0s3
version     2.1
iosize      min=512 (bytes) max=2048 (blocks)
public      slice=4 offset=0 len=1043840
private     slice=3 offset=1 len=1119
update      time=861801175 seqno=0.48
headers     0 248
configs     count=1 len=795
logs        count=1 len=120
Defined regions
config      priv 000017-000247[000231]:copy=01 offset=000000 enabled
config      priv 000249-000812[000564]:copy=01 offset=000231 enabled
log         priv 000813-000932[000120]:copy=01 offset=000000 enabled

Multipathing information:
numpaths:   2
c2t0d0s2    active
c1t0d0s2    failed
```

この出力には、メタデバイス c2t0d0s2 によって表された物理デバイスにパスが 2 つあることが示されています。パス c2t0d0s2 はアクティブ状態で、パス c1t0d0s2 は非アクティブ状態になっています。



vxdiskadm メニュー インタフェース処理

ここでは、Volume Manager Support Operation (vxdiskadm) と呼ばれるメニューから起動するインタフェースについて説明します。vxdiskadm は物理デバイスおよび論理デバイスの管理に使用します。

ここでは、vxdiskadm を使用したディスクおよびボリューム関連タスクの実行に関する次の情報について説明します。

- ◆ Volume Manager 管理下への物理ディスクの配置
- ◆ 今後の使用に備えたディスクの追加
- ◆ Volume Manager で使用するためのディスクの再初期化
- ◆ VM ディスクのホットリロケーション プールへの追加
- ◆ VM ディスクのホットリロケーション プールからの削除
- ◆ 物理ディスク上の情報の表示
- ◆ 物理ディスクの削除
- ◆ 物理ディスクの無効化 (物理ディスクのオフライン化)
- ◆ 物理ディスクの有効化
- ◆ 物理ディスクの交換
- ◆ 障害ディスクまたは削除済みディスクの交換
- ◆ ディスク グループへのディスクの追加
- ◆ ディスク グループの作成
- ◆ ディスク グループのデポート
- ◆ ディスク グループのインポート

vxdiskadm の起動

vxdiskadm を起動するには、次のコマンドを使用します。

```
# vxdiskadm
```

vxdiskadm メイン メニューが表示されます。

vxdiskadm メイン メニュー

vxdiskadm メイン メニューの表示内容は以下の通りです。

Volume Manager Support Operations

Menu:VolumeManager/Disk

- 1 1 つ以上のディスクの追加または初期化
- 2 1 つ以上のディスクのカプセル化
- 3 ディスクの削除
- 4 交換用ディスクの削除
- 5 障害が発生したディスクまたは削除したディスクの交換
- 6 ディスク上でのボリュームのミラー
- 7 ディスクからの複数ボリュームの移動
- 8 ディスク グループへのアクセスの有効化 (インポート)
- 9 ディスク グループへのアクセスの無効化 (デポート)
- 10 ディスク デバイスの有効化 (オンライン化)
- 11 ディスク デバイスの無効化 (オフライン化)
- 12 ディスク グループ用のスペア ディスクの設定
- 13 ディスク上のスペア フラグをオフにする
- 14 サブディスクのリロケート処理を解除してディスクへ回送
- 15 ディスクのホットリロケーション利用対象からの除外
- 16 ディスクのホットリロケーション設定

list ディスク情報の一覧表示

- ? メニューについてのヘルプを表示します。
- ?? メニュー システムについてのヘルプを表示します。
- q メニューの終了

実行する操作を選択してください :

- ◆ メニューの使用中はいつでも「?」を入力してヘルプを表示できます。「?」を入力すると、処理およびその定義が一覧表示されます。
- ◆ 「??」を入力すると、すべてのプロンプトに対して入力可能な項目が一覧表示されます。
- ◆ 操作の再起動が必要な場合に「q」を入力するとメイン メニューに戻ります。メイン メニューレベルで「q」を入力すると、Volume Manager Support Operation が終了します。



vxdiskadm メニューを使用すると、以下のタスクにアクセスできます。番号はメインメニューで一覧表示されている項目の番号を示します。

1. 1つ以上のディスクの追加または初期化

システムにフォーマット済みのディスクを追加できます。SCSI ディスクは既にフォーマットされています。ほかのディスクのフォーマット方法については、付属のマニュアルを参照してください。ディスク デバイスの指定を促すプロンプトが表示されます。ディスクを追加するディスク グループを指定できます。ここで「none」を選択すると、ディスクは将来の操作やディスクの交換用にスペアとして保持され、この時点での初期化は不要です。また、選択したディスクをディスク グループ用のホットリローテーション スペアとして指定することもできます。ディスクがまだ初期化されていない場合は、Volume Manager で使用するためにパーティション分割および初期化されます。

2. 1つ以上のディスクのカプセル化

Volume Manager をインストールする前に、システムに追加されているディスクを Volume Manager の管理下に置くことができます。1つ以上のディスク デバイス、ディスク グループ、または 1 つ以上のディスク名の指定を促すプロンプトが表示されます。選択したディスクは指定したディスク グループに追加されます。ディスクは検査され、ファイル システムやほかの目的で使用されているパーティションが検索されます。既存のデータにアクセスする手段として、ボリュームが作成されディスク パーティションと置き換えられます。カプセル化によってパーティションの用途が自動的に決定できない場合は、用途を指定する必要があります。用途として、パーティションのボリュームとの置き換え、パーティションの単独使用、またはパーティションの削除を選択できます。

パーティションがマウント済みのファイル システムまたはアプリケーションの起動に使用されている場合は、システムをリブートする必要があります。ディスク デバイスを直接使用するのではなく、ボリュームを使用してディスク パーティションにアクセスする場合は、アプリケーションの設定ファイルを変更する必要がある場合もあります。ファイル システムのマウント情報は自動的に調整されます。

3. ディスクの削除

ディスク グループからディスクを削除できます。削除するディスク名の指定を促すプロンプトが表示されます。そのディスク上のストレージを使用するボリュームがある場合、そのディスクは削除できません。ディスク上のストレージをボリュームが使用している場合は、そのストレージをディスク グループ内のほかのディスクに移動するように指定することもできます。

注 このタスクを使用して、ディスク グループ内の最後のディスクを削除することはできません。ディスク グループ内の残りのディスクをすべて使用するには、ディスク グループを無効にする（デポート）必要があります。これで、ディスクを再利用できます。ただし、rootdg はデポートできません。

4. 交換用ディスクの削除

ディスク名は保持したまま、物理ディスクをディスクグループから削除できます。このタスクを実行すると、指定した名前付きディスクの状態は `removed` に変更されます。ディスクグループに属さない未初期化ディスクがある場合は、それらのディスクのいずれかを交換ディスクとして使用することもできます。

5. 障害が発生したディスクまたは削除したディスクの交換

[交換用ディスクの削除] メニュー項目を使用して、削除したディスク、または使用中に障害が発生したディスク用の交換ディスクを指定できます。交換するディスクの名前および交換ディスクとして使用するディスク デバイスの指定を促すプロンプトが表示されます。未初期化ディスクを選択して操作中にディスクを初期化するか、または [ディスクの追加または初期化] メニュー操作を実行して既に初期化済みのディスクを選択することができます。

6. ディスク上でのボリュームのミラー化

ボリュームをディスク上でミラー化できます。ボリュームは、使用可能な領域を持つ別のディスクにミラー化することもできます。この方法でボリュームのミラーコピーを作成すると、ディスク障害時のデータ損失を防ぐことができます。ミラー化されたボリューム、または複数のサブディスクで構成されているボリュームは、この方法ではミラー化できません。ブートディスクのボリュームをミラー化すると、代替ブートディスクとして使用可能なディスクが作成されます。

7. ディスクからの複数ボリュームの移動

1 つのディスクを使用している任意のボリューム（またはボリュームの一部）をほかのディスクに移動できます。このメニュー項目は、ディスクを削除する（完全な削除またはディスク交換のための削除）直前に使用します。

注 ディスクからボリュームを移動するだけでそのディスクを削除しない場合は、将来、そのディスク上にほかのボリュームが移動してくることがあります。

8. ディスクグループへのアクセスの有効化（インポート）

システムからディスクグループへのアクセスを有効にすることができます。ディスクグループをあるシステムから別のシステムへ移動する場合は、初めに元のシステム上でディスクグループを無効化（デポート）します。次に、デポートされたディスクグループからほかのシステムにディスクを移動し、移動先でそのディスクグループを有効化（インポート）します。ディスクグループ名の指定を促すプロンプトが表示されます。

9. ディスクグループへのアクセスの無効化（デポート）

システムによって現在有効化（インポート）されているディスクグループへのアクセスを無効にすることができます。ディスクグループ内のディスクを別のシステムに移動する場合は、ディスクグループをデポートします。また、別の目的でディスクグループ内の残りのディスクをすべて使用する場合も、ディスクグループをデポートしてください。



ディスク グループ名の指定を促すプロンプトが表示されます。ディスクの無効化 (オフライン化) を確認するメッセージが表示されます。システム上のリムーバブル デバイスの場合は、ディスクを削除する前にディスクへのすべてのアクセスを無効にする必要があります。

10. ディスク デバイスの有効化 (オンライン化)

通常システム稼働中に、あるシステムから別のシステムへディスクを移動する場合、Volume Manager では移動したディスクは自動認識されません。このメニュー項目を使用して、ディスクを検索して識別し、そのディスクがディスク グループに属しているかどうかを判断するように Volume Manager を設定します。また、このタスクを実行すると、ディスク グループのデポート タスク、またはディスク デバイスの無効化 (オフライン) タスクによって無効化されていたディスクへのアクセスが再び有効化されます。

11. ディスク デバイスの無効化 (オフライン化)

Volume Manager を使用してディスク デバイスへのすべてのアクセスを無効にすることができます。このタスクは、現在ディスク グループに属していないディスクにのみ適用可能です。リブートせずにシステムからディスクを削除するには、このタスクを実行します。

システムによっては、稼働中はシステムからディスクを削除することができません。そのようなシステムでは、このオフライン操作の効果は望まれません。

12. ディスク グループ用のスペア ディスクの設定

ディスク グループ内のディスクに障害が発生した場合に、別のディスクを (ホットリロケーション用の) 交換ディスクとして自動的に使用するために確保しておくことができます。

13. ディスク上のスペア フラグをオフにする

ディスクのスペア フラグをオフにするのに使用することができます。

14. サブディスクのリロケート処理を解除してディスクへ回送

VxVM のホットリロケーションを使用すると、システムは冗長性のある VxVM オブジェクト上の入出力障害にサブディスク レベルで自動対応し、オブジェクトを使用可能な状態にするための処理を実行します。この機構によって、サブディスク上の入出力障害の検出、サブディスクのリロケート、サブディスクに関連付けられたプレックスのリカバリなどが行われます。ディスクの交換後は、ディスク障害発生前と同じ設定にシステムをリストアするための vxunreloc ユーティリティを使用できます。vxunreloc は、ホットリロケート済みのサブディスクを、ディスク障害によって交換されたディスクに戻します。

15. ディスクのホットリロケーション利用対象からの除外

空きプール内のディスク (非スペア) をホットリロケーションの利用対象から外します。

16. ディスクのホットリロケーション設定

手順 15 の設定を元に戻し、空きプール内のディスク（非スペア）をホットリロケーションの利用対象にします。

ディスク管理を行う場合は、デバイス名とディスク名の違いを認識しておくことが重要です。

注 システムによっては、例とは異なるデバイス名を使用している場合があります。デバイス名の詳細については、「Volume Manager とは(1 ページ)」を参照してください。

デバイス名（`devname` またはディスク アクセス名としても参照される）はディスクの場所を示します。デバイス名の基本形式は、`c#b#t#d#s#` です。

- ◆ `c#` は、ディスク デバイスが接続されているコントローラの番号です。
- ◆ `b#` は、対応するバス番号（システムで使用している場合）です。
- ◆ `t#` は、コントローラ上のターゲット ディスク番号です。
- ◆ `d#` は、ディスク番号です。
- ◆ `s#` は、ディスクスライス番号です。

デバイスのフルパスは、`/dev/vx/dmp/ デバイス名` となります。このマニュアルではデバイス名のみを表記し、デバイスは常に `/dev/vx/dmp` ディレクトリにあると想定します。たとえば、`c0t0d0s2` のようにデバイス名を表記します。

ディスク名（またはディスク メディア名）はディスクの管理上の名前です（`disk01` など）。ディスク名を割り当てていない場合、ディスクが `rootdg` に追加されると、ディスク名として `disk##` がデフォルトで設定されます（`##` は連番）。それ以外の場合、デフォルト ディスク名はグループ名 `##`（グループ名はディスクの追加先ディスク グループの名前）になります。

Volume Manager 管理下への物理ディスクの配置

Volume Manager を起動しているシステムにディスクを追加する場合は、そのディスクを Volume Manager の管理下に置き、Volume Manager でディスク上の領域割り当てを制御できるようにする必要があります。

別のディスク グループを指定しない限り、Volume Manager は新しいディスクをデフォルトのディスク グループ `rootdg` に配置します。追加ディスク グループの作成については、本章の後半で説明します。



ディスクを Volume Manager の管理下に置く方法は、次のように環境によって異なります。

- ◆ ディスクが新しい場合は、初期化してから Volume Manager の管理下に置く必要があります（「Volume Manager 管理下へのディスクの配置（123 ページ）」を参照）。
- ◆ ディスクがすぐに必要でない場合は、初期化し（ディスク グループには追加しない）、今後の使用に備えて確保しておくことができます（「今後の使用に備えたディスクの追加（134 ページ）」を参照）。
- ◆ 今後の Volume Manager での使用に備えて、ディスクが既に初期化されている場合は、そのディスクを再初期化して Volume Manager の管理下に置くことができます（「Volume Manager で使用するためのディスクの再初期化（134 ページ）」を参照）。
- ◆ Volume Manager の管理下でないディスクが使用されていた場合は、ディスク上の既存データを保持したまま、このディスクを Volume Manager の管理下に配置することもできます。この処理は、カプセル化によって実現できます（「Volume Manager で使用するディスクのカプセル化（131 ページ）」を参照）。
- ◆ 1つ以上のコントローラ上の複数のディスクを同時に Volume Manager の管理下に配置できます。環境によっては、すべてのディスクが同じ方法で処理されない場合もあります（「Volume Manager の管理下への複数ディスクの配置（126 ページ）」を参照）。

複数のディスクを同時に初期化またはカプセル化する場合は、特定のディスクまたはコントローラのみを対象外に指定することができます。ディスクを対象から外すには、初期化またはカプセル化を実行する前に、`/etc/vx ディレクトリ`にある `disks.exclude` ファイルに除外するディスク名を記述します。同様に、`/etc/vx ディレクトリ`にある `cntrls.exclude` ファイルにコントローラ名を記述すると、特定のコントローラ上のすべてのディスクを初期化またはカプセル化から除外することができます。

以下に、さまざまな方法および環境下で、Volume Manager の管理下にディスクを置くために `vxdiskadm` を使用する方法について詳しく説明します。

-
- 注** Volume Manager の管理下にディスクを置く前に、そのディスクをフォーマットする（`format` コマンドを使用）か、システムに追加する（`diskadd` を使用）必要があります。
- フォーマットしていないディスクを `vxdiskadm` を使用して Volume Manager の管理下に置こうとすると、初期化は通常どおり開始されますが、ディスクが不正でフォーマットできないことを示すメッセージが表示され、処理が中止されます。この場合は、ディスクを正しくフォーマットしてから、再度 Volume Manager の管理下に置く必要があります。
-

Volume Manager 管理下へのディスクの配置

ここでは、フォーマット済みディスクを Volume Manager の管理下に置く方法について説明します。新しいディスクや、以前に Volume Manager 以外で使用されていたディスクも使用できます。

注 初期化では、ディスク上の既存データは保持されません。

Volume Manager で使用するために 1 つのディスクを初期化するには、次の処理を実行します。

1. vxdiskadm メイン メニューの項目 1[1 つ以上のディスクの追加または初期化]を選択します。
2. 次に示すプロンプトで、Volume Manager の管理下に追加するディスクのディスク デバイス名を入力します（ディスク一覧を表示するには「**list**」と入力します）。

ディスクの追加または初期化

Menu:VolumeManager/Disk/AddDisks

この操作を実行すると、ディスクをディスクグループへ追加します。選択済みのディスクを既存のディスクグループへ追加することや、操作の他の部分で作成した新規のディスクグループへ追加することができます。また選択済みのディスクをスペアとしてディスク グループに追加することもできます。さらに、ディスクを初期化し、ディスク グループに追加しないで交換ディスク用に確保しておくこともできます。

プロンプトでは、複数のディスクまたはパターンを入力できます。以下にディスクの選択例を示します：

```
all:                すべてのディスク
c3 c4t2:           コントローラ 3、およびコントローラ 4 ターゲット 2 上の
                  すべてのディスク
c3t4d0:           単一のディスク
```

追加するディスクデバイスを選択してください：

```
[<pattern-list>,all,list,q,?] list
```

<pattern-list> には、1 つのディスク、または一連のディスクやコントローラ（オプションのターゲット）を指定できます。<pattern-list> に複数の項目を指定する場合は、各項目を空白で区切って入力してください。



プロンプトで「**list**」を入力すると、次のようなシステムで使用可能なディスクの一覧が表示され、続いて、追加するディスクのデバイス名の入力を促すプロンプトが表示されます。

DEVICE	DISK	GROUP	STATUS
c0t0d0	disk01	rootdg	online
c0t1d0	disk02	rootdg	online
c0t2d0	disk03	rootdg	online
c0t3d0	-	-	online
c1t0d0	disk10	rootdg	online
c1t0d1	-	-	error
.			
.			
.			
c3t0d0	-	-	error
c3t1d0	disk33	rootdg	online
c3t2d0	disk34	rootdg	online
c3t3d0	disk35	rootdg	online

追加するディスクデバイスを選択してください：

```
[<pattern-list>,all,list,q,?] c1t0d1
```

システムに接続されているすべてのディスクが **Volume Manager** で認識され、ここに表示されます。

STATUS 行に「error」と表示されている場合は、そのディスクがまだ **Volume Manager** の管理下に追加されていないことを示します。これらのディスクは、初期化済みの場合も初期化されていない場合もあります。ディスク名とディスクグループが一覧に表示されているディスクは、既に **Volume Manager** の管理下に置かれているため、このタスクでは使用できません。

3. 処理を続行するには、次に示すプロンプトで、「**y**」を入力するか、または **Return** キーを押します。

選択したディスクの一覧を以下に表示します。出力形式：[Device_Name]

```
c1t2d0
```

操作を続行しますか？[y,n,q,?] (デフォルト：y) **y**

4. 次に示すプロンプトで、ディスクの追加先ディスク グループを指定するか、または Return キーを押してデフォルトの rootdg を使用します。

ディスクを既存のディスク グループに追加するか、新しいディスク グループに追加するかを選択します。また、後の作業で追加や交換操作を行うためにディスクを確保しておくこともできます。新しいディスク グループを作成するには、使用していないディスク グループ名を指定します。ディスクを後の作業で使用するために確保しておくには、ディスク グループ名に「none」を指定します。

ディスク グループの指定 [`<group>`,none,list,q,?] (デフォルト :rootdg)

5. 次に示すプロンプトで、Return キーを押してデフォルト ディスク名を使用するか、または別のディスク名を入力します。

このディスクにデフォルトのディスク名を使用しますか?[y,n,q,?] (デフォルト :y)

6. ディスクをホットリロケーション スペア ディスクにするかどうかを確認するメッセージが表示されたら、「n」を入力するか、または Return キーを押します。

ディスクを rootdg のスペア ディスクとして追加しますか?[y,n,q,?] (デフォルト :n)n

7. ディスクをホットリロケーションの利用対象から除外するかどうかを確認するメッセージが表示されたら、「n」を入力するか、または Return キーを押します。

ディスクをホットリロケーションへの利用対象から除外しますか?[y,n,q,?] (デフォルト :n)

8. 処理を続行するには、次に示すプロンプトで、「y」を入力するか、または Return キーを押します。

選択したディスクは、スペアとしてデフォルトのディスク名でディスク グループ rootdg に追加されます。

c1t2d0

操作を続行しますか?[y,n,q,?] (デフォルト :y) y

9. ディスク上にデータがないことが確実な場合は、「n」を入力してカプセル化しないようにします。ディスクを初期化するかどうかを確認する次のようなメッセージが表示されたら、「y」を入力します。

次のディスク デバイスには有効な VTOC がありますが、Volume Manager 用に初期化されていません。ディスク上に重要なデータがある場合は、ディスクを新しいディスクとして追加するのではなく、既存のディスク パーティションをボリュームとしてカプセル化します。

出力形式 : [Device_Name]

c1t2d0



Encapsulate this device?[y,n,q,?] (デフォルト:y) **n**

c1t2d0

Instead of encapsulating, initialize?[y,n,q,?] (デフォルト:n) **y**

ディスク c1t2d0 が Volume Manager の管理下に置かれたことを示す次のようなメッセージが表示されます。システムによっては、表面分析を実行するかどうかを選択するオプションが表示される場合もあります。

デバイス c1t2d0 を初期化しています。

表面分析の実行 (推奨)

[y,n,q,?] (デフォルト:y) **n**

ディスク デバイス c1t2d0 を、disk39 というディスク名でディスク グループ rootdg に追加します。

10. 次に示すプロンプトで、ほかのディスクの初期化も続行するか (**y**)、vxdiskadm メインメニューに戻るか (**n**) を指定します。

他のディスクを追加または初期化しますか?[y,n,q,?] (デフォルト:n)

Volume Manager の管理下への複数ディスクの配置

ここでは、Volume Manager の管理下に複数のディスクを同時に配置する方法について説明します。ディスク セットには、システム上のすべてのディスク、コントローラ上のすべてのディスク、選択したディスク、またはそれらの組み合わせを指定できます。

環境によっては、すべてのディスクが同じ方法で処理されない場合もあります。たとえば、一部のディスクが初期化され、ほかのディスクがカプセル化される場合もあります。

注 初期化では、ディスク上の既存データは保持されません。

複数のディスクを同時に初期化またはカプセル化する場合は、特定のディスクまたはコントローラを対象外にすることができます。ディスクを対象から外すには、初期化またはカプセル化を実行する前に、/etc/vx ディレクトリにある disks.exclude ファイルに除外するディスク名を記述します。同様に、/etc/vx ディレクトリにある cntrls.exclude ファイルにコントローラ名を記述すると、特定のコントローラ上にあるすべてのディスクを初期化またはカプセル化の対象外にすることができます。

複数のディスクを Volume Manager の管理下に同時に配置するには、次の処理を実行します。

1. vxdiskadm メイン メニューの項目 1[1つ以上のディスクの追加または初期化]を選択します。
2. 次に示すプロンプトで、Volume Manager の管理下に追加するディスクのパターンリスト (pattern-list) を入力します。ここでは、コントロール 3 上のすべてのディスクを示す「c3」が入力されています。

ディスクの追加または初期化

Menu:VolumeManager/Disk/AddDisks

この操作を実行すると、ディスクをディスクグループへ追加します。選択済みのディスクを既存のディスクグループへ追加することや、操作の他の部分で作成した新規のディスクグループへ追加することができます。また選択済みのディスクをスペアとしてディスクグループに追加することもできます。さらに、ディスクを初期化し、ディスクグループに追加しないで交換ディスク用に確保しておくこともできます。

プロンプトでは、複数のディスクまたはパターンを入力できます。

以下にディスクの選択例を示します：

all:	すべてのディスク
c3 c4t2:	コントローラ 3、およびコントローラ 4 ターゲット 2 上の すべてのディスク
c3t4d0:	単一のディスク

追加するディスクデバイスを選択してください：

[<pattern-list>,all,list,q,?] c3

<pattern-list> には、1つのディスク、一連のディスクまたはコントローラ（オプションのターゲット）を指定できます。<pattern-list> に複数の項目を指定する場合は、各項目を空白で区切って入力してください。

追加するディスクのアドレス（デバイス名）が不明な場合は、プロンプトで「list」を入力すると、使用可能なディスクが一覧表示されます。

3. 処理を続行するには、次に示すプロンプトで、「y」を入力するか、または Return キーを押します。

選択したディスクの一覧を以下に表示します。出力形式:[Device_Name]

c3t0d0 c3t1d0 c3t2d0 c3t3d0

操作を続行しますか?[y,n,q,?] (デフォルト:y) y



4. これらのディスクをデフォルトのディスク グループ **rootdg** に追加するには、次に示すプロンプトで、「**y**」を入力するか、または **Return** キーを押します。

ディスクを既存のディスク グループに追加するか、新しいディスク グループに追加するかを選択します。また、後の作業で追加や交換操作を行うためにディスクを確保しておくこともできます。新しいディスク グループを作成するには、使用していないディスク グループ名を指定します。ディスクを後の作業で使用するために確保しておくには、ディスク グループ名に「none」を指定します。

ディスク グループの指定 [**<group>**,none,list,q,?] (デフォルト :rootdg) **y**

5. **vxdiskadm** で各ディスクにデフォルト ディスク名を使用できるようにするには、次に示すプロンプトで、「**y**」を入力するか、または **Return** キーを押します。

これらのディスクにデフォルトのディスク名を使用しますか?[y,n,q,?] (デフォルト :y) **y**

6. 次に示すプロンプトで、「**n**」を入力して、これらのディスクをホットリロケーション スペア ディスクとして使用しないように設定します。

ディスクを rootdg のスペア ディスクとして追加しますか?[y,n,q,?] (デフォルト :n) **n**

7. ディスクをホットリロケーションの利用対象から除外するかどうかを確認するメッセージが表示されたら、「**n**」を入力するか、または **Return** キーを押します。

ディスクをホットリロケーションへの利用対象から除外しますか?[y,n,q,?] (デフォルト :n)

8. 処理を続行するには、次に示すプロンプトで、「**y**」を入力するか、または **Return** キーを押します。

選択したディスクは、スペアとしてデフォルトのディスク名でディスク グループ rootdg に追加されます。

c3t0d0 c3t1d0 c3t2d0 c3t3d0

操作を続行しますか?[y,n,q,?] (デフォルト :y) **y**

9. **Volume Manager** で使用するために既に初期化済みのディスクの一覧を表示します。次に示すプロンプトで、「**y**」を入力して、すべてのディスクをすぐに使用するように設定します。

次のディスク デバイスは既に初期化されています。

現在このディスクを交換ディスクとして使用できます。

出力形式 :[Device_Name]

c3t1d0 c3t2d0 c3t3d0

Use these devices?[Y,N,S (elect) ,q,?] (デフォルト :Y) **y**

このプロンプトでは、すべてのディスクに対して「Yes」または「No」(Y または N) を選択することも、各ディスクを個別に処理する方法を選択する (s) こともできます。これらのディスクをすべて再初期化するには、次に示すプロンプトで、「Y」を入力します。

選択した次のディスクは、既に Volume Manager 用に初期化されています。ディスクが既に Volume Manager 用に初期化されていることが確実な場合は、ディスク デバイスを再初期化する必要はありません。

出力形式 : [Device_Name]

c3t1d0 c3t2d0 c3t3d0

Reinitialize these devices?[Y,N,S (elect) ,q,?] (デフォルト :Y) **Y**

- 10.** vxdiskadm はコントローラ 3 上のいずれかのディスクをカプセル化候補として表示します。「**y**」を入力するか、または Return キーを押して、このディスクをカプセル化します。

次のディスク デバイスには有効な VTOC がありますが、Volume Manager 用に初期化されていません。ディスク上に重要なデータがある場合は、ディスクを新しいディスクとして追加するのではなく、既存のディスク パーティションをボリュームとしてカプセル化します。

出力形式 : [Device_Name]

c3t0d0

Encapsulate this device?[y,n,q,?] (デフォルト :y) **y**

カプセル化では、アクティブなディスクを Volume Manager の管理下に置いて、そのディスク上のデータを保持することができます。

-
- 注** ディスクのカプセル化ではシステムをリブートする必要があります。場合によっては数回リブートする必要があります。そのような場合は、リブートを促すメッセージが表示されます。
-

ディスクが初期化中であり、Volume Manager の管理下に追加されることを示す次のようなメッセージが表示されます。

デバイス c3t1d0 を初期化しています。

デバイス c3t2d0 を初期化しています。

デバイス c3t3d0 を初期化しています。



ディスク デバイス c3t1d0 を、disk33 というディスク名でディスク グループ rootdg に追加します。

ディスク デバイス c3t2d0 を、disk34 というディスク名でディスク グループ rootdg に追加します。

ディスク デバイス c3t3d0 を、disk35 というディスク名でディスク グループ rootdg に追加します。

上記のメッセージに加えて、表面分析を実行するかどうかを選択するオプションが表示される場合もあります。

11. この後、カプセル化用に選択されたすべてのディスクが表示され、カプセル化を続行するかどうかを確認するメッセージが表示されます。「**y**」を入力するか、または **Return** キーを押してカプセル化を続行します。

次のディスクがカプセル化の対象として選択されています。

出力形式 : [Device_Name]

c3t0d0

カプセル化を続行しますか ? [y,n,q,?] (デフォルト :y) **y**

次に、カプセル化状態が表示され、ただちにシャットダウンおよびリブートする必要があることを示す次のようなメッセージが表示されます。

ディスク デバイス c3t0d0 はカプセル化され、disk38 という名前でディスク グループ rootdg に追加されます。

カプセル化の初期段階が問題なく完了しました。

できるだけ早い段階でシステムをリブートしてください。

カプセル化操作では数回リブートする必要があります。これからリブートすると、その後自動的に数回リブートされます。リブートするには、次のコマンドを実行します :

```
shutdown -g0 -y -i6
```

このコマンドにより /etc/vfstab ファイルが更新され、ボリューム デバイスを使用してファイル システムをこのディスク デバイス上にマウントできるようになります。バックアップ スクリプト、データベース、または手動で作成したスワップ デバイスなどの参照項目を更新する必要があります。

12. 次に示すプロンプトで、ほかのディスクの初期化も続行するか (**y**)、vxdiskadm メインメニューに戻るか (**n**) を指定します。

他のディスクを追加または初期化しますか ? [y,n,q,?] (デフォルト :n)

Volume Manager で使用するディスクのカプセル化

ここでは、Volume Manager で使用するディスクのカプセル化方法について説明します。ディスクが Volume Manager の管理下に置かれている場合は、カプセル化することによってディスク上の既存データが保持されます。

カプセル化の失敗を避けるため、ディスクが次の状態であることを確認してください。

- ◆ どのパーティションにも属していない空き領域が多少あること（ディスクの先頭または末尾）。
- ◆ 空きパーティションが2つあること。
- ◆ ディスク全体を表す s2（システムによっては s0）スライスがあること。

システムによっては、ブート（ルート）ディスクをカプセル化すると、後でスワップパーティションにダンプするために、そのパーティションに「swap」というタグが付けられます。

交換が必要な場合に、Volume Manager ブート ディスク エイリアスを使用してブート ディスクのミラーを識別するには、ブート ディスクをカプセル化する前に、EEPROM 変数 `use-nvramrc?` を `true` に設定しておく必要があります。この変数を `false` に設定した場合は、ブート可能ディスクを手動で特定する必要があります。この変数を `true` に設定するには、次のように入力します。

```
eeeprom use-nvramrc?=true
```

Volume Manager でディスクのカプセル化を使用するには、次の手順を実行します。

1. vxdiskadm メイン メニューの項目 2[1つ以上のディスクのカプセル化]を選択します。

注 システムによっては、例とは異なるデバイス名を使用している場合があります。デバイス名の詳細については、本章の「標準ディスク デバイス (70 ページ)」を参照してください。

2. 次に示すプロンプトで、カプセル化するディスクのデバイス名を入力します。

1つ以上のディスクのカプセル化

Menu:VolumeManager/Disk/Encapsulate

この操作は、Volume Manager で使用できるよう、1つまたは複数のディスクを変換するために使用します。この操作を実行すると、指定したディスクがディスクグループに追加され、既存のパーティションがボリュームで置き換えられます。

ディスクをカプセル化した場合、その変更を反映するにはシステムをリブートする必要があります。

プロンプトでは、複数のディスクまたはパターンを入力できます。



ディスクの選択例を次に示します。

```
all:          すべてのディスク
c3 c4t2:     コントローラ 3、およびコントローラ 4 ターゲット 2 上の
              すべてのディスク
c3t4d0:     単一のディスク
```

カプセル化するディスクデバイスの選択：

```
[<pattern-list>,all,list,q,?] c2t5d0
```

<pattern-list> には、1 つのディスク、一連のディスクまたはコントローラ（オプションのターゲット）を指定できます。<pattern-list> に複数の項目を指定する場合は、各項目を空白で区切って入力してください。

カプセル化するディスクのアドレス（デバイス名）が不明な場合は、プロンプトで「1」または「list」を入力すると、使用可能なディスクが一覧表示されます。

3. 処理を続行するには、次に示すプロンプトで、「y」を入力するか、または Return キーを押します。

選択したディスクの一覧を以下に表示します。出力形式：[Device_Name]

```
c2t5d0
```

操作を続行しますか？[y,n,q,?]（デフォルト：y） y

4. このディスクをデフォルトのディスクグループ rootdg に追加するには、Return キーを押します。

ディスクを既存のディスクグループに追加するか、新しいディスクグループに追加するかを選択します。新しいディスクグループを作成するには、使用していないディスクグループ名を指定します。

ディスクグループの指定 [<group>,list,q,?]（デフォルト：rootdg）

5. 次に示すプロンプトで、Return キーを押してデフォルトディスク名を使用するか、または別のディスク名を入力します。

このディスクにデフォルトのディスク名を使用しますか？[y,n,q,?]（デフォルト：y）

6. 処理を続行するには、次に示すプロンプトで、「y」を入力するか、または Return キーを押します。

選択したディスクは、カプセル化され、デフォルトのディスク名で rootdg ディスクグループに追加されます。

```
c2t5d0
```

操作を続行しますか?[y,n,q,?] (デフォルト :y) **y**

7. カプセル化を続行するには、次に示すプロンプトで、「**y**」を入力するか、または Return キーを押します。

次のディスクがカプセル化の対象として選択されています。出力形式:[Device_Name]

```
c2t5d0
```

カプセル化を続行しますか?[y,n,q,?] (デフォルト :y) **y**

Volume Manager で使用するためにカプセル化するディスクが表示され、リポートが必要であることを示す次のようなメッセージが表示されます。

ディスク デバイス c2t5d0 はカプセル化され、disk01 という名前でディスク グループ rootdg に追加されます。

ディスク c2t5d0 はカプセル化用に設定されています。

カプセル化の初期段階が問題なく完了しました。できるだけ早い段階でシステムをリブートしてください。

カプセル化操作では数回リブートする必要があります。これからリブートすると、その後自動的に数回リブートされます。リブートするには、次のコマンドを実行します：

```
shutdown -g0 -y -i6
```

このコマンドにより /etc/vfstab ファイルが更新され、ボリューム デバイスを使用してファイル システムをこのディスク デバイス上にマウントできるようになります。バックアップ スクリプト、データベース、または手動で作成したスワップ デバイスなどの参照項目を更新する必要があります。

8. 次に示すプロンプトで、ほかのディスクのカプセル化も続行するか (**y**)、vxdiskadm メイン メニューに戻るか (**n**) を指定します。

他のディスクをカプセル化しますか?[y,n,q,?] (デフォルト :n) **n**

環境によっては、ディスクのカプセル化に失敗する場合があります。カプセル化は、ディスク上に Volume Manager で使用できる十分な空き領域がない場合は失敗する可能性が高くなります。このような場合は、次のようなメッセージが表示され、上述の処理は中止されます。



ディスク デバイス c2t5d0 はカプセル化され、disk01 という名前でディスク グループ rootdg に追加されます。

次のエラーが原因で、カプセル化操作を実行できませんでした。

c2t5d0 は、次の理由で、カプセル化できません。

vxvm:vx slicer:ERROR: サポートされていないディスク レイアウトです。

続行するには、Return キーを押します。

今後の使用に備えたディスクの追加

今後の使用に備えて Volume Manager の管理下にディスクを追加するには、「Volume Manager 管理下への物理ディスクの配置 (121 ページ)」で説明されている手順に従って処理を実行します。ただし、ディスク グループ名の指定を促すメッセージが表示された場合は、rootdg を選択したりディスク グループ名を入力する代わりに「none」を入力してください。指定したディスクは今後の使用に備えて通常通り初期化されます。このディスクはディスク グループに追加されるまで、使用することはできません。

注 このような「スペア ディスク」とホットリロケーション スペア ディスクを混同しないでください。

Volume Manager で使用するためのディスクの再初期化

ここでは、既に初期化済みのディスクを Volume Manager で使用するために再初期化する方法について説明します。

追加するディスクが以前に Volume Manager 以外で使用されていた場合は、次の手順のいずれかを実行してください。

- ◆ ディスクをカプセル化してそのディスク情報を保持する（「Volume Manager で使用するディスクのカプセル化 (131 ページ)」を参照）。
- ◆ Volume Manager がディスクを Volume Manager 用に設定できるように、ディスクを再初期化する。再初期化では、ディスク上の既存データは保持されません。ディスクを再初期化する場合は、保持する必要のあるデータがディスクに含まれていないことを確認してください。

Volume Manager で使用するためにディスクを再初期化するには、次の手順を実行してください。

1. vxdiskadm メイン メニューの項目 1[1つ以上のディスクの追加または初期化]を選択します。

- 次に示すプロンプトで、Volume Manager の管理下に追加するディスクのディスク デバイス名を入力します。

ディスクの追加または初期化

Menu:VolumeManager/Disk/AddDisks

この操作は、1 つまたは複数のディスクをディスク グループに追加するのに使用します。選択したディスクを既存のディスク グループに追加するか、操作手順の一部として作成される新しいディスク グループに追加します。また、選択したディスクをホットスペアとしてディスク グループに追加することもできます。さらに、ディスクを初期化し、ディスク グループに追加しないで交換ディスク用に確保しておくこともできます。

プロンプトでは、複数のディスクまたはパターンを入力できます。以下にディスクの選択例を示します：

```
all:                すべてのディスク
c3 c4t2:           コントローラ 3、およびコントローラ 4 ターゲット 2 上の
                  すべてのディスク
c3t4d0:           単一のディスク
```

追加するディスクデバイスを選択してください：

[<pattern-list>,all,list,q,?] **c1t3d0**

<pattern-list> には、1 つのディスク、一連のディスクまたはコントローラ（オプションのターゲット）を指定できます。<pattern-list> に複数の項目を指定する場合は、各項目を空白で区切って入力してください。

カプセル化するディスクのアドレス（デバイス名）が不明な場合は、プロンプトで「1」または「list」を入力すると、使用可能なディスクが一覧表示されます。

- 処理を続行するには、次に示すプロンプトで、「y」を入力するか、または Return キーを押します。

選択したディスクの一覧を以下に表示します。出力形式:[Device_Name]

```
c1t3d0
```

操作を続行しますか?[y,n,q,?] (デフォルト:y) **y**

- 次に示すプロンプトで、ディスクの追加先ディスク グループを指定するか、または Return キーを押してデフォルトの rootdg を使用します。



ディスクを既存のディスク グループに追加するか、新しいディスク グループに追加するかを選択します。また、後の作業で追加や交換操作を行うためにディスクを確保しておくこともできます。新しいディスク グループを作成するには、使用していないディスク グループ名を指定します。ディスクを後の作業で使用するために確保しておくには、ディスク グループ名に「none」を指定します。

ディスク グループの指定 [`<group>,none,list,q,?`] (デフォルト :`rootdg`)

- 次に示すプロンプトで、**Return** キーを押してデフォルト ディスク名を使用するか、または別のディスク名を入力します。

このディスクにデフォルトのディスク名を使用しますか?`[y,n,q,?]` (デフォルト : `y`)

- ディスクをホットリロケーション スペア ディスクにするかどうかを確認するメッセージが表示されたら、「**n**」を入力するか、または **Return** キーを押します。

ディスクを `rootdg` のスペア ディスクとして追加しますか?`[y,n,q,?]` (デフォルト :`n`) `n`

- ディスクをホットリロケーションの利用対象から除外するかどうかを確認するメッセージが表示されたら、「**n**」を入力するか、または **Return** キーを押します。

ディスクをホットリロケーションへの利用対象から除外しますか?`[y,n,q,?]` (デフォルト :`n`)

- 処理を続行するには、次に示すプロンプトで、「**y**」を入力するか、または **Return** キーを押します。

選択したディスクは、スペアとしてデフォルトのディスク名でディスク グループ `rootdg` に追加されます。

`c1t3d0`

操作を続行しますか?`[y,n,q,?]` (デフォルト :`y`) `y`

- このディスクが既に **Volume Manager** での使用に備えて初期化済みであることを示す次のようなプロンプトが表示されます。「**y**」を入力して、このディスクの使用を確定します。

次のディスク デバイスは既に初期化されています。ディスクは現在、交換ディスクとして使用できます。出力形式 :`[Device_Name]`

`c1t3d0`

Use this device?`[y,n,q,?]` (デフォルト :`y`) `y`

- ディスクを再初期化するには、次に示すプロンプトで、「**y**」を入力するか、または **Return** キーを押します。

選択した次のディスクは、既に Volume Manager 用に初期化されています。ディスクが既に Volume Manager 用に初期化されていることが確実な場合、ディスク デバイスを再初期化する必要はありません。

出力形式 : [Device_Name]

c1t3d0

Reinitialize this device?[y,n,q,?] (デフォルト :y) **y**

Volume Manager で使用するためにこのディスクを再初期化することを示す次のようなメッセージが表示されます。

デバイス c1t3d0 を初期化しています。

ディスク デバイス c1t3d0 を、disk40 というディスク名でディスク グループ rootdg に追加します。

- 次に示すプロンプトで、ほかのディスクの初期化も続行するか (**y**)、vxdiskadm メイン メニューに戻るか (**n**) を指定します。

他のディスクを追加または初期化しますか?[y,n,q,?] (デフォルト :n)

VM ディスクのホットリロケーション プールへの追加

ホットリロケーションを使用すると、冗長性のあるサブディスクをほかのディスクにリロケートすることで、入出力障害に自動対応することができます。その後、影響を受けた Volume Manager オブジェクトおよびデータをリストアします。あらかじめ 1 つのディスクがディスク グループ内でスペアとして指定されている場合、障害ディスクのサブディスクはそのスペアディスクにリロケートされます。スペアが指定されていない場合は、利用可能なディスク グループ内の空き領域が使用されます。ただし、事前にホットリロケーションの利用対象から除外されていたディスク上の空き領域は使用されません。

スペア ディスクを指定するには、次の処理を実行します。

- vxdiskadm メイン メニューの項目 12 [ディスク グループ用のスペア ディスクの設定] を選択します。
- 次に示すプロンプトで、ディスク名 (disk01 など) を入力します。

ディスク グループ用のスペア ディスクの設定

Menu:VolumeManager/Disk/MarkSpareDisk

この操作は、ディスクをディスク グループのスペアとして設定するのに使用します。この操作ではディスク名を指定します。このディスク名は、ディスクをディスク グループに追加したときに付けた名前です。

ディスク名を入力してください。 [<disk>,list,q,?] disk01



次の確認メッセージが表示されます。

rootdg に属する disk01 をスペア ディスクとして設定できました

- 次に示すプロンプトで、ほかのディスクもスペアとして追加するか (**y**)、vxdiskadm メイン メニューに戻るか (**n**) を指定します。

他のディスクをスペアとして設定しますか? [**y,n,q,?**] (デフォルト :n)

ディスク グループ内のすべての VM ディスクで、障害時にこのディスクをスペアとして使用できます。ディスクに障害が発生すると、ホットリロケーションが自動的に実行されます (可能な場合)。障害の発生とリロケーションについては、電子メールで通知されます。リロケーションが正常に終了した後で、障害ディスクを入れ替える必要のある場合があります。

障害の発生したサブディスクをリロケートするための十分なスペア領域がない場合、ホットリロケーションは自動的に空き領域を使用します。ユーザは、ホットリロケーションで使用しない空きディスクを指定することによって、ホットリロケーションによる空き領域の使用量を制限できます。

VM ディスクのホットリロケーション プールからの削除

ディスクがスペアとして指定されている場合は、そのディスク上の領域はディスク グループ内の Volume Manager オブジェクトを作成するための空き領域として使用されることはありません。必要に応じて、ホットリロケーション ディスクのプールからスペア ディスクを削除することにより、スペア ディスクを通常通り使用できるように解放することができます。

現在スペアとして指定されているディスクを特定するには、vxdiskadm メイン メニューで list メニュー項目を選択します。スペア ディスクが spare フラグと共に一覧表示されます。

ホットリロケーション プールからディスクを削除するには、次の処理を実行します。

- vxdiskadm メイン メニューの項目13[ディスク上のスペア フラグをオフにする]を選択します。
- 次に示すプロンプトで、スペア ディスクの名前 (disk01 など) を入力します。

ディスク上のスペア フラグをオフにする

Menu:VolumeManager/Disk/UnmarkSpareDisk

この操作は、ディスクのスペア フラグをオフにするのに使用します。この操作ではディスク名を指定します。このディスク名は、ディスクをディスク グループに追加したときに付けた名前です。

ディスク名を入力してください。 [<disk>,list,q,?] disk01

次の確認メッセージが表示されます。

rootdg に属するディスク disk01 は現在スペアとして設定されていません。

- 次に示すプロンプトで、ほかのスペア ディスクも無効にするか (**y**)、vxdiskadm メインメニューに戻るか (**n**) を指定します。

他のディスクのスペア フラグをオフにしますか?[y,n,q,?] (デフォルト :n)

ホットリロケート済みサブディスクのディスクへの移動

このオプションを使用すると、障害のあったディスクを交換した後に、ホットリロケート済みのサブディスクを元のディスクに戻すことができます。

- vxdiskadm メインメニューの項目 14[ホットリロケート済みのサブディスクをディスクへ戻す]を選択します。
- 最初に、元のディスクメディア名が表示されます。

次に示すプロンプトで、ホットリロケート済みのサブディスクが存在した元のディスクメディア名を入力します。

Enter the original disk name [<disk>,list,q,?]

ただし、ホットリロケート済みサブディスクがシステム内にはない場合は、「現在ホットリロケート済みのディスクはありません。」というメッセージが出力され、Return キーを押すよう求められます。

- 新しいデスティネーションディスクに関する確認が表示されます。

While unrelocating the subdisks, you can choose to move the subdisks to a different disk from the original disk.

Unrelocate to new disk [y,n,q,?] (デフォルト :n)

- サブディスクを元のオフセットに移動できない場合は、「強制オプション」によりサブディスクのリロケート処理を解除して指定したディスクに移動できます。移動先が元のオフセットである必要はありません。

サブディスクのリロケート処理を解除して同一オフセットへの回送処理が失敗した場合は s-f オプションを使用して強制的に処理を実行しますか?[y,n,q,?] (デフォルト :n)

- 実行する操作の内容と、操作の続行を確認するメッセージが表示されます。

要求した操作は、ホットリロケート済みの全サブディスクを disk10 から rootdg ディスクグループの disk10 へ戻す操作です。

操作を続行しますか?[y,n,q,?] (デフォルト :y)

- 処理が終了すると、処理の状態が表示されます。

ディスク disk10 へのアンリロケート処理が完了しました。



ディスクのホットリロケーション利用対象からの除外

ディスクをホットリロケーションの利用対象から除外するには、次の手順を実行します。

1. vxdiskadm メイン メニューの項目 15[ディスクのホットリロケーション利用対象からの除外]を選択します。

2. 次に示すプロンプトで、ディスク名 (disk01 など) を入力します。

ディスクのホットリロケーション利用対象からの除外

Menu:VolumeManager/Disk/MarkNoHotUse

この操作によって、ディスクをホットリロケーションの利用対象から除外することができます。この操作ではディスク名を指定します。このディスク名は、ディスクをディスクグループに追加したときに付けた名前です。

ディスク名を入力してください。 [<disk>,list,q,?] **disk01**

次の確認メッセージが表示されます。

ディスク disk01 をホットリロケーションへの利用対象から除外する処理が完了しました。

3. 次に示すプロンプトで、ほかのディスクもホットリロケーションの利用対象から除外する場合は「**y**」を入力します。除外せずに vxdiskadm メイン メニューに戻る場合は、Return キーを押すか、または「**n**」を入力します。

他のディスクをホットリロケーションへの利用対象から除外しますか?[y,n,q,?]
(デフォルト :n}

ディスクのホットリロケーション設定

ディスクがホットリロケーションの利用対象から除外されている場合は、その設定を解除してディスクをホットリロケーションプールに戻すことができます。

ディスクをホットリロケーションの利用対象とするには、次の手順を実行します。

1. vxdiskadm メイン メニューの項目 16[ディスクのホットリロケーション設定]を選択します。

ディスクのホットリロケーション設定

2. 次に示すプロンプトで、ディスク名 (disk01 など) を入力します。

ディスクのホットリロケーション設定

Menu:VolumeManager/Disk/UnmarkNoHotUse

この操作によって、ディスクをホットリロケーションの利用対象に設定できます。この操作は、ホットリロケーションの利用対象から除外されていたディスクにのみ適用できません。

この操作ではディスク名を指定します。このディスク名は、ディスクをディスク グループに追加したときに付けた名前です。

ディスク名を入力してください。[<disk>,list,q,?] **disk01**

次の確認メッセージが表示されます。

rootdg の disk01 をホットリロケーションで利用できる状態に設定する処理が完了しました。

- 次に示すプロンプトで、ほかのディスクもホットリロケーションの利用対象に設定する場合は、「**y**」を入力します。設定せずに **vxdiskadm** メイン メニューに戻る場合は、**Return** キーを押すか、または「**n**」を入力します。

他のディスクをホットリロケーションで利用できる状態に設定しますか?[y,n,q,?]
(デフォルト :n)

物理ディスク上の情報の表示

ディスク情報を表示すると、初期化済みディスク、ディスクの属するディスク グループ、ディスクの状態などを確認できます。list コマンドを実行すると、認識されたすべてのディスクのデバイス名、ディスク名、各ディスクに関連付けられているディスク グループ名、および各ディスクの状態が表示されます。

ディスク情報を表示するには、次の処理を実行します。

- vxdiskadm** メイン メニューで **list** を選択します。
- 次に示すプロンプトで、参照するディスクのアドレスを入力します。すべてのディスクを一覧表示するには「**all**」を入力します。

ディスク情報の一覧表示

Menu:VolumeManager/Disk/ListDisk

このメニュー操作は、ディスクの一覧を表示する操作です。特定のディスクデバイスアドレスにあるディスクに関する詳細情報を一覧表示することもできます。

ディスク デバイス名を入力するか、「all」と入力します [<address>,all,q,?]
(デフォルト :all)

「all」を入力すると、次の情報が表示されます (システムによっては、c0t0d0s2 ではなく rootdisk と表示される場合もあります)。

DEVICE	DISK	GROUP	STATUS
c0t0d0	c0t0d0s2	rootdg	online
c1t0d0	disk01	rootdg	online
c1t1d0	-	-	online



詳細情報を表示するデバイス [`<address>,none,q,?`] (デフォルト :`none`)

- ◆ 情報を表示するデバイスのアドレスを入力すると、完全なディスク情報 (デバイス名、ディスクの種類、ディスクの共有パーティションおよび専有パーティションなど) が表示されます。

情報を確認したら、Return キーを押してメイン メニューに戻ります。

物理ディスクの削除

ディスクを削除する前に、そのディスクにデータが含まれていないこと、すべてのデータが不要であること、または別のディスクにデータを移動できることを確認してください。その後、次の手順でディスクを削除します。

1. vxdiskadm メイン メニューのメニュー項目 3[ディスクの削除] を選択します。

注 ディスク グループを無効化しないと、ディスク グループ内の最後のディスクは削除できません。

2. 次に示すプロンプトで、削除するディスクのディスク名を入力します。

ディスクの削除

Menu:VolumeManager/Disk/RemoveDisk

この操作は、ディスク グループからディスクを削除するのに使用します。

この操作ではディスク名を指定します。このディスク名は、ディスクをディスク グループに追加したときに付けた名前です。

ディスク名を入力してください。 [`<disk>,list,q,?`] **disk01**

3. 指定したディスクにボリュームが含まれている場合は、ディスクからそれらのボリュームを退避するかどうかを確認するメッセージが表示されます。ボリュームを保持する場合は、「**y**」を入力します。保持しない場合は、「**n**」を入力します。
4. 処理を続行するには、次に示すプロンプトで Return キーを押します。

要求した操作では、グループ rootdg からディスク disk01 を削除します。

操作を続行しますか? [`y,n,q,?`] (デフォルト :`y`)

vxdiskadm ユーティリティによりディスクがディスク グループから削除され、次の完了メッセージが表示されます。

ディスク disk01 は問題なく削除できました。

これで、指定したディスクを削除するか、または交換ディスクとしてシステム上に残すことができます。

- 次に示すプロンプトで、ほかのスペア ディスクも削除するか (**y**)、vxdiskadm メインメニューに戻るか (**n**) を指定します。

他のディスクを削除しますか?[y,n,q,?] (デフォルト:n)

物理ディスクの無効化 (物理ディスクのオフライン化)

ディスクが破損している場合は、オフラインにして削除します。物理ディスク デバイスを別の場所に移動し、別のシステムに接続する場合があります。ディスクをオフラインにするには、まずディスク グループから該当ディスクを削除してから、次の処理を実行します。

- vxdiskadm メイン メニューの項目 11[ディスク デバイスの無効化 (オフライン化)] を選択します。
- 次に示すプロンプトで、無効にするディスクのアドレスを入力します。

ディスク デバイスの無効化 (オフライン化)

Menu:VolumeManager/Disk/OfflineDisk

このメニュー操作は、Volume Manager からのディスクデバイスへのアクセスを無効にするのに使用します。この操作は、現在ディスクグループに属していないディスクにのみ適用できます。リポートせずにシステムからディスクを削除する場合は、この操作を実行します。

注： 多くのシステムでは、通常の動作中における、システムからのディスクの削除はサポートしていません。そのようなシステムではオフライン操作はあまり効果的ではありません。

無効にするディスク デバイスを選択してください。[<address>,list,q,?] **c1t1d0**
vxdiskadm プログラムによって、指定したディスクが無効になります。

- 次に示すプロンプトで、ほかのデバイスも無効にするか (**y**)、vxdiskadm メインメニューに戻るか (**n**) を指定します。

ほかのデバイスを無効にしますか?[y,n,q,?] (デフォルト:n)



物理ディスクの有効化

通常のシステム稼働中に、あるシステムから別のシステムへディスクを移動する場合、Volume Manager は移動したディスクを自動認識しません。ディスクの有効化タスクにより、ディスクを検索して識別し、そのディスクがディスクグループに属しているかどうかを判断するように Volume Manager を設定します。また、このタスクを実行すると、ディスクグループのデポート タスクまたはディスク デバイスの無効化（オフライン）タスクによって無効化されていたディスクへのアクセスが再び有効化されます。ディスクを有効化するには、次の処理を実行します。

1. vxdiskadm メイン メニューの項目 10[ディスク デバイスの有効化 (オンライン化)] を選択します。
2. 次に示すプロンプトで、有効化するディスクのデバイス名（デバイス一覧を表示する場合は「list」）を入力します。

ディスク デバイスの有効化 (オンライン化)

Menu:VolumeManager/Disk/OnlineDisk

この操作は、[ディスクデバイスの無効化 (オフライン化)] 操作で無効にしたディスクへのアクセスを有効にするのに使用します。

Volume Manager 以外で変更された可能性のあるディスクを再走査する場合にもこの操作を実行します。たとえば、1 台のディスクが 2 台のシステムで共有されていて、他のシステムで実行されている Volume Manager がディスクを変更した場合などが考えられます。その場合は、この操作を実行してディスクを再走査します

注： たいていの voldiskadm 操作では、ユーザの操作を必要とせずにディスクが再走査されます。これにより、ディスクを直接オンライン化する必要がなくなります（ディスクを直接オフラインにした場合を除く）。

有効にするディスク デバイスを選択してください。[<address>,list,q,?] **clt1d0**
指定したデバイスが有効化されます。

3. 次に示すプロンプトで、ほかのデバイスも有効にするか (y)、vxdiskadm メイン メニューに戻るか (n) を指定します。

他のデバイスを有効にしますか?[y,n,q,?] (デフォルト :n)

注 VxVM がディスクおよびパスを認識できるように、ホット ディスク スワップの後に vxdctl enable を実行する必要があります。

物理ディスクの交換

ディスクに障害が発生すると、そのディスクを別のディスクと交換する必要があります。このタスクでは、障害ディスクを無効化して削除し、代わりに新しいディスクを挿入する必要があります。ディスクを交換するには、次の処理を実行します。

1. vxdiskadm メイン メニューの項目 4[交換用ディスクの削除] を選択します。
2. 次に示すプロンプトで、交換するディスクの名前（ディスク一覧を表示する場合は「list」）を入力します。

交換用ディスクの削除

