



VERITAS Volume Manager™

Getting Started Guide

(日本語版)

リリース 3.0.1

Solaris

1999 年 11 月

P/N 100-001301

© 1998 - 1999 VERITAS® Software Corporation. All rights reserved.

商標

VERITAS、VxVM、VxVA、VxFS、および FirstWatch 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 Volume Manager、VERITAS File System、VERITAS NetBackup、VERITAS HSM、VERITAS Media Librarian、CVM、VERITAS Quick I/O、VxSmartSync、および VERITAS ロゴは、VERITAS Software Corporation の商標です。

本書に記載されているその他の製品は、各社の商標または登録商標です。

目次

まえがき	vii
1. Volume Manager とは	1
はじめに.....	1
データの格納方法.....	3
Volume Manager の概要.....	3
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	4
パーティション	4
ボリュームと仮想オブジェクト.....	5
Volume Manager ディスク.....	5
ディスク・グループ	6
サブディスク	6
プレックス	7
ボリューム	8
Volume Manager の仮想オブジェクト間の関係.....	10
仮想オブジェクトのデータ編成（ボリューム・レイアウト）.....	11
コンカチネーション	12
ストライピング（RAID-0）.....	14
RAID-5.....	18



ミラーリング (RAID-1).....	19
ミラーリング・プラス・ストライピング (RAID-1 + RAID-0).....	20
ストライピング・プラス・ミラーリング (RAID-0 + RAID-1).....	20
Volume Manager および RAID-5.....	20
従来の RAID-5 アレイ.....	21
Volume Manager RAID-5 アレイ	22
左対称レイアウト	23
ロギング	24
階層構成ボリューム.....	25
Volume Manager のユーザー・インタフェース.....	27
ユーザー・インタフェースの概要	27
Volume Manager の概念.....	28
Volume Manager を使用する目的.....	28
Volume Manager のオブジェクト.....	28
Volume Manager およびオペレーティング・システム.....	30
動的マルチパッシング (DMP)	30
Volume Manager のレイアウト.....	32
2. Volume Manager の機能	35
はじめに.....	35
オンライン再レイアウト	36
ストレージ・レイアウト	36
オンライン再レイアウトの動作	36
レイアウト変更の種類	37
レイアウト変更の特性	40
レイアウト変更およびボリュームの大きさ	40
サポートされていないレイアウト変更	40
ホットリロケーション	41
ホットリロケーションの仕組み	41



リロケート用スペースの選択方法	42
ボリュームの再同期化.....	43
ダーティ・リージョン・ロギング.....	44
Volume Manager のルート機能	45
ルート・ボリュームを使用するのブート	46
ブート時のボリュームの制限	46
動的マルチパッシング (DMP)	47
パス・フェイルオーバー機構	48
負荷の平準化	48
DMP デバイスからのブート	48
コントローラの有効化と無効化	49
DMP データベース情報の無効化	49
VxSmartSync Recovery Accelerator	50
データ・ボリュームの設定	50
REDO ログ・ボリュームの設定	51
Volume Manager Task Monitor	51
Volume Manager のクラスタ機能.....	52
3. Volume Manager の初期化と設定	53
はじめに.....	53
Volume Manager の初期化.....	54
vxinstall 実行後のリブート.....	55
Volume Manager のデーモン.....	55
設定デーモン vxconfigd.....	55
Volume Manager 設定デーモンの起動	55
ボリューム入出力デーモン vxiod.....	56
ボリューム入出力デーモンの起動	56
システムの設定.....	57
システム・セット設定手順の例	57



システム設定ガイドライン.....	58
ホットリロケーションのガイドライン	58
ストライピングのガイドライン	59
ミラーリングのガイドライン	60
ダーティ・リージョン・ロギング (DRL) のガイドライン	61
ミラーリング・アンド・ストライピングのガイドライン	62
ストライピング・アンド・ミラーリングのガイドライン	63
RAID-5 のガイドライン	63
システムの保護.....	64
索引.....	67

まえがき

『VERITAS Volume Manager™ Getting Started Guide(日本語版)』では、Volume Manager の構成要素、機能、およびインタフェースの概要を紹介します。また、ストレージ管理システムを計画する際に役立つ Volume Manager の設定方法も説明します。

対象読者

本書では、VERITAS Volume Manager を使用してシステムの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・経験を有していることを前提とします。

- システム管理の基本的な理解
- 使用するオペレーティング・システムに関する実務知識
- 使用するウィンドウ・システムの経験

本書の構成

本書では、Volume Manager の機能について概要を説明します。

本書の構成は以下のとおりです。

第1章「Volume Manager とは」では、VERITAS Volume Manager および関連する概念について簡単に説明します。

第2章「Volume Manager の機能」では、Volume Manager の機能の概要について説明します。

第3章「Volume Manager の初期化と設定」では、Volume Manager の初期化の方法(設定のガイドラインを含む)について説明します。



関連マニュアル

関連情報については、以下のマニュアルを参照してください。

- 『VERITAS Volume Manager Command Line Interface Administrator's Guide』には、Volume Manager コマンドを使用して管理作業を行う方法について記載されています。
- 『VERITAS Volume Manager Storage Administrator Administrator's Guide』には、管理者を対象に、グラフィカル・ユーザー・インタフェースを使用して Volume Manager のさまざまな処理を実行する方法について記載されています。
- 『VERITAS Volume Manager Administrator's Reference Guide』には、管理者を対象に、Volume Manager での回復手順に関する情報について記載されています。

表記法

本書で採用している表記規則を下の表に示します。

表記	用途	例
courier	コンピュータからの出力、ユーザーによる入力、コマンド、ファイル、およびディレクトリ	<code>\$You have mail.</code> <code>cat</code> コマンドを実行すると、ファイルが表示されます。 <code>\$ls -a</code>
斜体	新しい用語、マニュアルのタイトル、強調する用語、実際の名前または値が代入される変数	<code>\$cat</code> ファイル名 詳細については『ユーザーズ・ガイド』を参照してください。
太字	用語集に記載されている用語	

Volume Manager とは

1



はじめに

この『Getting Started Guide』では、VERITAS Volume Manager (VxVM) について、その機能、仕組み、ユーザー・インタフェースを介した操作方法、および概念などの概要を説明します。Volume Manager に関するさらに詳細な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」の「関連マニュアル」の項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、ユーザーに非常に不便をかけていました。分散型クライアント / サーバー環境においては、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
 - 物理ディスクおよびディスク名の命名
 - パーティション

- ボリュームと仮想オブジェクト
 - Volume Manager ディスク
 - ディスク・グループ
 - サブディスク
 - プレックス
 - ボリューム
 - Volume Manager の仮想オブジェクト間の関係
- 仮想オブジェクトのデータ編成（ボリューム・レイアウト）
 - コンカチネーション
 - ストライピング（RAID-0）
 - RAID-5
 - ミラーリング（RAID-1）
 - ミラーリング・プラス・ストライピング（RAID-1 + RAID-0）
 - ストライピング・プラス・ミラーリング（RAID-0 + RAID-1）
- Volume Manager および RAID-5
 - ロギング
- 階層構成ボリューム
- Volume Manager のユーザー・インタフェース
 - ユーザー・インタフェースの概要
- Volume Manager の概念
 - Volume Manager を使用する目的
 - Volume Manager のオブジェクト
 - Volume Manager およびオペレーティング・システム
 - Volume Manager のレイアウト

データの格納方法

データを物理ディスクに保存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。ディスク上のデータの編成方法は、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。Volume Manager でサポートされているストレージ・レイアウトの種類は本章で紹介しますが、詳細については『Volume Manager Administrator's Reference Guide』を参照してください。

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、物理ディスクとパーティションの2種類の物理オブジェクトを使用します。パーティションは、物理ディスク上（パーティションを使用するシステムにおいて）に作成されます。

- 仮想オブジェクト

Volume Manager は、ボリュームと呼ばれる仮想オブジェクトを作成します。各ボリュームに記録されそこから読み出されるデータは、1つ以上の物理ディスクにわたることができます。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。ボリュームを始めとする仮想構成要素を仮想オブジェクトと呼びます。

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（物理ディスクおよびパーティション）について説明します。

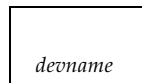
物理ディスクおよびディスク名の命名

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、パラメータ `c#t#d#` が含まれます。各パラメータの意味は次のとおりです。

- `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 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

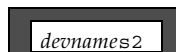


パーティション

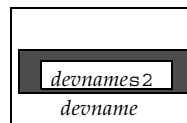
コンピュータ・システムによっては、物理ディスクを 1 つ以上のパーティションに分割することができます。パーティションに分割する場合、パーティション番号 (`s#`) をデバイス名 (*devname*) の末尾に追加します。図 2 に示すように、物理ディスク全体を 1 つのパーティションとすることもできます。

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 は仮想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、ストレージ管理の作業を行います。

ボリュームは仮想ディスク・デバイスで、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。仮想オブジェクトの詳細については、以降のセクションで説明します。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」セクションで説明します。

