

VERITAS Volume Manager™ 3.1

Guide de l'administrateur

Solaris



Octobre 2000
30-000110-011


VERITAS



Exclusion de responsabilité

Les informations contenues dans cette documentation peuvent être modifiées sans préavis. VERITAS Software Corporation ne fournit aucune garantie en ce qui concerne le présent manuel, y compris, mais sans restriction, en ce qui concerne les garanties implicites de commercialisation et d'adéquation à un objectif particulier. VERITAS Software Corporation ne peut être tenue pour responsable des erreurs qui pourraient figurer dans cette documentation ni des dommages fortuits ou indirects relatifs à la délivrance et à l'utilisation de ce manuel.

Copyright

Copyright © 1999 - 2000 VERITAS Software Corporation. Tous droits réservés. VERITAS est une marque déposée de VERITAS Software Corporation aux États-Unis d'Amérique et dans d'autres pays. Le logo VERITAS et VERITAS Volume Manager sont des marques de VERITAS Software Corporation. Toutes les autres marques ou marques déposées mentionnées dans cette documentation sont la propriété de leur détenteur respectif.

Imprimé aux États-Unis d'Amérique, octobre 2000.

VERITAS Software Corporation
1600 Plymouth St.
Mountain View, CA 94043
États-Unis d'Amérique

Téléphone +1 650-335-8000
Télécopie +1 650-335-8050
www.veritas.com

Table des matières

Préface	xiii
Utilisateurs	xiii
Domaine d'application	xiii
Organisation	xiii
Utilisation de ce guide	xiv
Documents connexes	xiv
Conventions	xv
Chapitre 1. Présentation de Volume Manager	17
Introduction	17
Méthodes de stockage des données	18
Présentation de Volume Manager	18
Objets physiques	19
Disques physiques et nom des disques	19
Partitions	19
Volumes et objets virtuels	20
Disques Volume Manager	21
Disk groups	21
Sous-disques	22
Plex	23
Volumes	24
Connexion entre des objets virtuels Volume Manager	25
Structures de volume	26
Concaténation	28

Répartition des données par bandes sur plusieurs disques (RAID 0)	30
RAID 5	34
Mise en miroir (RAID 1)	35
Mise en miroir avec répartition des données par bandes sur plusieurs disques (RAID 1 + RAID 0)	36
Répartition des données par bandes sur plusieurs disques avec mise en miroir (RAID 0 + RAID 1)	36
Volume Manager et RAID 5	37
Ensembles de disques RAID 5 classiques	37
Ensembles de disques RAID 5 Volume Manager	38
Journalisation	41
Volumes répartis en couches	42
Volume Manager et le système d'exploitation	43
Structures Volume Manager	44
Interfaces utilisateur Volume Manager	45
Présentation de l'interface utilisateur	45
Pourquoi utiliser Volume Manager	46
Chapitre 2. Mise en route	49
Introduction	49
Initialisation de Volume Manager	49
Démons Volume Manager	51
Démon de configuration vxconfigd	51
Démon d'E/S de volume vxiod	52
Configuration du système	52
Exemple de séquence de configuration du système	53
Recommandations pour la configuration du système	53
Recommandations pour la reconstruction automatique	53
Recommandations pour la répartition des données par bandes sur plusieurs disques	55
Recommandations pour la mise en miroir	56

Recommandations pour le journal des zones modifiées (DRL)	57
Recommandations pour la mise en miroir et la répartition des données par bandes sur plusieurs disques	58
Recommandations pour la répartition des données par bandes sur plusieurs disques et la mise en miroir	58
Recommandations pour RAID 5	59
Protection du système	59
Chapitre 3. Opérations Volume Manager	63
Introduction	63
Restructuration en ligne	63
Structure de stockage	64
Principes de fonctionnement de la restructuration en ligne	64
Types de restructuration	65
Caractéristiques de la restructuration	68
Restructurations et longueur de volume	68
Reconstruction automatique	69
Principes de fonctionnement de la reconstruction automatique	69
Sélection de l'espace pour la reconstruction	70
Utilitaire de reconstruction dans le disque d'origine	71
Resynchronisation des volumes	72
Journal des zones modifiées (DRL)	73
Resynchronisation rapide des miroirs	74
Composants de resynchronisation rapide des miroirs	75
Améliorations apportées par la resynchronisation rapide des miroirs à la fonction de prise de vue VxVM	76
Contrôle racine de Volume Manager	80
Démarrage avec des volumes racine	81
Limitations des volumes au démarrage	81
Fonction DMP	82
Mécanisme de basculement de chemin	83

Équilibrage de charge	84
Démarrage à partir des périphériques DMP	84
Activation et désactivation des contrôleurs	84
Affichage des informations de la base de données DMP	85
VxSmartSync Recovery Accelerator	85
Configuration des volumes de données	86
Configuration des volumes de journal redo	86
Moniteur de tâches Volume Manager	87
Fonction de cluster Volume Manager	87
Chapitre 4. Opérations de disque	89
Introduction	89
Périphériques de disque standard	90
Disk groups	91
Commandes de disque et de disk group	92
Initialisation et ajout de disques	93
Ajout d'un disque à Volume Manager	96
Ajout d'un disque à un disk group	99
Suppression d'un disque d'un disk group	102
Déplacement des disques	103
Attribution d'un nouveau nom à un disque	104
Réservation des disques	104
Mise hors ligne d'un disque	105
Mise en miroir d'un disque	105
Suppression d'un disque	107
Affichage des informations sur le disque	108
Détection et remplacement des disques qui ont échoué	110
Reconstruction automatique	110
Utilisation de vxunrelocate	115
Détection des disques qui ont échoué	119

Remplacement des disques	121
Création d'un disk group	123
Mise à niveau d'un disk group	125
Suppression d'un disk group	128
Déplacement des disk groups d'un système à un autre	128
Destruction d'un disk group	133
Utilisation de périphériques spéciaux	133
Utilisation de vxdisk pour des encapsulations spéciales	133
Utilisation de vxdisk pour les disques RAM	134
Utilisation de vxdisk pour afficher des informations de chemins d'accès multiples	135
Opérations de l'interface de menu vxdiskadm	136
Démarrage de vxdiskadm	137
Menu principal vxdiskadm	137
Placement des disques physiques sous le contrôle de Volume Manager	142
Ajout d'un disque pour une utilisation future	155
Réinitialisation d'un disque pour être utilisé par Volume Manager	155
Ajout d'un disque VM au pool de reconstruction automatique	158
Suppression d'un disque VM du pool de reconstruction automatique	160
Nouveau déplacement des sous-disques reconstruits automatiquement vers un disque	161
Exclusion d'un disque de la reconstruction automatique	162
Libération d'un disque pour la reconstruction automatique	162
Affichage d'informations sur les disques physiques	163
Suppression d'un disque physique	164
Désactivation d'un disque physique (mise hors ligne)	165
Activation d'un disque physique	166
Remplacement d'un disque physique	167
Remplacement d'un disque qui a échoué ou qui a été supprimé	168
Ajout d'un disque à un disk group	170
Création d'un disk group	170

Exportation d'un disk group	173
Importation d'un disk group	174
Fermeture de vxdiskadm	175
Chapitre 5. Opérations de volume	177
Introduction	177
Opérations de volume, de plex et de sous-disque, et sauvegarde en ligne	177
Création de volumes	179
Création d'un volume concaténé	181
Création d'un volume réparti par bandes	182
Création d'un volume RAID 5	183
Création d'un volume mis en miroir	183
Redimensionnement d'un volume	183
Redimensionnement des volumes au moyen de vxassist	184
Redimensionnement des volumes au moyen de vxresize	185
Suppression d'un volume	185
Mise en miroir d'un volume	186
Création d'un volume mis en miroir	186
Mise en miroir d'un volume existant	186
Création d'un volume avec activation du journal des zones modifiées	186
Mise en miroir de tous les volumes	187
Suppression d'un miroir	187
Ajout d'un journal RAID 5	188
Ajout d'un journal des zones modifiées	188
Suppression d'un journal des zones modifiées	188
Suppression d'un journal RAID 5	189
Arrêt d'un volume	189
Démarrage d'un volume	189
Liste des volumes impossibles à démarrer	190
Modification de la stratégie de lecture du volume	190

Préparation d'un volume pour la restauration à partir d'une sauvegarde	191
Récupération d'un volume	192
Affichage d'informations sur un volume	192
Opérations de plex	194
Création d'un plex	194
Association d'un plex	195
Dissociation et suppression d'un plex	196
Sauvegarde à l'aide de la mise en miroir	197
Affichage d'informations sur les plex	198
Modification des attributs d'un plex	198
Modification de l'état d'un plex : détachement et attachement d'un plex	199
Déplacement d'un plex	201
Copie d'un plex	202
Opérations de sous-disque	202
Création d'un sous-disque	202
Suppression d'un sous-disque	203
Affichage d'informations sur les sous-disques	203
Association d'un sous-disque	204
Dissociation d'un sous-disque	206
Modification des attributs d'un sous-disque	206
Déplacement d'un sous-disque	207
Division d'un sous-disque	207
Jonction de sous-disques	208
Exécution d'une sauvegarde en ligne	208
Resynchronisation rapide des miroirs	210
Mise en miroir d'un volume sur un disque VM	212
Déplacement d'un volume à partir d'un disque VM	213

Chapitre 6. Fonction de cluster de Volume Manager	217
Introduction	217
Présentation de la fonction de cluster	217
Objets Volume Manager partagés	218
Utilisation de la fonction de cluster de Volume Manager	219
Disques dans les clusters VxVM	225
Stratégies de détachement des disques	226
Activation des disk groups	226
Journal des zones modifiées et environnements cluster	227
Format et taille du journal	228
Compatibilité	228
Fonction DRL dans un environnement cluster	229
Mise à niveau de la fonction de cluster de Volume Manager	229
Utilitaires et démons Volume Manager relatifs au cluster	230
vxclust	231
vxclustadm	232
vxconfigd	232
Récupération vxconfigd	233
vxdg	234
vxdisk	236
vxrecover	237
vxdctl	238
vxstat	239
Gestion des clusters et resynchronisation rapide des miroirs	240
Chapitre 7. Récupération	243
Introduction	243
Procédure de démarrage UNIX	244
Démarrage après une défaillance	245
Configurations root (/), swap et usr possibles	245

Réparation des systèmes de fichiers root (/) ou /usr sur les volumes	246
Sauvegarde et restauration du système de fichiers racine	249
Défaillances et procédures de récupération	249
Défaillances dans le partitionnement UNIX	250
Défaillances lors de l'accès au périphérique de démarrage	250
Défaillances dues à des entrées incorrectes dans /etc/vfstab	251
Défaillances dues à un fichier /etc/system manquant ou endommagé	252
Défaillances dues à un démarrage à partir de plex inutilisables ou stale	254
Reconstruction automatique et défaillances du disque d'initialisation	256
Rajout et remplacement de disques d'initialisation	256
Rajout d'un disque d'initialisation qui a échoué	257
Rétablissement des sous-disques encapsulés reconstruits sur un nouveau disque	258
Remplacement d'un disque d'initialisation qui a échoué	259
Rattachement des disques	260
Récupération de réinstallation	260
Informations générales sur la réinstallation	261
Procédures de réinstallation et de reconfiguration	262
États des plex et des volumes	272
États des plex	272
Cycle des états des plex	275
État du noyau du plex	275
États des volumes	275
États des volumes RAID 5	276
État du noyau du volume	277
Récupération des volumes RAID 5	277
Structure des volumes RAID 5	277
Création de volumes RAID 5	280
Initialisation des volumes RAID 5	280
Défaillances et volumes RAID 5	281

Récupération RAID 5	284
Opérations RAID 5 diverses	287
Manipulation des journaux RAID 5	287
Manipulation des sous-disques RAID 5	288
Démarrage des volumes RAID 5	289
Volumes RAID 5 impossibles à démarrer	290
Démarrage forcé de volumes RAID 5	291
Récupération au démarrage des volumes RAID 5	292
Modification des attributs d'un volume RAID 5	293
Écriture dans les ensembles de disques RAID 5	294
Chapitre 8. Contrôle des performances de VxVM	301
Introduction	301
Recommandations pour les performances	301
Attribution des données	302
Répartition des données par bandes sur plusieurs disques	302
Mise en miroir	303
Mise en miroir et répartition des données par bandes sur plusieurs disques ...	305
Répartition des données par bandes sur plusieurs disques et mise en miroir ...	305
RAID 5	306
Contrôle des performances	306
Types de performances	306
Obtention des données de performances	307
Utilisation des données de performances	308
Réglage de Volume Manager	312
Recommandations générales pour le réglage	312
Éléments réglables	312
Réglage pour les gros systèmes	318
Glossaire	321

Préface

Le *Guide de l'administrateur de VERITAS Volume Manager™* contient des informations sur l'utilisation de Volume Manager.

Utilisateurs

Ce guide est destiné aux administrateurs système qui doivent assurer l'installation, la configuration et la maintenance des systèmes sous le contrôle de VERITAS Volume Manager.

Ce guide suppose que l'utilisateur :

- ◆ est familiarisé avec le système d'exploitation UNIX ;
- ◆ comprend les principes de base d'administration d'un système ;
- ◆ comprend les principes de base de la gestion des volumes.

Domaine d'application

Ce guide a pour but de fournir à l'administrateur système une connaissance détaillée des procédures et des concepts inhérents à la gestion des volumes et à l'administration des systèmes au moyen de Volume Manager. Il inclut des recommandations sur la façon d'utiliser les diverses fonctions de Volume Manager, des instructions sur l'utilisation des commandes Volume Manager pour créer et manipuler des objets, et des informations sur la récupération après une défaillance de disque.

Organisation

Ce guide est composé des chapitres suivants :

- ◆ Présentation de Volume Manager
- ◆ Mise en route
- ◆ Opérations Volume Manager

Utilisation de ce guide

- ◆ Opérations de disque
- ◆ Opérations de volume
- ◆ Fonction de cluster de Volume Manager
- ◆ Récupération
- ◆ Contrôle des performances de VxVM

Utilisation de ce guide

Ce guide contient des instructions qui décrivent l'exécution des tâches d'administration système Volume Manager dans une ou plusieurs des interfaces suivantes :

- ◆ un ensemble de commandes complexes ;
- ◆ une seule commande automatique (`vxassist`) ;
- ◆ une interface de menus (`vxdiskadm`) ;
- ◆ Storage Administrator (interface utilisateur graphique).

Ce guide décrit l'utilisation des diverses interfaces de lignes de commande Volume Manager pour l'administration Volume Manager. Pour plus d'informations sur l'utilisation de l'interface utilisateur graphique de Storage Administrator, voir le *Guide de l'administrateur de VERITAS Volume Manager Storage Administrator*. Les pages man de Volume Manager contiennent une description détaillée des utilitaires Volume Manager et de leurs options ainsi que des détails sur leur utilisation.

Remarque Les commandes Volume Manager nécessitent normalement des privilèges de super-utilisateur ou d'autres privilèges appropriés.

Documents connexes

Les documents suivants fournissent des informations complémentaires sur Volume Manager :

- ◆ *Guide d'installation de VERITAS Volume Manager*
- ◆ *Notes de version de VERITAS Volume Manager*
- ◆ *Notes sur le matériel de VERITAS Volume Manager*
- ◆ *Guide de référence de VERITAS Volume Manager*

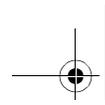
- ◆ *Guide de l'administrateur de VERITAS Volume Manager Storage Administrator*
- ◆ Pages man de VERITAS Volume Manager

Conventions

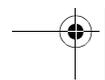
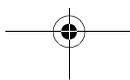
Le tableau suivant décrit les conventions typographiques utilisées dans ce guide.

Police	Utilisation	Exemples
non proportionnelle	Sorties d'ordinateur, fichiers, répertoires, éléments logiciels tels que des options de commande, des noms de fonction et des paramètres	Voir les paramètres dans le fichier <code>/etc/vx/tunefstab</code> . Voir la page man <code>ls(1)</code> pour plus d'informations.
non proportionnelle (caractères gras)	Entrée de l'utilisateur	# mount -F vxfs /h/filesys
<i>italique</i>	Nouveaux termes, titres de livre, accentuations, variables remplacées par un nom ou une valeur	Voir le <i>Guide de l'utilisateur</i> pour plus d'informations. La variable <i>ncsize</i> détermine la valeur de...

Symbole	Utilisation	Exemples
%	Invite de Shell C	
\$	Invite de Shell Bourne/Korn	
#	Invite de super-utilisateur (tous les Shells)	
\	Suite de l'entrée sur la ligne suivante ; ne tapez pas ce caractère.	# mount -F vxfs \ /h/filesys
[]	Dans un synopsis de commande, les crochets indiquent un argument facultatif.	<code>ls [-a]</code>
	Dans un synopsis de commande, une barre verticale sépare des arguments qui s'excluent mutuellement.	<code>mount [suid nosuid]</code>



Conventions



Présentation de Volume Manager

1

Introduction

Ce guide propose une vue d'ensemble de VERITAS Volume Manager (VxVM[®]). Il décrit Volume Manager, son fonctionnement, explique comment utiliser ses interfaces utilisateur et présente ses concepts. Les documents connexes contenant des informations spécifiques sur Volume Manager sont indiqués dans la Préface.

VERITAS Volume Manager facilite la gestion en ligne du stockage sur disque dans les environnements informatiques. La gestion de stockage sur disque nécessite souvent que les systèmes soient mis hors ligne, ce qui représente un inconvénient majeur pour les utilisateurs. Dans l'environnement client/serveur distribué, les bases de données et les autres ressources doivent pouvoir être disponibles à tout moment, facilement accessibles et protégées des dommages éventuels dus à un problème matériel.

VERITAS Volume Manager propose des outils qui permettent d'améliorer les performances et d'assurer la disponibilité et l'intégrité des données. Volume Manager permet également de *configurer dynamiquement le stockage sur disque sans arrêter le système*.

Ce chapitre présente les concepts VERITAS Volume Manager et décrit les outils utilisés par Volume Manager pour gérer le stockage sur disque.

Ce chapitre traite des sujets suivants :

- ◆ Méthodes de stockage des données
- ◆ Présentation de Volume Manager
- ◆ Objets physiques
- ◆ Volumes et objets virtuels
- ◆ Structures de volume
- ◆ Volume Manager et RAID 5
- ◆ Volumes répartis en couches
- ◆ Interfaces utilisateur Volume Manager

Méthodes de stockage des données

Il existe différentes méthodes pour stocker les données sur les disques physiques. Ces méthodes consistent à organiser les données sur le disque afin de les stocker et de les récupérer de façon efficace. On appelle *formatage* la méthode d'organisation de base des disques. Le formatage prépare le disque dur de façon à pouvoir y écrire des fichiers ou à les extraire en utilisant un modèle de gestion de stockage prédéfini.

Il existe deux méthodes de formatage des disques durs et de stockage des informations : une structure de stockage physique et une de stockage logique. Volume Manager utilise la méthode de *stockage logique*. Ce chapitre présente les différents types de structure de stockage pris en charge par Volume Manager.

Présentation de Volume Manager

Volume Manager gère le stockage à l'aide de deux types d'objets : les *objets physiques* et les *objets virtuels*.

- ◆ Objets physiques

Volume Manager stocke les données sur des disques physiques.

- ◆ Objets virtuels

Volume Manager crée des objets virtuels (les *volumes*) sur des disques physiques. Chaque volume enregistre et extrait les données contenues sur un ou plusieurs disques physiques. L'accès aux volumes se fait via les systèmes de fichiers, les bases de données ou d'autres applications de la même manière que l'accès aux disques physiques. Les volumes se composent également d'autres objets virtuels qui permettent de modifier leur configuration. Les volumes et leurs composants virtuels sont appelés *objets virtuels* ou objets Volume Manager. Vous pouvez utiliser les objets Volume Manager pour exécuter des tâches d'administration sur des disques sans interrompre les applications ni les utilisateurs.

Objets physiques

Cette section décrit les objets physiques (disques physiques) utilisés par Volume Manager.

Disques physiques et nom des disques

Un *disque physique* constitue le périphérique de stockage de base (support) sur lequel les données sont finalement stockées. Vous pouvez accéder aux données d'un disque physique en utilisant un *nom de périphérique* afin de localiser ce disque. Le nom de périphérique du disque physique varie en fonction du système informatique utilisé. Tous les paramètres ne sont pas toujours utilisés sur tous les systèmes. En général, un nom de périphérique contient les paramètres suivants : $c\#t\#d\#$, où :

- ◆ $c\#$ représente le contrôleur ;
- ◆ $t\#$ représente l'ID cible ;
- ◆ $d\#$ représente le numéro de disque.

La figure 1, « Exemple de disque physique », illustre la façon dont un disque physique et un nom de périphérique (*devname*) sont représentés dans ce document. Par exemple, le périphérique $c0t0d0$ correspond au contrôleur numéro 0 sur le système, qui a pour ID cible 0 et pour numéro de disque physique 0.

Figure 1. Exemple de disque physique

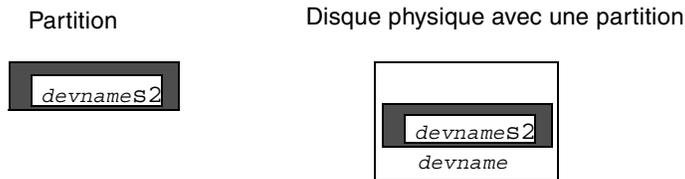


devname

Partitions

Sur certains systèmes informatiques, un disque physique peut être divisé en une ou plusieurs *partitions*. Le *numéro de partition*, ou $s\#$, est ajouté à la fin du nom de périphérique (*devname*). Une partition peut correspondre à un disque physique entier, comme illustré à la figure 2, « Exemple de partition ».

Figure 2. Exemple de partition



Volumes et objets virtuels

La connexion entre des objets physiques et des objets Volume Manager s'effectue lorsque vous placez un disque physique sous le contrôle de Volume Manager.

Volume Manager crée des objets virtuels (ou objets Volume Manager) et établit des connexions logiques entre ces objets. Volume Manager utilise ensuite ces objets virtuels pour exécuter des tâches de gestion de stockage. Les objets Volume Manager sont les suivants :

- ◆ disk groups ;
- ◆ disques VM ;
- ◆ volumes ;
- ◆ plex (miroirs) ;
- ◆ sous-disques.

Ces objets sont décrits dans les sections suivantes.

Un volume est un périphérique de disque virtuel considéré par les applications, les bases de données et les systèmes de fichiers comme un disque physique. Cependant, un volume n'est pas soumis aux mêmes limitations qu'un disque physique. Lorsque vous utilisez Volume Manager, les applications accèdent aux volumes créés sur les disques Volume Manager (disques VM) et non aux disques physiques.

Les volumes contiennent d'autres objets virtuels qui vous permettent de manipuler les données de ces volumes. Les objets virtuels contenus dans les volumes sont les sous-disques et les plex. Pour plus d'informations sur les objets virtuels, reportez-vous aux sections suivantes. La combinaison d'objets virtuels et leur manipulation par les volumes sont décrites dans la section « Volumes » à la page 24.

Disques Volume Manager

Lorsque vous placez un disque physique sous le contrôle de Volume Manager, un disque Volume Manager (ou disque VM) est attribué à ce disque physique. Un disque VM est placé sous le contrôle de Volume Manager et il appartient généralement à un disk group. Chaque disque VM correspond à au moins un disque physique.

Un disque VM comprend généralement une *zone publique* (stockage alloué) et une *zone privée* dans laquelle sont stockées les informations de configuration interne de Volume Manager.

Chaque disque VM possède un *nom d'accès logique au disque* unique (nom de disque virtuel). Vous pouvez spécifier le nom du disque ou laisser Volume Manager lui attribuer un nom par défaut de type `disk##`. La figure 3, « Exemple de disque VM », représente un disque VM dont le nom logique est `disk01` et qui est attribué au disque physique `devname`.

Figure 3. Exemple de disque VM



Disk groups

Un *disk group* est un ensemble de disques VM qui partagent une configuration commune. Une configuration de disk group est un ensemble d'enregistrements contenant des informations détaillées sur d'autres objets Volume Manager, leurs attributs et leurs connexions. Le disk group par défaut est `rootdg` (le disk group racine).

Le cas échéant, vous pouvez créer des disk groups supplémentaires. Ceux-ci vous permettent de regrouper les disques dans des ensembles logiques. Vous pouvez déplacer un disk group et ses composants sous la forme d'une unité, d'un ordinateur hôte à un autre.

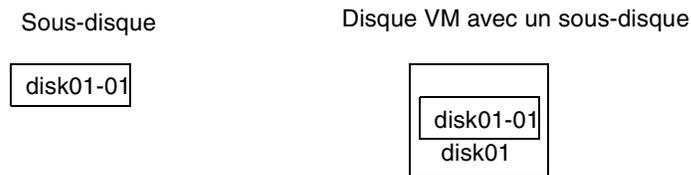
Les volumes sont créés dans un disk group. Un volume spécifique doit être configuré à partir des disques du même disk group.

Sous-disques

Un *sous-disque* est un ensemble de blocs de disque contigus. Un bloc est une unité d'espace sur le disque. Volume Manager alloue de l'espace disque au moyen des sous-disques. Un disque VM peut être divisé en un ou plusieurs sous-disques. Chaque sous-disque représente une partie spécifique d'un disque VM, qui est mappée sur une zone donnée d'un disque physique.

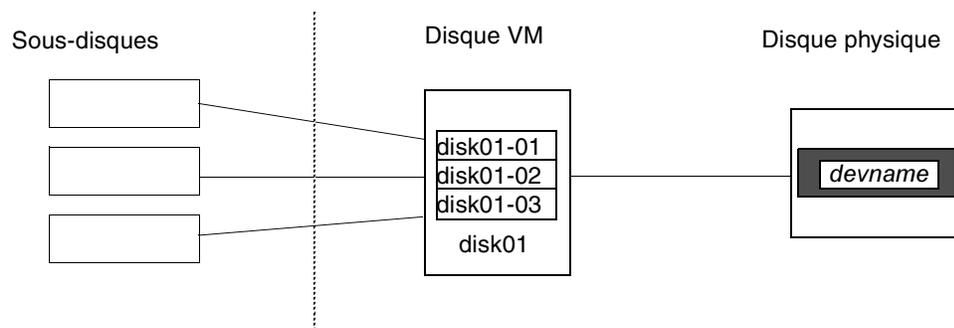
Le nom par défaut d'un disque VM est du type `disk##` (par exemple, `disk01`) et celui d'un sous-disque est `disk##-##`. Dans la figure 4, « Exemple de sous-disque », `disk01-01` représente le nom du premier sous-disque sur le disque VM `disk01`.

Figure 4. Exemple de sous-disque



Un disque VM peut contenir plusieurs sous-disques, mais ceux-ci ne peuvent pas se recouvrir ni partager les mêmes parties d'un disque VM. La figure 5, « Exemple de trois sous-disques attribués à un disque VM », représente un disque VM comprenant trois sous-disques. Le disque VM est attribué à un disque physique.

Figure 5. Exemple de trois sous-disques attribués à un disque VM



Tout espace d'un disque VM qui n'est pas attribué à un sous-disque est considéré comme de l'espace libre. Vous pouvez utiliser cet espace libre pour créer de nouveaux sous-disques.

Volume Manager version 3.0 ou ultérieure permet aux sous-disques de contenir des volumes. Dans les versions antérieures de Volume Manager, les sous-disques ne pouvaient pas contenir de volumes. Pour plus d'informations, voir « Volumes répartis en couches » à la page 42.

Plex

Volume Manager utilise les sous-disques pour créer des objets virtuels appelés *plex* (ou miroirs). Un plex comprend un ou plusieurs sous-disques situés sur un ou plusieurs disques physiques. Pour organiser les données sur les sous-disques afin de former un plex, procédez selon une des méthodes suivantes :

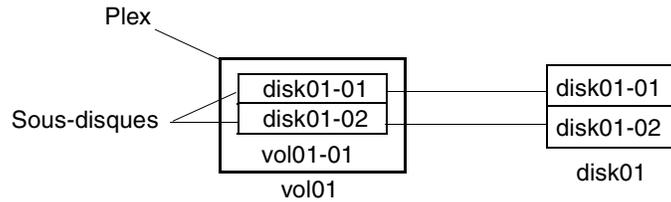
- ◆ concaténation ;
- ◆ répartition des données par bandes sur plusieurs disques (stripping, RAID 0) ;
- ◆ répartition des données par bandes sur plusieurs disques avec parité (RAID 5) ;
- ◆ mise en miroir (RAID 1).

Remarque Dans un ensemble de disques RAID (*Redundant Array of Independent Disks*), une partie de l'espace de stockage combiné sert à stocker des informations dupliquées sur les données de cet ensemble de disques ; vous pouvez ainsi régénérer ces données en cas de défaillance de disque.

La figure 6, « Exemple de plex avec deux sous-disques », représente un plex avec deux sous-disques.

La concaténation, la répartition des données par bandes sur plusieurs disques (RAID 0), la structure RAID 5 et la mise en miroir (RAID 1) sont décrites dans la section « Structures de volume » à la page 26.

Figure 6. Exemple de plex avec deux sous-disques



Volumes

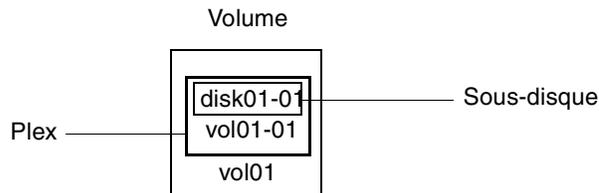
Un volume comprend un ou plusieurs plex, contenant chacun une copie des données du volume. En raison de sa nature virtuelle, un volume n'est pas limité à un disque particulier ou à une zone spécifique d'un disque. Vous pouvez modifier la configuration d'un volume à l'aide des interfaces utilisateur de Volume Manager sans perturber l'exécution des applications ou des systèmes de fichiers qui utilisent ce volume. Par exemple, vous pouvez mettre en miroir un volume sur des disques séparés ou vous pouvez le déplacer pour utiliser un stockage sur disque différent.

Un volume peut être composé de 32 plex maximum, contenant chacun un ou plusieurs sous-disques. Il doit y avoir au moins un plex associé comprenant une copie complète de ses données et au moins un sous-disque associé. Tous les sous-disques d'un volume doivent appartenir au même disk group.

Un volume comprenant au moins deux plex de données est « mis en miroir » et contient les images miroir de ces données. Chaque plex contient une copie identique des données du volume. Pour plus d'informations sur les volumes en miroir, voir « Mise en miroir (RAID 1) » à la page 35.

Volume Manager utilise les conventions de nom par défaut : `vol##` pour les volumes et `vol##-##` pour les plex d'un volume. Veuillez à choisir des noms significatifs pour les volumes. La figure 7, « Exemple de volume comprenant un plex », illustre un volume comprenant un seul plex.

Figure 7. Exemple de volume comprenant un plex

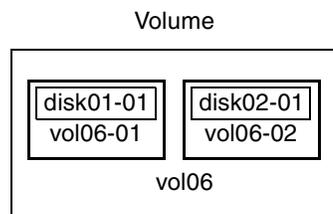


Le volume `vol01` de la figure 7 possède les caractéristiques suivantes :

- ◆ il contient le plex `vol01-01` ;
- ◆ le plex contient le sous-disque `disk01-01` ;
- ◆ le sous-disque `disk01-01` est attribué à partir du disque VM `disk01`.

La figure 8, « Exemple de volume comprenant deux plex », illustre un volume mis en miroir comprenant deux plex.

Figure 8. Exemple de volume comprenant deux plex



Le volume `vol06` de la figure 8 possède les caractéristiques suivantes :

- ◆ il contient les deux plex `vol06-01` et `vol06-02` ;
- ◆ chaque plex contient un sous-disque ;
- ◆ chaque sous-disque est attribué à partir d'un disque VM différent (`disk01` et `disk02`).

Connexion entre des objets virtuels Volume Manager

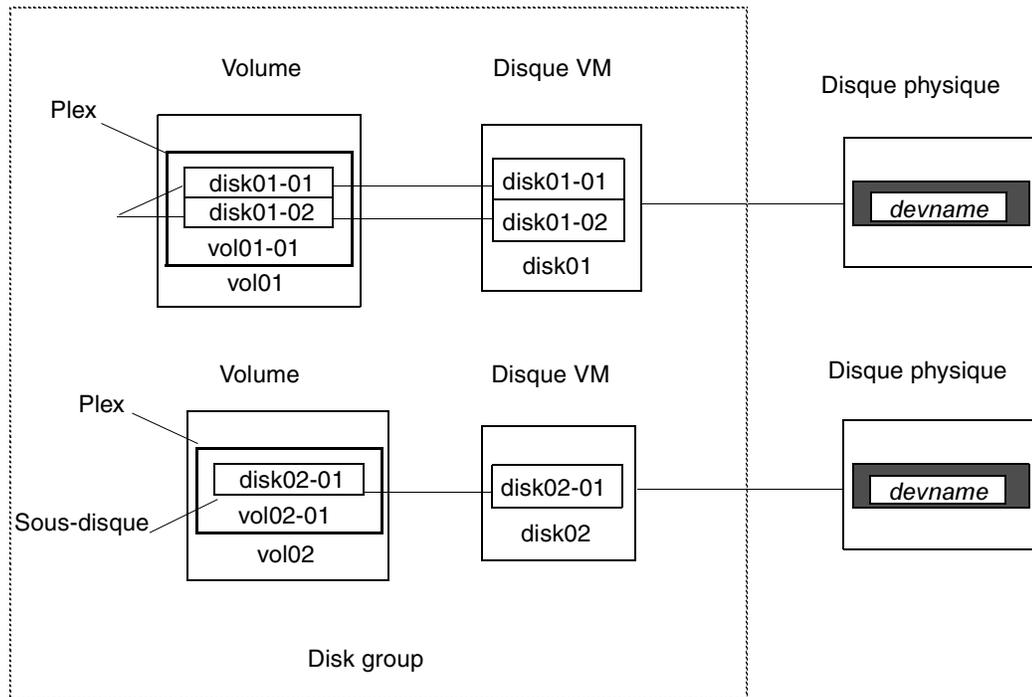
Les objets virtuels Volume Manager sont combinés pour créer des volumes. Il s'agit des objets virtuels suivants : disques VM, disk groups, sous-disques et plex. Les objets Volume Manager possèdent les connexions suivantes :

- ◆ les disques Volume Manager sont regroupés en disk groups ;
- ◆ un ou plusieurs sous-disques (représentant chacun une zone spécifique d'un disque) sont combinés pour former des plex ;
- ◆ un volume se compose d'un ou de plusieurs plex.

La figure 9, « Connexion entre objets Volume Manager », illustre les connexions qui existent entre les objets Volume Manager ainsi que leurs rapports avec les disques physiques. La figure 9 représente un disk group comprenant deux disques VM (`disk01` et `disk02`). `disk01` a un volume comprenant un plex et deux sous-disques. `disk02` a un volume comprenant un plex et un seul sous-disque.

Structures de volume

Figure 9. Connexion entre objets Volume Manager



Structures de volume

Vous pouvez organiser les données dans les objets virtuels pour créer des volumes en appliquant l'une des méthodes de structure suivantes :

- ◆ concaténation ;
- ◆ répartition des données par bandes sur plusieurs disques (RAID 0) ;
- ◆ RAID 5 (répartition des données par bandes sur plusieurs disques avec parité) ;
- ◆ mise en miroir (RAID 1) ;
- ◆ mise en miroir avec répartition des données par bandes sur plusieurs disques ;
- ◆ répartition des données par bandes sur plusieurs disques avec mise en miroir.

Les sections qui suivent décrivent chacune de ces méthodes de structure.

Un périphérique virtuel Volume Manager est défini par un volume. La structure d'un volume est définie par l'association d'un volume à un ou plusieurs plex, qui à leur tour, sont mappés sur des sous-disques. Le volume présente alors aux clients Volume Manager

une interface de périphérique virtuel pour l'accès aux données. Ces blocs de construction logique remapent l'espace d'adressage de volume via lequel les E/S sont redirigées à l'exécution.

Des structures de volume différentes offrent chacune des niveaux de service de stockage différents. Une structure de volume peut être configurée et reconfigurée pour refléter les niveaux voulus de service de stockage.

Dans les versions antérieures de Volume Manager, le sous-disque ne pouvait être mappé directement qu'à un disque VM. Cela lui permettait de définir une zone contiguë d'espace de stockage supportée par la zone publique d'un disque VM. Lorsque le disque VM est actif, il est associé à un disque physique sous-jacent ; c'est ainsi que les objets logiques de Volume Manager sont mappés à des objets physiques et que les données sont conservées dans une zone de stockage stable.

Par conséquent, la combinaison entre une structure de volume et les disques physiques qui assurent l'espace de stockage pour la sauvegarde détermine le service de stockage disponible à parti d'un périphérique virtuel donné.

Dans Volume Manager 3.0 et les versions ultérieures, des « volumes répartis en couches » peuvent être construits en permettant au sous-disque d'être mappé soit à un disque VM comme auparavant, soit à un nouvel objet logique (le *volume de stockage*). Un volume de stockage assure un niveau de mappage récursif avec des structures similaires à celle du volume de niveau supérieur. Finalement, le « bas » du mappage nécessite une association à un disque VM, et donc à un stockage physique attaché.

Les volumes répartis en couches permettent davantage de combinaisons de compositions logiques, dont certaines peuvent se révéler utiles pour configurer un périphérique virtuel. Certaines configurations de volumes répartis en couches « prêtes à l'emploi » ont été incluses dans Volume Manager 3.0 car la libre utilisation de ces volumes à l'aide des commandes rendraient les tâches d'administration beaucoup plus lourdes.

Ces configurations prêtes à l'emploi fonctionnent avec des règles intégrées qui assurent la correspondance automatique des niveaux de service voulus dans le cadre des contraintes spécifiées. La configuration automatique est effectuée « au mieux » afin d'assurer l'appel correct de la commande en cours dans la configuration actuelle.

Pour assurer le service de stockage voulu à partir d'un ensemble de périphériques virtuels, il peut être nécessaire d'inclure un jeu approprié de disques VM dans un disk group et d'exécuter plusieurs commandes de configuration.

Dans la mesure de ses possibilités, Volume Manager 3.0 gère la configuration initiale et la reconfiguration en ligne avec son jeu de structures et son interface d'administration pour faciliter les opérations et accroître l'autonomie des systèmes.

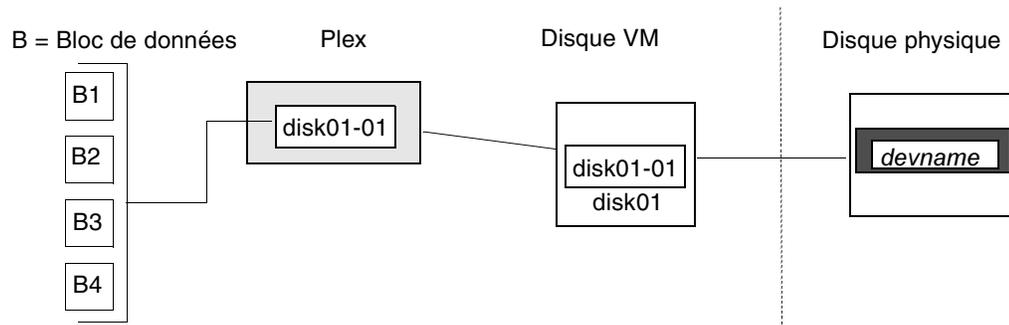
Concaténation

La *concaténation* mappe les données selon une procédure linéaire sur un ou plusieurs disques d'un plex. Pour accéder séquentiellement à toutes les données d'un plex concaténé, il est nécessaire de commencer par le premier sous-disque du début à la fin. L'accès aux données s'effectue ensuite séquentiellement dans les sous-disques restants, du début à la fin, jusqu'à la fin du dernier sous-disque.

Il n'est pas nécessaire que les sous-disques d'un plex concaténé soient physiquement contigus et appartiennent à plusieurs disques VM. On appelle *répartition* la concaténation effectuée au moyen de sous-disques qui se trouvent sur plusieurs disques VM.

La figure 10, « Exemple de concaténation », illustre la concaténation avec un sous-disque.

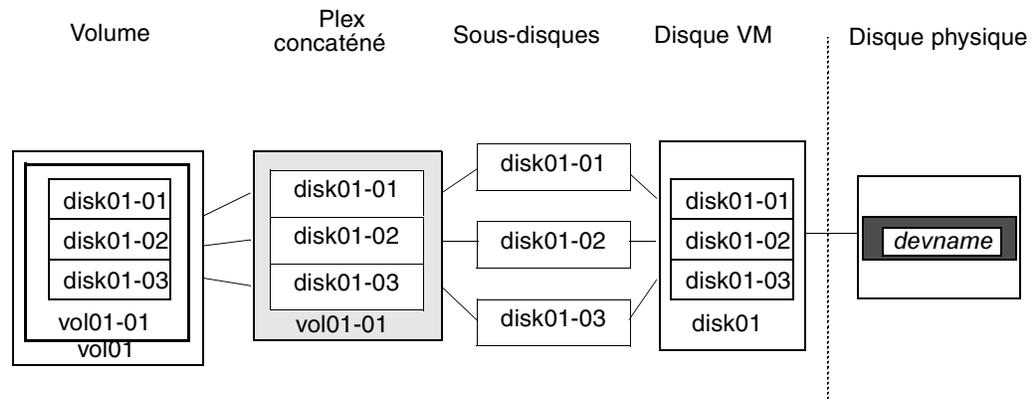
Figure 10. Exemple de concaténation



Vous pouvez utiliser la concaténation avec plusieurs sous-disques lorsque l'espace contigu d'un disque est insuffisant pour le plex. Ce type de concaténation permet d'équilibrer la charge entre les disques et d'optimiser le mouvement des têtes sur un disque spécifique.

La figure 11, « Exemple de volume dans une configuration concaténée », illustre un volume dans une configuration concaténée.

Figure 11. Exemple de volume dans une configuration concaténée

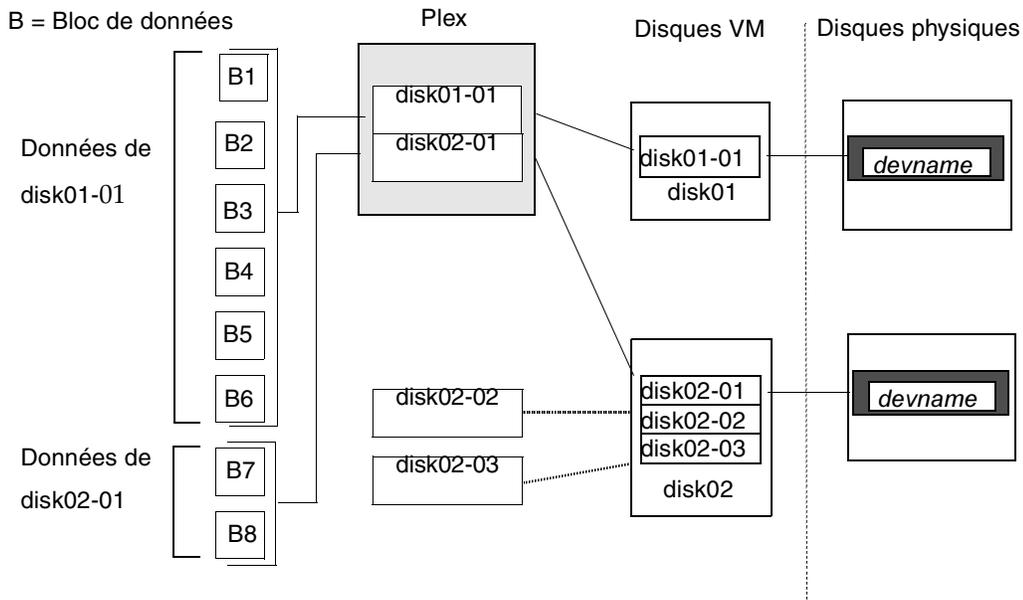


Dans l'exemple de la figure 12, les six premiers blocs de données (B1 à B6) utilisent la plus grande partie de l'espace sur le disque auquel le disque VM `disk01` est attribué. Cette procédure ne requiert que l'espace du sous-disque `disk01-01` sur le disque VM `disk01`. Cependant, les deux derniers blocs de données, B7 et B8, n'utilisent qu'une partie de l'espace sur le disque auquel le disque VM `disk02` est attribué.

Le reste de l'espace libre sur le disque VM `disk02` peut être affecté à d'autres usages. Dans cet exemple, les sous-disques `disk02-02` et `disk02-03` sont disponibles pour d'autres tâches de gestion de disque.

La figure 12, « Exemple de répartition », illustre les données réparties sur deux sous-disques dans un plex réparti.

Figure 12. Exemple de répartition



Attention La répartition d'un plex sur plusieurs disques augmente le risque que la défaillance d'un seul disque entraîne l'échec du volume attribué. Utilisez la mise en miroir ou RAID 5 (décrits dans les sections suivantes) pour réduire ce risque.

Répartition des données par bandes sur plusieurs disques (RAID 0)

La *répartition des données par bandes sur plusieurs disques* (RAID 0) mappe les données de façon à ce qu'elles soient intercalées entre au moins deux disques physiques. Un plex réparti par bandes contient au moins deux sous-disques, qui sont répartis sur deux disques physiques ou plus. Les données sont allouées de façon alternative et régulière aux divers sous-disques d'un plex réparti par bandes.

Les sous-disques sont regroupés en « colonnes », chaque disque physique étant limité à une colonne. Chaque colonne contient un ou plusieurs sous-disques et peut provenir d'un ou de plusieurs disques physiques. Le nombre et la taille des sous-disques par colonne peut varier. Si nécessaire, vous pouvez ajouter des sous-disques supplémentaires aux colonnes.

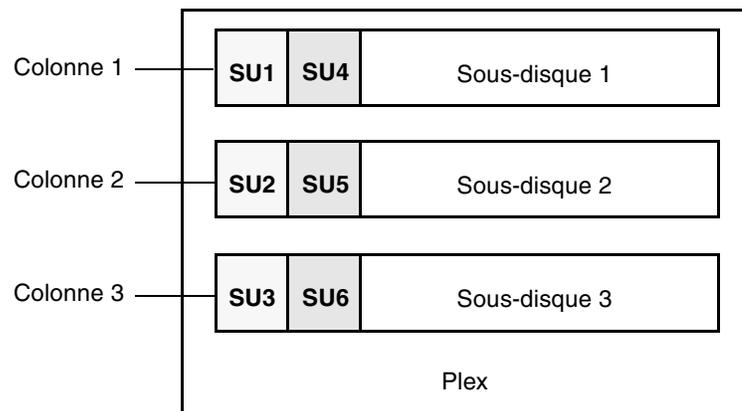
Attention La répartition par bandes ou la division d'un volume entre plusieurs disques augmente le risque que la défaillance d'un disque entraîne l'échec de ce volume. Par exemple, si cinq volumes sont répartis entre cinq mêmes disques, après la défaillance de l'un de ces disques, il faudra restaurer les cinq volumes à partir d'une sauvegarde. Si chaque volume se trouve sur un disque séparé, il ne faudra restaurer qu'un seul volume. La mise en miroir ou RAID 5 permet de réduire considérablement le risque que la défaillance d'un seul disque entraîne l'échec d'un grand nombre de volumes.