```
Menu:VolumeManager/Disk/RemoveForReplace
```

このメニュー操作は、ディスク名を変えずに、ディスクグループから物理ディスクを削除するのに使用します。この操作によって、ディスクの状態は、"removed" ディスクに変わります。ディスク グループに属していない初期化済みディスクがある場合、それらのディスクのいずれかを交換ディスクとして使用することができます。

ディスク名を入力してください。[<disk>,list,q,?] **disk02**

削除するディスクに関連付けられているボリュームも追加表示されます。ディスク交換時には、これらのボリュームに関連付けられているデータを保持するかどうかを指定する必要があります。必要に応じてプロンプトに対処してください。

3. 次に示すプロンプトで、交換ディスクのデバイス名を（表示されている一覧から）選択するか、または **Return** キーを押してデフォルト ディスクを使用します。

次のデバイスは交換用に使用できます：

```
c1t1d0
```

disk02 を交換するために、これらのディスクのいずれか 1 つを選択できます。交換ディスクを選択しない場合は、「none」を選択します。

デバイスを選択するか「none」を選択してください。

```
[>device>,none,q,?] (デフォルト :c1t1d0)
```

4. 次に示すプロンプトで、**Return** キーを押して処理を続行します。

要求した操作では、グループ rootdg からディスク disk02 を削除します。削除されたディスクはディスク デバイス c1t1d0 と交換されます。

操作を続行しますか？[y,n,q,?] (デフォルト :y)

処理が完了すると次のようなメッセージが表示されます。
ディスク disk02 は問題なく削除できました。



disk02 をデバイス c1t1d0 と交換します。

ディスクを問題なく交換できました。

- 次に示すプロンプトで、ほかのディスクも削除するか (y)、vxdiskadm メインメニューに戻るか (n) を指定します。

他のディスクを削除しますか?[y,n,q,?] (デフォルト :n)

障害ディスクまたは削除済みディスクの交換

ディスクを削除して、後で交換することができます。ディスクの交換を行うには、メニューの項目 4[交換用ディスクの削除] を選択してディスクを削除してから、次の手順を実行します。

- vxdiskadm メインメニューの項目 5[障害が発生したディスクまたは削除したディスクの交換] を選択します。
- 交換するディスクのディスク名を選択します。

障害が発生したディスクまたは削除したディスクの交換

Menu:VolumeManager/Disk/ReplaceDisk

このメニュー操作は、[交換用ディスクの削除] メニュー操作を実行して削除したか、使用中に障害が発生したディスクと交換するディスクを指定するのに使用します。交換するディスクの名前と、交換用に使用するディスクデバイスの入力を求めるプロンプトが表示されます。未初期化ディスクを選択して操作中にディスクを初期化するか、[ディスクの追加または初期化] メニュー操作を実行して既に初期化済みのディスクを選択することができます。

Select a removed or failed disk [<disk>,list,q,?] **disk02**

- 交換ディスクとして使用可能なディスク デバイスのデバイス名が表示されます。選択したデバイスのデバイス名を入力するか、または Return キーを押してデフォルト デバイスを使用します。

注 システムによっては、例とは異なるデバイス名を使用している場合があります。デバイス名の詳細については、「Volume Manager とは(1 ページ)」を参照してください。

次のデバイスは交換用に使用できます。

c1t0d0s2 c1t1d0s2

disk02 を交換するには、これらのディスクのいずれか 1 つを選択できます。

別のディスクを初期化して disk02 と交換する場合は、「none」を選択します。

デバイスを選択するか「none」を選択してください。

[<device>, none, q, ?] (デフォルト : c1t0d0s2)

4. 次に示すプロンプトで、Return キーを押してディスクを交換します。

要求した操作は、初期化されたデバイス c1t0d0s2 を、ディスクグループ rootdg 内の削除されたか障害が発生したディスク disk02 と交換します。

操作を続行しますか ? [y, n, q, ?] (デフォルト : y)

vxdiskadm プログラムによって、次の完了メッセージが表示されます。

グループ disk02 に属するディスク disk02 をデバイス c1t0d0s2 と問題なく交換できました。

5. 次に示すプロンプトで、ほかのディスクも交換するか (y)、vxdiskadm メインメニューに戻るか (n) を指定します。

他のディスクを交換しますか ? [y, n, q, ?] (デフォルト : n)

ディスクグループへのディスクの追加

新しいディスクを、作成済みのディスクグループに追加したい場合があります。たとえば、プロジェクトやワークグループの要件が変更されたため、現在のディスクに必要な領域が足りなくなった場合などにディスクを追加できます。ディスクグループにディスクを追加するには、ディスクの追加処理を実行します。「Volume Manager 管理下への物理ディスクの配置 (121 ページ)」を参照してください。

ディスクグループの作成

特定のアプリケーションの集まりや、特定のユーザのグループに関連するデータは、次のような場合に、別のシステムからアクセスできるように設定しておく必要があります。

- ◆ システムに障害が発生したため、データを別のシステムに移動する場合。
- ◆ 複数のシステムに作業負荷を分散する場合。

この場合、特定のアプリケーションまたはユーザに関連するデータを識別可能なディスクセットに保存しておくことが重要です。そのディスクセットを移動する場合は、その他のデータを移動することなく、アプリケーションまたはユーザに関連するディスク上のデータだけを移動できます。

注 Volume Manager はデフォルトディスクグループ rootdg をサポートしています。特に何も指定していない場合、すべてのボリュームがこのディスクグループ内に作成されます。すべてのコマンドで、デフォルトとして rootdg が使用されます。



ディスク グループを単独で作成することはできません。必ず、ディスクと共に作成します。また、ディスク グループには少なくとも 1 つのディスクが関連付けられている必要があります。

rootdg 以外のディスク グループを作成する必要がある場合は、次の処理を実行します。

1. vxdiskadm メイン メニューの項目 1[1 つ以上のディスクの追加または初期化]を選択します。
2. 次に示すプロンプトで、Volume Manager の管理下に追加するディスクのディスク デバイス名を入力します。

ディスクの追加または初期化

Menu:VolumeManager/Disk/AddDisks

この操作は、1 つまたは複数のディスクをディスク グループに追加するのに使用します。

この操作を実行すると、ディスクをディスクグループへ追加します。選択済みのディスクを既存のディスクグループへ追加することや、操作の他の部分で作成した新規のディスクグループへ追加することができます。さらに、ディスクを初期化し、ディスクグループに追加しないで交換ディスク用に確保しておくこともできます。

プロンプトでは、複数のディスクまたはパターンを入力できます。

以下にディスクの選択例を示します：

all: すべてのディスク
c3 c4t2: コントローラ 3、およびコントローラ 4 ターゲット 2 上のすべてのディスク
c3t4d0: 単一のディスク

追加するディスクデバイスを選択してください：

[<pattern-list>,all,list,q,?] **c1t2d0**

<pattern-list> には、1 つのディスク、一連のディスクまたはコントローラ（オプションのターゲット）を指定できます。<pattern-list> に複数の項目を指定する場合は、各項目を空白で区切って入力してください。

カプセル化するディスクのアドレス（デバイス名）が不明な場合は、プロンプトで「1」または「list」を入力すると、使用可能なディスクが一覧表示されます。

3. 処理を続行するには、次に示すプロンプトで、「y」を入力するか、または Return キーを押します。

選択したディスクの一覧を以下に表示します。出力形式:[Device_Name]

c1t2d0

操作を続行しますか?[y,n,q,?] (デフォルト:y) **y**

4. 次に示すプロンプトで、ディスクの追加先ディスク グループを指定します（ここでは anotherdg）。

ディスクを既存のディスクグループに追加するか、新しいディスクグループに追加するかを選択します。また、後の作業で追加や交換操作を行うためにディスクを確保しておくこともできます。新しいディスクグループを作成するには、使用していないディスクグループ名を指定します。ディスクを後の作業で使用するために確保しておくには、ディスク グループ名に「none」を指定します。

ディスク グループの指定 [`<group>`,none,list,q,?] (デフォルト :rootdg)

anotherdg

5. 同じ名前のアクティブなディスク グループが存在しないことが確認され、この新しいディスク グループを作成するかどうかを確認するメッセージが表示されます。

anotherdg という名前のアクティブなディスク グループはありません。

anotherdg という名前で新しいグループを作成しますか?[y,n,q,?] (デフォルト :y) **y**

6. 次に示すプロンプトで、Return キーを押してデフォルト ディスク名を使用するか、または別のディスク名を入力します。

このディスクにデフォルトのディスク名を使用しますか?[y,n,q,?] (デフォルト :y)

7. ディスクをホットリロケーションの利用対象から除外するかどうかを確認するメッセージが表示されたら、「n」を入力するか、または Return キーを押します。

ディスクをホットリロケーションへの利用対象から除外しますか?[y,n,q,?] (デフォルト :n)

8. ディスクをホットリロケーション スペア ディスクにするかどうかを確認するメッセージが表示されたら、「n」を入力するか、または Return キーを押します。

ディスクを anotherdg のスペア ディスクとして追加しますか?[y,n,q,?] (デフォルト :n) **y**

9. 処理を続行するには、次に示すプロンプトで、「y」を入力するか、または Return キーを押します。

新しいディスク グループ anotherdg が作成されます。選択したディスクはデフォルトのディスク名で新しいディスク グループに追加されます。

c1t2d0

操作を続行しますか?[y,n,q,?] (デフォルト :y) **y**

10. このディスクが既に Volume Manager での使用に備えて初期化済みである場合は、次のようなプロンプトが表示されます。「y」を入力してこのディスクの使用を確定します。



次のディスク デバイスは既に初期化されています。ディスクは現在、交換ディスクとして使用できます。

出力形式 : [Device_Name]

c1t2d0

Use this device?[y,n,q,?] (デフォルト :y) **y**

ディスクを再初期化する場合は、次に示すプロンプトで、「**y**」を入力するか、または **Return** キーを押します。

選択した次のディスクは、既に Volume Manager 用に初期化されています。ディスクが既に Volume Manager 用に初期化されていることが確実な場合は、ディスクデバイスを再初期化する必要はありません。

出力形式 : [Device_Name]

c1t2d0

Reinitialize this device?[y,n,q,?] (デフォルト :y) **y**

Volume Manager で使用するためにディスクを再初期化していることを示す次のようなメッセージが表示されます。

デバイス c1t2d0 を初期化しています。

anotherdg という名前のディスク デバイス c1t2d0 が属する新しいディスク グループ another01 を作成しています。

11. 次に示すプロンプトで、ほかのディスクの初期化も続行するか (**y**)、vxdiskadm メインメニューに戻るか (**n**) を指定します。

他のディスクを追加または初期化しますか ?[y,n,q,?] (デフォルト :n)

ディスク グループのデポート

システムによって現在有効化（インポート）されているディスク グループへのアクセスを無効にするためには、ディスク グループのデポート タスクを使用します。ディスク グループ内のディスクを別のシステムに移動する場合は、ディスク グループをデポートします。また、その他の目的でディスク グループ内の残りのディスクをすべて使用する場合も、ディスク グループをデポートしてください。

ディスク グループをデポートするには、次の処理を実行します。

1. vxdiskadm メイン メニューの項目 9[ディスク グループへのアクセスの無効化（デポート）]を選択します。
2. 次に示すプロンプトで、デポートするディスク グループの名前を入力します。

ディスク グループへのアクセスの無効化（デポート）

Menu:VolumeManager/Disk/DeportDiskGroup

このメニューでは、システムによって現在有効になっている（インポート済）のディスク グループへのアクセスを無効に設定することができます。ディスク グループ内のディスクを別のシステム上に移動する場合や、ディスク グループ内のすべてのディスクを別の目的に使用する場合に、ディスクをデポートします。

ディスク グループ名の入力を求めるプロンプトが表示され、またディスクを無効（オフライン）にするかどうかも指定します。一部のシステムでリムーバブル ディスク デバイスを使用している場合、ディスクを削除する前にディスクへのすべてのアクセスを無効にすることが重要です。

ディスク グループ名を入力してください。 [<group>,list,q,?] (デフォルト :list)

newdg

3. 次に示すプロンプトで、このディスク グループ内のディスクを削除する場合は、「**y**」を入力します。

要求した操作を実行すると、newdg という名前前のリムーバブルディスクグループへのアクセスを無効にします。このディスクグループは次のディスク上に格納されています：

newdg01 on device c1t1d0s2

これらのディスクへのアクセスを無効（" オフライン化 "）にするように選択できます。

これは、システムから実際にディスクを削除する場合に、エラーを防止するために必要です。

指定したディスクを無効（オフライン化）にしますか？[y,n,q,?] (デフォルト :n)

4. 次に示すプロンプトで、Return キーを押して処理を続行します。

操作を続行しますか？[y,n,q,?] (デフォルト :y)



ディスク グループのデポートが完了したことを示す次のようなメッセージが表示されます。

ディスク グループ newdg を問題なく削除できました。

- 次に示すプロンプトで、ほかのディスク グループも無効化するか (**y**)、vxdiskadm メイン メニューに戻るか (**n**) を指定します。

他のディスク グループを無効にしますか ?[y,n,q,?] (デフォルト :n)

ディスク グループのインポート

システムからディスク グループへのアクセスを有効にするには、次のメニュー タスクを使用します。ディスク グループをあるシステムから別のシステムへ移動する場合は、初めに元のシステム上でディスク グループを無効化 (デポート) します。次に、デポートされたディスク グループからほかのシステムにディスク を移動し、移動先でそのディスク グループを有効化 (インポート) します。

ディスク グループをインポートするには、次の処理を実行します。

- vxdiskadm メイン メニューの項目 8[ディスク グループへのアクセスの有効化(インポート)]を選択します。
- 次に示すプロンプトで、インポートするディスク グループの名前を入力します。

ディスク グループへのアクセスの有効化 (インポート)

Menu:VolumeManager/Disk/EnableDiskGroup

この操作は、ディスクグループへのアクセスを有効にするのに使用します。この操作は、ディスクグループをシステムからシステムへ移動する場合の最終過程で実行します。ディスクグループを移動する最初の過程では、元のホスト上で [ディスクグループへのアクセスの無効化 (デポート)] 操作を実行します。

ディスク グループを障害の発生した別のホストからインポートする場合、あらかじめディスク グループをデポートしておく必要はありません。ディスク グループ内のすべてのディスクがホスト間で移動したかどうかを確認してください。2 台のホストが SCSI バスを共有している場合は、他のホストに本当に障害が発生しているか、ほかのホストのディスク グループがデポートされているかを必ず確認してください。

2 台のアクティブなホストが同じディスク グループを同時にインポートすると、ディスク グループは破壊され、使用できなくなります。

インポートするディスク グループを選択してください。 [<group>,list,q,?] (デフォルト :list)

newdg

ディスク グループのインポートが完了したことを示す次のようなメッセージが表示されます。

newdg を問題なくインポートできました。

- 次に示すプロンプトで、ほかのディスク グループもインポートするか (**y**)、vxdiskadm メイン メニューに戻るか (**n**) を指定します。

他のディスク グループを選択しますか?[y,n,q,?] (デフォルト :n)

vxdiskadm の終了

すべてのディスクについて管理作業が完了したら、メイン メニューから「**q**」を選択して vxdiskadm を終了します。



ボリューム関連タスク

はじめに

本章では、Volume Manager の管理下にあるシステム設定の作成および管理方法について説明します。Volume Manager オブジェクト（ボリューム、プレックス、およびサブディスク）の作成、削除、および管理についての情報が含まれます。また、オンラインバックアップについても説明します。

ボリューム、プレックス、およびサブディスク関連タスクとオンラインバックアップ

本章では、次のタスクについて説明します。

ボリューム関連タスク

- ◆ ボリュームの作成
- ◆ ボリュームのサイズ変更
- ◆ ボリュームの削除
- ◆ ボリュームのミラーリング
- ◆ ミラーの削除
- ◆ RAID-5 ログの追加
- ◆ DRL ログの追加
- ◆ DRL ログの削除
- ◆ RAID-5 ログの削除
- ◆ ボリュームの停止
- ◆ ボリュームの起動
- ◆ ボリュームの読み取りポリシーの変更
- ◆ バックアップからリストアするためのボリュームの準備
- ◆ ボリュームのリカバリ
- ◆ ボリューム情報の表示



プレックス関連タスク

- ◆ プレックスの作成
- ◆ プレックスの関連付け
- ◆ プレックスの関連付けの解除とプレックスの削除
- ◆ ミラーを使用したバックアップ
- ◆ プレックス情報の表示
- ◆ プレックス属性の変更
- ◆ プレックスの状態の変更：プレックスの切り離しと接続
- ◆ プレックスの移動
- ◆ プレックスのコピー

サブディスク関連タスク

- ◆ サブディスクの作成
- ◆ サブディスクの削除
- ◆ サブディスク情報の表示
- ◆ サブディスクの関連付け
- ◆ サブディスクの関連付けの解除
- ◆ サブディスク属性の変更
- ◆ サブディスクの移動
- ◆ サブディスクの分割
- ◆ サブディスクの結合

オンライン バックアップの実行

- ◆ 高速ミラー再同期
- ◆ VM ディスク上でのボリュームのミラーリング
- ◆ VM ディスクからのボリュームの移動

注 Volume Manager コマンドによっては、スーパーユーザ権限またはその他の適切な権限が必要です。

ボリュームの作成

ボリュームを作成して、Volume Manager の仮想ディスク概念を活用します。ボリュームを作成しておく、ボリューム上にファイルシステムを配置して、ディスク領域をファイルとディレクトリで編成できます。また、データベースなどのアプリケーションを使用して、ボリューム上のデータを編成することもできます。

ディスクおよびディスクグループは、ボリュームを作成する前に、初期化し、Volume Manager に対して定義する必要があります。

次の基本または詳細のいずれかの方法でボリュームを作成できます。

- ◆ 基本—基本方法では、実行するタスクの情報を取得してから、必要な関連タスクを実行します。この方法では、ユーザ入力は最小限で済みますが、より詳細な指定を行うこともできます。基本方法は、`vxassist` コマンドを使用して最初に実行されます。`vxassist` は、任意のボリュームの基本属性のみを入力として使用し、必要なプレックスおよびサブディスクを作成します。`vxassist` は、既存のボリュームも変更できます。下位のオブジェクトや関連付けられているオブジェクトも自動的に変更されます。`vxassist` では、特定の値を指定しない限り、多くのボリューム属性でデフォルト値が使用されます。
- ◆ 詳細—詳細方法では、通常詳細な入力が必要とする多くのコマンドを使用します。これらのコマンドはブロックを組み立てていくように使用するの、特定のタスクを完了するために必要なコマンドを手動で実行するには、基本構造と構成要素の詳細な知識が必要となります。詳細方法は、複数の Volume Manager コマンドを使用して実行されます。

一般的なボリュームの手動作成手順は次の通りです。

- ◆ サブディスクの作成
- ◆ プレックスの作成
- ◆ サブディスクとプレックスの関連付け
- ◆ ボリュームの作成
- ◆ ボリュームとプレックスの関連付け
- ◆ ボリュームの初期化

ボリュームの作成は、プレックスおよびサブディスクなどの構成要素の作成を伴います。基本方法でボリュームを作成する場合、希望のボリューム特性を指定すると、下位のプレックスおよびサブディスクは自動的に作成されます。



ボリュームは、次のようなさまざまなレイアウトタイプで作成できます。

- ◆ コンカチネイテッドプレックス内でサブディスクが順次、連続して編成されているボリューム。ディスクの単一領域内のボリューム全体に空き領域がない場合は、コンカチネーションを使用して、1つ以上のディスクの複数領域からボリュームを作成できます。
- ◆ ストライプデータが複数ディスク間で均等に分散されているボリューム。ストライプは、1つのプレックスのサブディスクに交互に均等に割り当てられた均等サイズのフラグメントです。少なくともストライププレックスには2つのサブディスクが必要で、各サブディスクは異なるディスクに存在する必要があります。スループットは、プレックスがストライプされたディスクの数と共に増加します。ストライプは、特定のサブディスクにトラフィック量の多い領域が存在する場合に、入出力の負荷を分散するときに有効です。
- ◆ RAID-5ストライプを使用して、アレイ内の複数のディスクにデータとパリティを均等に分散するボリューム。各ストライプは、パリティストライプユニットとデータストライプユニットで構成されます。パリティは、ディスクの1つで障害が発生したときに、データを再編成するために使用できます。ストライプボリュームの性能と比べると、RAID-5ボリュームの場合はデータがアクセスされるたびにパリティ情報がアップデートされるため、書き込みスループットが低下します。しかし、ミラーと比べると、パリティの使用によって必要な領域を減らすことができます。
- ◆ ミラーボリュームに含まれる情報を複製した複数のプレックスを持つボリューム。ボリュームのプレックス数は1つでも問題ありませんが、完全なミラーリング（データの冗長性）のためには最低2つのプレックスが必要です。冗長性を有効にするために、各プレックスには異なるディスクのディスク領域を含める必要があります。
- ◆ 階層化ボリューム—ボリュームの最上位に構築されたボリューム。階層化ボリュームは、サブディスクをVMディスクまたはストレージボリュームにマッピングすることで構築できます。ストレージボリュームは、最上位ボリュームに類似したレイアウトを使用して繰り返し下層へとマッピングしていきます。階層化ボリュームでは、論理的レイアウトの複数の組み合わせが可能です。

注 VERITAS Volume Manager Storage Administrator では、ストライプ プロはストライプ ミラーを示す GUI 用語で、コンカチネイテッド プロはコンカチネイテッド ミラーを示す GUI 用語です。

vxassist コマンドを使用すると、ボリュームを最も簡単に作成できます。vxassist を次のように使用してボリュームを作成します。

vxassist make ボリューム名 サイズ [属性]

make はボリューム作成のキーワード、**ボリューム名**はボリュームに付けた名前、**サイズ**はボリューム内のセクタ数（デフォルト）を示します。サイズはキロバイト、メガバイト、またはギガバイト単位で指定でき、それぞれ k、m、または g のサ

フィックスを使用します。ボリューム作成時のサイズの単位指定についての詳細は、`vxintro (1M)` マニュアル ページを参照してください。必要に応じて、属性を追加して指定できます。

デフォルトでは、`vxassist` は `rootdg` ディスク グループ内にボリュームを作成します。別のディスク グループを指定するには、コマンド ラインで `-g ディスクグループ` を指定して `vxassist` を実行します。

`rootdg` ディスク グループ内にボリュームを作成すると、次の 2 つのデバイス ノード ファイルが作成されます。これらのファイルを使用してボリュームにアクセスすることができます。

- ◆ `/dev/vx/dsk/ ボリューム名` (ボリュームのブロック デバイス ノード)
- ◆ `/dev/vx/rdsk/ ボリューム名` (ボリュームの raw デバイス ノード)

ボリュームを `rootdg` やそれ以外のディスク グループ内に作成する場合、パスには次のようにディスク グループ名が含まれます。

- ◆ `/dev/vx/dsk/ ディスクグループ名 / ボリューム名`
- ◆ `/dev/vx/rdsk/ ディスクグループ名 / ボリューム名`

「コンカチネイテッド ボリュームの作成 (159 ページ)」では、デフォルト ボリュームを作成する最も簡単な方法について説明します。後続の節では、属性を指定してボリュームを作成する方法について説明します。

コンカチネイテッド ボリュームの作成

デフォルトでは、`vxassist` はディスク領域のセクションを 1 つ以上使用するコンカチネイテッド ボリュームを作成します。この方法により、ディスクが断片化されている場合でも、使用可能な空きディスク領域の個々のセクションより大きなボリュームを作成できます。

単一ディスク上に十分な領域がない場合、`vxassist` はスパン ボリュームを作成します。スパン ボリュームとは、ディスク領域のセクションが複数のディスクに分散されているコンカチネイテッド ボリュームのことです。複数のディスクの領域を使用するので、スパン ボリュームは最大サイズの単一ディスクより大きくすることができます。

ディスクへのコンカチネイテッド ボリュームの作成

ディスクが指定されていない場合、Volume Manager はボリュームの作成先ディスクを選択します。

デフォルトのコンカチネイテッド ボリュームを作成するには、次のコマンドを使用します。

```
# vxassist make ボリューム名 サイズ
```



ボリューム名にはボリュームの名前を、サイズにはボリュームのサイズをセクタ数で指定します（サフィックスで別の単位が指定されていない場合のみ）。ボリュームのサイズを変更する場合、変更後のサイズはセクタ、キロバイト、メガバイト、またはギガバイト単位で指定できます。単位を指定するには、サイズの後にサフィックス（s、k、m、またはg）を追加します。単位が指定されていない場合のデフォルトの単位はセクタです。

10 メガバイトのボリューム voldefault を作成するには、次のように入力します。

```
# vxassist make voldefault 10m
```

指定ディスクへのコンカチネイテッド ボリュームの作成

Volume Manager は、特に指定しない限り、各ボリュームが常駐するディスクを自動的に選択します。ボリュームを特定のディスクに常駐させる場合、Volume Manager にそのディスクを指定する必要があります。複数のディスクを指定できます。

特定のディスクにボリュームを作成するには、次のコマンドを使用します。

```
# vxassist make ボリューム名 サイズ ディスク名 [...]
```

disk03 にボリューム volspecific を作成するには、次のように入力します。

```
# vxassist make volspecific 3m disk03
```

ストライプ ボリュームの作成

ストライプ ボリュームには、複数の物理ディスク上の複数のサブディスクで構成されているプレックスが1つ以上含まれます。

ストライプ ボリュームを作成するには、次のコマンドを使用します。

```
# vxassist make ボリューム名 サイズ layout=stripe
```

ストライプ ボリューム volzebra を作成するには、次のように入力します。

```
# vxassist make volzebra 10m layout=stripe
```

このコマンドを実行すると、デフォルトの数のディスク上に、デフォルトのストライプユニット サイズでストライプ ボリュームが作成されます。

ボリュームの作成先ディスクを指定するには、コマンド ラインの最後にディスク名を指定します。たとえば、30 メガバイトのストライプ ボリュームを 3 つの指定ディスク (disk03、disk04、および disk05) 上に作成するには、次のコマンドを使用します。

```
# vxassist make stripevol 30m layout=stripe disk03 disk04\disk05
```

RAID-5 ボリュームの作成

RAID-5 ボリュームには、複数の物理ディスク上にある複数のサブディスクで構成されている RAID-5 プレックスが1つ含まれます。1つのボリュームに含めることのできる RAID-5 プレックスは1つのみです。また、RAID-5 ボリュームに、1つ以上の RAID-5 ログプレックスを含めることもできます。RAID-5 ログプレックスは、ボリュームに書き込まれるデータとパリティーについての情報を記録するために使用されます。

RAID-5 ボリュームを作成するには、次のコマンドを使用します。

```
# vxassist make ボリューム名 サイズ layout=raid5
```

RAID-5 ボリューム volraid を作成するには、次のように入力します。

```
# vxassist make volraid 10m layout=raid5
```

このコマンドにより、デフォルトで指定されている数のディスク上に、デフォルトのストライプユニット サイズで RAID-5 ボリュームが作成されます。また、デフォルトで RAID-5 ログも作成されます。

ミラー ボリュームの作成

新しいミラー ボリュームを作成するには、次のコマンドを使用します。

```
# vxassist make ボリューム名 サイズ layout=mirror
```

ミラー ボリューム volmir を作成するには、次のように入力します。

```
# vxassist make volmir 5m layout=mirror
```



ボリュームのサイズ変更

ボリュームのサイズを変更するには、`vxassist` または `vxresize` コマンドを使用します。

保存するデータ量に対してボリュームに十分な領域がない場合は、ボリュームのサイズを拡大する必要があります。ボリュームのサイズを拡大すると、`vxassist` では、使用可能なディスク領域が自動的に検出されます。

ボリュームのサイズを変更する場合、変更後のサイズはセクタ、キロバイト、メガバイト、またはギガバイト単位で指定できます。単位を指定するには、サイズの後にはサフィックス (`s`、`k`、`m`、または `g`) を追加します。単位が指定されていない場合のデフォルトの単位はセクタです。

注意 ボリュームのサイズは、ファイルシステムより小さく縮小しないようにしてください。VxFS ファイルシステムを使用している場合は、ファイルシステムを縮小してからボリュームを縮小します。先にファイルシステムを縮小しておかないと、データが失われ、リカバリできなくなる恐れがあります。

vxassist を使用したボリュームのサイズ変更

`vxassist` は、次の方法でボリュームのサイズを変更します。

- `growto` – 指定したサイズにボリュームを拡大
- `growby` – 指定した量だけボリュームを拡大
- `shrinkto` – 指定したサイズにボリュームを縮小
- `shrinkby` – 指定した量だけボリュームを縮小

注意 カプセル化されたブート ディスクに関連付けられたボリューム (`rootvol`、`usr`、`var`、`opt`、`swapvol` など) は、ディスク上の基礎となる物理パーティションにマッピングされていて、連続している必要があるため、拡大または縮小することはできません。`rootvol`、`usrvol`、`varvol`、または `swapvol` を拡大しようとする、ブートのためのスライスに戻る必要がある場合、システムをブートできなくなることがあります。また、これは Solaris のアップグレードに失敗する原因となり、再インストールの必要が生じる可能性があります。さらに、`upgrade_start` スクリプトにもエラーが発生することがあります。

指定サイズへの拡大

指定したサイズにボリュームを拡大するには、次のコマンドを使用します。

```
# vxassist growto ボリューム名 サイズ
```

ボリューム `volcat` を 2000 セクタに拡大するには、次のように入力します。

```
# vxassist growto volcat 2000
```

拡大するサイズの指定

指定したサイズだけボリュームを拡大するには、次のコマンドを使用します。

```
# vxassist growby ボリューム名 サイズ
```

ボリューム volcat を 100 セクタ拡大するには、次のように入力します。

```
# vxassist growby volcat 100
```

指定サイズへの縮小

指定したサイズにボリュームを縮小するには、次のコマンドを使用します。

```
# vxassist shrinkto ボリューム名 サイズ
```

ボリュームのサイズは、そのボリュームを使用しているファイルシステムやデータベースの現在のサイズより小さく設定しないようにしてください。このコマンドを安全に実行するためには、空のボリュームを使用してください。

ボリューム volcat を 1300 セクタに縮小するには、次のように入力します。

```
# vxassist shrinkto volcat 1300
```

縮小するサイズの指定

指定したサイズだけボリュームを縮小するには、次のコマンドを使用します。

```
# vxassist shrinkby ボリューム名 サイズ
```

ボリューム volcat を 300 セクタ縮小するには、次のように入力します。

```
# vxassist shrinkby volcat 300
```

vxresize を使用したボリュームのサイズ変更

vxresize コマンドを使用すると、ファイルシステムを含むボリュームのサイズを変更できます。別のコマンドでもファイルシステムを含むボリュームのサイズを変更できますが、vxresize では、ボリュームだけでなくファイルシステムのサイズも自動的に変更されるため便利です。vxresize の使用方法の詳細については、vxresize (1M) マニュアル ページを参照してください。vxresize で変更できるファイルシステムは vxfs および ufs のみです。

ボリュームが必要以上に大きい場合は、そのサイズを変更できます。



ボリユームの削除

ボリユームが不要になった場合（使用しないのでアーカイブした場合など）、`vxassist remove volume` コマンドを使用し、ボリユームを削除してディスクの空き領域を確保できます。

`vxassist` コマンドを使用して、ボリユーム全体を削除することもできます。次のように、コマンドラインでキーワード `remove` および `volume` を使用し、次にボリユーム名を指定します。

```
# vxassist remove volume ボリユーム名
```

ボリユームのミラーリング

ミラーはボリユームのコピーです。ミラーをボリユームの元のコピーとして同じディスク上に格納することはできません。ボリユームをミラーリングしておく、と、ディスクの1つに障害が発生した場合に、ボリユームデータの損失を防ぐことができます。

ミラー ボリユームの作成

ミラー ボリユームの作成には、`vxassist` コマンドを使用します。

```
# vxassist make ボリユーム名 サイズ layout=mirror
```

既存ボリユームのミラーリング

ミラー（プレックス）は既存のボリユームに追加できます。これには、`vxassist` コマンドを使用します。

```
# vxassist mirror ボリユーム名
```

たとえば、次のように入力します。

```
# vxassist mirror voltest
```

まずプレックスを作成してから、そのプレックスをボリユームに関連付ける方法でも既存のボリユームをミラーリングできます。次のコマンドを使用します。

```
# vxmake plex プレックス名 sd=サブディスク名 ...
```

```
# vxplex att ボリユーム名 プレックス名
```

ダーティ リージョン ロギングを使用したボリユームの作成

ダーティ リージョン ロギング（DRL）を有効にしてミラー ボリユームを作成するには、次のコマンドを使用して、ログ付きのミラー ボリユームを作成します。

```
# vxassist make ボリユーム名 サイズ layout=mirror,log
```

デフォルトでは、`vxassist` は各ログ サブディスクにログ プレックスを1つ作成します。

すべてのボリュームのミラーリング

システム上のすべての既存のボリュームを使用可能なディスク領域にミラーリングするには、次のコマンドを使用します。

```
# /etc/vx/bin/vxmirror -g ディスクグループ -a
```

デフォルトでミラー ボリュームを作成するように、Volume Manager を設定することもできます。この場合、次のように入力します。

```
# /etc/vx/bin/vxmirror -d yes
```

このように変更した場合も、vxassist コマンドの属性として nmirror=1 を指定すると、ミラーリングせずにボリュームを作成できます。たとえば、ミラーリングせずに 20 メガバイトのボリューム nomirror を作成するには、次のコマンドを使用します。

```
# vxassist make nomirror 20m nmirror=1
```

ミラーの削除

ミラーが不要になった場合は削除することができます。ミラーを削除すると、ディスクの空き領域を確保できます。

注 ボリュームに関連付けられた有効なプレックスのうち、最後のプレックスは削除できません。

注意 削除するミラー上のデータを保存するには、ミラーの設定を認識しておく必要があります。設定パラメータ（ストライプユニット サイズおよびサブディスクの順序）は、新しいミラーを作成して同じデータを格納する際に重要です。このタイプのミラーを削除する場合は、設定を必ず記録しておいてください。

ミラーの関連付けを解除してボリュームから削除するには、次のコマンドを使用します。

```
# vxplex -o rm dis plex_name
```

ミラー vol101-02 の関連付けを解除して削除するには、次のように入力します。

```
# vxplex -o rm dis vol101-02
```

上記のコマンドを実行すると、ミラー vol101-02 と、関連付けられたすべてのサブディスクが削除されます。



RAID-5 ログの追加

1つの RAID-5 ボリュームに含めることができる RAID-5 プレックスは1つだけです。追加プレックスは RAID-5 ログプレックスとなり、ボリュームに書き込まれるデータとパリティの情報を記録するために使用されます。vxassist を使用して RAID-5 ボリュームを作成すると、そのボリュームのログプレックスがデフォルトで作成されます。

RAID-5 ログを既存のボリュームに追加するには、次のコマンドを使用します。

```
# vxassist addlog ボリューム名
```

RAID-5 ボリューム volraid のログを作成するには、次のように入力します。

```
# vxassist addlog volraid
```

DRL ログの追加

ボリュームのダーティリージョン ロギング (DRL) を有効にするには、ログサブディスクをそのボリュームに追加し、ボリュームをミラーリングする必要があります。1つのプレックスに含めることができるログサブディスクは1つだけです。

DRL ログを既存のボリュームに追加するには、次のコマンドを使用します。

```
# vxassist addlog ボリューム名
```

ボリューム vol103 のログを作成するには、次のように入力します。

```
# vxassist addlog vol103
```

vxassist を使用してボリュームにログサブディスクを追加すると、そのログサブディスクを含むログプレックスもデフォルトで作成されます。

作成後、ログサブディスクを含むプレックスは通常のプレックスとして扱われます。データサブディスクは、ログプレックスに追加できます。ログプレックスとログサブディスクは、通常のプレックスやサブディスクを削除する場合と同じ手順で削除できます。

DRL ログの削除

vxassist コマンドを使用して、ログを削除することもできます。

```
# vxassist remove log ボリューム名
```

削除するログの数を指定するには、属性 nlog= を使用します。デフォルトでは、vxassist は1つのログを削除します。

RAID-5 ログの削除

vxassist コマンドを使用して、RAID-5 ログを削除することもできます。

```
vxassist remove log ボリューム名
```

削除するログの数を指定するには、属性 nlog= を使用します。デフォルトでは、vxassist は 1 つのログを削除します。

ボリュームの停止

ボリュームを停止すると、そのボリュームは使用できなくなります。ボリュームを停止すると、ボリュームの状態が ENABLED または DETACHED から DISABLED に変わります。コマンドを実行しても停止できない場合、ボリュームは現在の状態のままになります。ボリュームを停止するには、次のコマンドを使用します。

```
# vxvol stop ボリューム名 ...
```

たとえば、ボリューム vol01 を停止するには、次のように入力します。

```
# vxvol stop vol01
```

すべての ENABLED 状態のボリュームを停止するには、次のように入力します。

```
# vxvol stopall
```

ボリュームの起動

ボリュームを起動すると、ボリュームを使用できるようになります。

ボリュームを起動すると、ボリュームの状態が DISABLED または DETACHED から ENABLED に変わります。有効化できない場合、ボリュームは現在のままの状態になります。ボリュームを起動するには、次のコマンドを使用します。

```
# vxrecover -s ボリューム名 ...
```

すべての DISABLED 状態のボリュームを起動するには、次のように入力します。

```
# vxrecover -s
```



起動できないボリュームの一覧表示

ボリュームを起動できない原因として、設定が不正であることや、起動を妨げるその他のエラーまたは条件が発生していることが考えられます。起動できないボリュームを一覧表示するには、`vxinfo` コマンドを使用します。`vxinfo` を実行すると、1 つ以上のボリュームについてそのアクセス性や有用性に関する情報が表示されます。

```
# vxinfo [ ボリューム名
```

ボリュームの読み取りポリシーの変更

Volume Manager では、次の読み取りポリシーを選択できます。

- ◆ `round` は、検出された非連続的な入出力に対して、各プレックスを「ラウンドロビン」方式で順に読み取ります。連続的にアクセスできるのは、1 つのプレックスのみです。これは、デバイスまたはコントローラが読み取り前にキャッシュする方式を利用しています。
- ◆ `prefer` は、優先プレックスとして設定されたプレックスから最初に読み取ります。
- ◆ `select` は、プレックスのボリュームへの関連付けに基づいてデフォルトの読み取り方式を選択します。ボリュームに有効なストライププレックスが含まれている場合、`select` はそのプレックスをデフォルトで優先します。それ以外の場合、ラウンドロビン方式で読み取ります。

読み取りポリシーは、`round` から `prefer` (またはその逆)、または別の優先プレックスに変更できます。`vxvol rdpol` コマンドは、ボリュームの読み取りポリシーを設定します。

読み取りポリシーを `round` に設定するには、次のコマンドを使用します。

```
# vxvol rdpol round ボリューム名
```

たとえば、ボリューム `vol01` の読み取りポリシーをラウンドロビン方式に設定するには、次のように入力します。

```
# vxvol rdpol round vol01
```

読み取りポリシーを `prefer` に設定するには、次のコマンドを使用します。

```
# vxvol rdpol prefer ボリューム名 優先プレックス名
```

たとえば、ボリューム `vol01` の読み取りポリシーをプレックス `vol01-02` から優先的に読み取るように設定するには、次のように入力します。

```
# vxvol rdpol prefer vol01 vol01-02
```

読み取りポリシーを `select` に設定するには、次のコマンドを使用します。

```
# vxvol rdpol select ボリューム名
```

バックアップからリストアするためのボリュームの準備

ボリュームのバックアップ コピーを作成しておく、データのコピーをバックアップ時の状態のまま維持することができます。バックアップ コピーは、ディスクの障害によるボリュームの損失や、人為的なエラーによるデータの破損などが発生した場合に、ボリュームをリストアする際に使用します。**Volume Manager** を使用すると、最小限の手動操作でボリュームをバックアップできます。

`vxassist` を使用してボリュームをバックアップするには、次の処理を実行します。

1. バックアップするボリュームのスナップショット ミラーを作成します。

`vxassist snapstart` タスクを実行すると、書き込み専用のバックアップ ミラーが作成されます。このミラーは、バックアップするボリュームに関連付けられ、そのボリュームと同期します。ボリュームと同期しているバックアップ ミラーは、スナップショット プレックスとして使用できます。ただし、スナップショット処理中は、切り離されるまで継続して更新されるため、ボリュームサイズによっては時間のかかる場合があります。

ボリュームのスナップショット ミラーを作成するには、次のコマンドを使用します。

```
# vxassist snapstart ボリューム名
```

ボリューム `voldef` のスナップショット ミラーを作成するには、次のように入力します。