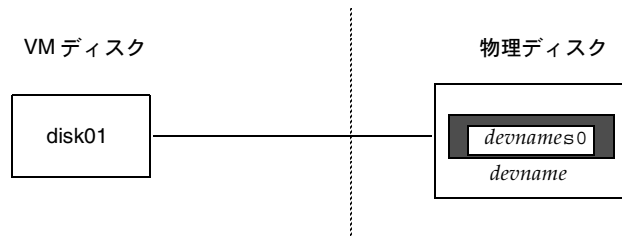
Volume Manager ディスク

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一つの 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。Volume Manager は、Volume Manager ディスク領域の連続する記憶領域を割り当てます。

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）を設定し、それを入力することも、Volume Manager にデフォルト名（通常は「disk##」の形式）を割り当てさせることもできます。図 3 に、ディスク *devnames0* 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、管理者は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 1 まとまりの単位として移動することができます。

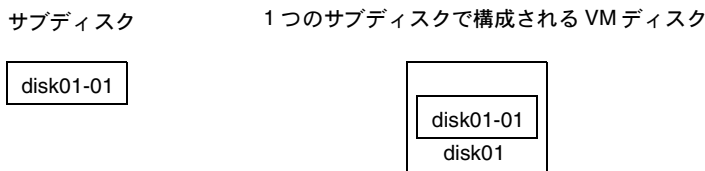
ボリュームはディスク・グループ内に作成されます。ボリュームを構成するディスクは、同じディスク・グループに属している必要があります。

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。ブロックとは、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。

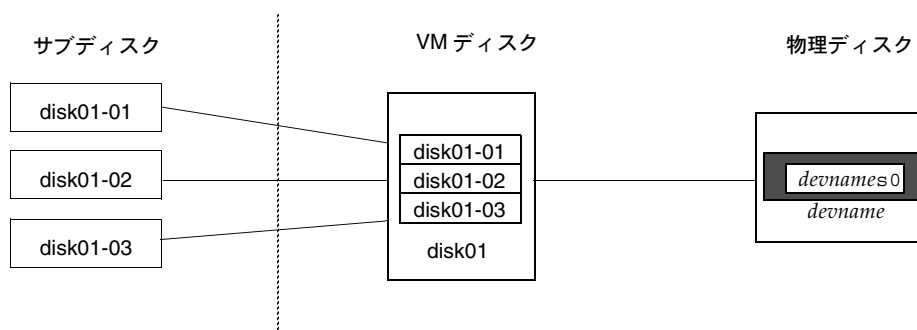
VM ディスクのデフォルト名は `disk##`（たとえば `disk01`）であり、サブディスクのデフォルト名は `disk##-##` です。図 4 に例を示します。この例では、VM ディスク名が `disk01`、その最初のサブディスク名が `disk01-01` と付けられています。

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ないし、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 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スペースです。空きスペース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、「階層構造ボリューム」を参照してください。

プレックス

Volume Manager では、サブディスクに基づいてプレックス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プレックスに編成するには下記の方法があります。

- コンカチネーション
- ストライピング (RAID-0)

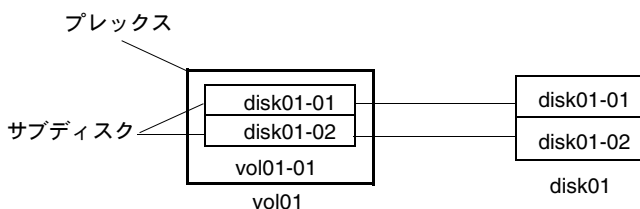
- パリティ付きストライピング (RAID-5)
- ミラーリング (RAID-1)

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 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、「仮想オブジェクトのデータ編成 (ボリューム・レイアウト)」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分が保持されます。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 のユーザー・インタフェースを使用して変更することができます。構成の変更は、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同一のコピーが含まれます。ミラーリングボリュームの詳細については、「ミラーリング (RAID-1)」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 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

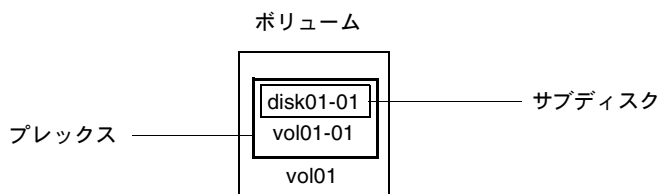


図 7 のボリューム vol01 には、次のような特徴があります。

- 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 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

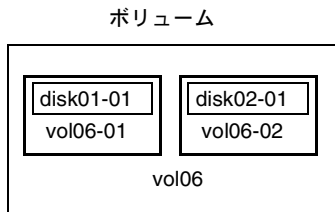


図 8 のボリューム vol06 には、次のような特徴があります。

- 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

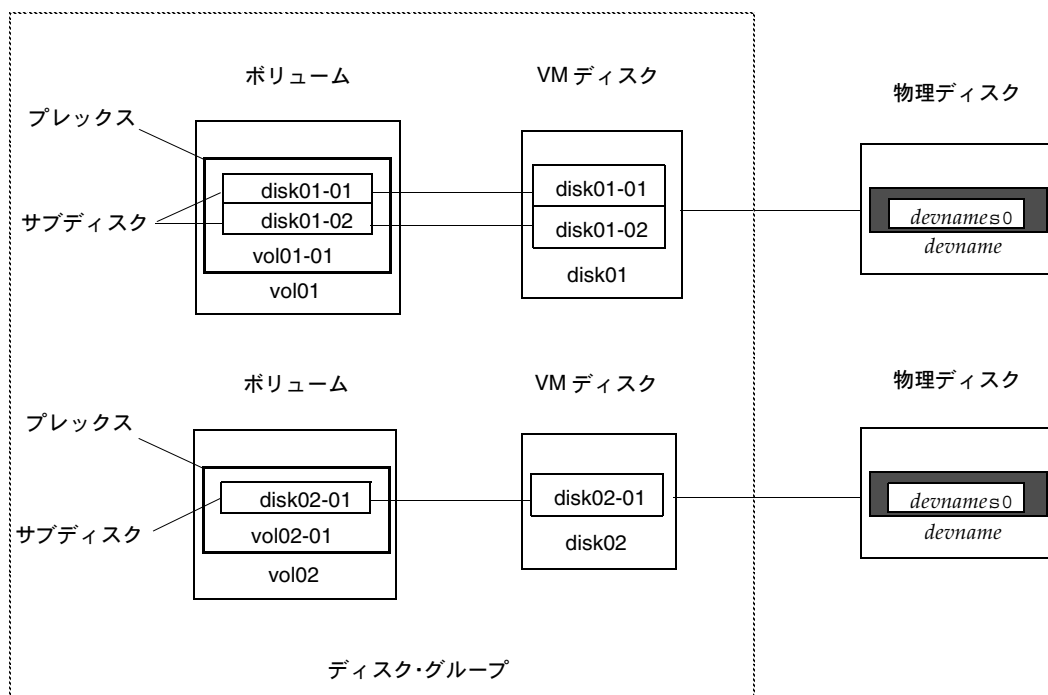
Volume Manager の仮想オブジェクト間の関係

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。ボリュームを構成する要素であ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、およびプレックスがあります。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 に、Volume Manager の仮想オブジェクト間の関係と、それらの物理ディスクとの関連の例を示します。この図は2つの VM ディスク (disk01 および disk02) を含むディスク・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



仮想オブジェクトのデータ編成（ボリューム・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。

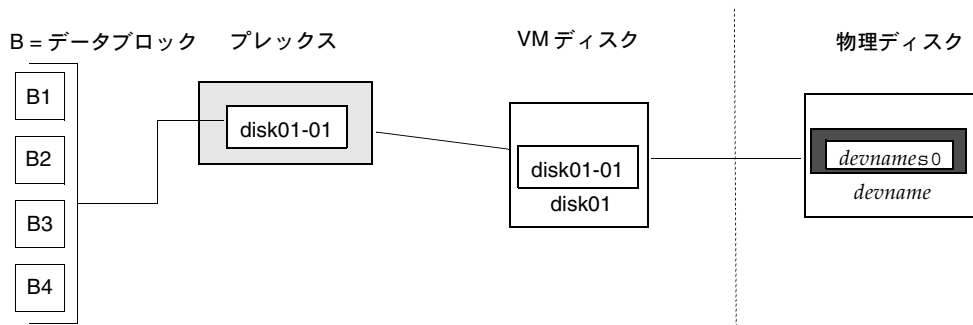
コンカチネーション

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づけることです。コンカチネーテッド・プレックスのデータすべて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。

コンカチネーテッド・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 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スペースが十分でない場合は、複数のサブディスクを使用して連結することもできます。このような連結形式は、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 は、連結構成内のボリュームを示したものです。