Les données sont allouées en unités de taille égale (appelées *unités de bande (stripe unit)*) intercalées entre les colonnes. Chaque unité de bande est un ensemble de blocs contigus sur un disque. La taille de l'unité de bande par défaut est 64 Ko.

Par exemple, si un plex réparti par bandes contient trois colonnes et six unités de bande, les données sont réparties sur trois disques physiques, comme illustré à la figure 13, « Répartition des données par bandes sur trois disques (colonnes) » :

- ◆ les première et quatrième unités de bande sont allouées dans la colonne 1 ;
- ◆ les deuxième et cinquième unités de bande sont allouées dans la colonne 2 ;
- ◆ les troisième et sixième unités de bande sont allouées dans la colonne 3.

Figure 13. Répartition des données par bandes sur trois disques (colonnes)



SU = Unité de bande

Une *bande* se compose de l'ensemble des unités de bande qui occupent les mêmes positions dans toutes les colonnes. Dans la figure 13, les unités de bande 1, 2 et 3 constituent une seule bande.

Structures de volume

Vue en séquence, la première bande se compose de :

- ◆ l'unité de bande 1 dans la colonne 1 ;
- ◆ l'unité de bande 2 dans la colonne 2 ;
- ◆ l'unité de bande 3 dans la colonne 3.

La seconde bande se compose de :

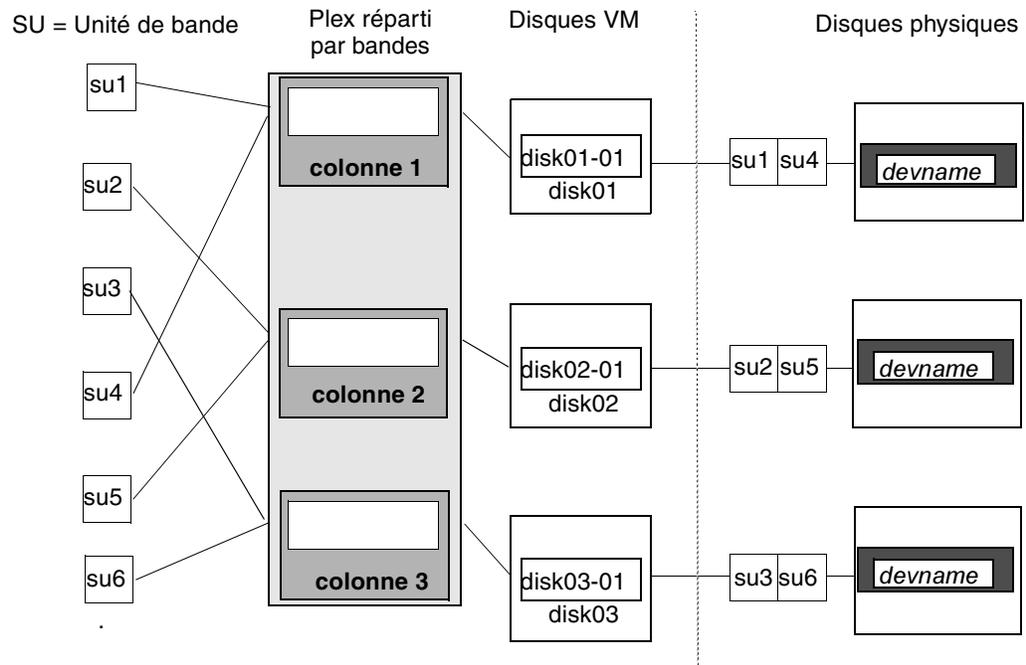
- ◆ l'unité de bande 4 dans la colonne 1 ;
- ◆ l'unité de bande 5 dans la colonne 2 ;
- ◆ l'unité de bande 6 dans la colonne 3.

Les données sont réparties par bandes sur plusieurs disques, sur toute la longueur des colonnes (si elles ont toutes la même longueur) ou jusqu'à la fin de la plus courte. Tout espace restant à la fin des sous-disques dans les colonnes les plus longues devient de l'espace inutilisé.

La répartition des données par bandes sur plusieurs disques est utile si vous devez écrire ou lire rapidement de grandes quantités de données sur les disques physiques en utilisant le transfert parallèle des données sur plusieurs disques. Elle permet également d'équilibrer entre plusieurs disques la charge d'E/S provenant d'applications multi-utilisateurs.

La figure 14, « Exemple d'un plex réparti par bandes comportant un sous-disque par colonne », illustre un plex réparti par bandes comportant trois colonnes de taille égale sur un seul sous-disque. Il existe une colonne par disque physique.

Figure 14. Exemple d'un plex réparti par bandes comportant un sous-disque par colonne

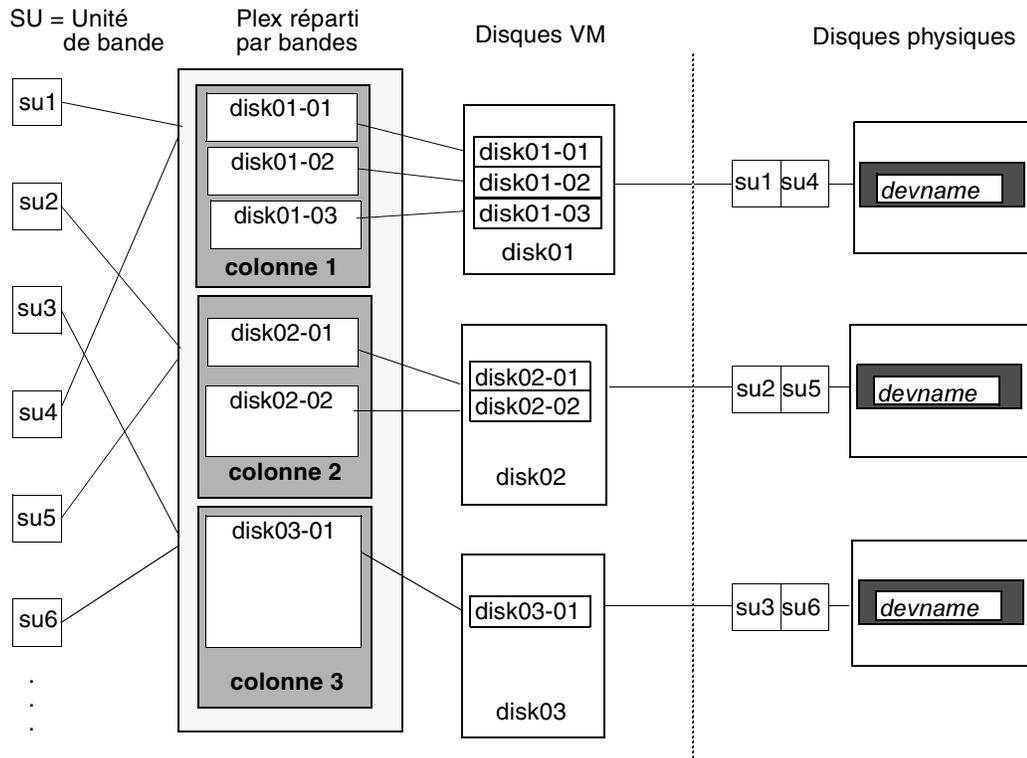


L'exemple de la figure 14 représente trois sous-disques qui occupent entièrement l'espace des disques VM. Il se peut également que chaque sous-disque d'un plex réparti par bandes occupe uniquement une partie du disque VM, ce qui laisse de l'espace libre pour les autres tâches de gestion de disque.

La figure 15, « Exemple d'un plex réparti par bandes avec des sous-disques concaténés par colonne », représente un plex réparti par bandes dont les trois colonnes contiennent des sous-disques de tailles différentes. Chaque colonne contient un nombre différent de sous-disques. Il existe une colonne par disque physique. Pour créer des plex répartis par bandes, vous pouvez utiliser un seul sous-disque provenant de chacun des disques VM en cours de répartition. Vous pouvez aussi allouer de l'espace à partir de différentes zones du même disque ou d'un autre disque (par exemple, si la taille du plex a augmenté). Les colonnes peuvent contenir des sous-disques provenant de disques VM différents.

Structures de volume

Figure 15. Exemple d'un plex réparti par bandes avec des sous-disques concaténés par colonne

**RAID 5**

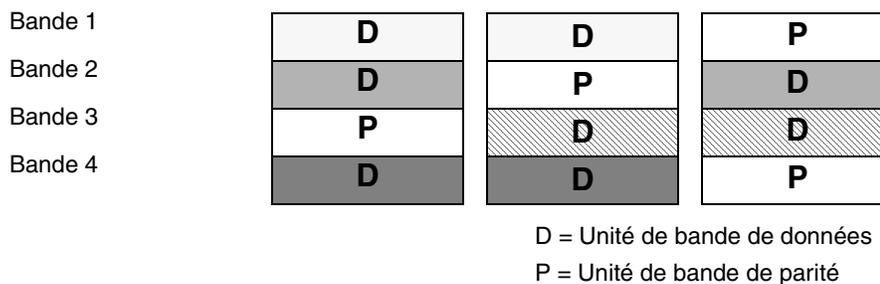
RAID 5 assure la redondance des données à l'aide de la *parité*. La parité est une valeur calculée qui permet de reconstruire les données après une défaillance. Cette valeur est calculée en effectuant une procédure XOR (*OU exclusif*) sur les données pendant qu'elles sont écrites sur un volume RAID 5. La parité ainsi obtenue est alors écrite sur le volume. Si une partie d'un volume RAID 5 échoue, il est possible de recréer les données qu'elle contenait à partir de celles restantes et des informations de parité.

Les volumes RAID 5 assurent la redondance des données dans un volume. Ils conservent une copie des données et de la parité calculée dans un plex « réparti par bandes » sur plusieurs disques. En cas de défaillance d'un disque, un tel volume utilise la parité pour reconstruire les données. La structure peut contenir un mélange de concaténation et de répartition des données par bandes sur plusieurs disques.

Les volumes RAID 5 peuvent utiliser des journaux RAID 5 pour réduire la durée de la récupération. Ces journaux leur permettent de conserver une copie des données et de la parité en cours d'écriture. Les journaux RAID 5 sont facultatifs et peuvent être recréés avec les volumes RAID 5 ou ajoutés par la suite.

La figure 16, « Emplacements de parité dans un modèle RAID 5 », indique les emplacements de parité dans une configuration d'ensemble de disques RAID 5. Chaque bande dispose d'une colonne comprenant une unité de bande de parité et de colonnes contenant des données. La parité est répartie sur tous les disques de l'ensemble de disques, ce qui réduit la durée des écritures indépendantes volumineuses (il n'est ainsi plus nécessaire que les écritures attendent qu'un seul disque de parité accepte les données).

Figure 16. Emplacements de parité dans un modèle RAID 5



Pour plus d'informations, voir « Volume Manager et RAID 5 » à la page 37.

Mise en miroir (RAID 1)

La *mise en miroir* utilise plusieurs miroirs (plex) pour dupliquer les informations contenues dans un volume. En cas de défaillance d'un disque physique, le plex du disque défectueux devient indisponible, mais le système continue à fonctionner à l'aide des miroirs intacts. Même s'il est possible qu'un volume n'ait qu'un seul plex, deux plex au moins sont cependant nécessaires pour assurer la redondance des données. Celle-ci nécessite que l'espace disque de chacun de ces plex provienne de disques différents.

Lors de la répartition des données par bandes ou sur un grand nombre de disques, la défaillance de n'importe lequel de ces disques peut rendre le plex totalement inutilisable. Le risque de défaillance de l'un des disques est suffisant pour envisager la mise en miroir afin d'améliorer la fiabilité (et la disponibilité) d'un volume réparti ou réparti par bandes.

Mise en miroir avec répartition des données par bandes sur plusieurs disques (RAID 1 + RAID 0)

Volume Manager prend en charge la combinaison de la mise en miroir avec la répartition des données par bandes sur plusieurs disques. Cette structure combinée est dite miroir-répartition par bandes (mirrored striped). Lorsque plusieurs structures de ce type sont utilisées sur le même volume, la mise en miroir avec répartition des données par bandes sur plusieurs disques présente les avantages d'une répartition des données sur plusieurs disques (répartition des données par bandes sur plusieurs disques) tout en assurant la redondance (mise en miroir) des données.

Pour que l'utilisation de la mise en miroir avec la répartition des données par bandes sur plusieurs disques soit efficace, il est nécessaire d'allouer le miroir et son plex réparti par bandes à partir de disques séparés. La structure du miroir peut être concaténée ou répartie par bandes.

Répartition des données par bandes sur plusieurs disques avec mise en miroir (RAID 0 + RAID 1)

Volume Manager prend en charge la combinaison de la répartition des données par bandes sur plusieurs disques avec la mise en miroir. Cette structure combinée est dite répartition par bandes-miroir (striped mirrored). Dans les versions antérieures, la mise en miroir devait se produire par dessus la répartition des données par bandes sur plusieurs disques. À présent, elle peut être effectuée au-dessus et au-dessous de la répartition des données par bandes sur plusieurs disques.

Lorsque la mise en miroir est effectuée au-dessous de la répartition des données par bandes sur plusieurs disques, chaque colonne de la bande est mise en miroir. Si la bande est suffisamment large pour comporter plusieurs sous-disques par colonne, chaque sous-disque peut être mis en miroir individuellement. Cette structure permet d'améliorer la redondance et de réduire la durée de récupération en cas d'erreur.

Si un disque échoue dans une structure de type miroir-répartition par bandes, le plex entier est détaché et la redondance est perdue sur tout le volume. Lorsque vous remplacez le disque, vous devez actualiser le plex entier. La récupération du plex entier peut être très longue. Si un disque échoue dans une structure de type répartition par bandes-miroir, seul le sous-disque défectueux doit être détaché et seule cette partie du volume perd la redondance. Lorsque vous remplacez le disque, seule une partie du volume doit être récupérée.

Par comparaison avec la mise en miroir avec répartition des données par bandes sur plusieurs disques, la répartition des données par bandes sur plusieurs disques avec mise en miroir permet au volume de mieux gérer les défaillances de disques. En cas de défaillance de disque, la durée de la récupération est plus courte avec la répartition des données par bandes sur plusieurs disques avec mise en miroir. Pour plus d'informations, voir « Volumes répartis en couches » à la page 42.

Volume Manager et RAID 5

Cette section décrit la façon dont Volume Manager implémente RAID 5.

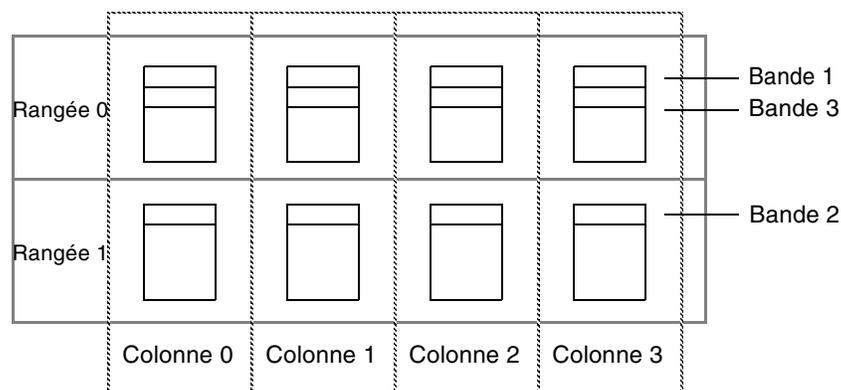
Bien que la mise en miroir (RAID 1) et RAID 5 assurent la redondance des données, ces deux modes reposent sur des méthodes différentes. La mise en miroir assure la redondance des données en conservant plusieurs copies complètes des données dans un volume. Les données en cours d'écriture sur un volume mis en miroir sont reflétées dans toutes les copies. Si une partie d'un volume mis en miroir échoue, le système continue à utiliser les autres copies des données.

RAID 5 assure la redondance des données par *parité*. La parité est une valeur calculée qui permet de reconstruire les données après une défaillance. Cette valeur est calculée en effectuant une procédure XOR (*OU exclusif*) sur les données pendant qu'elles sont écrites sur un volume RAID 5. La parité ainsi obtenue est alors écrite sur le volume. Si une partie d'un volume RAID 5 échoue, il est possible de recréer les données qu'elle contenait à partir de celles restantes et des informations de parité.

Ensembles de disques RAID 5 classiques

Un ensemble de disques RAID 5 *classique* se compose de plusieurs disques organisés en rangées et colonnes. Une *colonne* indique le nombre de disques placés dans la baie selon la même position ordinaire. Une *rangée* indique le nombre minimum de disques nécessaires pour prendre en charge la largeur totale d'une bande de parité. La figure 17, « Ensemble de disques RAID 5 classique », illustre la structure de rangées et de colonnes d'un ensemble de disques RAID 5 classique.

Figure 17. Ensemble de disques RAID 5 classique



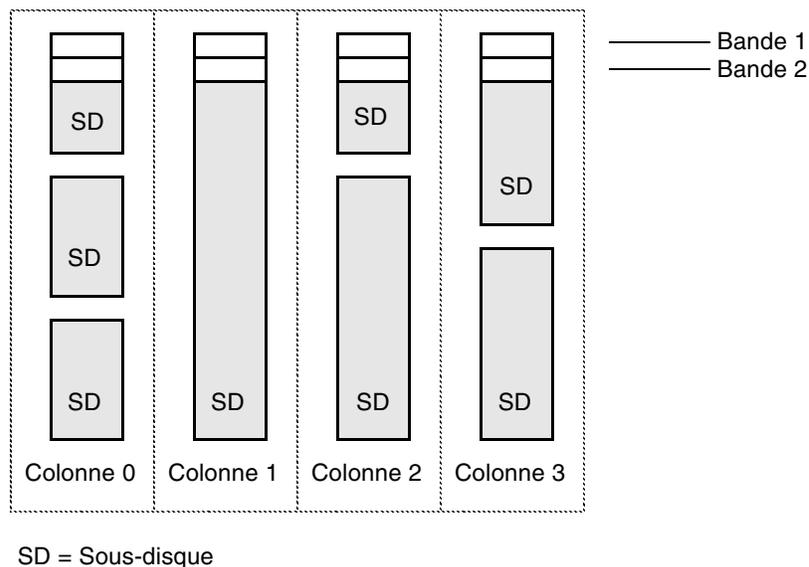
Volume Manager et RAID 5

Cette structure d'ensemble de disques classique offre des possibilités d'expansion en permettant d'ajouter des rangées dans une colonne. La répartition des données par bandes sur plusieurs disques s'effectue en appliquant la première bande sur les disques de la rangée 0, la seconde bande sur les disques de la rangée 1, puis la troisième bande sur les disques de la rangée 0, etc. Ce type d'ensemble de disques nécessite que toutes les colonnes de disque et les rangées aient la même taille.

Ensembles de disques RAID 5 Volume Manager

La structure de l'ensemble de disques RAID 5 Volume Manager diffère de la structure classique. Volume Manager n'utilise pas les rangées à cause de la nature virtuelle de ses disques et d'autres objets. À la place, Volume Manager utilise des colonnes composées de sous-disques de longueur variable (comme illustré à la figure 18, « Ensemble de disques RAID 5 Volume Manager »). Chaque sous-disque représente une zone spécifique d'un disque.

Figure 18. Ensemble de disques RAID 5 Volume Manager



Avec la structure d'ensemble de disques RAID 5 Volume Manager, chaque colonne peut comporter un nombre différent de sous-disques. Les sous-disques d'une colonne donnée peuvent provenir de disques physiques différents. Si nécessaire, vous pouvez ajouter des sous-disques supplémentaires aux colonnes. Les données sont réparties par bandes sur plusieurs disques (voir « Répartition des données par bandes sur plusieurs disques (RAID 0) » à la page 30) en appliquant la première bande sur chaque sous-disque en haut

de chaque colonne, puis une autre bande au-dessous etc., sur toute la longueur des colonnes. Pour chaque bande, une unité de bande de taille égale est placée dans chaque colonne. Dans le mode RAID 5, la taille des unités de bande est par défaut de 16 Ko.

Remarque Actuellement, la mise en miroir des volumes RAID 5 n'est pas prise en charge.

Structure symétrique gauche

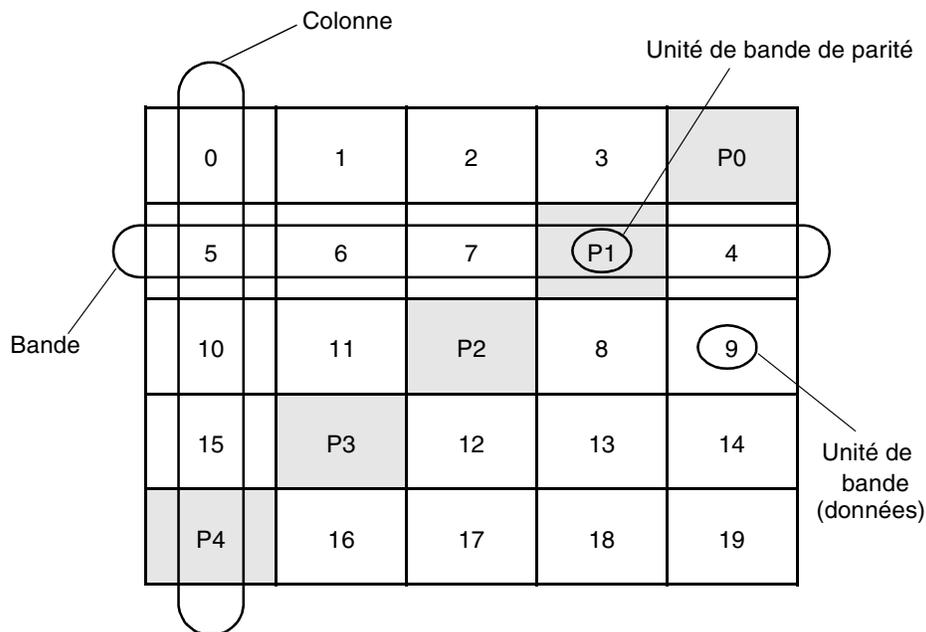
Plusieurs structures de données et de parité peuvent être utilisées pour la configuration d'un ensemble de disques RAID 5. La structure symétrique gauche représente l'implémentation du mode RAID 5 par Volume Manager. La structure de parité symétrique gauche assure des performances optimales à la fois pour les opérations d'E/S aléatoires et pour les opérations d'E/S séquentielles volumineuses. En termes de performances, la sélection de la structure ne revêt pas une importance aussi critique que celle du nombre de colonnes et de la taille d'unité de bande.

La structure symétrique gauche répartit par bandes les données et la parité sur les colonnes, en plaçant la parité dans une colonne différente pour chaque bande de données. La première unité de bande de parité se trouve dans la colonne la plus à droite de la première bande. Chaque unité de bande de parité successive se trouve dans la bande suivante (dans la colonne immédiatement à gauche de l'unité de bande de parité précédente). S'il existe davantage de bandes que de colonnes, le positionnement des unités de bande de parité commence à partir de la colonne la plus à droite.

Volume Manager et RAID 5

La figure 19, « Structure symétrique gauche », illustre une structure de parité symétrique gauche avec cinq disques (un par colonne).

Figure 19. Structure symétrique gauche



Pour chaque bande, les données sont organisées à partir de la droite de l'unité de bande de parité. Dans la figure 19, l'organisation des données pour la première bande commence à P0 et continue avec les unités de bande 0-3. L'organisation des données pour la seconde bande commence à P1 et continue avec l'unité de bande 4, puis les unités de bande 5-7. L'organisation des données se poursuit de la même façon pour les bandes restantes.

Chaque unité de bande de parité contient le résultat d'une procédure XOR (OU exclusif) effectuée sur les données des unités de bandes dans la même bande. Si les données d'un disque correspondant à une colonne sont inaccessibles à cause d'une défaillance matérielle ou logicielle, les données peuvent être restaurées. Pour cela, la procédure XOR doit être appliquée sur le contenu des unités de bande des données des colonnes restantes en fonction de leurs unités de bande de parité respectives (pour chaque bande).

Par exemple, si le disque correspondant à la colonne la plus à gauche dans la figure 19 échoue, le volume est placé en mode détérioré. Dans ce mode, les données provenant de la colonne peuvent être recrées en procédant comme suit : appliquez la procédure XOR sur les unités de bande 1-3 en fonction de l'unité de bande de parité P0 pour recréer l'unité de bande 0, puis appliquez la procédure XOR sur les unités de bande 4, 6 et 7 en fonction de l'unité de bande de parité P1 pour recréer l'unité de bande 5, etc.

Remarque L'échec de plusieurs colonnes dans un plex avec une structure RAID 5 entraîne le détachement du volume. Celui-ci ne peut plus répondre aux requêtes de lecture ou d'écriture. Une fois que les colonnes défectueuses ont été récupérées, vous devrez peut-être récupérer les données des utilisateurs à partir des sauvegardes.

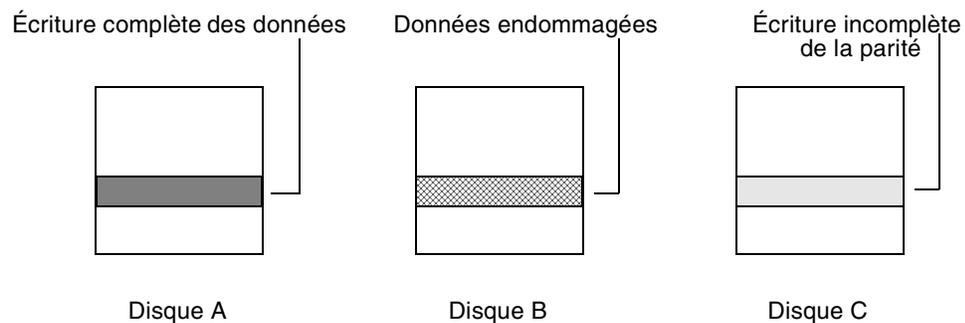
Journalisation

La *journalisation* (logging) permet d'éviter l'endommagement des données de récupération. Le journal des nouvelles données et parité est créé sur un périphérique permanent (par exemple, un volume sur disque ou une RAM permanente). Les nouvelles données et parité sont alors écrites sur les disques.

Sans l'enregistrement dans un journal, les données qui ne sont pas écrites de façon active risquent d'être perdues ou endommagées sans avertissement en cas de défaillance d'un disque, puis du système. Lorsque cette double défaillance se produit, il est impossible de savoir si les données qui étaient en cours d'écriture dans les zones de données des disques ou si la parité qui était en cours d'écriture dans les zones de parité ont réellement été écrites. Par conséquent, la récupération du disque endommagé peut elle-même être endommagée.

Dans la figure 20, « Écriture incomplète », la récupération du disque B dépend de l'écriture complète des données sur le disque A et de la parité sur le disque C. Le diagramme illustre une écriture de données complète et une écriture de parité incomplète provoquant une reconstruction incorrecte des données pour les données du disque B.

Figure 20. Écriture incomplète



Pour éviter une telle défaillance, enregistrez dans le journal toutes les écritures de données avant de les diriger vers l'ensemble de disques. Cela vous permet de relire le journal, et de mettre à jour les données et la parité avant de reconstruire le lecteur qui a échoué.

Volumes répartis en couches

Vous pouvez associer les journaux à un volume RAID 5 en les attachant en tant que plex de journal. Plusieurs plex de journal pour chaque volume RAID 5, auquel cas les zones du journal sont mises en miroir.

Volumes répartis en couches

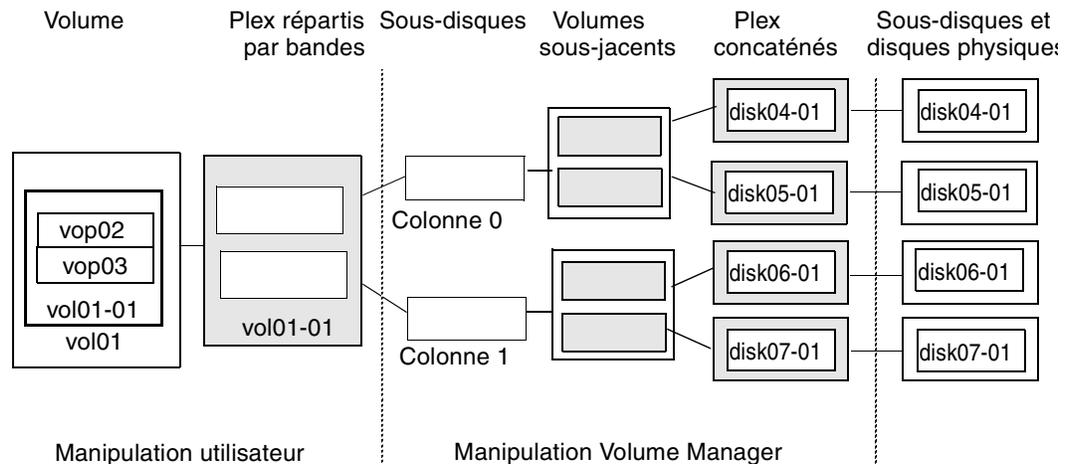
Un volume réparti en couches est un objet virtuel Volume Manager construit sur un ou plusieurs volumes. La structure de volume réparti en couches gère mieux les défaillances et offre une redondance supérieure à celle de la structure de volume standard. Par exemple, dans un volume réparti par bandes et mis en mémoire en couches, chaque miroir (plex) couvre une zone d'espace de stockage plus petite, si bien que la récupération est plus rapide qu'avec un volume mis en miroir standard. La figure 21, « Exemple de volume réparti en couches de type répartition par bandes-miroir », illustre un volume réparti en couches.

Remarque Volume Manager 3.0 et versions ultérieures prend en charge les volumes répartis en couches, mais pas les versions antérieures. Volume Manager 3.0 et versions ultérieures permet de construire les sous-disques sur des volumes (volumes de stockage) alors que les versions antérieures de Volume Manager ne gèrent pas les sous-disques.

Les tâches utilisateur peuvent uniquement être exécutées au niveau supérieur d'un volume réparti en couches. Vous ne pouvez pas détacher un volume réparti en couches ni exécuter d'autres opérations sur les volumes sous-jacents en manipulant la structure interne. Vous pouvez effectuer toutes les opérations nécessaires à partir de la zone de manipulation utilisateur qui inclut le volume et le plex réparti par bandes. Dans la figure 21, le volume et le plex réparti par bandes dans la zone « Manipulation utilisateur » vous permettent d'exécuter des tâches Volume Manager normales.

La zone « Manipulation Volume Manager » de la figure 21 représente des sous-disques à deux colonnes, construits sur des volumes sous-jacents, chacun de ces volumes étant mis en miroir de façon interne. Les volumes répartis en couches constituent une infrastructure dans Volume Manager et permettent d'ajouter certaines fonctions au programme. Les volumes sous-jacents sont utilisés exclusivement par Volume Manager et ils ne sont pas conçus pour être manipulés par l'utilisateur. La structure de volume sous-jacent est décrite dans ce document pour vous permettre de comprendre le fonctionnement des volumes répartis en couche et la raison pour laquelle ils sont utilisés par Volume Manager.

Figure 21. Exemple de volume réparti en couches de type répartition par bandes-miroir



Il se peut que les administrateurs système doivent manipuler la structure de volumes répartis en couches pour effectuer le dépannage ou d'autres opérations (par exemple, pour placer des données sur des disques spécifiques). Volume Manager utilise les volumes répartis en couches pour exécuter les tâches et opérations suivantes :

- ◆ répartition par bandes-miroir (voir la page `man vxassist`) ;
- ◆ miroirs concaténés (voir la page `man vxassist`) ;
- ◆ restructuration en ligne (voir les pages `man vxrelayout` et `man vxassist`) ;
- ◆ déplacement de sous-disques RAID 5 (voir la page `man vxsd`) ;
- ◆ prise de vue RAID 5 (voir la page `man vxassist`).

Volume Manager et le système d'exploitation

Volume Manager fonctionne en tant que sous-système entre le système d'exploitation et les systèmes de gestion des données, tels que les systèmes de fichiers et les systèmes de gestion de base de données.

Pour pouvoir être placé sous le contrôle de Volume Manager, le disque doit être accessible via l'interface du périphérique du système d'exploitation. Volume Manager est un sous-système réparti en couches situé au-dessus des services d'interface du système d'exploitation. Par conséquent, Volume Manager dépend de la façon dont le système d'exploitation accède aux disques physiques.

Volumes répartis en couches

Volume Manager dépend du système d'exploitation pour les composants suivants :

- ◆ les périphériques du système d'exploitation (disques) ;
- ◆ les gestionnaires de périphériques ;
- ◆ les disques VM ;
- ◆ le méta-périphérique de chemins d'accès multiples (DMP, *Dynamic Multipathing*) de Volume Manager.

Structures Volume Manager

Un périphérique virtuel Volume Manager est défini par un volume. La structure d'un volume est définie par l'association d'un volume à un ou plusieurs plex, qui à leur tour, sont mappés sur des sous-disques. Le volume présente alors aux clients Volume Manager une interface de périphérique virtuel pour l'accès aux données. Ces blocs de construction logique remappent l'espace d'adressage de volume via lequel les E/S sont redirigées à l'exécution.

Des structures de volume différentes offrent chacune des niveaux de service de stockage différents. Une structure de volume peut être configurée et reconfigurée pour refléter les niveaux voulus de service de stockage.

Dans les versions antérieures de Volume Manager, le sous-disque ne pouvait être mappé directement qu'à un disque VM. Cela lui permettait de définir une zone contiguë d'espace de stockage supportée par la zone publique d'un disque VM. Lorsque le disque VM est actif, il est associé à un disque physique sous-jacent ; c'est ainsi que les objets logiques de Volume Manager sont mappés à des objets physiques et que les données sont conservées dans une zone de stockage stable.

Par conséquent, la combinaison entre une structure de volume et les disques physiques qui assurent l'espace de stockage pour la sauvegarde détermine le service de stockage disponible à parti d'un périphérique virtuel donné.

Dans Volume Manager 3.0 et versions ultérieures, des « volumes répartis en couches » peuvent être construits en permettant au sous-disque d'être mappé soit à un disque VM comme auparavant, soit à un nouvel objet logique (le *volume de stockage*). Un volume de stockage assure un niveau de mappage récursif avec des structures similaires à celle du volume de niveau supérieur. Finalement, le « bas » du mappage nécessite une association à un disque VM, et donc à un stockage physique attaché.

Les volumes répartis en couches permettent davantage de combinaisons de compositions logiques, dont certaines peuvent se révéler utiles pour configurer un périphérique virtuel. Certaines configurations de volumes répartis en couches « prêtes à l'emploi » ont été incluses dans Volume Manager 3.0 car la libre utilisation de ces volumes à l'aide des commandes rendraient les tâches d'administration beaucoup plus lourdes.

Ces configurations prêtes à l'emploi fonctionnent avec des règles intégrées qui assurent la correspondance automatique des niveaux de service voulus dans le cadre des contraintes spécifiées. La configuration automatique est effectuée « au mieux » afin d'assurer l'appel correct de la commande en cours dans la configuration actuelle.

Pour assurer le service de stockage voulu à partir d'un ensemble de périphériques virtuels, il peut être nécessaire d'inclure un jeu approprié de disques VM dans un disk group et d'exécuter plusieurs commandes de configuration.

Dans la mesure de ses possibilités, Volume Manager 3.0 gère la configuration initiale et la reconfiguration en ligne avec son jeu de structures et son interface d'administration pour faciliter les opérations et accroître l'autonomie des systèmes.

Interfaces utilisateur Volume Manager

Cette section décrit brièvement les interfaces utilisateur de VERITAS Volume Manager.

Présentation de l'interface utilisateur

Volume Manager prend en charge les interfaces utilisateur suivantes :

- ◆ Volume Manager Storage Administrator (VMSA)
- ◆ Storage Administrator est l'interface utilisateur graphique de Volume Manager. Storage Administrator comprend des éléments visuels tels que des icônes, des menus et des boîtes de dialogue qui permettent de manipuler les objets Volume Manager. Storage Administrator sert également d'interface pour certaines opérations courantes de systèmes de fichiers. Pour plus d'informations, reportez-vous au *Guide de l'administrateur de VERITAS Volume Manager Storage Administrator*.
- ◆ Interface de ligne de commande
Volume Manager comprend une série de commandes, des plus simples aux plus complexes nécessitant une entrée détaillée de la part de l'utilisateur. Généralement, il est nécessaire d'avoir une bonne maîtrise des concepts Volume Manager pour pouvoir utiliser ses commandes. Ces concepts sont décrits dans ce chapitre. Les commandes Volume Manager nécessitent en général des privilèges de super-utilisateur ou d'autres privilèges appropriés. La ligne de commande est décrite dans le *Guide de référence de VERITAS Volume Manager*.
- ◆ Volume Manager Support Operations
L'interface Volume Manager Support Operations (`vxdiskadm`) est une interface de menus qui permet d'exécuter des fonctions d'administration des disques et des volumes. `vxdiskadm` utilise un menu principal à partir duquel vous pouvez sélectionner des tâches de gestion de stockage à exécuter. Les opérations `vxdiskadm` sont décrites dans *Guide de référence de VERITAS Volume Manager*.

Interfaces utilisateur Volume Manager

Les objets Volume Manager créés via une interface sont compatibles avec ceux créés via les autres interfaces.

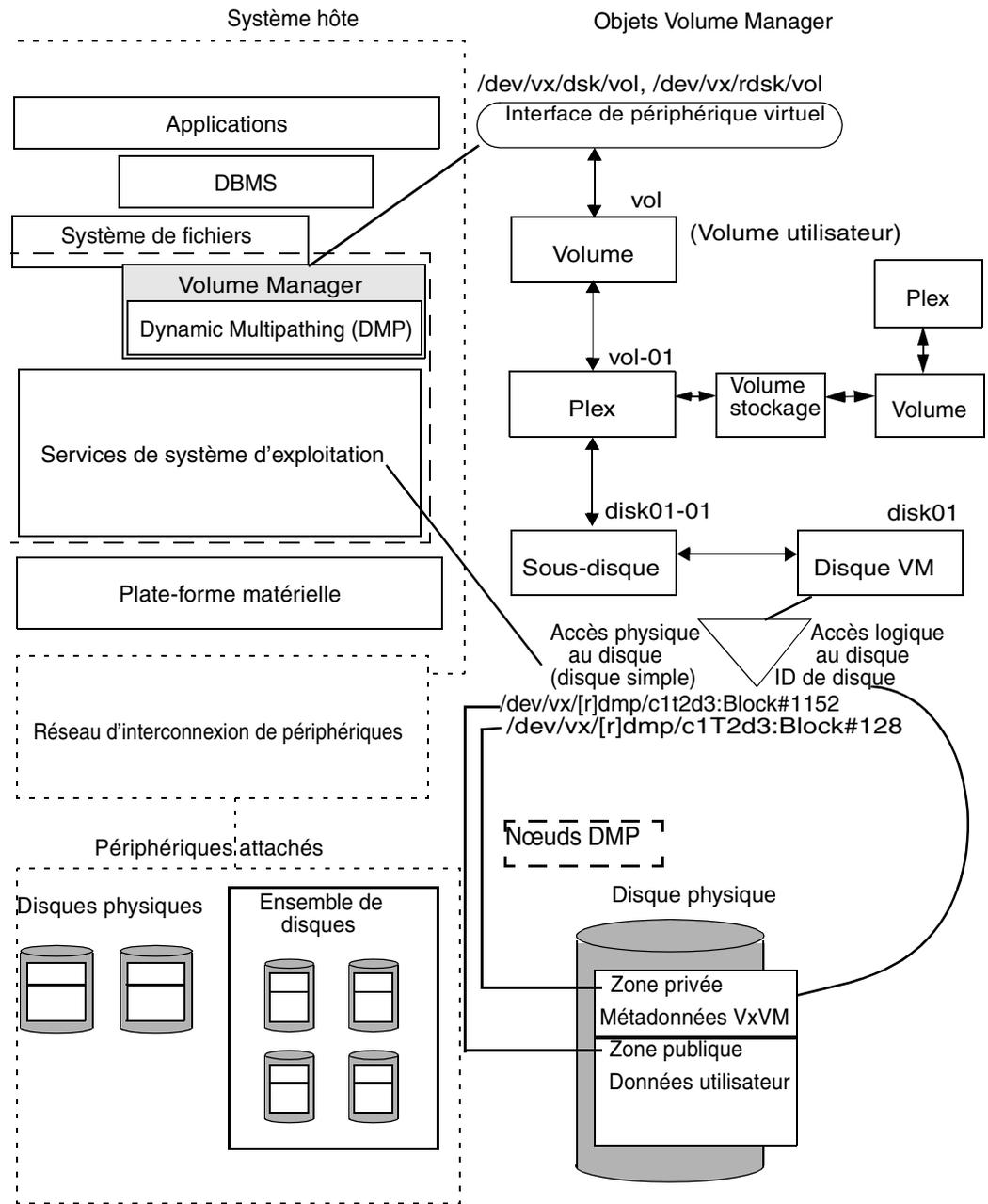
Pourquoi utiliser Volume Manager

Volume Manager offre un meilleur service de stockage des données en séparant les aspects physiques et logiques de la gestion de données. Volume Manager améliore le stockage des données en contrôlant les aspects suivants du stockage :

- ◆ espace — Allocation et utilisation ;
- ◆ performances — Transmission améliorée des données ;
- ◆ disponibilité des données — Fonctionnement continu et accès multisystème ;
- ◆ installation des périphériques — Gestion centralisée et optimisée ;
- ◆ système — Gestion multisystème et contrôle des systèmes privés/partagés.

Voir la figure 22, « Concepts système de Volume Manager », à la page 47.

Figure 22. Concepts système de Volume Manager



Interfaces utilisateur Volume Manager

Lorsque Volume Manager est installé sur un système hôte, procédez comme suit pour configurer et utiliser les objets Volume Manager :

- ◆ placez le contenu des disques physiques sous le contrôle de Volume Manager ;
- ◆ regroupez les disques Volume Manager dans des disk groups ;
- ◆ allouez l'espace des disk groups pour créer des volumes logiques.

Vous ne pouvez placer le contenu des disques physiques sous le contrôle de Volume Manager que si :

- ◆ vous permettez à Volume Manager de contrôler des disques physiques ;
- ◆ le disque n'est pas sous le contrôle d'un autre gestionnaire de stockage.

Volume Manager écrit les informations d'identification sur les disques physiques placés sous son contrôle (disques réclamés). Vous pouvez identifier ces disques, même après la déconnexion des disques physiques ou un arrêt du système. Volume Manager peut alors reconstituer les disk groups et les objets logiques pour permettre la détection des défaillances et récupérer le système plus rapidement.

Mise en route

2

Introduction

Cette chapitre décrit brièvement la procédure de configuration de Volume Manager et les démons nécessaires pour assurer son bon fonctionnement. Il contient également des instructions qui vous permettront de configurer un système avec la gestion du stockage.

Reportez-vous au *Guide d'installation de VERITAS Volume Manager* pour obtenir des informations détaillées sur l'installation et la configuration de Volume Manager et de Storage Administrator.

Ce chapitre traite des sujets suivants :

- ◆ Initialisation de Volume Manager
- ◆ Démons Volume Manager
- ◆ Configuration du système
- ◆ Recommandations pour la configuration du système
- ◆ Protection du système

Initialisation de Volume Manager

Le programme `vxinstall` vous permet d'initialiser Volume Manager. Il place les disques spécifiés sous le contrôle de Volume Manager. Par défaut, ces disques sont placés dans le disk group `rootdg`. Vous devez utiliser `vxinstall` pour initialiser au moins un disque dans `rootdg`. Vous pouvez ensuite utiliser `vxdiskadm` ou Storage Administrator pour initialiser ou encapsuler des disques supplémentaires dans des disk groups.

Lorsque l'installation du programme est terminée, procédez comme suit pour initialiser Volume Manager.

1. Connectez-vous en tant que super-utilisateur.
2. Créez un fichier `disks.exclude` pour exclure des disques du contrôle de Volume Manager. `vxinstall` ignore tous les disques indiqués dans ce fichier. Placez ce fichier dans : `/etc/vx/disks.exclude`.

Initialisation de Volume Manager

3. Créez un fichier `cntrls.exclude` pour exclure du contrôle de Volume Manager tous les disques situés sur un contrôleur spécifique. Placez ce fichier dans :
`/etc/vx/cntrls.exclude`.
4. Démarrez `vxinstall` en entrant la commande suivante : `vxinstall`.

`vxinstall` effectue alors les opérations suivantes :

- ◆ il exécute et affiche les informations de licence et vous demande d'entrer une clé ;
- ◆ il examine et liste tous les contrôleurs attachés au système ;
- ◆ il vous permet de choisir une procédure d'initialisation : Installation rapide ou personnalisée.

L'installation rapide vous permet d'initialiser ou d'encapsuler tous les disques. Choisissez l'installation personnalisée pour encapsuler certains disques sur un contrôleur déterminé et en initialiser d'autres.

L'installation personnalisée vous permet de contrôler les disques mis sous le contrôle de Volume Manager ainsi que la façon dont ils sont ajoutés. Vous pouvez initialiser ou encapsuler tous les disques placés sur un contrôleur ou vous pouvez en initialiser certains et en encapsuler d'autres.

Pour plus d'informations sur l'utilisation de l'installation rapide ou de l'installation personnalisée, reportez-vous au *Guide d'installation de VERITAS Volume Manager*.

La configuration choisie pour le système détermine si vous devrez arrêter le système et le redémarrer. Si vous choisissez d'encapsuler des disques, vous devez redémarrer le système. `vxinstall` vous indique si un redémarrage est nécessaire.

Vous pouvez utiliser cette commande pour confirmer que les principales procédures Volume Manager sont en cours d'exécution (`vxconfigd`, `vxnotify` et `vxrelocd`) après avoir installé et initialisé VM.

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

Démons Volume Manager

Il est nécessaire d'exécuter deux démons pour assurer le fonctionnement correct de Volume Manager :

- ◆ vxconfigd
- ◆ vxiod

Démon de configuration vxconfigd

Le démon de configuration Volume Manager (`vxconfigd`) permet de gérer la configuration des disques et des disk groups Volume Manager. `vxconfigd` transmet les modifications de configuration au noyau et modifie les informations de configuration stockées sur disque.

Démarrage du démon de configuration Volume Manager

`vxconfigd` est appelé par les scripts de démarrage lors de la procédure de démarrage.

Pour déterminer si le démon de volume est activé, entrez la commande suivante :

```
# vxdctl mode
```

Ce message s'affiche si `vxconfigd` est en cours d'exécution et activé :

```
mode: enabled
```

Ce message s'affiche si `vxconfigd` est en cours d'exécution, mais n'est pas activé :

```
mode: disabled
```

Pour activer le démon de volume, entrez la commande suivante :

```
# vxdctl enable
```

Ce message s'affiche si `vxconfigd` n'est pas en cours d'exécution :

```
mode: not-running
```

Pour démarrer `vxconfigd`, entrez la commande suivante :

```
# vxconfigd
```

Une fois démarré, `vxconfigd` devient automatiquement une procédure d'arrière-plan.