```
# vxassist snapstart voldef
```

2. スナップショット ボリュームを作成する適切なタイミングを選択します。

スナップショットを作成するタイミングとして、ボリュームにアクセスするユーザができる限り少ない時間帯を選択します。

3. スナップショット作成時の元のボリュームを反映するスナップショット ボリュームを作成します。

オンライン バックアップ処理は、スナップショット ミラーを持つボリューム上で `vxassist snapshot` コマンドを実行することによって完了します。このタスクは、作成されたスナップショット ミラーを切り離して、新しい標準ボリュームを作成し、そのボリュームにスナップショット ミラーを関連付けます。このスナップショットは読み取り専用ボリュームになります。この処理は数秒で終了します。



スナップショット ボリュームを作成するには、次のコマンドを使用します。

```
# vxassist snapshot ボリューム名 新しいボリューム名
```

スナップショット ボリューム voldef を作成するには、次のように入力します。

```
# vxassist snapshot voldef snapvol
```

これで、元のボリュームは引き続きアプリケーションやユーザが使用できる状態のまま、スナップショット ボリュームをバックアップ ユーティリティで使用できるようになります。

スナップショット ボリュームのバックアップは任意の方法で実行できます。領域を無駄に使用しないように、バックアップ完了後はスナップショット ボリュームを削除するようにしてください。スナップショット ボリュームは、元のボリュームと同じサイズの領域を使用します。

ボリュームのリカバリ

システムのクラッシュや I/O エラーが発生すると、ボリュームのプレックスが破壊され、CLEAN または ACTIVE 状態のプレックスがなくなる可能性があります。1 つ以上のプレックスを CLEAN 状態に設定し、それ以外のプレックスのリカバリ元として使用するよう、システムを設定することができます。

プレックスを CLEAN 状態にするには、次のコマンドを使用します。

```
# vxmend fix clean プレックス名
```

たとえば、プレックス vol101-02 を CLEAN 状態にするには、次のように入力します。

```
# vxmend fix clean vol101-02
```

vxmend の使用方法については、vxmend (1M) マニュアル ページを参照してください。

ボリューム情報の表示

vxprint コマンドを使用すると、ボリュームの設定に関する情報を表示できます。

システム内のすべてのボリュームに関するボリューム、プレックス、およびサブディスクレコード情報を表示するには、次のコマンドを使用します。

```
# vxprint -ht
```

vxprint コマンドを実行すると、次のような結果が表示されます。

```
Disk group:rootdg

DG NAME      NCONFIG  NLOG      MINORS    GROUP-ID
DM NAME      DEVICE   TYPE      PRIVLEN   PUBLLEN   STATE
V NAME       USETYPE  KSTATE    STATE     LENGTH    READPOL  PREFPLEX
PL NAME      VOLUME   KSTATE    STATE     LENGTH    LAYOUT   NCOL/WID  MODE
SD NAME      PLEX     DISK      DISKOFFS  LENGTH    [COL/]OFF  DEVICE    MODE
dg rootdg    759096729.1025.tweety

dm disk10    c1t0d0s2 sliced   559      1044400  -
dm disk20    c2t0d0s2 sliced   559      1044400  -

v pubs       fsgen    ENABLED   ACTIVE    2288     SELECT    -
pl pubs-01   pubs     ENABLED   ACTIVE    2288     CONCAT    -        RW
sd disk10-01 pubs-01   disk10    0          2288     0         c0t0d0  ENA

v voldef     sgen     ENABLED   ACTIVE    20480    SELECT    -
pl voldef-01 voldef    ENABLED   ACTIVE    20480    CONCAT    -        RW
sd disk10-02 voldef-0  disk10    2288      20480    0         c0t1d0  ENA
```

dg はディスクグループ、dm はディスク、v はボリューム、pl はプレックス、sd はサブディスクをそれぞれ示します。最初の数行は、下に続く出力行の各タイプに該当するヘッダーを示します。各ボリュームは、関連付けられたプレックスとサブディスクと一緒に一覧表示されます。

特定のボリュームについてのボリューム関連情報を表示するには、次のコマンドを使用します。

```
# vxprint -t ボリューム名
```

voldef についての情報を表示するには、次のように入力します。

```
# vxprint -t voldef
```

vxprint コマンドを実行すると、次のような結果が表示されます。

```
Disk group:rootdg

V NAME       USETYPE  KSTATE    STATE     LENGTH    READPOL  PREFPLEX

v voldef     fsgen    ENABLED   ACTIVE    20480    SELECT    -
```



プレックス関連タスク

プレックスは、物理ディスク サイズやほかの制限とは無関係に、ディスク領域を作成するサブディスクの論理的なグループ化です。ディスク データの複製（ミラーリング）は、1つのボリュームに複数のプレックスを作成することで実現されます。各プレックスには、ボリューム データの完全なコピーが含まれます。各プレックスは異なるディスク上に格納されるため、ミラーリングして複製しておくこと、ディスク サブシステムの 1 か所で障害が発生した場合にはデータの損失を防ぐことができます。複数のプレックスを作成することにより、データの整合性と信頼性も向上します。

プレックス関連タスクには、次のような処理が含まれます。

- ◆ プレックスの作成
- ◆ ミラーを使用したバックアップ
- ◆ プレックスの関連付け
- ◆ プレックスの関連付けの解除とプレックスの削除
- ◆ すべてのプレックスの一覧表示
- ◆ プレックスの表示
- ◆ プレックス属性の変更
- ◆ プレックスの状態の変更
- ◆ プレックスの移動
- ◆ プレックスのコピー

プレックスの作成

`vxmake` コマンドは、プレックスなどの Volume Manager オブジェクトを作成します。プレックスの作成時に、サブディスクを特定し、作成しようとしているプレックスに関連付けます。

既存のサブディスクからプレックスを作成するには、次のコマンドを使用します。

```
# vxmake plex プレックス名 sd= サブディスク名 ,...
```

たとえば、次のコマンドを実行すると、2 つの既存のサブディスク `disk02-01` および `disk02-02` を使用して、コンカチネイテッド プレックス `vol01-02` が作成されます。

```
# vxmake plex vol01-02 sd=disk02-01,disk02-02
```

ストライプ プレックスの作成

ストライプ プレックスを作成するには、追加属性を指定する必要があります。たとえば、次のコマンドを実行すると、ストライプ幅が 32 セクタで 2 カラムのストライプ プレックス `p1-01` が作成されます。

```
# vxmake plex p1-01 layout=stripe stwidth=32 ncolumn=2 \  
sd=disk01-01,disk02-01
```

プレックスを使用してボリュームを構築する場合は、該当するプレックスとボリュームを関連付ける必要があります（「プレックスの関連付け（173 ページ）」を参照）。

プレックスの関連付け

プレックスは、ボリュームと関連付けることにより、そのボリュームに属するプレックスになります。プレックスを既存のボリュームに関連付けるには、次のコマンドを使用します。

```
# vxplex att ボリューム名 プレックス名
```

たとえば、次のコマンドを実行すると、プレックス `vol101-02` がボリューム `vol101` に関連付けられます。

```
# vxplex att vol101 vol101-02
```

ボリュームがまだ作成されていない場合に、ボリュームを作成すると同時に 1 つ以上のプレックスと関連付けるには、次のように入力します。

```
# vxmake -U usetype vol ボリューム名 plex=プレックス名 1, プレックス名 2...
```

たとえば、次のコマンドを実行すると、`fsgen` タイプのミラー ボリューム `home` が作成され、既存のプレックス `home-1` および `home-2` と関連付けられます。

```
# vxmake -Ufsgen vol home plex=home-1,home-2
```

注 次のコマンドを既存のボリューム上で実行して、プレックスを追加および関連付けることもできます。 `# vxassist mirror ボリューム名`



プレックスの関連付けの解除とプレックスの削除

プレックスが不要になった場合は削除することができます。次のような場合に、プレックスを削除します。

- ◆ ディスクの空き領域を確保する場合。
- ◆ ボリューム内のミラー数を減らして、ほかのミラーや関連付けられているボリュームのサイズを増やす場合。プレックスおよびサブディスクを削除すると、それらの領域はほかのボリュームに追加することができます。
- ◆ ボリュームのバックアップ用に作成され、不要になった一時的なミラーを削除する場合。
- ◆ プレックスのレイアウトを変更する場合。

注意 削除するプレックス上のデータを保存するには、プレックスの設定を認識しておく必要があります。設定パラメータ（ストライプユニットサイズおよびサブディスクの順序）は、新しいプレックスを作成して同じデータを格納する際に重要です。プレックスを削除する前に、設定を必ず記録しておいてください。詳細については、「プレックス情報の表示（175 ページ）」を参照してください。

プレックスのボリュームとの関連付けを解除したり、関連付けられているボリュームから削除するには、次のコマンドを使用します。

```
# vxplex -o rm dis plex_name
```

プレックス vol101-02 の関連付けを解除して削除するには、次のように入力します。

```
# vxplex -o rm dis vol101-02
```

上記のコマンドを実行すると、プレックス vol101-02 と、関連付けられているサブディスクが削除されます。

次のように、まずプレックスとサブディスクの関連付けを解除してから削除することもできます。

```
# vxplex dis プレックス名
```

```
# vxedit -r rm プレックス名
```

これらを組み合わせると、vxplex -o rm dis コマンドと同じ処理内容になります。

ミラーを使用したバックアップ

ボリュームがミラーリングされている場合は、ボリューム ミラーの 1 つを一定期間オフライン状態にすることによって、そのボリュームのバックアップを実行できます。この場合、バックアップのみの目的で余分なディスク領域を使用する必要がありません。ただし、バックアップ実行中は、ボリュームの冗長性も失われています。

注 この節の説明は RAID-5 には適用されません。

ミラー ボリュームをアクティブなシステム上でバックアップするには、次の手順に従います。

1. バックアップの一貫性を維持するため、ユーザによる操作を短時間停止します（必要な場合のみ）。
2. ボリューム ミラーのいずれか1つの関連付けを解除します(この例ではvol101-01)。
vxplex dis vol101-01
3. 関連付けを解除したプレックスを使用して、新しい一時ボリュームを作成します。
vxmake -U gen vol tempvol plex=vol101-01
4. 一時ボリュームを起動します。
vxvol start tempvol
5. 一時ボリュームを使用して、適切なバックアップ手順を実行します。
6. 一時ボリュームを停止します。
vxvol stop tempvol
7. バックアップ プレックスの一時ボリュームへの関連付けを解除します。
vxplex dis vol101-01
8. バックアップ プレックスを元のボリュームに再度関連付け、ボリュームの冗長性を回復します。
vxplex att vol101 vol101-01
9. 一時ボリュームを削除します。
vxedit rm tempvol

vxassist コマンドを使用したもう 1 つのオンライン バックアップ方法の詳細については、「オンラインバックアップの実行 (186 ページ)」を参照してください。

プレックス情報の表示

プレックスを一覧表示すると、ボリュームを構築する際に必要な空きプレックスを特定できます。vxprint コマンドをプレックス オプション (-p) を指定して実行すると、すべてのプレックスに関する情報が一覧表示されます。

システム内のすべてのプレックスに関する詳細情報を表示するには、次のコマンドを使用します。

```
# vxprint -lp
```

特定のプレックスに関する詳細情報を表示するには、次のように入力します。

```
# vxprint -l プレックス名
```



Alt オプションを指定すると、プレックスに関する 1 行の情報が印刷されます。空きプレックスを一覧表示するには、次のように入力します。

```
# vxprint -pt
```

プレックス属性の変更

注意 プレックス属性の変更は、必要な場合にのみ、十分注意して行うようにしてください。

vxedit コマンドは、プレックスおよびほかの Volume Manager オブジェクトの属性を変更します。プレックス属性を変更するには、次のコマンドを使用します。

```
# vxedit set フィールド = 値 ... プレックス名 ...
```

comment フィールド、putil および tutil フィールドは、プレックスの作成後に Volume Manager コマンドで使用されます。putil 属性はリブート時に維持されますが、tutil フィールドは一時的なものでリブート時に維持されません。

putil および tutil は、両方とも 3 つの機能を持ち、それらの機能に従って番号が付けられています。これらのフィールドは必要に応じて変更できます。putil0 および tutil0 と記されているユーティリティ フィールドは、Volume Manager で使用されます。putil1 および tutil1 と記されているフィールドは、ほかの Veritas 製品で使用されます。putil2 および tutil2 と記されているフィールドはユーザ フィールドです。176 ページの表 4 は、putil および tutil フィールドの機能の一覧です。

表 4. putil[n] フィールドと tutil[n] フィールド

フィールド	ユーティリティ フィールドの説明
putil0	Volume Manager コマンドで使用。リブート時に保持されます。
putil1	グラフィカル ユーザ インタフェースなど、高度なユーティリティで使用。このフィールドはリブート時に保持されます。
putil2	システム管理者またはサイト特有のアプリケーションで使用。このフィールドはリブート時に保持されます。
tutil0	Volume Manager コマンドで使用。リブート時に消去されます。
tutil1	グラフィカル ユーザ インタフェースなど、高度なユーティリティで使用。リブート時に消去されます。
tutil2	システム管理者またはサイト特有のアプリケーションで使用。リブート時に消去されます。

コマンド例

```
# vxedit set comment="my plex" tutil2="u" user="admin" vol01-02
```

この例では `vxedit` を使用して次の属性を設定しています。

- ◆ `comment` フィールド（プレックスの用途を特定）を「`myplex`」に設定
- ◆ `tutil2` をサブディスクが使用中であることを示す「`u`」に設定
- ◆ ユーザ ID を `admin` に変更

特定のプレックスがボリュームと関連付けられないようにするには、次のコマンドで指定しているように、`putil0` フィールドを `NULL` 以外の文字列に設定します。

```
# vxedit set putil0="DO-NOT-USE" vol01-02
```

プレックスの状態の変更：プレックスの切り離しと接続

ボリュームを作成してオンライン状態（`ENABLED`）にすると、`Volume Manager` では一時的にプレックスをボリュームから切り離すことができます。この機能は、プレックスが存在しているハードウェアを修理する必要がある生じた場合や、ボリュームが起動できない状態になった場合に、そのボリュームの回復のためソースプレックスを手動で選択する必要が生じた場合などに役立ちます。

ディスクまたはシステムの障害を解決するには、ボリュームをオフライン状態にし、そのプレックスを接続および切り離すなどの処理が必要になります。ディスク障害を解決するには、`vxmend` および `vxplex` の 2 つのコマンドを使用します。

プレックスを `OFFLINE` にし、そのプレックスのサブディスクを含む物理ディスク上で修理や保守を実行するには、次のコマンドを使用します。

```
# vxmend off プレックス名 ..
```

ディスクのヘッドが破損している場合は、障害ディスク上のサブディスクと関連付けられているすべてのプレックスを `OFFLINE` にする必要があります。たとえば、プレックス `vol01-02` および `vol02-02` に修理するドライブ上のサブディスクが含まれる場合は、次のように入力します。

```
# vxmend off vol01-02 vol02-02
```

上記のコマンドを実行すると、`vol01-02` および `vol02-02` が `OFFLINE` 状態になり、再び変更されるまでそのままの状態が保持されます。

プレックスの切り離し

ミラーボリュームのプレックスを一時的に切り離すには、次のコマンドを使用します。

```
# vxplex det プレックス名
```

たとえば、次のコマンドを実行すると、プレックス `vol01-02` を一時的に切り離し保守モードにすることができます。

```
# vxplex det vol01-02
```



このコマンドではプレックスが一時的に切り離されますが、プレックスとボリューム間の関連付けは維持されます。ただし、プレックスを入出力操作に使用することはできません。上述のコマンドで切り離されたプレックスは、システムのリブート時にリカバリされます。プレックスの状態は **STALE** に設定されており、**vxvol start** コマンドが適切なボリューム上で実行されると（システムリブート時など）、プレックスの内容がリカバリされ **ACTIVE** 状態になります。

プレックスがボリュームのアクティブ パートとして復帰できるようになったら、次の手順に従います。

- ◆ ボリュームが有効 (**ENABLED**) でない場合は、次のコマンドを使用してボリュームを起動します。

```
# vxvol start ボリューム名
```

ボリュームを起動できない場合は、次のコマンドを使用してプレックスの 1 つを **CLEAN** に設定します。

```
# vxmend fix clean プレックス名
```

その後、ボリュームを起動します。

- ◆ プレックスのカーネル状態がまだ **ENABLED** でない場合は、次のコマンドを使用します。

```
# vxplex att ボリューム名 プレックス名 ...
```

OFFLINE 状態のプレックスを **ACTIVE** 状態に戻す場合、このコマンドは、プレックスの内容をリカバリしてからプレックス状態を **ACTIVE** に設定します。

プレックスの接続

ディスクを修理または交換し、再び使用できるようになると、プレックスをオンライン状態に戻す必要があります（プレックスの状態を **ACTIVE** に設定する）。

ボリュームが現在有効 (**ENABLED**) である場合は、次のコマンドを使用します。

```
# vxplex att ボリューム名 プレックス名 ...
```

たとえば、ボリューム **vol101** 上のプレックス **vol101-02** に対するコマンドは次の通りです。

```
# vxplex att vol101 vol101-02
```

上記のコマンドを実行すると、プレックス内容のリカバリが開始され、リカバリの完了後に、プレックスユーティリティの状態が **ACTIVE** に設定されます。

ボリュームが使用中でない (**ENABLED** でない) 場合は、次のコマンドを使用します。

```
# vxmend on プレックス名
```

たとえば、プレックス **vol101-02** の場合は、次のように入力します。

```
# vxmend on vol101-02
```

この例では、vol01-02 の状態は STALE に設定されています。ボリュームの次回リブート時に、プレックス上のデータはほかのプレックスから回復され、プレックス状態が ACTIVE に設定されてボリュームに反映されます。

プレックスの状態を手動で変更する場合は、本章の「ボリュームのリカバリ (170 ページ)」を参照してください。これらのコマンドの詳細については、vxmake (1M) および vxmend (1M) マニュアル ページを参照してください。

プレックスの移動

プレックスを移動する場合は、元のプレックスのデータを新しいプレックスにコピーします。あるプレックスから別のプレックスにデータを移動するには、次のコマンドを使用します。

```
# vxplex mv 元のプレックス 新しいプレックス
```

プレックスの移動を正しく実行するには、次の条件を満たす必要があります。

- ◆ 古いプレックスがアクティブな (ENABLED) ボリュームの一部で、アクティブ状態であること。
- ◆ 新しいプレックスのサイズが古いプレックスと同じか、またはそれ以上であること。
- ◆ 新しいプレックスが別のボリュームに関連付けられていないこと。

プレックスのサイズによって、次のように処理結果が異なります。

- ◆ 新しいプレックスが元のプレックスより小さいか、空白プレックスである場合は、元のプレックス データを完全にコピーできません。このような場合は、-o force オプションを指定する必要があります。
- ◆ 新しいプレックスが元のプレックスより大きい場合、空白領域が少ない場合、元のプレックス上の既存のデータは新しいプレックス上にコピーされます。元のプレックス上にはなく新しいプレックス上にはある領域は、同じボリュームに関連付けられているほかのプレックスから埋められます。
- ◆ 新しいプレックスがボリューム自体よりも大きい場合、新しいプレックスの余った領域は初期化されずに未使用のまま残ります。

プレックスのコピー

この処理は、ボリューム上の内容を指定したプレックスにコピーします。コピーされるボリュームは有効にしておくことはできません。また、プレックスをほかのボリュームに関連付けることはできません。プレックスをコピーするには、次のコマンドを使用します。

```
# vxplex cp ボリューム名 新しいプレックス
```

コピー処理の完了後、新しいプレックスは指定したボリュームボリューム名に関連付けられていません。新しいプレックスはボリューム データの完全なコピーを含んでいます。コピー元のプレックスのサイズは、ボリュームと同じまたはそれ以上である必要があります。



す。コピー元のプレックスがボリュームより大きい場合、データのコピーは不完全な状態で終了します。同じ理由から、**新しいプレックスには空白プレックスを指定しないでください。**

サブディスク関連タスク

サブディスクは、Volume Manager 設定の下位の構成ブロックで、ボリュームを構築するために使用されます。以下に、サブディスク上で実行できるタスクについて説明します。

サブディスクの作成

注 vxassist または Storage Administrator を使用してボリュームを作成すると、サブディスクが自動的に作成されます。

vxmake コマンドを使用すると、サブディスクなどの Volume Manager オブジェクトを作成できます。サブディスクの作成時には、サブディスクの名前とサイズ、ディスク内のサブディスクの開始ポイント（オフセット）、およびディスクメディア名を指定する必要があります。

サブディスクを作成するには、次のコマンドを使用します。

```
# vxmake sd サブディスク名 ディスク , オフセット , サイズ
```

たとえば、次のコマンドを実行すると、ディスク disk02 の先頭から開始する、サイズが 8000 セクタのサブディスク disk02-01 が作成されます。

```
# vxmake sd disk02-01 disk02,0,8000
```

デフォルトでは、Volume Manager コマンドのサイズはセクタ単位です。サフィックス (k、m、または g など) を追加すると、サイズの単位を変更できます。

新規サブディスクを使用してボリュームを構築する場合は、そのサブディスクをプレックスに関連付ける必要があります（「サブディスクの関連付け (181 ページ)」を参照）。サブディスクは、すべてのプレックスレイアウト（コンカチネイテッド、ストライプ、RAID-5）で同じ方法で作成されます。

サブディスクの削除

サブディスクを削除するには、次のコマンドを使用します。

```
# vxedit rm サブディスク名
```

たとえば、次のコマンドを実行すると、サブディスク disk02-01 が削除されます。

```
# vxedit rm disk02-01
```

サブディスク情報の表示

`vxprint` コマンドは、Volume Manager オブジェクトに関する情報を表示します。すべてのサブディスクに関する一般情報を表示するには、次のコマンドを使用します。

```
# vxprint -st
```

`-s` オプションは、サブディスクに関する情報を指定します。`-t` オプションは、一覧表示されるオブジェクトの種類に従って、1行の出力レコードを印刷します。

特定のサブディスクに関する完全な情報を表示するには、次のように入力します。

```
# vxprint -l サブディスク名
```

たとえば、次のコマンドを実行すると、サブディスク `disk02-01` に関するすべての情報を取得できます。

```
# vxprint -l disk02-01
```

上記のコマンドを実行すると、次のような結果が表示されます。

```
Disk group:rootdg

Subdisk:disk02-01

info:disk=disk02 offset=0 len=205632
assoc:vol=mvol plex=mvol-02 (offset=0)

flags:enabled

device:device=c2t0d1s2 path=/dev/vx/dmp/c2t0d1s4 diskdev=32/68
```

サブディスクの関連付け

サブディスクをプレックスに関連付けると、サブディスクで定義されている総ディスク領域をプレックス内の特定のオフセットに配置できます。サブディスクが占める全領域は、ほかのサブディスクとまったく重複しないようにする必要があります。サブディスクをプレックスに関連付ける方法は、全体的な設定状況に応じていくつかあります。

特定のプレックスに必要なサブディスクをすべて作成してある場合は、プレックスの作成時に次のコマンドを使用してサブディスクを関連付けます。

```
# vxmake plex プレックス名 sd= サブディスク名 ,...
```

たとえば、次のコマンドを実行すると、プレックス `home-1` が作成され、サブディスク `disk02-01`、`disk02-00`、および `disk02-02` がプレックス `home-1` に関連付けられます。

```
# vxmake plex home-1 sd=disk02-01,disk02-00,disk02-02
```



サブディスクは、オフセット 0 から指定された順序で関連付けられます。disk02-01 で定義されるディスク領域が 1 番目、ディスク領域 disk02-00 が 2 番目、disk02-02 が 3 番目になります。このタイプのコマンドを使用する場合は、プレックスを作成し、そのプレックスへ各サブディスクを関連付けるために、複数のコマンドを指定する必要はありません。このサブディスク関連付け方式は、初期設定中に利用すると便利です。

サブディスクは既存のプレックスに関連付けることもできます。次のコマンドを使用すると、1 つ以上のサブディスクを既存のプレックスと関連付けることができます。

```
# vxsd assoc プレックス名 サブディスク名 [サブディスク名2 サブディスク名3 ...]
```

たとえば、次のコマンドを実行すると、サブディスク disk02-01、disk02-00、および disk02-02 がプレックス home-1 に関連付けられます。

```
# vxsd assoc home-1 disk02-01 disk02-00 disk02-01
```

プレックスが空でない場合、新しいサブディスクは、そのプレックスに既に関連付けられているサブディスクの後に追加されます（このコマンドで -1 オプションが指定されていない場合）。-1 オプションは、プレックス内の特定のオフセットにサブディスクを関連付けます。

-1 オプションは、特定のボリュームに空白プレックス（サブディスク間に空き領域があるプレックス）を作成していて、そのプレックスを完全にさせる場合に指定します。プレックスを完全体にするには、空白プレックス内の空き領域と一致するサイズのサブディスクを作成する必要があります。次に、プレックス内の空き領域の開始オフセットを指定して、作成したサブディスクをプレックスに関連付けます。次のコマンドを使用します。

```
# vxsd -1 オフセット assoc 空白プレックス名 正確なサイズのサブディスク
```

注 Volume Manager では、2 つのサブディスクで定義された領域が 1 つのプレックス内で重複することは許されないため、サブディスクのサイズは正確に指定する必要があります。

ストライプ サブディスクの場合は、サブディスクにカラム番号とカラム オフセットを指定できます。

```
# vxsd -1 カラム数 / オフセット assoc プレックス名 サブディスク名 ...
```

-1 オプションでストライプ プレックスに 1 つの番号のみを指定すると、その番号はカラム番号として解釈され、サブディスクはそのカラムの末尾に関連付けられます。

ログ サブディスクの関連付け

ログ サブディスクは、ダーティ リージョン ロギングを使用してボリュームの一部となるプレックスに追加されます。ダーティ リージョン ロギングは、1 つ以上のログ サブディスクを含むミラー ボリュームに対して有効です。

ダーティ リージョン ロギングについては、第 3 章の「ダーティ リージョン ロギング (55 ページ)」を参照してください。ログ サブディスクは、通常のプレックス方式が採用されている場合は無視され、ダーティ リージョン ログを保持する目的でのみ使用されます。

注 プレックスに関連付けることができるログ サブディスクは1つのみです。このログ サブディスクは頻繁に書き込まれるため、あまり使用されていないディスク上に配置する必要があります。頻繁に使用されているディスク上にログ サブディスクを配置すると、システム パフォーマンスが低下する場合があります。

次のコマンドを使用して、既存のボリュームにログ サブディスクを追加することもできます。

```
# vxassist addlog ボリューム名 ディスク
```

このコマンドは、指定したボリュームのログ プレックス内にログ サブディスクを自動的に作成します。

ログ サブディスクを既存のプレックスに追加するには、次のコマンドを使用します。

```
# vxsd aslog プレックス サブディスク
```

サブディスクは、ログ サブディスクとして使用する名前です。プレックスは、DRL を有効にする前にミラー ボリュームに関連付けられている必要があります。

たとえば、次のコマンドを実行すると、サブディスク `disk02-01` がプレックス `vol101-02` (ボリューム `vol101` に関連付け済み) に関連付けられます。

```
# vxsd aslog vol101-02 disk02-01
```

サブディスクの関連付けの解除

サブディスクとそのサブディスクが属するプレックスとの関係を解除するには、サブディスクのプレックスへの関連付けを解除する必要があります。サブディスクの関連付けは、サブディスクが削除されるか、または別のプレックスで使用されている場合に解除されます。サブディスクの関連付けを解除するには、次のコマンドを使用します。

```
# vxsd dis サブディスク名
```

たとえば、次のコマンドを実行すると、サブディスク `disk02-01` と現在関連付けられているプレックスとの関連付けが解除されます。

```
# vxsd dis disk02-01
```

注 次のコマンドを使用して、サブディスクを削除することもできます。

```
# vxsd -orm dis サブディスク名
```



サブディスク属性の変更

注意 サブディスク属性の変更は、必要な場合にのみ、十分注意して行うようにしてください。

vxedit コマンドは、サブディスクの属性をほかの Volume Manager オブジェクトに変更します。サブディスクに関する情報を変更するには、次のコマンドを使用します。

```
# vxedit set フィールド = 値 ... サブディスク名
```

たとえば、次のコマンドを実行すると、サブディスク disk02-01 の comment フィールドが変更されます。

```
# vxedit set comment="新しいコメント" disk02-01
```

vxedit を使用して変更できるサブディスク フィールドは次の通りです。

- ◆ name
- ◆ putil[n] フィールド
- ◆ tutil[n] フィールド
- ◆ len (サブディスクの関連付けが解除されている場合のみ)
- ◆ comment

注 putil0 フィールドにデータを入力すると、サブディスクがまだプレックスの一部でない場合は、そのサブディスクがプレックスの一部としては使用されないようになります。

サブディスクの移動

サブディスクを移動すると、サブディスクの内容が別のサブディスクにコピーされます。移動するサブディスクがプレックスに関連付けられている場合、元のサブディスクに格納されたデータは新しいサブディスクにコピーされます。古いサブディスクはプレックスとの関連付けを解除され、新しいサブディスクがそのプレックスに関連付けられます。関連付けは、プレックス内の、元のサブディスクと同じオフセットで行われます。サブディスクを移動するには、次のコマンドを使用します。

```
# vxsd mv 古いサブディスク名 新しいサブディスク名
```

サブディスクの移動を正しく実行するには、次の条件を満たす必要があります。

- ◆ 関連するサブディスクのサイズがすべて同じであること。
- ◆ 移動するサブディスクが、アクティブなボリューム (ENABLED) 上のアクティブなプレックスの一部であること。
- ◆ 新しいサブディスクが別のプレックスに関連付けられないこと。

サブディスクの分割

サブディスクを分割すると、既存のサブディスクが2つに分割されます。サブディスクを分割するには、次のコマンドを使用します。

```
# vxsd -s size split サブディスク名 新サブディスク名 1 新サブディスク名 2
```

各オペランドの説明は次の通りです。

- ◆ サブディスク名は、元のサブディスクの名前です。
- ◆ 新サブディスク1は、作成する2つのサブディスクのうちの1つ目のサブディスクの名前です。
- ◆ 新サブディスク2は、作成する2つのサブディスクのうちの2つ目のサブディスクの名前です。

—s オプションは、作成する2つのサブディスクのうちの1つ目のサブディスクのサイズを指定するために使用します。2つ目のサブディスクでは、元のサブディスクで使用していた領域の残りを使用します。

分割前は、元のサブディスクはプレックスに関連付けられています。分割後は、作成された両方のサブディスクが同じプレックスに関連付けられます。

元のサブディスクを3つ以上のサブディスクに分割するには、前述のコマンドを必要な回数繰り返し実行します。

サブディスクの結合

サブディスクを結合すると、複数の既存のサブディスクが1つのサブディスクに結合されます。サブディスクを結合するには、対象となるサブディスクが同じディスク上で隣接している必要があります。選択したサブディスクが関連付けられている場合は、同じプレックスに関連付けられ、そのプレックス内で隣接している必要があります。サブディスクを結合するには、次のコマンドを使用します。

```
# vxsd join サブディスク 1 サブディスク 2 新しいサブディスク
```



オンライン バックアップの実行

Volume Manager では、ボリューム デバイスのスナップショット バックアップを提供しています。スナップショット バックアップは、`vxassist` およびその他のコマンドを使用して実現されます。バックアップを実行するには、ボリューム内容の整合性に関する要件によってさまざまな方法があります。いずれの方法の場合でも、ボリュームの内容を完全に格納できるサイズのプレックスが必要です。プレックスは必要なサイズより大きくても構いませんが、逆に小さい場合はバックアップを実行できません。

ボリューム バックアップは、`vxassist` コマンドを使用すると簡単に実行できます。`vxassist snapstart`、`snapwait`、および `snapshot` タスクを使用すると、最小限の手動操作でボリュームのオンライン バックアップを実行できます。

`vxassist snapshot` 処理は 2 段階に分かれています。

1. `vxassist snapstart` を実行してスナップショット ミラーを作成します。
2. `vxassist snapshot` を実行してスナップショット ボリュームを作成します。

注 `vxassist` コマンドを使用して、RAID-5 ボリュームのスナップショットを作成するには、この節で説明しているボリュームバックアップの手順に従ってください。

`vxassist snapstart` タスクを実行すると、書き込み専用のバックアップ プレックスが作成されます。このプレックスは、バックアップするボリュームに関連付けられ、そのボリュームと同期します。ボリュームと同期しているバックアップ プレックスは、スナップショット ミラーとして使用できます。新しいスナップショット ミラーの状態が `SNAPDONE` に変わると、更新処理は終了です。この変更は、1 つ以上のミラーの状態が `SNAPDONE` に変わるまでは、`vxassist snapwait` タスクによって参照することができます。関連付けに失敗した場合、スナップショット ミラーは削除され、使用されていた領域は解放されます。

スナップショット ミラーを同期させると、切り離されるまで継続して更新されます。その後ならば、スナップショット ボリュームを既存ボリュームのイメージとして作成するため適切な時間を設定できます。また、スナップショットの実行中（通常は 1 分未満）は、ユーザにシステムを使用しないように指示することもできます。スナップショット ミラーの作成に必要な時間は、スナップショット ボリュームを作成する時間よりも長くなります。

オンライン バックアップ処理は、`SNAPDONE` ミラーを持つボリューム上で `vxassist snapshot` コマンドを実行することによって完了します。このタスクでは、作成されたスナップショット（標準のミラーになる）を切り離し、新しい標準ボリュームを作成し、スナップショット ミラーをスナップショット ボリュームに関連付けます。その結果、スナップショットは標準のミラーとなり、スナップショットの状態は `ACTIVE` に設定されます。

スナップショット処理で割り込みが発生すると、スナップショット ミラーはボリュームの起動時に自動的に削除されます。

次の処理を実行して、完全な vxassist バックアップを実行します。

1. 次のコマンドを使用して、ボリュームのスナップショット ミラーを作成します。

```
# vxassist snapstart ボリューム名
```

2. snapstart 手順が完了し、ミラーが SNAPDONE 状態になったら、スナップショットタスクを実行する適切なタイミングを選択します。ユーザに、実行するスナップショットを通知し、ファイルを保存して、実行中はシステムを使用しないように指示します。
3. 次のコマンドを使用して、元のボリュームを反映するスナップショットボリュームを作成します。

```
# vxassist snapshot ボリューム名 一時ボリューム名
```

4. fsck (または、ボリューム上で実行するアプリケーションに適したユーティリティ) を使用して、一時的なボリュームの内容を消去します。たとえば、次のように入力します。

```
# fsck -y /dev/vx/rdisk/ 一時ボリューム名
```

5. 一時的なボリュームを、テープまたは適切なバックアップ メディアにコピーします。
6. 次のコマンドを使用して、新しいボリュームを削除します。

```
# vxedit -rf rm 一時ボリューム名
```

高速ミラー再同期

高速ミラー再同期 (FMR) 機能は、バックアップや意志決定支援などの支援処理に対する VxVM のスナップショット機構の効果を高めることにより、陳腐化したミラーに対して高速で効果的な再同期化処理を実行します。

FMR の有効化

vxassist を使用して新しいボリュームを作成する際に、属性に on または off を指定して FMR を有効または無効にできます。FMR をボリューム上で使用するかどうかを指定するには、fmr または fastresync というキーワードを属性として使用します。

FMR が有効なボリュームを作成するには、次の vxassist make コマンドを使用します。

```
# vxassist make ボリューム名 サイズ fmr=on
```

FMR のデフォルトは off ですが、このデフォルトは vxassist デフォルト ファイルの設定で変更できます。



FMR 機能を ON または OFF に設定するには、`vxvol` コマンドを使用することもできます。FMR を使用するには、スナップショットが作成される時点で FMR が有効化されていて、スナップバックが完了するまで FMR が有効な状態である必要があります。FMR を無効にすると、指定したボリュームに関するすべての追跡マップが消去されます。その後の再接続において FMR 機能は使用されませんが、ボリュームの完全な再同期化処理は行われます。FMR が後で有効化された場合でも同様です。

FMR を有効化するには、次のコマンドを使用します。

```
# vxvol set fmr=on ボリューム名
```

FMR を無効化するには、次のコマンドを使用します。

```
# vxvol set fmr=off ボリューム名
```

スナップショット ボリュームのマージ

ボリュームのスナップショット コピーを元のボリュームにマージして戻すことができます。スナップショット プレックスはスナップショット ボリュームから切り離され、元のボリュームに接続されます。スナップショット ボリュームは削除されます。このタスクにより、ボリューム内のデータは再同期化され、プレックスの一貫性が保たれます。

スナップショットを元のボリュームにマージするには、次のコマンドを使用します。

```
# vxassist snapback レプリカ - ボリューム
```

レプリカ-ボリュームは、ボリュームのスナップショット コピーです。

デフォルトでは、元のプレックス内のデータが、マージされたボリュームで使用されます。これを複製ボリュームからのコピー データにするには、次のコマンドを使用します。

```
# vxassist -o resyncfromreplica snapback レプリカ - ボリューム
```

スナップショット ボリュームの関連付けの解除

スナップショットと元のボリューム間のリンクを完全に削除すると、スナップショット ボリュームは独立したボリュームになります。

スナップショットと元のボリュームの関連付けを解除するには、次のコマンドを使用します。

```
# vxassist snapclear レプリカ - ボリューム
```

レプリカ-ボリュームは、ボリュームのスナップショット コピーです。

スナップショット ボリューム情報の表示

vxassist snapprint コマンドを使用すると、元のボリュームとそのレプリカ (スナップショット コピー) の関連を表示できます。

snapprint オプションの構文は次の通りです。