図 11 連結構成内のボリュームの例

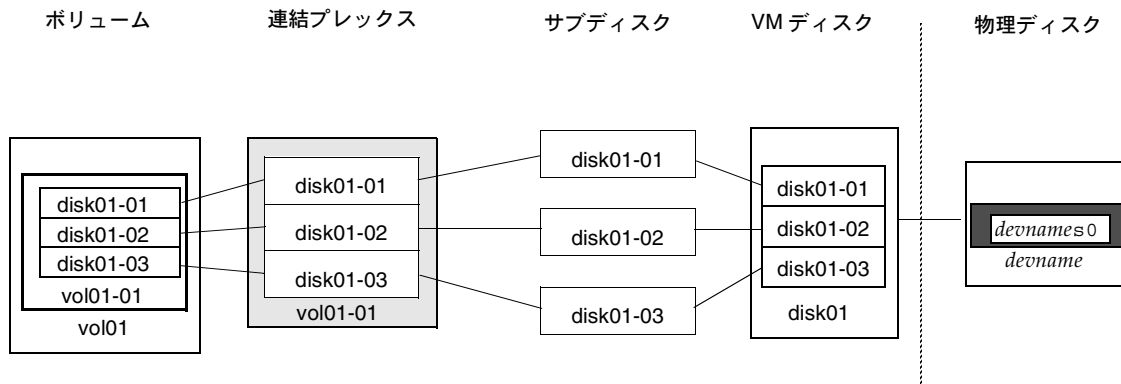
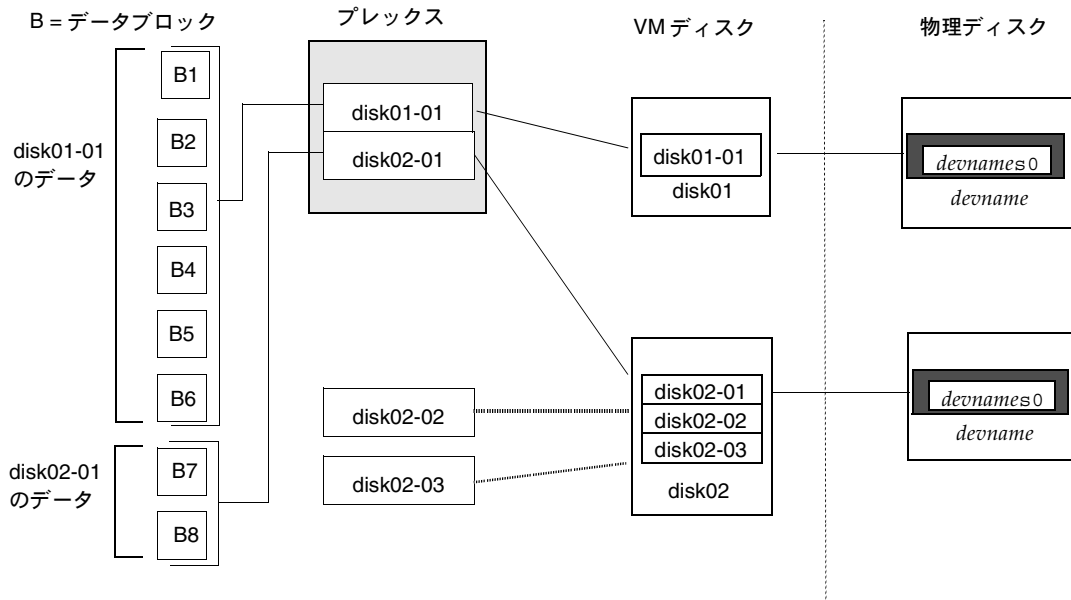


図 12 の例では、最初の 6 つのデータブロック (B1 から B6) が、VM ディスク `disk01` が割り当てられているディスク上のほとんどの領域を使用します。これには、VM ディスク `disk01` 上のサブディスク `disk01-01` 上の領域のみが必要となります。ただし、最後の 2 つのデータブロック (B7 および B8) は、VM ディスク `disk02` が割り当てられているディスク上の領域の一部のみを使用します。

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処理で使用可能になっています。

図 12 に、スパンド・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々につなぐように、データをマッピングすることです。ストライプト・プレックス には、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ト・プレックスのサブディスクの間で順々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。そのさい、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を含められます。それらのサブディスクの出所は、1つ以上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。

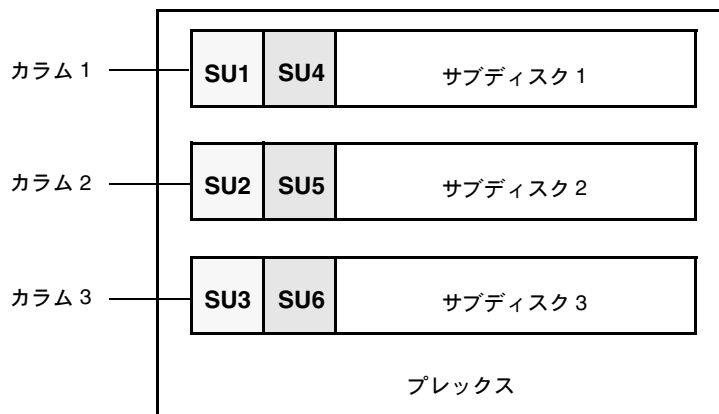
注意！ 1つのボリュームをストライピング、つまり、複数のディスクにわたってボリュームを分割すると、1件のディスクの障害がボリューム全体の障害につながる確率が高くなります。たとえば、5つのボリュームを5つのディスクにわたってストライピングしている場合に、いずれか1つのディスクに障害が発生したときは、5つのボリュームすべてをバックアップからリストアする必要が生じます。各ボリュームが別々のディスク上にある場合は、1つのボリュームをリストアするだけで済みます。1つのディスクの障害が多数のボリュームの障害を引き起こす可能性を低くするには、ミラーリングまたは RAID-5 を使用してください。

データは均等なサイズの単位（ストライプ・ユニット・サイズと呼びます）に分けられて、各カラムに順々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、ストライプト・プレックス内に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 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ー・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 に、サブディスクを 1 つずつ含む同じサイズのコラムが 3 つあるストライプト・プレックスを示します。物理ディスクにはコラムが 1 つずつ含まれます。

図 14 コラムごとに 1 つのサブディスクを含むストライプト・プレックスの例

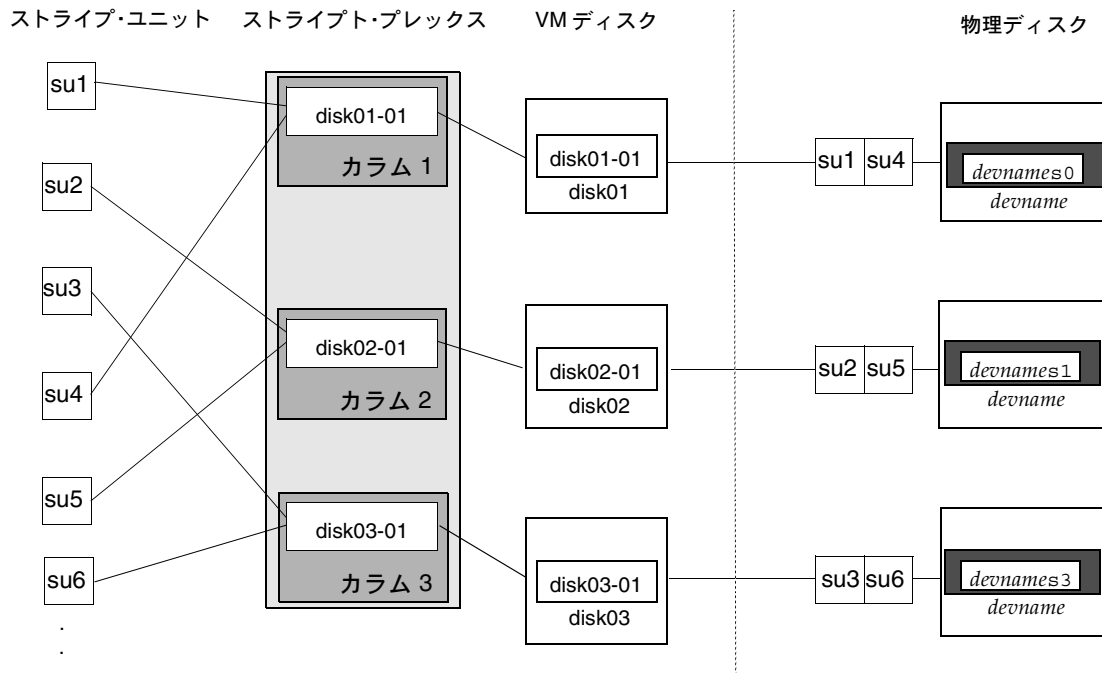
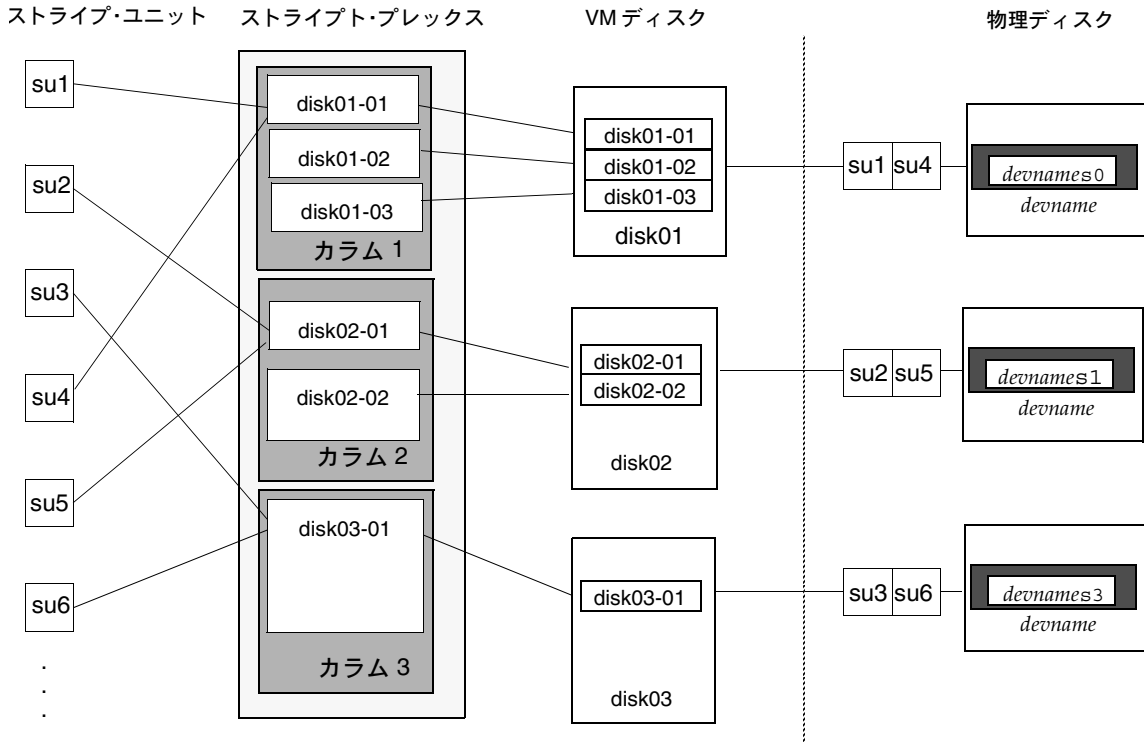