Par défaut, `vxconfigd` envoie les erreurs à la console. Cependant, vous pouvez configurer `vxconfigd` pour qu'il les envoie à un fichier journal.

Pour plus d'informations sur le démon `vxconfigd`, voir les pages `man vxconfigd(1M)` et `vxdctl(1M)`.

Démon d'E/S de volume vxiod

Le démon d'E/S étendues de volume (`vxiod`) permet d'exécuter des opérations d'E/S étendues sans bloquer les procédures d'appel.

Pour plus d'informations sur `vxiod`, voir la page `man vxiod (1M)`.

Démarrage du démon d'E/S de volume

Les démons `vxiod` sont démarrés en même temps que le système. Généralement, plusieurs démons `vxiod` sont en cours d'exécution à tout moment. `vxiod` est démarré lorsque le système est réinitialisé après l'installation initiale.

Pour vous assurer que les démons `vxiod` sont en cours d'exécution, entrez la commande suivante :

```
# vxiod
```

C'est la seule méthode vous permettant de déterminer si des démons `vxiod` sont en cours d'exécution car `vxiod` est un thread de noyau et il n'est pas visible via `ps` command.

Si des démons `vxiod` sont en cours d'exécution, le message suivant s'affiche :

```
10 Démons d'E/S du volume en cours d'exécution
```

où 10 représente le nombre de démons `vxiod` actuellement en cours d'exécution.

Si aucun démon `vxiod` n'est actuellement en cours d'exécution, démarrez-en certains en entrant la commande suivante :

```
# vxiod set 10
```

où vous pouvez remplacer 10 par le nombre voulu de démons `vxiod`. Il est recommandé de choisir au moins un démon `vxiod` pour chaque CPU du système.

Configuration du système

Cette section contient les informations qui vous aideront à configurer le système afin d'assurer une gestion efficace du stockage. Pour plus de détails sur des tâches de configuration spécifiques, reportez-vous au *Guide de référence de VERITAS Volume Manager* et au *Guide de l'administrateur de VERITAS Volume Manager Storage Administrator*.

La séquence de configuration du système suivante est une séquence type indiquée à titre d'exemple. Il se peut que vos spécifications système soient différentes. Les recommandations pour la configuration du système offrent des informations utiles pour des configurations spécifiques.

Exemple de séquence de configuration du système

La liste suivante décrit les étapes type que vous pouvez utiliser pour configurer votre système de gestion du stockage.

Configuration initiale

- ◆ Placez les disques sous le contrôle de Volume Manager.
- ◆ Créez de nouveaux disk groups (si vous ne voulez pas utiliser `root dg` ou que vous voulez d'autres disk groups).
- ◆ Créez des volumes.
- ◆ Placez le ou les systèmes de fichiers sur des volumes.

Options

- ◆ Encapsulez le disque `boot/root` et mettez-le en miroir afin de créer un autre disque d'initialisation.
- ◆ Spécifiez les disques de remplacement pour la reconstruction automatique.
- ◆ Ajoutez des miroirs aux volumes.

Maintenance

- ◆ Redimensionnez les volumes et les systèmes de fichiers.
- ◆ Ajoutez d'autres disques/disk groups.
- ◆ Créez des prises de vue.

Recommandations pour la configuration du système

Ces recommandations générales peuvent vous aider à comprendre et à planifier un système efficace de gestion du stockage. Vous pouvez utiliser les références croisées de chaque section pour obtenir davantage d'informations sur les recommandations indiquées.

Recommandations pour la reconstruction automatique

Vous pouvez suivre ces recommandations générales lors de l'utilisation de la reconstruction automatique. Pour plus d'informations, voir « Reconstruction automatique » à la page 69.

- ◆ La fonction de reconstruction automatique est activée par défaut. Vous pouvez la désactiver, mais il est recommandé de la laisser activée.

Recommandations pour la configuration du système

- ◆ Même si la reconstruction automatique ne nécessite pas de spécifier des disques de remplacement, vous pouvez indiquer au moins un disque de remplacement dans chaque disk group. Vous aurez ainsi un certain contrôle sur les disques utilisés pour la reconstruction. S'il n'existe pas de disque de remplacement, Volume Manager utilise l'espace disponible dans le disk group. Lorsque de l'espace libre est utilisé pour la reconstruction, il se peut que les performances aient diminué à l'issue de l'opération.
- ◆ Lorsque la reconstruction automatique est terminée, vous pouvez spécifier un ou plusieurs disques supplémentaires comme disques de remplacement afin d'accroître l'espace disponible (il se peut qu'une partie de l'espace de remplacement d'origine soit occupé par des sous-disques reconstruits).
- ◆ Si un disk group donné couvre plusieurs contrôleurs et contient plusieurs disques de remplacement, vous pouvez configurer ces disques sur des contrôleurs différents (au cas où l'un des contrôleurs échouerait).
- ◆ Dans le cas d'un volume mis en miroir, le disk group doit comprendre au moins un disque qui ne contient pas déjà un miroir du volume. Ce disque doit être soit un disque de remplacement ayant un espace libre, soit un disque normal qui n'a pas été exclu de la reconstruction automatique et qui a de l'espace disponible.
- ◆ Dans le cas d'un volume mis en miroir et réparti par bandes, le disk group doit comprendre au moins un disque qui ne contient pas déjà un des miroirs du volume ou un autre sous-disque dans le plex réparti par bandes. Ce disque doit être soit un disque de remplacement ayant un espace libre, soit un disque normal qui n'a pas été exclu de la reconstruction automatique et qui a de l'espace disponible.
- ◆ Dans le cas d'un volume RAID 5, le disk group doit comprendre au moins un disque qui ne contient pas déjà le plex RAID 5 (ou un de ses plex de journal) du volume. Ce disque doit être soit un disque de remplacement ayant un espace libre, soit un disque normal qui n'a pas été exclu de la reconstruction automatique et qui a de l'espace disponible.
- ◆ Si un sous-disque de journal des zones modifiées (DRL, *Dirty Region Logging*) fait partie du plex de données d'un volume mis en miroir, vous ne pouvez pas reconstruire ce plex. Vous pouvez placer les sous-disques de journal dans des plex qui ne contiennent aucune donnée (plex de journal).
- ◆ La reconstruction automatique ne conserve pas systématiquement les caractéristiques de performances d'origine ni la structure des données. Vous pouvez examiner l'emplacement des sous-disques que vous venez de reconstruire et déterminer s'il est préférable de les reconstruire sur des disques plus appropriés afin de rétablir les performances d'origine.
- ◆ La reconstruction automatique peut créer un nouveau miroir du disque racine si ce dernier est mis en miroir et qu'il échoue. Le disk group `rootdg` doit donc avoir un espace libre ou de remplacement contigu suffisant pour contenir les volumes du disque racine (`rootvol` et `swapvol` nécessitent un espace disque contigu).

- ◆ Vous pouvez construire des objets VxVM sur des disques de remplacement (au moyen de `vxmake` ou de l'interface de Storage Administrator), mais il est préférable d'utiliser des disques de remplacement réservés à la reconstruction automatique.

Recommandations pour la répartition des données par bandes sur plusieurs disques

Vous pouvez suivre ces recommandations générales lors de l'utilisation de la répartition des données par bandes sur plusieurs disques. Pour plus d'informations, voir « Répartition des données par bandes sur plusieurs disques (RAID 0) » à la page 30.

- ◆ Ne placez pas plusieurs colonnes d'un plex réparti par bandes sur le même disque physique.
- ◆ Calculez soigneusement les tailles d'unité de bande. En général, il est recommandé de choisir une taille d'unité de bande moyenne (par exemple, 64 Ko, qui correspond également à la valeur par défaut utilisée par `vxassist`). Si vous ne pouvez pas adapter la taille d'unité de bande à celle de la piste et que vous ne connaissez pas le modèle d'E/S de l'application, choisissez une taille d'unité de bande de 64 Ko.

Remarque Les unités de disque modernes ont souvent une « géométrie variable », c'est-à-dire que la taille de la piste varie d'un cylindre à l'autre (les pistes externes du disque ont davantage de secteurs que les pistes internes). Il n'est donc pas toujours approprié d'utiliser la taille de la piste comme taille d'unité de bande. Pour ces unités de disque, choisissez une taille d'unité de bande moyenne (par exemple, 64 Ko), sauf si vous connaissez le modèle d'E/S de l'application.

- ◆ Les volumes dont les unités de bande sont de petite taille peuvent présenter un temps de réponse médiocre pour les E/S séquentielles si les disques ne sont pas équipés d'axes synchronisés. Généralement, la répartition des données par bandes sur plusieurs disques à axes non synchronisés donne de meilleurs résultats avec des tailles d'unité de bande élevées et des plex d'E/S aléatoires multithread ou largement asynchrones.
- ◆ Généralement, plus la bande comprend de disques physiques, plus les performances d'E/S seront améliorées ; cependant, cela réduit le temps moyen de bon fonctionnement du volume. Si cela pose un problème, vous pouvez combiner la répartition des données par bandes sur plusieurs disques avec la mise en miroir pour obtenir un volume à performances élevées et plus fiable.
- ◆ Si un seul plex d'un volume mis en miroir est réparti par bandes, veillez à définir la stratégie du volume comme `prefer` pour le plex réparti par bandes. (La stratégie de lecture par défaut, `select`, définit automatiquement ce paramètre.)
- ◆ Si plusieurs plex d'un volume mis en miroir sont répartis par bandes, assurez-vous que la taille d'unité de bande est la même pour tous les plex répartis par bandes.

Recommandations pour la configuration du système

- ◆ Dans la mesure du possible, distribuez les sous-disques d'un volume réparti par bandes entre les unités de disque connectées à des contrôleurs et à des bus différents.
- ◆ Veillez à ne pas utiliser les contrôleurs qui ne prennent pas en charge les recherches superposées (ils sont rares).

La commande `vxassist` applique et implémente automatiquement la plupart de ces règles lorsqu'elle alloue de l'espace aux plex répartis par bandes d'un volume.

Recommandations pour la mise en miroir

Vous pouvez suivre ces recommandations générales lors de l'utilisation de la mise en miroir. Pour plus d'informations, voir « Mise en miroir (RAID 1) » à la page 35.

- ◆ Ne placez pas sur le même disque physique les sous-disques provenant de plex différents d'un volume mis en miroir. Cela réduit la disponibilité liée à la mise en miroir ainsi que les performances. L'utilisation de `vxassist` permet d'éviter ce problème.
- ◆ Pour garantir l'amélioration des performances via l'utilisation de la mise en miroir, il faut que 70 % minimum des opérations d'E/S physiques soient des opérations de lecture. Un pourcentage plus élevé d'opérations de lecture se traduit par un accroissement des performances. La mise en miroir risque de ne pas augmenter les performances, voire de les réduire, lorsque le pourcentage des opérations d'écriture est élevé.

Remarque Le système d'exploitation UNIX implémente un cache de système de fichiers. Les requêtes de lecture peuvent souvent être traitées à partir du cache. Cela peut entraîner une augmentation du nombre d'opérations d'E/S physiques de type écriture effectuées via le système de fichiers (par rapport au nombre obtenu au niveau de l'application).

- ◆ Dans la mesure du possible, utilisez des disques attachés à des contrôleurs différents pour la mise en miroir ou la répartition des données par bandes sur plusieurs disques. La plupart des contrôleurs de disques gèrent les recherches superposées qui sont lancées sur deux disques à la fois. Ne configurez pas deux plex du même volume sur des disques attachés à un contrôleur qui ne prend pas en charge les recherches superposées. Cela est important dans le cas des anciens contrôleurs ou des disques SCSI qui n'assurent pas de cache sur le lecteur, mais cela l'est moins pour la plupart des nouveaux modèles de contrôleurs et de disques SCSI utilisés dans presque tous les serveurs et stations de travail modernes. La mise en miroir au moyen des contrôleurs peut être utile car le système peut alors faire face à une défaillance de contrôleur. L'autre contrôleur peut continuer à fournir des données à partir de l'autre miroir.

- ◆ Les performances d'un plex peuvent être considérablement améliorées lorsqu'il est réparti par bandes ou concaténé sur plusieurs disques, ou bien lorsqu'il est situé sur un périphérique beaucoup plus rapide. La stratégie de lecture peut alors être configurée de manière à préférer le plex « le plus rapide ». Par défaut, un volume avec un plex réparti par bandes est configuré de façon à préférer la lecture de ce plex.

Recommandations pour le journal des zones modifiées (DRL)

Vous pouvez suivre ces recommandations générales lors de l'utilisation du journal des zones modifiées. Pour plus d'informations, voir « Journal des zones modifiées (DRL) » à la page 73.

Le journal des zones modifiées permet une récupération plus rapide des volumes mis en miroir à la suite d'une panne du système. Lorsque le journal des zones modifiées est activé, Volume Manager assure le suivi des zones d'un volume qui ont changé après l'écriture de données sur un plex. Volume Manager conserve une bitmap et stocke cette information dans un *sous-disque de journal*. Les sous-disques de journal sont définis pour un volume auxquels ils sont ajoutés pour créer un journal des zones modifiées. Les sous-disques de journal sont indépendants des plex, ils sont ignorés par les stratégies de plex et ils servent uniquement à contenir les informations du journal des zones modifiées.

Remarque L'utilisation du journal des zones modifiées risque d'affecter les performances du système lors de l'enregistrement de nombreuses opérations d'écritures.

Veillez suivre les recommandations suivantes lors de l'utilisation du journal des zones modifiées :

- ◆ Pour activer le journal des zones modifiées, le volume doit être mis en miroir.
- ◆ Il doit exister au moins un sous-disque de journal sur le volume pour que le journal des zones modifiées fonctionne. Cependant, il ne peut y avoir qu'un seul sous-disque de journal par plex.
- ◆ Le sous-disque utilisé comme sous-disque de journal ne doit pas contenir des données nécessaires.
- ◆ Il est possible de « mettre en miroir » les sous-disques de journal en plaçant plusieurs sous-disques de journal (mais un par plex uniquement) dans le volume. Cela garantit que l'enregistrement dans le journal continuera, même si une défaillance de disque rend un sous-disque de journal inaccessible.
- ◆ Vous devez configurer les sous-disques de journal avec un ou deux secteurs (choisissez de préférence un nombre pair car le dernier secteur d'un sous-disque de journal ayant un nombre impair de secteurs n'est pas utilisé). La taille du sous-disque de journal est généralement proportionnelle à celle du volume. Si la taille d'un volume est inférieure à 2 Go, un sous-disque de journal de 2 secteurs est suffisant. Il

Recommandations pour la configuration du système

faut alors augmenter la taille du sous-disque de journal de 2 secteurs par tranche supplémentaire de 2 Go de taille de volume. Cependant, `vxassist` choisit par défaut des tailles raisonnables. En général, il est recommandé d'utiliser la longueur de sous-disque de journal par défaut proposée par `vxassist`.

- ◆ Dans la mesure du possible, le sous-disque de journal ne devrait pas être placé sur un disque fréquemment utilisé.
- ◆ Des disques de stockage permanents (non volatiles) devraient être utilisés pour les sous-disques de journal.

Recommandations pour la mise en miroir et la répartition des données par bandes sur plusieurs disques

Vous pouvez suivre ces recommandations générales lors de l'utilisation de la mise en miroir et de la répartition des données par bandes sur plusieurs disques. Pour plus d'informations, voir « Mise en miroir avec répartition des données par bandes sur plusieurs disques (RAID 1 + RAID 0) » à la page 36.

- ◆ Assurez-vous que vous disposez d'un nombre suffisant de disques pour la configuration répartition par bandes-miroir. Deux disques minimum sont nécessaires pour le plex réparti par bandes et un ou plusieurs *autres* disques pour le miroir.
- ◆ Ne placez jamais les sous-disques d'un plex sur le même disque physique que ceux de l'autre plex. Suivez les recommandations pour la répartition des données par bandes sur plusieurs disques décrites dans « Recommandations pour la répartition des données par bandes sur plusieurs disques » à la page 55.
- ◆ Suivez les recommandations pour la mise en miroir décrites dans « Recommandations pour la mise en miroir » à la page 56.

Recommandations pour la répartition des données par bandes sur plusieurs disques et la mise en miroir

Vous pouvez suivre ces recommandations générales lors de l'utilisation de la répartition des données par bandes sur plusieurs disques et de la mise en miroir. Pour plus d'informations, voir « Répartition des données par bandes sur plusieurs disques avec mise en miroir (RAID 0 + RAID 1) » à la page 36.

- ◆ Assurez-vous que vous disposez d'un nombre suffisant de disques pour la configuration répartition par bandes-miroir. Deux disques minimum sont nécessaires pour le plex réparti par bandes et un ou plusieurs *autres* disques pour le miroir.
- ◆ Ne placez jamais les sous-disques d'un plex sur le même disque physique que ceux de l'autre plex. Suivez les recommandations pour la répartition des données par bandes sur plusieurs disques décrites dans « Recommandations pour la répartition des données par bandes sur plusieurs disques » à la page 55.

- ◆ Suivez les recommandations pour la mise en miroir décrites dans « Recommandations pour la mise en miroir » à la page 56.

Recommandations pour RAID 5

Vous pouvez suivre ces recommandations générales lors de l'utilisation de RAID 5. Pour plus d'informations, voir « RAID 5 » à la page 34.

En général, les recommandations relatives à l'utilisation conjointe de la mise en miroir et de la répartition des données par bandes sur plusieurs disques s'appliquent également à RAID 5. Vous devez également suivre les recommandations ci-dessous avec RAID 5 :

- ◆ Il ne peut exister qu'un seul plex RAID 5 par volume RAID 5 (mais il peut exister plusieurs plex de journal).
- ◆ Le plex RAID 5 doit provenir d'au moins deux sous-disques sur deux ou plusieurs disques physiques. S'il existe des plex de journal, ils doivent appartenir aux disques qui ne sont pas utilisés pour le plex RAID 5.
- ◆ Les journaux RAID 5 peuvent être mis en miroir et répartis par bandes.
- ◆ Si la longueur du volume n'est pas explicitement spécifiée, elle est définie comme étant la longueur de n'importe quel plex RAID 5 associé au volume ; sinon, la valeur zéro lui est attribuée. Si la longueur du volume est définie explicitement, elle doit correspondre, le cas échéant, à un multiple de la taille d'unité de bande du plex RAID 5 associé.
- ◆ Si la longueur du journal n'est pas spécifiée explicitement, elle est définie, le cas échéant, comme étant la longueur du plus petit plex de journal RAID 5 associé. Si aucun plex de journal RAID 5 n'est défini, la valeur zéro lui est attribuée.
- ◆ Les plex de journal RAID 5 clairsemés ne sont pas valides.

Protection du système

Les défaillances de disque peuvent causer deux types de problème : la perte des données du disque ayant échoué et l'impossibilité d'accéder au système. La perte d'accès peut être imputable à la défaillance d'un disque clé (un disque utilisé pour les opérations système). VERITAS Volume Manager peut protéger votre système contre ces problèmes.

Pour garantir la disponibilité du système, vous devez mettre en miroir les données importantes pour l'exécution et le démarrage du système. Ces données doivent être conservées pour pouvoir être utilisées en cas de défaillance.

Les suggestions suivantes vous permettront de protéger votre système et vos données :

Protection du système

- ◆ Placez le disque contenant le système de fichiers racine (le disque *racine* ou *d'initialisation*) sous le contrôle de Volume Manager au moyen de l'encapsulation. Cette opération convertit les périphériques `root` et `swap` en volumes (`rootvol` et `swapvol`). Mettez ensuite en miroir le disque racine afin de disposer d'un autre disque racine pour le démarrage. La mise en miroir des disques importants pour le démarrage assure que la défaillance d'un seul disque n'empêchera pas le démarrage et l'utilisation du système.

Pour assurer la disponibilité maximale du système, il est nécessaire de créer des miroirs pour les volumes `rootvol`, `swapvol`, `usr` et `var`. Pour plus d'informations, voir le chapitre « Récupération » à la page 243.

- ◆ Utilisez la mise en miroir pour protéger les données. Lorsque vous mettez en miroir les données, vous évitez tout risque de perte de données dû à une défaillance de disque. Pour conserver les données, vous devez créer et utiliser les volumes mis en miroir qui comprennent au moins deux plex de données. Ces plex doivent se trouver sur des disques différents. Si une défaillance de disque entraîne l'échec d'un plex, les données du volume mis en miroir continuent d'exister sur l'autre disque.

Lors de la création des miroirs, `vxassist mirror` permet de détecter les miroirs de sorte que la perte d'un disque ne se traduit pas par une perte des données. Par défaut, `vxassist` ne crée pas de volumes mis en miroir ; vous pouvez modifier le fichier `/etc/default/vxassist` afin de définir la structure par défaut comme étant mise en miroir.

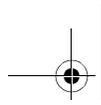
- ◆ Conservez la fonction de reconstruction automatique de Volume Manager activée afin qu'elle puisse détecter automatiquement les défaillances, vous informer de la nature de ces défaillances, essayer de reconstruire les sous-disques affectés redondants et lancer les procédures de récupération. Spécifiez au moins un disque de remplacement pour la reconstruction automatique par `disk group` afin de disposer d'un espace suffisant pour la reconstruction en cas de défaillance.

Si le disque `root` est mis en miroir, la reconstruction automatique peut créer automatiquement un autre miroir de ce disque si celui d'origine échoue. Le `disk group rootdg` devrait contenir suffisamment d'espace libre ou de remplacement contigu pour les volumes du disque racine (les volumes `rootvol` et `swapvol` nécessitent un espace disque contigu).

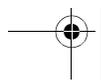
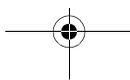
- ◆ Pour les volumes mis en miroir, utilisez la fonction de journal des zones modifiées pour récupérer plus rapidement les volumes mis en miroir après une panne du système. Assurez-vous que chaque volume mis en miroir comprend au moins un sous-disque de journal. (Les volumes `rootvol`, `swapvol` et `usr` ne peuvent pas être des volumes DRL.)
- ◆ Pour les volumes RAID 5, utilisez la fonction de journalisation (`logging`) pour éviter tout risque d'endommagement des données de récupération. Assurez-vous que chaque volume RAID 5 comprend au moins un plex de journal.

Effectuez régulièrement des sauvegardes afin de protéger vos données. Les sauvegardes sont nécessaires si toutes les copies d'un volume sont perdues ou endommagées. Par exemple, la surtension risque d'endommager plusieurs (voire tous les) disques de votre système. De plus, si vous entrez une commande incorrecte, vous risquez de supprimer des fichiers très importants ou d'endommager directement un système de fichiers.





Protection du système



Opérations Volume Manager

3

Introduction

Ce chapitre contient des informations détaillées sur les fonctions de VERITAS Volume Manager.

Il traite des sujets suivants :

- ◆ Restructuration en ligne
- ◆ Reconstruction automatique
- ◆ Resynchronisation des volumes
- ◆ Journal des zones modifiées (DRL)
- ◆ Resynchronisation rapide des miroirs
- ◆ Contrôle racine de Volume Manager
- ◆ Fonction DMP
- ◆ VxSmartSync Recovery Accelerator
- ◆ Moniteur de tâches Volume Manager
- ◆ Fonction de cluster Volume Manager

Restructuration en ligne

La *restructuration en ligne* vous permet de convertir n'importe quelle structure de stockage prise en charge dans Volume Manager en une autre, au même emplacement et *sans interrompre l'accès aux données*. Généralement, la structure de stockage dans Volume Manager est modifiée afin d'accroître la redondance ou les performances du stockage. Volume Manager rend le stockage plus redondant en dupliquant l'espace d'adressage (mise en miroir) ou en ajoutant la parité (RAID 5). Les performances de stockage dans Volume Manager peuvent être changées en modifiant les paramètres de répartition des données par bandes sur plusieurs disques (le nombre de colonnes et la largeur de bande).

Restructuration en ligne

Les changements de structure peuvent correspondre à l'un des types suivants :

- ◆ RAID 5 vers mise en miroir ;
- ◆ mise en miroir vers RAID 5 ;
- ◆ ajout ou suppression de la parité ;
- ◆ ajout ou suppression de colonnes ;
- ◆ modification de la largeur de bande.

Structure de stockage

La restructuration en ligne prend actuellement en charge les structures de stockage suivantes :

- ◆ concaténation ;
- ◆ répartition par bandes ;
- ◆ RAID 5 ;
- ◆ mise en miroir ;
- ◆ répartition par bandes-miroir ;
- ◆ concaténation-miroir.

Remarque Lorsque vous utilisez VERITAS Volume Manager Storage Administrator, les termes *Pro réparti par bandes (striped-pro)* et *Pro concaténé (concatenated-pro)* de l'interface graphique utilisateur désignent respectivement un miroir réparti par bandes et un miroir concaténé.

Principes de fonctionnement de la restructuration en ligne

La fonction de restructuration en ligne VERITAS vous permet de modifier les structures de stockage que vous avez déjà créées sans affecter l'accès aux données. Vous pouvez changer les performances d'une structure particulière en fonction de l'évolution des spécifications. Vous pouvez transformer une structure en une autre en appelant une seule commande.

Un plex avec miroir réparti par bandes est un plex réparti par bandes placé sur un volume mis en miroir, ce qui équivaut à un seul plex, à la fois mis en miroir et réparti par bandes sur plusieurs disques. Cette combinaison forme un plex appelé *plex réparti par bandes-miroir*. Un plex concaténé peut être mis en miroir de la même façon. La restructuration en ligne prend en charge les transformations des plex répartis par bandes-miroir vers les plex concaténés-miroir et vice versa.

Remarque Actuellement, il est impossible de modifier le nombre de miroirs pendant une restructuration.

Par exemple, vous utilisez une structure répartie par bandes avec une taille d'unité de bande de 128 Ko, qui n'offre peut-être pas des performances optimales. Vous pouvez modifier la taille d'unité de bande à l'aide de la fonction de restructuration.

Il est inutile de démonter les systèmes de fichiers montés sur les volumes pour effectuer cette transformation tant que le système de fichiers permet les opérations de réduction et de développement en ligne. VFS offre ces fonctions.

La restructuration en ligne réutilise l'espace de stockage existant et détermine les besoins de la nouvelle structure à l'aide de stratégies d'allocation d'espace. La procédure de restructuration convertit la structure d'un volume donné en structure de destination en utilisant un espace temporaire minimum.

La transformation s'effectue en déplaçant les données, portion par portion, de la structure source vers la structure de destination. Les données sont copiées du volume source vers l'espace temporaire. Elles sont supprimées de la zone de stockage source par portions. La zone de stockage du volume source est ensuite transformée en nouvelle structure et les données enregistrées dans l'espace temporaire sont écrites sur la nouvelle structure. Cette opération est répétée jusqu'à ce que le stockage et les données du volume source aient été transformés en fonction de la nouvelle structure.

Vous pouvez utiliser la restructuration en ligne pour modifier le nombre de colonnes et la largeur des bandes, supprimer et ajouter la parité et passer de RAID 5 à la mise en miroir.

Types de restructuration

Vous devez respecter au moins un des critères suivants pour réussir une opération de restructuration en ligne. Vous devez exécuter une ou plusieurs des opérations suivantes :

- ◆ Passer de RAID 5 à la mise en miroir ;
- ◆ Passer de miroir à RAID 5 ;
- ◆ Modifier le nombre de colonnes ;
- ◆ Modifier la largeur de bande ;
- ◆ Supprimer ou ajouter la parité.

Restructuration en ligne

Pour pouvoir faire l'objet d'une restructuration, les plex des volumes mis en miroir doivent avoir une structure identique, avec la même largeur de bande et le même nombre de colonnes. Voir le tableau 1, « Restructurations prises en charge ».

Tableau 1. Restructurations prises en charge

Depuis/Vers	Réparti par bandes Mis en miroir	Mis en miroir Concaténé	Mis en miroir normal	RAID 5	Concaténé	Réparti par bandes
Réparti par bandes Mis en miroir	Oui 1	Oui 2	Non 3	Oui 4	Oui 5	Oui 6
Concaténé Mis en miroir	Oui 7	Non 8	Non 9	Oui 10	Non 11	Oui 12
Mis en miroir normal	Oui 13	Oui 14	Non 15	Oui 16	Non 17	Non 18
RAID 5	Oui 4	Oui 10	Non 19	Oui 20	Oui 21	Oui 22
Concaténé	Oui 5	Non 11	Non 17	Oui 21	Non 23	Oui 24
Réparti par bandes	Oui 6	Oui 12	Non 18	Oui 22	Oui 24	Oui 25

Le tableau 1 contient les informations suivantes :

- ◆ *Oui* — Indique que la restructuration en ligne est possible.
- ◆ *Non* — Indique que l'opération *peut* être possible, mais vous ne pouvez pas utiliser la fonction de restructuration.
- ◆ *Nombres* — Indique une brève description des changements possibles pour cette restructuration particulière. Voir « Descriptions des nombres » ci-dessous.
- ◆ *Opérations* — Peuvent s'effectuer dans les deux directions.

Descriptions des nombres

Les nombres du tableau 1 décrivent l'opération de restructuration comme suit :

1. Change la largeur de bande ou le nombre de colonnes.
2. Supprime toutes les colonnes.
3. Opération de conversion, mais pas de restructuration.

4. Change la mise en miroir en RAID 5 et/ou la largeur de bande/le nombre de colonnes.
5. Change la mise en miroir en RAID 5 et/ou la largeur de bande/le nombre de colonnes.
6. Change la largeur de bande/le nombre de colonnes et supprime un miroir.
7. Ajoute des colonnes.
8. Pas une opération de restructuration.
9. Opération de conversion.
10. Change la mise en miroir en RAID 5. Voir la procédure `vxconvert`.
11. Supprime un miroir ; pas une opération de restructuration.
12. Supprime un miroir et ajoute la répartition des données par bandes sur plusieurs disques.
13. Convertit un ancien volume mis en miroir en un miroir réparti par bandes. La restructuration est uniquement valide si vous modifiez le nombre de colonnes/la largeur de bandes ; dans le cas contraire, il s'agit d'une opération de conversion. Voir la procédure `vxconvert`.
14. Convertit un ancien volume mis en miroir en un volume concaténé. La restructuration est uniquement valide si vous modifiez le nombre de colonnes ; dans le cas contraire, il s'agit d'une opération de conversion.
15. Aucun changement ; pas une opération de restructuration.
16. Convertit un ancien volume mis en miroir en RAID 5. Vous devez choisir un plex dans l'ancien volume mis en miroir pour utiliser la restructuration. L'autre plex est supprimé à l'issue de la restructuration.
17. À moins de choisir un plex dans le volume mis en miroir et de changer le nombre de colonnes/la largeur de bande, il ne s'agit pas d'une opération de restructuration.
18. À moins de choisir un plex dans le volume mis en miroir et de changer le nombre de colonnes/la largeur de bandes, il ne s'agit pas d'une opération de restructuration.
19. Pas une opération de restructuration.
20. Change la largeur de bande/le nombre de colonnes.

Restructuration en ligne

21. Supprime la parité et toutes les colonnes.
22. Supprime la parité.
23. Aucun changement ; pas une opération de restructuration.
24. Supprime les colonnes.
25. Change la largeur de bande/le nombre de colonnes.

Caractéristiques de la restructuration

La transformation des données d'une structure à l'autre implique la réorganisation des données de la structure existante dans la nouvelle structure. Pendant la transformation, la restructuration en ligne conserve la redondance des données en mettant en miroir l'espace temporaire utilisé. L'accès aux données en lecture/écriture n'est pas interrompu pendant la transformation.

Les données ne sont pas endommagées en cas de défaillance du système pendant une transformation. La transformation se poursuit après la restauration du système et l'accès en lecture/écriture est conservé.

Vous pouvez inverser à tout moment la procédure de restructuration, mais vous ne pourrez pas retourner les données à leur emplacement de stockage précédent. Veillez à arrêter toute transformation existante dans le volume avant d'inverser l'opération.

Pour déterminer la direction d'une transformation, utilisez la commande `vxrelayout status`.

Ces transformations éliminent les erreurs d'E/S à condition que la redondance et l'espace soient suffisants pour permettre le déplacement des données.

Restructurations et longueur de volume

Certaines restructurations risquent d'augmenter ou de réduire la longueur du volume. Dans ce cas, la restructuration en ligne utilise `vxresize` pour réduire ou développer un système de fichiers.

La restructuration en ligne ne transforme pas les plex clairsemés et elle ne permet pas non plus de rendre un plex clairsemé.

Remarque La restructuration en ligne ne peut être utilisée qu'avec les volumes créés à l'aide de la commande `vxassist` ou de Storage Administrator.

Les transformations suivantes ne sont pas prises en charge :

- ◆ Transformation des plex de journal
- ◆ Prise de vue d'un volume lorsqu'une opération de restructuration en ligne est en cours d'exécution sur le volume

Reconstruction automatique

La *reconstruction automatique* permet à un système de répondre automatiquement aux erreurs d'E/S sur les objets Volume Manager redondants (mis en miroir ou RAID 5) et de restaurer la redondance et l'accès à ces objets. Volume Manager détecte les erreurs d'E/S sur les objets et reconstruit les sous-disques affectés sur les disques spécifiés comme *disques de remplacement* et/ou sur l'espace libre du disk group. Volume Manager reconstruit alors les objets qui existaient avant la défaillance et rétablit leur redondance et leur accessibilité.

En cas de défaillance partielle du disque (une défaillance affectant uniquement certains sous-disques d'un disque), les données redondantes sont reconstruites sur la partie du disque qui a échoué. Les volumes existant dans les portions non affectées du disque restent accessibles.

Remarque La reconstruction automatique s'applique uniquement aux sous-disques redondants (mis en miroir ou RAID 5) d'un disque qui a échoué. Les sous-disques non redondants d'un disque qui a échoué ne sont pas reconstruits, mais l'administrateur système est notifié de la défaillance.

Principes de fonctionnement de la reconstruction automatique

La fonction de reconstruction automatique est activée par défaut. Aucune intervention de l'administrateur système n'est requise pour activer la reconstruction automatique en cas de défaillance.

Le démon de reconstruction automatique, `vxrelocd`, recherche dans Volume Manager les événements susceptibles d'affecter la redondance et exécute la reconstruction automatique pour restaurer cette dernière. `vxrelocd` notifie également l'administrateur système (par courrier électronique) des défaillances et des éventuelles opérations de reconstruction et de récupération. Pour plus d'informations sur `vxrelocd`, voir la page `man vxrelocd(1M)`.

Le démon `vxrelocd` démarre en même temps que le système et surveille sur Volume Manager les défaillances relatives aux disques, plex ou sous-disques RAID 5. La reconstruction automatique est tentée à la suite d'une défaillance.

Reconstruction automatique

Une reconstruction automatique réussit lorsque les opérations suivantes sont effectuées :

1. détection des événements Volume Manager dus à la défaillance d'un disque, d'un plex ou d'un sous-disque RAID 5 ;
2. notification de l'administrateur système (et des autres utilisateurs spécifiés) de l'échec et identification des objets Volume Manager affectés. Cette opération s'effectue via courrier électronique ;
3. détermination des sous-disques qui peuvent être reconstruits et de l'espace nécessaire dans le disk group, et reconstruction des sous-disques. L'administrateur système doit être notifié de ces actions et de leur résultat (réussite ou échec) ;
4. lancement des procédures de récupération nécessaires pour restaurer les volumes et les données. L'administrateur système doit être notifié du résultat de cette tentative de récupération.

Remarque La reconstruction automatique ne garantit pas la même structure de données ni les mêmes performances à l'issue de l'opération. L'administrateur système peut modifier la configuration à l'issue de la reconstruction automatique.

Sélection de l'espace pour la reconstruction

Vous devez initialiser un disque de remplacement et le placer dans un disk group en tant que disque de remplacement *avant* de pouvoir l'utiliser pour le remplacement. Si une défaillance se produit alors que vous n'aviez pas spécifié de disque de remplacement, Volume Manager utilise automatiquement n'importe quel espace disponible dans le disk group dans lequel la défaillance s'est produite. Si l'espace disponible est insuffisant, le programme utilise une combinaison d'espace de remplacement et d'espace libre.

L'espace libre mentionné dans la reconstruction automatique correspond toujours à l'espace libre non exclu de la reconstruction automatique. Pour exclure des disques de la reconstruction automatique, utilisez l'interface Storage Administrator : `vxdiskadm` ou `vxedit`.

L'administrateur système peut spécifier un ou plusieurs disques comme disques de remplacement pour la reconstruction automatique dans chaque disk group. Pour cela, il est nécessaire d'utiliser l'interface Storage Administrator, `vxdiskadm` ou `vxedit`. Les disques spécifiés comme disques de remplacement ne sont pas inclus dans le modèle d'espace libre et aucun espace de stockage ne devrait leur être alloué.

Lors de la sélection de l'espace pour la reconstruction, la reconstruction automatique conserve la redondance de l'objet Volume Manager auquel le sous-disque reconstruit appartient. Par exemple, la reconstruction automatique assure que les sous-disques d'un plex qui a échoué ne sont pas reconstruits sur un disque contenant un miroir de ce plex.

S'il est impossible de conserver la redondance au moyen des disques de remplacement disponibles et/ou de l'espace libre, la reconstruction automatique n'est pas effectuée. Si la reconstruction est impossible, l'administrateur système est notifié du problème et aucune action supplémentaire n'est effectuée.

Pour déterminer les disques qui doivent être utilisés, la reconstruction automatique utilise de préférence le disque « le plus proche » du disque qui a échoué. La « proximité » est déterminée en fonction du contrôleur, de la cible et du numéro du disque qui a échoué. Un disque situé sur le même contrôleur que le disque qui a échoué est plus proche qu'un disque situé sur un contrôleur différent ; un disque sous la même cible que le disque qui a échoué est plus proche qu'un disque sous une cible différente.

La reconstruction automatique tente, si possible, de déplacer tous les sous-disques d'un disque qui a échoué vers le même disque de destination.

Si le disque qui a échoué est un disque racine, la reconstruction automatique ne fonctionne que si les systèmes de fichiers sont sur le même disque. S'ils sont introuvables, l'administrateur système est notifié par courrier électronique.

Lors de la reconstruction automatique, le sous-disque qui a échoué est supprimé de la base de données de configuration et Volume Manager vérifie que l'espace disque occupé par ce sous-disque n'est pas recyclé en tant qu'espace disponible.

Pour plus d'informations sur la désactivation de la reconstruction automatique, voir le *Guide d'installation de VERITAS Volume Manager*.

Utilitaire de reconstruction dans le disque d'origine

La reconstruction automatique VxVM permet au système de répondre automatiquement aux erreurs d'E/S sur un objet VxVM redondant au niveau du sous-disque et de prendre les mesures nécessaires pour rendre l'objet de nouveau disponible. Cette opération détecte les erreurs d'E/S sur un sous-disque, le reconstruit et récupère le plex qui lui est associé. Une fois le disque remplacé, vous pouvez utiliser l'utilitaire `vxunrelloc` de Volume Manager pour restaurer sur le système la configuration qui existait avant la défaillance du disque. `vxunrelloc` vous permet de rétablir les sous-disques reconstruits automatiquement dans le disque remplacé à la suite d'une défaillance.

Resynchronisation des volumes

Lors du stockage redondant des données à l'aide des volumes mis en miroir ou RAID 5, Volume Manager vérifie la correspondance exacte de toutes les copies des données. Cependant, dans certains cas (le plus souvent en cas de défaillance complète du système), certaines données redondantes d'un volume deviennent incohérentes ou *non synchronisées*. Les données mises en miroir ne sont pas exactement identiques à celles d'origine. À l'exception des modifications normales de la configuration (par exemple, le détachement et le rattachement d'un plex), ce problème se produit uniquement lors de l'arrêt anormal d'un système alors que des données sont en cours d'écriture sur un volume.

Les données sont écrites en parallèle sur les miroirs d'un volume, à l'instar des données et de la parité dans un volume RAID 5. Si un système s'arrête de façon anormale avant la fin de toutes les écritures individuelles, certaines écritures pourront être terminées et d'autres pas. Par conséquent, les données ne seront plus synchronisées. Pour les volumes mis en miroir, ce problème peut se traduire par le fait que deux lectures effectuées à partir de la même zone du volume retourneront deux résultats différents, si la requête de lecture utilise des miroirs différents. Dans le cas des volumes RAID 5, cela peut entraîner l'endommagement de la parité et la reconstruction incorrecte de données.

Volume Manager doit s'assurer que tous les miroirs contiennent exactement les mêmes données et vérifier la correspondance entre les données et la parité dans les volumes RAID 5. Cette procédure est appelée *resynchronisation des volumes*. Lorsque les volumes appartiennent à des disk groups qui sont automatiquement importés au démarrage (par exemple, rootdg), la resynchronisation est exécutée à la réinitialisation du système.

Il n'est pas nécessaire de resynchroniser tous les volumes à la suite d'une défaillance du système. Les volumes qui n'ont jamais été écrits ou qui étaient au repos (c'est-à-dire, sans E/S actives) au moment de la défaillance du système ne peuvent pas avoir d'écritures en cours et ne doivent pas être resynchronisés.

Volume Manager enregistre le moment de la première écriture sur un volume et le marque comme *modifié (dirty)*. Lorsqu'un volume est fermé par tous les processus ou arrêté normalement par l'administrateur, les écritures sont terminées et Volume Manager supprime l'indicateur « modifié » du volume. Seuls les volumes marqués comme modifiés à la réinitialisation du système doivent être resynchronisés.

La resynchronisation dépend du type de volume. Les volumes RAID 5 qui contiennent des journaux RAID 5 peuvent « relire » ces journaux. Si aucun journal n'est disponible, le volume est placé en mode reconstruction-récupération et la parité est entièrement régénérée. Pour les volumes mis en miroir, la resynchronisation s'effectue en plaçant le volume en mode récupération (ou *mode de récupération en lecture-réécriture*). La resynchronisation des données dans le volume s'effectue en arrière-plan, ce qui permet d'utiliser le volume pendant la récupération.

La resynchronisation peut se révéler onéreuse et risque d'affecter les performances du système. La procédure de récupération permet de limiter ce problème en répartissant les récupérations de manière à éviter de surcharger un disque ou un contrôleur spécifique.

Si les volumes sont importants ou nombreux, la resynchronisation peut prendre un certain temps. Pour éviter cet inconvénient, vous pouvez utiliser le journal des zones modifiées pour les volumes mis en miroir ou vérifier que les volumes RAID 5 ont des journaux RAID 5 valides. Si les volumes sont utilisés par des applications de base de données, vous pouvez utiliser VxSmartSync™ Recovery Accelerator (voir « VxSmartSync Recovery Accelerator » à la page 85).

Journal des zones modifiées (DRL)

Le journal des zones modifiées (DRL, *Dirty Region Logging*) est une propriété facultative d'un volume, qui permet de récupérer rapidement des volumes mis en miroir après une défaillance du système. Ce journal assure le suivi des zones qui ont été modifiées par les écritures d'E/S sur un volume mis en miroir. Il utilise ces informations pour récupérer uniquement les portions du volume qui doivent l'être.

Si vous n'utilisez pas le journal des zones modifiées et qu'une défaillance du système se produit, vous devrez restaurer tous les miroirs des volumes avec un état cohérent. La restauration s'effectue en copiant tout le contenu du volume dans ses miroirs. Cette procédure peut prendre un certain temps et utiliser un grand nombre d'E/S. Vous devrez peut-être également récupérer les zones des volumes qui sont déjà cohérentes.

Le journal des zones modifiées divise de façon logique un volume en un ensemble de zones consécutives. Il assure le suivi des zones du volume qui sont l'objet d'écritures. Un journal des zones modifiées est conservé. Il contient un bit d'état qui représente chaque zone du volume. Pour n'importe quelle opération d'écriture sur le volume, les zones en cours d'écriture sont marquées comme modifiées dans le journal avant l'écriture des données. Si une écriture modifie une zone du journal qui était auparavant inaltérée, le journal est écrit de façon synchrone sur le disque avant que les données ne soient écrites. Au redémarrage du système, Volume Manager récupère uniquement les zones du volume qui sont marquées comme modifiées dans le journal des zones modifiées.

Les *sous-disques de journal* permettent de stocker le journal des zones modifiées activé d'un volume. Un volume avec un journal des zones modifiées contient au moins un sous-disque de journal ; vous pouvez utiliser plusieurs sous-disques de journal pour mettre en miroir le journal des zones modifiées. Chaque sous-disque de journal est associé à un plex du volume. Il ne peut exister qu'un seul sous-disque de journal par plex. Si le plex ne contient qu'un seul sous-disque de journal et aucun sous-disque de données, ce plex est appelé *plex de journal*.

Resynchronisation rapide des miroirs

Vous pouvez également associer le sous-disque de journal à un plex normal contenant des sous-disques de données. Dans ce cas, le sous-disque de journal risque de ne plus être disponible si vous devez détacher le plex à cause de la défaillance de l'un de ses sous-disques de données.

Si vous utilisez la commande `vxassist` pour générer un journal des zones modifiées, elle crée un plex de journal contenant par défaut un seul sous-disque de journal. Vous pouvez aussi générer manuellement un journal des zones modifiées en créant un sous-disque de journal, puis en l'associant à un plex. Le plex peut alors contenir à la fois un sous-disque de journal et des sous-disques de données.

Seul un nombre limité de bits peut être marqué comme modifié dans le journal à un moment donné. Le bit modifié d'une zone n'est pas effacé dès que les données y sont écrites. En revanche, la zone reste marquée comme modifiée jusqu'à ce que la zone de volume correspondante soit la moins récemment utilisée. Si un bit d'une zone donnée est déjà marqué comme modifié alors qu'une autre écriture est effectuée dans la même zone, il n'est pas nécessaire d'écrire le journal sur le disque pour permettre l'exécution de l'opération d'écriture.

Certains volumes, tels que ceux utilisés pour les journaux de relecture Oracle, sont écrits en mode séquentiel et ne bénéficient pas de ce nettoyage des bits du journal des zones modifiées. Pour ces volumes, vous pouvez utiliser le mode *DRL séquentiel* afin de réduire davantage le nombre de bits modifiés et de rendre la récupération plus rapide. Le nombre de bits modifiés autorisés pour le mode DRL séquentiel sera limité par la commande réglable `voldr1_max_dirty`. L'utilisation du mode DRL séquentiel sur les volumes écrits en mode séquentiel risque d'affecter considérablement le débit d'E/S.

Remarque Le journal des zones modifiées ajoute un petit délai d'E/S à la plupart des modèles d'accès en écriture.

Resynchronisation rapide des miroirs

La fonction de resynchronisation rapide des miroirs (FMR, *Fast Mirror Resynchronization*) effectue une resynchronisation rapide et efficace des miroirs stale (non valides) en augmentant l'efficacité de la prise de vue VxVM afin d'améliorer la prise en charge des opérations telles que la sauvegarde et la prise de décision. Généralement, ces opérations nécessitent que le volume de stockage des données soit au repos et/ou qu'un accès secondaire au stockage n'affecte ou n'empêche pas l'accès principal (débit, mise à jour, cohérence, etc.).