```
# vxassist snapprint [ ボリューム名 ]
```

出力結果は、次のように表示されます。

V NAME	USETYPE	LENGTH	RP NAME	VOLUME	LENGTH	RRPLEXID
v vol	fsgen	2048	rp vol-05	SNAP1-vol	3040	rp vol-04

ボリュームを指定した場合は、そのボリュームに関する情報が表示されます。ただし、ボリュームに有効な FMR マップがない場合は、エラー メッセージが表示されます。ボリュームを指定しない場合は、ディスクグループ内の全ボリュームに関する情報が表示されます。

VM ディスク上でのボリュームのミラーリング

ボリュームを VM ディスク上でミラーリングすると、ディスク上の別の場所にボリュームのコピーを 1 つ以上作成できます。ボリュームのミラー コピーを作成することによって、ディスクの障害発生時にシステムのデータ損失を防ぐことができます。このタスクをルート ディスク上で実行すると、有効な起動情報のコピーを代替ディスク上に作成できます。代替ディスクを作成しておくことで、ブート ディスクが破損した場合でもシステムをブートすることができます。

注 このタスクはコンカチネイテッド ボリュームのみをミラーリングします。既にミラーリング済みのディスクや複数のディスク上に常駐するサブディスクを含むディスクは対象外となります。

ボリュームをディスク上でミラーリングするには、ターゲット ディスクに元のディスクと同じか、またはそれ以上の領域があることを確認してから、次の処理を実行します。

1. vxdiskadm メイン メニューの項目 6 [ディスク上でのボリュームのミラー] を選択します。
2. 次に示すプロンプトで、ミラーリングするディスクの名前を入力します。

ディスク上でのボリュームのミラー

Menu:VolumeManager/Disk/Mirror

この操作は、ディスク上でボリュームのミラーを作成するのに使用します。これらのボリュームは別のディスクか、使用可能なディスク領域にミラーできます。既にミラーされている場合は、ボリュームはミラーされません。また、複数のサブディスクから成るボリュームもミラーすることはできません。



ブート ディスクからボリュームをミラーしておくと、代替ブート ディスクとして使用できるディスクを作成できます。

ディスク名を入力してください。 [`<disk>,list,q,?`] `disk02`

- 次に示すプロンプトで、ターゲット ディスクの名前を入力します (このディスクのサイズは元のディスクと同じかそれ以上であることが必要)。

ディスク `disk02` 上のボリュームを使用可能な任意のディスク領域にミラーするか、特定のディスク上へミラーするかを選択できます。特定のディスクにミラーするには、ディスク名を指定します。使用可能な任意のディスク領域にミラーする場合は、"`any`"を選択します。

Enter destination disk [`<disk>,list,q,?`] (デフォルト :`any`) `disk01`

注 代替ルート ディスクを作成する場合は、必ずデスティネーション ディスクを指定してください。指定しない場合、**Volume Manager** は代替ルート ディスクになるディスクを選択します。ただし、お使いのシステムがこのディスクからブートできない場合もあります。

- 次に示すプロンプトで、**Return** キーを押してミラーを作成します。

要求した操作は、ディスクグループ `rootdg` 内のディスク `disk02` 上のすべてのボリュームを、ディスク `disk01` 上の使用可能なディスク領域にミラーする操作です。

注：この操作には時間がかかります。

操作を続行しますか ? [`y,n,q,?`] (デフォルト :`y`)

ミラー操作の状態が表示されます。

ボリューム `voltest-bk00` のミラー ...

ディスク `disk01` をミラーできました。

- 次に示すプロンプトで、ディスクを別のディスク上にミラーリングするか (**y**)、`vxdiskadm` メイン メニューに戻るか (**n**) を指定します。

ボリュームを他のディスク上にミラーしますか ? [`y,n,q,?`] (デフォルト :`n`)

VM ディスクからのボリュームの移動

ディスクを無効化または削除する前に、ディスク内のデータをシステム上の別のディスクに移動することができます。データを移動するには、ターゲット ディスクに十分な領域があることを確認してから、次の処理を実行します。

1. vxdiskadm メイン メニューの項目 7[ディスクからの複数ボリュームの移動] を選択します。
2. 次に示すプロンプトで、移動するボリュームを含むディスクの名前を入力します。

ディスクからの複数ボリュームの移動

Menu:VolumeManager/Disk/Evacuate

このメニュー操作は、あるディスクを使用しているボリュームを他のディスク上に移動するのに使用します。ディスクを永久に、または交換のために一次的に削除する直前に、この操作を実行します。ボリュームの移動先ディスクを一覧形式で指定するか、同一のディスク グループに属する任意のディスク領域にボリュームを移動することが可能です。

注：ディスクを削除せずにディスクからボリュームを移動するだけでは、今後の操作によって移動したボリュームが元のボリュームに戻されてしまう場合があります。たとえば、移動操作を 2 回続けて実行すると、移動先のディスクからまた元のディスクにボリュームが戻されてしまいます。

ディスク名を入力してください。[<disk>,list,q,?] disk01

次の表示の後で、ディスク(1つ以上)の移動先のディスク一覧を指定することもできます。

移動先ディスクの一覧を指定できます。ディスクメディア名(例: disk01)を、空白で区切り、すべてのメディア名を 1 行の中に指定します。ディスクメディア名を入力しないと、ボリュームはディスクグループ内の使用可能な任意の領域に移動されます。

次に示すプロンプトで、Return キーを押してボリュームを移動します。

要求した操作では、グループ rootdg に属するディスク disk01 からすべてのボリュームを移動します。

注：この操作には時間がかかります。

操作を続行しますか?[y,n,q,?](デフォルト:y)

ディスクからのボリュームの移動時には、処理の状態が次のように表示されます。

ボリューム voltest の移動...

ボリューム voltest-bk00 の移動...



ボリュームの移動がすべて終了すると、次のようなメッセージが表示され、処理が完了したことを示します。

ディスク `disk01` を退避できました。

- 次に示すプロンプトで、ディスクを別のディスクからボリュームを移動するか (`y`)、`vxdiskadm` メイン メニューに戻るか (`n`) を指定します。

ボリュームを他のディスクから移動しますか?[`y,n,q,?`](デフォルト :`n`)

はじめに

この章では、VERITAS Volume Manager (VxVM) に付随して提供される、オプションのクラスタ機能について説明します。この機能を利用すると、クラスタ環境において VxVM を使用することができます。Volume Manager のクラスタ機能のライセンスは別立てとなっております。

この章では、次のトピックについて説明します。

- ◆ クラスタ機能の概要
- ◆ VxVM クラスタ内のディスク
- ◆ ダーティ リージョン ロギングとクラスタ環境
- ◆ Volume Manager クラスタ機能のアップグレード
- ◆ クラスタ関係の Volume Manager のユーティリティとデーモン
- ◆ クラスタ処理と高速ミラー再同期 (FMR)

クラスタ関連のエラー メッセージについては、『VERITAS Volume Manager Reference Guide』のエラー メッセージに関する章を参照してください。



クラスタ機能の概要

Volume Manager のクラスタ機能を利用すると、Volume Manager の管理下にある一連のディスク (VM ディスク) に、複数のホストから同時にアクセスして管理することができます。クラスタは、一連のディスクを共有するホストの集まりです。各ホストは、クラスタ内ではノードと呼ばれます。ノードはネットワークを通じて接続されます。1 台のノードが故障しても、ほかのノードは引き続きディスクにアクセスすることができます。Volume Manager のクラスタ機能では、すべてのノードにおいて、ディスクの構成の設定 (変更を含む) を同一の論理ビューとして見ることができます。

注 Volume Manager ではクラスタあたり 4 台のノードをサポートします。

以下の節で、Volume Manager のクラスタ機能についてさらに詳しく説明します。

Volume Manager におけるオブジェクトの共有

クラスタ機能を有効にすると、クラスタ内のすべてのノードによって Volume Manager オブジェクトを共有することができます。

Volume Manager のクラスタ機能では、2 種類のディスク グループがサポートされています。

- ◆ 専有ディスク グループ。単一のノードに属します。専有ディスク グループは単一のシステムのみインポートされます。物理的には複数のシステムから専有ディスク グループにアクセス可能であっても、実際にアクセスを許されるシステムは1 つだけです。
- ◆ クラスタで共有するディスク グループ。クラスタ共有ディスク グループとも言います。すべてのノードによって共有されます。クラスタで共有するディスク グループはすべてのクラスタ ノードにインポートされます。クラスタで共有するディスク グループ内のディスクは、クラスタに含まれる可能性のあるすべてのシステムから物理的にアクセスできる必要があります。

Volume Manager のクラスタにおいては、ほとんどのディスク グループが共有されます。ただし、ルート ディスク グループ (rootdg) は必ず専有ディスク グループとして扱われます。

共有ディスク グループ内のディスクはクラスタ内のすべてのノードからアクセス可能であるため、複数のクラスタ ノード上のアプリケーションが同一のディスクに同時にアクセスできます。ライセンスおよびディスク グループのアクティブ化モードの内容に応じて、共有ディスク グループ内のボリュームには、クラスタ内の複数のノードから同時にアクセスできます。

ノードからのアプリケーションの入出力に対してディスク グループ内のボリュームをアクセス可能にするには、共有ディスク グループをそのノード上でアクティブ化する必要があります。アプリケーションにディスクの読み取りまたは書き込みを許可するかどうかは、ディスク グループのアクティブ化モードで設定します。共有ディスク グループに関して有効なアクティブ化モードには、専有書き込み、共有書き込み、読み取り専用、共有読み取りおよびオフ (または非アクティブ) があります。各モードについては、表 5「共有ディスク グループのアクティブ化モード」で説明します。

表 5. 共有ディスク グループのアクティブ化モード

専有書き込み	ノードには、ディスク グループに対する排他的な書き込みが許可されます。ほかのノードは、書き込みのためにこのディスク グループをアクティブ化できません。
共有書き込み	ノードには、ディスク グループに対する書き込みが許可されます。
読み取り専用	ノードには、ディスク グループに対する読み取りが許可されます。クラスタ内のほかの全ノードの書き込みは拒否されます。このノードには、ディスク グループに対する書き込みは許可されません。ほかのノード上で書き込みモードのいずれかのディスク グループをアクティブ化すると、エラーになります。
共有読み取り	ノードには、ディスク グループに対する読み取りが許可されます。ほかのノードにはディスク グループに対する書き込みが許可されますが、このノードは許可されません。
オフ	ノードには、ディスク グループに対する読み取りと書き込みのどちらも許可されません。ディスク グループの照会操作は許可されます。

注 ディスク グループのアクティブ化は、VxVM 3.0 の新機能です。以前のリリースとの互換性を保つために、デフォルトでは、アクティブ化モードは **Volume Manager** ユーティリティに対して透過的になっています。共有ディスク グループは、自動的に共有書き込みモードでアクティブ化されます。

HA アプリケーション、オフホスト バックアップなど、クラスタの特殊な使用方法では、ディスク グループのアクティブ化を活用して、クラスタ内の異なるノードからのボリューム入出力を正確に制御することができます。アクティブ化モードの使用方法については、「ディスク グループのアクティブ化 (203 ページ)」を参照してください。

注 Volume Manager 3.0 の新機能ストライプ ミラー ボリューム、タスク モニタ、およびオンライン再レイアウトは専有ディスク グループでは使用できますが、共有ディスク グループではまだサポートされていません。

Volume Manager のクラスタ機能を通じてアクセスできるのは raw デバイスだけです。ファイル システムが組み込まれた共有ボリュームはサポートされていません。

Volume Manager のクラスタ機能では現在のところ、クラスタで共有するディスク グループ内の RAID-5 ボリュームをサポートしていません。ただし、クラスタの特定のノードに接続された専有ディスク グループ内で、RAID-5 ボリュームを使用することができます。

サポートされていないオブジェクトを含むディスク グループが共有ディスク グループとしてインポートされた場合は、このディスク グループをデポートしてください。保持するボリュームをサポートされているレイアウトに再構築してから、共有ディスク グループとして再度インポートしてください。

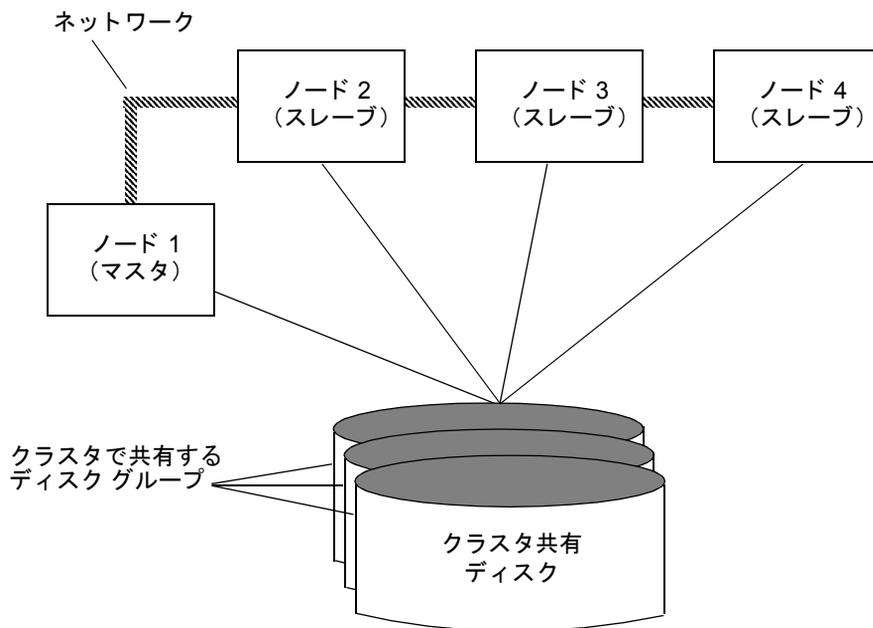


クラスタにおけるボリューム管理の仕組み

Volume Manager のクラスタ機能は外部的に提供されるクラスタ マネージャと連携して働きます。クラスタ マネージャはデーモンであって、クラスタ メンバシップの変化を VxVM に知らせる働きをします。各ノードは独立して起動され、それぞれがオペレーティングシステム、クラスタがサポートされている VxVM、およびクラスタ マネージャのコピーを保持しています。ノードをクラスタに追加すると、そのノードから共有ディスクにアクセスできるようになります。ノードをクラスタから取り外すと、そのノードから共有ディスクにアクセスできなくなります。システム管理者がノードをクラスタに追加するには、該当のノード上でクラスタ マネージャを起動します。

単純なクラスタの構成を 197 ページの図 24「4つのノードで構成されるクラスタの例」に示します。ノードはすべてネットワークに接続されています。各ノードはクラスタで共有するディスクグループにも接続されています。クラスタ マネージャにとっては、すべてのノードは同等です。ただし、Volume Manager のクラスタ機能では、1つのノードをマスターノードとし、残りのノードをスレーブノードとする必要があります。マスターノードは Volume Manager のある種の処理を調整する役割を負います。どのノードにマスターの機能を割り当てるかは、VxVM ソフトウェアによって決められます(どのノードもマスターノードとなりえます)。その割り当てはそのノードがクラスタに属している間ずっと変わりません。そのノードがクラスタから取り外されると、スレーブノードのいずれかひとつが新しいマスターノードに割り当てられます。図 24「4つのノードで構成されるクラスタの例」の例では、ノード 1 がマスターノードであり、ノード 2、ノード 3、ノード 4 がスレーブノードです。

図 24. 4つのノードで構成されるクラスタの例



システム管理者は、`vxdg` ユーティリティを使用して、ディスク グループをクラスタで共有するディスク グループとして設定します（詳細は、「`vxdg` (211 ページ)」を参照）。クラスタで共有するディスク グループをあるひとつのノードにインポートすると、ディスク ヘッダーにクラスタ ID が記録されます。そのクラスタにほかのノードが追加されると、それらのノードはディスク グループがクラスタで共有可能であることを認識して、そのディスク グループをインポートします。システム管理者は任意の時点で共有ディスク グループをインポートまたはデポートすることができます。その処理はすべてのノードに波及します。

各物理ディスクには一意のディスク ID が付けられます。マスタ ノード上でクラスタを起動すると、すべての共有ディスク グループがマスタ ノードにインポートされます（ただし、`noautoimport` 属性が設定されているものを除きます）。スレーブ ノードをクラスタに追加しようとする時、マスタにインポート済みのディスク ID の一覧がマスタからスレーブに送られます。スレーブ側では一覧をチェックし、一覧中のすべてのディスクにアクセスできるかどうかを確認します。スレーブからアクセスできないディスクが一覧中にひとつでもある場合、スレーブはクラスタに追加されません。一覧中のすべてのディスクにスレーブからアクセスできる場合は、マスタにインポートされている共有ディスク グループがスレーブにもインポートされて、スレーブがクラスタに追加されます。あるノードをクラスタから取り外すと、そのノードにインポートされていた共有ディスクはすべてデポートされますが、それらのディスクは残りのノードにはインポートされたままです。

共有ディスク グループの再設定はすべてのノードに対して一括して行われます。つまり、ディスク グループの構成の変更はすべてのノードに対して一斉に全く同様に行われます。この変更処理は全体としてひとまとまりのものです。すべてのノードで一斉に同様に変更できない場合には、その処理はまったく行われなかったものとして扱われます。

クラスタ内のすべてのメンバは、アクティブ化モードの設定に従って、クラスタで共有するどのディスク グループに対しても同時に読み取りおよび書き込みを行うことができます（「ディスク グループのアクティブ化（203 ページ）」を参照）。クラスタ内のアクティブなノードは、ほかのノードで障害が発生しても、その影響を受けません。クラスタで共有されているディスク グループ内に格納されているデータは、そのクラスタ内で少なくとも 1 台のノードがアクティブであるかぎり、利用可能です。クラスタで共有しているディスク グループにどのノードからアクセスしても、ディスク グループの構成の設定はまったく同様に見えます。各ノード上で稼働している複数のアプリケーションから VM ディスク上のデータに同時にアクセスすることができます。

注 VxVMでは、複数のノードからの共有ボリュームへの同時書き込みを防止しません。アプリケーションのレベルで（分散ロック マネージャを使用している場合など）、何らかの一貫性制御がなされているものと想定されています。

構成の設定と初期化

新しいクラスタにノードを追加するのに先立って、システム管理者は構成の設定に関する情報を登録する必要があります。それは通常はクラスタ マネージャを設定する間に指定します。指定したデータは通常はクラスタ マネージャの設定データベースに格納されます。必要な情報の詳細な内容とフォーマットは個々のクラスタ マネージャによって異なります。VxVM で必要とされる情報には下記のものがあります。

- ◆ クラスタ ID
- ◆ ノード ID
- ◆ 各ノードのネットワーク アドレス
- ◆ ポート アドレス

クラスタにノードを追加すると、ノードを起動したときに、そのノード上の VxVM にこの情報が自動的にロードされます。

ノードの初期化はクラスタ マネージャを起動する際に行われます。それを通じて、ノード上の種々のクラスタ コンポーネント（クラスタ サポートを伴う VxVM、クラスタ マネージャ、分散ロック マネージャなど）が起動します。これが完了すると、アプリケーションを起動することができます。システム管理者はクラスタに追加する各ノード上でクラスタ マネージャ起動手続きを呼び出します。

クラスタ環境にある VxVM の初期化では、クラスタの構成の設定情報を読み込み、クラスタへノードを追加します。最初にクラスタに追加したノードがマスタ ノードとなり、後から追加したノードがスレーブ ノードとなります。複数のノードを同時に追加した場合には、VxVM ソフトウェアによってマスタ ノードが選定されます。追加の完了したノードからは共有ディスクにアクセスできるようになります。

クラスタの再設定

クラスタの状態に変化（ノードの追加や取り外し）があった場合、クラスタの再設定が行われます。各ノードのクラスタ マネージャは、クラスタ内のほかのノードを監視し、クラスタ メンバシップに変化があると、それを VxVM に通知します。それに基づいて、VxVM は適切な処理を行います。

クラスタの再設定中は、共有ディスクへの入出力は一時停止されます。クラスタの再設定が完了すると、共有ディスクへの入出力が再開されます。したがって、少しの間アプリケーションは停止したかのように見えます。

Volume Manager による処理やリカバリなどが行われていると、それが完了するまで、クラスタの再設定が行われなことがあることがあります。ボリュームの再設定（後述）がクラスタの再設定と同時に進行することはありません。状況に応じて、処理は停止して後に再起動することがあります。通常は、クラスタの再設定が優先されます。ただし、ボリュームの再設定が確定段階に達している場合は、完了するまでそちらが先に進められます。

SunCluster_ をクラスタ マネージャとして使用している場合は、vxclust ユーティリティによってクラスタの再設定が調整され、VxVM と SunCluster_ 間の通信が可能となります。SunCluster_ と vxclust が関係して動作し、クラスタの再設定の各処理が正しい順序で完了します。

VCS をクラスタ処理のフレームワークとして使用している場合は、クラスタの再設定の全処理が VxVM カーネル内で行われます。

クラスタの再設定については、210 ページの「vxclust」および「vxclustadm (209 ページ)」を参照してください。

ボリュームの再設定

ボリュームの再設定とは、Volume Manager オブジェクトの構成（ディスク グループやボリュームやミラーなど）に、追加、変更、削除などを行う処理です。クラスタにおいては、それらの処理にはすべてのノードが関与します。ボリュームの再設定はすべてのノードに波及します。すべてのノードで設定が一斉に同様に変更されます。

注 ボリュームの再設定の起動と調整はマスタ ノードから行うため、システム管理者は Volume Manager オブジェクトを変更するユーティリティをマスタ ノード上で実行する必要があります。

ボリュームの再設定においては、vxconfigd デーモンが積極的な役割を演じます。ボリュームを正しく再設定するためには、vxconfigd がすべてのノードで稼働している必要があります。



マスタ ノード上のユーティリティはそのノードの `vxconfigd` デーモンに要求を伝え、`vxconfigd` デーモンが要求された変更の妥当性をノード上でチェックします。たとえば、新しいディスク グループを作成しようとしたが、同じ名前のディスク グループが既に存在する場合はエラーとなります。そのチェックが済むと、マスタ ノード上の `vxconfigd` デーモンはクラスタ内にあるほかのすべての `vxconfigd` デーモンに変更の詳細情報を記したメッセージを送信します。そのメッセージを受信した各スレーブ ノードの `vxconfigd` デーモンはそれぞれのノードにおけるチェックを行います。たとえば、作成されようとしているディスク グループと名前が同じ専有ディスク グループが存在しないことをチェックします。また、新しいディスクが作成される場合は、そのディスクにアクセスできることをチェックします。すべてのノード上の `vxconfigd` デーモンによって、提案された変更が妥当であると判定されると、各 `vxconfigd` デーモンはそのことをそのカーネルに伝えます。すると、各カーネルは協力して、トランザクションを確定させるかアボートさせるかを判定します。トランザクションを確定させる前に、すべてのカーネルは入出力処理が行われていないことを確認する必要があります。ボリュームの再設定を要求しトランザクションの確定を調整する役割はマスタ ノードが果たします。

ボリュームの再設定の処理中にいずれかのノードの `vxconfigd` デーモンが何らかのエラーを起こした場合、そのことがすべてのノードに通知され、処理は中止されます。処理中にいずれかのノードが取り外された場合も、処理は中止されます。ただし、マスタ ノードによってトランザクションが既に確定されている場合を除きます。マスタ ノードが取り外された場合は、新しいマスタ（それまではスレーブであった）によって、処理を完了させるか中止するかが決定されます。そのどちらになるかは、新しいマスタが前のマスタから処理が完了した通知を受けたかどうかによって異なります。この通知はすべてのスレーブに対して出されます。新しいマスタがその通知を受けていなければ、ほかのどのスレーブもその通知を受けていないこととなります。

ボリュームの再設定が行われている最中にクラスタにノードを追加しようとした場合、どのように処理されるかは、処理がどこまで進んでいたかによって異なります。まだカーネルが関与するに至っていなかった場合、ボリュームの再設定は一時中断されて、ノードの追加が完了した後で再開されます。カーネルが既に関与していた場合は、ボリュームの再設定が完了するまで、ノードは追加されません。

エラーが発生した場合（スレーブでのチェックで異常が検出された場合、クラスタからノードが取り外された場合など）、ユーティリティにエラー情報が戻され、エラーが発生したノードを知らせるメッセージが起動ノード上のコンソールに表示されます。

ノードのシャットダウン

システム管理者は指定したノードに関してクラスタをシャットダウンすることができます。そのためには、該当のノード上でクラスタ マネージャのシャットダウン プロシージャを呼び出します。すると、クラスタ アプリケーションが終了した後で、クラスタ コンポーネントが停止されます。VxVM ではノードのクリーン シャットダウンがサポートされています。これは、共有ボリュームへのアクセスがすべて終了してから、整合性を崩すことなくクラスタからノードを切り離します。ホストは稼働し続けていますが、クラスタ アプリケーションを実行することはできなくなります。

Volume Manager のクラスタ機能は、各ボリュームに関する全体的な状態情報を保持しています。このため、ノードがクラッシュした際は、リカバリする必要があるボリュームが正確に判定されます。クラッシュまたはほかの何らかの異常があったためにノードがクラスタから取り外された場合、書き込みが完了しなかった可能性のあるボリュームが VxVM によって判定され、そのボリュームがマスタによって再同期化されます。そのボリュームにデータ リージョン ロギングが適用されている場合は、それが利用されます。

ノードのクリーン シャットダウンは、すべてのクラスタ アプリケーションを停止するプロセスを実施した後またはそれと平行して行います。クラスタ アプリケーションの特性とシャットダウン プロシージャしだいで、シャットダウンが正常に終了するまでに長い時間がかかる可能性があります (数分から数時間)。たとえば、多くのアプリケーションに「掃き出し」という概念があります。それは、新しい仕事の受け付けを停止し、仕掛中の仕事をすべて完了させてから、処理を終了することです。それに長い時間がかかることがあります。たとえば、長時間を要するトランザクションを実行中の場合です。

VxVM シャットダウン プロシージャを起動すると、シャットダウン対象のノード上の共有ディスク グループがすべてチェックされ、その結果に基づいてシャットダウン処理が継続するか、またはエラーとなります。

- ◆ 共有ディスク グループ内のボリュームがすべて閉じられている場合、それらのディスク グループはアプリケーションで利用できなくなります。取り外されるノード上ではそれらのボリュームはすべて閉じられていることがすべてのノードに分かっているので、再同期化は行われません。
- ◆ 共有ディスク グループ内に開かれているボリュームがひとつでもあると、VxVM シャットダウン プロシージャはエラーとなります。ただし、正常に終了するまで、シャットダウン プロシージャを繰り返し試行することができます。この処理にはタイムアウトはありません。クラスタ アプリケーションが動作していないことを確認するためのサービスであるからです。

注 シャットダウンが正常に終了すると、対象のノードはクラスタから取り外されます。そのノードから共有ボリュームにアクセスすることはできなくなります。共有ボリュームにアクセスするためには、そのノードをクラスタに再び追加する必要があります。

シャットダウン処理には長時間かかることがあります。そのため、その進行中に、別の再設定処理が開始される可能性があります。通常は、別の再設定処理が完了するまで、シャットダウン処理は一時中断されます。ただし、シャットダウン処理が既に相当に進んでいる場合は、シャットダウン処理の方が先に完了することもあります。



ノードのアボート

ノードを完全に取り外せない場合もあります。その原因としては、ホストがクラッシュしたか、ほかの何らかのクラスタ コンポーネントによって該当ノードが緊急に取り外されたと判断されたことが考えられます。このような場合は、後続のクラスタ再設定機能によって、VxVM のアボート機能が呼び出されます。この機能によって、共有ボリュームへのすべてのアクセスが即座に停止する試みがなされます。ただし、既に開始されている入出力処理があれば、それが完了するのを待ちます。

まだ開始されていない入出力処理はエラーになり、共有ボリュームが削除されます。したがって、共有ボリュームにアクセスしていたアプリケーションはエラーとなります。

ノードがアボートまたはクラッシュした後は、共有ボリュームをリカバリする必要があります。ミラー間の同期が崩れている可能性が非常に高いからです。リカバリ方法としては、残っている正常なノードを利用するか、または後でクラスタを再起動します。

クラスタのシャットダウン

クラスタに属するノードをすべて取り外した場合は、次にクラスタを起動するときに、共有ボリュームをリカバリする必要があるかどうかを判定しなければなりません。すべてのノードを完全に取り外した場合は、リカバリする必要はありません。最後のノードを完全に取り外し、以前完全に取り外すことができなかったノードの再同期化処理が正常に完了している場合も、リカバリする必要はありません。最後のノードの取り外しが正常に完了しなかった場合、または以前の取り外しの再同期化処理が正常に完了しなかった場合は、リカバリを実施しなければなりません。

VxVM クラスタ内のディスク

クラスタ内の各ノードのディスクの状態は必ず一致していなければなりません。特に、あるノードからあるディスクに書き込みが行えない場合、書き込み処理の結果が呼び出し元に返される前に、すべてのノードからのそのディスクへのアクセスを停止する必要があります。したがって、あるノードがあるディスクにアクセスできない場合、そのノードはほかのノードに連絡して、ディスクの状態をチェックしてもらう必要があります。ディスクが故障している場合は、そのディスクにアクセスできるノードはないので、そのディスクを取り外すことにすべてのノードが合意できます。ディスクが故障しているのではなく、いくつかのノードからのアクセスパスが故障している場合、ノード間でディスクの状態が一致なくなります。このような不一致を解決するために、ポリシーを設定しておく必要があります。

ディスクの切り離しポリシー

上記の不一致に対処するには、次に示すポリシー（ディスクグループに対して設定）を使用できます。ポリシーの設定には `vxedit (1M)` コマンドを使用します。

共有ディスクグループに関するグローバル接続ポリシーでは、クラスタ内のいずれかのノードがディスク障害を通知した場合、クラスタ全体で（グローバルに）切り離しが行われます。このポリシーはデフォルトです。

ローカル接続ポリシーでは、ディスク障害が発生すると、障害を検知したノードだけが影響を受けます。ディスクの使用を継続できるよう、クラスタ内のすべてのノードに対して通信が試行されます。すべてのノードが障害を検知した場合、このディスクはクラスタ全体で切り離されます。

ディスクグループのアクティブ化

ディスクグループのアクティブ化によって、クラスタ内の異なるノードからのボリューム入出力が制御されます。指定したノードがクラスタ内のほかのノードと競合するモードでアクティブ化されている場合は、ディスクグループをアクティブ化できません。

表6「許可されるアクティブ化モードと競合するアクティブ化モード」に、共有ディスクグループで許可されるアクティブ化モードと競合するアクティブ化モードを示します。

表 6. 許可されるアクティブ化モードと競合するアクティブ化モード

クラスタ内のディスクグループのアクティブ化モード	ほかのノードにおけるディスクグループのアクティブ化モード			
	専有書き込み	共有書き込み	読み取り専用	共有読み取り
専有書き込み	失敗	失敗	失敗	成功
共有書き込み	失敗	成功	失敗	成功
読み取り専用	失敗	失敗	成功	成功
共有読み取り	成功	成功	成功	成功



ユーザがアクティブ化モードの設定を実行できるようにするには、デフォルト ファイル /etc/default/vxdg を作成して、次の行を挿入する必要があります。

default_activation_mode= アクティブ化モード

アクティブ化モードには、off (オフ)、shared-write (共有書き込み)、shared-read (共有読み取り)、read-only (読み取り専用)、または exclusive-write (専有書き込み) を設定します。

共有ディスク グループを作成してインポートすると、共有ディスク グループは指定したモードでアクティブ化されます。ノードをクラスタに追加すると、すべての共有ディスク グループが指定したモードでアクティブ化されます。

注:
◆ アクティブ化を有効にするためにデフォルト ファイルを使用する場合は、クラスタ内の全ノード上のデフォルト ファイルを等しくすることをお勧めします。このように設定しない場合は、アクティブ化の結果が予測できません。
◆ デフォルトのアクティブ化モードがオフ以外に設定されていると、クラスタ内のほかのノードがディスク グループを競合するモードでアクティブ化していた場合、クラスタ追加後のアクティブ化、またはディスク グループの作成やインポートがエラーになる可能性があります。

ダーティ リージョン ログとクラスタ環境

ダーティ リージョン ログ (DRL) はボリュームのリカバリのためのオプション機能です。この機能を利用すると、システムに障害が発生した後で、ミラー ボリュームのリカバリを迅速に行うことができます。ダーティ リージョン ログはクラスタで共有するディスク グループに関してサポートされています。この節では、DRL を簡単に概説し、クラスタ環境における DRL の仕組みを説明します。

DRL では、ミラー ボリュームへの書き込みにより変更のあった領域を記録します。そして、その情報を利用して、ボリューム中の該当する部分だけをリカバリの対象とするようにします。DRL では、ボリュームを連続する一連のリージョンに論理的に区分して、各リージョンに対応するステータスビットを記録したダーティ リージョン ログを維持します。ログ サブディスクは、DRL が有効になっているボリュームのダーティ リージョン ログを保存するために使用されます。DRL を伴うボリュームには、少なくともサブディスクがひとつ用意されます。そのサブディスクはそのボリュームのプレックスの1つに関連付けられます。

ボリュームにデータを書き込む直前に、ダーティ リージョン ログ内の該当するリージョンにダーティである印が付けられます。書き込みによって、以前はクリーンだったログ内のリージョンがダーティになる場合は、書き込み操作の開始前に、ログが同期的にディスクに書き込まれます。ミラーへの書き込み操作が完了すると、ログ内のリージョンが再度クリーンになります。システムを再起動する際にリカバリが必要な場合、Volume Manager はダーティ リージョン ログに書き込みがあったと記録されているリージョンだけを対象にします。

クラスタ環境においては、Volume Manager への DRL の実装の仕方は、通常の場合と少し違っています。後述の節では、その違いの一部を概説し、クラスタ環境における DRL の実装のいくつかの局面を説明します。

ログのフォーマットとサイズ

VxVM をクラスタ環境で使用しない場合は、クラスタ環境中のダーティ リージョン ログはミラー ボリューム内のログ サブディスクに格納されます。

VxVM のダーティ リージョン ログには、リカバリ マップとアクティブ マップが 1 つずつあります。クラスタ環境中のダーティ リージョン ログには、1 つのリカバリ マップと複数のアクティブ マップ（クラスタ内の各ノードごとに 1 つ）があります。VxVM とは異なり、クラスタ構成の場合、リカバリ マップはログの先頭に置かれています。

クラスタ環境中のダーティ リージョン ログのサイズは通常は VxVM のダーティ リージョン ログのサイズよりも大きくなります。その理由は、1 つのリカバリ マップに加え、すべてのノードの分のアクティブ マップを保持する必要があるからです。ダーティ リージョン ログ内の各マップのサイズは 1 ブロック以上です。ダーティ リージョン ログに十分なサイズが vxassist によって割り当てられます。

ログ サイズはボリューム サイズおよびノード数に応じて変わります。すべてのマップ（1 つのリカバリ マップに加え、各ノードごとに 1 つのアクティブ マップ）を収容できるだけの、十分な大きさが必要です。各マップには、ボリューム サイズ 2 ギガバイトごとに 1 ブロックが必要です。クラスタが 2 つのノードから成り、ボリューム サイズが 2 ギガバイトである場合、ログ サイズは 3 ブロック（各マップごとに 1 ブロック）で十分です。これが最低のログ サイズです。クラスタが 4 つのノードから成り、ボリューム サイズが 4 ギガバイトである場合、ログ サイズは 10 ブロックになります。

既存のクラスタにノードを追加するときは、既存の DRL ログを切り離して除去してから（`vxplex -o rm dis` を使用）、DRL を作成し直す（`vxassist addlog` を使用）必要があります。それによって、追加されたノードの分のマップを収容できるように、ログ サイズが拡大されます。

互換性

クラスタ環境の DRL ヘッダーにはクラスタに固有のマジック ナンバが追加されます。それ以外の点では、クラスタ環境のものも非クラスタ環境のものも、DRL ヘッダーは同じです。

VxVM のディスク グループ（およびそのボリューム）を共有ディスク グループとしてクラスタ環境にインポートすることが可能です。また、その逆も可能です。しかし、インポートされたディスク グループのダーティ リージョン ログは無効であるとみなされて全面的なリカバリが行われる可能性があります。

クラスタ サポートなしで VxVM によって共有ディスク グループがインポートされると、VxVM は共有ボリュームのログが無効であるとみなして、全面的なボリュームのリカバリを実行します。このリカバリが完了した後で、Volume Manager はクラスタ機能のダーティ リージョン ロギングを適用します。



Volume Manager のクラスタ機能は非共有 VxVM ボリューム上の DRL のリカバリを行えます。しかし、クラスタ サポートのある VxVM システムに VxVM ボリュームを移動して共有ボリュームとしてインポートすると、クラスタ内の全ノード分を収容するには、ダーティ リージョン ログが小さすぎる可能性があります。この場合、クラスタ機能はログを無効にし、全面的なリカバリを行います。同様に、2 ノード構成のクラスタから 4 ノード構成のクラスタへ DRL ボリュームを移動した場合も、ログ サイズが小さすぎるので、クラスタ機能によってボリュームの全面的なリカバリが行われます。どちらの場合も、システム管理者は十分な大きさの新しいログを割り当てる必要があります。

クラスタ環境における DRL の仕組み

クラスタ内のノードがいくつかクラッシュした場合は、その時点でそれらのノードで使用されていたすべてのボリュームを DRL がリカバリできる必要があります。最初にクラスタを起動したときに、アクティブ マップはすべてリカバリ マップに組み込まれます。これは volume start 処理の間に行われます。

クラッシュしたノード（「ダーティ」状態でクラスタから取り外されたノード）はそのままでは再びクラスタに戻すことはできません。まず、それらのアクティブ マップを、影響を受けたすべてのボリュームのリカバリ マップに組み込む必要があります。リカバリユーティリティはクラッシュしたノードのアクティブ マップをリカバリ マップと比較して、必要な更新を行います。それが済んだら、ノードをクラスタに再び追加して、ボリュームへの入出力を再開することができます（それによって、アクティブ マップが上書きされます）。その間に、ほかのノードは入出力を続行することができます。

VxVM のカーネルはどのノードがクラッシュしたのかを把握しています。ある時点でクラスタ内で複数のノードのリカバリが進行している場合、それぞれのリカバリおよびリカバリ マップの更新が相互に競合することがあります。そのために、VxVM のカーネルは DRL リカバリの状態の変化を追跡し、入出力処理が競合することを防止します。

マスタは各ボリュームごとの DRL リカバリ マップの更新の一時的な追跡を実行し、複数のユーティリティが同時にリカバリ マップを更新するのを防止します。

Volume Manager クラスタ機能のアップグレード

ローリング アップグレード機能によって、クラスタ全体をシャットダウンせずに、管理者はクラスタで実行中の Volume Manager のバージョンをアップグレードできます。新しいバージョンの Volume Manager をインストールしてクラスタ上で実行するには、システム管理者はクラスタからノードを1つずつ切り離し、アップグレードしてからクラスタにノードを再度追加します。これをクラスタ内の各ノードについて繰り返します。

リリース 3.1 以降の各 VxVM リリースには、クラスタ プロトコル バージョン番号が関連付けられています。この番号は、リリース番号とは異なります。クラスタ プロトコル バージョンは /etc/vx/volboot ファイルに格納されます。VxVM を新規にインストールすると、/etc/vx ディレクトリに volboot ファイルは作成されません。このファイルは vxdctl init によって作成され、サポートされる最新のクラスタ プロトコル バージョンが設定されます。

新しい VxVM リリースは、2 つのクラスタ プロトコル バージョンをサポートしています。古い方のバージョンは、既存の VxVM リリースに対応しています。このリリースには、機能および通信プロトコルの固定セットがあります。新しい方のバージョンは、新しい VxVM リリースに対応しています。このリリースには、新しい機能のセットがあります。新しい VxVM リリースの機能またはプロトコルに変更がない場合、クラスタ プロトコル バージョンは変更されません。たとえば、バグの修正や小さな変更の場合です。この場合、vxdctl upgrade を実行する必要はありません。

ローリング アップグレードの処理中は、各ノードをシャットダウンし、最新のクラスタ プロトコル バージョンを持つ VxVM リリースをインストールする必要があります。新しい VxVM リリースがインストールされているすべてのノードで、低いバージョンが継続して使用されます。新しいクラスタ プロトコル バージョンがインストールされているスレーブ ノードは、クラスタへの追加を試みます。新しいクラスタ プロトコル バージョンがマスタ ノードで使用されていない場合、マスタ ノードは追加を拒否し、現在のクラスタ プロトコル バージョンをスレーブ ノードに提供します。スレーブ ノードは、マスタ ノードによって提供されたクラスタ プロトコル バージョンを使用して、追加を再試行します。この時点で追加がエラーになる場合、マスタ ノードのクラスタ プロトコル バージョンは、スレーブの追加でサポートされるプロトコル バージョンの範囲外であることとなります。このような状況では、システム管理者は、サポートされている最新のクラスタ プロトコル バージョンになるまで、中間の各 VxVM リリースによってクラスタをアップグレードしていく必要があります。

すべてのノードが最新のクラスタ プロトコル バージョンにアップグレードされると、新しい機能が使用できるようになります。

すべてのノードに新しいリリースがインストールされたら、新しいクラスタ プロトコル バージョンに切り替えるために、マスタ ノード上で vxdctl upgrade コマンドを実行する必要があります。



クラスタ関係の Volume Manager のユーティリティとデーモン

クラスタ環境の Volume Manager で使用するために、下記のユーティリティ/デーモンが開発または修正されました。

- ◆ vxclust
- ◆ vxclustadm
- ◆ vxconfigd
- ◆ vxdg
- ◆ vxdisk
- ◆ vxrecover
- ◆ vxdctl
- ◆ vxstat

上記のユーティリティのそれぞれに節を割いて、クラスタ環境でのそれらの使用方法を以下に説明します。各ユーティリティの詳細については、対応するマニュアル ページを参照してください。

vxclust

クラスタの再設定が行われるたびに、その時点でクラスタ内にある各ノードで、既に設定済みの手順が実行され、**vxclust** ユーティリティが起動されます。クラスタ マネージャはすべてのノード上で同時に同じ手順が実行されるようにします。すべてのノード上で前の手順が完了しないと、次の手順は開始されません。再設定の手順ごとに、**vxclust** は Volume Manager のクラスタ機能が次に何を行うべきかを決定します。次の措置を VxVM に通知すると、**vxclust** はその結果（正常終了、異常終了、再試行）が出るのを待ち、それをクラスタ マネージャに通知します。

注：

- ◆ ほとんどの Volume Manager のコマンドには、スーパーユーザの特権が必要です。
- ◆ **vxclust** は SunCluster™ と共にクラスタ マネージャとして機能します。

vxclust からの要求にノードが一定のタイムアウト期限内に応答しないと、そのノードはアボート処理され、状況に応じて、**vxclust** は再設定を再試行するか断念するかを決定します。問題の原因がノード内の修正不可能なエラーである場合、**vxclust** は処理の続行を断念します。ほかのノードが取り外されたためにあるノードでの処理が完了できない場合、そのノードでの処理はタイムアウトとなります。その場合、別のノードが取り外されることを期待して、**vxclust** は再設定を要求します。ほかに取り外されるノードがないと、**vxclust** はローカル ノードを取り外します。

再設定の手順でエラーが発生すると、`vxclust` はクラスタ マネージャにエラーを返しません。クラスタ マネージャはノードをアボートさせる可能性があります。これにより、そのノードは直ちにクラスタから取り外されます。共有ディスクに対して実行中であった入出力があれば、エラーとされ、共有ディスクへのアクセスは停止されます。

クラスタに変更があったことを通知されると、`vxclust` は取るべき措置を決定します。前のマスタ ノードが故障したために新しいマスタ ノードが必要になった場合には、`vxclust` は新しいマスタにするノードを決定します。

vxclustadm

注 `vxclustadm` は VCS と共に動作し、単独では動作しません。

`vxclustadm` コマンドは、クラスタ内のノード上で、VxVM のクラスタ機能をアクティブ化または非アクティブ化するために使用します。このコマンドは、VCS クラスタの起動時およびシャットダウン時に、オンラインおよびオフライン スクリプトで呼び出されます。

`startnode` オプションを使用すると、クラスタ設定情報が VxVM カーネルに渡されます。このコマンドに対して、カーネルおよび設定デーモン `vxconfigd` が初期化を実行します。

`stopnode` オプションを使用すると、ノード上のクラスタ機能を停止できます。この処理は、未完了の入出力がすべて終了し、すべてのアプリケーションが共有ボリュームを閉じるまで待機します。`abortnode` オプションを使用すると、ノード上のクラスタ処理をアボートできます。これは緊急のシャットダウンであり、共有ボリューム上の未完了の入出力をすべて中断します。

`nodestate` オプションを使用すると、クラスタ内のノードの状態を判定することができます。

詳細については、`vxclustadm` (1M) マニュアル ページを参照してください。

vxconfigd

Volume Manager 設定デーモン `vxconfigd` は、VxVM オブジェクトの設定を維持します。SunCluster の配下または VCS 実行時のカーネル配下で、`vxconfigd` は `vxclust` ユーティリティからクラスタ関連の命令を受け取ります。各ノード上に `vxconfigd` のコピーがそれぞれ置かれます。それらのコピーはネットワーク機能を通じて相互に通信します。クラスタ内の各ノードに関して、Volume Manager のユーティリティは該当のノード上で稼働している `vxconfigd` と情報を交換します。それらのユーティリティがほかのノード上の `vxconfigd` デーモンに接続しようとすることはありません。クラスタの起動時、`vxclust` (SunCluster) またはカーネル (VCS) は、`vxconfigd` にクラスタ処理の開始を指示し、マスタ ノードとスレーブ ノードのいずれであるかを通知します。



ノードがクラスタ処理のために初期化されると、ノードがクラスタに追加されることが `vxconfigd` に通知され、次の情報（クラスタ マネージャの設定データベースに基づく）が示されます。

- ◆ クラスタ ID
- ◆ ノード ID
- ◆ マスタ ノード ID
- ◆ ノードの役割
- ◆ 各ノード上の `vxconfigd` のネットワーク アドレス

マスタ ノード上では、`vxconfigd` は共有の構成を設定（共有ディスク グループをインポート）し、スレーブを追加する準備が整ったことを `vxclust` に通知します。

スレーブ ノード上では、スレーブ ノードがクラスタを追加できるようになると、`vxconfigd` は通知を受けます。スレーブ ノードがクラスタに追加されると、`vxconfigd` と Volume Manager のカーネルはそれぞれマスタ ノード上の `vxconfigd` および Volume Manager と連絡を取り合って、共有の構成を設定します。

ノードがクラスタから切り離されると、`vxconfigd` デーモンはほかの全ノード上のカーネルに通知します。それを受けて、マスタ ノードは必要な整理処理を行います。マスタ ノードがクラスタから切り離されると、カーネルは新しいマスタ ノードを選択し、すべてのノード上の `vxconfigd` デーモンは新しいマスタ ノードが選択されたことを通知されます。

`vxconfigd` はボリュームの再設定にも関与します。ボリュームの再設定における `vxconfigd` の役割については、「ボリュームの再設定 (199 ページ)」を参照してください。

vxconfigd のリカバリ

Volume Manager の `vxconfigd` デーモンは任意の時点で停止したり再起動したりすることができます。 `vxconfigd` が停止している間は、ボリュームの再設定を行うことはできません。 `vxconfigd` が再起動するまで、ほかのノードをクラスタに追加することはできません。クラスタ環境においては、スレーブ ノード上の `vxconfigd` デーモンは常にマスタ ノード上の `vxconfigd` デーモンに接続されています。したがって、どのノード上の `vxconfigd` デーモンも停止させることは望ましくありません。

何らかの理由により `vxconfigd` が停止された場合、それがどのノードで発生したかによって、取られる措置が異なります。

- ◆ スレーブ ノード上の `vxconfigd` が停止された場合、マスタ側では何の措置も取られません。スレーブ上で `vxconfigd` が再起動したとき、スレーブの `vxconfigd` はマスタに再接続して共有の構成に関する情報を再び入手しようと試みます。（共有の構成に関するカーネルのビューは影響を受けません。同様に、共有ディスクへのアクセスも影響を受けません。マスタへの再接続に成功するまで、スレーブ上の `vxconfigd` には共有の構成に関する情報がほとんどありません。したがって、共有の構成を表示

したり修正したりしようとする試みは、いずれもエラーとなります。特に、共有ディスクグループを一覧表示すると (vxdg list を使用)、一覧のディスクグループの状態には disabled と表示されます。再追加が完了すると、enabled と表示されます。

- ◆ マスタ ノード上の vxconfigd が停止された場合、スレーブ上の vxconfigd は定期的にマスタへの再接続を試みます。マスタ上で vxconfigd が再起動するまで、その試みは成功しません。この場合、共有の構成に関するスレーブ上の vxconfigd の情報は失われていません。したがって、共有の構成に関して表示される内容は正確です。
- ◆ マスタ ノードとスレーブ ノードの両方の vxconfigd が停止された場合、両方のノード上で vxconfigd を起動して、両者を再接続する必要があります。それが完了するまで、共有の構成に関してスレーブ上に表示される内容は正しくありません。

vxconfigd がノード上で停止したことが、vxclust (SunCluster) またはカーネル (VCS) に通知されると、vxconfigd が再起動されます。

注 VxVM において vxconfigd に対して `-r reset` オプションを適用すると、vxconfigd が再起動され、すべての状態が初期化されます。ノードがクラスタ内にある間は、このオプションは利用できません。なぜならば、クラスタ情報が失われるからです。そのような状況のもとでこのオプションを使用しても、vxconfigd は起動しません。

vxdg

vxdg ユーティリティは Volume Manager のディスクグループを管理する働きをします。vxdg を使用して、ディスクグループをクラスタで共有するように指定することができます。オプションの `-s` を指定すると、ディスクグループを「共有」として、初期化またはインポートすることができます。

クラスタ ソフトウェアを実行してクラスタの設定が済んでいれば、下記のコマンドを使用して、共有ディスクグループを作成することができます。

```
vxdg -s init ディスクグループ [メディア名 =] アクセス名
```

ここで、**ディスクグループ**はディスクグループ名前、**メディア名**はディスクを管理するための名前、**アクセス名**はディスクアクセス名 (またはデバイス名) です。

vxdg `-s import` を使用すると、ディスクグループを「共有」するものとしてインポートすることができます。クラスタ ソフトウェアを実行する前にディスクグループを設定しておく、下記のコマンドを使用して、ディスクグループをクラスタ環境にインポートすることができます。

```
vxdg -s import ディスクグループ
```

ここで、**ディスクグループ**はディスクグループの名前または ID を表します。以降、クラスタを再起動する際に、ディスクグループは自動的に共有するものとしてインポートされます。このコマンドを実行する前に、ディスクグループをデポートしておく (vxdg `deport` ディスクグループを使用) 必要があることに注意してください。



共有ディスク グループを専有ディスク グループに変換することができます。そのためには、`vxvg deport` を使用してデポートしてから、`vxvg import` ディスクグループを使用してインポートし直します。

注 システムはディスクが共有されているかどうか認識できません。複数のシステムからディスクにアクセスできるときにデータの整合性を保つためには、ディスクをディスク グループに追加する際に、システム管理者は設定を慎重に正しく行う必要があります。物理的に共有されていないディスクをシステム管理者が共有ディスクグループに追加しようとすると、そのディスクにアクセス可能なノードがクラスタ内にそれだけしかないときに、**Volume Manager** は追加を認めます。しかし、ほかのノードをそのクラスタに追加することはできません。さらに、システム管理者が同一のディスクを同時に2つのノード上の異なるディスクグループに追加しようとすると、結果はどうなるか分かりません。したがって、構成の設定はすべて、1つのノードだけを対象にして行ってください。

`vxvg` には強制オプション (`-f`) があります。このオプションを使用すると、ディスクグループを強制的にインポートしたり、ディスクをディスクグループに強制的に追加したりすることができます。

注 強制オプション (`-f`) を使用するときは、注意が必要です。システム管理者は、可能性のある結果について熟知している場合にのみ、このオプションを使用するようにしてください。

クラスタを再起動したときに、ディスクグループの自動インポートが **VxVM** によって拒否されることがあります。その理由として、下記のいずれかが考えられます。

- ◆ 該当のディスクグループ内のディスクにハードウェア的なエラーがあったために、アクセスできなくなっている。この場合、システム管理者は下記のように強制オプションを指定して、そのディスクグループを再インポートすることができます。

```
# vxvg -s -f import ディスクグループ
```

- ◆ ディスクグループ内のディスクが接続されているノードのいくつかが現在はクラスタから取り外されているため、ディスクグループ内のすべてのディスクにはアクセスできない。このような場合は、強制的にインポートすることは危険なので、実施すべきではありません（ミラーの整合性が崩れる可能性があります）。

あるディスクが **VxVM** によって既存のディスクグループに追加されない場合（ディスクグループ内のほかのディスクが接続されているノードと同じノードにそのディスクが接続されていないため）、システム管理者は下記のように、そのディスクを強制的に追加することができます。

```
# vxvg -f adddisk -g ディスクグループ [medianame=] アクセス名
```

`vxvg` を使用して共有ディスクグループを一覧表示することもできます。下記のコマンドを実行すると、各ディスクグループごとに1行の情報が表示されます。

```
# vxvg list
```



次のような結果が表示されます。

```

NAME          STATE          ID
rootdg        enabled        774215886.1025.teal
group2        enabled,shared 774575420.1170.teal
group1        enabled,shared 774222028.1090.teal

```

共有ディスク グループの "STATE" 欄に `shared` と表示されます。

下記のコマンドを実行すると、各共有ディスク グループごとに 1 行の情報が表示されます。

```
# vxdg -s list
```

次のような結果が表示されます。

```

NAME          STATE          ID
group2        enabled,shared 774575420.1170.teal
group1        enabled,shared 774222028.1090.teal

```

下記のコマンドを実行すると、指定したディスク グループに関する情報が表示されます。その中に、共有されているか否かの表示も含まれます。

```
# vxdg list ディスクグループ
```

ここで、**ディスクグループ**はディスク グループの名前を表します。

マスタ上のディスク グループ `group1` を対象にしてこのコマンドを実行すると、次のような結果が表示されます。

```

Group:      group1
dgid:       774222028.1090.teal
import-id:  32768.1749
flags:      shared
copies:     nconfig=default nlog=default
config:     seqno=0.1976 permlen=1456 free=1448 templen=6 loglen=220
config disk c1t0d0s2 copy 1 len=1456 state=clean online
config disk c1t1d0s2 copy 1 len=1456 state=clean online
log disk   c1t0d0s2 copy 1 len=220
log disk   c1t1d0s2 copy 1 len=220

```

`flags`: フィールドに `shared` が設定されていることに注意してください。このコマンドをスレーブに適用した場合、出力は少し異なります。



vxdisk

vxdisk ユーティリティは Volume Manager のディスクを管理する働きをします。このユーティリティを使用して、クラスタで共有するディスクグループに、あるディスクが属しているかどうかを判定することができます。たとえば、次のように入力します。

```
# vxdisk list アクセス名
```

ここで、**アクセス名**はディスク アクセス名（またはデバイス名）を表します。

デバイス c1t0d0s2 にこのコマンドを適用すると、次のような結果が表示されます。

```
Device:      c1t0d0s2
devicetag:   c1t0d0
type:        sliced
clusterid:   cvm
disk:        name=disk01 id=774215890.1035.teal
group:       name=group1 id=774222028.1090.teal
flags:       online ready autoconfig shared imported
pubpaths:    block=/dev/dsk/c1t0d0s4 char=/dev/rdisk/c1t0d0s4
privpaths:   block=/dev/dsk/c1t0d0s3 char=/dev/rdisk/c1t0d0s3
version:     3.1
iosize:      min=512 (bytes) max=248 (blocks)
public:      slice=4 offset=0 len=2050272
private:     slice=3 offset=1 len=2015
update:      time=778564769 seqno=0.1614
headers:     0 248
configs:     count=1 len=1456
logs:        count=1 len=220
Defined regions:
  config  priv 000017-000247[000231]:copy=01 offset=000000 enabled
  config  priv 000249-001473[001225]:copy=01 offset=000231 enabled
  log     priv 001474-001693[000220]:copy=01 offset=000000 enabled
```

clusterid: フィールドに cvm(クラスタ名)、flags: フィールドに shared エントリが設定されていることに注意してください。ノードが追加されていない場合は、flags: フィールドには imported フラグではなく autoimport フラグが設定されます。

vxrecover

vxrecover ユーティリティはディスクの交換後にプレックスとボリュームをリカバリする働きをします。

ノードをクラスタから取り外したときに、いくつかのミラーが不整合な状態になることがあります。vxrecover ユーティリティはそのような状態にあるすべてのボリュームのリカバリを行います。-c オプションを指定すると、vxrecover はクラスタで共有するディスクグループ内のすべてのボリュームをリカバリします。必要に応じて、vxclust によって、vxrecover -c が自動的に呼び出されます。

注 vxrecover が動いている間は、システムのパフォーマンスがいくらか低下することがあります。

vxctl

vxctl ユーティリティはボリューム設定デーモンである vxconfigd の一部の局面を管理する働きをします。-c オプションを指定すると、クラスタ情報を要求することができます。vxconfigd が有効になっているか稼働しているかを判定するために、vxctl を下記のように使用することができます。

```
# vxctl -c mode
```

状況によって多少の違いはありますが、次のような結果が表示されます。

```
mode:enabled:cluster active - MASTER
mode:enabled:cluster active - SLAVE
mode:enabled:cluster inactive
mode:enabled:cluster active - role not set
```

注 vxconfigd が無効になっていると、クラスタ情報は表示されません。

vxctl に関する詳細の説明については、vxctl (1M) マニュアルページを参照してください。

vxctl は、クラスタ プロトコルバージョンおよびクラスタ プロトコルの範囲を表示します。クラスタ内のすべてのノードが新しいクラスタ プロトコルでアップグレードされたら、クラスタ全体を次のコマンドでアップグレードします。

```
# vxctl upgrade
```

vxctl protocolversion コマンドを使用すると、既存のクラスタ プロトコルバージョンをチェックできます。たとえば、次のように入力します。

```
# vxctl protocolversion
```

```
Cluster running at protocol 10
```



`vxdctl protocolrange` コマンドは、現在の VxVM リリースでサポートされているクラスタ プロトコル バージョンの最大値と最小値を表示します。たとえば、次のように入力します。

```
# vxdctl protocolrange
minprotoversion:10, maxprotoversion: 20
```

`vxdctl list` コマンドは、ノード上で実行しているクラスタ プロトコル バージョンを表示します。`vxdctl list` コマンドを実行すると、次のような結果が表示されます。

```
Volboot file
version: 3/1
seqno: 0.19
cluster protocol version: 20
hostid:giga
entries:
```

`vxdctl support` コマンドは、ノードおよび現在のプロトコルバージョンでサポートされているプロトコルバージョンの最大値と最小値を表示します。`vxdctl support` コマンドを実行すると、次のような結果が表示されます。

```
Support information:
vold_vrsn: 11
dg_minimum: 60
dg_maximum: 70
kernel: 10
protocol_minimum: 10
protocol_maximum: 20
protocol_current: 20
```

vxstat

vxstat は指定されたオブジェクトに関する統計を返すユーティリティです。クラスタ環境においては、vxstat はクラスタ内のすべてのノードから統計を収集します。指定したオブジェクトに関するすべてのノードでの合計使用量の統計が返されます。ローカルオブジェクトを指定すると、そのノードでの使用量が返されます。

vxstat では、オプションとして、ノードのサブセットを指定することができます。

```
# vxstat -g ディスクグループ -n ノード [, ノード ...]
```

ノードには整数を指定します。複数のノードをコンマで区切って列挙すると、それらのノードでの統計値の合計が表示されます。

次の例では、ノード 2、ボリューム vol1 の統計を入手します。