図 14 の例は、VM ディスク上の全領域を占めている 3 つのサブディスクを示しています。ストライプト・プレックスの各サブディスクに VM ディスクの一部分のみを割り当てることも可能です。そうすると、空き領域をほかのディスク管理処理用に残しておくことができます。

図 15 に、サイズの異なるサブディスクを含むコラムが 3 つあるストライプト・プレックスを示します。各コ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コラムが 1 つずつ含まれます。ストライプト・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（プレックスのサイズが大きくなった場合など）領域を割り当てることもできます。コラムには、必要に応じて、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ーテッド・サブディスクを含むストライプト・ブレイクの例



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ために使用される計算値です。RAID-5 ボリュームへのデータ書き込み中に、**排他的論理和 (XOR)** 演算によってパリティが計算され、その結果が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」ブレイ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 に、RAID-5 アレイ構成におけるパリティの位置を示します。各ストライプには、パリティ・ストライプ・ユニットを含む 1 つのカラムとデータを含むカラムがあります。パリ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 つのパリティ・ディスクがデータを受け入れ可能になるまで待たずに、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 モデルのパリティの位置

D	D	P
D	P	D
P	D	D
D	D	P

D = データストライプ・ユニット
P = パリティ・ストライプ・ユニット

RAID-5 の詳細と Volume Manager での実装方法については、「Volume Manager および RAID-5」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（ブックス）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、そのディスク上のブ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。ボリュームの中のブックスは 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 つ以上のブックスが必要です。各ブックスには、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、ブ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・プラス・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。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）を持たせることができます。

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ト・ブ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。このミラーのレイアウトタイプは、連結またはストライピングすることができます。

ストライピング・プラス・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ることができます。以前のリリースでは、ボリュームにミラーリングを併用するには、ミラーリングを主としストライピングを従と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。

ミラーリングを主としストライピングを従とすると、1つのディスクに障害が発生したときにブレックス全体が切り離されます。そのため、ボリューム全体の冗長性が失われます。ディスクを入れ替えた後で、ブ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うしなわれるだけです。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、「階層構成ボリューム」を参照してください。

Volume Manager および RAID-5

このセクションでは、Volume Manager で RAID-5 を実現する方法を説明します。RAID-5 の一般的な情報については、「RAID-5」を参照してください。

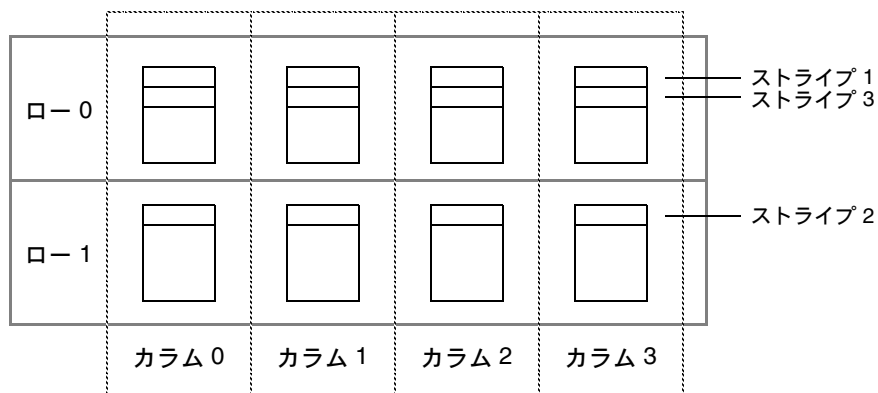
ミラーリング（RAID-1）と RAID-5 ではどちらもデータに冗長性を持たせますが、その方法が異なります。ミラーリングでは、1つのボリューム内のデータの完全なコピーを複数作成することによって冗長性を保持します。ミラーリングされているボリュームに書き込まれたデータは、すべてのコピーに反映されます。ミラーリングされているボリュームの一部に障害が発生した場合、他のコピーのデータが使用されます。

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、障害発生後のデータを復元するために使用される計算値です。RAID-5 ボリュームへのデータ書き込み中に、**排他的論理和（XOR）** 演算によってパリティが計算され、その結果が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。カラムとはアレイ内で同じ序数の位置にあるディスクの集まりを示し、ローとはパリティ・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を示します。図 17 に、従来の RAID-5 アレイのロー・カラム配置を示します。

図 17 従来の RAID-5 アレイ

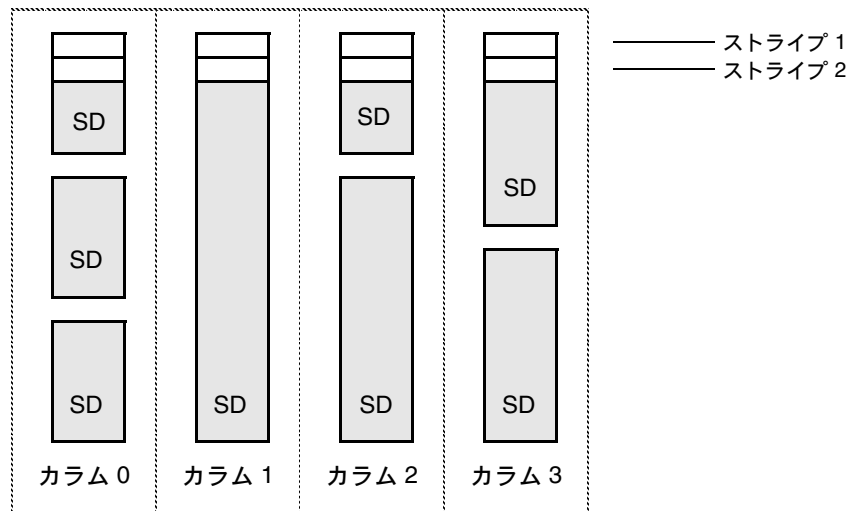


こうした従来のアレイ構造は、ローを追加することによってデータの増加に対処しています。ストライピングを行うには、1 番目のストライプをロー 0 のディスクに適用し、2 番目のストライプをロー 1 のディスクに適用し、3 番目のストライプをロー 0 のディスクに適用するといったぐあいにします。この種類のアレイでは、すべてのディスク（パーティション）、カラム、およびローのサイズを等しくする必要があります。

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 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)」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するといったぐあいに、カラムの末尾に達するまで適用していきます。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。

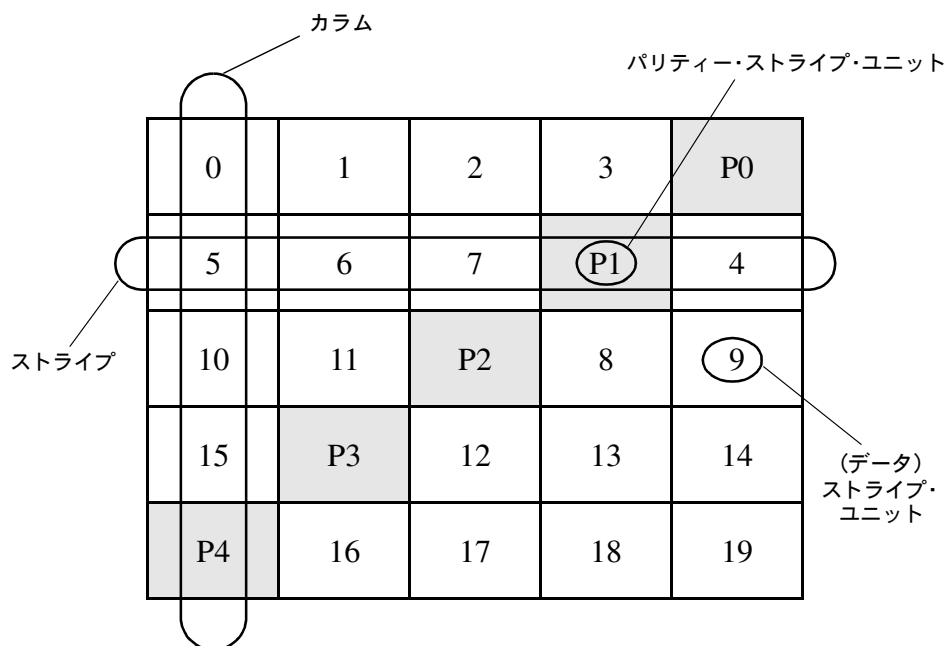
左対称レイアウト

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。そのうち、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と大量の順次入出力操作のどちらの場合でも、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。ストライプの数がカラムの数より多い場合、パリティ・ストライプ・ユニットは右端のカラムに戻って順に配置されます。