Pour parvenir à ces objectifs, VxVM offre un mécanisme de *prise de vue*, qui crée une copie exacte d'un volume principal à un moment donné dans le temps. La prise de vue est ensuite accessible indépendamment du volume à partir duquel elle a été effectuée. Dans

un environnement VxVM partagé/de clusters, il est possible d'éliminer le conflit de ressources et le délai d'utilisation de la prise de vue en y accédant à partir d'un autre système.

La resynchronisation rapide des miroirs permet d'éviter certains des inconvénients suivants qui caractérisaient la version antérieure du mécanisme de prise de vue :

- ◆ Une fois la prise de vue effectuée, des différences peuvent exister entre celle-ci et le volume principal. Ils ne sont plus cohérents. Par conséquent, la prise de vue doit être ignorée lorsqu'elle n'est plus utile et une nouvelle doit être exécutée pour obtenir une copie actualisée des données principales.
- ◆ La durée de configuration de la prise de vue risque de réduire l'utilité de la fonction. En effet, la période pendant laquelle une prise de vue peut être créée est directement proportionnelle à la taille du volume. Dans le cas de volumes importants utilisés à l'échelle de l'entreprise, cette période peut entraîner la mise en place de stratégies de mise hors ligne inacceptables.

Composants de resynchronisation rapide des miroirs

La resynchronisation rapide des miroirs apporte deux améliorations fondamentales à VxVM. La première consiste à optimiser la procédure de resynchronisation des miroirs (*Resynchronisation rapide des miroirs*) et la seconde étend le modèle de prise de vue (*Reconnexion rapide des miroirs*) afin de fournir une méthode permettant d'actualiser et de réutiliser les prises de vue, au lieu de les ignorer.

Composant Resynchronisation rapide des miroirs

La resynchronisation rapide des miroirs nécessite le suivi des mises à jour des stockages de données manquées par les miroirs qui n'étaient pas disponibles au moment de ces mises à jour. Lorsqu'un miroir est de nouveau en service, la resynchronisation rapide des miroirs doit utiliser uniquement les mises à jour que ce miroir a manquées. Avec la resynchronisation rapide des miroirs, cette procédure nécessite une opération de restauration nettement inférieure à celle de la méthode courante, dans laquelle le stockage de données est entièrement copié sur le miroir de nouveau mis en service.

Un miroir devient indisponible lorsqu'il a été *détaché* de son volume, soit automatiquement (par VxVM à la suite d'une erreur), soit directement par un administrateur (via un utilitaire VxVM tel que `vxplex` ou `vxassist`). Un *miroir redevient en service* lorsqu'après avoir été détaché, il est rattaché à son volume d'origine à la suite de l'exécution de `vxrecover` ou de `vxplex att`.

La resynchronisation rapide des miroirs ne modifie pas la défaillance du miroir actuelle et ne répare pas le modèle d'administration. Le seul effet visible est que les opérations classiques de réparation des miroir se terminent plus rapidement.

Resynchronisation rapide des miroirs

L'amélioration de la resynchronisation permet à l'administrateur d'activer/désactiver la resynchronisation rapide des miroirs volume par volume et de vérifier son état.

Composant Reconnexion rapide des miroirs

La *reconnexion rapide des miroirs* augmente le modèle existant d'utilisation des prises de vue. Sans la resynchronisation rapide des miroirs, une copie indépendante d'un volume est créée au moyen du mécanisme de prise de vue. Le volume d'origine et son volume réplica sont totalement indépendants l'un de l'autre et leurs données peuvent diverger.

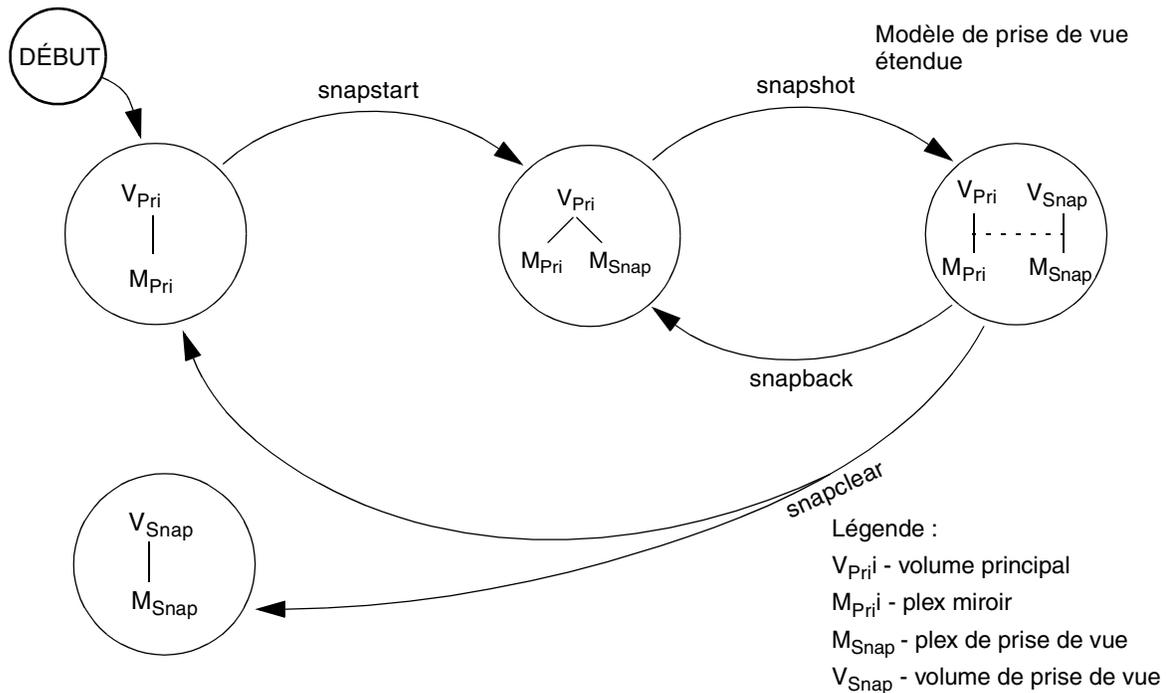
L'amélioration de la prise de vue avec *reconnexion rapide des miroirs* permet d'associer de nouveau un volume de prise de vue à son homologue d'origine afin de réduire spécifiquement la charge de travail nécessaire à l'exécution d'opérations cycliques reposant en grande partie sur la fonction de prise de vue VxVM.

Améliorations apportées par la resynchronisation rapide des miroirs à la fonction de prise de vue VxVM

Les améliorations de la prise de vue FMR apportées à la version 3.1 étendent le modèle de prise de vue comme illustré à la figure 23, « Prise de vue étendue FMR », à la page 77. À partir de la version 3.1, la commande `snapshot` peut être utilisée comme auparavant, sauf qu'elle associe le volume d'origine au volume de prise de vue. Une nouvelle commande, `vxassist snapshot`, utilise cette association pour renvoyer rapidement le plex de prise de vue (M_{Snap}) au volume à partir duquel la prise de vue a été effectuée (dans cet exemple, V_{Pri}).

La figure 23, « Prise de vue étendue FMR », illustre les transitions étendues du modèle de prise de vue créé par les commandes `snapshot` et `snapshot clear`.

Figure 23. Prise de vue étendue FMR



De plus, une nouvelle commande, `vxassist snapclear`, libère un volume de la charge administrative que représente le suivi d'une prise de vue en détruisant définitivement l'association créée par la commande `snapshot`. Cette fonction est utile lorsque vous savez qu'une prise de vue ne retournera jamais au volume à partir duquel elle a été créée.

Théorie de fonctionnement

La base de la resynchronisation rapide des miroirs repose sur le suivi des modifications. Le suivi des mises à jour manquées lorsqu'un miroir est hors ligne/détaché/en prise de vue, puis l'application de seulement ces mises à jour lorsque le miroir est de nouveau opérationnel réduit considérablement la durée de synchronisation du volume.

Ce suivi des modifications repose sur l'utilisation d'une bitmap. Chaque bit de la bitmap représente une zone contiguë (une étendue) de l'espace d'adressage d'un volume. On appelle *taille de zone* cette zone contiguë. Généralement, cette taille de zone occupe un seul bloc ; chaque bloc du volume est représenté par un bit de la bitmap. Cependant, la fonction réglable `vol_fmr_logsz` permet de limiter la taille maximale (en blocs) de la mappe FMR. Pour calculer la taille de la mappe, l'algorithme commence par une taille de

Resynchronisation rapide des miroirs

zone égale à 1 ; si la taille de mappe résultante est inférieure à `vol_fmr_logsz`, la valeur calculée devient la taille de mappe. Si elle est supérieure à `vol_fmr_logsz`, le système essaie d'utiliser `vol_fmr_logsz` avec une taille de zone de 2, et ainsi de suite jusqu'à ce que la taille de mappe soit inférieure à celle (réglable) de `vol_fmr_logsz`.

Par exemple :

taille de volume = 1 Go

`vol_fmr_logsz` = 4

sur un système ayant une taille de bloc de 512 octets, c'est-à-dire, $4 \times 512 = 2048$ octets ou $2048 \times 8 = 16384$ bits.

Par conséquent, pour un volume de 1 Go, une taille de zone de 1 correspond à 2048 bits, ce qui est inférieur à 4 blocs ; la taille de mappe est donc 2048 bits ou 256 octets.

Si la taille du volume augmente, le système effectue une nouvelle fois ce calcul afin de vérifier que la taille de mappe ne dépasse pas celle (réglable) de `vol_fmr_logsz`.

Suivi permanent et suivi volatile

Dans VxVM 3.1, les mappes FMR sont allouées en mémoire. Contrairement au journal des zones modifiées, elles ne se trouvent pas sur le disque ou dans un stockage permanent. Par conséquent, en cas d'arrêt anormal du système, ces informations sont perdues et il faut synchroniser entièrement la longueur du volume. Cette approche présente l'avantage suivant : les mises à jour FMR (les mises à jour de cette mappe) ne diminuent pas les performances car les disques ne doivent pas être mis à jour. Cependant, en cas d'arrêt anormal du système, ces informations sont perdues et vous devrez de nouveau resynchroniser entièrement les miroirs.

Prises de vue et resynchronisation rapide des miroirs

Pour tirer parti du suivi delta FMR avec les prises de vue, utilisez la nouvelle option de prise de vue. Une fois la prise de vue effectuée, l'option de prise de vue permet de rattacher le plex de prise de vue. Si la resynchronisation rapide des miroir est activée avant l'exécution de la prise de vue et qu'elle n'est pas désactivée avant que la prise de vue ne soit terminée, les modifications delta FMR reflétées dans la bitmap FMR permettent de resynchroniser le volume pendant le snapback. Pour faciliter la création simultanée de prises de vue pour plusieurs volumes, l'option de prise de vue a été améliorée de façon à accepter plusieurs volumes et un modèle d'attribution de nom a été ajouté. Par défaut, chaque volume réplica est nommé `SNAP-<volume d'origine>`. Vous pouvez utiliser les options de la ligne de commande pour remplacer cette valeur par défaut. Seuls les volumes du même disk group peuvent prendre une prise de vue de plusieurs volumes en même temps.

Pour créer plus facilement des prises de vue de tous les volumes d'un même disk group, une nouvelle option a été ajoutée à `vxassist`, mais elle échoue si l'un des volumes du disk group ne possède pas un plex de prise de vue complet. Vous pouvez effectuer plusieurs prises de vue du même volume. Une nouvelle bitmap FMR est créée pour chaque prise de vue, ce qui réduit la durée de resynchronisation pour chaque prise de vue.

Vous pouvez spécifier le plex de prise de vue comme ensemble de données préféré lorsque vous effectuez un `snapback`. L'ajout de `-o resyncfromreplica` à l'option de `snapback` copie les données du plex de prise de vue (réplica) sur tous les miroirs attachés au volume d'origine. Par défaut, les données sur le volume d'origine sont considérées comme préférées et elles sont copiées sur le plex de prise de vue.

Vous pouvez augmenter le volume réplica ou le volume d'origine, tout en continuant à utiliser la resynchronisation rapide des miroirs. L'augmentation du volume étend la bitmap qui est utilisée par FMR pour assurer le suivi des modifications delta. Cette opération peut modifier la taille de la bitmap ou celle de sa zone. Dans les deux cas, la partie de la bitmap qui correspond à la zone augmentée du volume est marquée comme modifiée et cette zone est donc resynchronisée. L'opération de prise de vue échoue si la prise de vue tente de créer un plex de prise de vue incomplet. Dans ce cas, il est nécessaire d'augmenter le volume réplica ou le volume d'origine avant d'exécuter l'option de `snapback`. L'augmentation séparée des deux volumes peut créer une prise de vue qui partage des disques physiques avec un autre miroir du volume. Pour éviter ce problème, augmentez le volume une fois que l'exécution de la commande `snapback` est terminée.

Toutes les opérations modifiant la structure du volume réplica peuvent marquer comme modifiée la mappe FMR pour cette prise de vue et nécessiter une resynchronisation complète pendant le `snapback`. Il s'agit des opérations suivantes : la division de sous-disques, leur déplacement et la restructuration en ligne du réplica. L'exécution de ces opérations est tout à fait sûre une fois la prise de vue effectuée. Pour plus d'informations, voir les pages `man vxvol (1M)`, `vxassist (1M)` et `vxplex (1M)`.

Resynchronisation rapide des miroirs et prises de vue accessibles en écriture

Vous pouvez choisir les deux options suivantes pour assurer le suivi des modifications apportées à une prise de vue accessible en écriture :

- ◆ créer une mappe séparée qui assure le suivi des modifications d'un volume de prise de vue ;
- ◆ mettre à jour la mappe du parent du volume de prise de vue. Utilisez uniquement cette méthode de raccourci si le volume de prise de vue a fait l'objet d'un nombre limité de mises à jour, par exemple, dans les applications de sauvegarde et de système de prise de décision (DSS, *Decision Support Systems*).

La deuxième méthode est implémentée dans le cas de VxVM 3.1 ; c'est-à-dire, la mappe du parent du volume de prise de vue est mise à jour pendant l'écriture sur la prise de vue.

Contrôle racine de Volume Manager

Avertissements et limitations

La resynchronisation rapide des miroirs n'est pas prise en charge sur les volumes RAID 5.

Lorsqu'un sous-disque est reconstruit, le plex entier est marqué modifié et une resynchronisation complète s'impose alors.

Contrôle racine de Volume Manager

Divers fichiers provenant de systèmes d'exploitation différents (par exemple, le système de fichiers racine, le périphérique `swap`, le système de fichiers `usr` et le système de fichiers `stand`) peuvent être placés sous le contrôle de Volume Manager. Le *disque racine* (c'est-à-dire, le disque contenant le système de fichiers racine) peut être placé sous le contrôle de Volume Manager via l'*encapsulation*.

L'encapsulation convertit en volumes des partitions existant sur ce disque. Lorsque les périphériques `root` et `swap` sont sous le contrôle de Volume Manager, ils apparaissent sous la forme de volumes et présentent les mêmes caractéristiques que les autres volumes Volume Manager. Un volume configuré de façon à être utilisé dans une zone de `swap` est appelé *volume de swap* ; un volume contenant le système de fichiers racine est appelé *volume racine* ; un volume contenant le système de fichiers `stand` est appelé *volume stand*.

Il est possible de mettre en miroir les volumes `rootvol`, `swapvol` et `standvol`, ainsi que d'autres parties du disque racine requises pour assurer le bon démarrage du système (par exemple, `/usr`). Cela permet la redondance et la récupération complète en cas de défaillance du disque. Sans le contrôle racine de Volume Manager, la perte de la partition `root`, `swap`, `usrroot` ou `stand` empêche de démarrer le système à partir des disques restants.

La mise en miroir des lecteurs de disques nécessaires à l'initialisation assure qu'aucune défaillance de disque rend le système inutilisable. Une configuration suggérée consiste à mettre en miroir le disque critique sur un autre disque disponible (au moyen de la commande `vxdiskadm`). Si le disque qui contient les partitions `root`, `stand` et `swap` échoue, vous pouvez redémarrer le système à partir du disque contenant le miroir racine. Pour plus d'informations sur les procédures de mise en miroir du disque d'initialisation (racine) et de récupération du système, reportez-vous au chapitre 7, « Récupération ».

Démarrage avec des volumes racine

Généralement, au démarrage du système d'exploitation, le système de fichier `root`, le système de fichiers `stand` et la zone `swap` peuvent être utilisés dès le début de la procédure, c'est-à-dire avant que les processus utilisateur puissent être exécutés pour charger la configuration Volume Manager et démarrer les volumes. Les configurations de périphériques `root`, `stand` et `swap` doivent être terminées avant le démarrage de Volume Manager. Le lancement du démon de Volume Manager `vxconfigd` dans le cadre de la procédure `init` est trop tardif pour configurer les volumes afin de les utiliser en tant que périphérique `root` ou `swap`.

Pour éviter cette limitation, le système peut accéder aux volumes `rootvol`, `standvol` et `swapvol` pendant le démarrage. Le système considère alors que ces volumes sont des partitions normales et il peut y accéder à l'aide de la numérotation de partition standard. Les volumes `rootvol`, `standvol` et `swapvol` sont créés à partir d'un espace disque contigu qui est également mappé par une seule partition pour chacun. Cette limitation ne permet pas de répartir par bandes ou de répartir le plex principal (le plex utilisé au démarrage) d'un volume `rootvol`, `standvol` ou `swapvol`. Les miroirs de ces volumes nécessaires au démarrage ne peuvent pas être répartis par bandes ou répartis.

Limitations des volumes au démarrage

Les volumes `rootvol`, `standvol`, `swapvol` et `usr` diffèrent des autres volumes dans la mesure où leur configuration est soumise à des restrictions très spécifiques :

- ◆ Le volume racine (`rootvol`) doit exister dans le `disk group` par défaut, `rootdg`. Bien que vous puissiez créer d'autres volumes nommés `rootvol` dans des `disk groups` différents de `rootdg`, vous ne pouvez utiliser que le volume `rootvol` de `rootdg` pour démarrer le système.
- ◆ Le volume `rootvol` a un mineur de périphérique spécifique : le mineur de périphérique 0. De même, `swapvol` a le mineur de périphérique 1. Le volume `usr` n'a aucun mineur de périphérique spécifique. Reportez-vous au chapitre 7, « Récupération ».
- ◆ Des partitions de « recouvrement » sont créées pour les miroirs limités des périphériques `rootvol`, `var`, `usrrootvol`, `standvol` et `swapvol`. Une partition de recouvrement inclut l'espace disque exact occupé par le miroir limité. Pendant le démarrage, avant que les volumes `rootvol`, `var`, `usrrootvol`, `standvol` et `swapvol` ne soient totalement configurés, la configuration de volume par défaut utilise la partition de recouvrement pour accéder aux données du disque.
- ◆ Bien qu'il soit possible d'ajouter un miroir réparti par bandes à un périphérique `rootvol` pour des raisons de performances, vous ne pouvez pas répartir par bandes le plex principal ni aucun miroir de `rootvol` nécessaires à la récupération du système ou au démarrage en cas de défaillance du plex principal.

Fonction DMP

- ◆ `rootvol`, `standvol` et `swapvol` ne peuvent pas être répartis ni contenir un plex principal comprenant plusieurs sous-disques non contigus. Vous ne pouvez ni augmenter ni réduire un volume associé à un disque de d'initialisation encapsulé (`rootvol`, `usr`, `var`, `opt`, `swapvol`, etc.) car ces disques sont mappés à une partition sous-jacente physique sur le disque et ils doivent être contigus. Une solution consiste à désencapsuler le disque d'initialisation, à le repartitionner en fonction de vos spécifications (en augmentant ou réduisant les partitions, selon le cas), puis à encapsuler de nouveau.
- ◆ Lorsque vous mettez en miroir des parties du disque d'initialisation, le disque mis en miroir doit disposer d'un espace suffisant pour recevoir les données du plex d'origine, faute de quoi la mise en miroir risque d'échouer.
- ◆ `rootvol`, `standvol`, `swapvol` et `usr` ne peuvent pas être utilisés en tant que volumes pour le journal des zones modifiées.

Outre ces spécifications, il est recommandé de disposer au moins d'un miroir contigu (avec des cylindres alignés, le cas échéant), pour chaque volume de `root`, `usr`, `var`, `opt`, `varadm`, `usrkvm` et `swap`. Cela facilite la conversion de ces volumes en partitions de disque normales (par exemple, pendant la mise à niveau d'un système d'exploitation).

Fonction DMP

Sur certains systèmes, Volume Manager prend en charge les ensembles de disques multiports. Le programme reconnaît automatiquement les chemins d'E/S multiples vers un périphérique de disque spécifique dans l'ensemble de disques. La fonction de chemins d'accès multiples (DMP, *Dynamic Multipathing*) de Volume Manager assure une meilleure fiabilité au moyen d'un mécanisme de basculement (failover) de chemin. Si le système perd une connexion à un disque, il continue d'accéder aux données critiques sur les autres connexions valides du disque. La fonction DMP assure également un débit supérieur pour les E/S en équilibrant leur charge de façon régulière sur les différents chemins d'E/S au périphérique de disque.

Dans Volume Manager, les disques physiques connectés au système sont tous représentés sous la forme de méta-périphériques ayant un ou plusieurs chemins d'accès physiques. Un seul disque physique connecté au système est représenté par un méta-périphérique avec un seul chemin. Un disque inclus dans un ensemble de disques est représenté par un méta-périphérique avec deux chemins d'accès physiques. Vous pouvez utiliser des utilitaires d'administration de Volume Manager tels que `vxdisk` pour afficher tous les chemins d'un méta-périphérique et les informations d'état des divers chemins.

Des chemins d'accès multiples peuvent exister lorsqu'un disque physique est accessible par plusieurs identificateurs de périphérique du système d'exploitation. Chaque identificateur de périphérique du système d'exploitation à chemins d'accès multiples permet d'accéder et de contrôler des données au moyen d'autres chemins hôte-périphérique.

Vous pouvez configurer Volume Manager avec son propre système DMP afin d'organiser l'accès à des périphériques à chemins d'accès multiples. Volume Manager détecte les systèmes à chemins d'accès multiples au moyen des identificateurs de périphérique universels (WWD ID, *Universal World-Wide-Device Identifiers*). Le disque physique doit fournir une identification sans équivoque au moyen de son ID WWD pour que la fonction DMP puisse accéder au périphérique.

Si la fonction DMP ne peut pas identifier le disque physique au moyen de son ID WWD, l'identification est effectuée à l'aide des méthodes de détection de périphérique de Volume Manager. La détection des périphériques dépend de la reconnaissance par Volume Manager des identificateurs de métadonnées sur le disque.

La fonction DMP de Volume Manager crée des métanœuds représentant les méta-périphériques pour chaque cible à chemins d'accès multiples qu'il a détectée. Chaque métanœud est mappé à un ensemble d'identificateurs de périphérique du système d'exploitation et est configuré à l'aide d'une stratégie appropriée de chemins d'accès multiples. La fonction DMP crée des métanœuds pour tous les disques physiques attachés et accessibles à l'aide d'un identificateur de périphérique du système d'exploitation.

La fonction DMP gère des cibles à chemins d'accès multiples, comme les ensembles de disques, qui définissent les stratégies permettant d'utiliser plusieurs chemins. Certains ensembles de disques permettent d'activer simultanément plusieurs chemins (Actif/Actif). D'autres ensembles de disques ne permettent d'activer qu'un seul chemin et conservent un autre chemin comme chemin de remplacement en cas de défaillance du chemin existant (Actif/Passif). Certains ensembles de disques reposent sur des stratégies plus évoluées.

Généralement, Volume Manager est conçu pour que le disque VM soit mappé sur un métanœud Volume Manager DMP. Pour simplifier les opérations logiques de VxVM, chaque disque VM est mappé sur un métanœud Volume Manager DMP unique. Le mappage est effectué que le périphérique de disque physique soit ou non connecté à une configuration à chemins multiples.

Mécanisme de basculement de chemin

La fonction DMP améliore la fiabilité du système lorsque vous l'utilisez avec des ensembles de disques multiports. En cas de perte d'une connexion à l'ensemble de disques, elle sélectionne automatiquement et de façon dynamique les chemins d'E/S suivants pour les requêtes d'E/S, sans aucune intervention de l'administrateur.

La fonction DMP permet à l'administrateur d'indiquer au sous-système DMP de Volume Manager si la connexion est réparée ou restaurée. Il s'agit de la reconfiguration DMP. La procédure de reconfiguration permet également de détecter les périphériques récemment ajoutés et ceux qui sont supprimés une fois que le système a redémarré (uniquement si le système d'exploitation les détecte).

Équilibrage de charge

Pour assurer un bon équilibrage de la charge entre les chemins, la fonction DMP suit le *mécanisme de chemin équilibré* pour les ensembles de disques de type actif/actif. Avec l'équilibrage de charge, le débit des E/S peut être augmenté à la valeur maximale en utilisant la largeur de bande maximale de tous les chemins. Cependant, les E/S séquentielles sur un disque utilisent le même chemin dans la fonction DMP afin d'optimiser le débit des E/S. Cela permet de tirer parti des données mises en cache sur le contrôleur de disques.

Pour les ensembles de disques de type actif/passif, les E/S utilisent le chemin principal jusqu'à ce qu'il échoue. Lorsque cela se produit, les E/S utilisent alors les autres chemins principaux ou secondaires disponibles. Pour éviter de transférer continuellement la propriété des numéros d'unités logiques d'un contrôleur à l'autre, ce qui se traduirait par un fort ralentissement des E/S, l'équilibrage de charge entre les chemins n'est pas utilisé pour les ensembles de disques de type actif/passif.

Démarrage à partir des périphériques DMP

Lorsque le disque racine est placé sous le contrôle de Volume Manager, il est automatiquement accessible en tant que périphérique DMP avec un chemin s'il ne s'agit que d'un seul disque ou de plusieurs chemins si le disque fait partie d'un ensemble de disques multiports. L'encapsulation du disque racine vous permet d'améliorer la fiabilité du système en cas de perte d'un ou de plusieurs des chemins physiques existant sur un disque.

Activation et désactivation des contrôleurs

La fonction DMP permet à l'administrateur de désactiver un contrôleur d'E/S hôte pour exécuter des opérations d'administration. Il permet de maintenir des contrôleurs attachés à l'hôte ou à un ensemble de disques pris en charge par Volume Manager. Les opérations d'E/S sur le contrôleur d'E/S hôte peuvent être activées une fois que la tâche de maintenance est terminée. Pour cela, vous pouvez utiliser la commande `vxdmpadm` de Volume Manager.

Par exemple, si le système est équipé d'un ensemble de disques StorEdge A5000TM et que l'utilisateur doit changer une carte d'interface A5000 connectée à cet ensemble de disques, la commande `vxdmpadm` vous permet d'obtenir la liste des contrôleurs d'E/S hôtes connectés à cette carte d'interface. Vous devez désactiver ces contrôleurs. D'autres E/S vers ces disques qui sont ensuite accessibles via ces contrôleurs sont arrêtées.

Vous pouvez ensuite remplacer la carte d'interface sans interrompre les E/S en cours sur les disques de cet ensemble de disques. Cela est nécessaire car, normalement, pour les ensembles de disques de type actif/actif (comme dans l'exemple ci-dessus),

Volume Manager utilise le mécanisme de chemin équilibré pour programmer les E/S vers un disque à chemins d'accès multiples. Par conséquent, les E/S peuvent suivre n'importe quel chemin à n'importe quel moment donné.

Pour les ensembles de disques de type actif/passif, Volume Manager programme toujours les E/S vers le chemin principal jusqu'à ce qu'une défaillance se produise. Par conséquent, pour modifier une carte d'interface sur l'ensemble de disques ou une carte sur l'hôte (le cas échéant) connectée à l'ensemble de disques, vous devez désactiver les opérations d'E/S vers le ou les contrôleurs d'E/S hôtes. Cela permet de rediriger toutes les E/S vers un chemin secondaire actif ou vers un chemin principal actif sur un autre contrôleur d'E/S avant de changer le composant matériel.

Une fois l'opération terminée, vous pouvez remettre en service les chemins utilisant ce ou ces contrôleurs à l'aide de l'option `enabled` de la commande `vxddmpadm`.

Volume Manager ne vous permet pas de désactiver le dernier chemin actif vers le disque racine.

Affichage des informations de la base de données DMP

Vous pouvez utiliser la commande `vxddmpadm` pour lister les informations de la base de données DMP et pour exécuter d'autres tâches d'administration. Cette commande vous permet de lister tous les contrôleurs sur les systèmes (connectés aux disques) et d'autres informations connexes stockées dans la base de données DMP. Vous pouvez utiliser ces informations pour déterminer l'emplacement des composants matériels du système et décider des contrôleurs à activer/désactiver.

`vxddmpadm` fournit également des informations utiles telles que le numéro de série de l'ensemble de disques et la liste des périphériques DMP (disques) connectés à l'ensemble de disques, la liste des chemins utilisant un contrôleur spécifique, etc.

VxSmartSync Recovery Accelerator

La fonction VxSmartSync™ Recovery Accelerator est disponible pour certains systèmes. VxSmartSync for Mirrored Oracle Databases est un ensemble de fonctions qui rendent la procédure de resynchronisation (aussi appelée « *resilvering* ») plus rapide pour les volumes utilisés avec Oracle Universal Database. Ces fonctions utilisent une interface étendue entre les volumes Volume Manager et le logiciel de base de données de façon à éviter les opérations inutiles pendant la resynchronisation des miroirs. Ces extensions peuvent se traduire par une réduction notable de la durée de récupération des volumes.

Oracle tire automatiquement parti de la fonction SmartSync lorsqu'elle est disponible.

VxSmartSync Recovery Accelerator

L'administrateur système doit configurer correctement les volumes pour utiliser VxSmartSync. Pour Volume Manager, la base de données utilise deux types de volumes :

- ◆ Les *volumes de journal redo* contiennent les journaux redo de la base de données.
- ◆ Les *volumes de données* sont tous les autres volumes utilisés par la base de données (fichiers de contrôle et fichiers tablespace).

VxSmartSync utilise de façon différente ces deux types de volumes, qui doivent être configurés correctement pour tirer le meilleur parti des interfaces étendues. La seule différence entre les deux types de volumes est que, contrairement aux volumes de données, les volumes de journal redo doivent avoir des journaux des zones modifiées.

Configuration des volumes de données

La durée de récupération des volumes de données peut être réduite en laissant le logiciel de base de données choisir les portions du volume qui nécessitent une récupération. La base de données conserve des journaux des modifications apportées aux données et elle peut déterminer les portions du volume qui nécessitent une récupération. En réduisant la quantité d'espace qui doit être récupérée et en permettant à la base de données de contrôler la récupération, il devient possible de réduire la durée totale de la récupération.

De plus, la récupération est effectuée au démarrage du logiciel de base de données, et non au démarrage du système. Cela réduit l'incidence générale de la récupération au redémarrage du système. La durée de récupération du volume correspond à la durée de resilvering pour la base de données (c'est-à-dire, la durée requise pour relire les journaux redo) car la récupération est contrôlée par la base de données.

Volume Manager peut éviter de procéder à la journalisation car la base de données conserve ses propres journaux. Les volumes de données doivent donc être configurés en tant que volumes mis en miroir *sans* journaux des zones modifiées. Outre une récupération plus rapide, cela permet d'éviter également les délais d'E/S à l'exécution dus à la fonction DRL, ce qui améliore l'accès en écriture normal à la base de données.

Configuration des volumes de journal redo

Un *journal redo* est un journal des modifications apportées aux données de la base de données. Celle-ci ne peut pas fournir des informations sur les sections qui nécessitent le resilvering car elle ne conserve pas de journaux des modifications apportées aux journaux redo. Les journaux redo sont également écrits de façon séquentielle. Les journaux des zones modifiées classiques sont davantage utiles pour les données écrites de façon aléatoire que pour diminuer la durée de récupération des journaux redo. Cependant, Volume Manager peut réduire le nombre de zones modifiées en modifiant le comportement de sa fonction DRL afin d'utiliser les modèles d'accès séquentiel. Cela réduit la quantité des données à récupérer ainsi que l'incidence de la durée de récupération sur le système.

Les interfaces améliorées des journaux redo permettent au logiciel de base de données d'informer Volume Manager qu'un volume sera utilisé comme journal redo. Volume Manager peut ainsi modifier le comportement DRL du volume afin d'utiliser des modèles d'accès. Les volumes de journal redo devraient être configurés comme des volumes mis en miroir *avec* des journaux des zones modifiées car la durée de récupération améliorée dépend des journaux des zones modifiées.

Moniteur de tâches Volume Manager

Le moniteur de tâches Volume Manager assure le suivi de la progression de la récupération du système en surveillant la création, la maintenance et la terminaison des tâches. Il vous permet de surveiller la progression des tâches et de modifier leurs caractéristiques, notamment l'interruption et le taux de récupération (par exemple, pour réduire l'incidence des performances sur le système). Vous pouvez également surveiller et modifier la progression de la fonction de restructuration en ligne. Pour plus d'informations, voir « Restructuration en ligne » à la page 63.

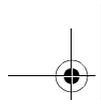
Fonction de cluster Volume Manager

Volume Manager offre une fonction de cluster *facultative* qui permet d'utiliser VxVM dans un environnement cluster. Cette fonction permet à plusieurs hôtes d'accéder et de gérer simultanément un jeu de disques donné sous le contrôle de Volume Manager (*disques VM*).

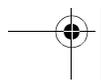
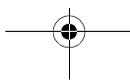
Un *cluster* est un groupe d'hôtes partageant un jeu de disques ; chaque hôte est un *nœud* du cluster. Les nœuds sont connectés sur un réseau. Si un nœud échoue, les autres nœuds peuvent continuer d'accéder aux disques. La fonction de cluster Volume Manager offre la même vue logique des configurations de disques (y compris les modifications) sur tous les nœuds. Lorsque la fonction de cluster est activée, les objets Volume Manager peuvent être partagés par tous les nœuds d'un cluster.

Pour plus d'informations sur la fonction de cluster dans Volume Manager, reportez-vous au chapitre approprié dans ce manuel.

Remarque La fonction de cluster Volume Manager fait l'objet d'une licence séparée.



Fonction de cluster Volume Manager



Opérations de disque

4

Introduction

Ce chapitre décrit les opérations de gestion de disques effectuées par Volume Manager. Il contient des informations sur les opérations de disk group.

Remarque Les commandes Volume Manager exigent généralement les privilèges de super-utilisateur ou d'autres privilèges appropriés.

Ce chapitre traite des sujets suivants :

- ◆ Périphériques de disque standard
- ◆ Disk groups
- ◆ Commandes de disque et de disk group
- ◆ Initialisation et ajout de disques
- ◆ Ajout d'un disque à Volume Manager
- ◆ Ajout d'un disque à un disk group
- ◆ Suppression d'un disque d'un disk group
- ◆ Déplacement des disques
- ◆ Attribution d'un nouveau nom à un disque
- ◆ Réserveation des disques
- ◆ Mise hors ligne d'un disque
- ◆ Mise en miroir d'un disque
- ◆ Suppression d'un disque
- ◆ Affichage des informations sur le disque
- ◆ Détection et remplacement des disques qui ont échoué
- ◆ Création d'un disk group
- ◆ Mise à niveau d'un disk group

Périphériques de disque standard

- ◆ Suppression d'un disk group
- ◆ Déplacement des disk groups d'un système à un autre
- ◆ Destruction d'un disk group
- ◆ Utilisation de périphériques spéciaux
- ◆ Opérations de l'interface de menu vxdiskadm

Périphériques de disque standard

Il est possible d'utiliser deux classes de périphériques de disque avec Volume Manager : des périphériques standard et des périphériques spéciaux. Les périphériques spéciaux sont décrits dans la suite de ce chapitre.

Volume Manager gère un maximum de huit partitions sur un disque physique. Ces partitions sont nommées, dans l'ordre, de 0 à 7. La partition 2 est réservée et indique le disque entier.

Remarque Sur certains systèmes, Volume Manager peut gérer un maximum de 16 partitions. Dans ce cas, les partitions sont nommées de 0 à 15 et la partition 0 est réservée pour indiquer le disque entier.

Lorsqu'une partition est placée sous le contrôle de Volume Manager, un disque VM est assigné à cette partition. Vous pouvez utiliser un nom symbolique (le *nom du disque* ou le *nom d'accès logique au disque*), par exemple, `disk0`, pour indiquer un disque VM.

Remarque Votre système peut utiliser un *nom de périphérique* différent des exemples.

Une partition est référencée à l'aide d'une adresse physique (généralement appelée *nom du périphérique* ou nom d'accès au disque) dont le format est `c#b#t#d#s#` et qui comprend les éléments suivants :

- ◆ `c#` — Le numéro du contrôleur auquel l'unité de disque est attachée ;
- ◆ `b#` — Le bus correspondant (s'il est utilisé sur le système) ;
- ◆ `t#` et `d#` — L'ID cible et le numéro de périphérique qui constituent l'adresse de l'unité de disque sur le contrôleur ;
- ◆ `s#` — Le numéro de partition sur l'unité de disque.

Par exemple, le nom d'un périphérique peut être `c0t0d0s2`. Par convention, `s2` fait référence à la méthode de partitionnement standard utilisée par Volume Manager. Sur certains systèmes, Volume Manager utilise `s0` comme méthode de partitionnement standard. Le disque physique est identifié auprès de Volume Manager sous la forme `c#b#t#d#s#` (`b#` est réservé aux systèmes qui utilisent un bus).

Pour la plupart des commandes, le suffixe `s#` est facultatif, même si les commandes d'affichage signalent les périphériques ayant ce suffixe. Les utilitaires `vxdiskadm` et `vxdiskadd` de Volume Manager utilisent les noms de périphériques sans le suffixe `s2` (ou `s0`). Par exemple, pour spécifier le second disque attaché au premier contrôleur de `vxdiskadd`, utilisez le nom `c0t1d0`.

Le disque d'initialisation (qui contient le système de fichiers racine et sert au démarrage du système) est souvent identifié par Volume Manager sous le nom de périphérique `c0t0d0`.

Un disque VM possède deux zones :

- ◆ *zone privée* — Petite zone contenant les informations de configuration. Cette zone stocke un libellé de disque et des enregistrements de configuration ;
- ◆ *zone publique* — Zone couvrant le reste du disque. Elle sert au stockage des sous-disques (et à l'allocation de l'espace de stockage).

Volume Manager utilise trois types de disques de base :

- ◆ *partitionné* — Les zones publique et privée se trouvent sur des partitions de disque différentes ;
- ◆ *simple* — Les zones publique et privée se trouvent sur la même partition de disque (la zone privée précède la zone publique) ;
- ◆ *nopriv* — Le disque ne contient pas de zone privée (uniquement une zone publique pour l'allocation des sous-disques).

Volume Manager initialise chaque nouveau disque en utilisant le moins possible de partitions (généralement deux partitions par disque physique). Pour les noms d'accès au disque se terminant par `s2` (ou `s0`), le type par défaut est `sliced` (partitionné).

Disk groups

Volume Manager organise les disques en disk groups. Un *disk group* est un ensemble de disques partageant une configuration commune ; son nom est spécifique. Les volumes sont créés dans un disk group et ils ne peuvent utiliser que les disques de ce disk group.

Tous les systèmes sur lesquels Volume Manager est installé comprennent le disk group par défaut `rootdg`. Les opérations sont dirigées par défaut vers ce disk group. Si nécessaire, l'administrateur système peut créer des disk groups supplémentaires. Le plus souvent, les systèmes n'utilisent qu'un disk group, sauf s'ils possèdent un grand nombre de disques. Les disques ne sont pas ajoutés aux disk groups, sauf s'ils sont nécessaires pour créer des objets Volume Manager. Les disques peuvent être par la suite initialisés, réservés et ajoutés aux disk groups. Cependant, il est nécessaire d'ajouter au moins un disque (partition) à `rootdg` pour pouvoir effectuer les procédures d'installation de Volume Manager.

Commandes de disque et de disk group

Lorsqu'il est ajouté à un disk group, le disque est nommé (par exemple, `disk02`). Ce nom l'identifie pour les opérations de volume, création ou mise en miroir, et fait directement référence au disque physique. Si un disque physique est déplacé vers une autre adresse cible ou un autre contrôleur, le nom `disk02` continue de renvoyer à ce disque. Pour remplacer un disque, vous devez d'abord associer un autre disque physique au nom du disque à remplacer, puis récupérer les données de volume éventuellement situées sur le disque d'origine (à partir de miroirs ou de copies de sauvegarde).

L'existence de disk groups volumineux risque de remplir la zone privée. Dans ce cas, vous devez configurer les disques en leur attribuant des zones privées plus importantes pour les opérations de journal. Une grande partie de la zone privée est réservée à la base de données de configuration des disk groups, qui contient les enregistrements propres à chaque objet Volume Manager inclus dans ce disk group. Le nombre d'enregistrements pouvant être créé dans un disk group correspond au double de la taille de la copie de la base de données de configuration car chaque enregistrement de configuration occupe 256 octets (ou un demi-bloc). La taille de la copie peut être obtenue à partir des informations affichées au moyen de la commande `vx dg list nom_diskgroup`.

Commandes de disque et de disk group

Volume Manager offre plusieurs interfaces qui vous permettent de gérer les disques :

- ◆ l'interface utilisateur graphique ;
- ◆ un ensemble de commandes pour la ligne de commandes ;
- ◆ l'interface basée sur les menus pour `vx diskadm` ;
- ◆ la commande « Rétablir un disque reconstruit dans le disque d'origine ».

Ce chapitre traite des commandes suivantes :

- ◆ `vx diskadm` — Il s'agit de l'interface de menu pour les opérations de gestion Volume Manager. Cette commande fournit le menu des opérations de disque. Chaque entrée du menu principal vous guide le long de l'exécution d'une tâche spécifique à l'aide d'informations et d'invites. Des réponses par défaut sont fournies pour la plupart des questions et vous pouvez choisir sans problème les réponses courantes. Pour plus d'informations sur l'utilisation de `vx diskadm`, voir la page man `vx diskadm(1M)`.
- ◆ `vx diskadd` — Cette commande permet d'ajouter des disques standard à Volume Manager. Elle vous guide le long de la procédure d'initialisation d'un nouveau disque par l'affichage d'informations et d'invites. Pour plus d'informations sur l'utilisation de `vx diskadd`, voir la page man `vx diskadd(1M)`.

- ◆ `vxdisk` — Cette commande gère les disques sous le contrôle de VxVM. Elle définit les périphériques de disque spéciaux, initialise les informations placées sur les disques (utilisées par Volume Manager pour identifier et gérer les disques) et exécute d'autres opérations spéciales. Pour plus d'informations sur l'utilisation de `vxdisk`, voir la page `man vxdisk(1M)`.
- ◆ `vx dg` — Cette commande est utilisée sur les disk groups. Elle crée de nouveaux disk groups et gère les disk groups existants. Pour plus d'informations sur l'utilisation de `vx dg`, voir la page `man vx dg(1M)`.
- ◆ `vxunrelocate` — Cette commande déplace les sous-disques qui ont été reconstruits sur leurs disques d'origine à l'aide de la fonction de reconstruction automatique.

Vous ne pouvez utiliser l'utilitaire `vx diskadd` et la plupart des opérations `vx diskadm` qu'avec les périphériques de disque standard.

Les commandes Volume Manager vous permettent souvent de spécifier un disk group au moyen de l'option `-g`. Par exemple, pour créer un volume dans le disk group `mkt dg`, utilisez la commande suivante :

```
# vxassist -g mkt dg make mktvol 50m
```

Le périphérique de volume (bloc) pour ce volume est :

```
/dev/vx/dsk/mkt dg/mktvol
```

Il est inutile de spécifier le disk group si les objets ont des noms uniques. Les commandes de Volume Manager utilisent souvent les noms d'objet spécifiés sur la ligne de commande pour déterminer le disk group pour l'opération. Par exemple, pour créer un volume sur le disque `mkt dg01` sans spécifier le nom du disk group, entrez la commande suivante :

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

Les commandes fonctionnent de cette façon tant que les objets de deux disk groups ont des noms différents. Par exemple, Volume Manager vous permet de créer des volumes nommés `mktvol` dans `root dg` et dans `mkt dg`. Dans ce cas, vous devez ajouter `-g mkt dg` à toute commande destinée à manipuler le volume placé dans le disk group `mkt dg`.

Initialisation et ajout de disques

Les disques sont soit initialisés, soit encapsulés lorsque vous les ajoutez à VxVM. L'encapsulation conserve les données existant sur les disques alors que l'initialisation les détruit.

Initialisation et ajout de disques

Il existe deux niveaux d'initialisation des disques dans Volume Manager :

1. Formatage du support de disque. Cette opération doit être effectuée en dehors de Volume Manager.
2. Stockage des informations d'identification et de configuration sur le disque pour que Volume Manager puisse les utiliser. Volume Manager offre des interfaces qui permettent d'effectuer ce niveau d'initialisation du disque.

Vous pouvez ajouter un disque totalement initialisé à un disk group et l'utiliser pour remplacer un disque qui a échoué ou pour créer un nouveau disk group. Ces sujets sont traités dans la suite de ce chapitre.

Formatage du support de disque

Pour exécuter la première phase d'initialisation, utilisez la commande interactive `format` (sur certains systèmes, `diskadd`) pour lancer le formatage du support de disque.

Remarque Les disques SCSI sont généralement préformatés. La commande `format` (ou `diskadd`) n'est nécessaire que si le format est gravement endommagé.

Installation de disques avec Volume Manager

Vous pouvez utiliser les menus `vxdiskadm` ou `vxdiskadd` pour lancer la phase d'initialisation du disque. Cette section décrit la procédure d'utilisation de `vxdiskadd`. Pour plus d'informations sur l'utilisation de `vxdiskadm` afin d'initialiser un seul disque ou tous les disques sur un contrôleur, voir « Opérations de l'interface de menu `vxdiskadm` » à la page 136.

Vous pouvez utiliser `vxdiskadd` pour initialiser un disque spécifique. Par exemple, pour initialiser le deuxième disque sur le premier contrôleur, utilisez la commande suivante :

```
# vxdiskadd c0t1d2
```

`vxdiskadd` examine le disque pour déterminer s'il a été initialisé et affiche des invites en fonction des résultats. `vxdiskadd` vérifie les disques qui peuvent être encapsulés (voir « Opérations de l'interface de menu `vxdiskadm` » à la page 136), ceux qui ont déjà été ajoutés à Volume Manager, et d'autres conditions.

Remarque Si le disque que vous ajoutez n'est pas initialisé, des messages d'avertissement et d'erreur s'affichent sur la console pendant l'exécution de `vxdiskadd`. Vous pouvez ignorer ces messages. En effet, ils ne devraient plus s'afficher après l'initialisation complète du disque ; à l'issue de l'initialisation, `vxdiskadd` affiche un message indiquant la réussite de l'opération.