```
# vxstat -g group1 -n 2 vol1
```

次のような結果が表示されます。

		OPERATIONS		BLOCKS		AVG TIME(ms)	
TYP	NAME	READ	WRITE	READ	WRITE	READ	WRITE
vol	vol1	2421	0	600000	0	99.0	0.0

vxstat を使用して、クラスタ全体の統計を表示することもできます。それには下記のように指定します。

```
# vxstat -b
```

すべてのノードの統計が合算されます。たとえば、ノード 1 で 100 回の入出力が行われ、ノード 2 では 200 回の入出力が行われた場合、上記のコマンドの結果は 300 となります。

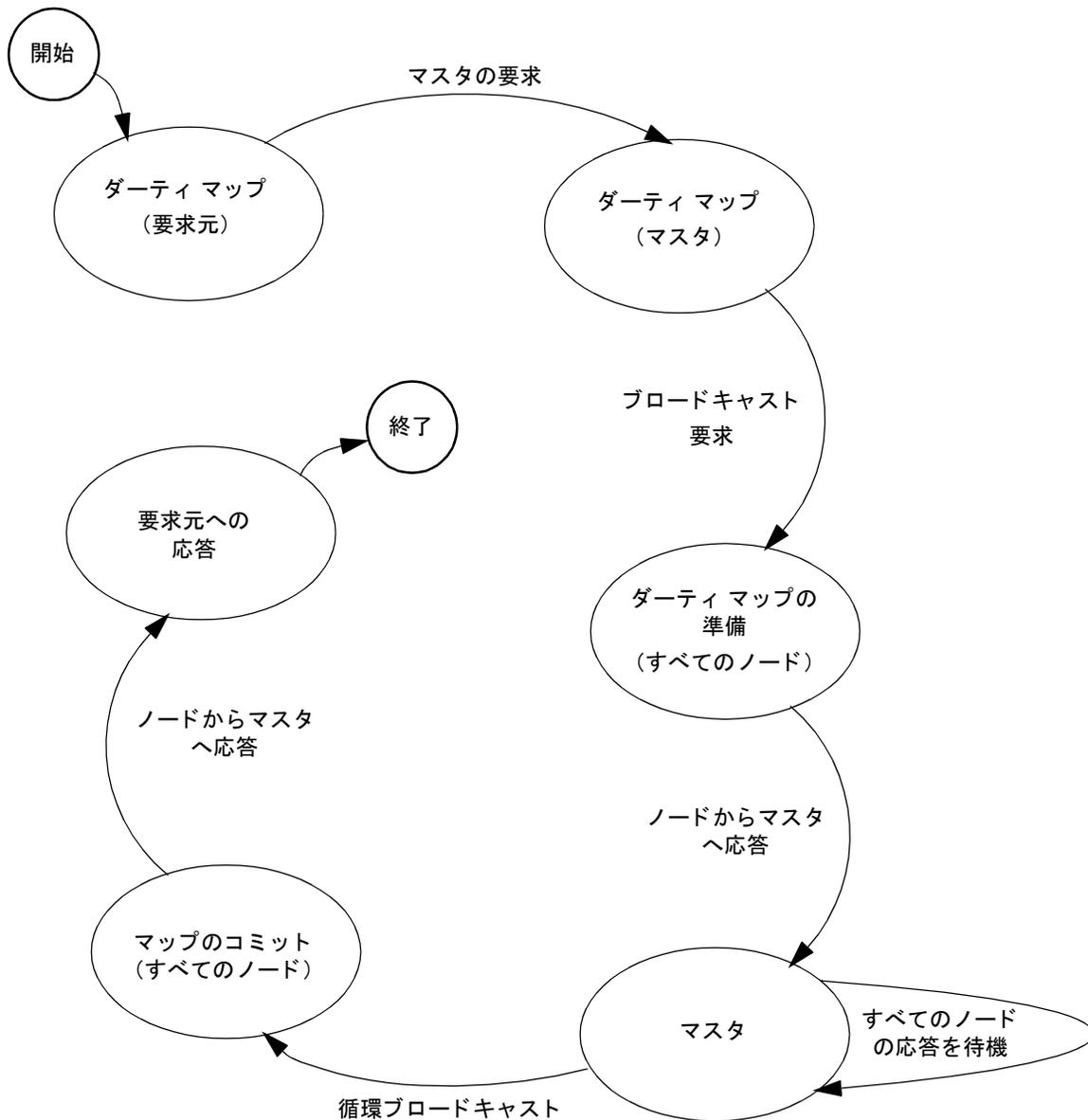
クラスタ処理と高速ミラー再同期 (FMR)

高速ミラー再同期 (FMR) は共有ボリュームでサポートされています。更新マップ (FMR マップ) はクラスタ内で配布されます。持続性のないマップが存在しても、クラスタ環境ではさほど問題にはなりません。これは、クラスタ内で 1 つのノードさえ FMR マップを使用可能であればよいからです。1 つのノードに障害が発生しても FMR マップが失われることはありません。

マップの更新は、マスタ ノードの補助を受けて適用されます。マスタ ノードからすべてのノードに更新が配布されると、すべての更新が全ノードに適用されるか、どのノードにも適用されないかのいずれかになります。更新が必ず適用されるように、マスタ ノードは 2 段階のコミットメントを組み合わせています。218 ページの図 25「ビットマップのクラスタ処理」を参照してください。



図 25. ビットマップのクラスタ処理



はじめに

VERITAS Volume Manager は、ディスク障害からシステムを保護およびリカバリする役に立ちます。この章では、リカバリ手順と、ディスク障害によりデータが失われたりシステムへアクセスできなくなるなどの問題を回避するために役立つ情報について説明します。また、プレックスやボリュームの状態についても説明します。

システムの保護方法については、「Volume Manager の初期化 (32 ページ)」および「システムの設定 (35 ページ)」を参照してください。

この章では、次のトピックについて説明します。

- ◆ UNIX のブート プロセス
- ◆ ルート (/)、swap、および usr 構成の設定
- ◆ 障害とリカバリ手順
- ◆ ホットリロケーションおよびブート ディスク障害
- ◆ ブート ディスクの再追加と交換
- ◆ ディスクの再接続
- ◆ 再インストールによるリカバリ
- ◆ プレックスとボリュームの状態
- ◆ RAID-5 ボリュームのリカバリ
- ◆ RAID-5 のさまざまな処理



UNIX のブート プロセス

Sun SPARC システムでは、ファーム ウェアで使用されている不揮発性ストレージ領域に autoboot フラグが設定されていないと、boot コマンドのプロンプトが表示されます。古い PROM を搭載しているマシンでは、新しいバージョン V2 および V3 の PROM を搭載したマシンとは異なるプロンプトが表示されます。これらの新しいバージョンの PROM は、OpenBoot PROM (OBP) と呼ばれます。この 2 つのタイプの PROM の場合、boot コマンドには次のように別の構文を用います。

```
# ok boot [OBP 名] [ファイル名] [ブートフラグ]
```

OBP 名には、OBP (Open Boot PROM) の場所を指定します。たとえば、デスクトップ SPARC システムでは、次のように指定します。

```
# /sbus/esp@0,800000/sd@3,0:a
```

この場所は、SCSI バス上のターゲット 3、論理ユニット番号 (LUN) 0 に配置されている SCSI ディスク (sd) を示しています。ESP ホスト アダプタ (esp) はスロット 0 に差し込まれています。

注 Volume Manager では、ブート ディスクのエイリアス名を使用できます。このエイリアスには、Volume Manager が提供する形式 (vx-rootdisk、vx-disk01 など) またはオペレーティング システムが提供する形式 (disk1) のいずれも使用できます。OpenBoot の OK プロンプトに応じて devalias コマンドを使用すると、有効なブート可能デバイスの一覧を表示できます。

ファイル名には、ブート プログラムに対するスタンドアロン プログラムの名前を指定します。デフォルトでは、/kernel/unix をルート パーティションからブートします。コマンド ラインから、別のプログラム (/stand/diag など) を指定できます。ファームウェアのバージョンによっては、デフォルトのファイル名をシステムの不揮発性ストレージ領域に保存しておくことができます。

ブート プログラムでは、-a フラグは「ask me」と解釈され、ブートするスタンドアロン プログラムの名前を入力するようにプロンプトが表示されます。-a フラグは、そのスタンドアロン プログラムに渡されます。

注 ルート機能を有効にした状態で Volume Manager を実行しているシステムは、-a フラグで指されるデフォルトではブートできません。boot -a コマンドへの正しい応答の仕方については、「/etc/system のコピーが使用できる場合 (229 ページ)」を参照してください。

ブート プログラムはフラグを解釈しません。ブート プログラムは、すべてのブート フラグをファイル名で識別できるファイルに送ります。ブートします。デフォルトのスタンドアロン プログラム /kernel/unix に有効なオプションについては、kernel (1) および kadb (1M) マニュアル ページを参照してください。

障害発生後のブート

ルート ディスクをミラーリングしている場合、プライマリ ブート ディスクに障害が発生しても、代替ブート ディスクを使用してシステムをブートできます。プライマリ ブート ディスクに障害が発生した後にシステムをブートする手順は下記の通りです。

1. OpenBootコマンド プロンプトに応じて`devalias`コマンドを実行し、エイリアスを付けられている Volume Manager ディスクを確認します。

ルート ディスクのミラーに適しているディスクが `vx メディア名` と表示されます。メディア名は、ルート ファイル システムの候補が含まれているディスクのディスク メディア名を示しています。

2. 次のコマンドを入力します。

```
# ok boot 別名
```

エイリアス名には、選択したディスクのエイリアス名を指定します。

選択したディスクに含まれているルート ミラーが無効な場合、`vxconfigd` により、ミラーが使用できないことを通知するエラー メッセージと、無効ではないブート可能な代替ディスクの一覧が表示されます。

ルート (/)、swap、および usr 構成の設定

インストール中に、ルート、swap、およびusrの各ファイル システムの構成を設定することができます。次の場合が考えられます。

- ◆ `usr` は、ルートの下ディレクトリとして設定され、単独のパーティションが割り当てられることはありません。この場合、ルート ディスクがカプセル化されて Volume Manager の管理下に置かれると、`usr` は `rootvol` ボリュームの一部になります。
- ◆ `usr` は、ルート ディスク上の単独のパーティションに設定されます。この場合、単独のボリュームが `usr` パーティション用に作成されます。`vxmirror` は、デスティネーション ディスク上に `usr` ボリュームをミラーリングします。
- ◆ `usr` は、ルート ディスク上以外の単独のパーティションに設定されます。この場合、Volume Manager を使用してそのディスクをカプセル化した場合のみ、ボリュームが `usr` パーティション用に作成されます。ただし、ルート ディスクをカプセル化し、ルート ボリュームをミラーリングしていても、何らかの理由で `usr` パーティションにアクセスできなくなると役には立ちません。システムの可用性を最大限に維持するためには、`usr` パーティションが含まれているディスクとルート ディスクの両方をカプセル化し、`usr`、`rootvol`、および `swapvol` ボリュームのミラーを作成しておくことをお勧めします。

`rootvol` ボリュームは、`rootdg` ディスク グループに存在している必要があります。`rootvol`および`usr`ボリュームの制限事項については、「ブート時のボリュームの制限(62 ページ)」を参照してください。



Volume Manager では、任意のディスクに swap パーティションを設定でき、ブートプロセスの初期の段階でスワップ領域を必要としません。デフォルトでは、Volume Manager のインストール時に、選択したルート ディスク上のパーティション 0 がルート パーティション、パーティション 1 が swap パーティションとして選択されます。ただし、ルート ディスク上にないパーティションに swap パーティションを置くこともできます。この場合は、そのディスクをカプセル化し、swap ボリュームのミラーを作成しておくことをお勧めします。この処理を行わないと、swap パーティションに発生した障害が原因でシステムが破壊される恐れがあります。その場合でもシステムをブートすることはできますが、swapvol ボリュームのミラーを作成しておくこと、システム障害を防ぐことができます。

ボリューム上のルート (/) ファイルシステムまたは /usr ファイルシステムの修復

ルート (/) ファイルシステムまたは /usr ファイルシステムが使用できなくなった場合、ネットワークにマウントされているルート ファイルシステムまたは有効なバックアップからブートすることをお勧めします。使用するバックアップには、ルート ディスク上の関連するすべてのファイルシステム パーティションが含まれている必要があります。また、ルート ディスクをカプセル化する前に、ルート ディスクのパーティションテーブルを印刷しておくことも必要です。

ルート ファイルシステムまたは /usr ファイルシステムがミラー ボリューム上に定義されている場合、この処理はさらに複雑になります。いずれか 1 つのミラーの基盤であるパーティションが変更された場合、後から Volume Manager を起動したときにミラーが正しく同期していると想定され、ファイルシステムが破壊されることがあります。

この問題を回避するには、次の 2 つの方法があります。

- ◆ 最も簡単な方法は、ルート ファイルシステムまたは /usr ファイルシステムのプレックスを 1 つマウントし、修復してからマウント解除します。次に、dd を使用して、修正されたプレックスをほかのすべてのプレックスにコピーします。ただし、この方法ではエラーを起こす可能性が高くなります。
- ◆ もう 1 つの方法は、システムを有効なバックアップ テープからリストアすることです。その手順は次の節で説明します。この手順では、オペレーティングシステムを CD-ROM からインストールする必要はありません。

次の手順で、ルート ディスクとミラーの両方が破損した場合の、簡単で、効率よく、信頼できるリカバリ方法を説明します。

バックアップテープからの Volume Manager ルート ディスク (ルート ミラー) のリカバリ

この手順では、次の条件を前提とします。

- ◆ 元の Volume Manager ルート ディスク上の全ファイル システムの完全な最新バックアップがある。
- ◆ 元のブート ディスクが物理的に破損している場合、それに代わる新しいブート ディスクがインストールされている。

この処理では、Volume Manager ルート ディスクをインストールし直す必要があります。再インストールに関係のないディスク上のデータの損失を防ぐため、再インストール処理では、Volume Manager ルート ディスクのみを扱うようにしてください。

インストールの自動オプションの中には、システム管理者による確認を必要とすることなく、ルート ディスク以外のディスクにアクセスするものがあります。そのため、処理を開始する前に、ボリュームを含んでいるほかのすべてのディスクをシステムから切り離しておくことをお勧めします。切り離しておけば、ほかのディスクは再インストールの影響を受けずに済みます。処理の完了後、これらのディスクを再接続します。

ルート ディスクのリカバリ

次に説明する手順では、新しいブート ディスクを c0t0d0、リカバリする必要のある / ファイルシステムおよび /usr ファイルシステムをそれぞれ s0 および s6 と想定します。

1. オペレーティング システムを CD-ROM からブートします。
2. `format` コマンドを使用して、元のブート ディスク上にあったファイルシステムを収めるために、新しいブート ディスク (c0t0d0) 上に同じパーティションを作成します。

注 ファイルシステムをリストアする必要があるので、最大7つのパーティションを作成できます。また、このディスクを一時的に再カプセル化するため、ディスク上に専有領域と共有領域の 2 つのパーティションが必要となり、作成できるパーティションは5つだけとなります。

3. /dev/rdisk/c0t0d0s0 を /a/root 上にマウントします。ルート ファイル システムをテープからリストアします。installboot を使用して、ブートブロック デバイスを /a/root 上にインストールします。
4. /dev/rdisk/c0t0d0s6 を /a/usr 上にマウントします。/usr ファイル システムをテープからリストアします。
5. リストアしたルート ファイルを次のように変更します。
 - ◆ touch を使用して /a/root/etc/vx/reconfig.d/state.d/install-db を作成します。
 - ◆ /a/root/etc/system の中の次の 2 行を削除します。


```
rootdev:/pseudo/vxio@0:0
set vxio:vol_rootdev_is_volume=1
```
 - ◆ /a/root/etc/vfstab の Volume Manager ボリューム デバイスのエントリ /dev/vx/dsk を標準のディスク デバイス /dev/dsk/c0t0d0s0 および /dev/dsk/c0t0d0s6 で置き換えます。
6. システムを新しいブート ディスクからリブートします。リブートすると、ブート ディスクは Volume Manager がインストールされていないと判断します。



これに続く手順は、ルート ディスク ミラーが古い rootdg にあるかどうかによって異なります。

- ◆ 古い rootdg 内にルート ディスク ミラーとして使用されないディスクがある場合は、手順 7 に進みます。
 - ◆ 古い rootdg 内にルート ディスク ミラーだけがある場合は、手順 8 に進みます。
7. 古い rootdg 内にルート ディスク ミラーとして使用されないディスクがある場合は、次の手順を実行して、Volume Manager で障害が発生していると認識されるブート ディスクを古い rootdg から外し、新しいブート ディスクを設定します。

- a. インストール処理で扱われたファイルで、不要になったものを削除します。次のように入力します。

```
rm -r /etc/vx/reconfig.d/state.d/install-db
```

- b. VXVM 入出力デーモンを起動します。次のように入力します。

```
# vxiod set 10
```

- c. VXVM 設定デーモンを無効モードで起動します。次のように入力します。

```
# vxconfigd -m disable
```

- d. vxconfigd デーモンを初期化します。次のように入力します。

```
# vxdctl init
```

- e. vxconfigd を有効にします。次のように入力します。

```
# vxdctl enable
```

上記の手順によって、VxVM Volume Manager で障害が発生していると認識されるルート ディスクが古い rootdg から外されます。

- f. vxedit コマンド（または Volume Manager Storage Administrator）を使用して、古いルート ディスク ボリュームとルート ディスク自体を削除します。
- g. 次に、vxdiskadm コマンドを使用して、新しいブート ディスクをカプセル化し、ルート ディスクのミラーとして使用するディスクを初期化します。必要なリブートを実行後、ルート ディスクをルート ディスク ミラーにミラーリングします。
8. 古い rootdg 内にルート ディスク ミラーだけがある場合は、次の手順を実行します。

vxinstall コマンドを実行して、新しいルート ディスクをカプセル化し、ルート ディスク ミラーを初期化します。

必要なリブートを実行後、ルート ディスクをルート ディスク ミラーにミラーリングします。

ルート ファイル システムのバックアップとリストア

破損した場合にリストアできるように、ルート ファイル システムをバックアップしておくことをお勧めします。

ufs ファイル システムを使用している場合は、次のコマンドを使用して、ルート ファイル システムをバックアップできます。

```
# /usr/lib/fs/ufs/ufsdump [ ダンプ オプション ] /dev/vx/rdsk/rootvol
```

これで、障害が発生した後でルート ファイル システムをリストアできます。次の手順に従います。

1. CD-ROM またはネットワークにマウントされているルート ファイル システムからブートしてから、Volume Manager を起動します（「ボリューム上のルート (/) ファイル システムまたは /usr ファイル システムの修復 (222 ページ)」を参照）。
2. 次のコマンドを使用して、ルート ファイル システムをマウントしてリストアします。

```
# newfs /dev/vx/rdsk/rootvol
```

```
# mount /dev/vx/dsk/rootvol /mnt
```

```
# cd /mnt
```

```
# /usr/lib/fs/ufs/ufsrestore [ リストア オプション ]
```



障害とリカバリ手順

システムがブートできなくなる障害にはいくつかの種類がありますが、どの場合も同じ基本的な手順でシステムをリカバリできます。システムをブートできなくなった場合、まず画面に表示された内容から障害を特定し、問題の解決を試みてください（誤って電源を切ってしまったドライブの電源を入れるなど）。解決できないような問題（ブート ディスク上のデータ エラーなど）の場合は、障害の修復または障害が発生したディスクの交換を行えるように、代替ブート ディスク（ルート ボリュームのミラーを含む）からシステムをブートします。

この節では、発生する可能性がある障害と、その対処法について説明します。

UNIX パーティションでの障害

ブート プログラムが読み込まれると、ブート プログラムは通常の UNIX パーティション情報を使用してブート ディスクへのアクセスを試みます。この情報に障害が発生している場合は、ブート プログラムで次のエラーが発生します。

```
File just loaded does not appear to be executable
```

ブート時にこのメッセージが表示された場合は、システムを代替ブート ディスクからブートする必要があります。ブート中は、ほとんどのディスク ドライバで、障害が発生しているディスク上の無効な UNIX パーティション情報に関するエラーがコンソールに表示されます。次のようなメッセージが表示されます。

```
WARNING:unable to read label
```

```
WARNING:corrupt label_sdo
```

このメッセージは、障害の原因がディスク パーティションが無効であることを示しています。「障害を起こしたブート ディスクの再追加 (233 ページ)」の説明を参照して、ディスクを再追加することができます。ただし、再追加に失敗した場合は、「障害を起こしたブート ディスクの交換 (236 ページ)」の説明に従って、ディスクを交換する必要があります。

ブート デバイスへのアクセス時の障害

ブート プロセスの初期の段階で、システムの初期化直後に、次のようなメッセージが表示される場合があります。

```
SCSI device 0,0 is not responding
```

```
Can't open boot device
```

このメッセージは、システム PROM でブート ドライブからブート プログラムを読み込めないことを示しています。この問題の一般的な原因として、次のことが考えられます。

- ◆ ブート ディスクの電源が入っていない。
- ◆ SCSI バスが終端されていない。
- ◆ コントローラに何らかの障害が発生している。
- ◆ あるディスクに障害が発生してバスがロックされているため、コントローラでディスクを識別することができず、ディスクが接続されていないと認識されている。

この問題の原因を特定するには、まず SCSI バス上のすべてのデバイスが正しく設定されているかどうかを注意深く確認します。ディスクの電源が入っていない場合や、バスが終端されていない場合は、これらの問題を解決してからシステムをリブートします。故障しているディスクがある場合は、バスからそのディスクを取り外して交換します。

ハードウェアに問題が見つからない場合は、エラーの原因はブート ディスク上のデータ障害である可能性があります。この問題を解決するには、代替ブート ディスク（ルートボリュームのミラーを含む）からシステムをブートしてみます。代替ブート ディスクからブートできない場合は、ハードウェア側に何らかの問題があります。障害が発生したブート ディスクを代替ブート ディスクに切り替えてもシステムがブートできない場合も、ハードウェアに問題があります。

/etc/vfstab 中の不正なエントリによる障害

ルート ディスクがカプセル化されて Volume Manager の管理下に置かれている場合、通常のカプセル化処理の一部として、ディスク上のすべてのパーティションに対してボリュームが作成されます。Volume Manager は /etc/vfstab を変更して、ディスクパーティションの代わりに対応するボリュームを使用できるようにします。/etc/vfstab ファイルを手動で編集する場合は、慎重に行う必要があります。もっとも重要なエントリは、/ および /usr に関するエントリです。Volume Manager のインストール前の vfstab は、/etc/vfstab.prevm という名前で保存されます。

/etc/vfstab 中の / エントリの障害

/etc/vfstab 中の / に関するエントリが失われたか不正な場合、システムはシングルユーザ モードでブートします。次のようなメッセージが表示されます。

```
File just loaded does not appear to be executable
```

この時点で、fsck を実行することをお勧めします。次のように入力します。

```
# fsck /dev/vx/rdisk/rootvol
```



ブートプロセスのこの段階では、/ は read/write としてマウントされていません。/etc/vfstabのエントリが不正であるか削除されているので、次のコマンドを使用して、/ を read/write として手動でマウントする必要があります。

```
# mount -o remount /dev/vx/dsk/rootvol
```

/ を read/write でマウントした後に、シェルを終了します。実行レベルの指定を求めるプロンプトが表示されます。マルチユーザモードの場合は、実行レベル 3 を入力します。

```
ENTER RUN LEVEL (0-6,s or S): 3
```

システムのブート後に、/etc/vfstab の / エントリをリストアします。

/etc/vfstab 中の /usr エントリの障害

/etc/vfstab には、/usr が別のディスクパーティションにある場合にだけ、/usr に関するエントリが記述されます。/usr パーティションを含むディスクをカプセル化した後で、Volume Manager は /etc/vfstab のエントリを変更し、対応するボリュームを使用できるようにします。

/etc/vfstab の /usr エントリが失われると、システムをブートすることはできません (/usr ボリュームのミラーがある場合でも同様)。この場合、システムを CD-ROM からブートして、/etc/vfstab をリストアします (「ボリューム上のルート (/) ファイルシステムまたは /usr ファイルシステムの修復 (222 ページ)」を参照)。

/etc/system の欠損または破損による障害

注 Volume Manager によって /etc/system に追加されたエントリは編集しないでください。それらのエントリはすべて、*vxvm_START と *vxvm_END に囲まれています。

ルートファイルシステムの /etc/system は、変更する前にコピーを作成しておくことをお勧めします。これにより、新しい /etc/system ファイルに加えた変更が不正な場合、保存しておいたシステムファイルをブートプログラムに指定できます。ブートプログラムに対して保存しておいたシステムファイルを指定するには、boot -a コマンドを使用してシステムをブートします。システムファイル名の指定を求めるプロンプトが表示されるので、保存してあるシステムファイルのパスを入力します。

/etc/system のコピーが使用できない場合

/etc/system ファイルが破損し、保存しておいたシステムファイルのコピーが使用できない場合は、Volume Manager のルート機能を有効にした状態でシステムをブートすることはできません。/usr がボリュームでない場合は、Volume Manager のルート機能を使用せずに (rootvol が / になっていない状態)、システムをブートできます。

Volume Manager のルート機能を使用しないでシステムをブートするには、「ボリューム上のルート (/) ファイルシステムまたは /usr ファイルシステムの修復 (222 ページ)」の説明を参照して次の手順を実行します。

1. Volume Manager を CD-ROM から起動します。詳細については、『VERITAS Volume Manager Installation Guide』を参照してください。
2. `fixmountroot` コマンドを実行します。
3. `/tmp/rootvol` を作成してマウントします。

これで、`/tmp/rootvol/etc/system` ファイルを編集し、その他の必要な修復を行えるようになります。

システムのブート後、`/etc/system` に次のエントリを入力します。

```
* vxvm_START
rootdev:/pseudo/vxio@0:0
set vxio:vol_rootdev_is_volume=1
* vxvm_END
```

また、`forceload` を使用してルート ミラー ディスクに必要なすべてのドライバを強制的に読み込む必要があります。この処理を行うには、`/etc/system` ファイルを編集し、ドライバごとに次の形式の行を追加します。

```
# forceload:drv/ドライバ名
```

これらのディスクのドライバ名は、`/dev/dsk/` ルートデバイスでロング形式を指定して `list` を実行することにより取得できます。 `io-unit` などのドライバ名があります。

`/etc/system` のコピーが使用できる場合

`/etc/system` ファイルが破損したが、保存しておいた `/etc/system` ファイルのコピーが使用できる場合は、Volume Manager のルート機能を使用してシステムをブートすることができます。

Volume Manager のルート機能を使用してシステムをブートするには、次のコマンドを実行してシステムをブートします。すべてのプロンプトで、Return キーを押してデフォルトの値を使用します (ルート デバイス名を除く)。

```
ok boot -a
.
.
Rebooting with command:-a
Boot device:/iommu/sbus/espdma/esp/sd@5,0   File and args:-a
Enter filename [/kernel/unix]:
```



```

Name of system file [/etc/system.sav]:
Name of default directory for modules [/kernel /usr/kernel]:
Enter name of device instance number file [/etc/path_to_inst]:
root file system type [ufs]:
Enter physical name of root device
[/iommu.....]:/pseudo/vxio@0:0

```

/etc/system が使用できず /usr がボリュームの場合

etc/system ファイルが破損または欠損し、バックアップ コピーも使用できず、/usr がボリュームである場合は、システムをCD-ROMからブートする必要があります(「ボリューム上のルート (/) ファイル システムまたは /usr ファイル システムの修復 (222 ページ)」で説明している手順を参照)。この処理を実行後、ルート ボリュームをマウントし、そのボリューム上の etc/system ファイルを編集します。etc/system に次のエントリを入力します。

```

* vxvm_START
rootdev:/pseudo/vxio@0:0

set vxio:vol_rootdev_is_volume=1
set vxio:vol_swapdev_is_volume=1

* vxvm_END

```

また、前述のように、forceload を使用してルート ミラー ディスクに必要なすべてのドライバを強制的に読み込む必要があります。ファイルの変更後、システム ファイルがリストアされたルート パーティションからシステムをリブートします。

使用できない陳腐化したプレックスからブートしたことによる障害

システムの稼働中にディスクが使用できなくなった場合、そのディスク上にあるボリュームのミラーは陳腐化 (正しいデータが格納されていない状態) します。つまり、そのディスク上のデータは、同じボリュームのほかのミラーと一致しくなくなります。ブート プロセス中、システムは、ルート ボリュームの設定を完全に取得するまで、ルート ボリュームの1つのコピー (ブート ディスク上のコピー) にのみアクセスします。

ブートに使用されたこのボリュームのプレックスが陳腐化している場合、最新のプレックスを含んでいる代替ブート ディスクからシステムをリブートする必要があります。たとえば、元のブート ディスクの電源を切った状態で、Volume Manager によって作成されたブート可能ディスクの1つからシステムをブートした場合、この問題が発生する可能性があります。システムは通常通りブートしますが、電源の入っていないディスク上にあるプレックスが陳腐化します。元のブート ディスクの電源を入れ、そのディスクからシステムをリブートすると、システムは陳腐化したプレックスを使用してブートします。

このほかにも、ブート ディスク上の Volume Manager ヘッダーにエラーが発生したことにより、Volume Manager が正しくディスクを識別できなかった場合にも問題が発生します。この場合、Volume Manager はそのディスクの名前を判別できません。これは、プレックスがディスク名と関連付けられており、識別できなかったディスク上のプレックスが使用できなくなるため、問題となります。

また、ルート ボリューム プレックスに影響を及ぼす障害がルート ディスクで発生した場合にも問題が発生します。次にブートしようとするときも、システムは依然として障害が発生したルート プレックスをブート処理に使用するものと想定します。障害が発生した時点のルート ディスクがミラーリングされている場合は、代替ルート ディスク（有効なルート プレックス）をブート処理用に指定できます。

これらの問題が発生した場合、Volume Manager の vxconfigd ユーティリティは、ブート手順の init プロセスの一部として、いつシステムの構成を設定するときに、そのことを記録します。vxconfigd は、エラーおよび対処法を説明するメッセージを表示し、システムを停止します。たとえば、ディスク rootdisk 上のルート ボリューム rootvol にあるプレックス rootvol-01 が陳腐化している場合、vxconfigd は次のメッセージを表示します。

```
vxvm:vxconfigd:Warning Plex rootvol-01 for root volume is stale or
unusable.

vxvm:vxconfigd:Error: システム ブート ディスクに有効なプレックスがありません。
次のディスクからブートしてください：

Disk:disk01                Device:c0t1d0s2

vxvm:vxconfigd:Error: システムを起動できません。

The system is down.
```

このメッセージは、代替ブート ディスク disk01 にルート プレックスの有効なコピーが含まれており、そのディスクをブート処理に使用すべきであることを管理者に通知します。このメッセージが表示された場合は、システムを代替ブート ディスクからリブートする必要があります。

システムがブートしたら、問題そのものを解決する必要があります。ブート ディスク上のプレックスが単に陳腐化している場合は、システムのブート時に自動的に修正されます。一方、ディスク上の専有領域に問題がある場合や、ディスクに障害が発生している場合は、ディスクを再追加または交換する必要があります。



ブート ディスク上のプレックスを使用できない場合、問題を通知するメールが **Volume Manager** ユーティリティから送信されます。 **vxdisk** ユーティリティを使用してディスクを一覧表示して問題を特定することもできます。上の例で、問題が **rootdisk** の専有領域で発生している障害（メディア障害やディスク上の **Volume Manager** の専有領域が偶発的に上書きされたなど）である場合は、次のような一覧が表示されます。

DEVICE	TYPE	DISK	GROUP	STATUS
-	-	rootdisk	rootdg	failed was:c0t3d0s2
c0t1d0s2	sliced	disk01	rootdg	ONLINE

ホットリロケーションおよびブート ディスク障害

ブート（ルート）ディスクに障害が発生した場合、そのディスクがミラーリングされていれば、障害が発生しているルートディスクのミラーがホットリロケーションによって自動的に新しいミラーに交換されます。この処理を行うために、ホットリロケーションは障害の影響を受けていないルートディスクのミラーを使用して、スペアディスクまたは十分な空き領域があるディスクのいずれかに新しいミラーを作成します。これにより、ブート時に使用できるルートディスクのミラーが、少なくとも2つは常に存在することになります。また、ホットリロケーションデーモンは **vxbootsetup** ユーティリティも呼び出します。このユーティリティは、新しいミラーを含むディスクをブート可能ディスクとして設定します。

障害が発生したルートディスクのボリュームを保存するスペアディスクまたは十分な空き領域が **rootdg** ディスクグループにない場合、ルートディスクのホットリロケーションは異常終了します。**rootvol** および **swapvol** ボリュームには、連続したディスク領域が必要です。障害が発生したルートディスク上のルートボリュームやその他のボリュームを新しい同一のディスクにリロケートできない場合は、各ボリュームを別のディスクにリロケートできます。

rootvol および **swapvol** ボリュームのミラーはシリンダ単位で割り当てる必要があります。したがって、それらのサブディスクがシリンダ境界で開始または終了できる十分な空き領域を備えたディスク上にのみ、ミラーを作成することができます。これらのディスクが使用できない場合、ホットリロケーションは異常終了します。

ブート ディスクの再追加と交換

システムのブートに必要なではないデータは、システムが完全に動作するようになってからでないと Volume Manager によりアクセスされないため、特別な領域に配置しておく必要はありません。Volume Manager はこれらのデータを探し出すことができます。しかし、ブートに不可欠なデータは、ブート プロセスにより検出できるように、ブート可能ディスク上の特定の領域に配置しておく必要があります。

いくつかのシステムでは、ディスク コントローラにより実行されるコントローラ固有の処理やシステム BIOS により、この重要データの位置が制限されます。

ディスク障害が発生した場合、問題を解決するには次の 2 つの方法があります。

- ◆ エラーが一過性または修正可能なものである場合は、同じディスクを再利用できます。これをディスクの再追加と呼びます。場合によっては、障害を起こしたディスクを再フォーマットするか、表面分析を行って代替セクタ マッピングを行うだけで、ディスクを再利用または再追加できるようになります。
- ◆ ディスクが完全に破損している場合は、交換する必要があります。

次の節では、障害が発生したブート ディスクの再追加と交換方法について説明します。

障害を起こしたブート ディスクの再追加

ディスクの再追加は、同じ物理ディスクが使用されるという点を除けば、ディスクの交換と手順は同じです。通常、再追加が必要なディスクは切り離されています。つまり、Volume Manager でディスク障害が検出され、そのディスクへのアクセスが停止されています。

注 システムによっては、例とは異なるデバイス名またはパスを使用している場合があります。デバイス名に関する詳細は、「Volume Manager とは (1 ページ)」を参照してください。

たとえば、システムに disk01 および disk02 の 2 つのディスクがあり、ブート時に通常は、それぞれディスク c0t0d0s2 および c0t1d0s2 に対応づけられていると想定します。障害が発生すると disk01 が切り離されます。これは、次のコマンドを使用して、vxdisk ユーティリティでディスクの一覧を表示すると確認できます。

```
# vxdisk list
```

次のような結果が表示されます。

DEVICE	TYPE	DISK	GROUP	STATUS
c0t0d0s2	sliced	-	-	error
c0t1d0s2	sliced	disk02	rootdg	online
-	-	disk01	rootdg	failed was:c0b0t0d0s0



ディスク `disk01` には関連付けられているデバイスがなく、状態が `failed` と表示されていることに注意してください。このことは、このデバイスが切り離されていることを意味します。デバイス `c0b0t0d0s0` は表示されない場合もあります。ディスクが完全に破損していて、ディスク コントローラがバス上でそのディスクを検出しない場合は、表示されません（バスを使用するシステムの場合）。

`vxdisk list` の出力が異なる場合もあります。たとえば、UNIX パーティション テーブルに関連する修復できない障害がブート ディスクにある場合です。この場合、ルート パーティションが欠落していてそれを修正できない可能性があります。Volume Manager 専有領域ではエラーは発生していません。`vxdisk list` コマンドにより、次のような結果が表示されます。

DEVICE	TYPE	DISK	GROUP	STATUS
<code>c0t0d0s2</code>	<code>sliced</code>	<code>disk01</code>	<code>rootdg</code>	<code>online</code>
<code>c0t1d0s2</code>	<code>sliced</code>	<code>disk02</code>	<code>rootdg</code>	<code>online</code>

ただし、説明されている手順ではエラーを修正できないため、ディスクは障害が発生したものと扱われます。この場合は、障害が発生しているディスクをデバイスから手動で切り離す必要があります。デバイスを切り離すには、`vxdiskadm` ユーティリティの「交換用ディスクの削除」機能を使用します（`vxdiskadm` の詳細については、`vxdiskadm (1M)` マニュアル ページを参照）。ディスクがデバイスから切り離されると、問題を解決するための特別な処理（デバイスの再フォーマットなど）を実行できます。

ディスクを再追加するには、`vxdiskadm` ユーティリティの「障害が発生したディスクまたは削除したディスクの交換」機能を使用してディスクを交換しますが、交換用に同じデバイスを選択します。前述の例では、`disk01` をデバイス `c0t0d0s2` または `c0b0t0d0s2` として（バスを使用するシステムの場合）交換します。

ミラーリングされているブート ディスクに障害が発生した場合は、ホットリロケーションが有効になっていると、新しいミラーが作成され、障害が発生しているブート ディスクから問題のサブディスクが削除されます。ホットリロケーションの完了後に再追加も完了した場合、再追加されたディスクにはディスク障害による影響を受けたルート ボリュームやほかのボリュームは存在しません。この場合、`vxunreloc` を呼び出して、ホットリロケート済みサブディスクを、新しく置き換えられたディスクに戻すことができます。

カプセル化されたサブディスクの新しいディスクへのリロケート処理の解除

ブート ディスクがカプセル化されると、ルート ファイル システムやブート ディスク上のスワップ デバイスなどのシステム領域がボリュームに作成されます。Volume Manager はスワップ領域の一部を使用して専有領域を作成するため、専有領域はディスクの中間に配置されます。Volume Manager ディスクに設定するために、ディスクがカプセル化されずに初期化された場合、専有領域はディスクの先頭から作成されます。

カプセル化されたブート ディスクに障害が発生すると、すべてのサブディスクがホットリロケーションによってほかのディスクにコピーされます。リロケーション処理の一部として、元のディスク名およびオフセットは、サブディスクレコード内に保存されます。障害の発生したブート ディスクが同じ記憶容量を持つディスクと置換されると、新しいディスクは「初期化」されてディスクグループに戻されます。vxunreloc が自動的に実行され、すべてのサブディスクが新しいディスクに戻されます。ただし、初期化されたディスクとカプセル化されたディスクのディスクレイアウトが異なるため、リロケート処理が解除された各サブディスクに対するディスクのオフセットのカプセル化方法が影響を受けます。サブディスクをディスクに移動するには、-f オプションを使用する必要がありますが、正確なオフセットに移動できる保証はありません。交換されたディスクが元のブート ディスクよりも少なくとも 2 メガ以上大きい場合は、すべてのサブディスクをディスク上の正確なオフセットに移動することができます。

vxunreloc によって、すべてのサブディスクをディスクに移動した後は、新しいディスクからのブートが可能になります。

注：通常、ダンプ デバイスがルート ディスクのスワップ パーティションとして設定されます。スワップ サブディスクをディスク間で移動すると（ホットリロケーション、リロケート処理の解除、手動操作など）、ダンプ デバイスを新しいディスクに再設定する必要があります。

Solaris 2.6 では、ダンプ デバイスは dumpfile 構造に保存されます。デバイスの設定を確認するには、次のコマンドを使用します。

```
# echo dumpfile+0x10/s | adb -k /dev/ksyms /dev/mem
```

この設定はリブートによって変更可能で、ダンプ デバイスを 1 番目のスワップパーティションに設定できます。

Solaris 2.7 以降では、

ダンプ デバイスの参照または設定に、dumpadm コマンドを使用します。詳細については、次のように入力して結果を参照してください。

```
# man dumpadm
```



障害を起こしたブート ディスクの交換

ブート ディスクを交換する際は、まずシステムを代替ブート ディスクからブートする必要があります。障害が発生しているディスクがデバイスから切り離されていない場合、`vxdiskadm` の「交換用ディスクの削除」機能を使用して手作業で切り離す必要があります。`vxdiskadm` の詳細については、`vxdiskadm (1M)` マニュアル ページを参照してください。ディスクが切り離されたら、システムをシャットダウンしてハードウェアを交換します。

交換用ディスクには、少なくとも交換前のディスクで使用されていた記憶領域と同じ容量が必要です。交換用ディスクには、サブディスクを格納するためのディスク領域に、元のディスクのすべてのサブディスクを現在の各ディスク オフセットの位置で格納できる大きさが必要です。交換用ディスクに必要な最小サイズを判断するには、障害が発生したディスク上で使用されていた領域を確認する必要があります。

交換用ディスクに必要なサイズの概数を知るには、次のコマンドを使用します。

```
# vxprint -st -e 'sd_disk="ディスク名"'
```

出力結果から、一覧表示された最後のサブディスクの `DISKOFFS` カラムと `LENGTH` カラムの値を加算します。その値は、512 バイトの単位です。合計を 2 で割ってキロバイト単位の値を得ます。

注 通常、製造元が報告しているディスク サイズは使用可能な容量を示しているわけではありません。また、百万バイト単位で表示している製造元もありますが、それはメガバイト単位と等しくありません。

交換用ディスクが検出されたら、システムを完全にシャットダウンして必要なハードウェアを交換します。ハードウェアの交換後、システムをブートします。`vxdiskadm` ユーティリティの[障害が発生したディスクまたは削除したディスクの交換]機能を使用して、障害のあるディスクを追加した新しいデバイスと交換します。

ディスクの再接続

ディスクが完全に破損し、ホットリロケーションが使用できない場合、または Volume Manager がいくつかのディスクドライバを読み込まないか読み込めない状態で起動した場合（ディスクが **failed** 状態になります）、ディスクの再接続を実行できます。問題の解決後に、`vxreattach` コマンドを実行して、プレックスに陳腐化フラグを設定せずにディスクを再接続できます。ただし、必ずディスク上のボリュームを起動する前に再接続する必要があります。

`vxreattach` コマンドは、`vxdiskadm` メニューから実行するディスクのリカバリの一部として、ブート プロセス中に呼び出されます。可能な場合、`vxreattach` は、障害が発生したディスク メディア レコードを同じデバイス名を持つディスクに再接続します。再接続は、ディスクが以前配置されていたのと同じディスク グループ内で行われ、元のディスク メディア名が維持されます。

再接続を実行後、リカバリ処理が必要ない場合があります。ディスク障害の元の（または別の）原因が残っている場合、再接続は失敗することがあります。

`vxreattach -c` コマンドを実行すると、再接続できるかどうかを確認されるだけで、処理自体は行われません。処理を行う代わりに、ディスクを接続できるディスク グループとディスク メディア名を表示します。

`vxreattach` コマンドの詳細については、`vxreattach (1M)` マニュアル ページを参照してください。



再インストールによるリカバリ

ルート（ブート）ディスクのすべてのコピーが破損した場合、またはファイルシステムの破損により重要なファイルが欠損した場合は、再インストールする必要があります。これらの種類の障害が発生した場合は、現時点でルート ファイルシステムをバックアップからリストアする方法がほかにはないので、システム全体を再インストールする必要があります。

これらの障害が発生した場合は、元の **Volume Manager** 設定を可能な限り維持するようにしてください。障害の影響を直接受けていないボリュームは保存できます。それらのボリュームの構成を設定し直す必要はありません。

一般的な再インストール情報

ここでは、**Volume Manager** の再インストール手順、および障害後に元の構成の設定をできる限り保全する方法を説明します。

システムの再インストールでは、再インストールに使用するすべてのディスクの内容が破棄されます。

すべての **Volume Manager** 関連情報が、再インストール中に削除されます。削除されるデータには、ディスク識別子や **Volume Manager** の設定コピーなどの、削除されるディスクの専有領域内のデータが含まれます。この情報が削除されると、ディスクを **Volume Manager** ディスクとして使用できなくなります。

システム ルート ディスクは、常に再インストールの対象となります。ほかのディスクも対象になることがあります。**Volume Manager** のインストール中、または後から実行するカプセル化によってルート ディスクが**Volume Manager**の管理下に置かれている場合、ルート ディスクとそのディスク上のボリュームまたはミラーが再インストール中に失われます。ほかの再インストール対象のディスク、削除および交換されたディスクでも、**Volume Manager** 設定データ（ボリュームおよびミラーを含む）が失われます。

ルート ディスクを含むディスクが障害の発生前に **Volume Manager** の管理下に置かれていない場合は、再インストールによって **Volume Manager** 設定データが失われることはありません。その他のディスクは、第4章の「ディスク コマンドおよびディスク グループ コマンド」で説明する処理で置き換えることができます。

ルート ディスクを **Volume Manager** の管理下に置いていない場合は、再インストール後のリカバリ処理を簡略化できますが、再インストールの必要が生じる可能性は高くなります。ルート ディスクを **Volume Manager** の管理下に置き、そのミラーを作成しておく、システムの再インストールが必要となる問題の多くを回避できます。

再インストールが必要になった場合、障害や再インストールに直接関係のないディスクにあるボリューム、またはそのディスク上にコピーを持つボリュームだけが保存されます。ルート ディスクおよび障害や再インストールに関係するその他のディスク上のボリュームはいずれも、再インストール中に失われます。これらのボリュームのバックアップコピーが使用できる場合、再インストール後にボリュームをリストアできます。システムによっては、**root**、**stand**、および **usr** ファイルシステムはバックアップからリストアすることができないものがあります。

再インストールおよび構成の再設定の手順

システムを再インストールし、Volume Manager 設定をリカバリするには、次の手順に従います。これらの手順の詳細は、以降の節で説明します。

1. インストールするためにシステムを準備します。

障害が発生したディスクまたはその他のハードウェアの交換、再インストール対象でないディスクの切り離しなどを行います。

2. オペレーティング システムをインストールします。

基本システムや Volume Manager と関係のないパッケージを再インストールします。

3. Volume Manager をインストールします。

Volume Manager パッケージを追加しますが、`vxinstall` コマンドは実行しないでください。

4. Volume Manager 設定をリカバリします。

5. Volume Manager 設定を補整します。

障害または再インストールによって影響を受けたボリューム中の情報のリストア、およびシステム ボリューム (`rootvol`、`swapvol`、`usr` など) の再作成を行います。

再インストールのためのシステムの準備

再インストールの対象でないディスク上のデータを損なわないようにするため、ルートディスクのみを再インストールの対象とする必要があります。

注 インストールに関する自動オプションの中には、システム管理者による確認を必要とすることなく、ルート ディスク以外のディスクにアクセスするものがあります。このため、オペレーティング システムを再インストールする前に、ボリュームを含んでいるほかのすべてのディスクをシステムから切り離しておくことをお勧めします。

切り離しておけば、ほかのディスクは再インストールの影響を受けずに済みます。たとえば、元のオペレーティング システムをインストールしたときに、ディスク上に `home` ファイルシステムを配置したとします。2 番目のディスクを切り離しておけば、再インストールしても `home` ファイルシステムはそのまま残ります。

オペレーティング システムの再インストール

障害を起したか起しているディスクを交換し、再インストール対象でないディスクを切り離したら、オペレーティング システムのマニュアルに従ってオペレーティング システムを再インストールします。Volume Manager のインストール前に、オペレーティング システムをインストールしてください。



オペレーティング システムのインストール中は、いかなる方法でもルート ディスク以外のディスクにアクセスできないようにする必要があります。ルート ディスク以外のディスクに何らかの書き込みが行われると、そのディスク上にある Volume Manager 設定が壊される可能性があります。

注 再インストール中に、ホスト ID (またはホスト名) を変更できます。以降の記述では、ホスト ID (ホスト名) を変更していないと想定しているため、既存のホスト ID (ホスト名) をそのまま維持することをお勧めします。

Volume Manager の再インストール

Volume Manager のインストールは、次の 2 段階で構成されます。

- ◆ CD-ROM からの Volume Manager の読み込み
- ◆ Volume Manager の初期化

Volume Manager を再インストールするには、『VERITAS Volume Manager Installation Guide』で CD-ROM から Volume Manager を読み込むための手順を参照してください。

注 ルート以外のディスクに保持されている Volume Manager 設定をリカバリする場合は、再インストール後に (vxinstall を使用して) Volume Manager を初期化しないでください。

システムによっては、vxserial を使用して、Volume Manager のライセンス キーをインストールできます (vxserial (1M) マニュアル ページを参照)。

Volume Manager 設定のリカバリ

Volume Manager パッケージの読み込みが完了したら、次の手順に従って Volume Manager 設定をリカバリします。

1. システムをシャットダウンします。
2. システムから切り離されたディスクを再接続します。
3. システムをリブートします。
4. システムのリブート後、次のコマンドを使用して、システムをシングル ユーザ モードにします。