図 19 に、1 カラムにつき 5 つのディスクがある左対称パリティ・レイアウトを示します。

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、1 番目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 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 4 および 5 から 7 の順に続きます。残りのストライプも同様にして編成されます。

各パリティ・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・ストライプ・ユニット (ストライプごとに) で 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。そうすると、ストライプ・ユニット 1 から 3 をパリティ・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・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 を復元する、といったぐあいに処理がなされます。

注：RAID-5 レイアウトでブ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。障害から回復した後で、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るでしょう。

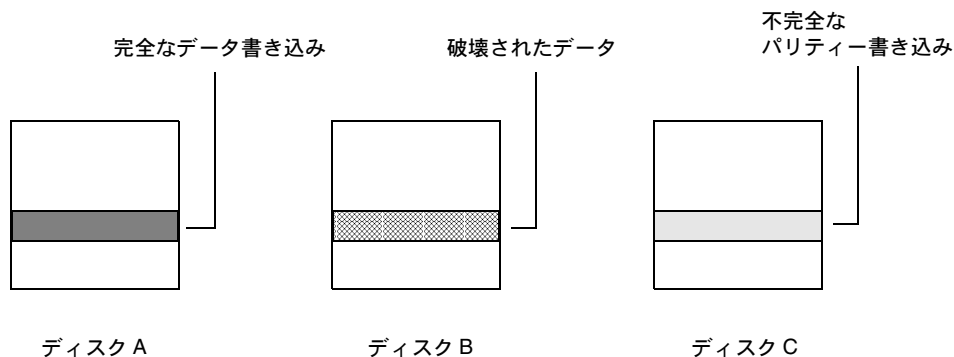
ロギング

ロギング (記録) は、復元のために使用する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) 上に取れます。その後で、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、誰も解らないうちに破壊される場合があります。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 は、ディスク B のリカバリの成否が、ディスク A のデータとディスク C のパリティの両方の書き込みが完了していることに依存す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、追加の非 RAID-5 レイアウト・プレックスとして RAID-5 ボリュームに関連付けられます。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。

階層構成ボリューム

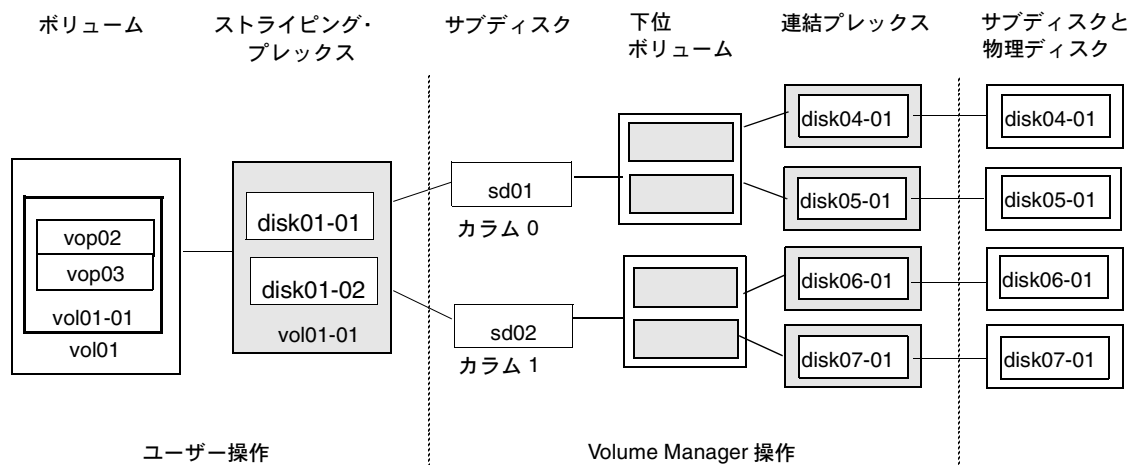
Volume Manager のもう 1 つの仮想オブジェクトとして階層構成ボリュームがあります。階層構成ボリュームは、ボリュームの上に構築されます。階層構成ボリューム構造は標準のボリューム構造に比べて、障害に対する耐性や冗長性が高められています。たとえば、ストライピングおよびミラーリングした階層構成ボリュームでは、各ミラー（プレックス）が扱うストレージの容量が少ないため、標準ミラー・ボリュームよりすばやくリカバリできます。図 21 に階層構成ボリューム設計の例を示します。

注：階層構成ボリュームは、Volume Manager 3.0 ではサポートされますが、それ以前のバージョンではサポートされません。また、ボリューム（ストレージ・ボリューム）の上にサブディスクを構築することも、Volume Manager 3.0 以降でサポートされますが、それ以前のバージョンではサポートされません。

図 21 に示すように、「ユーザー操作」領域においてボリュームおよびストライプト・ブレイクスに対して、通常の Volume Manager 処理を実行することができます。ユーザー処理は、階層構成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。階層構成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。必要な操作（カラム幅の変更、カラムの追加など）はすべて、ボリュームおよびストライプト・ブレイクスが位置しているユーザー操作領域から実行できます。

「Volume Manager 操作」領域は、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を示します。階層構成ボリュームは 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、ユーザーが操作するようには設計されてはいません。ここでは、階層構成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構成ボリュームの例



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、階層構成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 は階層構成ボリュームを使用して、次の処理や操作を実行します。

- ストライプト・ミラー (vxassist マニュアル・ページを参照)
- コンカチネーテッド・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 のユーザー・インタフェース

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

ユーザー・インタフェースの概要

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

- Volume Manager Storage Administrator

Storage Administrator は、Volume Manager へのグラフィカル・ユーザー・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、『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ー入力が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、スーパーユーザー権限または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、『Command Line Interface Administrator's Guide』を参照してください。

- Volume Manager 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の管理機能を実行するためのメニュー方式インタフェースです。vxdiskadm では、コマンドライン・メインメニューから、実行するストレージ管理処理を選択することができます。vxdiskadm については、『Command Line Interface Administrator's Guide』の第 5 章「メニュー・インタフェース処理」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 の概念

この節では、主な用語と Volume Manager の概念について説明します。図 22 に、この節で説明する用語と概念のまとめを示します。

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します。

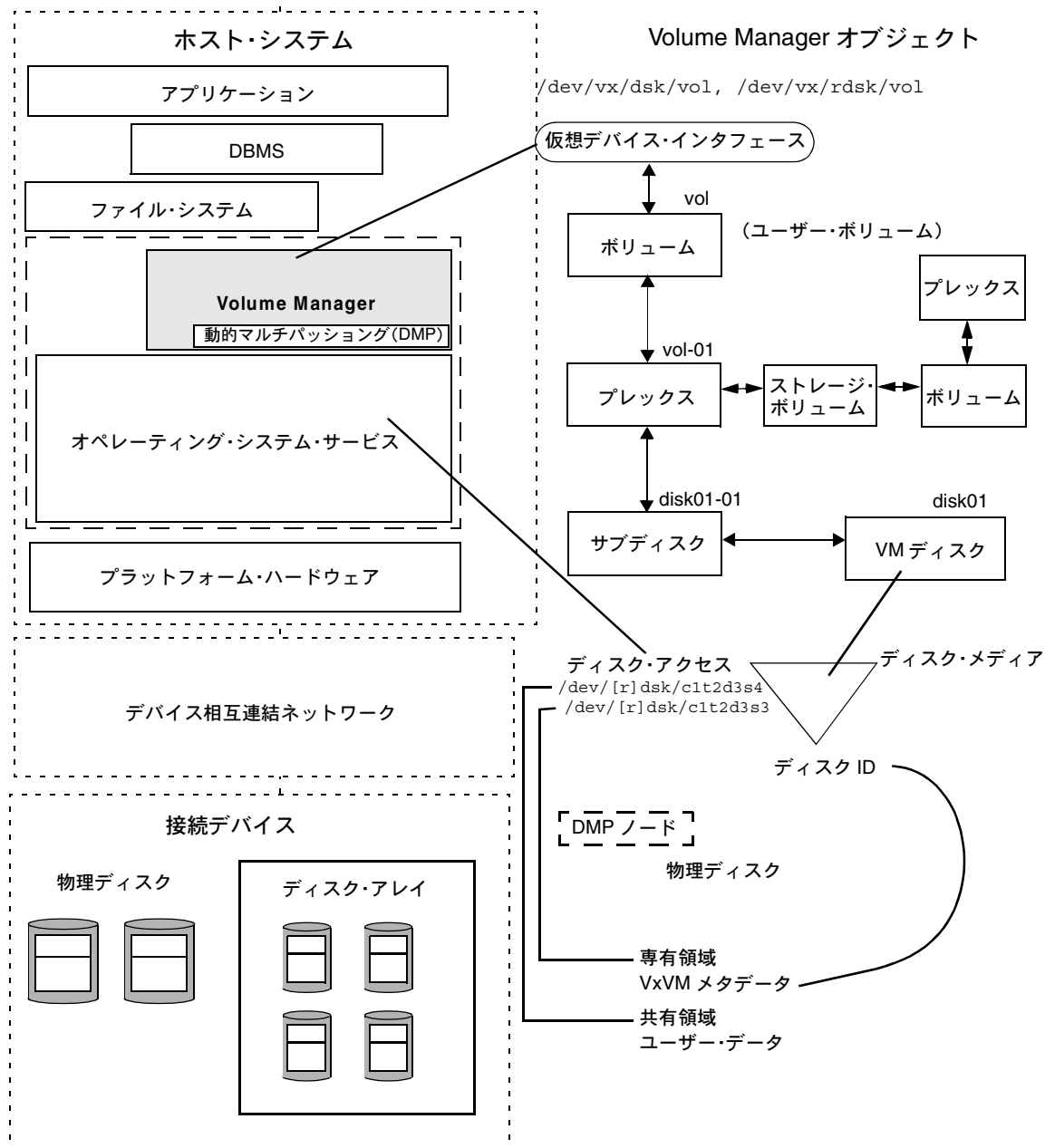
- スペース — 割り当ておよび使用
- パフォーマンス — データ伝達機能の向上
- データの可用性 — 連続操作および複数システムへのアクセス
- デバイスのインストーラー — サポートの集中化および最適化
- システム — 複数システムのサポートおよび専用 / 共有システムの監視