À l'invite suivante, entrez **y** (ou appuyez sur la touche de retour) pour continuer :

```
Ajouter ou initialiser les disques
Menu : VolumeManager/Disk/AddDisks
```

```
Le disque sélectionné est le suivant. Format de sortie :
[Nom_périphérique]
```

```
c0t1d0
```

```
Continuer l'opération ? [y,n,q,?] (default: y) y
```

Si le disque n'est pas initialisé ou si vous avez choisi de le réinitialiser, l'invite suivante s'affiche :

```
Vous pouvez choisir d'ajouter ce disque à un disk group
existant, à un nouveau disk group ou de le laisser disponible
afin qu'il puisse être utilisé pour des opérations futures
d'ajout ou de remplacement. Pour créer un nouveau disk group,
sélectionnez le nom d'un disk group qui n'existe pas. Pour
que le disque soit disponible plus tard, nommez le disk group
"none".
```

```
Quel disk group [<group>,none,list,q,?] (default: rootdg)
```

Pour ajouter ce disque au disk group par défaut rootdg, appuyez sur la touche de retour. Pour laisser le disque disponible afin de l'utiliser comme disque de remplacement (ajouté à aucun disk group), entrez **none**. À l'issue de cette opération, vous êtes invité à sélectionner le nom que vous voulez attribuer au disque dans le disk group :

```
Utiliser un nom de disque par défaut pour le disque ?
[y,n,q,?] (default: y) y
```

Généralement, vous devriez accepter le nom du disque par défaut (sauf si vous préférez entrer un nom spécial).

À l'invite suivante, entrez **n** pour indiquer que ce disque ne doit pas servir de disque de remplacement pour la reconstruction automatique :

```
Ajouter un disque comme disque de remplacement pour rootdg ?
[y,n,q,?] (default: n) n
```

À l'invite suivante, entrez **n** pour indiquer que ce disque ne doit pas être exclu de la reconstruction automatique :

```
Exclure un autre disque de la reconstruction automatique ?
[y,n,q,?] [default:n] n
```

Ajout d'un disque à Volume Manager

Entrez **y** ou appuyez sur la touche de retour pour continuer l'opération après ce message :

```
Les disques sélectionnés seront ajoutés au disk group rootdg
avec les noms de disque définis par défaut.
```

```
c0t1d0
```

```
Continuer l'opération ? [y,n,q,?] (default: y) y
```

Si vous êtes certain que vous ne voulez pas enregistrer des données de ce disque, entrez **n** à l'invite suivante :

```
Le périphérique de disque suivant a un VTOC valide, mais ne
semble pas avoir été initialisé pour Volume Manager. Si le
disque contient des données, elle NE devraient PAS être
détruites ; vous devriez encapsuler les partitions du disque
existantes comme volumes au lieu d'ajouter le disque comme
nouveau disque. Format de sortie : [Nom_périphérique]
```

```
c0t1d0
```

```
Encapsuler ce périphérique ? [y,n,q,?] (default: y)
```

Lorsque vxdiskadm vous invite à initialiser le disque à la place, entrez **y** :

```
Au lieu d'encapsuler, initialiser ? [y,n,q,?]
(default: n) y
```

Des messages du type suivant devraient à présent confirmer que le disque c1t0d1 est placé sous le contrôle de Volume Manager. Selon le système, vous pouvez également choisir d'effectuer une analyse de surface.

```
Initialisation du périphérique c0t1d0.
```

```
Exécuter l'analyse de surface (fortement recommandé)
```

```
[y,n,q,?] (default: y) n
```

```
Ajout du périphérique de disque c0t1d0 au disk group rootdg
avec le nom de disque disk33.
```

Ajout d'un disque à Volume Manager

Vous devez placer un disque sous le contrôle de Volume Manager et l'ajouter à un disk group avant de pouvoir utiliser son espace pour les volumes. Si le disque était utilisé auparavant, mais n'était pas sous le contrôle de Volume Manager, vous pouvez conserver les données de ce disque tout en laissant Volume Manager prendre le contrôle de ce dernier. Pour cela, vous pouvez utiliser la fonction d'encapsulation de Volume Manager. L'encapsulation conserve toutes les données du disque. S'il s'agit d'un nouveau disque,

vous devez l'initialiser. L'initialisation détruit toutes les données du disque. Si le disque n'était pas auparavant sous le contrôle de Volume Manager, mais que vous ne voulez conserver aucune donnée, vous devez l'initialiser.

Pour ajouter un disque, utilisez la commande suivante :

```
# vxdiskadd nom_périphérique
```

où *nom_périphérique* représente le nom de périphérique du disque à ajouter.

Pour placer le périphérique `c1t0d0` sous le contrôle de Volume Manager, procédez comme suit :

1. Entrez la commande suivante pour démarrer `vxdiskadd` :

```
# vxdiskadd c1t0d0
```

2. Pour poursuivre la tâche, entrez `y` (ou appuyez sur la touche de retour) à l'invite suivante :

```
Ajouter ou initialiser les disques  
Menu : VolumeManager/Disk/AddDisks
```

```
Le disque sélectionné est le suivant. Format de sortie :
```

```
[Nom_périphérique]  
c1t0d0
```

```
Continuer l'opération ? [y,n,q,?] (default: y) y
```

3. À l'invite suivante, spécifiez le disk group auquel ajouter le disque ou appuyez sur la touche de retour pour accepter `rootdg` (pour plus d'informations, voir « Mise hors ligne d'un disque » à la page 105) :

```
Vous pouvez choisir d'ajouter ce disque à un disk group  
existant, à un nouveau disk group ou de le laisser disponible  
afin qu'il puisse être utilisé pour des opérations futures  
d'ajout ou de remplacement. Pour créer un nouveau disk group,  
sélectionnez le nom d'un disk group qui n'existe pas. Pour  
que le disque soit disponible plus tard, nommez le disk group  
"none".
```

```
Quel disk group [<group>,none,list,q,?] (default: rootdg)
```

4. À l'invite suivante, appuyez sur la touche de retour pour accepter le nom du disque par défaut ou entrez le nom voulu :

```
Utiliser un nom de disque par défaut pour le disque ?  
[y,n,q,?] (default: y)
```

Ajout d'un disque à Volume Manager

5. Lorsque le système vous demande si ce disque doit devenir un disque de remplacement pour la reconstruction automatique, entrez **n** (ou appuyez sur la touche de retour) :

```
Ajouter un disque comme disque de remplacement pour rootdg ?
[y,n,q,?] (default: n) n
```

Pour plus d'informations, voir « Opérations de l'interface de menu vxdiskadm » à la page 136.

6. À l'invite suivante, entrez **n** pour indiquer que ce disque ne doit pas être exclu de la reconstruction automatique :

```
Exclure un autre disque de la reconstruction automatique ?
[y,n,q,?] [default:n] n
```

7. Pour poursuivre la tâche, entrez **y** (ou appuyez sur la touche de retour) à l'invite suivante :

Les disques sélectionnés seront ajoutés au disk group rootdg avec les noms de disque définis par défaut.

```
c1t0d0
```

```
Continuer l'opération ? [y,n,q,?] (default: y) y
```

8. Si vous voulez conserver certaines données de ce disque, entrez **y** pour sélectionner l'encapsulation :

Le périphérique de disque suivant a un VTOC valide, mais ne semble pas avoir été initialisé pour Volume Manager. Si le disque contient des données, elles NE devraient PAS être détruites ; vous devriez encapsuler les partitions du disque existantes comme volumes au lieu d'ajouter le disque comme nouveau disque. Format de sortie : [Nom_périphérique]

```
c1t0d0
```

```
Encapsuler ce périphérique ? [y,n,q,?] (default: y) y
```

9. Pour continuer la tâche, appuyez sur la touche de retour à l'invite suivante :

```
Le disque suivant a été sélectionné pour l'encapsulation.
Format de sortie : [Nom_périphérique c1t0d0]
```

```
Continuer l'encapsulation ? [y,n,q,?] (default: y)
```

Un message du type suivant indique que le disque est en cours d'encapsulation pour être utilisé par Volume Manager :

```
Le périphérique de disque c1t0d0 va être encapsulé et ajouté au
disk group rootdg avec le nom de disque disk01.
```

```
Le disque c1t0d0 a été configuré pour l'encapsulation.
```

La première étape d'encapsulation a réussi. Vous devriez redémarrer le système dès que possible. L'encapsulation nécessitera deux ou trois redémarrages qui se produiront automatiquement après le démarrage suivant. Pour redémarrer, exécutez la commande :

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

Elle permet de mettre à jour le fichier /etc/vfstab pour pouvoir utiliser les périphériques de volume afin de monter les systèmes de fichiers sur ce périphérique de disque. Vous devrez mettre à jour toutes les autres références, telles que les scripts de sauvegarde, les bases de données ou les périphériques de swap créés manuellement.

Au revoir.

N'oubliez pas d'arrêter le système et de le redémarrer dès que possible.

Ajout d'un disque à un disk group

Vous pouvez ajouter un nouveau disque à un disk group déjà créé. Par exemple, l'espace des disques actuels est insuffisant pour les spécifications des applications ou des groupes de travail, surtout si elles ont été modifiées.

Pour ajouter un disque initialisé à un disk group, utilisez la commande suivante :

```
# vxdiskadd nom_périphérique
```

Pour ajouter le périphérique c1t1d0 à rootdg, procédez comme suit :

1. Entrez la commande suivante pour démarrer vxdiskadd :

```
# vxdiskadd c1t1d0
```

Ajout d'un disque à un disk group

vxdiskadd affiche le message suivant :

```
Ajouter ou initialiser les disques
Menu : VolumeManager/Disk/AddDisks
```

```
Le disque sélectionné est le suivant. Format de sortie :
[Nom_périphérique]
```

```
c1t1d0
```

```
Continuer l'opération ? [y,n,q,?] (default: y) y
```

2. À l'invite suivante, spécifiez le disk group auquel ajouter le disque ou appuyez sur la touche de retour pour accepter rootdg:

```
Vous pouvez ajouter ce disque à un disk group existant, à un
nouveau disk group ou le laisser disponible pour qu'il soit
utilisé pour des opérations d'ajout ou de remplacement futures.
Pour créer un nouveau disk group, sélectionnez le nom d'un disk
group qui n'existe pas. Pour que le disque soit disponible plus
tard, nommez le disk group "none".
```

```
Quel disk group [<group>,none,list,q,?] (default: rootdg)
```

3. À l'invite suivante, appuyez sur la touche de retour pour accepter le nom du disque par défaut ou entrez le nom voulu :

```
Utiliser un nom de disque par défaut pour le disque ? [y,n,q,?]
(default: y)
```

4. Lorsque vxdiskadd vous demande si ce disque doit devenir un disque de remplacement pour la reconstruction automatique, entrez **n** (ou appuyez sur la touche de retour) :

```
Ajouter un disque comme disque de remplacement pour rootdg ?
[y,n,q,?] (default: n) n
```

5. À l'invite suivante, entrez **n** pour indiquer que ce disque ne doit pas être exclu de la reconstruction automatique :

```
Exclure un autre disque de la reconstruction automatique ?
[y,n,q,?] [default:n] n
```

6. Pour continuer la tâche, entrez **y** (ou appuyez sur la touche de retour) à l'invite suivante :

```
Les disques sélectionnés seront ajoutés au disk group rootdg
avec les noms de disque définis par défaut.
```

```
c1t1d0
```

Continuer l'opération ? [y,n,q,?] (default: y) **y**

7. L'invite suivante indique que ce disque a déjà été initialisé pour une utilisation ultérieure par Volume Manager ; entrez **y** pour confirmer que vous voulez utiliser ce disque maintenant :

Le périphérique de disque suivant semble avoir déjà été initialisé. Le disque est actuellement disponible comme disque de remplacement.

Format de sortie : [Nom_périphérique]

c1t1d0

Utiliser ce périphérique ? [y,n,q,?] (default: y) **y**

8. Pour réinitialiser le disque, entrez **y** (ou appuyez sur la touche de retour) à l'invite suivante :

Le disque suivant, que vous avez sélectionné, semble déjà avoir été initialisé pour Volume Manager. Si vous êtes sûr que c'est le cas, vous n'avez pas à réinitialiser le périphérique de disque.

Format de sortie : [Nom_périphérique]

c1t1d0

Réinitialiser ce périphérique ? [y,n,q,?] (default: y) **y**

Des messages du type suivant confirment que ce disque est en cours de réinitialisation pour être utilisé par Volume Manager. Selon le système, vous pouvez également choisir d'effectuer une analyse de surface.

Initialisation du périphérique c1t1d0.

Exécuter l'analyse de surface (fortement recommandé)

[y,n,q,?] (default: y) **n**

Ajout du périphérique de disque c1t1d0 au disk group rootdg avec le nom de disque disk03.

Pour confirmer que le disque a été ajouté au disk group, entrez :

```
# vxdisk list
```

Volume Manager renvoie une liste du type suivant :

PÉRIPHÉRIQUE	TYPE	DISQUE	GROUPE	ÉTAT
c0t0d0s2	sliced	disk04	rootdg	en ligne
c1t0d0s2	sliced	disk01	rootdg	en ligne
c1t1d0s2	sliced	disk03	rootdg	en ligne

Suppression d'un disque d'un disk group

Utilisez la commande suivante pour supprimer de son disk group un disque qui ne contient pas de sous-disque :

```
# vxdg [-g nom_groupe] rmdisk nom_disque
```

où le nom du disk group n'est spécifié que pour les autres disk groups que celui par défaut (rootdg).

Par exemple, pour supprimer disk02 de rootdg, utilisez la commande suivante :

```
# vxdg rmdisk disk02
```

Si vous tentez de supprimer un disque qui contient des sous-disques, le message d'erreur suivant s'affiche :

```
vxdg:Le disque nom_disque est utilisé par un ou plusieurs  
sous-disques
```

Utilisez l'option -k sur vxdg pour supprimer l'assignation de périphérique. Cette option vous permet de supprimer le disque même s'il contient des sous-disques. Pour plus d'informations, voir la page man vxdg(1M).

Remarque L'utilisation de l'option -k sur vxdg peut entraîner une perte de données.

Lorsque le disque a été supprimé de son disk group, vous pouvez le retirer entièrement (facultatif) du contrôle de Volume Manager, en procédant comme suit :

```
# vxdisk rm nom_périphérique
```

Par exemple, pour retirer c1t0d0 (ou c1b0t0d0) du contrôle de Volume Manager, utilisez les commandes suivantes :

```
# vxdisk rm c1t0d0s2
```

ou, si le système a un bus :

```
# vxdisk rm c1b0t0d0s0
```

Vous pouvez supprimer un disque sur lequel des sous-disques sont définis. Par exemple, vous pouvez consolider tous les volumes sur un seul disque. Si vous utilisez vxdiskadm pour supprimer un disque, vous pouvez déplacer les volumes de ce disque. Pour cela, exécutez vxdiskadm et sélectionnez la commande 3 (Supprimer un disque) du menu principal.

Si le disque est utilisé par des sous-disques, un message du type suivant s'affiche :

```
Les volumes suivants utilisent actuellement une partie du
disque disk02:
home usrvol
Ils doivent être déplacés de disk02 avant de pouvoir supprimer
ce dernier.
```

```
Déplacer les volumes sur d'autres disques ? [y,n,q,?] (default: n)
```

Si vous choisissez **y**, tous les sous-disques sont déplacés du disque, lorsque cette opération est possible. Certains sous-disques ne pourront peut-être pas être déplacés. Généralement, cela est dû aux conditions suivantes :

- ◆ L'espace disponible sur les disques restants est insuffisant.
- ◆ Les plex ou les sous-disques répartis par bandes ne peuvent pas être alloués sur des disques différents à partir des plex ou des sous-disques répartis par bandes existants dans le volume.

Si `vxddiskadm` ne peut pas déplacer certains sous-disques, vous devrez peut-être supprimer des plex de certains disques pour libérer davantage d'espace avant de continuer l'opération de suppression du disque.

Déplacement des disques

Pour déplacer un disque d'un disk group à l'autre, supprimez-le du premier disk group et ajoutez-le dans l'autre. Par exemple, pour déplacer le disque physique `c0t3d0` (attaché en tant que `disk04`) du disk group `rootdg` pour l'ajouter au disk group `mktgdg`, entrez les commandes suivantes :

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

Remarque Cette procédure n'enregistre pas les configurations ni les données sur les disques.

Vous pouvez aussi utiliser `vxddiskadm` pour déplacer un disque. Sélectionnez la commande 3 (Supprimer un disque) du menu principal, puis choisissez la commande 1 (Ajouter ou initialiser un ou plusieurs disques).

Attribution d'un nouveau nom à un disque

Attribution d'un nouveau nom à un disque

Si vous ne spécifiez pas de nom Volume Manager pour le disque, Volume Manager lui attribue un nom par défaut lorsque vous le placez sous son contrôle. Ce nom permet à Volume Manager d'identifier l'emplacement ou le type du disque. Vous pouvez changer le nom du disque pour traduire un changement de propriétaire ou d'utilisation ; pour ce cela, entrez :

```
# vxedit rename old_diskname new_diskname
```

Pour renommer `disk01` par `disk03`, entrez :

```
# vxedit rename disk01 disk03
```

Pour confirmer le changement de nom, entrez :

```
# vxdisk list
```

Volume Manager affiche la liste suivante :

PÉRIPHÉRIQUE	TYPE	DISQUE	GROUPE	ÉTAT
c0t0d0s2	sliced	disk04	rootdg	en ligne
c1t0d0s2	sliced	disk03	rootdg	en ligne
c1t1d0s2	sliced	-	-	en ligne

Remarque Par défaut, Volume Manager nomme les objets des sous-disques en fonction du disque VM sur lequel ils sont placés. Lorsque vous renommez un disque VM, les sous-disques qu'il contient ne sont pas automatiquement renommés.

Réservation des disques

Par défaut, `vxassist` attribue l'espace libre de n'importe quel disque. Vous pouvez réserver un jeu de disques à des fins spéciales, par exemple, pour éviter l'utilisation générale d'un disque particulièrement lent ou rapide.

Entrez la commande suivante afin de réserver un disque à des fins spéciales :

```
# vxedit set reserve=on nom_disque
```

Une fois que vous avez entré cette commande, `vxassist` n'alloue pas l'espace disponible du disque sélectionné, sauf si ce dernier est spécifiquement mentionné sur la ligne de commande `vxassist`. Par exemple, si le disque `disk03` est réservé, la commande :

```
# vxassist make vol103 20m disk03
```

ignore la réservation et crée un volume de 20 Mo sur `disk03`. Cependant, la commande :

```
# vxassist make vol104 20m
```

n'utilise pas `disk03`, même s'il n'existe pas d'espace libre sur un autre disque.

Pour désactiver la réservation d'un disque, entrez :

```
# vxedit set reserve=no nom_disque
```

Mise hors ligne d'un disque

Vous devrez peut-être mettre un disque physique hors ligne. Si le disque est endommagé, vous devez le désactiver et le supprimer. Vous devez aussi désactiver un disque avant de déplacer le périphérique de disque physique vers un autre emplacement afin de le connecter à un autre système.

Pour mettre un disque physique hors ligne, vous devez d'abord supprimer le disque de son `disk group` ; pour plus d'informations, voir « Suppression d'un disque » à la page 107. Mettez ensuite le disque dans l'état « hors ligne » en entrant la commande suivante :

```
# vxdisk offline nom_périphérique
```

Remarque Dans ce cas, le nom du périphérique doit être utilisé car, puisque le disque n'appartient plus à un `disk group`, il ne possède aucun nom administratif.

Mise en miroir d'un disque

La mise en miroir du disque racine met en miroir le volume racine et les autres zones nécessaires au démarrage sur un autre disque.

Pour mettre en miroir le disque d'initialisation (racine) sur un autre disque, utilisez soit la procédure décrite ici, soit `vxdiskadm`. Vous pourrez ainsi effectuer la récupération après la défaillance du disque d'initialisation en le remplaçant par son miroir.

Avant de mettre en miroir le disque d'initialisation, vous devez spécifier la valeur `true` pour la variable EEPROM `use-nvramrc?` afin d'utiliser les alias de disque d'initialisation Volume Manager pour identifier le miroir du disque d'initialisation si un remplacement est nécessaire. Si la variable a la valeur `false`, vous devrez déterminer vous-même le disque d'initialisation. Pour définir la variable comme `true`, procédez comme suit :

```
eprom use-nvramrc?=true
```

Mise en miroir d'un disque

Pour mettre en miroir le disque d'initialisation, procédez comme suit :

1. Sélectionnez un disque d'une taille au moins égale à celle du disque d'initialisation.
2. Utilisez la commande `vxdiskadd` pour ajouter (le cas échéant) le disque sélectionné en tant que nouveau disque.
3. Exécutez la commande suivante :

```
# /etc/vx/bin/vxrootmir autre_disque
```

où *autre_disque* représente le nom de l'autre disque.

`vxrootmir` crée un miroir pour `rootvol` (le volume pour le système de fichiers racine sur un autre disque).

L'autre disque d'initialisation est configuré pour activer le démarrage en cas de défaillance du disque d'initialisation principal.

Le disque d'initialisation peut comprendre d'autres volumes (par exemple, pour les systèmes de fichiers `/home` ou `/tmp`). Vous pouvez les mettre en miroir séparément au moyen de l'utilitaire `vxassist`. Par exemple, vous pouvez mettre en miroir un système de fichiers `/home` sur le volume `homevol` sur l'*autre_disque* en entrant la commande suivante :

```
# vxassist mirror homevol autre_disque
```

Si l'espace est insuffisant pour copier certains systèmes de fichiers sur l'autre disque d'initialisation, vous pouvez mettre ces derniers en miroir sur d'autres disques. Vous pouvez aussi répartir ou répartir par bandes ces autres volumes sur les autres disques attachés au système.

Pour lister tous les volumes sur le disque d'initialisation principal, utilisez la commande :

```
# vxprint -t -v -e\'aslist.aslist.sd_disk="disque_initialisation"'
```

Pour mettre en miroir tous les volumes concaténés sur ce disque vers l'autre disque d'initialisation, utilisez la commande :

```
# /etc/vx/bin/vxmirror disque_initialisation autre_disque
```

Suppression d'un disque

Vous pouvez supprimer un disque d'un système et le déplacer vers un autre système, s'il est en train d'échouer ou a échoué. Avant de supprimer le disque du système actuel, vous devez :

1. Démontez tous les systèmes de fichiers des volumes.
2. Arrêtez les volumes sur le disque.
3. Déplacez les volumes vers d'autres disques ou les sauvegardez. Pour déplacer un volume, mettez-le en miroir sur un ou plusieurs autres disques, puis supprimez la copie d'origine du volume. Si les volumes ne sont plus nécessaires, vous pouvez les supprimer au lieu de les déplacer.

Pour supprimer un disque, procédez comme suit :

4. Supprimez le disque de son disk group en entrant la commande suivante :

```
# vxdg [-g nom_groupe] rmdisk nom_disque
```

où *nom_groupe* représente le nom du groupe auquel le disque appartient et *nom_disque*, le nom du disque à supprimer.

Par exemple, pour supprimer `disk01` de `rootdg`, entrez :

```
# vxdg rmdisk disk01
```

Il n'est pas nécessaire de spécifier `rootdg` car il correspond au disk group par défaut.

5. Retirez le disque du contrôle de Volume Manager en entrant la commande suivante :

```
# vxdisk rm nom_périphérique
```

Par exemple, pour retirer `c1t0d0` du contrôle de Volume Manager, entrez :

```
# vxdisk rm c1t0d0s2
```

Affichage des informations sur le disque

Pour utiliser un disque, vous devez préalablement savoir s'il a été initialisé et placé sous le contrôle de Volume Manager. Vous devez également savoir s'il appartient à un disk group, car vous ne pouvez pas créer de volumes sur un disque si ce dernier n'appartient pas à un disk group. La commande `vxdisk list` permet d'afficher les noms de périphériques pour tous les disques reconnus, les noms de disque, les noms de disk group associés à chaque disque et l'état de chaque disque.

Entrez la commande suivante pour afficher des informations pour tous les disques définis sur Volume Manager :

```
# vxdisk list
```

Volume Manager affiche ce qui suit :

PÉRIPHÉRIQUE	TYPE	DISQUE	GROUPE	ÉTAT
c0t0d0s2	sliced	disk04	rootdg	en ligne
c1t0d0s2	sliced	disk01	rootdg	en ligne
c1t1d0s2	sliced	-	-	en ligne

Pour afficher des détails sur un disque spécifique définis sur Volume Manager, entrez :

```
# vxdisk list disk01
```

L'utilitaire `vxdisk` permet d'afficher des informations sur les chemins d'accès multiples d'un méta-périphérique spécifique. Un méta-périphérique représente le périphérique d'un disque physique spécifique qui possède plusieurs chemins d'accès physiques à partir du contrôleur d'E/S du système. Dans Volume Manager, les disques physiques du système sont tous représentés sous la forme de méta-périphériques possédant un ou plusieurs chemins d'accès physiques.

Pour afficher les informations de chemins d'accès multiples pour un méta-périphérique spécifique, entrez la commande :

```
# vxdisk list périphérique
```

où *périphérique* représente le méta-périphérique formé par le sous-système DMP de Volume Manager.

Les informations qui s'affichent indiquent deux chemins d'accès à un périphérique physique représenté par le méta-périphérique `c2t0d0s2`. Le chemin `c2t0d0s2` est actif (`state=enabled`) et l'autre chemin `c1t0d0s2` a échoué (`state=disabled`).

Volume Manager affiche ce qui suit :

```
Device      périphérique
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
```

D'autres informations sont fournies sous la forme de *type* pour les disques des ensembles de disques de type actif/passif. Ces informations indiquent les chemins d'accès *principal* et *secondaire* au disque. Par exemple, DG Clariion, Hitachi DF350, etc.

Les informations de *type* ne s'affichent pas pour les disques des ensembles de disques de type actif/actif, car le concept de chemin principal et de chemin secondaire ne s'applique pas aux disques de ce type d'ensemble (par exemple, les ensembles de disques StorEdge A5000 et Sparc Storage Array (SSA)).

Détection et remplacement des disques qui ont échoué

Cette section décrit comment détecter et remplacer les disques qui ont échoué. Elle présente d'abord la fonction de reconstruction automatique, qui tente de restaurer automatiquement les objets Volume Manager redondants en cas de défaillance.

Reconstruction automatique

La reconstruction automatique répond automatiquement aux erreurs d'E/S sur les objets Volume Manager redondants (en miroir ou RAID 5) et restaure la redondance ainsi que l'accès à ces objets. Volume Manager détecte les erreurs d'E/S sur les objets et reconstruit les sous-disques suspects sur les disques spécifiés comme disques de remplacement et/ou sur l'espace disponible dans le disk group. Volume Manager reconstruit alors les objets qui existaient avant la défaillance et rétablit leur redondance et leur accessibilité. Voir « Reconstruction automatique » à la page 69 pour obtenir une description de la reconstruction automatique.

Remarque La reconstruction automatique s'applique uniquement aux sous-disques redondants (en miroir ou RAID 5) d'un disque qui a échoué. Les sous-disques non redondants d'un disque qui a échoué ne sont pas reconstruits, mais l'administrateur système est informé de leur défaillance.

La reconstruction automatique est activée par défaut et est exécutée en cas de défaillance sans nécessiter l'intervention de l'administrateur système. Le démon de reconstruction automatique, `vxrelocd`, détecte et répond aux événements Volume Manager qui indiquent les types de défaillance suivants :

- ◆ Défaillance de disque — Elle est généralement détectée à la suite d'une erreur d'E/S à partir d'un objet Volume Manager. Volume Manager tente de corriger l'erreur. S'il n'y parvient pas, il tente d'accéder aux informations de configuration situées dans la zone privée du disque. S'il ne peut pas accéder à cette zone, il considère que le disque a échoué.
- ◆ Défaillance de plex — Elle est généralement détectée à la suite d'une erreur d'E/S irrécupérable dans le plex (qui affecte les sous-disques du plex). Le plex est détaché pour les volumes mis en miroir.
- ◆ Défaillance de sous-disque RAID 5 — Elle est généralement détectée à la suite d'une erreur d'E/S irrécupérable. Le sous-disque est détaché.

Lorsqu'une défaillance de ce type est détectée, `vxrelocd` en informe l'administrateur système par courrier électronique et indique également les objets Volume Manager affectés. `vxrelocd` détermine ensuite les sous-disques à reconstruire (le cas échéant). Si la reconstruction est possible, `vxrelocd` recherche l'espace approprié et reconstruit les sous-disques.

L'espace nécessaire à la reconstruction automatique est choisi sur les disques réservés à cet effet dans le disk group affecté par la défaillance. Si aucun disque de remplacement n'est disponible ou si l'espace disponible est insuffisant, le système utilise n'importe quel espace libre de ce même disk group, sauf sur les disques marqués `nohotuse` (ils sont exclus de la reconstruction automatique). Chaque sous-disque reconstruit est ensuite rattaché à son plex.

Enfin, `vxrelocd` lance les procédures de récupération appropriées. Par exemple, la récupération inclut la resynchronisation des miroirs pour les volumes mis en miroir ou la récupération des données pour les volumes RAID 5. L'administrateur système est informé de la reconstruction automatique et les opérations de récupération sont exécutées.

Si la reconstruction est impossible, l'administrateur système en est informé et aucune opération n'est exécutée. La reconstruction est impossible dans les cas suivants :

- ◆ Si les sous-disques ne sont pas redondants (c'est-à-dire, s'ils n'appartiennent pas à des volumes mis en miroir ou RAID 5), ils ne peuvent pas être reconstruits.
- ◆ Si l'espace disponible est insuffisant dans le disk group (sur les disques de remplacement ou l'espace libre, et s'il n'est pas exclu de la reconstruction automatique au moyen de l'indicateur `nohotuse`), les sous-disques qui ont échoué ne peuvent pas être reconstruits.
- ◆ Si l'espace disponible se trouve uniquement sur un disque qui contient déjà un miroir du plex qui a échoué, les sous-disques de ce plex ne peuvent pas être reconstruits.
- ◆ Si l'espace disponible se trouve uniquement sur un disque qui contient déjà un plex de journal pour le volume RAID 5 ou de l'un de ses sous-disques en bon état, il est impossible de reconstruire le sous-disque défectueux du plex RAID 5.
- ◆ Si le plex de données d'un volume mis en miroir inclut un sous-disque de journal des zones modifiées, les sous-disques de ce plex ne peuvent pas être reconstruits.
- ◆ En cas de défaillance d'un plex de journal RAID 5 ou d'un plex de journal des zones modifiées sur un volume mis en miroir, un nouveau plex de journal est créé ailleurs (en fait, le plex de journal n'est pas réellement reconstruit).

Vous pouvez préparer la reconstruction automatique en configurant un ou plusieurs disques comme disques de remplacement pour la reconstruction automatique pour chaque disk group. Pour plus d'informations sur la façon de configurer un disque en tant que disque de remplacement, voir « Ajout d'un disque VM au pool de reconstruction automatique » à la page 158. Si aucun disque de remplacement n'est disponible lorsqu'une défaillance se produit ou si l'espace disponible sur les disques de remplacement est insuffisant, l'espace libre est automatiquement utilisé, à condition qu'il ne soit pas exclu de la reconstruction automatique au moyen de l'indicateur `nohotuse`.

Détection et remplacement des disques qui ont échoué

Lorsque vous configurez les disques de remplacement, vous pouvez choisir l'espace qui sera utilisé pour la reconstruction en cas de défaillance. Si l'espace total obtenu par la combinaison entre l'espace disponible et l'espace des disques de remplacement est insuffisant ou ne respecte pas les contraintes de redondance, les sous-disques ne sont pas reconstruits.

Lorsque vous configurez un disque `nohotuse`, vous pouvez choisir les disques libres qui ne pourront pas être utilisés comme disques de remplacement.

Vous pouvez utiliser deux méthodes pour exclure un disque de la reconstruction automatique. Par exemple :

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

Dans le menu principal de `vxdiskadm`, sélectionnez l'option 15 (Exclure un disque de la reconstruction automatique).

Vous pouvez utiliser deux méthodes pour inclure un disque dans la reconstruction automatique. Par exemple :

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

Dans le menu principal de `vxdiskadm`, sélectionnez l'option 16 (Libérer un disque pour l'utilisation de la reconstruction automatique).

Vous pouvez utiliser deux méthodes pour déterminer les disques qui sont configurés comme des disques de remplacement ou qui sont exclus de la reconstruction automatique. Par exemple :

```
# vxdisl list
# vxprint
```

Lorsque la reconstruction est terminée, vous devez supprimer et remplacer le disque qui a échoué (voir « Remplacement des disques » à la page 121). Selon l'emplacement des sous-disques reconstruits, vous pouvez les déplacer lorsque la reconstruction automatique est terminée (voir « Déplacement des sous-disques reconstruits » à la page 114).

Modification de `vxrelocd`

La reconstruction automatique est activée tant que `vxrelocd` est en cours d'exécution. Vous devez laisser la reconstruction automatique activée pour pouvoir utiliser au mieux cette fonction en cas de défaillance. Cependant, si vous choisissez de désactiver cette fonction (si vous ne voulez pas utiliser l'espace libre de certains des disques pour la reconstruction), vous devez empêcher `vxrelocd` de démarrer en même temps que le système.

Reportez-vous au *Guide d'installation de VERITAS Volume Manager* pour plus d'informations sur la façon de désactiver la reconstruction automatique au démarrage du système. Pour arrêter la reconstruction automatique à tout moment, arrêtez la procédure `vxrelocd` (vous ne devez pas procéder ainsi tant qu'une tentative de reconstruction automatique est en cours).

Pour modifier quelque peu `vxrelocd`, éditez la ligne `vxrelocd` dans le fichier de démarrage `vxrelocd (/etc/rc2.d/S95vxvm-recover)` ou arrêtez la procédure `vxrelocd` existante et redémarrez-la avec des options différentes. Une fois que vous avez modifié la procédure qui permet d'appeler `vxrelocd` dans le fichier de démarrage, vous devez réinitialiser le système pour que les modifications prennent effet. Si vous préférez arrêter et redémarrer le démon, assurez-vous que la reconstruction automatique est en cours lorsque vous arrêtez la procédure `vxrelocd`. Vous devez aussi redémarrer immédiatement le démon pour que la reconstruction automatique soit lancée en cas de défaillance.

Vous pouvez modifier le comportement de `vxrelocd` comme suit :

- ◆ Par défaut, `vxrelocd` envoie un courrier électronique à `root` lorsque les défaillances sont détectées et les actions de reconstruction exécutées. Vous pouvez indiquer à `vxrelocd` de notifier d'autres utilisateurs en ajoutant les noms d'utilisateurs appropriés et en appelant `vxrelocd` comme suit :

```
# vxrelocd root nom_utilisateur1 nom_utilisateur2 &
```

- ◆ Pour limiter l'effet de la récupération sur les performances du système, vous pouvez indiquer à `vxrelocd` d'augmenter le délai séparant la récupération de chaque zone du volume, comme suit :

```
# vxrelocd -o slow[=/Odelay] root &
```

où le paramètre facultatif `/Odelay` indique le délai voulu (en millisecondes). La valeur par défaut est de 250 millisecondes. Pour plus d'informations, voir la page `man vxrelocd(1M)`.

`-O` Cette option permet de rétablir une ancienne version. Lorsque vous spécifiez `VxVM_version -O`, vous indiquez à `vxrelocd` d'utiliser le motif de reconstruction dans cette version.

`-s` Avant que `vxrelocd` n'active la reconstruction, une prise de vue de la configuration actuelle est enregistrée dans `/etc/vx/saveconfig.d`. Cette option spécifie le nombre maximal de configurations à conserver pour chaque `disk group` (valeur par défaut : 32).

Détection et remplacement des disques qui ont échoué

Affichage d'informations sur les disques de remplacement

La commande `vxldg spare` permet d'afficher des informations sur tous les disques de remplacement disponibles pour la reconstruction. Elle affiche les informations suivantes :

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

Dans cet exemple, `disk02` est le seul disque spécifié comme disque de remplacement. Le champ `LENGTH` indique l'espace de remplacement actuellement disponible sur ce disque pour la reconstruction.

Vous pouvez aussi utiliser les commandes suivantes pour afficher des informations sur les disques actuellement spécifiés comme disques de remplacement :

- ◆ `vxldisk list` — Affiche les informations sur les disques, en indiquant les disques de remplacement par l'indicateur `spare`.
- ◆ `vxprint` — Liste les disques et affiche d'autres informations, en indiquant les disques de remplacement par l'indicateur `spare`.

Déplacement des sous-disques reconstruits

Lors de la reconstruction automatique, les sous-disques sont reconstruits sur les disques de remplacement et/ou sur l'espace disponible dans ce `disk group`. Il se peut que les nouveaux emplacements des sous-disques ne présentent pas les mêmes performances ou structure de données que celles d'avant la reconstruction automatique. Vous pouvez déplacer les sous-disques reconstruits (à l'issue de la reconstruction automatique) pour améliorer les performances.

Vous pouvez aussi déplacer les sous-disques reconstruits des disques de remplacement pour conserver de l'espace en vue de futures reconstructions automatiques ou pour recréer la configuration qui existait avant l'exécution de la reconstruction automatique.

Pendant la reconstruction automatique, des messages électroniques sont envoyés à `root` (voir l'exemple suivant) :

```
To: root
Subject: Défaillances de Volume Manager sur l'hôte teal
```

```
Tentative de reconstruction du sous-disque disk02-03
depuis le plex home-02.
Dev_offset 0 length 1164 dm_name disk02 da_name c0t5d0s2.
Le plex disponible home-01 sera utilisé pour récupérer les
données.
```

Ce message contient des informations sur le sous-disque avant la reconstruction automatique et vous permet de choisir ou déplacer le sous-disque à l'issue de la reconstruction automatique.

Autre exemple de message indiquant le nouvel emplacement du sous-disque reconstruit :

```
To: root
Subject: Tentative de reconstruction de VxVM sur l'hôte teal

Volume home Sous-disque disk02-03 reconstruit dans disk05-01,
mais pas encore récupéré.
```

Avant de déplacer un sous-disque reconstruit, réparez ou remplacez le disque qui a échoué (en suivant les procédures décrites dans les sections précédentes). Vous pouvez ensuite déplacer un sous-disque reconstruit sur le disque d'origine. Par exemple, vous pouvez entrer la commande suivante pour déplacer le sous-disque reconstruit `disk05-01` sur `disk02` :

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

Remarque Les volumes RAID 5 ne sont pas redondants pendant le déplacement des sous-disques.

Utilisation de `vxunrelocate`

La reconstruction automatique VxVM permet au système de répondre automatiquement aux erreurs d'E/S sur un objet VxVM redondant au niveau du sous-disque, puis de prendre les actions nécessaires pour rendre l'objet de nouveau disponible. Ce mécanisme détecte les erreurs d'E/S sur un sous-disque, reconstruit ce sous-disque et récupère le plex qui lui est associé. Une fois le disque remplacé, `vxunreloc` vous permet de restaurer la configuration du système active avant la défaillance du disque. `vxunreloc` vous permet de déplacer les sous-disques reconstruits sur un disque remplacé à la suite d'une défaillance.

Lorsque vous appelez `vxunreloc`, vous devez spécifier le nom d'accès logique au disque qui contenait initialement les sous-disques reconstruits. `vxunreloc` déplace les sous-disques vers les décalages d'origine. Si vous tentez de rétablir dans son disque d'origine un disque plus petit que le disque d'origine qui a échoué, `vxunreloc` renvoie simplement une erreur.

`vxunreloc` offre une option permettant de déplacer les sous-disques vers un disque différent de celui sur lequel ils ont été initialement reconstruits. Cette fonction comprend également une option permettant de rétablir les sous-disques vers un décalage différent, à condition que l'espace du disque de destination soit suffisant pour contenir tous les sous-disques.

Détection et remplacement des disques qui ont échoué

Si `vxunreloc` ne peut pas déplacer les sous-disques vers les décalages d'origine, vous pouvez utiliser l'option `force` pour les déplacer vers un disque spécifique sans utiliser les décalages d'origine. Pour plus d'informations, voir la page `man vxunreloc(1M)`.

Les exemples suivants illustrent l'utilisation de `vxunreloc`.

Exemple 1 :

Un `disk01` a échoué et tous les sous-disques ont été reconstruits. Après le remplacement de `disk01`, vous pouvez utiliser `vxunreloc` pour déplacer vers `disk01` tous les sous-disques reconstruits.

```
vxunreloc -g newdg disk01
```

Exemple 2 :

L'utilitaire `vxunreloc` comprend l'option `-n` qui permet de déplacer les sous-disques vers un disque différent du disque de reconstruction d'origine.

Si `disk01` a échoué et que tous les sous-disques qu'il contenait ont été reconstruits sur d'autres disques, `vxunreloc` comprend une option permettant de déplacer les sous-disques vers un disque différent de celui sur lequel ils se trouvaient à l'origine. Une fois réparé, le disque est ajouté de nouveau au `disk group`, mais sous un autre nom, par exemple `disk05`. Pour déplacer tous les sous-disques reconstruits automatiquement vers le nouveau disque, vous pouvez utiliser la commande suivante :

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

Exemple 3 :

`disk01` a échoué, les sous-disques ont été reconstruits et vous voulez déplacer les sous-disques reconstruits automatiquement sur `disk05`, qui contient déjà des sous-disques. Vous pouvez utiliser l'option `force` pour déplacer les sous-disques reconstruits automatiquement sur `disk05`, mais pas aux décalages exacts :

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

Exemple 4 :

Si un sous-disque a été reconstruit automatiquement plusieurs fois à la suite de diverses défaillances de disque, vous pouvez tout de même le rétablir à son emplacement d'origine. Par exemple, si `disk01` a échoué, que le sous-disque `disk01-01` a été déplacé sur `disk02`, puis que `disk02` a enregistré une défaillance de disque, tous les sous-disques qu'il contient, y compris celui qui a fait l'objet d'une reconstruction automatique, seront déplacés une nouvelle fois. Lors du remplacement de `disk02`, l'opération `vxunreloc` pour `disk02` n'a effectué aucune opération sur le sous-disque `disk01-01` reconstruit

automatiquement. Cependant, le remplacement de `disk01` suivi d'une opération `vxunreloc` déplace `disk01-01` vers `disk01` si vous exécutez `vxunreloc` dès la fin du remplacement.

Une fois que vous avez réparé ou remplacé le disque qui a échoué, vous pouvez utiliser `vxunreloc` pour déplacer vers le disque tous les sous-disques reconstruits automatiquement. Lorsqu'un sous-disque est reconstruit automatiquement, son nom d'accès logique au disque d'origine et le décalage dans le disque sont enregistrés dans la base de données de configuration. Lorsqu'un sous-disque est déplacé vers son disque d'origine ou sur un nouveau disque au moyen de `vxunreloc`, ces informations sont effacées. Le nom `dm` et le décalage d'origine sont enregistrés dans les enregistrements du sous-disque. Pour imprimer tous les sous-disques qui ont été reconstruits automatiquement à partir de `disk01` dans le disk group `rootdg`, utilisez la commande suivante :

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

Pour déplacer vers le disque d'origine tous les sous-disques reconstruits automatiquement à partir de `disk01`, entrez la commande suivante :

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

L'utilitaire `vxunreloc` comprend l'option `-n` qui permet de déplacer les sous-disques vers un disque différent de celui sur lequel ils se trouvaient à l'origine. Par exemple, lorsque `disk01` a échoué, tous les sous-disques qu'il contenait ont été reconstruits automatiquement sur d'autres disques. Une fois réparé, le disque est ajouté une nouvelle fois au disk group, mais sous un autre nom, par exemple `disk05`. Pour déplacer vers le nouveau disque tous les sous-disques reconstruits automatiquement, vous pouvez utiliser la commande suivante :

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

Le disque de destination devrait avoir une capacité de stockage au moins équivalente à celle du disque d'origine. Si l'espace disponible est insuffisant, l'opération `unrelocate` échoue et aucun des sous-disques n'est déplacé.

`vxunreloc` déplace les sous-disques reconstruits automatiquement vers les décalages d'origine. Cependant, si certains sous-disques occupaient toute ou une partie de la zone correspondante sur le disque de destination, `vxunreloc` échoue. Généralement, vous disposez de deux options : (1) déplacer les sous-disques existants, puis relancer `vxunreloc`, ou (2) utiliser l'option `-f` de `vxunreloc` pour déplacer les sous-disques sur le disque de destination, mais laisser `vxunreloc` trouver de l'espace sur le disque. Tant que la capacité du disque de destination est suffisante pour permettre à la zone du disque affectée au stockage des sous-disques de recevoir tous les sous-disques, les sous-disques reconstruits automatiquement seront tous « rétablis dans leur disque d'origine », sans tenir compte des décalages d'origine.

Redémarrage de `vxunreloc` après des erreurs

`vxunreloc` déplace les sous-disques en trois phases par une opération interne. La première phase crée autant de sous-disques sur le disque de destination spécifié qu'il faut en rétablir dans le disque d'origine. Pendant la création des sous-disques, `vxunreloc` entre dans le champ de commentaire de l'enregistrement de sous-disque la chaîne d'identification « UNRELOC ». La seconde phase correspond au déplacement des données à proprement parler. Si les déplacements de sous-disques sont tous réussis, la troisième phase efface le champ de commentaire des enregistrements de sous-disque.

La création du sous-disque est une opération de type « tout ou rien ». Si `vxunreloc` ne parvient pas à créer tous les sous-disques, l'utilitaire ne crée aucun sous-disque et se ferme. En revanche, le fonctionnement des sous-disques n'est pas une opération de type « tout ou rien ». Le déplacement d'un sous-disque est indépendant des autres ; par conséquent, si le déplacement d'un sous-disque échoue, l'utilitaire `vxunreloc` imprime un message d'erreur, puis se ferme. Cependant, les sous-disques suivants restent tous sur le disque où ils ont été reconstruits automatiquement et ne seront pas déplacés. Le champ de commentaire des enregistrements des sous-disques rétablis dans leur disque d'origine reste marqué « UNRELOC », car la phase de nettoyage n'est jamais exécutée.

Si le système est arrêté après la création des nouveaux sous-disques sur le disque de destination, mais avant leur rétablissement, vous pouvez exécuter de nouveau l'utilitaire `unrelocate` au redémarrage du système. Comme décrit précédemment, lors de la création d'un nouveau sous-disque, `vxunreloc` indique « UNRELOC » dans le champ de commentaire du sous-disque. Lorsque `vxunreloc` est de nouveau exécuté, il vérifie les champs `offset`, `len` et `comment` des sous-disques existants sur le disque de destination afin de déterminer s'il a été laissé sur le disque lors d'une exécution antérieure de `vxunreloc` qui l'utilisera alors au mieux.

Vous ne devez pas modifier manuellement la chaîne « UNRELOC » du champ `comment`.

Si l'un des déplacements de sous-disques échoue dans une série, `vxunreloc` se ferme. Dans ce cas, vous devez vérifier l'erreur qui a entraîné l'échec du déplacement de sous-disque et déterminer si le rétablissement dans le disque d'origine peut se poursuivre. Lorsque `vxunreloc` est réexécuté pour reprendre les déplacements de disques, il utilise les sous-disques créés lors d'une exécution précédente.