```
# shutdown -g0 -is -y
```
5. プロンプトが表示されたら、パスワードを入力し、Return キーを押して処理を続行します。
6. Volume Manager の読み込み時に作成されたインストール処理に関連したファイルで、以降は不要であるファイルを削除します。次のコマンドを使用します。

```
# rm -rf /etc/vx/reconfig.d/state.d/install-db
```

7. ファイルの削除後、いくつかの Volume Manager 入出力デーモンを起動します。次のコマンドを使用してデーモンを起動します。

```
# vxiod set 10
```
8. Volume Manager 設定デーモン `vxconfigd` を無効(disable)モードで起動します。次のコマンドを使用します。

```
# vxconfigd -m disable
```
9. `vxconfigd` デーモンを初期化します。次のコマンドを使用します。

```
# vxdctl init
```
10. DMP サブシステムを初期化します。次のコマンドを使用します。

```
# vxdctl initdmp
```
11. `vxconfigd` を有効にします。次のコマンドを使用します。

```
# vxdctl enable
```

これで、再インストールの対象でないディスク上に保持されていた設定がリカバリされました。しかし、ルート ディスクは再インストールされたので、Volume Manager で VM ディスクとしては認識されません。退避されていたディスクの設定には、Volume Manager 設定の一部としてのルート ディスクが含まれていません。

障害発生時または再インストール時に、システムのルート ディスクと再インストール対象のほかのディスクが Volume Manager の管理下に置かれていない場合、これで再設定は完了です。ディスクの交換方法は複数あるので、適した方法を選択してください。

ルート ディスク（または別のディスク）が再インストールの対象となる場合、そのディスク（またはシステムに接続されていない別のディスク）のボリュームやミラーはこの時点でアクセスできなくなります。ボリュームに、再インストール、削除、または交換されたディスク上のプレックスが 1 つのみ含まれる場合、ボリュームのデータは失われるので、バックアップからリストアする必要があります。

また、システムのルート ファイルシステム、スワップ領域（一部のシステムでは、stand 領域）、および /usr ファイルシステムは、ボリューム上から削除されています。これらの問題を解決するには、「設定の補整 (242 ページ)」の手順に従ってください。

`vxdctl enable` が正常に終了すると、ホットリロケーション機能を起動することができますが、実際には、管理者がサービスを起動および作動しても、ほかの再設定手順に影響がないことを確認した後で、ホットリロケーション機能を起動する必要があります。「最終的な再設定」の手順を完了した後でホットリロケーション機能を起動することをお勧めします。ホットリロケーションの起動方法については、「ホットリロケーションの起動 (248 ページ)」を参照してください。



設定の補整

ここでは、Volume Manager の再インストール後にシステム設定を補整する手順について以降に説明します。

次の種類の補整処理を説明します。

- ◆ ルート機能の補整
- ◆ ボリュームの補整
- ◆ ディスクの補整

その後、下記の再設定情報を示します。

- ◆ ルート機能の再設定
- ◆ 最終的な再設定

ルート機能の補整

Volume Manage 設定の補整を開始するには、ルート機能に関連するボリュームをすべて削除します。ルート ディスク（およびシステム ブート プロセスに関連するほかのディスク）が Volume Manage の管理下にある場合は、この処理を行う必要があります。次のボリュームを削除します。

- ◆ `rootvol`。ルート ファイルシステムが含まれています。
- ◆ `swapvol`。スワップ領域が含まれています。
- ◆ （一部のシステム）`standvol`。`stand` ファイルシステムが含まれています。
- ◆ `usr`。/`usr` ファイルシステムが含まれています。

ルート ボリュームを削除するには、`vxedit` コマンドを使用します。

```
# vxedit -fr rm rootvol
```

`swap`、`stand`、および `usr` ボリュームを削除するには、上記の `rootvol` の代わりそれぞれ `swapvol`、`standvol`、および `usr` を指定します。

ボリュームの補整

ルート機能の補整が完了したら、バックアップからリストアするボリュームを指定する必要があります。リストアの対象となるボリュームには、再インストールまたは削除されたディスク上にすべてのミラー（ボリュームのすべてのコピー）があるものが含まれます。これらのボリュームは無効なため、削除、再作成、およびバックアップからのリストアを行う必要があります。ボリュームのミラーの一部のみが再初期化または削除されたディスク上にある場合は、それら一部のミラーを削除する必要があります。ミラーは後で再追加できます。

ボリュームをリストアするには、次の手順に従います。

1. 削除または再インストールされた VM ディスクを確認します。次のコマンドを使用します。

```
# vxdisk list
```

システム ディスク デバイスと、これらのデバイスの状態が一覧表示されます。たとえば、再インストールしたシステムに 3 つのディスクと再インストール済みルートディスクがある場合、`vxdisk list` コマンドを実行すると、次のような結果が表示されます。

DEVICE	TYPE	DISK	GROUP	STATUS
c0t0d0s2	sliced	-	-	error
c0t1d0s2	sliced	disk02	rootdg	online
c0t2d0s2	sliced	disk03	rootdg	online
-	-	disk01	rootdg	failed was:
c0t0d0s2				

注 システムによっては、例とは異なるデバイス名を使用している場合があります。デバイス名の詳細については、「Volume Manager とは(1 ページ)」を参照してください。

この出力は、再インストールされたルート デバイス `c0t0d0s2` が VM ディスクに関連付けられておらず、`error` 状態であることを示しています。`disk02` および `disk03` は再インストールの対象ではないので、Volume Manager によって認識され、それぞれデバイス `c0t1d0s2` および `c0t2d0s2` に関連付けられています。`disk01` は、交換されたディスク デバイスに関連付けられた VM ディスクでしたが、この時点ではそのデバイス (`c0t0d0s2`) に関連付けられていません。

ほかのディスク (ディスク上のボリュームやミラーを含む) が再インストール中に削除または交換された場合は、これらのディスクについても、ディスク デバイスが `error` 状態で、VM ディスクがデバイスに関連付けられていないと表示されます。

2. 削除または交換されたディスクを確認したら、障害が発生したディスク上のミラーをすべて特定します。次のコマンドを使用します。

```
# vxprint -sF "%vname" -e'sd_disk = "ディスク名"
```

ディスク名には、`failed` 状態のディスクの名前を指定します。コマンド内ではディスク名を必ず引用符で囲ってください。囲まないと、エラーメッセージが返されません。`vxprint` コマンドによって、障害を起したディスク上にミラーがあるボリュームの一覧が表示されます。`failed` 状態のすべてのディスクに対して、このコマンドを繰り返し実行します。



3. 各ボリュームの状態を確認します。次のコマンドを使用してボリューム情報を印刷します。

```
# vxprint -th ボリューム名
```

ボリューム名には、確認するボリュームの名前を指定します。vxprint コマンドは、ボリュームの状態、そのプレックス、およびこれらのプレックスを構成するディスクの部分を表示します。たとえば、プレックスを1つのみ含むボリューム v01 が、再インストール済みディスク disk01 上にあるとします。vxprint -th v01 コマンドを実行すると、次のような結果が表示されます。

```
V  NAME          USETYPE  KSTATE   STATE    LENGTH  READPOL
   PREFPLEX

PL NAME          VOLUME   KSTATE   STATE    LENGTH  LAYOUT
   NCOL/WID MODE

SD NAME          PLEX     DISK     DISKOFFS LENGTH  [COL/]OFF
   DEVICE MODE

v  v01           fsgen    DISABLED ACTIVE    24000   SELECT   -
pl v01-01       v01      DISABLED NODEVICE  24000
   CONCAT      -   RW
sd disk01-06   v0101    disk01   245759   24000   0
   c1t5d1 ENA
```

p1 で始まる行には、ボリュームの唯一のプレックスが表示されています。プレックス v01-01 の STATE フィールドは NODEVICE です。このプレックスは、交換、削除、または再インストールされたディスク上の領域を占めています。このプレックスは有効ではないので、削除する必要があります。

v01-01 はボリュームの唯一のプレックスなので、ボリュームをバックアップからリストアする方法以外でボリュームの内容をリカバリすることはできません。ボリュームも削除する必要があります。ボリュームのバックアップ コピーがある場合は、後でボリュームをリストアできます。バックアップ手順で必要となるため、ボリューム名とサイズを記録しておいてください。

4. ボリューム v01 を削除するには、vxedit コマンドを使用します。

```
# vxedit -r rm v01
```



障害が発生したディスク上にプレックスの一部だけが存在する場合があります。ボリュームにストライププレックスが関連付けられている場合、そのボリュームは複数のディスクにまたがっています。たとえば、ボリューム `v02` に、3つのディスクにわたってストライピングされている1つのストライププレックスが含まれ、3つのうち1つのディスクが再インストールされたディスク `disk01` であるとして、`vxprint -th v02` コマンドを実行すると、次のような結果が表示されます。

```

V   NAME           USETYPE   KSTATE    STATE     LENGTH
  READPOL  PREFPLEX
PL  NAME           VOLUME    KSTATE    STATE     LENGTH
  LAYOUT    NCOL/WID  MODE
SD  NAME           PLEX      DISK      DISKOFFS  LENGTH
  [COL/]OFF DEVICE    MODE
v   v02            fsgen     DISABLED  ACTIVE    30720
  SELECT    v02-01
pl  v02-01         v02       DISABLED  NODEVICE  30720
  STRIPE    3/128     RW
sd  disk02-02     v02-01    disk02    424144    10240
  0/0      c1t2d0    ENA
sd  disk01-05     v02-01    disk01    620544    10240
  1/0      c1t2d1    DIS
sd  disk03-01     v02-01    disk03    620544    10240
  2/0      c1t2d2    ENA

```

この出力は、プレックス `v02-01` がストライピングされている3つのディスクを示しています (`sd` で始まる行はストライプを表します)。ストライプ領域の1つが、障害が発生したディスク上にあります。このディスクは有効ではないため、プレックス `v02-01` の状態は `NODEVICE` になっています。このプレックスはボリュームの唯一のプレックスなので、ボリュームは無効となり、削除する必要があります。 `v02` のコピーがバックアップメディア上にある場合は、後からをリストアできます。後でバックアップからリストアする場合は、ボリュームの名前とサイズを記録しておいてください。

5. `vxedit` コマンドを使用して、前述の通りボリュームを削除します。

障害が発生したディスク上にミラーが1つ含まれてるボリュームが、有効なディスク上にほかのミラーを保持している場合があります。このような場合、有効なディスク上のデータは依然として有効なので、ボリュームをバックアップからリストアする必要はありません。



`vxprint -th` コマンドを実行すると、障害が発生したディスク (`disk01`) 上にプレックスが 1 つ含まれ、有効なディスク (`disk02`) 上に別のプレックスが含まれているボリュームについて、次のような結果が表示されます。

V	NAME	USETYPE	KSTATE	STATE	LENGTH
	READPOL	PREFPLEX			
PL	NAME	VOLUME	KSTATE	STATE	LENGTH
	LAYOUT	NCOL/WID	MODE		
SD	NAME	PLEX	DISK	DISKOFFS	LENGTH
	[COL/]OFF	DEVICE	MODE		
v	v03	fsgen	DISABLED	ACTIVE	30720
	SELECT	-			
pl	v03-01	v03	DISABLED	ACTIVE	30720
	CONCAT	-	RW		
sd	disk02-01	v03-01	disk01	620544	30720
	0	clt3d0	ENA		
pl	v03-02	v03	DISABLED	NODEVICE	30720
	CONCAT	-	RW		
sd	disk01-04	v0302	disk03	262144	30720
	0	clt2d2	DIS		

このボリュームには、`v03-01` と `v03-02` の 2 つのプレックスが含まれています。最初のプレックス (`v03-01`) は、無効なディスク上の領域を使用していないので、この時点でも使用できます。2 番目のプレックス (`v03-02`) は、無効なディスク `disk01` 上の領域を使用しているため、`NODEVICE` 状態となっています。プレックス `v03-02` は削除する必要があります。ただし、このボリュームには有効なデータを含む有効なプレックスが 1 つあります。このボリュームをミラーリングする必要がある場合、別のプレックスを後から追加できます。後から別のプレックスを作成する場合は、ボリューム名を記録しておいてください。

- 無効なプレックスを削除するには、まずプレックスとボリュームの関連付けを解除してから削除します。この処理には、`vxplex` コマンドを使用します。プレックス `v03-02` を削除するには、次のように入力します。

```
# vxplex -o rm dis v03-02
```

- すべてのボリュームの補正が完了したら、「ディスクの補正 (247 ページ)」の説明に従ってディスク設定を補正します。

ディスクの補整

無効なボリュームとプレックスをすべて削除したら、ディスク設定を補整できます。削除、再インストール、または交換された各ディスク (vxdisk list コマンドの出力で確認) は、設定から削除する必要があります。

ディスクを削除するには、vxdg コマンドを使用します。障害が発生したディスク disk01 を削除するには、次のように入力します。

```
# vxdg rmdisk disk01
```

vxdg コマンドによりエラー メッセージが返された場合は、無効なミラーが存在します。「ボリュームの補整 (242 ページ)」で説明されている手順を繰り返して、無効なボリュームとミラーをすべて削除してください。

ルート機能の再設定

無効なディスクをすべて削除したら、交換または再インストールされたディスクを Volume Manager の管理下に追加することができます。ルート ディスクが本来 Volume Manager の管理下にあった場合や、ルート ディスクを新たに Volume Manager の管理下に置く場合は、まず、そのディスクを追加します。

ルート ディスクを Volume Manager の管理下に追加するには、Volume Manager Support Operations (vxdiskadm) を使用します。次のコマンドを使用します。

```
# vxdiskadm
```

vxdiskadm のメイン メニューから、メニュー項目 2[1 つ以上のディスクのカプセル化] を選択します。指示に従って、システムのルート ディスクをカプセル化します。

カプセル化の完了後、システムをリブートしてマルチユーザ モードにします。

最終的な再設定

ルート ディスクをカプセル化したら、vxdiskadm を使用して、交換されたほかのディスクを追加する必要があります。オペレーティング システムの再インストール中に再インストールしたディスクは、カプセル化または追加する必要があります。

すべてのディスクをシステムに追加したら、設定の補整によって完全に削除されたボリュームを再作成し、バックアップからその内容をリストアできます。ボリュームの再作成は、vxassist またはグラフィカル ユーザ インタフェースを使用して実行できます。

次のコマンドを使用して、ボリューム v01 および v02by を再作成できます。

```
# vxassist make v01 24000
```

```
# vxassist make v02 30720 layout=stripe nstripe=3
```

ボリュームを作成したら、通常のバックアップ / リストア手順で、バックアップからボリュームをリストアできます。



ボリュームの補整の一環としてプレックスが削除されたボリュームのミラーは、本書の `vxassist` を使用したボリュームのミラーリングの手順に従って再作成できます。ボリューム `v03` から削除されたボリュームを交換するには、次のように `vxassist` コマンドを使用します。

```
# vxassist mirror v03
```

再インストール中に失われたボリュームとプレックスをリストアしたらリカバリ処理は完了です。システムは障害が発生する前と同じ設定になりました。

ホットリロケーションの起動

この時点で、管理者はシステムをリブートするか、必要に応じて手作業でホットリロケーション機能を起動する必要があります。どちらを行っても、リロケーションデーモン（およびその `vxnotify` プロセス）が起動します。

ホットリロケーションを起動するには、次のコマンドを使用します。

監視デーモンを起動します。これによって、何らかの問題が検出されたときに、管理者に電子メールで通知されるようになります。問題の通知の宛先のアドレスを変更するには、`vxrelocd` の引数を変更します。

```
# nohup /usr/lib/vxvm/bin/vxrelocd root &
```

次のコマンドも有効です。

```
# nohup /usr/lib/vxvm/bin/vxrelocd root > /dev/null 2>&1 &
```

次のコマンドを使用すると、ホットリロケーションが起動しているかどうかを確認することができます。

```
# ps -ef | grep vxrelocd | grep -v grep
```

プレックスとボリュームの状態

以降の節では、プレックスとボリュームの状態について説明します。

プレックスの状態

プレックスの状態は、プレックスがボリュームの内容と同一で完全なコピー（ミラー）であるかどうかを反映しています。Volume Manager のユーティリティは自動的にプレックスの状態を維持します。ただし、プレックスと関連付けられているボリュームに対する変更をプレックスに書き込みたくない場合は、プレックスの状態を変更することができます。たとえば、あるプレックスを含むディスクに障害の兆候が現れた場合、そのプレックスを一時的に無効状態にすることができます。

注 プレックスは、必ずしもボリュームに関連付ける必要はありません。プレックスは、`vxmake plex` コマンドで作成できます。このコマンドを使用して作成されたプレックスは、後でボリュームに接続できます。

Volume Manager のユーティリティは、プレックス状態を使用して次の処理を行います。

- ◆ ボリューム内容が既知の状態に初期化されているかどうかを示す。
- ◆ プレックスにボリュームの内容の有効なコピー（ミラー）が含まれているかどうかを判別する。
- ◆ システムが障害を起した時に、プレックスが使用されていたかどうかを確認する。
- ◆ プレックス上の処理を監視する。

この節では、プレックスの状態の詳細な知識が必要とされる管理者を対象として、プレックスの状態について詳しく説明します。

ボリュームに関連付けられているプレックスは、次のいずれかの状態になります。

- ◆ EMPTY
- ◆ CLEAN
- ◆ ACTIVE
- ◆ STALE
- ◆ OFFLINE
- ◆ TEMP
- ◆ TEMPRM
- ◆ TEMPRMSD
- ◆ IOFAIL

ダーティ リージョン ロギングまたは RAID-5 ログのプレックスは特殊なケースであり、状態は常に LOG に設定されます。



EMPTY プレックス状態

ボリュームの作成時に、ボリュームに関連付けられているすべてのプレックスは EMPTY 状態に設定され、プレックスが初期化されていないことを示します。

CLEAN プレックス状態

プレックスにボリュームの内容の一貫性のあるコピー（ミラー）が含まれており、ある操作によってそのボリュームを無効にする操作が行われた場合、そのプレックスは CLEAN 状態になります。結果的に、1つのボリュームのすべてのプレックスが CLEAN 状態であれば、ボリュームの起動時に、すべてのプレックスが同一であることを確認するために操作を行う必要はありません。

ACTIVE プレックス状態

プレックスは、次の2つの場合に ACTIVE 状態になります。

- ◆ ボリュームが起動し、プレックスが通常のボリューム入出力に全面的に使用される場合（ボリュームの内容が変更されるとプレックスの内容も変更される場合）。
- ◆ システムクラッシュによりボリュームが停止されたときにプレックス状態が ACTIVE だった場合。

後者の場合、システム障害により、プレックスの内容の整合性が失われる可能性があります。ボリュームの起動時に、ACTIVE 状態のプレックスの内容が同一となるように Volume Manager はリカバリ処理を行います。

注 正常に稼働しているシステムでは、通常、すべてのボリュームのプレックスが ACTIVE 状態になっています。

STALE プレックス状態

プレックスに含まれるボリュームの内容が完全かつ最新でない可能性がある場合、そのプレックスは STALE 状態になります。また、プレックスで入出力エラーが発生すると、カーネルはそのプレックスの内容の使用と更新を停止し、何らかの操作によってプレックスの状態は STALE に設定されます。

`vxplex att` を実行すると、ACTIVE 状態のプレックスから STALE 状態のプレックスの内容をリカバリできます。アトミックコピー操作によって、ボリュームの内容を STALE 状態のプレックスにコピーします。システム管理者は、`vxplex det` を使用して、プレックスを強制的に STALE 状態にすることができます。

OFFLINE プレックスの状態

`vxmend off` を実行すると、プレックスの状態が OFFLINE になり、プレックスがボリュームから切り離されます。切り離されたプレックスとボリュームとの関連付けは維持されますが、ボリュームが変更されても OFFLINE 状態のプレックスは更新されません。この状

態は、プレックスをオンライン化し、`vxplex att` を実行してボリュームに再接続するまで続きます。再接続されると、プレックスは `STALE` 状態になり、プレックスの内容は `vxvol start` を次に実行したときにリカバリされます。

TEMP プレックスの状態

プレックスを `TEMP` 状態にすると、アトミックには実行できない一部のプレックス処理を簡単に実行できます。たとえば、有効なボリュームにプレックスを接続する場合、ボリュームの内容をプレックスにコピーしてからでないと、完全に接続することはできません。

ユーティリティは、このような処理の開始時にプレックスの状態を `TEMP` に設定し、処理が終了すると適切な状態に設定します。何らかの理由でシステムに障害が発生し、プレックスの状態が `TEMP` になっている場合は、処理が完了していません。その後 `vxvol start` を実行すると、`TEMP` 状態にあるプレックスの関連付けが解除されます。

TEMPRM プレックスの状態

プレックスの `TEMPRM` 状態は、`TEMP` 状態と似ていますが、`TEMPRM` 状態のプレックスは処理が完了すると削除されます。サブディスクの処理には、一時プレックスを必要とするものがあります。たとえば、サブディスクをプレックスに関連付ける場合は、サブディスクをボリュームの内容で更新してから実際に関連付けを行います。この更新処理では、`TEMPRM` 状態の一時プレックスにサブディスクを関連付ける必要があります。処理が完了すると、`TEMPRM` 状態のプレックスが削除されます。

何らかの理由でシステムに障害が発生し、プレックスの状態が `TEMPRM` になっている場合は、処理が完了していません。後の処理で、`TEMPRM` 状態のプレックスの関連付けが解除され、プレックスは削除されます。

TEMPRMSD プレックス状態

プレックスの `TEMPRMSD` 状態は、新しいプレックスを接続するときに `vxassist` によって使用されます。処理が完了しないと、そのプレックスとサブディスクは削除されます。

IOFAIL プレックスの状態

プレックスの `IOFAIL` 状態は、パーシステント ステート ロギングに関連付けられています。ACTIVE 状態のプレックスに障害が検出されると、そのプレックスは `vxconfigd` によって `IOFAIL` 状態に設定され、ボリューム起動時に行われるリカバリ選択プロセスの対象外になります。



プレックスの状態のサイクル

プレックスの状態の変化は、正常な処理の一環として行われます。プレックスの状態の変化の中には、異常が発生しているため、Volume Manager が正常な状態に戻す必要があることを示しているものもあります。システムの起動時に、ボリュームが自動的に起動し、`vxvol start`によってCLEAN状態のプレックスはすべてACTIVE状態に変わります。システムを終了するまで何も問題が起これなければ、ボリュームの停止時にすべてのACTIVE状態のプレックスがCLEAN状態に設定されます。このサイクルが繰り返されます。起動時にすべてのプレックスがCLEAN状態であれば(`vxvol staçit`によってプレックスがACTIVE状態になる前)、システムが通常通り終了したことを示し、起動プロセスが最適化されます。

プレックスのカーネル状態

プレックスのカーネル状態は、プレックスへアクセスできるかどうかを示します。プレックスのカーネル状態はボリュームドライバで監視されます。この状態によって、プレックスの操作モードがオフライン(DISABLED)、保守(DETACHED)、またはオンライン(ENABLED)に区分されます。

プレックスのカーネル状態には、次の種類があります。

- ◆ DISABLED – プレックスにアクセスすることはできません。
- ◆ DETACHED – ボリュームへの書き込みはプレックスに反映されません。ボリュームからの読み取り要求にプレックスは応じません。プレックスの操作およびioctl関数は処理されます。
- ◆ ENABLED – ボリュームへの書き込み要求はプレックスに反映されます。ボリュームからの読み取り要求にプレックスは応じます。

注 これらのプレックス状態はシステム内部で管理されているため、ユーザの介入を必要とせずに設定されます。正常に稼働しているシステムでは、すべてのプレックスがENABLED状態になっています。

ボリュームの状態

ボリュームの状態には複数の種類があり、プレックスの状態と似ているものもあります。

- ◆ CLEAN – ボリュームは起動しておらず(カーネル状態がDISABLED)、プレックスは同期がとれています。
- ◆ ACTIVE – マシンがリブートされたときに、ボリュームが起動していた(現在のカーネル状態がENABLED)か、または使用されていた(そのときのカーネル状態がENABLED)ことを示します。ボリュームが現在ENABLED状態である場合は、ボリュームが使用中なのでプレックスの状態は固定されません。ボリュームが現在DISABLED状態である場合は、プレックスの内容に整合性がない可能性があります。ボリュームを起動すると整合性が確保されます。

- ◆ **EMPTY** – ボリュームの内容が初期化されていません。ボリュームが **EMPTY** 状態の場合、カーネル状態は常に **DISABLED** です。
- ◆ **SYNC** – ボリュームが読み取り/ライトバックリカバリモードである（現在のカーネル状態が **ENABLED**）か、またはシステムのリブート時に読み取り/ライトバックリカバリモードだった（カーネル状態が **DISABLED**）ことを示します。読み取り/ライトバックリカバリでは、1つのプレックスのブロックからデータを読み取り、ほかのすべての書き込み可能なプレックスへ書き込むことで、プレックス間の整合性がリカバリされます。ボリュームが **ENABLED** 状態の場合、プレックスは読み取り/ライトバックリカバリによっ再同期化されていることを示します。ボリュームが **DISABLED** 状態の場合は、プレックスがマシンのリブート時に読み取り/ライトバックリカバリによって再同期化されることを示します。
- ◆ **NEEDSYNC** – 次のボリューム起動時に、再同期処理を行う必要があります。

ボリューム起動中のこれらのフラグの解釈は、ボリュームのパーシステントステートロギングにより変更されます（**DIRTY/CLEAN** フラグなど）。**CLEAN** フラグが設定されている場合、**ACTIVE** ボリュームにはどのプロセスからも書き込みされていないか、またはリブート時に開かれていません。このため、そのボリュームは **CLEAN** 状態であると判断されます。ボリュームが **CLEAN** 状態になっている場合は、常に **CLEAN** フラグが設定されます。

RAID-5 ボリュームの状態

RAID-5 ボリュームには、独自のボリューム状態があります。

- ◆ **CLEAN** – ボリュームは起動しておらず（カーネル状態が **DISABLED**）、パリティは正常です。RAID-5 プレックスには一貫性があります。
- ◆ **ACTIVE** – システムがリブートされたときに、ボリュームが起動していた（現在のカーネル状態が **ENABLED**）か、または使用されていた（そのときのカーネル状態が **ENABLED**）ことを示します。ボリュームが現在 **ENABLED** 状態である場合は、ボリュームが使用中なので RAID-5 プレックスの状態は固定されません。ボリュームが現在 **DISABLED** 状態である場合は、パリティの同期がとれていない可能性があります。
- ◆ **EMPTY** – ボリュームの内容が初期化されていません。ボリュームが **EMPTY** 状態の場合、カーネル状態は常に **DISABLED** です。
- ◆ **SYNC** – ボリュームがパリティの再同期化中である（現在のカーネル状態が **ENABLED**）か、システムのリブート時にパリティの再同期化が行われていた（カーネル状態が **DISABLED**）ことを示します。
- ◆ **NEEDSYNC** – 次のボリューム起動時に、パリティの再同期化を行う必要があります。
- ◆ **REPLAY** – ボリュームはログ再適用の一段階として一次的な状態にあります。ログの再適用は、ログが記録されたパリティおよびデータを使用する必要が生じた場合に行われます。



ボリュームのカーネル状態

ボリュームのカーネル状態は、ボリュームへアクセスできるかどうかを示します。ボリュームのカーネル状態によって、ボリュームの操作モードがオフライン (DISABLED)、保守 (DETACHED)、またはオンライン (ENABLED) に区分されます。

ボリュームのカーネル状態には、次の種類があります。

- ◆ DISABLED – ボリュームにアクセスすることはできません。
- ◆ DETACHED – ボリュームの読み書きはできませんが、プレックス デバイスの操作と ioctl 関数は処理することができます。
- ◆ ENABLED – ボリュームの読み書きを行うことができます。

RAID-5 ボリュームのリカバリ

ここでは、RAID-5 ボリュームの操作とリカバリについて説明します。RAID-5 ボリュームの詳細については、「Volume Manager とは (1 ページ)」を参照してください。

RAID-5 ボリュームのレイアウト

RAID-5 ボリュームは、1 つ以上のプレックスで構成され、各プレックスには 1 つ以上のサブディスクが含まれています。ミラー ボリュームとは異なり、RAID-5 ボリューム内のすべてのプレックスがボリューム データのミラーを保持していません。

RAID-5 ボリュームには、次の 2 つの種類プレックスが含まれます。

- ◆ RAID-5プレックス。ボリュームのデータおよびパリティの両方を保持するために使用します。
- ◆ ログ プレックス。より速く効率的にリカバリを行うために、ボリュームに書き込まれるデータのログを保持します。

RAID-5 プレックス

RAID-5 ボリュームは、単一の RAID-5 プレックス内にデータとパリティの両方を保持しています。RAID-5 プレックスは、ストライピング モデルと同様に、カラムに配列されたサブディスクで構成されています。次の表示は、RAID-5 プレックスに関する vxprint の出力結果です。

```
PL  NAME VOLUME KSTATE STATE LENGTH LAYOUT NCOL/WID MODE
SD  NAME PLEX DISK DISKOFFS LENGTH [COL/]OFF DEVICE MODE

pl  rvol-01 rvol ENABLED ACTIVE 20480 RAID 3/16 RW
sd  disk00-00 rvol-01 disk00 0 10240 0/0 c1t4d1 ENA
sd  disk01-00 rvol-01 disk01 0 10240 1/0 c1t2d1 ENA
sd  disk02-00 rvol-01 disk02 0 10240 2/0 c1t3d1 ENA
```



プレックスの行には、プレックスレイアウトが RAID、カラム数が 3、ストライプユニットサイズが 16 セクタと表示されています。各サブディスクの行には、プレックス内のカラム、そのカラムが位置するサブディスク内のオフセットが表示されています。

注 システムによっては、例とは異なるデバイス名を使用している場合があります。デバイス名の詳細については、「Volume Manager とは(1 ページ)」を参照してください。

RAID-5 ログ

各 RAID-5 ボリュームには、データとパリティが格納されている RAID-5 プレックスが 1 つ含まれます。ボリュームに関連付けられているその他のプレックスは、そのボリュームに書き込まれるデータとパリティについての情報を記録するために使用します。これらのプレックスを、RAID-5 ログプレックスまたは RAID-5 ログと呼びます。

RAID-5 ログは、コンカチネイテッド プレックスまたはストライププレックスのどちらかの編成になります。RAID-5 ボリュームに関連付けられている各 RAID-5 ログには、そのボリュームのログ情報の完全なコピーが保持されます。各 RAID-5 ボリュームには最低 2 つの RAID-5 ログプレックスを含めることをお勧めします。これらのログプレックスは、別々のディスクに配置する必要があります。各 RAID-5 ボリュームに 2 つの RAID-5 ログプレックスがあると、単一のディスクに障害が発生した場合でも、ログ情報が失われることはありません。

RAID-5 ディスクアレイに同時にアクセスするには、ログの大きさをプレックスのストライプサイズの数倍にしておく必要があります。

RAID-5 ボリュームの RAID-5 ログプレックスと RAID-5 プレックスを見分けるには、`vxprint` による出力を参照してください。ログプレックスの STATE フィールドは LOG と表示されています。次に示す表示は、RAID-5 ボリュームに対する `vxprint` の出力結果です。

```
V NAME USETYPE KSTATE STATE LENGTH READPOL PREFPLEX
PL NAME VOLUME KSTATE STATE LENGTH LAYOUT NCOL/WID MODE
SD NAME PLEX DISK DISKOFFS LENGTH [COL/]OFF DEVICE MODE

v r5vol raid5 ENABLED ACTIVE 20480 RAID -
pl r5vol-01 r5vol ENABLED ACTIVE 20480 RAID 3/16 RW
sd disk00-00 r5vol-01 disk00 0 10240 0/0 c1t4d1 ENA
sd disk01-00 r5vol-01 disk01 0 10240 1/0 c1t2d1 ENA
sd disk02-00 r5vol-01 disk02 0 10240 2/0 c1t3d1 ENA
pl r5vol-11 r5vol ENABLED LOG 1024 CONCAT - RW
sd disk03-01 r5vol-11 disk00 0 1024 0 c1t3d0 ENA
pl r5vol-12 r5vol ENABLED LOG 1024 CONCAT - RW
sd disk04-01 r5vol-12 disk02 0 1024 0 c1t1d1 ENA
```



ボリューム `r5vol` は、読み取りポリシー (READPOL) に RAID と表示されているため、RAID-5 ボリュームであることがわかります。このボリュームには、前述のように、RAID-5 プレックス (`r5vol-01`) が 1 つ含まれています。このボリュームは 2 つの RAID-5 ログを保持しており、それぞれプレックス `r5vol-11` および `r5vol-12` 内に含まれています。これらのログ プレックスは、STATE フィールドに LOG と表示されているため識別できます。これらは RAID-5 ボリュームに関連付けられていますが、レイアウトは RAID ではありません。

RAID-5 ボリュームの作成

`vxassist` (推奨) または `vxmake` を使用して RAID-5 ボリュームを作成できます。どちらのコマンドを使用する場合についても、この節で説明します。

RAID-5 ボリュームには、複数の物理ディスク上にある複数のサブディスクで構成されている RAID-5 プレックスが含まれています。1 つのボリュームに含めることのできる RAID-5 プレックスは 1 つのみです。

`vxassist` と RAID-5 ボリューム

`vxassist` コマンドを使用して RAID-5 ボリュームを作成できます。

```
# vxassist make ボリューム名 サイズ layout=raid5
```

たとえば、10M の RAID-5 ボリューム `volraid` を作成するには、次のように入力します。

```
# vxassist make volraid 10m layout=raid5
```

このコマンドにより、デフォルトで指定されている数のディスク上に、デフォルトのストライプユニット サイズで RAID-5 ボリュームが作成されます。

RAID-5 ボリュームの初期化

`vxmake` を使用して作成されたが、初期化されていないか、または未初期化状態に設定されている RAID-5 ボリュームは、初期化する必要があります。

RAID-5 ボリュームは、`vxvol` により初期化できます。次のいずれかのコマンドを実行します。

```
# vxvol init zero ボリューム名
```

または

```
# vxvol start ボリューム名
```

`vxvol init zero` コマンドは、すべての RAID-5 ログ プレックスとボリューム全体にゼロを書き込みます。その後、ボリュームは ACTIVE 状態になります。

`vxvol start` コマンドは、ほかのすべてのコラム内にある対応するデータ ストライプ ユニットの排他的論理和 (XOR) 演算することによってパリティをリカバリします。`vxvol start` コマンドを実行すると、`vxvol init zero` コマンドを実行する場合よりも時間がかかりますが、RAID-5 ボリュームをすぐに使用することができます。

障害と RAID-5 ボリューム

障害には 2 つの種類があります。システム障害とディスク障害です。システム障害とは、オペレーティング システムや電源に問題が発生したことにより、システムが突然稼働しなくなる状態をいいます。ディスク障害とは、システム障害 (ヘッドの破損、ディスク上の電子的障害、またはディスク コントローラの故障など) により、いくつかのディスク上のデータが使用できなくなる状態をいいます。

システム障害

RAID-5 ボリュームは、ディスク障害が発生した場合でも、最小限のディスク領域のオーバーヘッドで使用できるように設計されています。しかし、システム障害が発生した場合は、多くの形式の RAID-5 でデータが失われる可能性があります。システム障害が発生すると、ボリューム内のデータとパリティが同期しなくなるため、データの損失が発生します。障害発生時に未完了であった書き込み処理の状態を判断できなくなると、同期しなくなります。

RAID-5 ボリュームへのアクセス中に同期が失われると、ボリュームに陳腐化したパリティが含まれていると判断されます。この場合、各ストライプ内のパリティ以外のすべてのコラムを読み取り、パリティを再計算し、ストライプにパリティ ストライプ ユニットを書き込んでパリティを復元する必要があります。ボリューム内のすべてのストライプに対してこの処置を行う必要があるため、処理を完了するまでにはかなりの時間がかかります。

注意 再同期化処理中に、ディスク アレイ内のディスクに障害が発生すると、ボリュームのデータが失われます。この事態は、ログ ブレックスを含んでいない RAID-5 ボリュームのみで発生します。

再同期化処理は障害に対する耐性が弱いだけでなく、処理を行うことによりシステム リソースにかなりの負荷がかかり、処理速度が遅くなる場合があります。

RAID-5 ログには障害時に書き込まれたデータのコピーが保存されているため、このログにより、システム障害による被害を軽減することができます。再同期化処理には、このログから問題のデータとパリティを読み取り、ボリュームの適切な領域に書き込む操作も含まれます。これにより、データとパリティの再同期化に要する時間を大幅に短縮できます。さらに、このログがあれば、ボリュームがまったく陳腐化するということはありません。ボリューム内のすべてのストライプのデータとパリティが常に保持されているため、単一のディスクに障害が発生しても、ボリューム内のデータが失われることはありません。



ディスク障害

ディスク障害が発生すると、ディスク上のデータが使用できなくなることがあります。RAID-5 ボリュームの場合は、サブディスクが使用できなくなります。

ディスク障害は、ディスクへの書き込み中に修正不可能な入出力エラーが発生したことにより発生します。入出力エラーが原因で、システムのブート時にサブディスクがディスクアレイから切り離されたり、ディスクが使用できなくなったりする場合があります（ケーブル接続に問題がある場合や、ドライブの電源が切れている場合など）。

このような場合は、サブディスクを使用してデータを保持することができず、サブディスクは陳腐化していると判断されて切り離されます。下位のディスクが使用できるようになるか、または交換された場合も、サブディスクは依然として陳腐化していると判断されたままなので使用できません。

陳腐化したサブディスクに含まれるデータを読み取ろうとすると、ストライプ内のほかのすべてのストライプユニットのデータから該当するデータが復元されます。この処理を、復元読み取りと呼びます。これは、単にデータを読み取る以上に複雑な処理なので、読み取りパフォーマンスが低下する場合があります。RAID-5 ボリュームに陳腐化したディスクが含まれている場合、そのボリュームは縮退モードであるとみなされます。

縮退モードにある RAID-5 ボリュームは、以下に示すように、`vxprint` の出力で確認できます。

```
V NAME USETYPE KSTATE STATE LENGTH READPOL PREFPLEX
PL NAME VOLUME KSTATE STATE LENGTH LAYOUT NCOL/WID MODE
SD NAME PLEX DISK DISKOFFS LENGTH [COL/]OFF DEVICE MODE
v r5vol RAID-5 ENABLED DEGRADED 20480 RAID -
pl r5vol-01 r5vol ENABLED ACTIVE 20480 RAID 3/16 RW
sd disk00-00 r5vol-01 disk00 0 10240 0/0 c1t4d1
sd disk01-00 r5vol-01 disk01 0 10240 1/0 c1t2d1 dS
sd disk02-00 r5vol-01 disk02 0 10240 2/0 c1t3d1 -
pl r5vol-11 r5vol ENABLED LOG 1024 CONCAT - RW
sd disk03-01 r5vol-11 disk00 10240 1024 0 c1t3d0 -
pl r5vol-12 r5vol ENABLED LOG 1024 CONCAT - RW
sd disk04-01 r5vol-12 disk02 10240 1024 0 c1t1d1 -
```

ボリューム `r5vol` は、`STATE` フィールドに `DEGRADED` と表示されているため、縮退モードであることがわかります。最後のカラムにフラグが表示されているため、障害が発生しているサブディスクは `disk01-00` であることも確認できます。d はサブディスクが切り離されていることを示し、s はサブディスクのデータが陳腐化していることを示します。

RAID-5 ログが記録されているディスクに障害が発生する場合があります。この場合、ボリュームの運用に直接影響はありません。ただし、ボリューム上のすべての RAID-5 ログが失われると、全体的な障害に対するボリュームの耐性がなくなります。vxprint -ht の出力では、RAID-5 ログ プレックス内の障害は、BADLOG であるプレックスの状態によって示されます。次の表示では、RAID-5 ログ プレックス r5vol-11 に障害が発生していることが示されています。

```
V NAME USETYPE KSTATE STATE LENGTH READPOL PREFPLEX
PL NAME VOLUME KSTATE STATE LENGTH LAYOUT NCOL/WID MODE
SD NAME PLEX DISK DISKOFFS LENGTH [COL/]OFF DEVICE MODE
v r5vol RAID-5 ENABLED ACTIVE 20480 RAID -
pl r5vol-01 r5vol ENABLED ACTIVE 20480 RAID 3/16 RW
sd disk00-00 r5vol-01 disk00 0 10240 0/0 c1t4d1 ENA
sd disk01-00 r5vol-01 disk01 0 10240 1/0 c1t2d1 dS
sd disk02-00 r5vol-01 disk02 0 10240 2/0 c1t3d1 ENA
pl r5vol-11 r5vol DISABLED BADLOG 1024 CONCAT - RW
sd disk03-01 r5vol-11 disk00 10240 1024 0 c1t3d0 ENA
pl r5vol-12 r5vol ENABLED LOG 1024 CONCAT - RW
sd disk04-01 r5vol-12 disk02 10240 1024 0 c1t1d1 ENA
```

RAID-5 のリカバリ

通常、RAID-5 ボリュームのリカバリには、次のような処理が必要です。

- ◆ パリティーの再同期化
- ◆ 無効なサブディスクのリカバリ
- ◆ ログ プレックスのリカバリ

これらのリカバリ処理については、以降に説明します。一般に、パリティーの再同期化および陳腐化したサブディスクのリカバリは、次の場合に実行されます。

- ◆ RAID-5 ボリュームのブート時
- ◆ システムのブート直後
- ◆ vxrecover コマンドを呼び出して実行

RAID-5 ボリュームの起動の詳細については、「RAID-5 ボリュームの起動 (264 ページ)」を参照してください。



ディスク障害の発生時にホットリロケーション機能が有効になっている場合は、リロケーションに使用できるディスク領域が不足している場合を除いて、システム管理者による介入は必要ありません。障害が発生するとホットリロケーションが自動的に行われ、システム管理者に障害の発生を通知する電子メールが送信されます。

ホットリロケーションは、障害が発生した RAID-5 プレックスのサブディスクを自動的にリロケートします。リロケートが行われた後で、ホットリロケーション デーモン (vxrelocd) によりパリティの再同期化が開始されます。

RAID-5 ログ プレックスに障害が発生している場合は、そのログ プレックスがミラーリングされている場合にのみリロケートが行われます。その後、vxrelocd はミラーの再同期化処理を開始して RAID-5 ログ プレックスを作成し直します。障害発生時にホットリロケーションが無効になっている場合は、システム管理者が再同期化処理またはリカバリ処理を開始する必要があります。

パリティのリカバリ

通常は、RAID-5 ディスク アレイには陳腐化したパリティは含まれていません。陳腐化したパリティが含まれるのは、RAID-5 ボリュームのすべての RAID-5 ログ プレックスに障害が発生し、さらにシステム障害が発生した場合だけです。RAID-5 ボリュームに陳腐化したパリティが含まれている場合でも、通常は、ボリュームの起動中に修復されます。

有効な RAID-5 ログがないボリュームを起動し、そのボリュームが再同期化される前にプロセスを強制終了すると、アクティブなボリュームに陳腐化したパリティが含まれることとなります。これは、次に示すように、vxprint -ht コマンドの出力でボリューム状態をチェックすると確認できます。

```
V NAME USETYPE KSTATE STATE LENGTH READPOL PREFPLEX
PL NAME VOLUME KSTATE STATE LENGTH LAYOUT NCOL/WID MODE
SD NAME PLEX DISK DISKOFFS LENGTH [COL/]OFF DEVICE MODE
v r5vol RAID-5 ENABLED NEEDSYNC 20480 RAID -
pl r5vol-01 r5vol ENABLED ACTIVE 20480 RAID 3/16 RW
sd disk00-00 r5vol-01 disk00 0 10240 0/0 c1t4d1 ENA
sd disk01-00 r5vol-01 disk01 0 10240 1/0 c1t2d1 ENA
sd disk02-00 r5vol-01 disk02 0 10240 2/0 c1t3d1 ENA
```



この出力には、ボリュームの状態が `NEEDSYNC` と表示されています。これは、パリティを再同期化する必要があることを示しています。ボリューム状態には `SYNC` もあり、起動時に同期化処理が開始され、その処理が進行中であることを示しています。こうした処理が行われていない場合、またはボリューム状態が `NEEDSYNC` である場合は、`resync` キーワードを指定して `vxvol` コマンドを実行し、手作業で再同期化処理を開始することができます。たとえば、上記の RAID-5 ボリュームを再同期化するには、次のコマンドを実行します。

```
# vxvol resync r5vol
```

`VOL_R5_RESYNC ioctl` を RAID-5 ボリュームに対して発行すると、パリティが再生成されます。再同期化処理は RAID-5 ボリュームの先頭から開始され、`-o 入出力サイズオプション` で指定したセクタ番号と同じ番号の領域を再同期化します。`-o 入出力サイズオプション` が指定されていない場合、デフォルトの最大入出力サイズが使用されます。RAID-5 ボリューム全体が再同期化されるまで、再同期化処理は次の領域で続行されます。

大きなボリュームの場合、パリティの再生成には長い時間がかかります。処理が完了しないうちに、システムがシャットダウンまたはクラッシュする可能性があります。システムがシャットダウンした場合、リブートを挟んでパリティ再生成の処理が継続されなければ、処理を最初からやり直す必要があります。

処理を始めからやり直さなくて済むように、パリティ再生成の処理ではチェックポイントが保存されます。これは、パリティの再生成がどこまで進められたかを示すオフセットを設定データベースに保存するということです。`-o checkpoint=size` オプションによって、チェックポイントを保存する間隔を設定します。このオプションを指定していない場合、デフォルトのチェックポイント サイズが使用されます。

チェックポイントのオフセットを保存するにはトランザクションが必要となるため、チェックポイント サイズを小さく設定しすぎると、パリティの再生成にかかる時間が長くなる場合があります。システムのリポート後、ボリュームのサイズよりも小さいチェックポイント オフセットを持つ RAID-5 ボリュームでは、チェックポイントのオフセットからパリティの再同期化が開始されます。

サブディスクのリカバリ

陳腐化したサブディスクのリカバリは、通常、ボリュームの起動時に行われます。ただし、リカバリ処理を行うプロセスが突然停止したり、サブディスクのリカバリを行わないようにするオプションが指定された状態でボリュームが起動される場合も考えられます。さらに、サブディスクが含まれているディスクが、リカバリ処理を行わずに交換されていることもあります。いずれの場合も、`recover` キーワードを指定して `vxvol` コマンドを実行すると、サブディスクをリカバリできます。たとえば、265 ページの図 26 「無効な RAID-5 ボリューム」に示すような、RAID-5 ボリューム内の陳腐化したサブディスクをリカバリするには、次のコマンドを実行します。

```
# vxvol recover r5vol disk01-00
```



複数の無効なサブディスクを持つ RAID-5 ボリュームをすべて同時に修正することができます。複数の無効なサブディスクを確認するには、次のようにボリューム名だけを指定して `vxvol recover` を呼び出します。

```
# vxvol recover r5vol
```

障害後のログのリカバリ

RAID-5 ログプレックスは、ディスク障害によって、切り離されることがあります。切り離された RAID-5 ログは、`att` キーワードを指定して `vxplex` コマンドを実行すると、再び追加することができます。270 ページの図 27「読み取り - 変更 - 書き込み」に示されているような障害が発生した RAID-5 ログプレックスを再追加するには、次のコマンドを実行します。

```
# vxplex att r5vol r5vol-l1
```

RAID-5 のさまざまな処理

RAID-5 ボリュームおよび関連オブジェクトを操作するためのさまざまな処理があります。通常これらの処理は、ディスクを取り外すなどの大規模な処理を行う場合に、その一部として `vxassist` や `vxrecover` などのコマンドにより実行されます。このようなコマンドラインからの処理は、Volume Manager の基本操作では必要ありません。

RAID-5 ログの操作

RAID-5 ログは RAID-5 ボリュームのプレックスとして扱われるので、`vxplex` コマンドを使用して操作します。RAID-5 ログを追加するには、`vxplex att` を使用します。

```
# vxplex att r5vol r5log
```

ログの追加は、新しいログがストライプ上のデータをすべて保持するのに十分な大きさである場合にのみ実行されます。RAID-5 ボリュームに既にログがある場合、新しいログのサイズには、少なくともそれらのログと同等のサイズが必要です。これは、新しいログを古いログのミラーとして使用するためです。

RAID-5 ボリュームが使用可能でない場合、新しいログは BADLOG とみなされ、ボリュームの起動時に有効になります。ただし、ログの内容は無視されます。

RAID-5 ボリュームが使用可能で、ほかにも有効な RAID-5 ログがある場合、新しいログの内容は `ATOMIC_COPY ioctl` によってほかのログと同期がとられます。

RAID-5 ボリュームに現在有効なログがない場合は、新しいログが有効になる前に、その内容が消去されます。

ログプレックスをボリュームから削除するには、`vxplex dis` コマンドを実行します。