Volume Manager のオブジェクト

Volume Manager はストレージ管理サブシステムです。このシステムを利用して、物理ディスクをボリュームと呼ばれる論理デバイスとして管理することができます。Volume Manager インタフェースを通じて、ボリュームにもとづく高度なデータ・アクセスとデータ・ストレージ管理を実現することができます。ボリュームは論理デバイスですが、データ管理システムからは物理ディスク・パーティション・デバイスのように見えます。ボリュームは、高度な回復機能、データの可用性、パフォーマンス、およびストレージ設定オプションの基です。

Volume Manager のボリュームは論理オブジェクトです。Volume Manager によって作成されるオブジェクトには他にもあります。それら进行操作、制御、監視したり、それらを対象にして照会を実行することにより、ストレージ管理を最適化できます。使用するボリュームを設定および維持するため、物理ディスクは Volume Manager の管理下に置かれ、ディスク・スペースを集めてディスク・グループが編成されます。ディスク・グループとは、提示されたディスクをまとめて論理ボリュームに編成したものです。Volume Manager では、それらのディスク上のスペースを論理ボリュームに割り当てます。

図 22 Volume Manager システムの概念



たとえば、ホスト・システムに Volume Manager をインストールした後で、Volume Manager オブジェクトの構成を設定して使用するには、先に次の処理を実行する必要があります。

- 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。

- 物理ディスクの管理を Volume Manager に移せる
- 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、システムの復旧を早めることができます。

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する必要があります。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 動的マルチパッシング（DMP）メタデバイス

動的マルチパッシング（DMP）

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、マルチパッシング条件が存在します。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にすることができます。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・ハンドルに対応づけられ、適切なマルチパス・ポリシーを設定されます。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。ディスク・アレイの中には、同時に複数のパスをアクティブにできる (アクティブ / アクティブ) ものもあれば、1 つのパスのみをアクティブにして、既存のパスに障害が発生した時の予備のパスとして別のパスを確保している (アクティブ / パッシブ) もものもあります。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、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 Volume Manager は、設計されています。VxVM 論理操作を単純化するために、各 VM ディスクは一意の Volume Manager DMP メタノードに対応づけられます。それは、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。

Volume Manager DMP を使用する場合は、デバイス認識の階層構成について心得ている必要があります。

- オペレーティング・システムからはパスがどのように見えるか
- Volume Manager DMP からはパスがどのように見えるか
- マルチパスのターゲットでパスをどのように扱っているか

DMP に関するその他の情報については、次のドキュメントを参照してください。

- Getting Started Guide (本書)、第 2 章
- Command Line Interface Administrator's Guide
 - 第 2 章、VM ディスクのマルチパスの表示
 - 第 4 章、ボリューム設定デーモン vxdctl
 - 第 5 章、ディスク処理

- Administrator's Reference Guide
 - 第1章、Volume Manager 設定の回復
 - 第3章、vxdisk を使用したマルチパス情報の表示
 - 付録 A、DMP エラー・メッセージ

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づけられます。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・スペースが再構成され、それに基づいて実行時に入出力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づ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スペース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づけ、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、「階層構成ボリューム」を構成するさいには、サブディスクを従来どおりに VM ディスクに関係づけることも、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づ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。

階層構成ボリュームでは、論理構成要素を様々に組み合わせることが可能です。それ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、階層構成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おそれがあります。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構成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 VM 一連の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。

Volume Manager の機能

2



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
- Volume Manager のルート機能
 - ルート・ボリュームを使用するのブート
 - ブート時のボリュームの制限
- 動的マルチパッシング (DMP)
 - パス・フェイルオーバー機構
 - 負荷の平準化
 - DMP デバイスからのブート
 - コントローラの有効化と無効化
 - DMP データベース情報の無効化
- VxSmartSync Recovery Accelerator
 - データ・ボリュームの設定
 - REDO ログ・ボリュームの設定

- Volume Manager Task Monitor
- 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・レイアウトをその他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・スペース（ミラーリング）を複製するか、パリティ（RAID-5）を追加することによって、ストレージの冗長性を高めます。Volume Manager でストレージの性能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 からミラーリングへの変更、およびミラーリングから RAID-5 への変更
- 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。

- コンカチネーション
- ストライピング
- RAID-5
- ミラーリング（複数のストレージ・デバイスにわたって複製されたデータもサポート）
- ストライプト・ミラー
- コンカチネーテッド・ミラー

オンライン再レイアウトの動作

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、既に設定しているストレージ・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。

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ピング・レイアウトがあり、そのパフォーマンスが十分でないと仮定します。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。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。

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。

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先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にコピーされ、そのデータはソース・ボリューム・ストレージ領域から削除されます。その後で、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、またはミラーリングから 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、ミラーリングされた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 内の項目について次に説明します。

- 「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、再レイアウトは行えないことを示します。

- 数字は、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です。
- 説明されている処理は、両方向で実行可能です。

表 1 は、サポートされているレイアウト変更を示したものです。

表 1 サポートされているレイアウト変更

変更元 / 変更先	ストライピング ミラーリング	コンカチネー テッド・ミラー リング	通常の ミラーリング	RAID-5	コンカチネー ション	ストライピング
ストライピング ミラーリング	可 1	可 2	不可 3	可 4	可 5	可 6
コンカチネーテッ ド・ミラーリング	可 7	不可 8	不可 9	可 10	不可 11	可 12
通常のミラーリング	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
コンカチネーション	可 5	不可 11	不可 17	可 21	不可 23	可 24
ストライピング	可 6	可 12	不可 18	可 22	可 24	可 25

表 1 に示す番号は、次のレイアウト処理を示します。

1. ストライプ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. これは再レイアウト処理ではなく、変換処理です。
4. ミラーリングから RAID-5 への変更、およびストライプ幅 / カラムのいずれかまたは両方の変更を行います。
5. ミラーリングから RAID-5 への変更、およびストライプ幅 / カラムのいずれかまたは両方の変更を行います。
6. ストライプ幅 / カラムを変更して、ミラーを削除します。
7. カラムを追加します。
8. これは再レイアウト処理ではありません。
9. これは変換処理です。
10. ミラーリングから RAID-5 に変更します。vxconvert 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. 既存のミラーリングされているボリュームがストライプト・ミラーに変更されます。これは、カラムまたはストライプ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 を参照してください。
14. 既存のミラーリングされているボリュームがコンカチネーテッド・ミラーに変更されました。これは、カラムまたはストライプ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. 既存のミラーリングされているボリュームを RAID-5 に変更します。既存のミラーリングされているボリュームのプレックスを選択して、再レイアウトを使用する必要があります。その他のプレックスは、再レイアウト処理の最後に削除されます。
17. ミラーリングされているボリュームのプレックスを選択して、カラムまたはストライプ幅を変更した場合にのみ、これは再レイアウト処理になります。
18. ミラーリングされているボリュームのプレックスを選択して、カラムまたはストライプ幅を変更した場合にのみ、これは再レイアウト処理になります。
19. これは再レイアウト処理ではありません。
20. ストライプ幅 / 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. パリティを削除します。
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. ストライプ幅 / カラム数を変更します。

ストライプト・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ト・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、ストライプト・ミラー・プレックスと呼ばれます。コンカチネーテッド・プレックスも同じように作成されます。オンライン再レイアウトでは、ストライピングミラー・プレックスからコンカチネーテッド・ミラー・プレックスへの変更と、コンカチネーテッド・ミラー・プレックスからストライピングミラー・プレックスへの変更の両方をサポートしています。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。

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編成替えされます。

レイアウト変更中、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、レイアウト変更中でも中断されません。