La phase de nettoyage est exécutée avec une seule transaction. `vxunreloc` attribue une chaîne NULL au champ `comment` de tous les sous-disques marqués « UNRELOC », qui se trouvent sur le disque de destination. Cette procédure inclut le nettoyage des sous-disques qui ont été rétablis dans leur disque d'origine lors d'un appel précédent de `vxunreloc` dont l'exécution a échoué.

Détection des disques qui ont échoué

Remarque La fonction de reconstruction automatique de Volume Manager détecte automatiquement les défaillances de disque et en informe l'administrateur système à l'aide d'un message électronique. Si la reconstruction automatique est désactivée ou si vous n'avez pas lu le message électronique, vous pouvez afficher les défaillances de disque à l'aide de la commande `vxprint` ou consulter l'état des disques dans l'interface utilisateur graphique. Vous pouvez aussi voir les messages d'erreur des pilotes sur la console ou dans le fichier de messages système.

Si une erreur d'E/S de disque se produit sur un volume (par exemple, parce que le disque présente une erreur irrécupérable), Volume Manager peut détacher le plex faisant partie de la défaillance.

Lorsqu'un plex est détaché, les E/S s'arrêtent sur ce plex, mais elles se poursuivent sur les autres plex du volume. En cas de défaillance complète d'un disque, Volume Manager peut le détacher de son disk group.

Lorsqu'un disque est détaché, tous ses plex sont désactivés. S'il contient des volumes qui ne sont pas mis en miroir lorsqu'il est détaché, ces derniers sont également désactivés.

Défaillance partielle du disque

Si la reconstruction automatique est activée lorsqu'un plex ou un disque est détaché à la suite d'une défaillance, un message électronique indiquant la défaillance des objets est envoyé à `root`. En cas de défaillance partielle du disque, le message identifie les plex qui ont échoué. Par exemple, si un disque contenant des volumes mis en miroir échoue, vous pouvez recevoir le message électronique illustré dans cet exemple :

```
To: root
Subject: Défaillances de Volume Manager sur l'hôte teal

Des erreurs ont été détectées par VERITAS Volume Manager :

plex ayant échoué :
  home-02
  src-02
```

Pour plus d'informations sur la façon d'envoyer le message à d'autres utilisateurs que `root`, voir « Modification de `vxrelocd` » à la page 112.

Entrez la commande suivante pour déterminer le disque à l'origine des défaillances dans le message ci-dessus :

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

Détection et remplacement des disques qui ont échoué

L'exemple ci-dessous représente un affichage type :

TYP NOM	ÉCHEC	
	LECTURES	ÉCRITURES
sd disk01-04	0	0
sd disk01-06	0	0
sd disk02-03	1	0
sd disk02-04	1	0

Cet affichage indique que les défaillances concernent `disk02` (et que les sous-disques `disk02-03` et `disk02-04` sont affectés).

Les sous-disques affectés sont reconstruits automatiquement et toutes les procédures de récupération nécessaires sont lancées. Cependant, si la reconstruction n'est pas possible ou si la fonction de reconstruction automatique est désactivée, vous devez analyser le problème et tenter de récupérer les plex. Ces erreurs peuvent être dues à des câbles défectueux ; il est recommandé de vérifier les câbles qui connectent les disques au système. Si les problèmes sont évidents, corrigez-les et récupérez les plex en entrant la commande suivante :

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

Cette commande lance la récupération des plex qui ont échoué en arrière-plan (la commande s'affiche de nouveau à l'écran avant la fin de l'opération). Si un message d'erreur s'affiche par la suite ou si les plex sont de nouveau détachés alors que les câbles ne présentent aucune défaillance évidente, remplacez le disque (voir « Remplacement des disques » à la page 121).

Défaillance complète du disque

Si un disque échoue complètement alors que la reconstruction automatique est activée, le message électronique indique le disque qui a échoué, ainsi que tous les plex qui utilisent ce disque. Par exemple, vous pouvez recevoir le message électronique suivant :

```
To: root
Subject: Défaillances de Volume Manager sur l'hôte teal

Des erreurs ont été détectées par VERITAS Volume Manager :

disques ayant échoué :
  disk02
```

```
plex ayant échoué :
  home-02
  src-02
  mktng-01

sous-disques ayant échoué :
  disk02
```

Ce message indique que `disk02` a été détaché à la suite d'une défaillance. Lorsqu'un disque est détaché, les E/S ne peuvent pas lui parvenir. Les plex `home-02`, `src-02` et `mktng-01` ont également été détachés (probablement à la suite de la défaillance du disque).

De nouveau, le problème peut provenir d'un câblage défectueux. Dans le cas contraire, remplacez le disque (voir « Remplacement des disques » à la page 121).

Remplacement des disques

Pour remplacer les disques qui ont entièrement échoué (détachés à la suite de la défaillance), exécutez `vxdiskadm` et sélectionnez l'option 5 (Remplacer un disque qui a échoué ou a été supprimé) dans le menu principal. Si des disques initialisés n'ont pas été ajoutés, vous pouvez en sélectionner un comme disque de remplacement.

Remarque Ne choisissez pas l'ancien disque pour le remplacement, même s'il apparaît dans la liste de sélection. Si vous ne trouvez pas de disque initialisé approprié, vous pouvez initialiser un nouveau disque.

Si un volume a été désactivé à cause d'une défaillance de disque, vous devez restaurer ce volume à partir de la sauvegarde une fois que vous avez remplacé le disque. Pour identifier les volumes qui se trouvent entièrement sur les disques désactivés à la suite d'une défaillance, entrez la commande suivante :

```
# vxinfo
```

Vous devez restaurer tous les volumes marqués comme `Unstartable` à partir de la sauvegarde. L'exemple ci-dessous illustre l'affichage de `vxinfo` :

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

Détection et remplacement des disques qui ont échoué

Pour redémarrer le volume `mkting` afin de pouvoir le restaurer à partir de la sauvegarde, utilisez la commande suivante :

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

La combinaison d'options `-o bg` resynchronise les plex en arrière-plan.

Si des défaillances commencent à se produire sur un disque, mais que ce dernier n'a pas encore entièrement échoué, remplacez-le. Cette procédure comprend deux étapes :

1. Détachez le disque de son disk group.
2. Remplacez le disque par un nouveau disque.

Pour détacher le disque, exécutez `vxdiskadm` et sélectionnez l'option 4 (Supprimer un disque pour le remplacer) du menu principal. Si des disques initialisés sont disponibles comme disques de remplacement, vous pouvez spécifier le disque voulu dans au cours de cette opération. Dans le cas contraire, vous devez spécifier le disque de remplacement par la suite en sélectionnant l'option 5 (Remplacer un disque qui a échoué ou a été supprimé) du menu principal.

Lorsque vous sélectionnez le disque à supprimer pour le remplacer, le programme affiche tous les volumes que cette opération est susceptible d'affecter. Exemple d'affichage :

```
Les volumes suivants perdront des miroirs à la suite de
cette opération :
```

```
home src
```

```
Aucune donnée de ces volumes ne sera perdue.
```

```
Les volumes suivants sont utilisés et seront désactivés
à la suite de cette opération :
```

```
mkting
```

```
L'accès pour toutes les applications utilisant ces volumes
sera refusé.
```

```
Ces volumes devront être restaurés à partir des sauvegardes.
```

```
Êtes-vous sûr de vouloir effectuer cette opération ? [y,n,q,?]
(default: n)
```

Si des volumes risquent d'être désactivés, quittez à partir de `vxdiskadm` et enregistrez le volume. Sauvegardez le volume ou déplacez-le du disque. Pour déplacer le volume `mkting` vers un autre disque que `disk02`, entrez la commande suivante :

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

Une fois le volume sauvegardé ou déplacé, exécutez de nouveau `vxdiskadm` et continuez à supprimer le disque pour le remplacer.

Une fois que vous avez supprimé un disque pour le remplacer, vous pouvez spécifier le disque de remplacement en sélectionnant l'option 5 (Remplacer un disque qui a échoué ou a été supprimé) du menu principal de `vxdiskadm`.

Création d'un disk group

Les disk groups sont généralement créés pour un ensemble particulier d'utilisateurs ou d'applications. Les disques doivent être inclus dans des disk groups avant que Volume Manager puisse les utiliser pour les volumes. Volume Manager utilise toujours le disk group par défaut `rootdg`, mais vous pouvez ajouter d'autres disk groups si nécessaire.

Remarque Les volumes sont tous créés par défaut dans `rootdg`. De même, toutes les commandes s'appliquent par défaut à `rootdg`.

Pour créer le disk group `newdg`, procédez comme suit :

1. Entrez la commande suivante pour démarrer `vxdiskadd` :

```
# vxdiskadd c1t1d0
```

2. À l'invite suivante, appuyez sur la touche de retour pour continuer :

```
Ajouter ou initialiser les disques
```

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

```
Le disque sélectionné est le suivant. Format de sortie :
```

```
[Nom_périphérique]
```

```
c1t1d0
```

```
Continuer l'opération ? [y,n,q,?] (default: y)
```

3. À l'invite suivante, spécifiez le disk group auquel ajouter le disque (`newdg`, dans ce cas) :

```
Vous pouvez choisir d'ajouter ce disque à un disk group existant, à un nouveau disk group ou de le laisser disponible pour qu'il soit utilisé pour des opérations d'ajout ou de remplacement futures. Pour créer un nouveau disk group, sélectionnez le nom d'un disk group qui n'existe pas. Pour que le disque soit disponible plus tard, nommez le disk group "none".
```

```
Quel disk group [<group>,none,list,q,?] (default: rootdg) newdg
```

Création d'un disk group

4. Lorsque `vxdiskadd` confirme qu'il n'existe actuellement aucun disk group actif ayant le même nom et vous demande de confirmer que vous voulez créer ce nouveau disk group, entrez **y** pour continuer :

Il n'existe pas de disk group actif nommé newdg.

Créer un nouveau groupe nommé newdg? [y,n,q,?] (default: y) **y**

5. À l'invite suivante, appuyez sur la touche de retour pour accepter le nom du disque par défaut ou entrez le nom voulu :

Utiliser un nom de disque par défaut pour le disque ? [y,n,q,?] (default: y)

6. Lorsque `vxdiskadd` vous demande si ce disque doit devenir un disque de remplacement pour la reconstruction automatique, entrez **n** (ou appuyez sur la touche de retour):

Ajouter un disque comme disque de remplacement pour rootdg (ou newdg) ? [y,n,q,?] (default: n) **n**

À l'invite suivante, entrez **n** pour indiquer que ce disque ne doit pas être exclu de la reconstruction automatique :

Exclure un autre disque de la reconstruction automatique ? [y,n,q,?] [default:n] **n**

7. Pour continuer la tâche, entrez **y** (ou appuyez sur la touche de retour) à l'invite suivante :

Un nouveau disk group nommé newdg sera créé et les disques sélectionnés seront ajoutés au disk group avec les noms de disque définis par défaut.

`c1t1d0`

Continuer l'opération ? [y,n,q,?] (default: y) **y**

Des messages du type suivant devraient confirmer que ce disque est en cours d'initialisation pour être utilisé par Volume Manager :

Initialisation du périphérique `c1t1d0`.

Création d'un nouveau disk group nommé newdg contenant le périphérique de disque `c1t1d0` avec le nom newdg01.

8. Pour vérifier que le disk group a été créé, entrez la commande suivante :

```
# vxdisk list
```

Volume Manager affiche ce qui suit :

PÉRIPHÉRIQUE	TYPE	DISQUE	GROUPE	ÉTAT
c0t0d0s2	sliced	disk04	rootdg	en ligne
c1t0d0s2	sliced	disk03	rootdg	en ligne
c1t1d0s2	sliced	newdg01	newdg	en ligne

Vous pouvez aussi utiliser l'opération `vxvg init` pour créer des disk groups. Pour créer un disk group au moyen de l'utilitaire `vxvg`, entrez la commande suivante :

```
# vxvg init diskgroup nom_disque=nom_périphérique
```

Par exemple, pour créer le disk group `mkt dg` sur le périphérique `c1t0d0s2`, entrez la commande suivante :

```
# vxvg init mkt dg mkt dg01=c1t0d0
```

Le nom de périphérique de disque attribué à `vxvg` doit déjà avoir été initialisé au moyen de `vxdiskadd`. Le disque ne doit pas appartenir à un disk group.

Mise à niveau d'un disk group

Avant Volume Manager 3.0, la version de disk group était automatiquement mise à niveau (le cas échéant) lorsque le disk group était importé.

La fonction de mise à niveau d'un disk group Volume Manager sépare les deux opérations d'importation d'un disk group et de mise à niveau de sa version. Vous pouvez importer un disk group appartenant à une version de niveau inférieur et l'utiliser sans le mettre à niveau.

Si vous voulez utiliser les nouvelles fonctions, vous pouvez mettre à niveau le disk group. La mise à niveau est une opération explicite. Une fois qu'elle est exécutée, le disk group n'est plus compatible avec les versions antérieures de VxVM qui ne prennent pas en charge la nouvelle version.

Avant de mettre à niveau le disk group importé, vous ne devez pas le modifier afin de ne pas empêcher son utilisation sur la version à partir de laquelle il a été importé tant que l'administrateur ne le met pas explicitement à niveau vers la version actuelle.

Tant que la mise à niveau n'est pas terminée, vous pouvez utiliser le disk group « tel quel », à condition que vous ne tentiez pas d'utiliser les fonctions de la version actuelle.

Mise à niveau d'un disk group

Toute tentative d'utilisation d'une fonction de la version actuelle ne correspondant à aucune fonction de la version à partir de laquelle le disk group a été importé entraîne l'affichage d'un message d'erreur du type suivant :

```
vxvm:vxedit : ERREUR : La version du disk group ne prend pas
en charge la fonction
```

Pour utiliser l'une de ces nouvelles fonctions, l'administrateur doit exécuter une commande afin de mettre explicitement à niveau le disk group vers une version qui les prend en charge.

Un numéro de version est associé à tous les disk groups. Les versions de Volume Manager prennent en charge un ensemble spécifique de versions de disk groups. Volume Manager peut effectuer des importations et des opérations sur un disk group de cette version. Les opérations sont limitées par les fonctions et les opérations que la version du disk group prend en charge.

Le tableau 2 récapitule les versions de VxVM qui offrent et prennent en charge des versions de disk group spécifiques, comme suit :

Tableau 2. Assignation des versions de disk groups

Version de VxVM	Offre la version	Prend en charge les versions
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

Remarque À l'exception de VxVM 3.0, toutes les versions de VxVM assurent la mise à niveau des versions de disk groups prises en charge *lorsqu'elles* sont importées.

L'importation du disk group d'une version antérieure à VxVM 3.0 empêche l'utilisation des fonctions ajoutées depuis cette version. Le tableau 3 récapitule les fonctions qui ne sont pas prises en charge par des versions de disk groups spécifiques, comme suit :

Tableau 3. Fonctions de disk groups non prises en charge

Version des disk group	Fonctions non prises en charge
40	
30	Reconstruction automatique
20	VxSmartSync Recovery Accelerator
10, 15	Volumes RAID 5, nouvelles bandes, point de reprise de récupération, limite de la copie de configuration du disk group, journal des zones modifiées, journalisation (logging) des volumes mis en miroir

Pour obtenir la liste des versions de disk groups, spécifiez un nom de disk group au moyen de la commande suivante :

```
# vxdg list nom_dg
```

Pour déterminer la version du disk group, utilisez la commande `vxprint(1M)` avec l'option de format `-l`.

VxVM met à niveau le disk group vers la version la plus élevée prise en charge par la version de VxVM en cours d'exécution. Pour mettre à jour un disk group, utilisez la commande :

```
# vxdg upgrade nom_dg
```

Par défaut, VxVM crée un disk group spécifique à la version la plus élevée prise en charge par la version de VxVM. Par exemple, VxVM 3.0 crée des disk groups de la version 60.

Il peut être nécessaire de créer un disk group spécifique à une version antérieure. Un disk group créé par défaut sur un système avec VxVM 3.0 serait un disk group de la version 60. Vous ne pourriez pas l'importer sur un système avec VxVM 2.5, qui prend uniquement en charge les disk groups jusqu'à la version 50. Par conséquent, pour créer sur un système exécutant VxVM 3.0 un disk group susceptible d'être importé par un système exécutant VxVM 2.5, vous devriez créer ce disk group avec une version 40 ou inférieure.

Pour créer un tel disk group, vous pouvez spécifier l'option `-T version` dans la commande `vxdg init`. Le paramètre `version` doit correspondre à la version de disk group voulue. Par exemple, pour créer un disk group susceptible d'être importé par un système avec VxVM 2.5, entrez la commande suivante :

```
# vxdg -T 40 init newdg newdg01=c0t3d0s2
```

Suppression d'un disk group

Cette commande crée le disk group `newdg` avec la version de disk group 40. Ce disk group peut être importé par VxVM 2.5. Cependant, les tentatives d'utilisation avec VxVM 3.0 échoueront.

Suppression d'un disk group

Pour supprimer un disk group, démontez et arrêtez ses volumes, puis entrez la commande suivante :

```
# vxdg deport diskgroup
```

Lorsque vous exportez un disk group, vous ne le supprimez pas vraiment. Cette opération désactive l'utilisation du disk group par le système. Cependant, vous pouvez réutiliser, réinitialiser ou ajouter à d'autres disk groups les disques appartenant à un disk group exporté.

Déplacement des disk groups d'un système à un autre

Les disk groups comportent une fonction importante : ils peuvent être déplacés d'un système à un autre. Si tous les disques d'un disk group sont déplacés d'un système à l'autre, le second système peut utiliser ce disk group. Il n'est pas nécessaire de spécifier de nouveau la configuration.

Procédez comme suit pour déplacer un disk group d'un système à un autre :

1. Sur le premier système, arrêtez tous les volumes du disk group, puis exportez le disk group (désactivez l'accès local) en entrant la commande suivante :

```
# vxdg deport diskgroup
```

2. Déplacez tous les disques vers le second système et exécutez les opérations nécessaires (propres au système) pour que le second système et Volume Manager puissent reconnaître les nouveaux disques.

Vous devrez peut-être réinitialiser le système, auquel cas le démon `vxconfigd` est redémarré et reconnaît les nouveaux disques. Si vous ne réinitialisez pas le système, utilisez la commande `vxdctl enable` pour redémarrer `vxconfigd` afin que Volume Manager puisse également reconnaître les disques.

3. Importez le disk group (activer l'accès local) sur le second système en entrant la commande suivante :

```
# vxdg import diskgroup
```

4. Une fois le disk group importé, démarrez tous les volumes du disk group en entrant la commande suivante :

```
# vxrecover -g diskgroup -sb
```

Vous pouvez déplacer les disques d'un système arrêté de façon anormale. Dans ce cas, vous ne pouvez pas exporter le disk group à partir du premier système. Lorsqu'un disk group est créé ou importé sur un système, ce dernier verrouille tous les disques du disk group.

Remarque Ce verrou est destiné à garantir que les *disques à double port* (les disques accessibles simultanément à partir de deux systèmes) ne sont pas utilisés par les deux systèmes à la fois. Si deux systèmes essaient de gérer les mêmes disques au même moment, les informations de configuration situées sur le disque seront endommagées. Le disque et les données qu'il contient deviennent alors inutilisables.

Si vous déplacez des disques à partir d'un système qui s'est arrêté de façon anormale ou qui n'a pas pu détecter le groupe avant le déplacement du disque, les verrous placés sur les disques demeurent en place et vous devrez les effacer. Le système renvoie le message d'erreur suivant :

```
vxdbg:disk group nom_groupe : import a échoué : Le disque est  
utilisé par un autre hôte
```

Pour effacer les verrous placés sur un ensemble spécifique de périphériques, entrez la commande suivante :

```
# vxdisk clearimport nom_périphérique ...
```

Vous pouvez effacer les verrous pendant l'importation en entrant la commande suivante :

```
# vxdbg -C import diskgroup
```

Remarque Veillez à utiliser avec précaution la commande `vxdisk clearimport` ou `vxdbg -C import` sur des systèmes comprenant des disques à double port. L'effacement des verrous permet à plusieurs hôtes d'accéder simultanément à ces disques, ce qui risque d'entraîner l'endommagement des données.

Dans certains cas, vous voudrez peut-être importer un disk group alors que certains disques ne sont pas disponibles. L'opération `import` échoue généralement s'il est impossible de trouver certains disques du disk group dans les unités de disques attachées au système. Si l'opération `import` échoue, l'un des messages d'erreur suivants s'affiche :

```
vxdbg: Disk group nom_groupe : import a échoué : Le disk group n'a  
pas de copies de configuration valides
```

Déplacement des disk groups d'un système à un autre

Ce message indique une erreur fatale qui nécessite une réparation matérielle ou la création d'un nouveau disk group.

```
vxldg: Disk group nom_groupe : import a échoué : Disque  
introuvable pour le disk group
```

Ce message indique une erreur récupérable.

Si certains disques du disk group ont échoué, vous pouvez forcer l'importation du disk group en entrant la commande suivante :

```
# vxldg -f import diskgroup
```

Remarque Vous devez utiliser l'option `-f` avec précaution. Elle peut entraîner la double importation du même disk group à partir de jeux de disques différents, ce qui provoquera une incohérence du disk group.

Vous pouvez effectuer ces opérations au moyen de `vxldiskadm`. Pour exporter un disk group au moyen de `vxldiskadm`, sélectionnez l'option de menu 9 (Désactiver l'accès à un disk group (exporter)). Pour importer un disk group, sélectionnez l'option de menu 8 (Activer l'accès à un disk group (importer)). L'opération d'importation `vxldiskadm` vérifie les verrous d'importation hôte et vous demande si vous voulez effacer les verrous trouvés. Elle démarre également les volumes dans le disk group.

Renommer les disk groups

Il ne peut y avoir qu'un seul disk group avec un nom donné par système. Vous ne pouvez pas importer ou exporter un disk group lorsque le système cible contient déjà un disk group du même nom. Pour éviter ce problème, Volume Manager vous permet de renommer un disk group pendant l'importation ou l'exportation.

Par exemple, puisque chaque système exécutant Volume Manager doit posséder un seul disk group par défaut `rootdg`, l'importation ou l'exportation de `rootdg` d'un système à l'autre risque de poser des problèmes. Il ne peut y avoir deux disk groups `rootdg` sur le même système. Pour éviter ce problème, renommez le disk group `rootdg` au cours de l'importation ou de l'exportation.

Pour renommer le disk group pendant l'importation, entrez la commande suivante :

```
# vxldg [-t] -n nom_nouveau_dg import diskgroup
```

Si l'option `-t` est utilisée, l'importation est provisoire et ne continue pas après les redémarrages. Dans ce cas, le disk group conserve son nom enregistré sur son hôte d'origine, mais il est nommé `nom_nouveau-dg` sur l'hôte d'importation. Si vous n'utilisez pas l'option `-t`, le nom est définitivement changé.

Vous pouvez également renommer un disk group pendant une exportation en entrant la commande suivante :

```
# vxdg [-h nom_hôte] -n nom_nouveau_dg deport diskgroup
```

Lorsque vous choisissez de renommer le disk group pendant son exportation, vous pouvez spécifier l'option `-h nom_hôte` pour assigner un verrou à un autre hôte. Cette procédure garantit que le disk group est automatiquement importé au redémarrage de l'autre hôte.

Pour déplacer provisoirement le disk group `rootdg` d'un hôte vers un autre (par exemple, pour effectuer une réparation sur le volume racine), puis le ramener sur son hôte d'origine, procédez comme suit :

1. Sur l'hôte d'origine, identifiez l'ID du disk group `rootdg` à importer vers l'autre hôte en entrant la commande suivante :

```
# vxdisk -s list
```

L'affichage de sortie inclut les informations suivantes sur les disk groups :

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

2. Sur l'hôte d'importation, importez et renommez le disk group `rootdg` en entrant la commande suivante :

```
# vxdg -tC -n newdg_name import diskgroup
```

où `-t` indique un nom d'importation provisoire ; `-C` efface les verrous d'importation ; `-n` spécifie un nom provisoire pour le disk group `rootdg` à importer (afin d'éviter tout conflit avec le disk group `rootdg` existant) ; et *diskgroup* est l'ID du disk group en cours d'importation (par exemple, `774226267.1025.tweety`).

En cas de redémarrage ou d'arrêt anormal du système, le disk group provisoirement importé ne l'est plus et il doit être réimporté.

3. Une fois que vous avez effectué les opérations nécessaires sur le disk group `rootdg` importé, exportez-le sur son hôte d'origine en entrant la commande suivante :

```
# vxdg -h nom_hôte deport diskgroup
```

où *nom_hôte* représente le nom du système dont le disk group `rootdg` est en cours de renvoi (vous pouvez confirmer le nom du système au moyen de la commande `uname -n`).

Déplacement des disk groups d'un système à un autre

Cette commande supprime le `rootdg` importé de l'hôte de destination et renvoie les verrous sur l'hôte d'origine. Ce dernier importe alors automatiquement son `rootdg` au redémarrage suivant.

Réservation des mineurs pour les disk groups

Volume Manager vous permet de sélectionner une plage de mineurs pour un disk group spécifié. Vous pouvez utiliser cette plage de mineurs pendant la création d'un volume. Cette opération garantit que chaque volume utilise le même mineur tout le long des redémarrages et des reconfigurations. Si les plages de deux disk groups se chevauchent, une collision d'importation est détectée et un mécanisme de contournement ou de renumérotation est alors nécessaire.

Si vous attribuez des codes de périphérique de volume appartenant à des plages séparées pour chaque disk group, vous pouvez déplacer tous les disk groups d'un groupe de système sans provoquer de collision entre les codes de périphérique.

Entrez la commande suivante afin de définir un mineur pour un périphérique de volume de base pour un disk group :

```
# vxdg init diskgroup minor=mineur_base nom_périphérique
```

Les numéros de périphérique de volume pour un disk group ont un mineur commençant au numéro *mineur_base*. Les mineurs (sur la plupart des systèmes) peuvent aller jusqu'à 131071. Une plage de taille raisonnable peut être laissée à la fin pour les remappages des numéros de périphérique provisoires (si deux numéros de périphérique sont toujours en conflit).

Si vous ne spécifiez pas l'argument `minor` sur la ligne de commande `vxdg init`, Volume Manager choisit un numéro aléatoire. Le numéro choisi est au moins égal à 1000 ou il est un multiple de 1000, et il renvoie une plage utilisable de 1000 numéros de périphérique. Le numéro choisi par défaut ne recouvre pas la plage de 1000 des disk groups actuellement importés. Il ne recouvre pas les numéros de périphérique de volume actuellement attribués.

Remarque La politique par défaut garantit qu'un petit nombre de disk groups peut être fusionné entre un ensemble d'ordinateurs. Cependant, lorsque les disk groups sont fusionnés automatiquement au moyen de mécanismes de basculement (failover), vous devriez sélectionner des plages qui évitent les chevauchements.

Pour plus d'informations sur la réservation des mineurs, voir la page `man vxdg(1M)`.

Destruction d'un disk group

La commande `vxdg` comprend une option de destruction qui supprime un disk group du système et libère ses disques afin que d'autres disk groups puissent les utiliser. Pour supprimer les disk groups inutiles afin que d'autres disk groups puissent utiliser leurs disques, utilisez la commande suivante :

```
# vxdg destroy diskgroup
```

La commande `vxdg deport` peut toujours être utilisée pour rendre les disques inaccessibles. Volume Manager empêche d'autres disk groups d'utiliser les disques d'un disk group exporté.

Utilisation de périphériques spéciaux

Cette section décrit les périphériques spéciaux utilisés par Volume Manager pour effectuer des tâches d'administration.

Utilisation de `vxdisk` pour des encapsulations spéciales

L'*encapsulation* est une procédure qui convertit en volumes des partitions d'un disque spécifié. Si des partitions contiennent des systèmes de fichiers, les entrées `/etc/vfstab` sont modifiées de façon à ce que les systèmes de fichiers soient montés à la place sur des volumes.

L'encapsulation nécessite que le disque ait de l'espace disponible pour stocker les informations d'identification et de configuration de Volume Manager. Cet espace disponible ne doit être inclus dans aucune autre partition. (Pour plus d'informations, reportez-vous au *Guide d'installation de VERITAS Volume Manager* et à la page man `vxencap(1M)`.)

Vous pouvez encapsuler un disque qui n'a pas d'espace disponible pour la partition de zone privée de Volume Manager. L'utilitaire `vxdisk` encapsule les disques dépourvus d'espace disponible. Cette procédure repose sur l'utilisation de périphériques spéciaux, appelés `nopriv`, qui sont dépourvus de zones privées.

Pour utiliser `vxdisk`, créez une partition sur le périphérique de disque qui mappe toutes les parties du disque auxquelles vous voulez accéder. Ajoutez ensuite le périphérique pour cette partition en entrant la commande suivante :

```
# vxdisk define périphérique_partition type=nopriv
```

Utilisation de périphériques spéciaux

Où *périphérique_partition* représente le nom de base du périphérique dans le répertoire /dev/dsk. Par exemple, pour utiliser la partition 3 du périphérique de disque c0t4d0, entrez la commande suivante :

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

Vous pouvez créer des volumes pour les autres partitions du périphérique de disque en :

- ◆ ajoutant le périphérique à un disk group ;
- ◆ déterminant l'emplacement de ces partitions dans la partition d'encapsulation ;
- ◆ utilisant *vxassist* pour créer un volume ayant ce décalage et cette longueur.

Par défaut, *vxassist* réinitialise la zone de données d'un volume qu'il crée. Si la partition contient des données à sauvegarder, n'utilisez pas *vxassist*. Créez le volume avec *vxmlake* et démarrez-le avec *vxvol init active*.

L'utilisation des périphériques *nopriv* présente un inconvénient : Volume Manager ne peut pas assurer le suivi des modifications de l'adresse ou du contrôleur du disque. Généralement, Volume Manager utilise les informations d'identification stockées dans la zone privée du disque physique pour assurer le suivi des modifications à l'emplacement d'un disque physique. Le suivi est impossible car les périphériques *nopriv* sont dépourvus de zones privées et ne comprennent aucune information d'identification sur le disque physique.

L'utilisation optimale des périphériques de partition d'encapsulation spéciale consiste à encapsuler un disque de façon à ce que Volume Manager puisse déplacer de l'espace sur le disque. Lorsque vous libérez de l'espace sur le disque, vous pouvez supprimer le périphérique de partition spéciale, puis encapsuler le disque en tant que périphérique de disque standard.

Vous ne pouvez pas constituer un disk group exclusivement de périphériques *nopriv*. En effet, les périphériques *nopriv* ne fournissent pas d'espace pour le stockage des informations de configuration des disk groups. Vous devez placer ces informations sur au moins un disque du disk group.

Utilisation de vxdisk pour les disques RAM

Remarque Cette section ne s'applique qu'aux systèmes équipés de disques RAM.

Certains systèmes prennent en charge la création de disques RAM. Un disque RAM est un périphérique composé d'une RAM système similaire à un petit périphérique de disque. Le contenu d'un disque RAM est souvent effacé au redémarrage du système. Dans ce cas, Volume Manager ne peut pas identifier les disques physiques. En effet, les informations situées sur les disques physiques (qui ont été effacés au redémarrage) permettent d'identifier le disque.

Les périphériques `nopriv` comprennent une fonction spéciale qui permet de prendre en charge les disques RAM : l'option `volatile` qui indique à Volume Manager que le contenu du périphérique ne « survit » pas aux redémarrages. Les périphériques volatiles font l'objet d'un traitement spécial au démarrage du système. Si un volume est mis en miroir, les plex composés de ces périphériques volatiles sont toujours récupérés à l'aide d'une copie des données effectuée à partir des plex non volatiles.

Pour utiliser un disque RAM, vous devez créer un nœud de périphérique pour le disque dans les répertoires `/dev/dsk` et `/dev/rdisk` (par exemple, `/dev/dsk/ramd0` et `/dev/rdisk/ramd0`). Pour définir le périphérique de disque RAM sur Volume Manager, utilisez la commande suivante :

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

Généralement, Volume Manager ne démarre pas les volumes qui sont exclusivement constitués de plex contenant des sous-disques volatiles. En effet, il n'est pas garanti que le contenu le plus récent du volume se trouve dans un plex.

Certains disques RAM sont utilisés lorsque le contenu de tous les volumes est recréé après le redémarrage. Dans ce cas, vous pouvez forcer le démarrage des volumes constitués de disque RAM au redémarrage en entrant la commande suivante :

```
# vxvol set startopts=norecov nom_volume
```

Vous pouvez uniquement utiliser cette option avec les volumes de type `gen`. Pour plus d'informations sur l'opération `vxvol set` et l'option `norecov`, voir la page `man vxvol(1M)`.

Utilisation de `vxdisk` pour afficher des informations de chemins d'accès multiples

Remarque Cette section s'applique uniquement aux systèmes équipés de la fonction de chemin d'accès multiple (DMP, *Dynamic Multipathing*).

Dans Volume Manager, des disques physiques connectés au système sont représentés sous la forme de méta-périphériques ayant un ou plusieurs chemins d'accès physique. Les chemins d'accès dépendent de la nature du disque : un disque unique ou un composant d'un ensemble de disques à plusieurs ports connecté au système. L'utilitaire `vxdisk` vous permet d'afficher les chemins d'accès d'un méta-périphérique et d'afficher l'état de chaque chemin (par exemple, activé ou désactivé). Par exemple, pour afficher les détails du disque `disk01`, entrez la commande suivante :

```
# vxdisk list disk01
```

Opérations de l'interface de menu vxdiskadm

Volume Manager affiche les informations suivantes :

```

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

```

Cet affichage indique deux chemins d'accès à un périphérique physique représenté par le méta-périphérique c2t0d0s2. Le chemin c2t0d0s2 est actif et l'état de l'autre chemin c1t0d0s2 est échoué.

Opérations de l'interface de menu vxdiskadm

Cette section décrit l'interface de menu vxdiskadm (Volume Manager Support Operations). vxdiskadm permet d'exécuter les opérations d'administration des périphériques physiques et logiques.

Cette section contient les informations suivantes sur l'exécution des opérations de disque et de volume au moyen de vxdiskadm :

- ◆ Placement des disques physiques sous le contrôle de Volume Manager
- ◆ Ajout d'un disque pour une utilisation future
- ◆ Réinitialisation d'un disque pour être utilisé par Volume Manager

- ◆ Ajout d'un disque VM au pool de reconstruction automatique
- ◆ Suppression d'un disque VM du pool de reconstruction automatique
- ◆ Affichage d'informations sur les disques physiques
- ◆ Suppression d'un disque physique
- ◆ Désactivation d'un disque physique (mise hors ligne)
- ◆ Activation d'un disque physique
- ◆ Remplacement d'un disque physique
- ◆ Remplacement d'un disque qui a échoué ou qui a été supprimé
- ◆ Ajout d'un disque à un disk group
- ◆ Création d'un disk group
- ◆ Exportation d'un disk group
- ◆ Importation d'un disk group

Démarrage de vxdiskadm

Pour démarrer vxdiskadm, entrez la commande suivante :

```
# vxdiskadm
```

Le menu principal vxdiskadm s'affiche.

Menu principal vxdiskadm

Le menu principal vxdiskadm est le suivant :

```
Volume Manager Support Operations
Menu : VolumeManager/Disk

1  Ajouter ou initialiser un ou plusieurs disques
2  Encapsuler un ou plusieurs disques
3  Supprimer un disque
4  Supprimer un disque pour le remplacer
5  Remplacer un disque qui a échoué ou a été supprimé
6  Mettre en miroir des volumes sur un disque
7  Déplacer les volumes d'un disque
8  Activer l'accès à un disk group (importer)
9  Désactiver l'accès à un disk group (exporter)
10 Activer un périphérique de disque (mettre en ligne)
11 Désactiver un périphérique de disque (mettre hors ligne)
12 Marquer un disque comme disque de remplacement pour un disk
    group
```

Opérations de l'interface de menu vxdiskadm

- 13 Désactiver l'indicateur de remplacement pour un disque
- 14 Reconstruire les sous-disques sur un disque
- 15 Exclure un disque de la reconstruction automatique
- 16 Libérer un disque pour la reconstruction automatique

Lister les informations sur le disque

- ? Afficher l'aide sur le menu
- ?? Afficher l'aide sur le système de menus
- q Quitter les menus

Sélectionnez une opération à exécuter

- ◆ Vous pouvez entrer ? à tout moment pour afficher l'aide sur le menu. ? affiche la liste des opérations et leur définition.
- ◆ ?? liste les entrées utilisables à n'importe quelle invite.
- ◆ q renvoie au menu principal si vous devez redémarrer un processus ; cependant, l'utilisation de q au niveau du menu principal ferme Volume Manager Support Operations.

Le menu vxdiskadm permet d'accéder aux tâches suivantes. Les numéros correspondent aux options listées dans le menu principal:

1. Ajouter ou initialiser un ou plusieurs disques.

Vous pouvez ajouter des disques formatés au système. Les disques SCSI sont déjà formatés. Pour les autres disques, reportez-vous à la documentation du constructeur pour obtenir les instructions de formatage. Le système vous invite à indiquer le ou les périphériques de disque. Vous pouvez spécifier le disk group auquel ajouter le ou les disques ; si vous sélectionnez none, le disque est conservé comme disque de remplacement pour les opérations futures ou les remplacements de disque sans redémarrage. Vous pouvez également indiquer que les disques sélectionnés soient marqués comme disques de remplacement pour la reconstruction automatique pour un disk group. Si le disque n'a pas encore été initialisé, il est partitionné et initialisé pour être utilisé avec Volume Manager.

2. Encapsuler un ou plusieurs disques.

Vous pouvez mettre sous le contrôle de Volume Manager un disque ajouté au système avant l'installation de Volume Manager. Le système vous invite à indiquer le ou les périphériques de disque, le disk group et les noms des disques. Le disque est ajouté au disk group spécifié. Le programme recherche dans le disque les partitions utilisées pour les systèmes de fichiers ou à d'autres fins. Des volumes sont créés pour remplacer les partitions de disque afin d'accéder aux données existantes. Si l'encapsulation ne peut pas déterminer automatiquement le but d'une partition, le système vous demande ce que vous voulez en faire. Vous pouvez remplacer la partition par un volume, la laisser telle quelle ou la supprimer.

Vous devez redémarrer le système si des partitions sont utilisées pour les systèmes de fichiers montés ou pour exécuter des applications. Vous devrez peut-être modifier les fichiers de configuration des applications afin d'utiliser les volumes, plutôt que les périphériques de disque directs, pour accéder aux partitions de disque. Les informations de montage du système de fichiers sont ajustées automatiquement.

3. Supprimer un disque.

Vous pouvez supprimer un disque d'un disk group. Le programme vous demande le nom du disque à supprimer. Vous ne pouvez pas supprimer un disque s'il sert de support de stockage à un ou plusieurs volumes. Dans ce cas, vous pouvez demander à Volume Manager de déplacer cet espace de stockage sur d'autres disques du disk group.

Remarque Vous ne pouvez pas utiliser cette option pour supprimer le dernier disque d'un disk group. Si vous souhaitez utiliser tous les disques restants d'un disk group, vous devez désactiver (exporter) le disk group. Vous pourrez alors réutiliser les disques. Cependant, vous ne pouvez pas exporter rootdg.

4. Supprimer un disque pour le remplacer.

Vous pouvez supprimer un disque physique d'un disk group, tout en conservant son nom. L'état du disque spécifié devient `removed`. Si des disques initialisés n'appartiennent pas à un disk group, vous pouvez utiliser l'un de ces disques comme disque de remplacement.

5. Remplacer un disque qui a échoué ou a été supprimé.

Vous pouvez spécifier un disque de remplacement pour un disque que vous avez supprimé au moyen de la commande `Supprimer un disque` pour le remplacer ou pour un disque qui a échoué en cours d'utilisation. Le système vous demande d'indiquer le nom du disque à remplacer ainsi que le périphérique de disque à utiliser comme remplacement. Vous pouvez choisir un disque non initialisé, auquel cas le disque le sera ou vous pouvez choisir un disque que vous avez déjà initialisé au moyen de la commande `Ajouter` ou `initialiser un ou plusieurs disques`.

6. Mettre en miroir des volumes sur un disque

Vous pouvez mettre en miroir des volumes sur un disque. Ces volumes peuvent être mis en miroir sur un autre disque ayant de l'espace disponible. Cette création de copies miroir des volumes protège les données contre tout risque de perte en cas de défaillance de disque. Cette option ne permet pas de mettre en miroir les volumes qui le sont déjà ou qui sont composés de plusieurs sous-disques. La mise en miroir des volumes à partir du disque d'initialisation créera un disque qui pourra être utilisé comme autre disque d'initialisation.

Opérations de l'interface de menu vxdiskadm

7. Déplacer les volumes d'un disque.

Vous pouvez déplacer vers d'autres disques les volumes (ou des parties du volume) qui utilisent un disque. Utilisez cette option du menu juste avant de supprimer un disque soit de façon définitive, soit à titre de remplacement.

Remarque Lorsque vous déplacez simplement les volumes d'un disque, sans supprimer ce dernier, vous n'empêchez pas le déplacement des autres volumes sur le disque pour les opérations futures.

8. Activer l'accès à un disk group (importer).

Vous pouvez activer l'accès de ce système à un disk group. Pour déplacer un disk group d'un système à un autre, vous devez d'abord le désactiver (exporter) sur le système d'origine. Vous devez ensuite déplacer les disques du disk group exporté sur l'autre système et activer (importer) le disk group. Le système vous demande le nom du disk group.

9. Désactiver l'accès à un disk group (exporter)

Vous pouvez désactiver l'accès de ce système à un disk group actuellement activé (importé). Vous devez exporter un disk group si vous souhaitez déplacer les disques d'un disk group vers un autre système. D'autre part, vous pouvez exporter un disk group si vous voulez utiliser tous les disques restants à d'autres fins.

Le système vous demande le nom du disk group. Vous devez indiquer si les disques doivent être désactivés (mis hors ligne). Pour les périphériques de disque amovibles sur certains systèmes, il est important de désactiver tout accès au disque avant de le supprimer.

10. Activer un périphérique de disque (mettre en ligne).

Si vous déplacez un disque d'un système vers un autre pendant un fonctionnement normal, Volume Manager ne reconnaît pas automatiquement le disque. Utilisez cette option de menu pour indiquer à Volume Manager d'analyser le disque pour l'identifier et pour déterminer s'il appartient à un disk group. De plus, cette option permet de réactiver l'accès à un disque désactivé soit via l'exportation du disk group, soit via la désactivation (mise hors ligne) du périphérique de disque.

11. Désactiver un périphérique de disque (mettre hors ligne).

Vous pouvez désactiver tous les accès à un périphérique de disque à l'aide de Volume Manager. Cette option s'applique uniquement aux disques qui ne se trouvent pas actuellement dans un disk group. Vous pouvez utiliser cette option pour supprimer un disque sans redémarrer le système.

Certains systèmes ne prennent pas en charge les disques qui peuvent être supprimés pendant un fonctionnement normal. Sur ces systèmes, la mise hors ligne est rarement utile.

12. Marquer un disque comme disque de remplacement pour un disk group.

Vous pouvez réserver un disque en tant que disque de remplacement automatique (pour la reconstruction automatique) dans le cas où un autre disque du disk group échouerait.

13. Désactiver l'indicateur de remplacement pour un disque.

Cette option permet de supprimer un disque de l'ensemble de disques qui peuvent être utilisés pour le remplacement et rendre l'espace de nouveau disponible.

14. Reconstruire les sous-disques sur un disque.

La reconstruction automatique de VxVM permet au système de répondre automatiquement aux erreurs d'E/S sur un objet VxVM redondant au niveau des sous-disques et de prendre les actions nécessaires pour rendre l'objet de nouveau disponible. Ce mécanisme détecte les erreurs d'E/S sur un sous-disque, reconstruit ce dernier et récupère le plex qui lui est associé. Une fois le disque remplacé, l'utilitaire `vxunrelloc` de Volume Manager permet de restaurer sur le système la configuration existant avant la défaillance du disque. `vxunrelloc` vous permet de déplacer les sous-disques reconstruits automatiquement sur un disque remplacé à la suite d'une défaillance de disque.

15. Exclure un disque de la reconstruction automatique.

Vous pouvez exclure de l'ensemble de disques disponibles (les disques ne servant pas au remplacement) les disques à utiliser pour la reconstruction automatique.

16. Libérer un disque pour la reconstruction automatique.

Cette opération annule l'étape 15 et permet de créer dans l'ensemble de disques disponibles (disques ne servant pas au remplacement) des disques disponibles pour la reconstruction automatique.

Lorsque vous effectuez l'administration de disques, il est important que vous puissiez distinguer le *nom du périphérique* et le *nom du disque*.

Remarque Il se peut que votre système utilise un nom de périphérique différent de ceux donnés dans les exemples. Voir « Présentation de Volume Manager » à la page 17 pour plus d'informations sur les noms des périphériques.

Opérations de l'interface de menu vxdiskadm

Le *nom du périphérique* (parfois appelé *devname* ou *nom d'accès physique au disque*) représente l'emplacement du disque. Sa syntaxe est `c#b#t#d#s#`, où :

- ◆ `c#` représente le numéro du contrôleur auquel l'unité de disque est reliée ;
- ◆ `b#` représente le numéro de bus correspondant (s'il est utilisé sur le système) ;
- ◆ `t#` représente le numéro du disque cible sur ce contrôleur ;
- ◆ `d#` représente le numéro du disque ;
- ◆ `s#` représente le numéro de la partition de disque.

Le nom de chemin d'accès complet d'un périphérique est `/dev/vx/dmp/nom_périphérique`. Ce document indique uniquement le nom du périphérique et suppose `/dev/vx/dmp`. Par exemple, le nom du périphérique peut être `c0t0d0s2`.

Le *nom du disque* (parfois appelé *nom d'accès logique au disque*) est le nom d'administration du disque, par exemple `disk01`. Lorsque vous n'attribuez pas de nom de disque, le disque est nommé par défaut `disk##` s'il est ajouté à `rootdg` (où `##` représente un numéro de séquence). Dans le cas contraire, le disque est nommé par défaut `groupname##`, où *groupname* représente le nom du disk group auquel le disque est ajouté.

Placement des disques physiques sous le contrôle de Volume Manager

Lorsque vous ajoutez un disque à un système qui exécute Volume Manager, vous devez le placer sous le contrôle de ce dernier de façon à ce qu'il puisse contrôler l'allocation d'espace sur le disque.

À moins qu'un autre disk group ne soit spécifié, Volume Manager place les nouveaux disques dans le disk group par défaut, `rootdg`. Les instructions permettant de créer des disk groups supplémentaires sont fournies dans la suite de ce chapitre.

La méthode que vous utilisez pour placer un disque sous le contrôle de Volume Manager dépend des circonstances :

- ◆ S'il s'agit d'un nouveau disque, vous devez l'initialiser et le placer sous le contrôle de Volume Manager (voir « Placement d'un disque sous le contrôle de Volume Manager » à la page 143).
- ◆ Si vous n'avez pas besoin du disque dans l'immédiat, vous pouvez l'initialiser (mais pas l'ajouter à un disk group) et le réserver pour une utilisation future (voir « Ajout d'un disque pour une utilisation future » à la page 155).
- ◆ Si le disque a déjà été initialisé pour une utilisation future par Volume Manager, vous pouvez le réinitialiser et le placer sous le contrôle de Volume Manager (voir « Réinitialisation d'un disque pour être utilisé par Volume Manager » à la page 155).