```
# vxplex dis r5log3
```



ログを削除することによりボリューム内に有効なログが 1 つ以下になってしまう場合は、警告メッセージが表示され、処理は続行できなくなります。処理を強制的に続行するには、`-o force` オプションを使用します。

RAID-5 サブディスクの操作

ほかのサブディスクと同様、RAID-5 ボリュームの RAID-5 ブレックスのサブディスクを操作するには、`vxsd` コマンドを実行します。ストライプ ブレックスの場合と同様にキーワード `assoc` を使用して関連付けを行います。たとえば、前述の RAID-5 ボリュームの各カラム末尾にサブディスクを追加するには、次のコマンドを実行します。

```
# vxsd assoc r5vol-01 disk10-01:0 disk11-01:1 disk12-01:2
```

サブディスクがブレックス内の「ホール」を埋めている（ボリュームの論理アドレス空間の一部がサブディスクに対応づけられている）場合、そのサブディスクは陳腐化していると解釈されます。RAID-5 ボリュームが使用可能である場合、関連付け処理によって、`VOL_R5_RECOVER ioctl` を使用してサブディスク上にあるデータが再生成されます。それ以外の場合は、データは陳腐化しているとみなされ、ボリュームの起動時にリカバリされます。

サブディスクを RAID-5 ブレックスから削除するには、`vxsd dis` を実行します。

```
# vxsd dis disk10-01
```

注意 サブディスクが RAID-5 ボリュームのアドレス空間の一部をマップしている場合、ボリュームは DEGRADED（縮退）モードに設定されます。この場合、`dis` を使用して処理を行うと、警告メッセージが表示されます。強制的に処理を続行するには、`-o force` オプションを使用します。また、同じストライプ内の別のサブディスクが使用できないか見つからず、ボリューム状態が `DISABLED` または空ではない場合、サブディスクを削除すると RAID-5 ボリュームが使用できなくなります。この場合は、サブディスクを削除できません。

RAID-5 ボリュームが割り当てられているディスクを変更するためにサブディスクを移動するには、`vxsd mv` を使用します。たとえば、`disk03` の内容を退避する必要があり、`disk22` の 2 つのサブディスクを合せれば十分な空き領域がある場合、次のコマンドを使用できます。

```
# vxsd mv disk03-01 disk22-01 disk22-02
```

このコマンドはストライプ ブレックスに使用するものと似ていますが、実際行われる処理は異なります。



RAID-5 サブディスクの移動

RAID-5 サブディスクを移動する場合、現在のサブディスクが RAID-5 ブレックスから削除され、新しいサブディスクに交換されます。新しいサブディスクは陳腐化しているとみなされ、VOL_R5_RECOVER を使用してリカバリ処理が行われます。リカバリは、vxsd によって、またはボリュームの起動時（ボリュームがアクティブでない場合）に実行されます。つまり、移動処理中の RAID-5 ボリュームの機能は低下します。

移動中にストライプ内で別の障害が発生すると、ボリュームが使用できなくなります。また、パリティが陳腐化した場合も、RAID-5 ボリュームは無効になります。

このような事態を避けるため、次の場合には、vxsd ユーティリティを使用してサブディスクを移動することはできません。

- ◆ 移動対象のサブディスク中のストライプと同じものが陳腐化したサブディスクに含まれている場合
- ◆ RAID-5 ボリュームが停止しているが、正常な状態で終了していない場合（パリティが陳腐化しているとみなされます）
- ◆ RAID-5 ボリュームがアクティブだが、有効なログ領域がない場合

上記の 3 番目の場合のみ、`-o force` オプションを使用して処理を強制的に続行することができます。

RAID-5 ボリュームのサブディスクは、`vxsd split` を使用して分割したり、`vxsd join` を使用して結合したりすることができます。これらの処理は、ミラー ボリュームの場合と同様に行われます。

注 RAID-5 サブディスクは、冗長性を損なうことなく、ほかのサブディスクの場合と同様に移動できます。

RAID-5 ボリュームの起動

RAID-5 ボリュームを起動すると、さまざまな状態になります。システムを通常通り停止した後は、ボリュームは正常な状態であるためリカバリの必要はありません。ただし、システムがクラッシュする前にボリュームを終了していないか、マウント解除していない場合は、次の起動時にリカバリ処理を行う必要があります。ボリュームは、リカバリ後に使用できるようになります。ここでは、特定の条件下でとるべき処置について説明します。

通常は、ボリュームはシステムのリブート後に自動的に起動し、リカバリ処理は自動的に行われるか、または `vxrecover` コマンドによって行われます。

RAID-5 ボリュームを起動できない場合

次のように、RAID-5 プレックスの一部がボリューム サイズに対応しない場合は、RAID-5 ボリュームを使用することはできません。

- ◆ RAID-5 プレックスのサイズが、RAID-5 ボリュームのサイズより小さく、空白があってははいけません。
- ◆ ストライプ内の2つのサブディスクに障害が発生している領域にRAID-5プレックスを対応づけることはできません。そのようなサブディスクの障害の原因としては、陳腐化または基となるディスクの故障が考えられます。

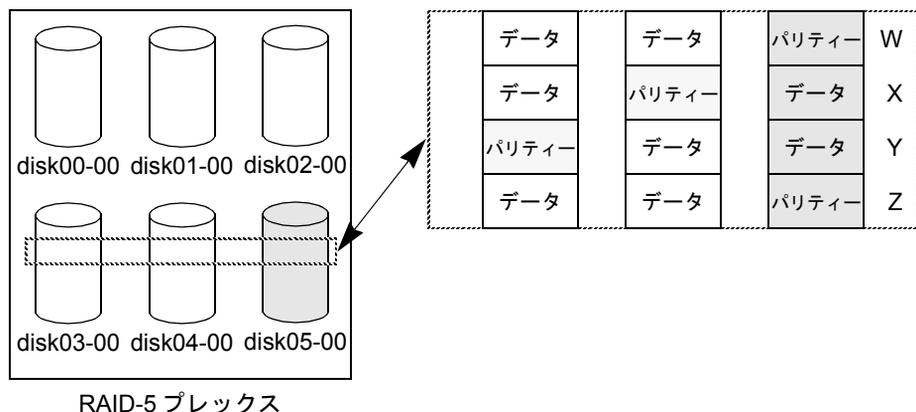
このような場合は、`vxvol start` コマンドを実行すると、次のエラーメッセージが返されます。

```
vxvm:vxvol:ERROR: ボリューム r5vol を起動できません。RAID5 プレックスにボリューム全体の長さがマップされていません。
```

この時点では、RAID-5 ボリュームの内容を使用することはできません。

ほかに RAID-5 ボリュームを起動できなくなる原因として、パリティが陳腐化し、さらにサブディスクが切り離されているか陳腐化している場合が考えられます。この状態では、障害が発生しているサブディスクに対応するストライプ内で、パリティストライプユニットが無効（パリティが陳腐化しているため）であり、かつ不良なサブディスク上のストライプユニットも無効になっています。図 26「無効な RAID-5 ボリューム」に、この状態を示します。パリティが陳腐化し、さらにサブディスクが障害を起したために、ボリュームが使用不能になっています。

図 26. 無効な RAID-5 ボリューム



この例では、RAID-5 ディスク アレイ内に 4 つのストライプがあります。すべてのパリティが陳腐化しており、サブディスク disk05-00 には障害が発生しています。ストライプ X および Y 内では 2 つの障害が発生しているため、これらのストライプは使用できません。

ストライプ内で障害が 2 つ発生したとみなされ、ボリュームは使用できなくなります。この場合、`vxvol start` コマンドを実行すると、次のように表示されます。

```
vxvm:vxvol:ERROR: ボリューム r5vol は起動できません。一部のサブディスクが使用できず、パリティが無効です。
```

こうした事態を回避するために、RAID-5 ボリューム内では常に複数の RAID-5 プレックスを使用してください。複数の RAID-5 ログプレックスにより、ボリューム内のパリティが陳腐化することを防止できるため、このような事態は発生しなくなります（「システム障害 (257 ページ)」を参照）。

RAID-5 ボリュームを強制的に起動する場合

サブディスクが陳腐化していると判断されている場合でも、ボリュームを起動することができます。たとえば、停止しているボリュームに無効なパリティが陳腐化しており、RAID-5 ログがないディスクを切り離された後で再び追加した場合などが想定されます。

このような場合は、サブディスクが使用できないときにボリュームが使用されたため、データが最新でもサブディスクは陳腐化していると判断され、RAID-5 ボリュームも無効とみなされます。このような問題を避けるため、ディスク アレイに関連付けられている有効な RAID-5 ログを常に複数保持するようにしてください。ただし、これが不可能な場合もあります。

陳腐化しているサブディスクが含まれている RAID-5 ボリュームを起動するには、`-f` オプションを指定して `vxvol start` コマンドを実行します。このようにすると、すべての陳腐化したサブディスクが陳腐化していないとみなされます。これは、`start` 処理によって RAID-5 ボリュームの有効性および開始に必要な条件が評価される前に行われます。また、`vxmend fix unstale サブディスク` コマンドを使用すると、個々のサブディスクが陳腐化していないとみなされます。

RAID-5 ボリューム起動時のリカバリ

RAID-5 ボリュームの内容を完全にリストアして使用できるようにするには、いくつかの処理を行う必要があります。ボリュームが起動する場合は常に、RAID-5 ログ ブレックスの内容がすべて消去されてから、ボリュームが起動します。これにより、攪乱データがログ エントリとして解釈されないので、ボリュームの内容も破損されません。また、いくつかのサブディスクのリカバリまたはパリティの再同期化が必要な場合もあります (RAID-5 ログに障害が発生している場合)。

RAID-5 ボリュームの起動時には、次の処理が実行されます。

1. RAID-5 ボリュームが正常に終了されていない場合、有効な RAID-5 ログ ブレックスがあるかどうかを確認されます。

- ◆ 有効なログ ブレックスがある場合は、再適用されます。この処理は、ボリュームのカーネル状態を `DETACHED`、ボリューム状態を `REPLAY` に設定し、RAID-5 ログ ブレックスを有効にすることによって行われます。ログの読み取りと再適用が正常に行われた場合は、手順 2 の処理が行われます。
- ◆ 有効なログがない場合は、パリティの再同期化処理を行う必要があります。再同期化処理は、ボリュームのカーネル状態を `DETACHED`、ボリューム状態を `SYNC` にして実行します。ログはすべて `DISABLED` 状態になります。

パリティの再同期化処理が行われている間は、ボリュームを使用することはできません。再同期化処理中にサブディスクに障害が発生すると、ボリュームを使用することができなくなるからです。ただし、`-o unsafe` オプションを指定して `vxvol start` を実行すると、強制的にボリュームを起動することができます。陳腐化しているサブディスクがある場合は、RAID-5 ボリュームを使用できません。

注意 `-o unsafe start` オプションを使用すると、ボリュームの内容が使用できなくなる可能性が高いため危険です。このオプションはできる限り使用しないでください。

2. 既存のログの内容が消去されて使用可能になります。この処理中にすべてのログで障害が発生すると、ボリュームの起動プロセスが異常終了します。
3. 陳腐化しているサブディスクがない場合や、陳腐化しているサブディスクをリカバリできる場合、ボリュームのカーネル状態は `ENABLED` に、ボリューム状態は `ACTIVE` に設定されます。これで、ボリュームが起動します。
4. 陳腐化しているのにリカバリする必要のあるサブディスクがあるが、有効なログがある場合は、カーネル状態を `ENABLED` にすることによってボリュームを有効にすると、そのボリュームをサブディスクのリカバリ中に使用できるようになります。それ以外の場合、ボリュームのカーネル状態は `DETACHED` に設定され、サブディスクのリカバリ中にそのボリュームを使用することはできません。



システムがクラッシュしたり、アクティブなボリュームが稼働中に強制的に停止されたりすると、パリティが陳腐化し、ボリュームを使用できなくなるため、この処理が行われます。この処理を行わない場合は、`-o unsafe start` オプションを指定して強制的にボリュームを起動できます。

注意 `-o unsafe start` オプションを使用すると、ボリュームの内容が使用できなくなる可能性が高いため危険です。このオプションはできる限り使用しないでください。

5. ボリューム状態が **RECOVER** に設定され、陳腐化しているボリュームがリストアされます。各サブディスク上のデータが有効になるので、サブディスクは陳腐化していないとみなされるようになります。
6. サブディスクのリカバリに失敗したが有効なログがない場合、サブディスクは陳腐化したままで、システムがクラッシュすると **RAID-5** ボリュームが使用できなくなるため、ボリュームの起動処理は異常終了します。強制的にボリュームを起動するには、`-o unsafe` オプションを使用します。

注意 `-o unsafe start` オプションを使用すると、ボリュームの内容が使用できなくなる可能性が高いため危険です。このオプションはできる限り使用しないでください。

ボリュームに有効なログがある場合は、サブディスクのリカバリに失敗しても、ボリュームの起動処理は停止しません。

7. すべてのサブディスクのリカバリが完了すると、ボリュームのカーネル状態が **ENABLED** になり、ボリューム状態が **ACTIVE** になります。これで、ボリュームが起動しました。

RAID-5 ボリュームの属性の変更

RAID-5 ボリュームの属性は変更することができます。RAID-5 ボリュームでは、`vxvol set` コマンドを使用すると、ボリュームおよび RAID-5 ログのサイズを変更できます。RAID-5 ボリュームのサイズを変更するには、次のコマンドを実行します。

```
# vxvol set len=10240 r5vol
```

ボリュームのサイズは、RAID-5 プレックスに割り当てられた領域（連続長）より大きくすることはできません。ボリュームが使用できなくなるため、ボリュームサイズを拡大することはできません。RAID-5 ボリュームがアクティブなときにサイズを縮小するには、`-o force` 使用タイプ オプションを使用して強制的に実行する必要があります。そのボリュームを使用しているアプリケーションから領域が削除されないようにするためです。

RAID-5 ログのサイズを変更する場合も、`vxvol set` コマンドを実行します。

```
# vxvol set loglen=2M r5vol
```

RAID-5 ログ ブレックスは、RAID-5 ボリュームのログ サイズ全体をカバーする場合のみ有効です。ログのサイズを拡大すると RAID-5 ログが無効になるような場合、その処理は実行できません。また、ボリュームが停止していて、正常に終了されていない場合は、ログ サイズを変更できません。これは、ログの内容が失われること（ログ サイズが縮小された場合）、または攪乱データがログに記録されること（ログ サイズが拡大された場合）を避けるためです。

RAID-5 ディスク アレイへの書き込み

ここでは、RAID-5 ディスク アレイへの書き込み処理を説明します。

読み取り - 変更 - 書き込み

RAID-5 ディスク アレイへの書き込みでは、入出力に関わる各ストライプに対して、次の処理が実行されます。

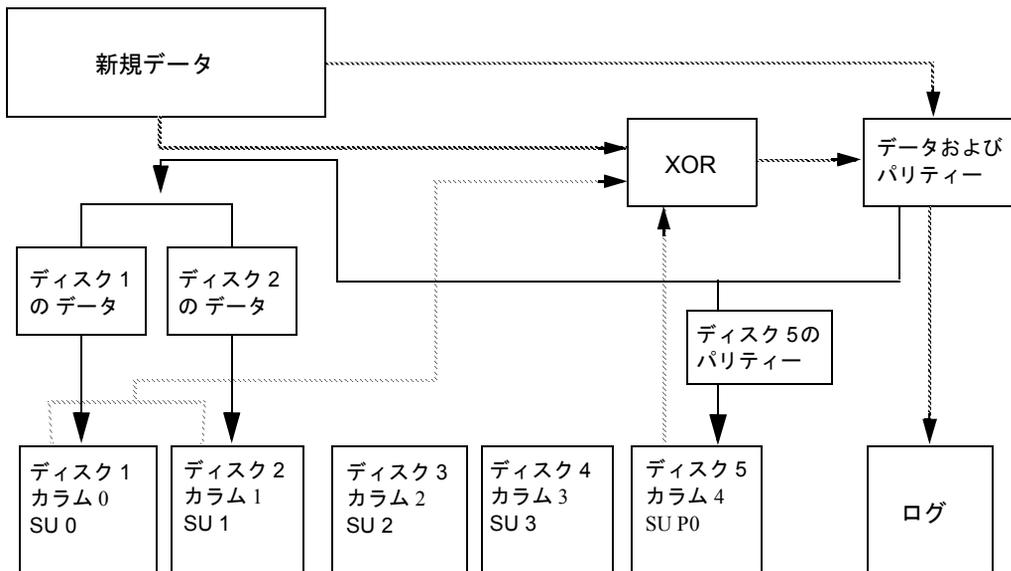
1. 新しい書き込みデータで更新されるデータ ストライプ ユニットがアクセスされ、内部バッファに読み取られます。パリティ ストライプ ユニットも内部バッファに読み取られます。
2. パリティが新しいデータ領域の内容を反映して更新されます。まず、古いデータ内容がパリティと排他的論理和 (XOR) 演算されます (論理的に古いデータを削除)。次に、新しいデータがパリティと排他的論理和 (XOR) 演算されます (論理的に新しいデータを追加)。新しいデータとパリティがログに書き込まれます。
3. 新しいパリティがパリティ ストライプ ユニットに書き込まれます。新しいデータはデータ ストライプ ユニットに書き込まれます。すべてのストライプ ユニットに、1 回の処理で書き込みが行われます。

この処理を読み取り - 変更 - 書き込みサイクルと言います。これが RAID-5 のデフォルトの書き込み処理です。ディスクに障害が発生した場合、そのディスク上のデータとパリティは使用できなくなります。その場合、ディスク アレイは縮退モードで動作します。



図 27 「読み取り - 変更 - 書き込み」は、読み取り - 変更 - 書き込みサイクルを示します。

図 27. 読み取り - 変更 - 書き込み



SU = ストライプ ユニット

- 手順 1: データをパリティ ストライプ ユニット P0 およびデータ ストライプ ユニット 0 および 1 から読み取ります。
- 手順 2: データとパリティを排他的論理和 (XOR) 演算して新しいパリティを計算します。新しいデータとパリティを記録します。
- 手順 3: 新しいパリティ (XOR から算出) をパリティ ストライプ ユニット P0 へ、新しいデータをデータ ストライプ ユニット 0 および 1 に書き込みます。

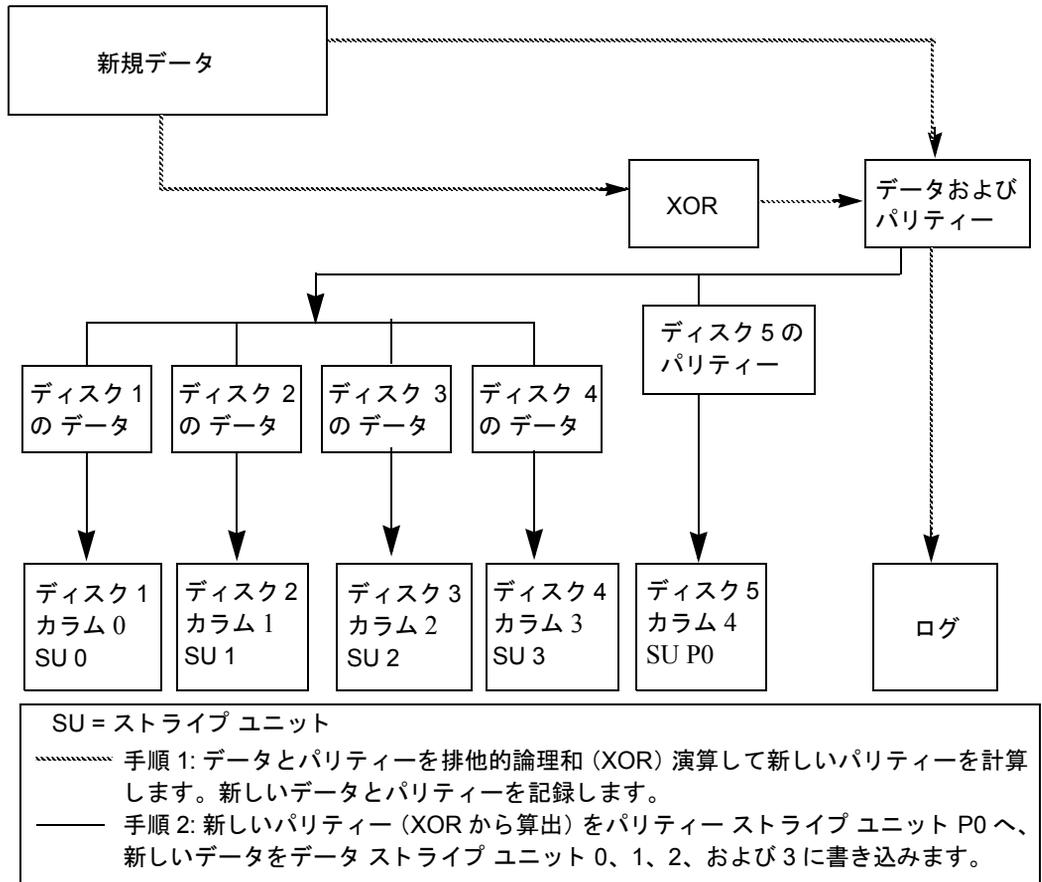
全ストライプ書き込み

大規模な書き込み（データ ストライプ全体を使用する書き込み）が要求された場合、読み取り - 変更 - 書き込みサイクルではなく、全ストライプ書き込みが行われます。全ストライプ書き込みでは、読み取りが必要ないため、読み取り - 変更 - 書き込みサイクルよりも処理速度が向上します。読み取りサイクルを省くことにより、ディスクへの書き込みに必要な入出力時間が短縮されます。全ストライプ書き込みは、次の手順で実行されます。

1. 新しいデータ ストライプ ユニットがすべて一緒に排他的論理和（XOR）され、新しいパリティ値が算出されます。新しいデータとパリティがログに書き込まれます。
2. 新しいパリティがパリティ ストライプ ユニットに書き込まれます。新しいデータはデータ ストライプ ユニットに書き込まれます。ストライプ全体に、1 回の処理で書き込みが行われます。

図 28 「全ストライプ書き込み」は、全ストライプ書き込みの流れを示します。

図 28. 全ストライプ書き込み



復元書き込み

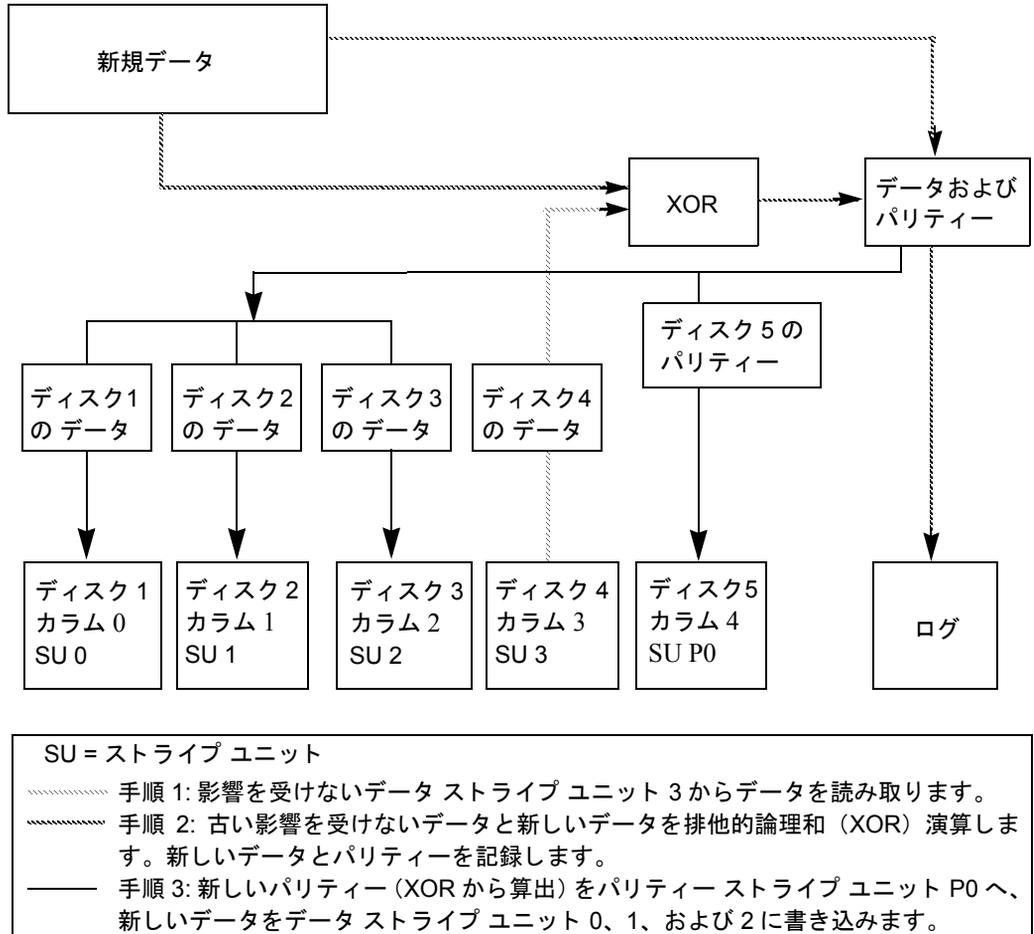
1 回の入出力でデータ ディスクの 50 パーセント以上が書き込み対象になる場合、復元書き込みが行われます。復元書き込みでは、排他的論理和 (XOR) 演算により入出力時間が短縮されます。排他的論理和 (XOR) 演算では、パリティ領域の読み取りが不要なく、影響を受けないデータのみが読み取られます。影響を受けていないデータは、ストライプ内のストライプユニットの 50 パーセント未満です。

復元書き込みは、次の手順で実行されます。

1. 影響を受けないデータが変更されていないデータ ストライプ ユニットから読み取られます。
2. 新しいデータが古い影響を受けないデータと排他的論理和 (XOR) 演算され、新しいパリティ ストライプ ユニットが生成されます。新データと算出されたパリティがログに記録されます。
3. 新しいパリティがパリティ ストライプ ユニットに書き込まれます。新しいデータはデータ ストライプ ユニットに書き込まれます。すべてのストライプ ユニットに、1 回の処理で書き込みが行われます。

図 29 「復元書き込み」は、復元書き込みの流れを示します。この例では、ディスクとパリティディスクではなく必要なデータディスクのみが読み取られるため、復元書き込みの方が読み取り - 変更 - 書き込みサイクルよりも適しています。

図 29. 復元書き込み



はじめに

論理ボリューム管理は、システムのパフォーマンスを全体的に向上させる手段です。この章では、パフォーマンス管理と構成の設定に関するガイドラインを紹介します。これらのガイドラインは、Volume Manager が提供するさまざまな機能を活用するのに役に立ちます。この章では、パフォーマンスの優先順位を決定するための情報を提供し、適切なデータを収集および使用する方法を説明します。

この章では、次のトピックについて説明します。

- ◆ パフォーマンスのガイドライン
- ◆ パフォーマンスの監視
- ◆ Volume Manager のチューニング

パフォーマンスのガイドライン

Volume Manager では、柔軟にストレージを設定してシステムのパフォーマンスを向上させることができます。パフォーマンスを最適化するには、次の2つの基本的な方法があります。

- ◆ データを物理ディスクに割り当て、使用可能なディスク間に入出力負荷を均等に分散する。
- ◆ 最もアクセス頻度の高いデータを識別し、ストライピングやミラーリングによって、そのデータへのアクセス帯域幅を広く確保する。

Volume Manager では、ミラーリングおよび RAID-5 によりデータの冗長性が提供されるため、ディスクに障害が発生した場合もデータへのアクセスを続行できます。



データの配置

ファイルシステムの配置場所を決定する際、通常システム管理者は、複数の使用可能なディスクドライブに入出力負荷を分散させようと考えます。しかし、このように考えて配置したとしても、将来の利用パターンの予測は困難であり、また複数のファイルシステムを複数のドライブに分割できないため、その効果には制限があります。たとえば、単一のファイルシステムにディスクアクセスが集中している場合、そのファイルシステムを別のドライブに配置すると、ボトルネックもそのドライブに移ります。

Volume Manager では、ボリュームを複数のドライブにわたって分割できるため、詳細レベルでデータを配置できます。システム管理者は実際のアクセスパターンを確認してから、ファイルシステムの配置を調整することができます。ボリュームの可用性に悪影響を与えることなく、パフォーマンスパターンを設定または変更した後でボリュームをオンラインで再設定できます。

ストライピング

ストライピングは、データを「スライス」し、複数のデバイスに分散して格納することにより、アクセスのパフォーマンスを向上させる手法です。ストライピングを適用すると、プレックスへのアクセス帯域幅を広く確保できます。ストライププレックスでは、読み取りおよび書き込みの両方でアクセスのパフォーマンスが向上します。

アクセスが集中するボリューム（ファイルシステムやデータベースを含むボリューム）を特定できれば、パフォーマンスの向上による利点を実感できます。「高トラフィック」のデータを複数ディスクの領域へストライピングすることにより、そのデータへのアクセス帯域幅を広く確保できるようになります。

図 30 「ストライピングを使用したデータアクセスの最適化」に、データアクセスのボトルネックとなっている単一のボリューム（Hot Vol）を示します。このボリュームは、4つのディスクにストライピングされており、その4つのディスクの残りの領域は使用頻度の低いボリュームで使用できるように空けられています。

図 30. ストライピングを使用したデータアクセスの最適化



ミラーリング

ミラーリングは、システム上に複数のコピーを格納する手法です。ミラーリングを正しく適用すれば、物理メディアに発生した障害によるデータの損失を回避できるため、障害時でもデータへのアクセスを続行できます。ミラーリングを適用すると、システムクラッシュやディスク障害の発生時に、データをリカバリできる可能性が高くなります。

システムのパフォーマンスを向上させるためにミラーリングを適用することもあります。アクセス頻度の高いデータをミラーリングすると、ディスク障害によるデータの損失を防げるだけでなく、入出力パフォーマンスを向上させることもできます。ただし、ストライピングとは異なり、ミラーリングのパフォーマンスはディスクアクセスの読み取り/書き込み比率によって左右されます。システムの作業負荷が主に書き込みに集中している場合（たとえば書き込み比率が 30% 超）は、ミラーリングで実現できるパフォーマンスはいくらか低下します。

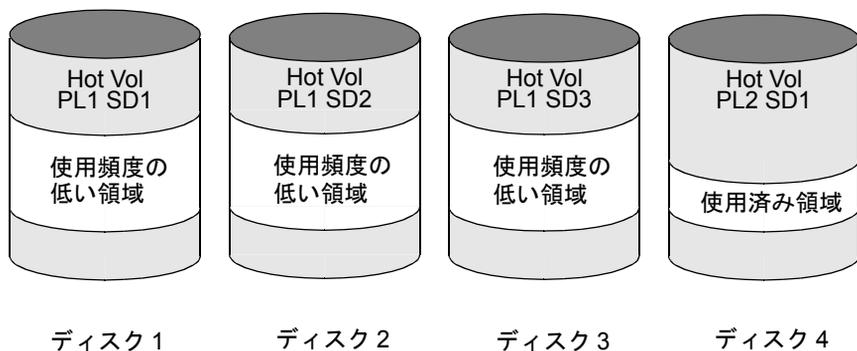
異なる種類のミラー ボリュームのパフォーマンスを最適化するために、**Volume Manager** では次の読み取りポリシーがサポートされています。

- ◆ ラウンドロビン読み取りポリシー (**round**) では、ボリューム内のすべてのプレックスを順に読み取ります（これは、プレックスが交替で使用されることを意味します）。
- ◆ 優先プレックス読み取りポリシー (**prefer**) では、指定したプレックス（通常はパフォーマンスが最も高いプレックス）に障害が発生しない限り、そのプレックスのみを読み取ります。指定プレックスに障害が発生した場合は、別のプレックスがアクセスされます。
- ◆ デフォルトの読み取りポリシー (**select**) は、設定に応じて適切な読み取りポリシーを自動的に選択します。たとえば、ボリュームに関連付けられているストライププレックスが 1 つしかない場合は優先プレックス方式が選択され、それ以外の場合は通常ラウンドロビン方式が選択されます。



図 31 「ミラーリング アンド ストライピングを応用したパフォーマンスの向上」の設定例では、ボリューム Hot Vol の読み取りポリシーに prefer を設定し、ストライプ プレックス PL1 を優先させる必要があります。このように PL1 を通して読み取ることで、単一のディスクに負荷が集中せず、使用頻度の低い複数のディスクに負荷を分散できます。

図 31. ミラーリング アンド ストライピングを応用したパフォーマンスの向上



作業負荷が読み取りに集中している場合、パフォーマンスを向上させるために、最大 32 のプレックスを同じボリュームに関連付けることができます。ただし、この方法ではディスク領域の有効利用度が低下します。使用可能なディスクの半分にわたってストライピングしてプレックスを1つ形成し、残りの半分を使用してプレックスをもう1つ作成する方法でもパフォーマンスを向上させることができます。通常、この方法が **Volume Manager** の構成を一連のディスクセット上で設定し、適切な信頼性と最高のパフォーマンスを実現するには最も適しています。可能な場合はこの方法をお勧めします。

ミラーリング アンド ストライピング

ミラーリングとストライピングを併用すると、複数のディスクにデータを分散でき、同時にデータの冗長性を持たせることができます。

複数の入出力ストリームがある場合、ミラーリング アンド ストライピングを併用すると、パフォーマンスを大幅に向上させることができます。ストライピングした場合、入出力が1つのストライプのストライプユニット全体に満遍なく行き渡ると、直列的なアクセスが向上します。入出力ストリームが複数のデバイスでパラレルに同時進行できるため、より優れたスループットを実現できます。

ミラーリングはディスク障害時のデータ損失を回避する場合によく使用されますが、作業負荷が書き込みに集中している場合にもミラーリングが必要になることがあります。この場合、ミラーリングをストライピングと組み合わせて、高い可用性とパフォーマンスの両方を実現することが可能です。

ストライピング アンド ミラーリング

ストライピングとミラーリングを併用すると、複数のディスクにデータを分散でき、同時にデータに冗長性を持たせることができます。

複数の入出力ストリームがある場合、ストライピングとミラーリングを併用するとパフォーマンスを大幅に向上させることができます。ストライピングした場合、入出力が1つのストライプのストライプ ユニット全体に満遍なく行き渡ると、直列的なアクセスが向上します。入出力ストリームが複数のデバイスでパラレルに同時進行できるため、より優れたスループットを実現できます。

ミラーリングはディスク障害時のデータ損失を回避する場合によく使用されますが、作業負荷が書き込みに集中している場合にもミラーリングが必要になることがあります。この場合、ミラーリングをストライピングと組み合わせ、高い可用性とパフォーマンスの両方を実現することが可能です。本書の「ミラーリング (277 ページ)」を参照してください。

RAID-5

RAID-5 を使用すると、ミラーリングとストライピングを併用した場合の利点の多くを享受できるだけでなく、必要なディスク領域が少なく済みます。RAID-5 の読み取りパフォーマンスはストライピングのパフォーマンスとほぼ等しく、さらにパリティによる冗長性もミラーリングによる冗長性とほぼ同等です。RAID-5 の弱点としては、書き込み速度が比較的遅いことなどが挙げられます。

注 アプリケーションのアクセス パターンとして、書き込みよりも読み取りの比率の方が非常に高い場合以外は、一般的に、パフォーマンスを向上させるために RAID-5 が使用されることはありません。



パフォーマンスの監視

システム管理者が優先すべき事項には、次の2種類あります。まず1つは、ハードウェアに関連する物理的な優先事項。もう1つはソフトウェアおよびその操作に関連する論理的な優先事項です。

パフォーマンスの優先事項

物理的なパフォーマンス特性により、各ドライブにおける入出力のバランスを取りドライブ内の入出力を局所化することで、シーク時間を最小限に抑えます。監視結果に基づいて、サブディスクの位置を移動させ、ディスク間でバランスを図ることができます。

論理的な優先事項には、ソフトウェア操作やその管理方法などが含まれます。監視結果に基づいて、特定のボリュームをミラーリングまたはストライピングして、そのパフォーマンスを向上させることが可能です。重要なボリュームのパフォーマンスを向上させるために、全体的なスループットが犠牲となる場合もあります。システムのどの事項を最優先し、それと引き換えにどの事項を切り捨てるかは、システム管理者のみが決定できます。

一般的に、すべてのボリュームを適当な数のディスクにわたってストライピングアンドミラーリングし、可能な場合はコントローラ間をまたがってミラーリングを行うと、最高のパフォーマンスを実現できます。これにより、すべてのディスクに負荷が均一に分散されます。ただし、そうすると **Volume Manager** の管理が一層難しくなります。ディスク数が多い（数百または数千）場合は、（ディスクグループを使用して）ディスクを10ずつのグループに分け、各グループをボリュームセットのストライピングアンドミラーリングに使用することができます。この方法を使用すると、高いパフォーマンスを発揮させながら、管理作業を容易にすることができます。

パフォーマンスデータの収集

Volume Manager は、入出力統計と入出力トレースの2種類のパフォーマンス情報を提供します。いずれの種類の情報もパフォーマンスを監視する上で役立ちます。**vxstat** ユーティリティを使用して入出力統計データを、**vxttrace** ユーティリティを使用して入出力トレースデータを収集します。これらのユーティリティについて以下に簡単に説明します。

入出力統計の収集（vxstat）

vxstat ユーティリティは、**Volume Manager** の管理下にあるボリューム、ブックス、サブディスク、およびディスクに関する活動情報を収集します。**vxstat** は、ブート時以降のオブジェクトの活動レベルを反映した統計を報告します。特定の **Volume Manager** オブジェクトまたはすべてのオブジェクトの統計を一度に表示できます。ディスクグループを指定して、それらのディスクグループ内にあるオブジェクトの統計のみを表示することもできます。ディスクグループを指定しない場合は、デフォルトで **rootdg** が指定されます。

表示される情報内容は、**vxstat** で指定するオプションによって異なります。指定可能なオプションの詳細については、**vxstat** (IM) マニュアルページを参照してください。

Volume Manager は、次の入出力統計を記録します。

- ◆ 処理回数
- ◆ 転送ブロック数 (1 回の処理で複数ブロックの処理が可能)
- ◆ 平均処理時間 (Volume Manager インタフェースを使用した処理の合計時間を反映した数値のため、ほかの統計プログラムとの比較には適しません)

Volume Manager は、上述の統計を論理入出力に対して記録します。統計には、各ボリュームに対する読み取り、書き込み、アトミックコピー、検証読み取り、検証書き込み、プレックスの読み取り、プレックスの書き込みなどが含まれます。2つのプレックスを含むボリュームに書き込みを1回行うと、少なくとも5回 (プレックスごとに1回ずつ、サブディスクごとに1回ずつ、およびボリュームに対して1回) の処理が実行されます。また、2つのサブディスクに渡って1回読み取る場合も、少なくとも4回 (各サブディスクごとに1回、プレックスに1回、ボリュームに1回) の読み取りが行われます。

Volume Manager では、ほかの統計データを保持することもできます。各プレックスごとに、読み取りおよび書き込みの失敗データが保持されます。ボリュームの場合、読み取りおよび書き込みの失敗データと共に修正済みのデータが保持されます。

vxstat では、統計情報をゼロにリセットすることもできます。コマンド `vxstat -r` を使用すると、すべての統計を消去できます。この操作は、すべてのオブジェクトに適用することも、指定したオブジェクトにのみ適用することもできます。処理の直前にリセットしておく、その処理による影響を測定することができます。

vxstat を実行すると次のような結果が表示されます。

TYP	NAME	OPERATIONS		BLOCKS		AVG TIME(ms)	
		READ	WRITE	READ	WRITE	READ	WRITE
vol	blop	0	0	0	0	0.0	0.0
vol	foobarvol	0	0	0	0	0.0	0.0
vol	rootvol	73017	181735	718528	1114227	26.8	27.9
vol	swapvol	13197	20252	105569	162009	25.8	397.0
vol	testvol	0	0	0	0	0.0	0.0

RAID-5 構成では、これ以外のボリューム統計も利用できます。詳細については、vxstat (1M) マニュアル ページを参照してください。

入出力のトレース (vxtrace)

vxtrace コマンドは、ボリューム上の処理をトレースします。vxtrace は、カーネル入出力エラーまたは入出力トレース レコードを標準出力に印刷するか、またはバイナリ形式のファイルに書き込みます。トレースの対象として、特定のカーネル入出力オブジェクト タイプか、指定したオブジェクトまたはデバイスを指定できます。詳細については、vxtrace (1M) マニュアル ページを参照してください。



パフォーマンス データの使用

収集したパフォーマンス データは、システム リソースを効率よく使用するために最適なシステム構成を設定する材料として活用できます。以下に、パフォーマンス データの使用方法を簡単に説明します。

入出力統計の使用

入出力統計を調べると、再設定の必要性が認められる場合があります。主要な統計には、ボリューム入出力活動とディスク入出力活動の2つがあります。

統計を取得する前に、既存の統計をすべて消去（リセット）してください。コマンド `vxstat -r` を使用すると、すべての統計を消去できます。統計を消去すると、ボリュームの作成時に生じるボリュームまたはディスク間の数値の差、およびブート時からの統計値（通常重要度が低い）が削除されます。

統計の消去後は、システムを通常通りに稼働させます。特定のアプリケーションまたは作業負荷による影響を測るには、そのアプリケーションを実行するかまたはシステムにその作業負荷をかける必要があります。複数の目的で使用しているシステムを監視する場合は、特定のアプリケーションを通常以上に使用しないようにしてください。複数のユーザで使用する時分割システムを監視する場合は、一日の通常利用時間中の統計値を累積するようにしてください。

ボリューム統計を表示するには、引数を指定せずに `vxstat` コマンドを実行します。次のような統計結果が表示されます。

TYP	NAME	OPERATIONS		BLOCKS		AVG TIME(ms)	
		READ	WRITE	READ	WRITE	READ	WRITE
vol	archive	865	807	5722	3809	32.5	24.0
vol	home	2980	5287	6504	10550	37.7	221.1
vol	local	49477	49230	507892	204975	28.5	33.5
vol	rootvol	102906	342664	1085520	1962946	28.1	25.6
vol	src	79174	23603	425472	139302	22.4	30.9
vol	swapvol	22751	32364	182001	258905	25.3	323.2

この出力を参照すると、処理回数が異常に多いボリュームや、読み取りまたは書き込みにかなり時間がかかっているボリュームを特定できます。

ディスク統計を表示するには、`vxstat -d` コマンドを実行します。次のような統計結果が表示されます。

TYP	NAME	OPERATIONS		BLOCKS		AVG TIME(ms)	
		READ	WRITE	READ	WRITE	READ	WRITE
dm	disk01	40473	174045	455898	951379	29.5	35.4
dm	disk02	32668	16873	470337	351351	35.2	102.9
dm	disk03	55249	60043	780779	731979	35.3	61.2
dm	disk04	11909	13745	114508	128605	25.0	30.7

ここでは、あるディスク上の複数のボリュームを別のディスク上に移動する状況を想定します。ボリューム `archive` を別のディスク上に移動するには、まず次のコマンドを実行して、ディスクが現在配置されているディスクを特定します。

```
# vxprint -tvh archive
```

次のような結果が表示されます。

```
V NAME USETYPE KSTATE STATE LENGTH READPOL PREFPLEX
PL NAME VOLUME KSTATE STATE LENGTH LAYOUT NCOL/WDTH  MODE
SD NAME PLEX PLOFFS DISKOFFS LENGTH [COL/]OFF FLAGS
v archive fsgen ENABLED ACTIVE 204800 SELECT -
pl archive-01 archive ENABLED ACTIVE 204800 CONCAT - RW
sd disk03-03 archive-01  0 409600 204800 0 デバイス名
```

注 システムによっては、例とは異なるデバイス名を使用している場合があります。デバイス名の詳細については、「Volume Manager とは」を参照してください。

関連付けられているサブディスクの一覧に、ボリューム `archive` がディスク `disk03` 上にあることが示されています。このボリュームを `disk03` から移動するには、次のコマンドを実行します。

```
# vxassist move archive !disk03 移動先ディスク
```

移動先ディスクには、ボリュームの移動先のディスクを指定します。移動先ディスクは指定しなくても構いません。指定しない場合は、ボリュームを格納するための十分な空き領域を備えた使用可能ディスクがデフォルトで指定されます。

たとえば、ボリュームを `disk03` から `disk04` へ移動するには、次のように入力します。

```
# vxassist move archive !disk03 disk04
```

このコマンドを実行すると、ボリュームは再編成され、`disk03` 上からは完全に削除されます。

注 グラフィカル ユーザ インタフェースを使用すると、コマンド ラインを使用するよりも簡単にボリュームの一部をディスク間で移動できます。

アクセス頻度が非常に高いボリューム（ルート ボリュームを除く）が 2 つある場合は、それらを移動して別々のディスク上に配置してください。

1 つのボリュームが特にアクセス頻度が非常に高い場合（特に読み取りまたは書き込み時間が異常に長い場合）、ボリュームをストライピング（ボリュームを複数の領域に分割し、各領域を別のディスク上に配置）してください。オンラインで実行する場合は、ストライピングできるようにボリュームを変換するために、ボリュームの追加コピーを格納する十分な空き領域が必要になります。使用可能な空き領域が十分でない場合は、代わりにバックアップ コピーを作成できます。ストライピングのために変換するには、ボリュームの



ストライプ プレックスを作成してから、古いプレックスを削除します。たとえば、ボリューム archive をディスク disk02、disk03、および disk04 にわたってストライピングするには、次のコマンドを実行します。

```
# vxassist mirror archive layout=stripe disk02 disk03 disk04  
# vxplex -o rm dis archive-01
```

特にアクセス頻度が非常に高い状態のボリュームを再編成した後で、ディスク統計を確認してください。いくつかのボリュームが再編成されている場合、まず統計を消去してから、統計の累積値を適当な期間にわたって収集します。

ディスクのアクセス頻度が過度に高い場合（読み取りや書きこみ時間が異常に長い場合）は、いくつかのボリュームを再編成することができます。1つのディスク上にアクセス頻度が比較的高いボリュームが2つある場合、それらを近くに配置して、シーク タイムを短縮させます。アクセス頻度が比較的高いボリュームが1つのディスク上に多数ある場合は、それらのボリュームをアクセス頻度が比較的低いディスクに移動してください。

入出力監視データ（またはサブディスク統計）を参照して、ボリュームの特定領域に活動が極端に集中していないかどうかを判断します。活動が集中している領域を特定した後、そのボリュームのサブディスクを分割し、アクセス頻度が比較的低いディスクへそれらの領域を移動します。

注意 ボリュームをストライピングまたは複数のディスクにわたって分割すると、ディスク障害によってボリューム全体に障害が発生する可能性が高くなります。たとえば、5つのボリュームを同じ5つのディスクにわたってストライピングしている場合、いずれか1つのディスクに障害が発生すると、5つのボリュームすべてをバックアップからリストアする必要があります。各ボリュームが別々のディスク上にあれば、1つのボリュームのみをリストアするだけで済みます。ミラーリングまたは RAID-5 を使用すると、1つのディスクの障害が複数のボリュームの障害を引き起こす可能性を低減できます。

一般にファイルシステムとデータベースでは、割り当てられた領域の位置が時間の経過に伴って変化するため、このようにボリューム内の位置を指定した情報はあまり有用ではありません。データベースの場合、アクセス頻度が特に高いインデックスまたはテーブルが使用している領域を特定することが可能です。特定できた領域は、アクセス頻度が高くないディスクへ移動することができます。

読み取りと書き込みの比率を調べると、ミラーリングによりパフォーマンスが向上するボリュームを特定することができます。書き込みよりも読み取りの比率の方が高い場合は、ミラーリングによりパフォーマンスと信頼性が向上します。ミラーリングによりパフォーマンスの向上が可能な読み取りと書き込み比率は、ディスク、ディスク コントローラ、複数のコントローラを使用できるかどうか、およびシステム バスの処理速度により大きく左右されます。アクセス頻度が特に高いボリュームで、書き込み比率よりも読み取り比率の方が高いと、ミラーリングによりそのボリュームのパフォーマンスを大幅に向上できる可能性が高くなります。

入出力トレース データの使用

入出力統計から基本的なパフォーマンス分析に必要なデータが得られる一方、入出力トレース データはより詳細な分析に役立ちます。入出力トレース データを使用すると、焦点を絞った特定の作業負荷のイベント トレース データを取得できます。これにより、ホットスポットの場所やサイズ、その原因であるアプリケーションを確実に特定できるようになります。

入出力監視データを使用すると、ディスク上の実際の作業負荷をシミュレートし、その結果を監視できます。システム管理者は、これらの統計を使用してシステムの限界を予測したり、リソースの追加を計画することができます。

Volume Manager のチューニング

ここでは、Volume Manager で使用するリソースを制御する技法を説明します。使用可能なシステム リソースの種類によっては、最高のパフォーマンスを実現するために、いくつかのチューニング可能パラメータの値を調整する必要があります。

一般的なチューニングのガイドライン

Volume Manager では、小型システムから大型サーバまで幅広い構成に対応したチューニングを行えます。高価なリソース（メモリなど）を使用した大型システムのパフォーマンスを向上させる場合、通常 Volume Manager はサポートされている最小の構成で稼働するようにチューニングされます。チューニングを行うと、全体的なシステムのパフォーマンスに悪影響を与えたり、Volume Manager を使用できない状態になる可能性があるため、注意して行う必要があります。

Volume Manager をチューニングするには、さまざまな技法があります。システムによっては、グローバルなチューニング可能ファイル `/etc/system` を使用してパラメータをチューニングすることができます。Volume Manager へのコマンド ライン インタフェースを使用しないと変更できないパラメータもあります。

チューニング可能パラメータ

いくつかのシステムでは、`id tune` コマンドを使用してチューニング可能パラメータを変更します。詳細については、`id tune (1M)` マニュアル ページを参照してください。

それ以外のシステムでは、`/etc/system` ファイルに行を追加してチューニング可能パラメータを変更します。その後システムのリブートが必要です。変更されたチューニング可能パラメータは、リブート後に有効となります。

たとえば、チューニング可能パラメータ `vol_tuneme` のデフォルト値を 5000 に変更するには、`/etc/system` ファイルの適切なセクションに次の行を挿入します。