レイアウト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、システムの復旧後に続行され、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、空のプレックスは変更されません。また、プレックスを空にすることもありません。

オンライン再レイアウトは、`vxassist` コマンドを使用して作成されたボリュームにのみ適用できます。

サポートされていないレイアウト変更

ログ・プレックスのレイアウト変更は、サポートされていません。

ボリュームでオンライン再レイアウト処理が実行されている場合、ボリュームのスナップショットはサポートされていません。

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それらのオブジェクトへのアクセスと冗長性を復元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て、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復元します。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、および回復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ー）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スペースの探し出し、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ことは保証されません。システム管理者は、ホットリロケーション実行後に、設定の変更を行う必要があります。

リロケート用スペースの選択方法

スベア・ディスクは交換処理で使用する前に、初期化し、スベアとしてディスク・グループに配置しておく必要があります。障害発生時に、スベ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スペースを使用します。スベア・ディスク領域が不足している場合は、スベア領域と空きスペースを組み合わせ使用します。

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・スベアとして指定できます。そのためには、Storage Administrator インタフェースの `vxdiskadm` や `vxedit` (VERITAS Volume Manager 『VERITAS Volume Manager Administrator's Reference Guide』を参照) を使用します。スベ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。

リロケート用スペースの選択時、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。使用可能なスベア・ディスクや空きスペース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・スペースが空きスペース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有効に活用するための情報は、第3章「Volume Manager の初期化と設定」を参照してください。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。

ボリュームの再同期化

ミラーリングされているボリュームまたは 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、データのすべてのコピーが正確に一致することが Volume Manager によって保証されます。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ること、または同期しないことがあります。ミラーリングされたデータは、元のデータと正確に同じではありません。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、ボリュームのミラーにデータは、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、一部書き込みが失敗して、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場合があります。ミラーリングされているボリュームで、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発生します。RAID ボリュームの場合、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、すべてのミラーに完全に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保証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、ボリュームが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ー・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ー・フラグが削除されます。システムのリポート時にダーティー・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 - 回復モードになり、すべてのパリティが再生成されます。ミラーリングされているボリュームの場合は、ボリュームを回復モード（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実行されるため、回復処理中でもボリュームを使用することができます。

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。回復処理は、回復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のかかることがあります。これらの影響に対しては、ミラーリングされているボリュームではダーティ・リージョン・ロギングを適用するか、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」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リングされているボリュームを迅速に回復するために使用されます。DRL では、入出力のミラーリングされているボリュームへの書き込みによって変更された領域の記録を取ります。DRL はこの情報を使用して、ボリュームのうちの回復が必要な部分のみを回復します。

DRL を使用していない場合に、システムに障害が発生したときは、ボリュームのすべてのミラーを整合性のとれた状態に復元する必要があります。復元処理では、ボリュームのすべての内容をミラー間でコピーします。この処理に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回復する必要があります。

DRL では、ボリュームを一連のリージョンに論理的に分割し、書き込みがなされている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ー・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ーになる場合は、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ー・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、使用できなくなるおそれがあります。

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・サブディスクを含むことができます。

ログ内では、常にある限られた数のビットのみにダーティ・フラグが設定されます。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。

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。

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、システムを正常に起動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回復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再起動することができなくなります。

起動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再起動できます。起動 (ルート) ディスクのミラーリングとシステム回復手順の詳細については、『Administrator's Reference Guide』の第1章「リカバリ」を参照してください。

ルート・ボリュームを使用時のブート

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。これは、ユーザー処理を実行して Volume Manager の設定を読み込みボリュームを始動させる前になります。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 デーモンを 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設定することができなくなります。

この制限に引っかからないように、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づけられた連続するディスク・スペースから作成されます。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 (起動時に使用するプレックス) をストラッピングまたはスパン化することはできません。

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、システムの起動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。『Administrator's Reference Guide』の「ディスク・グループ用のマイナー番号の予約」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・スペースとちょうど同じ大きさのパーティションのことです。起動時は、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（「ルート・ボリュームを使用するのブート」を参照）。
- パフォーマンスを向上させるために、ストライプト・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、システムの回復や起動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。
- 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スペースがないと、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）。

動的マルチパッシング（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 の動的マルチパッシング機能は、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 1 つの物理ディスクは、1 つのパスを持つメタデバイスとして、ディスク・アレイに含まれる 1 つのディスクは、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。

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復元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。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一次パスに送られます。一次パスに障害が発生すると、入出力は、使用可能なほかの一次パスまたは二次パスに引き継がれます。アクティブ / パッシブ・ディスク・アレイでは、複数パスにわたる負荷の平準化は実行されません。これは、あるコントローラから別のコントローラに LUN の所有権が連続的に移転され、その結果的として入出力速度が著しく低下することのないようにするためです。

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でも、システムの信頼性を向上します。

コントローラの有効化と無効化

この機能を使用すると、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。処理の完了後に、ホスト入出力コントローラへの入出力操作を有効にして、再びこのコントローラを通じて入出力操作が行われるようにすることができます。この操作は、Volume Manager の `vxddmpadm (1M)` ユーティリティを使用して実行できます。

たとえば、システムに StorEdge A5000TM 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・ボードをユーザーが変更する必要がある場合は、`vxddmpadm (1M)` 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 はパス不可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一次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二次パスまたは別の入出力コントローラ上のアクティブな一次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、`vxddmpadm (1M)` 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無効化

`vxddmpadm (1M)` ユーティリティを使用して、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、ユーザー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、この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™ と一緒に使用されます。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、ボリューム回復時間を大幅に短縮できます。

VxSmartSync 機能を有効にするには、Oracle インスタンス起動パラメータ・ファイルの `_ENABLE_ORACLE_RESILVERING` パラメータを TRUE に設定する必要があります。そのため、`initSID.ora` ファイルに次の行を追加します。

```
_ENABLE_ORACLE_RESILVERING      = TRUE
```

このファイルの `SID` は、Oracle データベース・インスタンスのシステム識別子です。このファイルは通常、ディレクトリ `${ORACLE_HOME}/dbs` に保存されています。

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表領域ファイル）。

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 つのボリュームの唯一の違いは、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。回復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、システムの再起動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（つまり、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ド・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ので、データベース自体は再同期（resilvering）が必要なセクション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 ログとして使用する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ド・ボリュームとして設定する必要があります。

Volume Manager Task Monitor

Volume Manager Task Monitor 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。Task Monitor 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）。Task Monitor のコマンドライン操作については、「Volume Manager Task Monitor のオペレーション」を参照してください。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 にはオプションのクラスタ機能が備わっています。この機能を利用すると、VxVM をクラスタ環境で使用することができます。そうすると、複数のホストから同時に、Volume Manager の管理下にある一覧のディスク (VM ディスク) にアクセスして管理することができます。

クラスタとは一組のディスクを共有する一組のホストです。クラスタ内の各ホスト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。ある一つのノードが障害を起こしても、他のノードは依然として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 の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、『VERITAS Volume Manager Administrator's Reference Guide』の第 5 章「Volume Manager のクラスタ機能」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。

Volume Manager の初期化と 設定

3



はじめに

この章では、Volume Manager の初期化に必要な手順および Volume Manager が稼動するために実行する必要の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 をインストール方法および設定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
 - vxinstall 実行後のリブート
- Volume Manager のデーモン
 - 設定デーモン vxconfigd
 - ボリューム入出力デーモン vxiod
- システムの設定
- システム・セット設定手順の例
- システム設定ガイドライン
 - ホットリロケーションのガイドライン
 - ストライピングのガイドライン
 - ミラーリングのガイドライン
 - ダーティ・リージョン・ロギング (DRL) のガイドライン
 - ミラーリング・アンド・ストライピングのガイドライン

- ストライピング・アンド・ミラーリングのガイドライン
- RAID-5 のガイドライン
- 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。`vxinstall` を使用して、1 つ以上のディスクを `rootdg` に含める必要があります。その後で `vxdiskadm` または `Storage Administrator` を使用して、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. スーパーユーザー権限（または適切なアクセス権限）でログインします。
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。`disks.exclude` ファイルは `/etc/vx/` の下に置いてください。
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、`cntrl.exclude` ファイルを作成します。`cntrl.exclude` ファイルは `/etc/vx/` の下に置いてください。
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
- クイック・インストールまたはカスタム・インストール手順の説明

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、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・インストールの詳細な手順については、『*VERITAS Volume Manager Installation Guide*』を参照してください。

vxinstall 実行後のリブート

選択したシステムの設定方法によって、シャットダウンと再起動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。

次のコマンドを使用すると、Volume Manager の主要プロセス (vxconfigd、vxnotify、および vxrelocd) が稼動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 つのデーモンが稼動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。

```
vxctl mode
```

vxconfigd が既に動いており、有効な場合は、次のメッセージが表示されます。

```
mode: enabled
```

vxconfigd が既に動いているのが、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。

```
vxctl enable
```

vxconfigd が動いていない場合は、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、起動されると、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・ブート時に起動されます。通常、複数の vxiod デーモンが常に動いています。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 デーモンが動い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、前述の方法が vxiod デーモンが動いていることを確認する唯一の方法です。

vxiod デーモンが動いている場合は、次のメッセージが表示されます。

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

「10」は動いている vxiod デーモンの数を示します。

vxiod デーモンが1つも動いていない場合は、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。具体的な設定手順については、『Volume Manage Command Line Interface Administrator's Guide』および『Volume Manager Storage AdministratorAdministrator's Guide』を参照してください。