- ◆ Si le disque était déjà en cours d'utilisation, mais pas sous le contrôle de Volume Manager, vous pouvez conserver les données qu'il contient tout en le plaçant sous le contrôle de Volume Manager. Pour cela, vous devez utiliser l'encapsulation (voir « Encapsulation d'un disque pour être utilisé par Volume Manager » à la page 152).
- ◆ Vous pouvez placer simultanément sous le contrôle de Volume Manager plusieurs disques sur un ou plusieurs contrôleurs. Suivant les circonstances, il se peut que ces disques ne fassent pas tous l'objet du même traitement (voir « Placement de plusieurs disques sous le contrôle de Volume Manager » à la page 147).

Lorsque vous initialisez ou encapsulez plusieurs disques en même temps, vous pouvez exclure certains disques ou certains contrôleurs. Pour exclure des disques, indiquez leur nom dans le fichier `/etc/vx/disks.exclude` avant l'initialisation ou l'encapsulation. De même, vous pouvez exclure tous les disques placés sur des contrôleurs spécifiques à partir de l'initialisation ou de l'encapsulation en indiquant ces contrôleurs dans le fichier `/etc/vx/cntrls.exclude`.

Les sections suivantes contiennent des exemples détaillés des diverses façons dont vous pouvez utiliser `vxdiskadm` afin de placer les disques sous le contrôle de Volume Manager dans diverses circonstances.

Remarque Vous devez formater (par exemple, au moyen de la commande `format`) ou ajouter (au moyen de la commande `diskadd`) un disque au système avant de pouvoir le placer sous le contrôle de Volume Manager.

Si vous tentez de placer un disque non formaté sous le contrôle de Volume Manager à l'aide de `vxdiskadm`, l'initialisation commence normalement, mais se ferme en affichant un message vous indiquant que le disque ne semble pas valide et ne peut pas être formaté. Dans ce cas, vous devez formater le disque correctement avant de tenter de nouveau de le placer sous le contrôle de Volume Manager.

Placement d'un disque sous le contrôle de Volume Manager

Cette section décrit la méthode à utiliser pour placer un disque formaté sous le contrôle de Volume Manager. Ce disque peut être nouveau ou avoir été précédemment utilisé en dehors de Volume Manager.

Remarque L'initialisation ne conserve pas les données situées sur les disques.

Opérations de l'interface de menu vxdiskadm

Initialisez un seul disque pour que Volume Manager puisse l'utiliser en procédant comme suit :

1. Sélectionnez l'option de menu 1 (Ajouter ou initialiser un ou plusieurs disques) dans le menu principal de vxdiskadm.
2. À l'invite suivante, entrez le nom de périphérique du disque à placer sous le contrôle de Volume Manager (ou entrez **list** pour une liste de disques) :

```
Ajouter ou initialiser les disques
Menu : VolumeManager/Disk/AddDisks
```

Utilisez cette opération pour ajouter un ou plusieurs disques dans un disk group. Vous pouvez ajouter les disques sélectionnés dans un disk group existant ou dans un nouveau disk group qui sera créé lors de l'opération. Les disques sélectionnés peuvent également être ajoutés dans un disk group comme disques de remplacement. Ils peuvent aussi être initialisés sans être ajoutés à un disk group, ce qui laisse les disques disponibles comme disques de remplacement.

Plusieurs disques ou modèles peuvent être entrés à l'invite. Voici des exemples de sélections de disque :

```
all :          tous les disques
c3 c4t2 :     tous les disques sur les contrôleurs 3 et 4,
              cible 2
c3t4d0 :     un seul disque
```

Sélectionnez les périphériques de disque à ajouter :
 [<pattern-list>,all,list,q,?] **list**

où <pattern-list> peut représenter un seul disque ou une série de disques et/ou de contrôleurs (avec des cibles facultatives). Si <pattern-list> comprend plusieurs éléments, vous devez les séparer par un espace.

Si vous entrez **list** à l'invite, vxdiskadm affiche la liste des disques disponibles sur le système, suivi d'une invite à entrer le nom de périphérique du disque à ajouter :

PÉRIPHÉRIQUE	DISQUE	GROUPE	ÉTAT
c0t0d0	disk01	rootdg	en ligne
c0t1d0	disk02	rootdg	en ligne
c0t2d0	disk03	rootdg	en ligne
c0t3d0	-	-	en ligne
c1t0d0	disk10	rootdg	en ligne
c1t0d1	-	-	erreur
.			
.			
c3t0d0	-	-	erreur
c3t1d0	disk33	rootdg	en ligne
c3t2d0	disk34	rootdg	en ligne
c3t3d0	disk35	rootdg	en ligne

Sélectionnez les périphériques de disque à ajouter :
[<pattern-list>,all,list,q,?] c1t0d1

Volume Manager reconnaît tous les disques attachés au système et les affiche ici.

Le mot **erreur** dans la colonne **ÉTAT** vous indique que le disque correspondant n'a pas encore été mis sous le contrôle de Volume Manager. Il est possible que ces disques aient été initialisés avant. Vous ne pouvez pas utiliser pour cette opération les disques listés avec un nom et un disk group car ils sont déjà sous le contrôle de Volume Manager.

3. Pour continuer l'opération, entrez **y** (ou appuyez sur la touche de retour) à l'invite suivante :

Le disque sélectionné est le suivant. Format de sortie :
[Nom_périphérique]

c1t2d0

Continuer l'opération ? [y,n,q,?] (default: y) **y**

Opérations de l'interface de menu vxdiskadm

4. À l'invite suivante, spécifiez le disk group auquel ajouter le disque ou appuyez sur la touche de retour pour accepter rootdg :

Vous pouvez ajouter ce disque à un disk group existant, à un nouveau disk group ou de le laisser disponible afin qu'il soit utilisé pour des opérations d'ajout ou de remplacement futures. Pour créer un nouveau disk group, sélectionnez le nom d'un disk group qui n'existe pas. Pour que le disque soit disponible plus tard, nommez le disk group "none".

Quel disk group [<group>,none,list,q,?] (default: rootdg)

5. À l'invite suivante, appuyez sur la touche de retour pour accepter le nom du disque par défaut ou entrez le nom voulu :

Utiliser un nom de disque par défaut pour le disque ? [y,n,q,?] (default: y)

6. Lorsque vxdiskadm vous demande si ce disque doit devenir un disque de remplacement pour la reconstruction automatique, entrez **n** (ou appuyez sur la touche de retour) :

Ajouter un disque comme disque de remplacement pour rootdg ? [y,n,q,?] (default: n) **n**

7. Lorsque vxdiskadm vous demande s'il faut exclure ce disque de la reconstruction automatique, entrez **n** (ou appuyez sur la touche de retour) :

Exclure le disque de la reconstruction automatique ? [y,n,q,?] (default: n)

8. Pour continuer l'opération, entrez **y** (ou appuyez sur la touche de retour) à l'invite suivante :

Les disques sélectionnés seront ajoutés au disk group rootdg avec les noms de disque définis par défaut.

c1t2d0

Continuer l'opération ? [y,n,q,?] (default: y) **y**

9. Si vous êtes certain que ce disque ne contient pas de données, entrez **n** pour éviter l'encapsulation. Lorsque vxdiskadm vous invite à initialiser le disque à la place, entrez **y** :

Le périphérique de disque suivant a un VTOC valide, mais ne semble pas avoir été initialisé pour Volume Manager. Si le disque contient des données, elles NE devraient PAS être

détruites ; vous devriez encapsuler les partitions du disque existantes comme volumes, au lieu d'ajouter le disque comme nouveau disque.

Format de sortie : [Nom_périphérique]

c1t2d0

Encapsuler ce périphérique ? [y,n,q,?] (default: y) **n**

c1t2d0

Au lieu d'encapsuler, initialiser ? [y,n,q,?] (default: n) **y**

Des messages du type suivant devraient confirmer que le disque c1t2d0 est placé sous le contrôle de Volume Manager. Selon le système, vous pouvez également choisir d'effectuer une analyse de surface.

Initialisation du périphérique c1t2d0.

Exécuter l'analyse de surface (fortement recommandé)
[y,n,q,?] (default: y) **n**

Ajout du périphérique de disque c1t2d0 au disk group rootdg avec le nom de disque disk39.

- 10.** À l'invite suivante, indiquez si vous voulez initialiser d'autres disques (**y**) ou revenir au menu principal de vxdiskadm (**n**) :

Ajouter ou initialiser d'autres disques ? [y,n,q,?] (default: n)

Placement de plusieurs disques sous le contrôle de Volume Manager

Cette section décrit la procédure à suivre pour placer simultanément plusieurs disques sous le contrôle de Volume Manager. Le jeu de disques peut inclure tous les disques du système, tous les disques d'un contrôleur, des disques sélectionnés ou une combinaison de tous ces éléments.

Suivant les circonstances, il est possible que tous les disques ne fassent pas l'objet du même traitement. Par exemple, certains peuvent être initialisés, alors que d'autres peuvent être encapsulés.

Remarque L'initialisation ne conserve pas les données situées sur les disques.

Lorsque vous initialisez ou encapsulez plusieurs disques en même temps, vous pouvez exclure certains disques ou certains contrôleurs. Pour exclure des disques, indiquez leur nom dans le fichier `/etc/vx/disks.exclude` avant l'initialisation ou l'encapsulation.

Opérations de l'interface de menu vxdiskadm

De même, vous pouvez exclure tous les disques placés sur des contrôleurs spécifiques à partir de l'initialisation ou de l'encapsulation en indiquant ces contrôleurs dans le fichier `/etc/vx/cntrls.exclude`.

Pour placer plusieurs disques, un par un, sous le contrôle de Volume Manager, procédez comme suit :

1. Sélectionnez l'option de menu 1 (Ajouter ou initialiser un ou plusieurs disques) dans le menu principal de vxdiskadm.
2. À l'invite suivante, entrez le paramètre *pattern-list* approprié pour les disques à mettre sous le contrôle de Volume Manager. Dans ce cas, entrez **c3** afin d'indiquer tous les disques sur le contrôleur 3 :

```
Ajouter ou initialiser les disques
Menu : VolumeManager/Disk/AddDisks
```

```
Utilisez cette opération pour ajouter un ou plusieurs disques
dans un disk group. Vous pouvez ajouter les disques
sélectionnés dans un disk group existant ou dans un nouveau
disk group qui sera créé lors de l'opération. Les disques
sélectionnés peuvent également être ajoutés dans un disk group
comme disques de remplacement. Ils peuvent aussi être initialisés
sans être ajoutés à un disk group, ce qui laisse les disques
disponibles comme disques de remplacement.
```

```
Plusieurs disques ou modèles peuvent être entrés à l'invite.
Exemples de sélections de disque :
```

```
all :           tous les disques
c3 c4t2 :      tous les disques sur les contrôleurs 3 et 4,
               cible 2
c3t4d0 :      un seul disque
```

```
Sélectionnez les périphériques de disque à ajouter :
[<pattern-list>,all,list,q,?] c3
```

où *<pattern-list>* peut représenter un seul disque ou une série de disques et/ou de contrôleurs (avec des cibles facultatives). Si *<pattern-list>* comprend plusieurs éléments, vous devez les séparer par un espace.

Si vous ne connaissez pas l'adresse (nom du périphérique) du disque à ajouter, entrez **list** à l'invite pour afficher la liste complète des disques disponibles.

3. Pour continuer l'opération, entrez **y** (ou appuyez sur la touche de retour) à l'invite suivante :

```
Les disques sélectionnés sont les suivants.  Format de sortie :  
[Nom_périphérique]
```

```
c3t0d0 c3t1d0 c3t2d0 c3t3d0
```

```
Continuer l'opération ? [y,n,q,?] (default: y) y
```

4. Pour ajouter ces disques au disk group par défaut, rootdg, entrez **y** (ou appuyez sur la touche de retour) à l'invite suivante :

```
Vous pouvez ajouter ce disque à un disk group existant, à un  
nouveau disk group ou de le laisser disponible afin qu'il puisse  
être utilisé pour des opérations d'ajout ou de remplacement  
futures. Pour créer un nouveau disk group, sélectionnez le nom  
d'un disk group qui n'existe pas. Pour que le disque soit  
disponible plus tard, nommez le disk group "none".
```

```
Quel disk group [<group>,none,list,q,?] (default: rootdg) y
```

5. Pour permettre à vxdiskadm d'utiliser des noms de disque par défaut pour chaque disque, entrez **y** (ou appuyez sur la touche de retour) à l'invite suivante :

```
Utiliser les noms de disque définis par défaut pour ces  
disques ? [y,n,q,?] (default: y) y
```

6. À l'invite suivante, entrez **n** pour indiquer que ces disques ne doivent pas être utilisés comme disques de remplacement pour la reconstruction automatique :

```
Ajouter les disques comme disques de remplacement pour rootdg ?  
[y,n,q,?] (default: n) n
```

7. Lorsque vxdiskadm vous demande s'il faut exclure ce disque de la reconstruction automatique, entrez **n** (ou appuyez sur la touche de retour) :

```
Exclure le disque de la reconstruction automatique ?  
[y,n,q,?] (default: n)
```

8. Pour continuer l'opération, entrez **y** (ou appuyez sur la touche de retour) à l'invite suivante :

```
Les disques sélectionnés seront ajoutés au disk group rootdg  
avec les noms de disque définis par défaut.
```

```
c3t0d0 c3t1d0 c3t2d0 c3t3d0
```

```
Continuer l'opération ? [y,n,q,?] (default: y) y
```

Opérations de l'interface de menu vxdiskadm

9. L'invite suivante liste les disques qui ont déjà été initialisés pour que Volume Manager puisse les utiliser ; entrez **Y** pour indiquer que tous ces disques doivent être utilisés maintenant :

Les périphériques de disque suivants semblent avoir déjà été initialisés.

Les disques sont actuellement disponibles comme disques de remplacement.

Format de sortie : [Nom_périphérique]

c3t1d0 c3t2d0 c3t3d0

Use these devices? [Y,N,S(elect),q,?] (default: Y) **Y**

Cette invite vous permet d'indiquer « yes » ou « no » pour *tous* ces disques (**Y** ou **N**) ou de sélectionner le traitement voulu pour chacun de ces disques sur une base individuelle (**S**).

Si vous voulez réinitialiser tous ces disques, entrez **Y** à l'invite suivante :

Les disques suivants sélectionnés semblent avoir déjà été initialisés pour Volume Manager. Si vous êtes sûr que c'est le cas, vous n'avez pas besoin de réinitialiser ces périphériques de disque.

Format de sortie : [Nom_périphérique]

c3t1d0 c3t2d0 c3t3d0

Réinitialiser ces périphériques ? [Y,N,S(elect),q,?] (default: Y) **Y**

10. vxdiskadm vous indique que l'un des disques sur le contrôleur 3 est un candidat à l'encapsulation ; entrez **y** (ou appuyez sur la touche de retour) pour encapsuler ce disque :

Le périphérique de disque suivant a un VTOC valide, mais ne semble pas avoir été initialisé pour Volume Manager. Si le disque contient des données, elles NE devraient PAS être détruites ; vous devriez encapsuler les partitions du disque existantes comme volumes, au lieu d'ajouter le disque comme nouveau disque.

Format de sortie : [Nom_périphérique]

c3t0d0

Encapsuler ce périphérique ? [y,n,q,?] (default: y) **y**

L'encapsulation vous permet de mettre un disque actif sous le contrôle de Volume Manager et conserve les données situées sur ce disque.

Remarque L'encapsulation de disque nécessite que vous redémarriez le système et peut requérir plusieurs autres redémarrages. Le cas échéant, le système vous invitera à effectuer ces opérations.

vxdiskadm confirme à présent les disques en cours d'initialisation et d'ajout au contrôle de Volume Manager, en affichant des messages du type suivant :

```
Initialisation du périphérique c3t1d0.
```

```
Initialisation du périphérique c3t2d0.
```

```
Initialisation du périphérique c3t3d0.
```

```
Ajout du périphérique de disque c3t1d0 au disk group rootdg  
avec le nom de disque disk33.
```

```
Ajout du périphérique de disque c3t2d0 au disk group rootdg  
avec le nom de disque disk34.
```

```
Ajout du périphérique de disque c3t3d0 au disk group rootdg  
avec le nom de disque disk35.
```

Outre l'affichage ci-dessus, des invites vous permettant d'exécuter l'analyse de surface peuvent s'afficher.

11. vxdiskadm confirme alors les disques sélectionnés pour l'encapsulation et vous demande l'autorisation de continuer l'encapsulation ; entrez **y** (ou appuyez sur la touche de retour) pour continuer l'encapsulation :

```
Le disque suivant a été sélectionné pour l'encapsulation.  
Format de sortie : [Nom_périphérique]
```

```
c3t0d0
```

```
Continuer l'encapsulation ? [y,n,q,?] (default: y) y
```

vxdiskadm affiche à présent un état de l'encapsulation et vous indique que vous devez arrêter le système et le redémarrer dès que possible :

```
Le périphérique de disque c3t0d0 sera encapsulé et ajouté au  
disk group rootdg avec le nom de disque disk38.
```

```
La première étape d'encapsulation a réussi. Vous devriez  
redémarrer le système dès que possible.
```

```
L'encapsulation nécessitera deux ou trois redémarrages qui se  
produiront automatiquement après le démarrage suivant. Pour  
redémarrer, exécutez la commande :
```

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

Opérations de l'interface de menu vxdiskadm

Elle permet de mettre à jour le fichier `/etc/vfstab` pour\n pour pouvoir utiliser les périphériques de volume afin de monter les systèmes de fichiers sur ce périphérique de disque. Vous devrez mettre à jour toutes les autres références, telles que les scripts de sauvegarde, les bases de données ou les périphériques de swap créés manuellement.

12. À l'invite suivante, indiquez si vous voulez initialiser d'autres disques (**y**) ou revenir au menu principal de vxdiskadm (**n**) :

Ajouter ou initialiser d'autres disques ? [y,n,q,?] (default: n)

Encapsulation d'un disque pour être utilisé par Volume Manager

Cette section décrit la procédure à suivre pour encapsuler un disque afin qu'il soit utilisé par Volume Manager. L'encapsulation conserve les données existantes enregistrées sur le disque lorsque ce dernier est placé sous le contrôle de Volume Manager.

Pour réduire les risques d'échec de l'encapsulation, assurez-vous que le disque :

- ◆ possède une petite quantité d'espace libre (au début ou à la fin) qui n'appartient à aucune partition ;
- ◆ possède deux partitions libres ;
- ◆ possède une partition `s2` (ou, sur certains systèmes, `s0`) représentant le disque entier.

Pendant l'encapsulation du disque d'initialisation (racine) sur certains systèmes, vous devez marquer la partition de swap comme `swap` afin de pouvoir par la suite y effectuer des vidages.

Avant d'encapsuler le disque d'initialisation, vous devez attribuer la valeur `true` à la variable EEPROM `use-nvramrc?` pour utiliser les alias de disque d'initialisation de Volume Manager afin d'identifier le miroir du disque d'initialisation au cas où un remplacement serait nécessaire. Si vous attribuez la valeur `false`, vous devez déterminer par vous-même les disques d'initialisation. Pour définir la valeur `true`, procédez comme suit :

```
eeprom use-nvramrc?=true
```

Pour encapsuler un disque afin qu'il soit utilisé par Volume Manager, procédez comme suit :

1. Sélectionnez l'option de menu 2 (Encapsuler un ou plusieurs disques) dans le menu principal de vxdiskadm.

Remarque Il se peut que votre système utilise un nom de périphérique différent de ceux donnés dans les exemples. Reportez-vous au *VERITAS Volume Manager Getting Started Guide* pour plus d'informations sur les noms de périphériques.

2. À l'invite suivante, entrez les noms de périphériques des disques à encapsuler :

Encapsuler un ou plusieurs disques
Menu : VolumeManager/Disk/Encapsulate

Utilisez cette opération pour convertir un ou plusieurs disques pour utiliser Volume Manager. Les disques sont alors ajoutés dans un disk group et les partitions existantes sont remplacées par des volumes. L'encapsulation des disques nécessite que le système soit redémarré pour que les modifications prennent effet.

Plusieurs disques ou modèles peuvent être entrés à l'invite. Exemples de sélections de disque :

```
all :          tous les disques
c3 c4t2 :     tous les disques sur les contrôleurs 3 et 4,
              cible 2
c3t4d0 :     un seul disque
```

Sélectionnez les périphériques de disque à encapsuler :
[<pattern-list>,all,list,q,?] c2t5d0

où <pattern-list> peut représenter un seul disque ou une série de disques et/ou de contrôleurs (avec des cibles facultatives). Si <pattern-list> comprend plusieurs éléments, vous devez les séparer par un espace.

Si vous ne connaissez pas l'adresse (nom du périphérique) du disque à encapsuler, entrez **l** ou **list** à l'invite pour afficher la liste complète des disques disponibles.

3. Pour continuer l'opération, entrez **y (ou appuyez sur la touche de retour) à l'invite suivante :**

Le disque sélectionné est le suivant. Format de sortie :
[Nom_périphérique]

```
c2t5d0
```

Continuer l'opération ? [y,n,q,?] (default: y) **y**

4. Pour ajouter le disque au disk group par défaut, rootdg, appuyez sur la touche de retour à l'invite suivante :

Vous pouvez ajouter ce disque à un disk group existant ou à un nouveau disk group. Pour créer un nouveau disk group, sélectionnez le nom d'un disk group qui n'existe pas.

Quel disk group [<group>,list,q,?] (default: rootdg)

Opérations de l'interface de menu vxdiskadm

5. À l'invite suivante, appuyez sur la touche de retour pour accepter le nom du disque par défaut ou entrez le nom voulu :

```
Utiliser un nom de disque par défaut pour le disque ? [y,n,q,?]
(default: y)
```

6. Pour continuer l'opération, entrez **y** (ou appuyez sur la touche de retour) à l'invite suivante :

Les disques sélectionnés seront encapsulés et ajoutés au disk group rootdg avec les noms de disque définis par défaut.

```
c2t5d0
```

```
Continuer l'opération ? [y,n,q,?] (default: y) y
```

7. Pour confirmer la poursuite de l'encapsulation, entrez **y** (ou appuyez sur la touche de retour) à l'invite suivante :

Le disque suivant a été sélectionné pour l'encapsulation.

Format de sortie : [Nom_périphérique]

```
c2t5d0
```

```
Continuer l'encapsulation ? [y,n,q,?] (default: y) y
```

Un message du type suivant confirme que le disque est en cours d'encapsulation pour être utilisé par Volume Manager et il vous indique qu'un redémarrage est nécessaire :

Le périphérique de disque c2t5d0 sera encapsulé et ajouté au disk group rootdg avec le nom de disque disk01.

Le disque c2t5d0 a été configuré pour l'encapsulation.

La première étape d'encapsulation a réussi. Vous devriez redémarrer le système dès que possible.

L'encapsulation nécessitera deux ou trois redémarrages qui se produiront automatiquement après le démarrage suivant. Pour redémarrer, exécutez la commande :

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

Elle permet de mettre à jour le fichier /etc/vfstab pour pouvoir utiliser les périphériques de volume afin de monter les systèmes de fichiers sur ce périphérique de disque. Vous devez mettre à jour toutes les autres références, telles que les scripts de sauvegarde, les bases de données ou les périphériques de swap créés manuellement.

8. À l'invite suivante, indiquez si vous voulez encapsuler d'autres disques (**y**) ou revenir au menu principal de vxdiskadm (**n**) :

```
Encapsuler d'autres disques ? [y,n,q,?] (default: n) n
```

Dans certains cas, l'encapsulation d'un disque peut échouer. Cela se produit souvent lorsque l'espace disponible sur le disque est insuffisant pour Volume Manager. Dans ce cas, la procédure ci-dessus s'interrompt brutalement et un message d'erreur du type suivant s'affiche :

```
Le périphérique de disque c2t5d0 sera encapsulé et ajouté au
disk group rootdg avec le nom de disque disk01.
```

```
L'encapsulation a échoué avec l'erreur suivante :
```

```
Il est impossible d'encapsuler c2t5d0, pour la raison suivante :
<vxvm:vx slicer: ERREUR : Structure de disque non prise en
charge.>
```

```
Hit RETURN to continue.
```

Ajout d'un disque pour une utilisation future

Pour mettre un disque sous le contrôle de Volume Manager en vue d'une utilisation future, suivez la procédure décrite dans « Placement des disques physiques sous le contrôle de Volume Manager » à la page 142. Cependant, lorsque le programme vous demande d'entrer le nom du disk group, entrez **none** au lieu de sélectionner rootdg ou d'entrer le nom d'un disk group. Le disque est alors initialisé comme précédemment, mais il est réservé pour une utilisation ultérieure. Vous devez l'ajouter à un disk group pour pouvoir l'utiliser.

Remarque Vous ne devez pas confondre ce type de « disque de remplacement » avec un disque de remplacement pour la reconstruction automatique.

Réinitialisation d'un disque pour être utilisé par Volume Manager

Cette section décrit la procédure à suivre pour réinitialiser un disque précédemment initialisé pour être utilisé par Volume Manager.

Si le disque que vous voulez ajouter a déjà été utilisé, mais pas avec Volume Manager, procédez de l'une des façons suivantes :

- ◆ Encapsulez le disque et conservez ses informations (voir « Encapsulation d'un disque pour être utilisé par Volume Manager » à la page 152).
- ◆ Réinitialisez le disque pour permettre à Volume Manager de configurer le disque. La réinitialisation ne conserve pas les données situées sur le disque. Si vous voulez réinitialiser le disque, assurez-vous qu'il ne contient pas les données à conserver.

Opérations de l'interface de menu vxdiskadm

Pour réinitialiser un disque afin qu'il soit utilisé par Volume Manager, procédez comme suit :

1. Sélectionnez l'option de menu 1 (Ajouter ou initialiser un ou plusieurs disques) dans le menu principal de vxdiskadm.
2. À l'invite suivante, entrez le nom du périphérique de disque spécifique au disque à mettre sous le contrôle de Volume Manager :

```
Ajouter ou initialiser les disques
Menu : VolumeManager/Disk/AddDisks
```

Utilisez cette opération pour ajouter un ou plusieurs disques dans un disk group. Vous pouvez ajouter les disques sélectionnés dans un disk group existant ou dans un nouveau disk group qui sera créé lors de l'opération. Les disques sélectionnés peuvent également être ajoutés dans un disk group comme disques de remplacement. Ils peuvent aussi être initialisés sans être ajoutés à un disk group, ce qui laisse les disques disponibles comme disques de remplacement.

Plusieurs disques ou modèles peuvent être entrés à l'invite. Exemples de sélections de disque :

```
all :          tous les disques
c3 c4t2 :     tous les disques sur les contrôleurs 3 et 4,
              cible 2
c3t4d0 :     un seul disque
```

Sélectionnez les périphériques de disque à ajouter :
[<pattern-list>,all,list,q,?] **c1t3d0**

où <pattern-list> peut représenter un seul disque ou une série de disques et/ou de contrôleurs (avec des cibles facultatives). Si <pattern-list> comprend plusieurs éléments, vous devez les séparer par un espace.

Si vous ne connaissez pas l'adresse (nom du périphérique) du disque à ajouter, entrez **1** ou **list** à l'invite pour afficher la liste complète des disques disponibles.

3. Pour continuer l'opération, entrez **y** (ou appuyez sur la touche de retour) à l'invite suivante :

```
Le disque sélectionné est le suivant. Format de sortie :
[Nom_périphérique]
```

```
c1t3d0
```

```
Continuer l'opération ? [y,n,q,?] (default: y) y
```

4. À l'invite suivante, spécifiez le disk group auquel ajouter le disque ou appuyez sur la touche de retour pour accepter rootdg:

Vous pouvez ajouter ce disque à un disk group existant, à un nouveau disk group ou le laisser disponible pour qu'il soit utilisé pour des opérations d'ajout ou de remplacement futures. Pour créer un nouveau disk group, sélectionnez le nom d'un disk group qui n'existe pas. Pour que le disque soit disponible plus tard, nommez le disk group "none".

```
Quel disk group [<group>,none,list,q,?] (default: rootdg)
```

5. À l'invite suivante, appuyez sur la touche de retour pour accepter le nom du disque par défaut ou entrez le nom voulu :

```
Utiliser un nom de disque par défaut pour le disque ? [y,n,q,?]  
(default: y)
```

6. Lorsque vxdiskadm vous demande si ce disque doit devenir un disque de remplacement pour la reconstruction automatique, entrez **n** (ou appuyez sur la touche de retour) :

```
Ajouter un disque comme disque de remplacement pour rootdg ?  
[y,n,q,?] (default: n) n
```

7. Lorsque vxdiskadm vous demande s'il faut exclure ce disque de la reconstruction automatique, entrez **n** (ou appuyez sur la touche de retour) :

```
Exclure le disque de la reconstruction automatique ?  
[y,n,q,?] (default: n)
```

8. Pour continuer l'opération, entrez **y** (ou appuyez sur la touche de retour) à l'invite suivante :

Les disques sélectionnés seront ajoutés au disk group rootdg avec les noms de disque définis par défaut.

```
c1t3d0
```

```
Continuer l'opération ? [y,n,q,?] (default: y) y
```

Opérations de l'interface de menu vxdiskadm

- 9.** L'invite suivante indique que ce disque a déjà été initialisé pour être utilisé ultérieurement par Volume Manager ; entrez **y** pour confirmer que vous voulez utiliser ce disque maintenant :

Le périphérique de disque suivant semble avoir déjà été initialisé. Le disque est actuellement disponible comme disque de remplacement.

Format de sortie : [Nom_périphérique]

c1t3d0

Utiliser ce périphérique ? [y,n,q,?] (default: y) **y**

- 10.** Pour réinitialiser le disque, entrez **y** (ou appuyez sur la touche de retour) à l'invite suivante :

Le disque suivant, que vous avez sélectionné, semble avoir déjà été initialisé pour Volume Manager. Si vous êtes sûr que c'est le cas, vous n'avez pas à réinitialiser le périphérique de disque.

Format de sortie : [Nom_périphérique]

c1t3d0

Réinitialiser ce périphérique ? [y,n,q,?] (default: y) **y**

Des messages du type suivant confirment que ce disque est en cours de réinitialisation pour être utilisé par Volume Manager :

Initialisation du périphérique c1t3d0.

Ajout du périphérique de disque c1t3d0 au disk group rootdg avec le nom de disque disk40.

- 11.** À l'invite suivante, indiquez si vous voulez initialiser d'autres disques (**y**) ou revenir au menu principal de vxdiskadm (**n**) :

Ajouter ou initialiser d'autres disques ? [y,n,q,?] (default: n)

Ajout d'un disque VM au pool de reconstruction automatique

La reconstruction automatique permet au système de répondre automatiquement aux erreurs d'E/S en reconstruisant les sous-disques redondant sur d'autres disques. La reconstruction automatique restaure ensuite les objets et les données Volume Manager affectés. Si un disque a déjà été spécifié comme disque de remplacement dans le disk group, les sous-disques du disque qui a échoué sont reconstruits sur le disque de remplacement. Dans le cas contraire, le programme utilise n'importe quel espace disponible approprié dans le disk group, à l'exception de l'espace disponible sur les disques qui ont été précédemment exclus de la reconstruction automatique.

Pour désigner un disque comme disque de remplacement, procédez comme suit :

1. Sélectionnez l'option de menu 12 (Marquer un disque comme disque de remplacement pour un disk group) dans le menu principal de vxdiskadm.

2. À l'invite suivante, entrez le nom d'un disque (par exemple, disk01) :

```
Marquer un disque comme disque de remplacement pour un disk
group
```

```
Menu : VolumeManager/Disk/MarkSpareDisk
```

Utilisez cette opération pour marquer un disque comme étant de remplacement pour un disk group. Vous devez entrer le nom d'un disque. Il s'agit du nom que vous avez donné au disque lorsque vous l'avez ajouté au disk group.

```
Entrez le nom du disque [<disk>,list,q,?] disk01
```

vxdiskadm affiche la confirmation suivante :

```
disk01 a été marqué dans rootdg comme disque de remplacement.
```

3. À l'invite suivante, indiquez si vous voulez ajouter d'autres disques comme disques de remplacement (**y**) ou revenir au menu principal de vxdiskadm (**n**) :

```
Marquer un autre disque comme disque de remplacement ?
[y,n,q,?] (default: n)
```

Tous les disques VM de ce disk group utiliseront à présent ce disque comme disque de remplacement en cas de défaillance. En cas de défaillance d'un disque, la reconstruction automatique devrait s'exécuter automatiquement (si possible). Vous devriez être informé de la défaillance et de la reconstruction par un courrier électronique. À l'issue de la reconstruction, vous pourrez remplacer le disque qui a échoué.

La reconstruction automatique utilise l'espace disponible si l'espace du disque de remplacement est insuffisant pour reconstruire les sous-disques qui ont échoué. L'utilisateur peut limiter l'utilisation de l'espace disponible par la reconstruction automatique en spécifiant les disques libres qu'il doit ignorer.

Suppression d'un disque VM du pool de reconstruction automatique

Lorsqu'un disque est spécifié comme disque de remplacement, son espace ne sert pas d'espace libre pour la création d'objets Volume Manager dans ce disk group. Si nécessaire, vous pouvez libérer un disque de remplacement pour une utilisation générale en le supprimant du pool des disques de reconstruction automatique.

Pour vérifier les disques qui sont actuellement spécifiés comme disques de remplacement, sélectionnez l'option de menu `list` dans le menu principal de `vxdiskadm`. L'indicateur `remplacement` est attribué aux disques de remplacement.

Pour supprimer un disque du pool de reconstruction automatique, procédez comme suit :

1. Sélectionnez l'option de menu `13` (Désactiver l'indicateur de remplacement pour un disque) dans le menu principal de `vxdiskadm`.

2. À l'invite suivante, entrez le nom du disque de remplacement (par exemple, `disk01`):

```
Désactiver l'indicateur de remplacement pour un disque
Menu : VolumeManager/Disk/UnmarkSpareDisk
```

Utilisez cette opération pour désactiver l'indicateur de remplacement sur un disque. Vous devez entrer le nom d'un disque. Il s'agit du nom que vous avez donné au disque lorsque vous l'avez ajouté au disk group.

```
Entrez le nom du disque [<disk>,list,q,?] disk01
```

`vxdiskadm` affiche la confirmation suivante :

```
Le disque disk01 dans rootdg n'est plus marqué comme disque de
remplacement.
```

3. À l'invite suivante, indiquez si vous voulez désactiver d'autres disques de remplacement (**y**) ou revenir au menu principal de `vxdiskadm` (**n**) :

```
Désactiver l'indicateur de remplacement sur un autre disque ?
[y,n,q,?] (default: n)
```

Nouveau déplacement des sous-disques reconstruits automatiquement vers un disque

Cette option permet de déplacer de nouveau les sous-disques reconstruits automatiquement vers le disque sur lequel ils se trouvaient à l'origine une fois que celui-ci a été remplacé à la suite d'une défaillance.

1. Sélectionnez l'option de menu 14 (Déplacer de nouveau les sous-disques reconstruits automatiquement vers un disque) dans le menu principal de vxdiskadm.

2. Cette option vous demande d'abord de fournir le nom d'accès logique au disque d'origine.

Entrez le nom d'accès logique au disque qui contenait initialement les sous-disques reconstruits automatiquement à l'invite suivante :

```
Enter the original disk name [<disk>,list,q,?]
```

Cependant, si le système ne comprend aucun sous-disque reconstruit automatiquement, vxdiskadm affiche le message `Aucun disque n'est actuellement reconstruit automatiquement` et vous invite à appuyer sur la touche de retour pour continuer.

3. Vous êtes ensuite invité à indiquer le nouveau disque de destination.

```
While unrelocating the subdisks, you can choose to move the subdisks to a different disk from the original disk.
```

```
Reconstruire dans un nouveau disque [y,n,q,?] (default: n)
```

4. Si vous ne pouvez pas déplacer de nouveau les sous-disques vers les décalages d'origine, vous pouvez également choisir l'option `force` pour reconstruire les sous-disques dans le disque spécifié, mais pas obligatoirement vers les décalages d'origine exacts.

```
Utilisez l'option -f pour reconstruire les sous-disques dans le disque d'origine si le déplacement vers le décalage exact échoue ? [y,n,q,?] (default: n)
```

5. Cette option confirme l'opération demandée et vous invite à confirmer la poursuite de l'opération.

```
L'opération demandée consiste à déplacer de nouveau tous les sous-disques qui ont été reconstruits automatiquement de disk10 vers disk10 du disk group rootdg.
```

```
Continuer l'opération ? [y,n,q,?] (default: y)
```

6. À la fin, l'utilisateur est informé de l'état de l'opération.

```
La reconstruction dans le disque disk10 est terminée.
```

Exclusion d'un disque de la reconstruction automatique

Pour exclure un disque de la reconstruction automatique :

1. Sélectionnez l'option de menu 15 (Exclure un disque de la reconstruction automatique) dans le menu principal de vxdiskadm.

2. À l'invite suivante, entrez le nom du disque (par exemple, disk01) :

```
Exclure un disque de la reconstruction automatique
Menu : VolumeManager/Disk/MarkNoHotUse
```

Utilisez cette opération pour exclure un disque de la reconstruction automatique. Vous devez entrer le nom d'un disque. Il s'agit du nom que vous avez donné au disque lorsque vous l'avez ajouté au disk group.

```
Entrez le nom du disque [<disk>,list,q,?] disk01
```

vxdiskadm affiche la confirmation suivante :

```
L'exclusion de disk01 dans rootdg de la reconstruction
automatique est terminée.
```

3. À l'invite suivante, indiquez si vous voulez ajouter d'autres disques à exclure de la reconstruction automatique (y) ou appuyez sur la touche de retour pour revenir au menu principal de vxdiskadm (n) :

```
Exclure un autre disque de la reconstruction automatique ?
[y,n,q,?] (default: n)
```

Libération d'un disque pour la reconstruction automatique

Si un disque a été précédemment exclu de la reconstruction automatique, la procédure suivante annulera cette exclusion et rajoutera le disque dans le pool de reconstruction automatique :

Pour libérer un disque pour la reconstruction automatique :

1. Sélectionnez l'option de menu 16 (Libérer un disque pour la reconstruction automatique) dans le menu principal de vxdiskadm.

```
Libérer un disque pour la reconstruction automatique
```

2. À l'invite suivante, entrez le nom du disque (par exemple, `disk01`) :

```
Libérer un disque pour la reconstruction automatique
Menu : VolumeManager/Disk/UnmarkNoHotUse
```

Utilisez cette opération pour que le disque puisse être utilisé par la reconstruction automatique. Cela ne s'applique qu'aux disques qui en étaient auparavant exclus. Vous devez entrer le nom d'un disque. Il s'agit du nom que vous avez donné au disque lorsque vous l'avez ajouté au disk group.

```
Entrez le nom du disque [<disk>,list,q,?] disk01
```

vxdiskadm affiche la confirmation suivante :

```
disk01 dans rootdg est maintenant disponible pour la
reconstruction automatique.
```

3. À l'invite suivante, indiquez si vous voulez rendre d'autres disques disponibles pour la reconstruction automatique (**y**) ou revenir au menu principal de vxdiskadm (**n**) :

```
Libérer un autre disque pour la reconstruction automatique ?
[y,n,q,?] (default: n)
```

Affichage d'informations sur les disques physiques

Des informations sur les disques telles que les disques initialisés, les disk groups auxquels ils appartiennent et leur état sont affichées. La commande `list` affiche les noms de périphériques pour tous les disques reconnus, les noms des disques, les noms des disk groups associés à chaque disque et l'état de chaque disque.

Pour afficher les informations sur les disques, procédez comme suit :

1. Sélectionnez `list` dans le menu principal de vxdiskadm.
2. À l'affichage suivant, entrez l'adresse du disque que vous voulez afficher ou entrez `a11` pour lister tous les disques :

```
Lister les informations sur le disque
Menu : VolumeManager/Disk/ListDisk
```

Utilisez cette opération du menu pour afficher la liste des disques. Vous pouvez également lister des informations détaillées sur le disque à une adresse de périphérique de disque spécifique.

```
Entrez un périphérique de disque ou "all" [<address>,all,q,?]
(default: all)
```

Si vous entrez `a11`, Volume Manager affiche les informations suivantes (certains systèmes peuvent afficher `rootdisk` à la place de `c0t0d0s2`) :

Opérations de l'interface de menu vxdiskadm

PÉRIPHÉRIQUE	DISQUE	GROUPE	ÉTAT
c0t0d0	c0t0d0s2	rootdg	en ligne
c1t0d0	disk01	rootdg	en ligne
c1t1d0	-	-	en ligne

Périphérique à lister en détails
 [<address>,none,q,?] (default: none)

- ◆ Si vous entrez l'adresse du périphérique sur lequel vous voulez des informations, le programme affiche des informations complètes sur le disque (y compris le nom de périphérique, le type de disque et des informations sur les partitions publiques et privées du disque).

Après avoir examiné ces informations, appuyez sur la touche de retour pour revenir au menu principal.

Suppression d'un disque physique

Avant de supprimer un disque, assurez-vous qu'il ne contient aucune donnée, que les données ne sont plus utiles ou que vous pouvez les déplacer vers d'autres disques. Supprimez ensuite le disque en procédant comme suit :

1. Sélectionnez l'option de menu 3 (Supprimer un disque) dans le menu principal de vxdiskadm.

Remarque Vous devez désactiver le disk group pour pouvoir supprimer le dernier disque de ce groupe.

2. À l'invite suivante, entrez le nom du disque à supprimer :

```
Supprimer un disque
Menu : VolumeManager/Disk/RemoveDisk
```

Utilisez cette opération pour supprimer un disque d'un disk group. Vous devez entrer le nom d'un disque. Il s'agit du nom que vous avez donné au disque lorsque vous l'avez ajouté au disk group.

Entrez le nom du disque [<disk>,list,q,?] **disk01**

3. Si le disque contient des volumes, Volume Manager vous demande s'il faut les déplacer du disque. Pour conserver les volumes, répondez **y**. Dans le cas contraire, répondez **n**.

4. À l'invite de vérification suivante, appuyez sur la touche de retour pour continuer :

L'opération demandée consiste à supprimer le disque disk01 du groupe rootdg.

Continuer l'opération ? [y,n,q,?] (default: y)

L'utilitaire vxdiskadm supprime le disque du disk group et affiche le message de réussite suivante :

La suppression du disque disk01 est terminée.

Vous pouvez à présent supprimer le disque ou le conserver sur le système en tant que disque de remplacement.

5. À l'invite suivante, indiquez si vous voulez supprimer d'autres disques (y) ou revenir au menu principal de vxdiskadm (n) :

Supprimer un autre disque ? [y,n,q,?] (default: n)

Désactivation d'un disque physique (mise hors ligne)

Si un disque est endommagé, vous devez le mettre hors ligne et le supprimer. Vous pouvez déplacer le périphérique de disque physique vers un autre emplacement qui sera connecté à un autre système. Pour mettre un disque hors ligne, supprimez-le d'abord de son disk group, puis procédez comme suit :

1. Sélectionnez l'option de menu 11 (Désactiver un périphérique de disque (mettre hors ligne)) dans le menu principal de vxdiskadm.**2. À l'invite suivante, entrez l'adresse du disque à désactiver :**

Désactiver un périphérique de disque (mettre hors ligne)
Menu : VolumeManager/Disk/OfflineDisk

Utilisez cette opération du menu pour désactiver l'accès à un périphérique de disque par Volume Manager. Cette opération ne peut être exécutée que sur les disques qui ne se trouvent actuellement pas dans un disk group. Utilisez cette opération si vous voulez supprimer le disque d'un système sans redémarrer celui-ci.

REMARQUE : de nombreux systèmes ne prennent pas en charge les disques qui peuvent être supprimés lors d'une opération normale. Sur ces systèmes, l'opération de mise hors ligne est rarement utile.

Opérations de l'interface de menu vxdiskadm

Sélectionnez un périphérique de disque à désactiver
[<address>,list,q,?] **c1t1d0**

Le programme vxdiskadm désactive le disque spécifié.

3. À l'invite suivante, indiquez si vous voulez désactiver un autre périphérique (**y**) ou revenir au menu principal de vxdiskadm (**n**) :

Désactiver un autre périphérique ? [y,n,q,?] (default: n)

Activation d'un disque physique

Si vous déplacez un disque d'un système vers un autre pendant le fonctionnement normal du système, Volume Manager ne reconnaît pas automatiquement le disque. Utilisez l'opération d'activation de disque pour indiquer à Volume Manager d'analyser le disque afin de l'identifier et de déterminer s'il appartient à un disk group. De plus, cette opération vous permet également de réactiver l'accès à un disque qui a été désactivé soit via l'exportation du disk group, soit via la désactivation (mise hors ligne) du périphérique de disque. Pour activer un disque, procédez comme suit :

1. Sélectionnez l'option de menu 10 (Activer un périphérique de disque (mettre en ligne)) dans le menu principal de vxdiskadm.
2. À l'invite suivante, entrez le nom de périphérique du disque à activer (ou entrez **list** pour afficher la liste des périphériques) :

Activer un périphérique de disque (mettre en ligne)
Menu : VolumeManager/Disk/OnlineDisk

Utilisez cette opération pour activer l'accès à un disque qui était désactivé à l'aide de l'opération "Désactiver un périphérique de disque (mettre hors ligne)".

Vous pouvez également utiliser cette opération pour analyser de nouveau un disque qui a peut-être été modifié en dehors de Volume Manager. Par exemple, si un disque est partagé par deux systèmes, Volume Manager exécuté sur l'autre système a peut-être modifié le disque. Dans ce cas, vous pouvez utiliser cette opération pour analyser de nouveau le disque.

REMARQUE : de nombreuses opérations vxdiskadm effectuent une nouvelle analyse des disques sans que l'utilisateur n'ait à intervenir. Cela élimine la plupart des opérations de mise en ligne directe d'un disque, sauf lorsque celui-ci est mis hors ligne directement.

Sélectionnez un périphérique de disque à activer :
[<address>,list,q,?] **c1t1d0**

vxdiskadm active le périphérique spécifié.

3. À l'invite suivante, indiquez si vous voulez activer un autre périphérique (**y**) ou revenir au menu principal de vxdiskadm (**n**) :

Activer un autre périphérique ? [y,n,q,?] (default: n)

Remarque Vous devez entrer `vxdctl enable` après un swap de disque automatique pour permettre à VxVM de reconnaître le disque et les chemins.