```
set vxio:vol_tuneme=5000
```

多くの場合、チューニング可能パラメータは `volinfo` 構造体に含まれます。詳細については、`vxio (7)` マニュアル ページを参照してください。



以下に、具体的なチューニング可能パラメータについて説明します。

vol_maxvol

このパラメータは、システムで作成できるボリュームの最大数を制御します。設定可能な値の範囲は、1 からシステムで表示できる最大のマイナー番号までです。

デフォルト値は、システム上の最大マイナー番号の半分です。

voliomem_maxpool_sz

このパラメータは、1 回の入出力においてシステム内のすべてのメモリが使用されることを防止します。

VxVM は 2 つの `voliomem_maxpool_sz` プールを割り当てます。1 つは RAID-5 用で、もう 1 つはミラー ボリューム用です。

RAID-5 ボリュームに対して `volio_maxpoll_sz/10` を超える大きさの書き込みがあると、この書き込みは分割され、`volio_maxpoll_sz/10` のサイズで処理されます。

ミラー ボリュームに対して `volio_maxpoll_sz/2` を超える大きさの書き込みがあると、この書き込みは分割され、`volio_maxpoll_sz/2` のサイズで処理されます。

vol_subdisk_num

このパラメータは、単一のプレックスに含められるサブディスクの最大数を制御します。この値には理論上の制限はありませんが、実用的にはデフォルト値 4096 が上限です。必要に応じて、このデフォルト値を変更できます。

vol_maxioctl

このパラメータは、`ioctl` の呼び出しにより Volume Manager に渡すことのできるデータの最大サイズを制御します。この制限値を上げると、大規模な処理を実行できます。この値を下げることは一般にお勧めできません。これは、いくつかのユーティリティが特定サイズの処理の実行に依存しており、指定サイズ以上の `ioctl` コマンドが発行されると異常終了する恐れがあるためです。

デフォルト値は 32678 バイト (32K) です。

vol_maxspecialio

このパラメータは、`ioctl` の呼び出しによって発行できる入出力の最大サイズを制御します。`ioctl` 要求そのものは小規模でも、大規模な入出力を実行を要求している場合があります。このチューニング可能パラメータは、こうした入出力のサイズを制限します。必要な場合は、この制限を超える要求を異常終了させるか、またはこのような入出力を分散して同時に実行することができます。

デフォルト値は 512 セクタ (256K) です。

vol_maxio

このパラメータは、要求を分割せずに実行できる論理入出力処理の最大サイズを制御します。この値より大きな物理入出力要求は、分割されて同時に実行されます。物理入出力はディスク デバイスの能力に基づいて分割され、論理要求を制限するこの値の変更には左右されません。

デフォルト値は 512 セクタ (256K) です。

この制限値を増加すると入出力サイズが増加し、処理で使用されるメモリやカーネルのマッピング領域が既存領域を超過するため、デッドロックが発生することがあります。vol_maxio の最大値は、物理メモリとカーネル仮想メモリの小さい方に対して 20% の値となります。この値を超えると、デッドロックが発生する恐れがあるため、お勧めできません。

vol_maxio を超える大きさのストライプがある場合は、フルストライプの入出力が分割されるため、全ストライプの読み取り / 書き込みは実行されません。その結果、それ以降の入出力や大規模入出力に関するボリューム入出力のスループットは低下します。

このパラメータは、最小値として、最も大きいストライプのサイズに設定する必要があります。このガイドラインは、Raid0 ストライピングと Raid5 ストライピングの両方に適用されます。

vol_maxkiocount

このパラメータは、Volume Manager でパラレルに処理できる入出力の最大数を制御します。ボリューム デバイスを使用する入出力がこの値を超えて発生すると、現在の処理数がこの値以下に落ちるまでの間、その入出力はキューに入れられます。

デフォルト値は 2048 です。

多くのプロセス スレッドでは一度に発行できる入出力が 1 つのみなので、カーネル内で処理中の入出力が上限に達するには、2K の入出力要求をパラレルで処理する必要がある計算になります。大規模なシステムの場合を除いて、この制限値を上げても、大きな利点はありません。

vol_default_iodelay

ユーティリティが入出力の発行速度を落とすように求められているが、特に遅延時間が指定されていない場合、このパラメータは入出力の発行ごとにユーティリティが一時停止するクロック数を制御します。ミラーの再同期または RAID-5 カラムの再構築などの処理を実行するユーティリティでこの値を使用します。

デフォルト値は 50 クロック サイクルです。

この値を上げると、リカバリ処理が低速化し、結果としてリカバリ処理がシステムに与える影響は小さくなります。



voldrl_min_regionsz

Volume Manager では、ダーティリージョン ロギングを使用して、ボリュームを論理的に一連のリージョンに分割します。このチューニング可能パラメータ `voldrl_min_regionsz` は、DRL ボリューム リージョンの最小セクタ数を指定します。

Volume Manager カーネルでは、このパラメータのデフォルト値は 1024 セクタに設定されています。

リージョン サイズを大きくすると、リージョンのキャッシュ ヒット率が向上する傾向にあります。これにより書き込みパフォーマンスは向上しますが、リカバリに要する時間は長くなります。

voldrl_max_dirty

Oracle のログの再適用に使用されるボリュームなどでは、順次書き込みが行われるため、このような時間のかかる DRL ビットのクリーニングは効果がありません。これらのボリュームには、シーケンシャル DRL を使用して、ダーティビット数を制限し、リカバリ時間を短縮することができます。シーケンシャル DRL で処理できるダーティビット数を制限するには、チューニング可能パラメータ `voldrl_max_dirty` を使用します。順次書き込みのボリューム上でシーケンシャル DRL を使用すると、入出力のスループットに重大な影響を与える可能性があります。

voldrl_max_drtregs

このパラメータには、任意の時点でシステムに存在できるダーティリージョンの最大数を指定します。これはグローバルな値で、システム内のアクティブなボリュームの数に関係なく、システム全体に適用されます。

デフォルト値は 2048 です。

このチューニング可能パラメータ `voldrl_max_drtregs` を使用して、障害発生後に最悪な状態のシステムをリカバリする時間をチューニングできます。値を大きくすると、システムパフォーマンスは向上しますが、リカバリ時間は長くなります。

vol_maxparallelio

このパラメータは、`vxconfigd (IM)` デーモンが `VOL_VOLDIO_WRITE ioctl` 呼び出しごとに1回の `VOL_VOLDIO_READ` でカーネルから要求できる入出力処理の数を制御します。

デフォルト値は 256 です。この値を変更することはお勧めできません。

vol_mvr_maxround

このパラメータは、ミラーからの読み取り時のラウンドロビン ポリシーの精密度を制御します。読み取りは、そのオフセットが前回の読み取り時にこのパラメータで指定されたセクタ数内にある場合、前回の読み取りと同じミラーから行われます。

デフォルト値は 512 セクタ (256K) です。

この値を上げると、読み取り先を代替ミラーに切り替える頻度が少なくなります。これは、実行中の入出力が主として順次的で、入出力間のシークタイムが少なく短い場合は理想的です。ボリュームにランダムに分散された多数の読み取りを行う場合は、一般に代替ミラーから読み取る方が効率的です。

`voliot_iobuf_limit`

このパラメータは、カーネル内にトレースバッファを格納するために使用できるメモリサイズの制限を設定します。トレースバッファは、Volume Manager カーネルがトレースイベントレコードを格納するために使用します。トレースバッファをカーネル内に格納するように要求があると、バッファ用メモリがこのプールから引き出されます。

このサイズを大きくすると、システムメモリの使用と引き換えに、さらに多くのトレースを実行できます。この値を、システム上ですぐに用意できる値より大きいサイズに設定することはお勧めできません。

デフォルト値は 131072 バイト (128K) です。

`voliot_iobuf_max`

このパラメータは、単一のトレースバッファで使用できる最大バッファサイズを制御します。このサイズより大きなバッファの要求も、自動的にこのサイズに切り詰められます。監視インタフェースから最大バッファサイズが要求された場合も、(使用法の限界に応じて) 結果的にこのサイズのバッファが使用されます。

デフォルトサイズは 65536 バイト (64K) です。

このバッファを大きくすると、特に使用頻度の高いボリュームのパフォーマンスを下げることなく、大規模なトレースを実行することができます。この値が、チューニング可能パラメータ `voliot_iobuf_limit` の値を超えないように注意する必要があります。

`voliot_iobuf_default`

このパラメータは、トレース `ioct1` の一部として任意のカーネルバッファサイズがほかに指定されていない場合、作成するトレースバッファのデフォルトサイズを制御します。

デフォルトサイズは 8192 バイト (8K) です。

このバッファサイズが小さすぎるためにトレースデータがたびたび失われる場合は、この値をより適当なサイズにチューニングできます。

`voliot_errbuf_default`

このパラメータは、エラートレースイベント用に保持されているバッファのデフォルトサイズを制御します。このバッファはドライバの読み込み時に割り当てられ、Volume Manager の実行中にサイズをチューニングすることはできません。

デフォルトサイズは 16384 バイト (16K) です。



このバッファを大きくすると、システム メモリと引き換えに、多くのエラー イベント用にストレージを確保できます。このバッファ サイズを小さくすると、トレース デバイスでエラーを検出できない状態になる可能性があります。エラーのトレースに基づいて何らかの応答アクションを実行するアプリケーションは、このバッファに依存します。

`voliot_max_open`

このパラメータは、同時に開くことのできるトレース チャネルの最大数を制御します。トレース チャネルは、トレース デバイスドライバへのクローン エントリ ポイントです。`vxtrace` コマンドをシステム上で実行するたびに、1 つのトレース チャネルが消費されます。

デフォルトのチャネル数は 32 です。各チャネルの割り当てには、使用中でない場合でも約 20 バイト必要です。

`vol_checkpoint_default`

システムに障害が発生しても全体的なリカバリを行う必要がなく、最後に到達したチェックポイントからリカバ리를続行できるように、リカバリまたは再同期化処理を実行するユーティリティが現在のオフセットをカーネルに読みこむ間隔を制御します。

デフォルトのチェックポイントサイズは 20480 セクタ (10M) です。

このサイズを大きくすると、リカバリ処理の間にチェックポイントを取るための作業によるオーバーヘッドが低減されますが、リカバリ中にシステム障害が発生すると、その後のリカバりに時間がかかります。

`volraid_rsrtransmax`

この RAID-5 チューニング可能パラメータは、パラレル処理できる一時的な再構築処理の最大数を制御します。一時的な再構築操作は、縮退モードでない RAID-5 ボリューム上で行われるため予測できません。同時に実行できるこれらの処理数を制限することにより、同時に多くの再構築処理が実行されてシステムが飽和状態になる可能性と、メモリ不足状態になる危険性を低減できます。

パラレル処理できる一時的な再構築のデフォルト数は 1 です。

このサイズを大きくすると、初めて障害が発生した場合および障害オブジェクトの切り離しが行われる前に、システム上の初期パフォーマンスを向上できますが、メモリ不足状態に陥る可能性があります。

`voliomem_chunk_size`

この値は、Volume Manager に対するシステム メモリの割り当ておよび解放の単位に使用します。この値を大きくすると、Volume Manager に割り当てられるメモリの単位量が多くなり、メモリ割り当てにかかるオーバーヘッドが (多少) 低減されます。

デフォルト サイズは 64K です。



大型システムのチューニング

小型システム（100 ドライブ未満）では、Volume Manager をチューニングする必要がなく、すべての設定パラメータに適切なデフォルトが適用されます。しかし、大型システムの場合、容量およびパフォーマンスの両方の理由から、構成によっては設定パラメータのチューニングをさらに制御する必要があります。

一般に、Volume Manager を大型システム上で設定する場合、決定すべき重大事項はわずかに過ぎません。まず、ディスクグループのサイズ、各ディスクグループで保持する設定コピー数を決定します。そして、ディスクグループ内のすべてのディスクの専有領域サイズを選択します。

大きなディスクグループには、vxassist (1M) コマンドで選択される空き領域プールを大きく確保でき、大型のレイを作成できるという利点があります。しかし、小さなディスクグループの場合、それほど大きな設定データベースは必要なく、専有領域も小さくて済みます。超大型のディスクグループでは、結果的にディスクグループ内の専有領域を消費し尽くしてしまって、そのディスクグループにはそれ以上の設定オブジェクトを追加できなくなることが起こりえます。この場合、構成を複数のディスクグループに分割するか、専有領域を大きくする必要があります。これには、ディスクグループ内の各ディスクを再初期化する必要があります（すべてオブジェクトの構成を再設定し、バックアップからリストアする場合があります）。

一般に、ディスクアレイサブシステムを使用している場合は、各アレイに単一のディスクグループを作成し、ディスクグループをシステム間で1つの単位として物理的に移動できるようにしておくことをお勧めします。

ディスクグループの設定コピー数

ディスクグループの設定コピーの数は、冗長性とパフォーマンスを秤にかけて選択します。一般に、ディスクグループ内に存在する設定コピーの数が少ないと、グループへの初回アクセス、vxconfigd (1M) の初回起動の処理速度、ディスクグループ上でのトランザクションの実行などが速くなります。

注意 データベースコピーの冗長性を低くすると、設定データベースが損われる危険が高くなります。データベースが失われると、データベース内のすべてのオブジェクトおよびディスクグループに含まれているすべてのデータが失われます。

ディスクグループ内の設定コピーを処理するデフォルトのポリシーは、設定コピーをディスクグループで認識されている各コントローラか、同じターゲット上で複数のアドレスで呼び出せるディスクを含む各ターゲットに割り当てることです。この方法は、冗長性の観点からは十分ですが、環境によっては設定コピーの数が膨大になる可能性があります。

この場合、設定コピーの数を最小値の 4 に制限することをお勧めします。コピーの位置は、最大コントローラまたはターゲット領域に従って、前述のように選択されます。



ディスク グループのコピー数を設定する技法では、新しいグループの設定に `vx dg init` コマンドを使用します (詳細については、`vx dg (1M)` マニュアル ページを参照)。また、`vx edit set` コマンドを使用して、既存のグループのコピーを変更することもできます (詳細については、`vx edit (1M)` マニュアル ページを参照)。たとえば、ディスク グループ `foodg` が 5 つのコピーを含むように設定するには、次のコマンドを実行します。

```
# vxedit set nconfig=5 foodg
```

用語集

RAID

RAID(Redundant Array of Independent Disks) は、ディスクのアレイ(集合) であり、全体の記憶容量の一部を、ディスク アレイに格納されたデータに関する情報の複製を格納するために使用します。これにより、ディスクに障害が発生した場合に、データを復元できます。

VM ディスク

Volume Manager の管理下に置かれ、ディスク グループに割り当てられているディスク。Volume Manager ディスクまたは単にディスクともいいます。グラフィカル ユーザ インタフェースでは、VM ディスクは D というラベルの付いた円柱形(シリンダ) のアイコンで表示されます。

volboot ファイル

ルート 設定のコピーの場所を特定するために使用される小さなファイル。このファイルには、標準の場所にある設定コピーが含まれているディスクの一覧が収められています。設定コピーの場所へのダイレクト ポインタが含まれる場合もあります。volboot が保存される場所は、システムによって異なります。

vxconfigd

Volume Manager 設定デーモン。Volume Manager の構成の設定を変更する役割を担っています。このデーモンが稼働していないと、Volume Manager の処理を行うことができません。

空きサブディスク

プレックスに関連付けられておらず、putil[0] フィールド が空であるサブディスク。

空き領域

Volume Manager の管理下にあるディスク 領域で、サブディスクに割り当てられていないもの、または Volume Manager のほかのオブジェクト 用の予備とされているもの。

アクティブ/アクティブ型のディスク アレイ

複数のパスを持つディスク アレイの種類の一つ。このディスク アレイでは、ディスクへのすべてのパスを同時に使用してディスクにアクセスできます。パフォーマンスが低下することはありません。



アクティブ/パッシブ型のディスク アレイ

複数のパスを持つディスク アレイの種類の一つ。このディスク アレイでは、ディスクのプライマリ パスが指定され、アクセス時には常時そのパスが使用されます。指定されているアクティブなパス以外のパスを使用すると、ディスク アレイによってはパフォーマンスが著しく低下する場合があります。「パス」、「プライマリ パス」、「セカンダリ パス」を参照してください。

アトミックな処理

問題なく完了するか、失敗した場合は何も変更されずに開始前の状態が保たれる処理。処理が正常に完了した場合、処理すべての結果が直ちに有効となり、処理途中の状態は確認できません。処理の一部が正常に終了しなかった場合、処理はアボートされ、開始前の状態に戻されます。

オブジェクト

Volume Manager が内部的に定義し認識するエンティティ。VxVM オブジェクトには、ボリューム、プレックス、サブディスク、およびディスク グループがあります。ディスク オブジェクトには、ディスクの物理面を扱うオブジェクトと、論理面を扱うオブジェクトの2種類があります。

開始ノード

システム管理者が、Volume Manager オブジェクトへ変更を要求するユーティリティを実行しているノード。このノードでボリュームの再設定が開始されます。

カプセル化

指定されたディスク上の既存のパーティションをボリュームに変換する処理。パーティションにファイルシステムが含まれている場合、`/etc/vfstab` エントリが変更され、そのファイルシステムはボリューム上にマウントされます。カプセル化を適用できないシステムもあります。

カラム

ストライプ プレックス内の1つまたは複数のサブディスク。ストライピングするには、プレックス内のカラムにデータを交互に均等に割り当てます。

関連付け

Volume Manager オブジェクト間に関係を確立する処理。たとえば、あるサブディスクがプレックスの中に開始点を持つように作成および定義されている場合、そのサブディスクはプレックスに関連付けられているといえます。

関連付けが解除されたサブディスク

プレックスとの関連付けが解除されたサブディスク。

関連付けが解除されたプレックス

ボリュームとの関連付けが解除されたプレックス。



関連付けの解除

Volume Manager オブジェクト 間に確立されている 関係を解除する処理。たとえば、サブディスクとプレックスの関連付けを解除すると、サブディスクがプレックスから削除され、空き領域プールに追加されます。

関連付けられたサブディスク

プレックスに関連付けられているサブディスク。

関連付けられたプレックス

ボリュームに関連付けられているプレックス。

共有 VM ディスク

共有ディスク グループに属する VM ディスク。

共有ディスク グループ

複数のホスト によってディスクが共有されるディスク グループ。クラスタで共有するディスク グループともいいます。

共有ボリューム

共有ディスク グループに属していて、同時に複数のノード で開かれているボリューム。

共有領域

Volume Manager によって管理されている物理ディスクの領域。使用可能な空き領域があり、サブディスクの割り当てに使用されます。

切り離し

Volume Manager オブジェクト がほかのオブジェクトと関連付けられているが、使用できない状態。

空白プレックス

ボリュームより小さいプレックス、またはホール(サブディスクが関連付けられていない領域)があるプレックス。

クラスタ

一連のディスクを共有するホストのセット。

クラスタで共有するディスク グループ

複数のホスト によって共有されるディスクのグループ。共有ディスク グループともいいます。

クラスタ マネージャ

クラスタ内の各ノード 上で稼働する、外部から提供されるデーモン。各ノード 上のクラスタ マネージャは相互に通信して、クラスタ メンバーシップの変更を VxVM に通知します。



コンカチネーション

サブディスクを順々に連続して連結したレイアウト。

サブディスク

連続するディスク ブロックが集まって論理ディスク セグメント を形成したもの。サブディスクはプレックスと 関連付けられてボリュームを形成します。

ストライピング

ストライプを使用して複数の物理ディスクにデータを分散させるレイアウト 技法。データは、各プレックスのサブディスク内のストライプに交互に均等に割り 当てられます。

ストライプ

複数のカラムの同じ位置を占めているストライプ ユニットの集まり。

ストライプ サイズ

1つのストライプを構成しているすべてのカラムにわたる、ストライプ ユニット サイズの合計。

ストライプ ユニット

各ストライプ プレックスのサブディスク (カラム内) に交互に割り 当てられた等サイズの領域。ディスク アレイでは、次のディスクによる割り 当て前に、各ディスクに存在している論理的に連続するブロックの集まりです。ストライプ ユニットは、ストライプ要素ともいいます。

ストライプ ユニット サイズ

各ストライプ ユニットのサイズ。デフォルト のストライプ ユニット サイズは 32 セクタ (16K) です。以前はストライプ幅と呼ばれていました。

スパンニング

大き過ぎて1つのディスクに収まりきらないボリューム (およびファイル システムまたはデータベース) を複数の物理ディスクに分散させるレイアウト 技法。

スライス

論理ディスク デバイスの標準的な分割単位。パーティションとスライスは同じ意味で使用されることがあります。

スレーブ ノード

マスタ ノードに指定されていないノード。

スワップ ボリューム

スワップ領域として使用するために設定された Volume Manager ボリューム。



スワップ領域

システムのページング処理によりスワップされたメモリー ページのコピーを格納するディスク領域。

セカンダリ パス

アクティブ / パッシブ型のディスク アレイでは、プライマリ パス以外のディスクへのパスはセカンダリ パスと呼ばれます。プライマリ パスに障害が発生し、ディスクの所有権が二次ディスクの1つに移転されない限り、ディスクへのアクセスはプライマリ パスに限定されます。「パス」および「セカンダリ パス」を参照してください。

セクタ

サイズの単位。システムによって異なる場合があります。通常、セクタは512バイトです。

接続

Volume Manager オブジェクト が別のオブジェクトと関連付けられていて、使用できる状態。

設定データベース

Volume Manager の既存のオブジェクトの詳細な情報（ディスクやボリュームの属性など）を含むレコードの集まり。設定データベースの単一のコピーを設定コピーと呼びます。

専有ディスク グループ

単一の特定のホストによってのみアクセスされるディスク グループ。

専有領域

Volume Manager に固有の構成情報を保存するための物理ディスクの領域。専有領域には、ディスク ヘッダー、テーブル、および設定データベースが含まれます。テーブルは、ディスクの内容をマップしています。ディスク ヘッダーにはディスク ID が記録されています。専有領域のすべてのデータは、信頼性を高めるために複製されます。

ダーティ リージョン ロギング

プレックスの変更を監視し、ログを作成する Volume Manager の機能。変更されたリージョンのビットマップが、ログ サブディスクと呼ばれる関連付けられたサブディスクに格納されます。

ディスク

インデックスが付けられ高速アクセスができる、読み取り / 書き込みデータ ブロックの集まり。各ディスクには、全体的に一意的 ID があります。

ディスク ID

各ディスクに与えられた汎用一意識別子。ディスクが移動された場合でも、そのディスクを識別するために使用できます。



ディスク アクセス名

物理ディスクへのアクセスに使用する名前。c0t0d0 など。c#t#d#s# の構文は、コントローラ、ターゲット アドレス、ディスク、およびパーティションを示します。デバイス名とも言います。

ディスク アクセス レコード

特定のディスクへのアクセス パスを指定する設定レコード。各ディスク アクセス レコードには、名前とタイプのほかに、タイプ固有の情報を含めることもできます。この情報は、ディスク アクセス レコードによって定義されているディスクへのアクセス方法および操作方法を判断するために、Volume Manager によって使用されます。

ディスク アレイ

論理的に1つのオブジェクトとして配列されているディスクの集まり。ディスク アレイを使用すると、冗長性を高め、パフォーマンスを向上させることができます。

ディスク アレイ シリアル番号

ディスク アレイのシリアル番号。通常、ディスク アレイのキャビネット上に印字されています。または、ディスク アレイ上のディスクに対してベンダ指定の SCSI コマンドを発行すると表示されます。この番号は、ディスク アレイを識別するために、DMP サブシステムにより使用されます。

ディスク グループ

同じ設定を共有するディスクの集まり。ディスク グループの設定は、Volume Manager オブジェクトに関する詳細な情報(ディスクやボリュームの属性など)とオブジェクト間の関係を記録したレコードの集まりです。各ディスク グループには管理者が割り当てた名前と、内部的に定義された一意の ID があります。ルート ディスク グループ(rootdg) は、特殊な専有ディスク グループであり、システムに必ず存在します。

ディスク グループ ID

ディスク グループを識別するための一意の識別子。

ディスク コントローラ

ホストまたはオペレーティング システム上でディスクの親ノードとなっているディスク アレイに接続されているコントローラ(HBA) を、Volume Manager のマルチパス サブシステムではディスク コントローラと呼びます。

たとえば、あるディスクが次のデバイス名で表されているとします。

```
/devices/sbus@1f,0/QLGC,isp@2,10000/sd@8,0:c
```

この場合、ディスク sd@8,0:c のディスク コントローラは次のとおりです。

```
QLGC,isp@2,10000
```

このコントローラ(HBA) はホストに接続されています。



ディスク メディア名

ディスクに設定された論理名または管理名です。disk03 のように付けられます。ディスク名ともいいます。

ディスク メディア レコード

特定のディスクを一意的 ID で識別し、そのディスクの論理(管理)名を保持している設定レコード。

データ ストライプ

ストライプの中の使用可能データ部分。ストライプからパリティ領域を差し引いた部分と等しくなります。

デバイス名

物理ディスクへのアクセスに使用するデバイス名またはアドレス。c0t0d0 など。c##t##d##s## の構文は、コントローラ、ターゲット アドレス、ディスク、および領域(システムによってはパーティション)を示します。

トランザクション

個々にではなく、グループとして成功または失敗する設定変更の集まり。トランザクションは設定の一貫性を維持するために内部的に使用されます。

ノード

クラスタ内のホストのひとつ。

ノード切断

障害が発生したときに、実行中の処理を停止させずに、クラスタからノードを切り離すこと。

ノード追加

ノードをクラスタに追加し、共有ディスクにアクセスできるようにする処理。

ノードのクリーン シャットダウン

共有ボリュームへのアクセスがすべて終了するのを待って、整合性を崩すことなくクラスタからノードを切り離すこと。

パーシステント ステート ロギング

アクティブなミラーのみをリカバリに使用し、障害が発生しているミラーがリカバリ用を選択されないようにする、ログの方式カーネルログともいいます。

パーティション

物理ディスク デバイスの標準的な分割単位。オペレーティング システムとディスクドライブによって直接サポートされます。



パス

ディスクがホストに接続されている場合、そのディスクへのパスは、ホスト上の HBA (Host Bus Adapter)、SCSI またはファイバ ケーブルのコネクタ、ディスクまたはディスク アレイ上のコントローラで構成されます。これらの構成要素がディスクへのパスを形成しています。構成要素のいずれかに障害が発生すると、DMP はディスクへのすべての入出力を残りのパス (代替パス) を使用して処理しようとします。

パリティ

障害発生後にデータを復元するために使用される計算値。RAID-5 ボリュームへのデータの書き込み時にデータを排他的論理和 (XOR) 演算することによって、パリティを計算します。算出されたパリティはボリュームに書き込まれます。RAID-5 の一部に障害が発生した場合は、残りのデータとパリティを使って、該当する部分のデータが復元されます。

パリティ ストライプ ユニット

RAID-5 ボリューム中のパリティ情報が格納されている記憶領域。入出力障害またはディスク障害により RAID-5 ボリュームの領域の一部が失われた場合、パリティ ストライプ ユニットに含まれているデータが復元のために使用されます。

ファイル システム

1 つの構造として編成されたファイルの集まり。UNIX のファイル システムは、ディレクトリとファイルで構成された階層構造になっています。

ブート ディスク

ブートのために使用されるディスク。このディスクは Volume Manager の管理下に置くことができます。

物理ディスク

基本的な記憶デバイス。Volume Manager の管理下に置くことも、置かないことも可能です。

プライマリ パス

アクティブ / パッシブ型のディスク アレイでは、ディスクはディスク アレイ上の 1 つの特定のコントローラに結合されるか、またはあるコントローラによって所有されます。ディスクには、この特定のコントローラを介したパスを使用してアクセスできます。「パス」、「セカンダリ パス」を参照してください。

プレックス

ボリュームとその中のデータ (順番に並べられたサブディスクの集まり) のコピー。各プレックスは対応するもとのボリュームのコピーです。ミラーとプレックスは同義です。

ブロック

ディスクまたはディスク アレイへのデータ転送の最小単位。



分散ロック マネージャ

複数のシステム上で稼働し、分散リソースへのアクセスの一貫性を保証する働きをするロック マネージャ。

ホスト ID

Volume Manager がホストを識別するための文字列。ホストのホスト ID は、そのホストの volboot ファイルに保存され、ディスクとディスク グループの所有権を定義するために使用されます。

ホットリロケーション

ディスクの障害発生時に、ミラー ボリュームや RAID-5 ボリュームの冗長性およびそれらへのアクセスを自動的にリストアする機能。この機能により、障害の影響を受けたサブディスクが、スペア ディスクとして設定されているディスク、または同じディスク グループ内の別の空き領域にリロケートされます。

ボリューム

ファイルシステムやデータベースなどのアプリケーションで使用される、アクセス可能な一連のディスク ブロックを表す仮想ディスク。1つのボリュームは、1～32のプレックスの集まりです。

ボリューム イベント ログ

ボリューム イベント ログ デバイス (/dev/vx/event) は、ボリューム ドライバ イベントをユーティリティに報告する場合に使用するインタフェースです。

ボリューム設定デバイス

ボリューム設定デバイス (/dev/vx/config) は、ボリューム デバイス ドライバへのすべての設定変更を行う場合に使用するインタフェースです。

ボリューム デバイス ドライバ

アプリケーション レベルと物理デバイス ドライバ レベルの間に仮想ディスク ドライブを形成するドライバ。ボリューム デバイス ドライバには、キャラクタ デバイス ノードが /dev/vx/rdsk に示されていて、ブロック デバイス ノードが /dev/vx/dsk に示されている仮想ディスク デバイス ノードからアクセスします。

マスタ ノード

ソフトウェアによって「マスタ」ノードに指定されているノード。どのノードもマスタ ノードに設定することが可能です。マスタ ノードは Volume Manager のある種の処理を調整する働きをします。

マスタリング ノード

ディスクが接続されているノード。ディスク所有者とも呼びます。



マルチパス機能

システムに接続されているディスクにアクセスする物理パスが複数ある場合、そのディスクにはマルチパス機能があります。マルチパス機能を提供するソフトウェアはホスト上に常駐して、複数のパスの存在をユーザから隠します。

ミラー

ボリュームとその中のデータ（順番に並べられたサブディスクの集まり）のコピー。各ミラーは対応する元のボリュームのコピーです。ミラーとプレックスは同義です。

ミラーリング

ボリュームの内容を複数のプレックス上にコピーするレイアウト技法。各プレックスは、ボリュームに格納されているデータの複製ですが、プレックス自体のレイアウトは異なる場合もあります。

無効なパス

入出力に使用できないディスクへのパス。パスは、ハードウェアに障害が発生した場合や、コントローラ上で `vxdmpadm disable` コマンドを使用した場合に、無効になります。

有効なパス

入出力に使用できる、ディスクへのパス。

読み取り / ライトバック モード

読み取りごとに、読み取る領域のプレックスの整合性をリカバリするモード。プレックスの整合性は、1つのプレックスのブロックからデータを読み取り、そのデータをほかのすべての書き込み可能なプレックスに書き込むことでリカバリされます。

ルート機能

ルート ファイルシステムとスワップ デバイスを **Volume Manager** の管理下に組み込む機能。この機能によって、ボリュームをミラーリングしてデータを冗長化し、ディスク障害時のリカバリを図ることが可能になります。

ルート設定

ルート ディスク グループ用の設定データベース。ほかのディスク グループの情報が必ず含まれており、バックアップのためにのみ使用される特殊なデータベースです。また、システム上のすべてのディスク デバイスを定義するディスク レコードも含まれます。

ルート ディスク

ルート ファイルシステムが収められているディスク。このディスクは **Volume Manager** の管理下に置くことができます。

ルート ディスク グループ

システム上に必ずある特殊な専用ディスク グループ。ルート ディスク グループには `rootdg` という名前が付けられています。



ルート パーティション

ルート ファイル システムがあるディスク領域。

ルート ファイル システム

UNIX カーネルの起動シーケンスの一部としてマウントされる初期ファイル システム。

ルート ボリューム

ルート ファイル システムが含まれている **Volume Manager** ボリューム。システム設定でこのボリュームが指定されたときのみ存在します。

ログ サブディスク

ダーティ リージョン ログを格納するために使用するサブディスク。詳細については、「ダーティ リージョン ロギング」を参照してください。

ログ プレックス

RAID-5 ログを格納するためのプレックス。ログ プレックスは、ダーティ リージョン ロギング プレックスともいいます。





索引

- A
 - autoboot フラグ 220
- D
 - DMP
 - DMP デバイスからのブート 65
 - 動的マルチパス 64
 - パス フェイルオーバー機構 65
 - 負荷の平準化 65
 - DMP デバイスからのブート
 - DMP 65
 - DMP 表示 115
 - DRL 166
- F
 - format ユーティリティ 74
- N
 - nopriv 112、113
 - デバイス 114
- O
 - OFFLINE 177
- P
 - putil 176
- R
 - RAID- 14
 - RAID-1 18
 - RAID-5 254、255、257、258、260、261、262、263、264、265、266、268、279
 - ガイドライン 41
 - サブディスクの移動 264
 - スナップショット 186
 - リカバリ 259、267
 - RAID-5 ブレックス 254
 - RAID-5 ボリューム 267
 - 作成 161
 - RAID-5 ログ 166
 - resilvering 67
 - rootdg 5、72
 - 名前変更 110
- S
 - Storage Administrator 28
- T
 - tutil 176
- U
 - UNIX パーティション 121
- V
 - VM ディスク 4
 - 定義 4
 - vm ディスク オペレーション 88
 - Volume Manager 28、45
 - オペレーティングシステム 26
 - デーモン 33
 - ルート機能 61
 - レイアウト 26
 - Volume Manager Support Operations 71
 - Volume Manager グラフィカル ユーザ インタフェース 73
 - vxassist 42、73、102、164、186、187
 - growby 162
 - growto 162
 - shrinkby 162
 - shrinkto 162
 - バックアップ 186
 - ボリュームの作成 157
 - vxassist addlog 166
 - vxassist growby 163
 - vxassist growto 162
 - vxassist make 86、160
 - vxassist snapshot 170
 - vxassist snapstart 169
 - vxclust 208
 - vxconfigd 33、62、209
 - vxctl 215
 - vxdg 73、83、105、108、109、211
 - ディスク グループの移動 108
 - ディスク グループの削除 107
 - vxdg rmdisk 88
 - vxdisk 73、83、112、113、214
 - rm 83



特殊なカプセル化 112
vxdisk list 89
vxdisk rm 88
vxdiskadd 71、73、74、77、89、103
vxdiskadm 62、71、73、74、83、84、101、109
 起動 116
 ディスクの交換 101
vxedit 174、176、180、184、242
 サブディスクの削除 180
vxedit rename 85
vxedit set 86
vxinfo 101
vxiod 33、34
vxmake 164、172、173、180、181
 サブディスクの関連付け 181
 サブディスクの作成 180
 ミラーの関連付け 173
 ミラーの作成 172
vxmend 170、177、178
vxmirror 221
vxplex 164、165、173、174、175、177、178、179
 ミラーの移動 179
 ミラーの関連付けの解除 165、174
 ミラーのコピー 180
vxprint 99、171、175、176、181
 サブディスクの表示 181
 ミラーの一覧表示 176
vxreattach 237
vxrecover 100、215
 ディスクグループの移動 108
vxrelocd 51、93
 変更 93
vxsd 182、183、184、185
 サブディスクの移動 184
 サブディスクの関連付け 182
 サブディスクの結合 185
 サブディスクの分割 185
 ログ サブディスクの関連付け 183
VxSmartSync 67
vxstat 100、217、280、282
vxtrace 280、281
VxVM 28、45
vxvol 102、178

い

移動
 ディスクからのボリュームの移動 119、191
インポート
 ディスクグループ 108

お

オフライン
 ディスク 86、120、143
オンラインバックアップ 186
オンライン再レイアウト 46
 障害のリカバリ 50
 動作 47
 変更の種類 48
 変更の特性 50
 レイアウト変更およびボリュームの大きさ 50

か

ガイドライン
 RAID-5 41
 ダーティリージョン ロギング 39
 ミラーリング 38
 ミラーリング アンド ストライピング 40
書き込み
 全ストライプ 271
 復元 272
 読み取り - 変更 269
カプセル化 42、61、112、118、131、221
カラム、ストライピング 14

き

起動
 vxdiskadm 116
共有オブジェクト 194
共有領域 71

く

クラスタ
 共有オブジェクト 194
 ディスク 203
クラスタプロトコルバージョン 215
クラスタプロトコルの範囲 215
クラスタ環境 68、193
クラスタ機能 68、193、194
クラスタで共有するディスクグループ 194
クラスタの再設定 199
グラフィカル ユーザ インタフェース 28、73

こ

交換するディスク 233
高速ミラー再同期 (FMR) 56
コマンドライン ユーティリティ 73
コンカチネイテッド ボリューム
 作成 159
コンカチネーション 11
コントローラ 121

さ

再インストール 239、240



- サイズ
 - ボリューム
 - 拡大 162
 - 縮小 163
 - 再設定手順 239
 - 再同期化
 - Oracle データベース 67
 - ボリューム 53
 - 削除
 - DRL 166
 - ディスク 118、142
 - 物理ディスク 88
 - ボリューム 164
 - 作成
 - RAID-5 ボリューム 161
 - VM ディスク上のボリューム 160
 - コンカチネイテッド ボリューム 159
 - ストライプ ボリューム 160
 - スパン ボリューム 159
 - ディスク グループ 147
 - サブディスク
 - 移動 184
 - 関連付け 181、182
 - 関連付けの解除 183
 - 結合 185
 - 削除 180
 - 情報の変更 184
 - 表示 181
 - 分割 185
 - ログ 55、182、204
 - サブディスクの移動
 - RAID-5 264
 - vxsd 184
 - サブディスクの関連付け
 - vxmake 181
 - vxsd 182
 - サブディスクの結合
 - vxsd 185
 - サブディスクの削除
 - vxedit 180
 - サブディスクの作成 180
 - vxmake 180
 - サブディスクの操作 263
 - サブディスクの表示
 - vxprint 181
 - サブディスクの分割
 - vxsd 185
 - システム障害 257
 - 終了
 - vxdiskadm 153
 - 縮退モード 258
 - 障害 226、261
 - システム 257
 - ディスク 258
 - リカバリ手順 226
 - 障害ディスク 91
 - 検出 99
 - 状態
 - プレックス 249
 - ボリューム 252
 - 使用できないボリューム 265
 - 情報 184
- す
- ストライピング 14、276、278
 - ストライプ カラム 14
 - ストライプ プレックス 14
 - ストライプ ボリューム
 - 作成 160
 - ストライプ ユニット 14
 - ストレージ レイアウト
 - 変換 46
 - スナップショット 186、187
 - RAID-5 186
 - スパン ボリューム
 - 作成 159
 - スパンニング 11
 - スライス 121
 - スレーブ ノード 196
 - スワップ ボリューム制限 62
- せ
- 設定ガイドライン 275
 - 専有ディスク グループ 194
 - 専有領域 71
- た
- ダーティ リージョン ロギング 55、166
 - ガイドライン 39
 - クラスタ環境で 204
 - ログ サブディスク 182
- ち
- チェックポイント 261
 - チューニング
 - Volume Manager 285
 - チューニング可能パラメータ 285
- つ
- 追加
 - DRL ログ 166
 - RAID-5 ログ 166
 - 今後の使用に備えたディスク 134
 - ディスク 77、118
 - ディスクをディスク グループに追加 80、147



て

- ディスク 121
 - VM
- ボリュームの作成 160
 - VM ディスク 4
 - volatile volatile 114
 - 移動 90
 - カプセル化 42、61、112、118、131、221
 - 切り離された 99
 - クラスタ内の 203
 - 交換 101、102、233
 - 今後の使用に備えた追加 134
 - 再初期化 134
 - 再接続 237
 - 再追加 233
 - 削除 142
 - 障害 91、258
- ホットリロケーション 51、91
- リカバリ 219
 - 情報の表示 141
 - 初期化 74
 - 追加 74
 - ディスクからのボリュームの移動 191
 - ブート ディスク 61、87、221、232、233
 - 物理
- VxVM の管理下への配置 77
- オフライン 86
- 交換 119、145
- 削除 88、118
- 情報の表示 89
- 追加 77、118
- ディスク グループへの追加 80
- ディスクからのボリュームの移動 119
- 無効化 120
- 有効化 120
- 予約 86
 - ホットリロケーション スペア 94、137
 - ボリュームのミラーリング 189
 - 無効 143
 - 有効化 144
 - ルート ディスク 42、87、232
- ディスク アクセス名 121
- ディスク グループ 5、72
 - 移動 108、109
 - インポート 108、109、152
 - 削除 107
 - 作成 147
 - システム間での移動 108
 - 初期化 147
 - ディスクの追加 147
 - デフォルト 147

- デポート 108、109、151
- 名前変更 110
- 無効化 151
- 有効化 119
- ディスク グループ ユーティリティ 73
- ディスク グループの移動
 - vx dg 108
 - vxrecover 108
- ディスク グループのインポート 119、152
- ディスク グループの削除 107
 - vx dg 107
- ディスク グループの作成
 - vx dg 105
- ディスク グループの使用
 - vxassist 73
- ディスク グループ名の変更 110
- ディスク メディア名 4、70、121
- ディスク障害 258
 - リカバリ 219
- ディスク情報、表示 141
- ディスクの移動 90
- ディスクのカプセル化 118、131
- ディスクの交換 101、119、145
 - vx diskadm 101
- ディスクの再初期化 134
- ディスクの再接続 237
- ディスクの再追加 233
- ディスクの削除
 - vx dg 83
- ディスクの初期化 74
- ディスクの追加 74
 - フォーマット 74
- ディスク名 70、72
- データ
 - 冗長性 18
 - 保護 42
- データの配置 276
- デーモン 33
 - Volume Manager 33
 - vxrelocd 93
 - ホットリロケーション 51
- デバイスのパス 121
- デバイス名 70、121
- デポート
 - ディスク グループ 108、119、151

と

- 動的マルチパス
 - DMP 64
- 特殊デバイス
 - 使用 112
 - 特殊デバイスの使用 112
- 特殊なカプセル化
 - vx disk 112



特別な目的に使用するディスクの予約 86

な

名前

- ディスク メディア 4、121
- ディスクへのアクセス 3、121
- デバイス 121

に

入出力

- 統計 282
- 収集 280
 - トレース データ 281、285
- 入出力デーモン 34
- 入出力統計の使用 282
- 入出力のトレース 281
 - vxtrace 281

の

ノード 194

は

- パーティション 3
 - UNIX 121
- パス フェイルオーバー
 - DMP 65
- バックアップ 43、186
 - vxassist 186
 - ミラー 174
- パフォーマンス 279
 - ガイドライン 275
 - 監視 280
 - 管理 275
 - 最適化 275
 - 優先事項 280
- パフォーマンス データ 280
 - 収集 280
 - 使用 282
- パフォーマンス データの収集 280
- パフォーマンス データの使用 282
- パリティ 17、20、261
- パリティのリカバリ 260、261

ひ

表示

- ディスク情報 89、141
- ボリューム設定 171
- 標準ディスク デバイス 70

ふ

ブート

- 障害後 221
- ブート ディスク 61、221
 - 交換 233、236
 - 再追加 233
 - 障害 233

障害およびホットリロケーション 232

ミラー 87

ブート プロセス 220

負荷の平準化

DMP 65

復元読み取り 258

ブレックス 6、254

一覧表示 176

移動 179

切り離し 177

コピー 179

作成 172

情報の変更 176

ストライプ 14

接続 177、178

定義 6

表示 175

ボリューム 7

ミラー 7

ブレックスのカーネル状態 252

DETACHED 252

DISABLED 252

ENABLED 252

ブレックスの関連付け

vxmake 173

ブレックスの状態 249

ACTIVE 250

CLEAN 250

EMPTY 249

IOFAIL 251

OFFLINE 250

STALE 250

TEMP 251

TEMPRM 251

ブレックスの状態のサイクル 252

へ

変換

ストレージレイアウト 46

ほ

ホスト

複数 194

ホットリロケーション 51、91、120、137

vxrelocd の変更 93

スペアとして指定 52、137

スペアの削除 138

ブート ディスク 232

ボリューム 7

RAID-5

作成 161

起動 167

コンカチネイテッド、作成 159

サイズの拡大 162



- サイズの縮小 163
 - 削除 164
 - 作成
 - コンカチネイテッド 159
 - ストライプ
 - 作成 160
 - スパン
 - 作成 159
 - すべての既存ボリュームのミラーリング 165
 - 設定の表示 171
 - 定義 2、7
 - 停止 167
 - ディスクからの移動 191
 - ディスク上のミラーリング 189
 - バックアップ 169
 - ブレックス7
 - 補整 242
 - ミラーリング 164
 - 読み取りポリシー 168
 - リカバリ 170
 - レイアウト 254
 - ボリュームサイズの拡大 162
 - ボリュームサイズの縮小 163
 - ボリューム制限
 - ブート時 62
 - ボリューム属性の変更 268
 - ボリュームのカーネル状態 254
 - DETACHED 254
 - DISABLED 254
 - ENABLED 254
 - ボリュームの起動 264
 - ボリュームの強制起動 266
 - ボリュームの再設定 199
 - ボリュームの再同期化 53
 - ボリュームの作成
 - vxassist 157
 - ボリュームの状態 252
 - ACTIVE 252、253
 - CLEAN 252
 - EMPTY 252、253
 - SYNC 252、253
 - ボリュームのバックアップ 169
- ま**
- マイナー番号
 - 予約 111
 - マスタ ノード 196
- み**
- ミラー 6、7、119
 - オフライン 177
 - 関連付けの解除 165、174
 - 削除 165、174
 - 作成 172
 - バックアップに使用 174
 - 表示 175
 - ブート ディスク 87
 - リカバリ 100
 - ルート ディスク 87
 - ミラーの一覧表示
 - vxprint 176
 - ミラーの移動 179
 - vxplex 179
 - ミラーの関連付け
 - vxmake 173
 - ミラーの関連付けの解除
 - vxplex 165、174
 - ミラーのコピー
 - vxplex 180
 - ミラーの作成
 - vxmake 172
 - ミラーリング 18、277、278
 - ガイドライン 38
 - すべてのボリューム 165
 - ディスク上のボリューム 189
- む**
- 無効
 - ディスク 143
 - 無効化
 - ディスク 120
 - ディスク グループ 119、151
- め**
- メニュー インタフェース 116
- ゆ**
- 有効化
 - ディスク 120、144
 - ディスク グループ 119
 - ディスク グループへのアクセス 152
- よ**
- 読み取り
 - ポリシー 277
- り**
- リカバリ 219、267
 - RAID-5 ボリューム 259、267
 - 手順 226
 - ボリューム 170
 - ログ 261
- る**
- ルート ディスク 42、232
 - ミラー 87
 - ルート ボリューム
 - ブート 62
 - ルート ボリューム制限 62

-
- ルート機能 61
補整 242
- れ
レイアウト
左対称 22
- ろ
ロギング 23
- ログ 255、262
RAID-5 166
追加 166
ログ サブディスク 39、55、182、204
関連付け 182
ログ サブディスクの関連付け 182
vxsd 183
ログ プレックス 254