次のシステム設定手順は一般的なもので、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・セット設定手順の例

ストレージ管理システムの設定時に使用可能な一般的な処理を以下に順に説明します。

▼ 初期設定

1. Volume Manager の管理下にディスクを組み入れます。
2. 新しいディスク・グループを作成します（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）。
3. ボリュームを作成します。
4. ボリュームに 1 つ以上のファイル・システムを収めます。

▼ 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。
- 必要に応じて、ミラーをボリュームに追加する。

▼ 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。各セクションで示される参照先では、それぞれのガイドラインについての詳細情報が紹介されています。

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、第2章「Volume Manager の機能」の「ホットリロケーション」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ではスベア・ディスクを指定することは必須ではありませんが、各ディスク・グループ内の少なくとも1つのディスクをスベアとして指定しておく、再配置に使用するディスクをある程度制御することができます。スベア・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スペースが使用されます。再配置に空きスペースを使用すると、再配置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ベアとして指定し、スベア・スペースを増やすことができます（元のスベア・スペースの一部が再配置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、複数のスベア・ディスクがある場合は、スベア・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、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、使用可能なスペースを含むスベア・ディスクか、空きスペースを含む通常のディスクのいずれかである必要があります。
- ミラーリング・アンド・ストライピング・ボリュームの場合、ディスク・グループに、そのボリュームのいずれかのミラー、またはストライプド・プ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、使用可能なスペースを含むスベア・ディスクか、空きスペースを含む通常のディスクのいずれかである必要があります。
- 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 プレックス（またはそのログ・プレックス）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、使用可能なスペースを含むスベア・ディスクか、空きスペースを含む通常のディスクのいずれかである必要があります。

- ミラーリングされているボリュームのデータブレッक्सの一部に DRL ログ・サブディスクが含まれる場合、そのブレックスを再配置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ブレックス（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。再配置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ベア・スペースか空きスペースが必要です（rootvol および swapvol では連続したディスク・スペースが必要）。
- VxVM オブジェクトをスベア・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）、用意したスベア・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、第1章「Volume Manager とは」の「ストライピング（RAID-0）」を参照してください。

- 同じ物理ディスク上に、ストライプト・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、中くらいのストライプ・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ドラ間で異なります（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む）。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言えません。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、中くらいのストライプ・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・サイズが大きくてマルチスレッドであるか、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・アンド・ミラーリングストライピング・アンド・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つのプレックスのみがストライピングされている場合は、ボリュームのポリシーを、そのストライプト・プレックスに対して `prefer` に設定してください（デフォルトのポリシー `select` ではそれが自動的に行われます）。
- 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、各ストライプト・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ト・プレックスにスペースを割り当てる際に、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、第1章「Volume Manager とは」の「ミラーリング (RAID-1)」を参照してください。

- ミラード・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、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つのブ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ー・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動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連結したり、より高速なデバイス上に配置すると、ブ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」ブ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、ストライプト・ブレックスからなるボリュームでは、ストライプト・ブ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、第2章「Volume Manager の機能」の「ダーティ・リージョン・ロギング」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ド・ボリュームの回復を高速化できます。DRL が有効になっている場合、Volume Manager はブ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スペースを監視します。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・サブディスクはブレックスとは独立しており、ブ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。

- ダーティ・リージョン・ロギング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、各ブ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、必要なデータを含め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。そうしておく、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（偶数が望ましいですが - セクタ数が奇数の場合、ログ・サブディスク中の最後のセクタは使用されない）。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、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、第1章「Volume Manager とは」の「ストライピング・プラス・ミラーリング (RAID-0 + RAID-1)」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。ストライプト・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」で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」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・アンド・ミラーリングを併用できます。詳細については、第1章「Volume Manager とは」の「ストライピング・プラス・ミラーリング (RAID-0 + RAID-1)」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が十分に十分に確保します。ストライプト・ブレイクスには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。
- あるブレイクスのサブディスクを、別のブレイ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」で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」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。詳細については第1章「Volume Manager とは」の「RAID-5」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 ボリュームにつき、1つの RAID-5 ブレイクスのみ含むことができます（ただしログ・ブレイクスは複数含むことも可能）。
- RAID-5 ブレイクスは、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・ブレイクスがある場合、それらのログ・ブレイクスは RAID-5 ブレイ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は、ボリュームに関連付けられている RAID-5 ブレイクスがあればその長さに設定されています。そうでなければ、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 ブレイクスがあるときは、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・ブレイクスがあるときは、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・ブレイクスがない場合は、ゼロに設定されます。
- 空の RAID-5 ログ・ブレイクスは無効です。

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、システムへのアクセスが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、システムの稼動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。

ここでは、システムおよびデータを保護する方法について助言します。

- 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で、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、『Administrator's Reference Guide』の「何通りかのルート（/）、*swap*、および *usr* の構成の設定」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ド・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と、いずれかのディスクが損傷した場合でもデータを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ド・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。

`vxassist` デフォルト・ファイルの詳細については、『Command Line Interface Administrator's Guide』の第4章を参照してください。

- 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再配置の試行、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、障害発生時の再配置に必要な十分なスペースを確保します。

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するスペア・スペースか空きスペースが必要です (rootvol および swapvol では連続したディスク・スペースが必要)。

- 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・ロギング機能を活用して、システム障害後のミラード・ボリュームのリカバリ処理を高速化できます。ミラード・ボリュームに 1 つ以上のログ用のサブディスクを用意します (rootvol、swapvol、および usr の各ボリュームは、ダーティ・リージョン・ロギングを適用することはできません)。
- 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。

索引

D

DMP

- DMP デバイスからのブート 48
- 動的マルチパッシング 47
- パス・フェイルオーバー機構 48
- 負荷の平準化 48

- DMP デバイスからのブート
 - DMP 48

G

GUI

- ストレージ管理者 20

R

- RAID-0 14
- RAID-1 19
- RAID-5
 - ガイドライン 63
- resilvering 50
- rootdg 6

S

SA

- GUI インタフェース 20
- Storage Administrator 27

V

- VM ディスク 5
 - 定義 5
- Volume Manager 27、28、35
 - オブジェクト 28
 - オペレーティング・システム 30
 - デーモン 55
 - ルート機能 45
 - レイアウト 32
- vxassist 64
- vxconfigd 46、55
- vxdiskadm 46
- vxiod 55、56
- vxrelocd 41
- VxSmartSync 50
- VxVM 27、35



オ

- オンライン再レイアウト 36
 - サポートされていないレイアウト変更 40
 - 障害の回復 40
 - 動作 36
 - 変更の種類 37
 - 変更の特性 40
 - レイアウト変更および
ボリュームの大きさ 40

カ

- ガイドライン
 - RAID-5 63
 - ダーティ・リージョン・ロギング 61
 - ミラーリング 60
 - ミラーリング・アンド・
ストライピング 62
- 概念 28
- カプセル化 45、64
- カラム、ストライピング 14

ク

- グラフィカル・ユーザー・インタフェース 27

コ

- コンカチネーション 12

サ

- 再同期化
 - Oracle データベース 50
 - ボリューム 43
- サブディスク
 - ログ 44

ス

- ストライピング 14
- ストライプ・カラム 14
- ストライプ・ユニット 15
- ストライプト・ブックス 14
- ストレージ・レイアウト
変換 36
- スパンニング 12
- スワップ・ボリューム制限 46

タ

- ダーティリージョンロギング 44
- ダーティ・リージョン・ロギング
ガイドライン 61

テ

- ディスク
 - VM ディスク 5
 - カプセル化 45、64
 - 障害とホットリロケーション 41
 - ブート・ディスク 64
 - ルート・ディスク 45
- ディスク・グループ 6
- ディスク・メディア名 5
- データ
 - 冗長性 19
 - 保護 64
- デーモン 55
 - Volume Manager 55
 - ホットリロケーション 41

ト

- 動的マルチパッシング
DMP 47



ナ

- 名前
 - ディスクへのアクセス 4
 - ディスク・メディア 5

ニ

- 入出力デーモン 56

ハ

- パーティション 4
- パス・フェイルオーバー
 - DMP 48
- バックアップ 65
- パリティ 18、21

ヒ

- 表示
 - マルチパス情報 31

フ

- ブート・ディスク 64
- 負荷の平準化
 - DMP 48
- プレックス 7
 - ストライプピング 14
 - 定義 7
 - ボリューム 8
 - ミラー 8

ヘ

- 変換
 - ストレージ・レイアウト 36

ホ

- ホットリロケーション 41
 - スベアとして指定 42
- ボリューム 8
 - 定義 3、8
 - プレックス 8
- ボリューム制限
 - ブート時 46
- ボリュームの再同期化 43

ミ

- ミラー 7、8
- ミラーリング 19
 - ガイドライン 60

ル

- ルート機能 45
- ルート・ディスク 45
- ルート・ボリューム
 - ブート 46
- ルート・ボリューム制限 46

レ

- レイアウト
 - 左対称 23

ロ

- ロギング 24
- ログ・サブディスク 44、61