Remplacement d'un disque physique

Si un disque échoue, vous devez le remplacer par un autre. Pour cela, il est nécessaire de désactiver et de supprimer le disque qui a échoué, et de le remplacer par un nouveau disque. Pour remplacer un disque, procédez comme suit :

1. Sélectionnez l'option de menu 4 (Supprimer un disque pour le remplacer) dans le menu principal de vxdiskadm.
2. À l'invite suivante, entrez le nom du disque à remplacer (ou entrez **list** pour afficher la liste des disques) :

Supprimer un disque pour le remplacer
Menu : VolumeManager/Disk/RemoveForReplace

Utilisez cette opération du menu pour supprimer un disque physique d'un disk group, tout en conservant le nom du disque. Cela change l'état du nom du disque et le remplace par disque "supprimé". Si des disques initialisés ne font pas partie d'un disk group, vous pouvez utiliser l'un de ces disques comme disque de remplacement.

Entrez le nom du disque [<disk>,list,q,?] **disk02**

Des affichages supplémentaires indiquent les volumes associés au disque que vous voulez supprimer. Vous devez décider si vous voulez conserver les données associées aux volumes ou les perdre lors du remplacement du disque. Répondez en conséquence aux invites éventuelles.

3. À l'invite suivante, sélectionnez le nom de périphérique du disque de remplacement (dans la liste affichée) ou appuyez sur la touche de retour pour choisir le disque par défaut :

Les périphériques suivants sont disponibles comme remplacement :
c1t1d0

Vous pouvez choisir l'un de ces disques maintenant pour

Opérations de l'interface de menu vxdiskadm

remplacer disk02. Sélectionnez "none" si vous ne désirez pas sélectionner de disque de remplacement.

Choisissez un périphérique ou sélectionnez
[<device>,none,q,?] (default: c1t1d0)

4. À l'invite suivante, appuyez sur la touche de retour pour continuer :

L'opération demandée consiste à supprimer le disque disk02 du groupe rootdg.

Le disque supprimé sera remplacé par le périphérique de disque c1t1d0.

Continuer l'opération ? [y,n,q,?] (default: y)

vxdiskadm affiche le message de réussite suivant :

La suppression du disque disk02 a réussi.

Remplacement de disk02 par le périphérique c1t1d0.

Remplacement de disque terminé sans problème.

5. À l'invite suivante, indiquez si vous voulez supprimer un autre disque (**y**) ou revenir au menu principal de vxdiskadm (**n**) :

Supprimer un autre disque ? [y,n,q,?] (default: n)

Remplacement d'un disque qui a échoué ou qui a été supprimé

Vous pouvez supprimer les disques et les remplacer par la suite. Pour cela, utilisez l'option de menu 4 (Supprimer un disque pour le remplacer), puis procédez comme suit :

1. Sélectionnez l'option de menu 5 (Remplacer un disque qui a échoué ou a été supprimé) dans le menu principal de vxdiskadm.

2. Sélectionnez le nom du disque à remplacer :

Remplacer un disque qui a échoué ou a été supprimé
Menu : VolumeManager/Disk/ReplaceDisk

Utilisez cette opération du menu pour spécifier un disque de remplacement pour un disque que vous avez supprimé avec l'opération du menu "Supprimer un disque pour le remplacer" ou qui a échoué lors de son utilisation. Vous devrez entrer le nom du disque à remplacer et un périphérique de disque à utiliser comme remplacement.

Vous pouvez choisir un disque non initialisé, et dans ce

cas il le sera, ou un disque déjà initialisé à l'aide de l'opération du menu "Ajouter ou initialiser un ou plusieurs disques".

```
Select a removed or failed disk [<disk>,list,q,?] disk02
```

- vxdiskadm affiche les noms des périphériques de disque disponibles pour le remplacement ; entrez le nom du périphérique voulu ou appuyez sur la touche de retour pour sélectionner le périphérique par défaut :

Remarque Il se peut que votre système utilise un nom de périphérique différent de ceux donnés dans les exemples. Pour plus d'informations sur les noms de périphériques, voir « Présentation de Volume Manager » à la page 17.

Les périphériques suivants sont disponibles comme remplacement :
c1t0d0s2 c1t1d0s2

Vous pouvez choisir l'un de ces disques pour remplacer disk02.

Choisissez "none" pour initialiser un autre disque et remplacer disk02.

Choisissez un périphérique ou sélectionnez "none"
[<device>,none,q,?] (default: c1t0d0s2)

- À l'invite suivante, appuyez sur la touche de retour pour remplacer le disque :

L'opération demandée consiste à utiliser le périphérique initialisé c1t0d0s2 pour remplacer le disque supprimé ou échoué disk02 dans le disk group rootdg.

```
Continuer l'opération ? [y,n,q,?] (default: y)
```

Le programme vxdiskadm affiche le message de réussite suivant :

Le remplacement du disque disk02 dans le groupe rootdg par le périphérique de disque c1t0d0s2 a réussi.

- À l'invite suivante, indiquez si vous voulez remplacer un autre disque (**y**) ou revenir au menu principal de vxdiskadm (**n**) :

```
Remplacer un autre disque ? [y,n,q,?] (default: n)
```

Ajout d'un disque à un disk group

Il se peut que vous vouliez ajouter un nouveau disque à un disk group déjà établi. Par exemple, les disques actuels ne possèdent pas un espace suffisant pour les spécifications du projet ou du groupe de travail, surtout si ces spécifications ont changé. Pour ajouter un disque à un disk group, suivez la procédure permettant d'ajouter un disque. Voir « Placement des disques physiques sous le contrôle de Volume Manager » à la page 142.

Création d'un disk group

Il se peut que vous vouliez que toutes les données relatives à un ensemble spécifique d'applications d'un groupe d'utilisateurs particulier soient accessibles sur un autre système. Par exemples :

- ◆ Un système est en panne et vous devez déplacer ses données sur d'autres systèmes.
- ◆ La charge de travail doit être équilibrée entre les divers systèmes.

Dans ces circonstances, il importe que les données relatives à des applications ou à des utilisateurs spécifiques soient placées sur un jeu identifiable de périphériques de disques. Seules les données appartenant aux applications ou aux utilisateurs situés sur ces disques sont déplacées avec ces derniers, et pas les autres données.

Remarque Volume Manager prend en charge un disk group par défaut, *rootdg*, dans lequel tous les volumes sont créés en l'absence de toute autre spécification. De même, toutes les commandes sont dirigées par défaut vers *rootdg*.

Un disk group doit être créé avec un disque et il doit être associé à au moins un disque.

Si vous devez créer un disk group en plus de *rootdg*, procédez comme suit :

1. Sélectionnez l'option de menu 1 (Ajouter ou initialiser un ou plusieurs disques) dans le menu principal de vxdiskadm.
2. À l'invite suivante, entrez le nom du périphérique spécifique au disque que vous voulez placer sous le contrôle de Volume Manager :

```
Ajouter ou initialiser les disques
Menu : VolumeManager/Disk/AddDisks
```

Utilisez cette opération pour ajouter un ou plusieurs disques dans un disk group. Vous pouvez ajouter les disques sélectionnés dans un disk group existant ou dans un nouveau disk group qui sera créé lors de l'opération. Les disques sélectionnés peuvent également être ajoutés dans un disk group comme disques de remplacement. Ils peuvent aussi être initialisés sans être ajoutés à un disk group, ce qui laisse les disques disponibles comme disques de remplacement.

Plusieurs disques ou modèles peuvent être entrés à l'invite.
Exemples de sélections de disque :

```
all :           tous les disques
c3 c4t2 :      tous les disques sur les contrôleurs 3 et 4,
                cible 2
c3t4d0 :       un seul disque
```

Sélectionnez les périphériques de disque à ajouter :
[<pattern-list>,all,list,q,?] **c1t2d0**

où <pattern-list> peut représenter un seul disque ou une série de disques et/ou de contrôleurs (avec des cibles facultatives). Si <pattern-list> comprend plusieurs éléments, vous devez les séparer par un espace.

Si vous ne connaissez pas l'adresse (nom du périphérique) du disque à ajouter, entrez **1** ou **list** à l'invite pour afficher la liste complète des disques disponibles.

3. Pour continuer l'opération, entrez **y** (ou appuyez sur la touche de retour) à l'invite suivante :

Le disque sélectionné est le suivant. Format de sortie :
[Nom_périphérique]

```
c1t2d0
```

Continuer l'opération ? [y,n,q,?] (default: y) **y**

4. À l'invite suivante, spécifiez le disk group auquel ajouter le disque (anotherdg, dans ce cas) :

Vous pouvez ajouter ce disque à un disk group existant, à un nouveau disk group ou le laisser disponible pour qu'il soit utilisé pour des opérations d'ajout ou de remplacement futures. Pour créer un nouveau disk group, sélectionnez le nom d'un disk group qui n'existe pas. Pour que le disque soit disponible plus tard, nommez le disk group "none".

Quel disk group [<group>,none,list,q,?] (default: rootdg)
anotherdg

5. vxdiskadm confirme qu'il n'existe actuellement aucun disk group actif avec le même nom et vous demande de confirmer que vous voulez vraiment créer ce nouveau disk group :

Il n'existe pas de disk group actif nommé anotherdg.

Créer un nouveau groupe nommé anotherdg? [y,n,q,?]
(default: y) **y**

Opérations de l'interface de menu vxdiskadm

6. À l'invite suivante, appuyez sur la touche de retour pour accepter le nom du disque par défaut ou entrez le nom voulu :

```
Utiliser un nom de disque par défaut pour le disque ?
[y,n,q,?] (default: y)
```

7. Lorsque vxdiskadm vous demande s'il faut exclure ce disque de la reconstruction automatique, entrez **n** (ou appuyez sur la touche de retour) :

```
Exclure le disque de la reconstruction automatique ?
[y,n,q,?] (default: n)
```

8. Lorsque vxdiskadm vous demande si ce disque doit devenir un disque de remplacement pour la reconstruction automatique, entrez **n** (ou appuyez sur la touche de retour) :

```
Ajouter un disque comme disque de remplacement pour anotherdg ?
[y,n,q,?] (default: n) y
```

9. Pour continuer l'opération, entrez **y** (ou appuyez sur la touche de retour) à l'invite suivante :

```
Un nouveau disk group sera créé, nommé anotherdg, et les
disques sélectionnés seront ajoutés au disk group avec les
disque définis par défaut.
```

```
c1t2d0
```

```
Continuer l'opération ? [y,n,q,?] (default: y) y
```

10. L'invite suivante indique que ce disque a déjà été initialisé pour être utilisé ultérieurement par Volume Manager ; entrez **y** pour confirmer que vous voulez utiliser ce disque maintenant :

```
Le périphérique de disque suivant semble avoir déjà été
initialisé. Le disque est actuellement disponible comme disque
de remplacement.
```

```
Format de sortie : [Nom_périphérique]
```

```
c1t2d0
```

```
Utiliser ce périphérique ? [y,n,q,?] (default: y) y
```

Si vous êtes certain de vouloir réinitialiser le disque, entrez **y** (ou appuyez sur la touche de retour) à l'invite suivante :

```
Le disque suivant, que vous avez sélectionné, semble avoir
déjà été initialisé pour Volume Manager. Si vous êtes sûr
que c'est le cas, vous n'avez pas à réinitialiser le
périphérique de disque.
```

Format de sortie : [Nom_périphérique]

c1t2d0

Réinitialiser ce périphérique ? [y,n,q,?] (default: y) **y**

Des messages du type suivant confirment que ce disque est en cours de réinitialisation pour être utilisé par Volume Manager :

Initialisation du périphérique c1t2d0.

Création d'un nouveau disk group nommé anotherdg contenant le périphérique de disque c1t2d0 avec le nom another01.

11. À l'invite suivante, indiquez si vous voulez initialiser d'autres disques (**y**) ou revenir au menu principal de vxdiskadm (**n**) :

Ajouter ou initialiser d'autres disques ? [y,n,q,?] (default: n)

Exportation d'un disk group

Utilisez l'opération d'exportation d'un disk group pour désactiver l'accès à un disk group actuellement activé (importé) par le système. Vous pouvez exporter un disk group si vous voulez déplacer vers un autre système les disques de ce disk group ou si vous voulez utiliser tous les disques restants de ce disk group à d'autres fins.

Pour exporter un disk group, procédez comme suit :

1. Sélectionnez l'option de menu 9 (Désactiver l'accès à un disk group (exporter)) dans le menu principal de vxdiskadm.
2. À l'invite suivante, entrez le nom du disk group à exporter :

Désactiver l'accès à un disk group (exporter)

Menu : VolumeManager/Disk/DeportDiskGroup

Utilisez cette opération du menu pour supprimer l'accès à un disk group qui est actuellement activé (importé) par ce système. Exportez un disk group si vous voulez déplacer les disques d'un disk group vers un autre système ou exportez un disk group pour utiliser tous les disques restants à d'autres fins.

Vous devrez entrer le nom d'un disk group et spécifier si les disques doivent être désactivés (mis hors ligne). Pour les périphériques de disques amovibles de certains systèmes, il est important de désactiver l'accès au disque avant de le supprimer.

Entrez le nom du disk group [<group>,list,q,?] (default: list)

newdg

Opérations de l'interface de menu vxdiskadm

3. À l'invite suivante, entrez **y** si vous voulez supprimer les disques de ce disk group :

L'opération demandée consiste à désactiver l'accès au disk group amovible nommé newdg. Ce disk group est stocké sur les disques suivants :

newdg01 sur le périphérique c1t1d0s2

Vous pouvez désactiver l'accès (opération appelée également "mise hors ligne") à ces disques. Cela permet d'éviter des erreurs si vous supprimez des disques du système.

Désactiver les disques indiqués (mettre hors ligne) ? [y,n,q,?] (default: n)

4. À l'invite suivante, appuyez sur la touche de retour pour continuer l'opération :

Continuer l'opération ? [y,n,q,?] (default: y)

Une fois le disk group exporté, vxdiskadm affiche le message suivant :

La suppression du disk group newdg a réussi.

5. À l'invite suivante, indiquez si vous voulez désactiver un autre disk group (**y**) ou revenir au menu principal de vxdiskadm (**n**) :

Désactiver un autre disk group ? [y,n,q,?] (default: n)

Importation d'un disk group

Utilisez cette option de menu pour activer l'accès de ce système à un disk group. Si vous voulez déplacer un disk group d'un système à un autre, vous devez d'abord le désactiver (exporter) sur le système d'origine, puis déplacer le disque vers le système voulu et activer (importer) le disk group.

Pour importer un disk group, procédez comme suit :

1. Sélectionnez l'option de menu 8 (Activer l'accès à un disk group (importer)) dans le menu principal de vxdiskadm.

2. À l'invite suivante, entrez le nom du disk group à importer :

Activer l'accès à un disk group (importer)
Menu : VolumeManager/Disk/EnableDiskGroup

Utilisez cette opération pour activer l'accès à un disk group. Cette opération peut être utilisée à la fin du déplacement d'un disk group d'un système vers un autre. La première partie du déplacement du disk group consiste à utiliser l'opération

"Supprimer l'accès à un disk group (Exporter)" sur l'hôte d'origine.

Un disk group peut être importé à partir d'un autre hôte qui a échoué sans qu'il soit d'abord exporté. Vérifiez que tous les disques du disk group sont déplacés entre les hôtes.

Si deux hôtes partagent un bus SCSI, vérifiez que l'autre hôte a vraiment échoué ou a exporté le disk group. Si deux hôtes actifs un disk group en même temps, le disk group sera endommagé et deviendra inutilisable.

Sélectionnez un disk group à importer [`<group>,list,q,?`]
(default: list) **newdg**

À l'issue de l'importation, vxdiskadm affiche le message de réussite suivant :

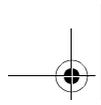
L'importation de newdg a réussi.

3. À l'invite suivante, indiquez si vous voulez importer un autre disk group (**y**) ou revenir au menu principal de vxdiskadm (**n**) :

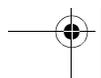
Sélectionnez un autre disk group ? [`y,n,q,?`] (default: n)

Fermeture de vxdiskadm

Une fois que toutes les activités d'administration de disque sont terminées, quittez vxdiskadm en sélectionnant **q** dans le menu principal.



Opérations de l'interface de menu vxdiskadm



Opérations de volume

5

Introduction

Ce chapitre décrit les procédures à suivre pour créer et actualiser la configuration d'un système sous le contrôle de Volume Manager. Il contient des informations sur la création, la suppression et l'actualisation d'objets Volume Manager (volumes, plex et sous-disques) ainsi que sur la sauvegarde en ligne.

Opérations de volume, de plex et de sous-disque, et sauvegarde en ligne

Ce chapitre contient des informations sur les opérations suivantes :

Opérations de volume

- ◆ Création de volumes
- ◆ Redimensionnement d'un volume
- ◆ Suppression d'un volume
- ◆ Mise en miroir d'un volume
- ◆ Suppression d'un miroir
- ◆ Ajout d'un journal RAID 5
- ◆ Ajout d'un journal des zones modifiées
- ◆ Suppression d'un journal des zones modifiées
- ◆ Suppression d'un journal RAID 5
- ◆ Arrêt d'un volume
- ◆ Démarrage d'un volume
- ◆ Modification de la stratégie de lecture du volume
- ◆ Préparation d'un volume pour la restauration à partir d'une sauvegarde
- ◆ Récupération d'un volume
- ◆ Affichage d'informations sur un volume

Introduction

Opérations de plex

- ◆ Création d'un plex
- ◆ Association d'un plex
- ◆ Dissociation et suppression d'un plex
- ◆ Sauvegarde à l'aide de la mise en miroir
- ◆ Affichage d'informations sur les plex
- ◆ Modification des attributs d'un plex
- ◆ Modification de l'état d'un plex : détachement et attachement d'un plex
- ◆ Déplacement d'un plex
- ◆ Copie d'un plex

Opérations de sous-disque

- ◆ Création d'un sous-disque
- ◆ Suppression d'un sous-disque
- ◆ Affichage d'informations sur les sous-disques
- ◆ Association d'un sous-disque
- ◆ Dissociation d'un sous-disque
- ◆ Modification des attributs d'un sous-disque
- ◆ Déplacement d'un sous-disque
- ◆ Division d'un sous-disque
- ◆ Jonction de sous-disques

Exécution d'une sauvegarde en ligne

- ◆ Resynchronisation rapide des miroirs
- ◆ Mise en miroir d'un volume sur un disque VM
- ◆ Déplacement d'un volume à partir d'un disque VM

Remarque Certaines commandes Volume Manager nécessitent des privilèges de super-utilisateur et d'autres privilèges appropriés.

Création de volumes

Les volumes sont créés afin de tirer le meilleur parti du concept de disques virtuels de Volume Manager. Une fois que le volume est créé, vous pouvez y placer un système de fichiers pour organiser les fichiers et les répertoires dans l'espace disque. De plus, vous pouvez utiliser certaines applications, comme les bases de données, pour organiser les données sur les volumes.

Pour pouvoir créer des volumes, vous devez d'abord initialiser et définir les disques et les disk groups avec Volume Manager.

Vous pouvez utiliser une méthode de base ou une méthode avancée pour créer les volumes :

- ◆ **Méthode de base** — Cette méthode utilise les informations relatives à l'action que vous voulez effectuer, puis exécute les tâches sous-jacentes nécessaires. Elle ne nécessite qu'une intervention minimale de votre part et elle rend aussi possible des spécifications plus détaillées. Les opérations de base sont exécutées principalement à l'aide de la commande `vxassist`. `vxassist` crée les plex et les sous-disques requis en n'utilisant comme entrée que les attributs de base du volume voulu. `vxassist` peut également modifier les volumes existants. Cette commande modifie automatiquement tous les objets sous-jacents ou associés. Elle conserve les valeurs par défaut pour la plupart des attributs de volumes, sauf si vous indiquez des valeurs spécifiques.
- ◆ **Méthode avancée** — Cette méthode consiste en un certain nombre de commandes qui nécessitent généralement une entrée détaillée de votre part. Ces commandes utilisent le principe de « bloc de construction », qui nécessite une connaissance approfondie de la structure et des composants sous-jacents pour pouvoir exécuter manuellement les commandes nécessaires à l'exécution d'une tâche donnée. Les opérations avancées sont exécutées via plusieurs commandes de Volume Manager.

Les étapes suivantes sont normalement incluses dans la création manuelle des volumes :

- ◆ création des sous-disques ;
- ◆ création des plex ;
- ◆ association des sous-disques et des plex ;
- ◆ création des volumes ;
- ◆ association des volumes et des plex ;
- ◆ initialisation des volumes.

La création d'un volume comprend la création des composants de plex et de sous-disques. Si vous choisissez la méthode de base pour créer les volumes, vous devez indiquer les caractéristiques voulues du volume ; les plex et les sous-disques sous-jacents sont alors créés automatiquement.

Création de volumes

Vous pouvez attribuer divers types de structure aux volumes créés :

- ◆ Concaténé — Volume dont les sous-disques sont organisés de manière séquentielle et contiguë dans un plex. La concaténation permet de créer un volume à partir de plusieurs zones d'un ou de plusieurs disques, si une même zone de disque ne contient pas un espace suffisant pour un volume entier.
- ◆ Réparti par bande — Volume dont les données sont également réparties sur plusieurs disques. Les *bandes* sont des fragments de taille identique alloués en alternance et régulièrement aux sous-disques d'un même plex. Un plex réparti par bande doit avoir au moins deux sous-disques, appartenant chacun à un disque différent. La capacité de traitement augmente avec le nombre de disques sur lesquels un plex est réparti. La répartition des données par bandes sur plusieurs disques permet d'équilibrer les E/S lorsque certains sous-disques ont des zones de trafic élevé.
- ◆ RAID 5 — Volume qui utilise la répartition des données par bandes sur plusieurs disques pour répartir régulièrement les données et la parité sur plusieurs disques d'un ensemble de disques. Chaque bande contient une unité de bande de parité et des unités de bande de données. La parité permet de reconstruire les données en cas de défaillance de l'un des disques. Par rapport aux performances des volumes répartis par bandes, la capacité de traitement des écritures est plus limitée pour les volumes RAID 5, puisqu'il faut mettre à jour les informations de parité lors de chaque accès aux données. Cependant, par rapport à la mise en miroir, l'utilisation de la parité réduit l'espace requis.
- ◆ Mis en miroir — Volume ayant plusieurs plex qui dupliquent les informations d'un volume. Un volume peut n'avoir qu'un plex, mais deux plex au moins sont requis pour une véritable mise en miroir (redondance des données). Pour optimiser la redondance, chaque plex doit avoir un espace disque provenant de disques différents.
- ◆ Réparti en couches — Volume créé sur les volumes. Vous pouvez construire des volumes répartis en couches en mappant un sous-disques sur un disque VM ou sur un volume de *stockage*. Un volume de stockage assure un niveau de structure récurrent similaire aux volumes de niveau supérieur. Les volumes répartis en couches augmentent les possibilités de combinaison pour les structures logiques.

Remarque Dans l'interface utilisateur graphique de VERITAS Volume Manager Storage Administrator, le terme *Pro réparti par bande* désigne un miroir réparti par bandes et *Pro concaténé*, un miroir concaténé.

La commande `vxassist` offre la méthode la plus simple pour créer de nouveaux volumes. Procédez comme suit :

```
# vxassist make nom_volume longueur [attributes]
```

où `make` représente le mot-clé pour la création de volume, *nom_volume* le nom attribué au volume et *longueur* le nombre de secteurs (par défaut) du volume. Vous pouvez exprimer la longueur en Ko, Mo ou Go au moyen d'un caractère de suffixe (respectivement k, m ou g). Voir la page `man vxintro (1M)` pour plus d'informations sur la définition des unités de longueur lors de la création des volumes. Le cas échéant, vous pouvez spécifier d'autres attributs.

Par défaut, `vxassist` crée les volumes dans le disk group `rootdg`. Pour spécifier un autre disk group, ajoutez `-g diskgroup` dans la ligne de commande `vxassist`.

La création d'un volume dans le disk group `rootdg` crée deux fichiers de nœud de périphérique qui permettent d'accéder à ce volume :

- ◆ `/dev/vx/dsk/nom_volume` (le nœud de périphérique en mode bloc pour le volume)
- ◆ `/dev/vx/rdsk/nom_volume` (le nœud de périphérique en mode raw pour le volume)

Pour les volumes appartenant à `rootdg` et les disk groups différents de `rootdg`, ces noms comprennent le nom du disk group, comme suit :

- ◆ `/dev/vx/dsk/nom_diskgroup/nom_volume`
- ◆ `/dev/vx/rdsk/nom_diskgroup/nom_volume`

Voir « Création d'un volume concaténé » à la page 181 pour obtenir une description de la méthode la plus simple pour créer un volume (par défaut). Les sections suivantes décrivent comment créer des volumes qui ont des attributs spécifiques.

Création d'un volume concaténé

Par défaut, `vxassist` crée un volume concaténé qui utilise une ou plusieurs sections de l'espace disque. Sur un disque fragmenté, cette procédure vous permet de constituer un volume dont la capacité dépasse celle de n'importe quelle section individuelle de l'espace disque disponible.

Si l'espace disponible sur un même disque est insuffisant, `vxassist` crée un volume réparti. Un volume réparti est un volume concaténé dont les sections de l'espace disque sont réparties sur plusieurs disques. Sa capacité peut dépasser celle du plus grand disque, car son espace provient de plusieurs disques.

Création de volumes

Création d'un volume concaténé sur un disque

Si vous ne spécifiez pas de disque, Volume Manager sélectionne le disque sur lequel créer le volume.

Pour créer un volume concaténé par défaut, entrez la commande suivante :

```
# vxassist make nom_volume longueur
```

où *nom_volume* représente le nom du volume et *longueur*, la longueur du volume en secteurs (sauf si vous avez spécifié une autre unité de taille au moyen d'un caractère de suffixe). Lorsque vous créez un volume, vous pouvez spécifier sa longueur en secteurs, Ko, Mo ou Go. L'unité de mesure est ajoutée à la longueur sous la forme d'un suffixe (respectivement, s, m, k ou g). Si vous n'ajoutez pas d'unité, le système utilise par défaut les secteurs.

Vous pouvez utiliser la commande suivante pour créer le volume `voldefault`, d'une longueur de 10 Mo :

```
# vxassist make voldefault 10m
```

Création d'un volume concaténé sur un disque spécifique

Volume Manager sélectionne automatiquement le ou les disques sur lesquels se trouve chaque volume, sauf spécification contraire de votre part. Si vous voulez qu'un volume se trouve sur un disque spécifique, vous devez indiquer ce disque à Volume Manager. Vous pouvez spécifier plusieurs disques.

Pour créer un volume sur un disque spécifique, entrez la commande suivante :

```
# vxassist make nom_volume longueur nom_disque [...]
```

Pour créer le volume `volspecific` sur `disk03`, entrez la commande suivante :

```
# vxassist make volspecific 3m disk03
```

Création d'un volume réparti par bandes

Un volume réparti par bandes contient au moins un plex comportant au minimum deux sous-disques situés sur deux disques physiques ou plus.

Pour créer un volume réparti par bandes, entrez la commande suivante :

```
# vxassist make nom_volume longueur layout=stripe
```

Pour créer le volume réparti par bandes `volzebra`, entrez la commande suivante :

```
# vxassist make volzebra 10m layout=stripe
```

Cette commande crée un volume réparti par bandes dont la taille d'unité de bande est celle par défaut sur le nombre de disques par défaut.

Pour indiquer les disques sur lesquels créer les volumes, spécifiez leur nom à la fin de la ligne de commande. Par exemple, pour créer un volume réparti par bandes de 30 Mo sur trois disques spécifiques (disk03, disk04 et disk05), entrez la commande suivante :

```
# vxassist make stripevol 30m layout=stripe disk03 disk04\disk05
```

Création d'un volume RAID 5

Un volume RAID 5 contient un plex RAID 5 avec au moins deux sous-disques appartenant à deux disques physiques ou plus. Il ne peut exister qu'un seul plex RAID 5 par volume. Un volume RAID 5 peut également contenir un ou plusieurs plex de journal RAID 5, qui contiennent des informations sur les données et la parité écrites sur le volume.

Pour créer un volume RAID 5, entrez la commande suivante :

```
# vxassist make nom_volume longueur layout=raid5
```

Pour créer le volume RAID 5 volraid, entrez la commande suivante :

```
# vxassist make volraid 10m layout=raid5
```

Cette commande crée un volume RAID 5 ayant la taille d'unité de bande par défaut sur le nombre de disques par défaut. Elle crée également un journal RAID 5 par défaut.

Création d'un volume mis en miroir

Pour créer un volume mis en miroir, entrez la commande suivante :

```
# vxassist make nom_volume longueur layout=mirror
```

Pour créer le volume mis en miroir volmir, entrez la commande suivante :

```
# vxassist make volmir 5m layout=mirror
```

Redimensionnement d'un volume

Le redimensionnement d'un volume modifie la taille du volume. Pour cela, vous pouvez utiliser `vxassist` ou `vxresize`.

Si le volume n'est pas suffisant pour conserver la quantité de données prévue, vous devez étendre sa longueur. Dans ce cas, `vxassist` recherche automatiquement l'espace disque disponible.

Lorsque vous redimensionnez un volume, vous pouvez spécifier sa longueur en secteurs, Ko, Mo ou Go. L'unité de mesure est ajoutée à la longueur sous la forme d'un suffixe (respectivement, s, m, k ou g). Si vous n'ajoutez pas d'unité, le système utilise par défaut les secteurs.

 Redimensionnement d'un volume

Attention Ne réduisez pas la taille d'un volume en dessous de celle du système de fichiers. Si vous disposez d'un système de fichiers VxFS, vous pouvez réduire d'abord le système de fichiers, puis le volume. Si vous ne réduisez pas d'abord le système de fichiers, vous risquez de perdre définitivement des données.

Redimensionnement des volumes au moyen de vxassist

vxassist permet de redimensionner un volume comme suit :

- growto — Augmente le volume à la longueur spécifiée
- growby — Augmente le volume de la quantité spécifiée
- shrinkto — Réduit le volume à la longueur spécifiée
- shrinkby — Réduit le volume à la quantité spécifiée

Attention Vous ne pouvez ni augmenter ni réduire un volume associé à un disque d'initialisation encapsulé (rootvol, usr, var, opt, swapvol, etc.) ; en effet, ces volumes sont mappés sur une partition physique sous-jacente située sur le disque et ils doivent être contigus. Lorsque vous tentez d'augmenter rootvol, usrvol, varvol ou swapvol, vous ne pourrez peut-être plus démarrer le système si vous devez effectuer le démarrage à partir des partitions. De plus, cette opération risque également d'empêcher les mises à niveau de Solaris et vous devrez peut-être effectuer une installation complète. Par ailleurs, le script upgrade_start risque d'échouer.

Augmentation à une longueur donnée

Pour étendre un volume à une longueur spécifique, entrez la commande suivante :

```
# vxassist growto nom_volume longueur
```

Pour étendre volcat à 2000 secteurs, entrez la commande suivante :

```
# vxassist growto volcat 2000
```

Augmentation d'une longueur donnée

Pour étendre un volume d'une longueur spécifique, entrez la commande suivante :

```
# vxassist growby nom_volume longueur
```

Pour étendre volcat de 100 secteurs, entrez la commande suivante :

```
# vxassist growby volcat 100
```

Réduction à une longueur donnée

Pour réduire un volume à une longueur spécifique, entrez la commande suivante :

```
# vxassist shrinkto nom_volume longueur
```

Assurez-vous que vous ne réduisez pas le volume en dessous de la taille actuelle du système de fichiers ou de la base de données qui utilisent ce volume. Vous pouvez utiliser sans problème cette commande sur des volumes vides.

Pour réduire `volcat` à 1300 secteurs, entrez la commande suivante :

```
# vxassist shrinkto volcat 1300
```

Réduction d'une longueur donnée

Pour réduire un volume d'une longueur spécifique, entrez la commande suivante :

```
# vxassist shrinkby nom_volume longueur
```

Pour réduire `volcat` de 300 secteurs, entrez la commande suivante :

```
# vxassist shrinkby volcat 300
```

Redimensionnement des volumes au moyen de `vxresize`

Vous pouvez utiliser la commande `vxresize` pour redimensionner un volume contenant un système de fichiers. Vous pouvez choisir d'autres commandes, mais `vxresize` présente l'avantage de redimensionner automatiquement le système de fichiers et le volume. Pour plus d'informations sur l'utilisation de `vxresize`, voir la page `man vxresize(1M)`. `vxresize` ne vous permet de redimensionner que les systèmes de fichiers `vxfs` et `ufs`.

Si la taille du volume excède vos besoins, vous pouvez la réduire.

Suppression d'un volume

Lorsqu'un volume n'est plus nécessaire (par exemple, s'il est inactif et archivé), vous pouvez le supprimer afin de libérer l'espace disque à d'autres fins ; pour cela, vous pouvez utiliser la commande `vxassist remove volume`.

La commande `vxassist` permet de supprimer un volume entier. Utilisez les mots-clé `remove` et `volume` et indiquez le nom du volume sur la ligne de commande comme illustré dans cet exemple :

```
# vxassist remove volume nom_volume
```

Mise en miroir d'un volume

Un miroir est une copie d'un volume qui n'est pas stockée sur le ou les mêmes disques que la copie d'origine du volume. La mise en miroir d'un volume protège ses données contre tout risque de perte en cas de défaillance de l'un des disques.

Création d'un volume mis en miroir

Pour créer un volume en miroir, utilisez la commande `vxassist` :

```
# vxassist make nom_volume longueur layout=mirror
```

Mise en miroir d'un volume existant

Vous pouvez ajouter un miroir (plex) à un volume existant, au moyen de la commande `vxassist` :

```
# vxassist mirror nom_volume
```

Par exemple :

```
# vxassist mirror voltest
```

Pour mettre en miroir un volume existant, vous pouvez aussi créer un plex, puis l'associer à un volume au moyen des commandes suivantes :

```
# vxmake plex nom_plex sd=nom_sous-disque ...  
# vxplex att nom_volume nom_plex
```

Création d'un volume avec activation du journal des zones modifiées

Pour créer un volume mis en miroir avec activation du journal des zones modifiées (DRL, *Dirty Region Logging*), créez un volume mis en miroir avec un journal au moyen de la commande suivante :

```
# vxassist make nom_volume longueur layout=mirror,log
```

`vxassist` crée par défaut un plex de journal pour chaque sous-disque de journal.

Mise en miroir de tous les volumes

Pour mettre en miroir tous les volumes existants du système sur un espace disque disponible, entrez la commande suivante :

```
# /etc/vx/bin/vxmirror -g diskgroup -a
```

Vous pouvez aussi configurer Volume Manager pour créer par défaut des volumes mis en miroir. Pour cela, entrez la commande suivante :

```
# /etc/vx/bin/vxmirror -d yes
```

Si vous effectuez cette modification, vous pouvez tout de même créer des volumes non mis en miroir en spécifiant `nmirror=1` comme attribut de la commande `vxassist`. Par exemple, pour créer un volume non mis en miroir de 20 Mo nommé `nomirror`, entrez la commande suivante :

```
# vxassist make nomirror 20m nmirror=1
```

Suppression d'un miroir

Lorsqu'un miroir n'est plus nécessaire, vous pouvez le supprimer afin de libérer de l'espace.

Remarque Vous ne pouvez pas supprimer le dernier plex valide associé à un volume.

Attention Pour enregistrer les données à supprimer d'un miroir, vous devez connaître la configuration de ce miroir. Les paramètres de cette configuration (taille d'unité de bande et ordre des sous-disques) sont très importants pour garantir que le nouveau miroir contiendra les mêmes données. Avant de supprimer ce type de miroir, vous devez enregistrer sa configuration.

Pour dissocier et supprimer un miroir du volume associé, entrez la commande suivante :

```
# vxplex -o rm dis nom_plex
```

Pour dissocier et supprimer le miroir `vol01-02`, entrez la commande suivante :

```
# vxplex -o rm dis vol01-02
```

Cette commande supprime le miroir `vol01-02` et tous les sous-disques associés.

Ajout d'un journal RAID 5

Il ne peut exister qu'un plex RAID 5 par volume RAID 5. Les plex supplémentaires sont tous des plex de journal RAID 5, qui permettent d'enregistrer des informations sur les données et la parité écrite dans le volume. Lorsque vous créez un volume RAID 5 au moyen de `vxassist`, le système crée par défaut un plex de journal pour ce volume.

Pour ajouter un journal RAID 5 à un volume existant, entrez la commande suivante :

```
# vxassist addlog nom_volume
```

Pour créer un journal pour le volume RAID 5 `volraid`, entrez la commande suivante :

```
# vxassist addlog volraid
```

Ajout d'un journal des zones modifiées

Pour activer le journal des zones modifiées d'un volume, vous devez ajouter un sous-disque de journal à ce volume, qui doit de plus être mis en miroir. Il ne peut exister qu'un sous-disque de journal par plex.

Pour ajouter un journal des zones modifiées à un volume existant, entrez la commande suivante :

```
# vxassist addlog nom_volume
```

Pour créer un journal pour le volume `vol03`, entrez la commande suivante :

```
# vxassist addlog vol03
```

Lorsque vous utilisez `vxassist` pour ajouter un sous-disque de journal à un volume, le système crée également par défaut un plex de journal contenant le sous-disque de journal par défaut.

Une fois que vous avez créé le plex contenant un sous-disque de journal, vous pouvez le traiter comme un plex normal. Vous pouvez lui ajouter des sous-disques de données, le supprimer, ainsi que le sous-disque de journal, en utilisant les mêmes procédures de suppression que pour les plex et sous-disques ordinaires.

Suppression d'un journal des zones modifiées

Pour supprimer un journal, utilisez la commande `vxassist` comme suit :

```
# vxassist remove log nom_volume
```

Utilisez l'attribut `nlog=` pour spécifier le nombre de journaux à supprimer. Par défaut, `vxassist` supprime un journal.

Suppression d'un journal RAID 5

Pour supprimer un journal RAID 5, utilisez la commande `vxassist` comme suit :

```
# vxassist remove log nom_volume
```

Utilisez l'attribut `nlog=` pour spécifier le nombre de journaux à supprimer. Par défaut, `vxassist` supprime un journal.

Arrêt d'un volume

Lorsque vous arrêtez un volume, il n'est plus disponible. Son état passe de `ENABLED` ou `DETACHED` à `DISABLED`. Si la commande entrée ne peut pas l'arrêter, le volume conserve son état actuel. Pour arrêter un volume, entrez la commande suivante :

```
# vxvol stop nom_volume ...
```

Par exemple, entrez la commande suivante pour arrêter le volume `vol01` :

```
# vxvol stop vol01
```

Pour arrêter tous les volumes dont l'état est `ENABLED`, entrez la commande suivante :

```
# vxvol stopall
```

Démarrage d'un volume

Lorsque vous démarrez un volume, celui-ci peut alors être utilisé.

Son état passe de `DISABLED` ou `DETACHED` à `ENABLED`. S'il est impossible de démarrer un volume, il conserve son état actuel. Pour démarrer un volume, entrez la commande suivante :

```
# vxrecover -s nom_volume ...
```

Pour démarrer tous les volumes dont l'état est `DISABLED`, entrez la commande suivante :

```
# vxrecover -s
```

Modification de la stratégie de lecture du volume

Liste des volumes impossibles à démarrer

Si vous ne pouvez pas démarrer un volume, celui-ci n'est peut-être pas configuré correctement ou il présente d'autres erreurs ou conditions qui l'empêchent de démarrer. Pour afficher les volumes impossibles à démarrer, utilisez la commande `vxinfo`. Cette commande affiche des informations sur l'accessibilité et l'utilisation possible d'un ou de plusieurs volumes :

```
# vxinfo [nom_volume]
```

Modification de la stratégie de lecture du volume

Volume Manager vous permet de choisir l'une des stratégies de lecture suivantes :

- ◆ `round` lit chaque plex de façon alternative pour chaque E/S non séquentielle détectée. Un accès séquentiel ne permet d'accéder qu'à un seul plex. Ce mode tire parti des stratégies de mise en cache des lectures anticipées du lecteur ou du contrôleur.
- ◆ `prefer` commence la lecture à partir du plex défini comme plex préféré.
- ◆ `select` choisit une stratégie par défaut en fonction des associations de plex au volume. Si un plex réparti par bandes est activé pour le volume, `select` choisit par défaut ce plex ; dans le cas contraire, il choisit par défaut le mode de lecture alternative.

Vous pouvez modifier la stratégie de lecture en passant de `round` à `prefer` (ou inversement) ou en choisissant un autre plex préféré. La commande `vxvol rdpol` définit la stratégie de lecture pour un volume.

Pour définir la stratégie de lecture `round`, entrez la commande suivante :

```
# vxvol rdpol round nom_volume
```

Par exemple, la commande suivante permet de définir la stratégie de lecture alternative pour le volume `vol01` :

```
# vxvol rdpol round vol01
```

Pour définir la stratégie de lecture `prefer`, entrez la commande suivante :

```
# vxvol rdpol prefer nom_volume nom_plex_préféré
```

Par exemple, la commande suivante permet de définir pour `vol01` la stratégie de lecture préférentielle à partir du plex `vol01-02` :

```
# vxvol rdpol prefer vol01 vol01-02
```

Pour définir la stratégie de lecture `select`, entrez la commande suivante :

```
# vxvol rdpol select nom_volume
```

Préparation d'un volume pour la restauration à partir d'une sauvegarde

Il est important de faire des copies de sauvegarde des volumes. Vous disposez ainsi d'une copie des données au moment de la sauvegarde. Les copies de sauvegarde permettent de restaurer les volumes perdus à la suite d'une défaillance de disque ou les données détruites à la suite d'une erreur humaine. Volume Manager vous permet de sauvegarder des volumes tout en garantissant une interruption minimale des activités.

Pour sauvegarder un volume au moyen de `vxassist`, procédez comme suit :

1. Créez un miroir de prise de vue du volume à sauvegarder.

`vxassist snapstart task` crée un miroir de sauvegarde défini en écriture seule, qui est attaché et synchronisé avec le volume à sauvegarder. Lorsque la synchronisation est effectuée, le miroir de sauvegarde est prêt pour être utilisé en tant que miroir de prise de vue. Cependant, sa mise à jour se poursuit tant que vous ne le détachez pas lors de l'étape de prise de vue de la procédure. Cette opération peut prendre un certain temps, selon la taille du volume.

Pour créer un miroir de prise de vue pour un volume, entrez la commande suivante :

```
# vxassist snapstart nom_volume
```

Pour créer un miroir de prise de vue du volume `voldef`, entrez la commande suivante :

```
# vxassist snapstart voldef
```

2. Choisissez l'heure appropriée pour créer un volume de prise de vue.

Si possible, choisissez de créer la prise de vue lorsque le moins possible d'utilisateurs accèdent au volume.

3. Créez un volume de prise de vue qui reflète le volume d'origine au moment de la prise de vue.

Pour terminer la procédure de sauvegarde en ligne, exécutez la commande `vxassist snapshot` sur le volume ayant un miroir de prise de vue. Cette opération détache le miroir de prise de vue terminé, crée un nouveau volume normal et lui attache le miroir de prise de vue. La prise de vue devient alors un volume en lecture seule. Cette étape ne devrait pas prendre plus de quelques minutes.

Récupération d'un volume

Pour créer un volume de prise de vue, entrez la commande suivante :

```
# vxassist snapshot nom_volume nouveau_nom_volume
```

Pour créer un volume de prise de vue de `voldef`, entrez la commande suivante :

```
# vxassist snapshot voldef snapvol
```

Le volume de prise de vue peut à présent être utilisé par les utilitaires de sauvegarde tandis que le volume d'origine reste à la disposition des applications et des utilisateurs.

Vous pouvez employer la méthode de votre choix pour sauvegarder le volume de prise de vue. Pour utiliser au mieux l'espace, pensez à supprimer le volume de prise de vue à l'issue de la sauvegarde. En effet, ce volume occupe autant d'espace que le volume d'origine.

Récupération d'un volume

Une panne du système ou une erreur d'E/S peuvent endommager un ou plusieurs plex d'un volume et n'en laisser aucun dans l'état CLEAN ou ACTIVE. Vous pouvez marquer l'un des plex comme CLEAN et indiquer au système d'utiliser ce plex comme source pour réactiver les autres.

Pour mettre un plex dans l'état CLEAN, entrez la commande suivante :

```
# vxmend fix clean nom_plex
```

Par exemple, la commande suivante permet de mettre le plex `vol01-02` dans l'état CLEAN :

```
# vxmend fix clean vol01-02
```

Pour plus d'informations sur l'utilisation de `vxmend`, voir la page `man vxmend (1M)`.

Affichage d'informations sur un volume

Vous pouvez utiliser la commande `vxprint` pour afficher des informations sur la configuration d'un volume.

Vous pouvez entrer la commande suivante afin d'afficher des informations sur les enregistrements de volumes, de plex et de sous-disques pour tous les volumes du système :

```
# vxprint -ht
```

La commande `vxprint` peut par exemple contenir les informations suivantes :

Disk group: rootdg

DG NOM	CONFIG	JOURN	MINEURS	ID GROUPE				
DM NOM	PÉRIPH	TYPE	LONGPRIV	LONGPUB	ÉTAT			
V NOM	TYPEUTIL	KSTATE	ÉTAT	LONG	RDPOL	PLEXPRÉF		
PL NOM	VOLUME	KSTATE	ÉTAT	LONG	STRUCT	COL/LARG	MODE	
SD NOM	PLEX	DISQUE	DÉCDISQ	LONG	COL/]DÉC	PÉRIPH.	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	

où `dg` représente un disk group, `dm` un disque, `v` un volume, `pl` un plex et `sd` un sous-disque. Les premières lignes indiquent les en-têtes correspondant à chaque type de ligne de sortie qui suit. Chaque volume est listé avec son ou ses plex et sous-disques associés.

Pour afficher les informations sur un volume spécifique, entrez la commande suivante :

```
# vxprint -t nom_volume
```

Pour afficher des informations sur `voldef`, entrez la commande suivante :

```
# vxprint -t voldef
```

La commande `vxprint` peut par exemple afficher les informations suivantes :

Disk group: rootdg

V NOM	TYPEUTIL	KSTATE	ÉTAT	LONG	RDPOL	PLEXPRÉF
v voldef	fsgen	ENABLED	ACTIVE	20480	SELECT	-

Opérations de plex

Les plex sont des groupements logiques de sous-disques constituant une zone d'espace disque sans restriction de taille du disque physique ou autres. La réplication (mise en miroir) des données d'un disque s'effectue en créant plusieurs plex pour un seul volume. Chaque plex contient une copie identique des données du volume. La réplication assurée par la mise en miroir empêche toute perte de données en cas de défaillance d'un disque car chaque plex doit se trouver sur un disque différent. L'utilisation de plusieurs plex permet également d'accroître l'intégrité et la fiabilité des données.

Les opérations de plex sont les suivantes :

- ◆ création d'un plex ;
- ◆ sauvegarde à l'aide de la mise en miroir ;
- ◆ association d'un plex ;
- ◆ dissociation et suppression d'un plex ;
- ◆ liste de tous les plex ;
- ◆ affichage de tous les plex ;
- ◆ modification des attributs de plex ;
- ◆ modification de l'état d'un plex ;
- ◆ déplacement des plex :
- ◆ copie des plex.

Création d'un plex

La commande `vxmake` crée des objets Volume Manager, par exemple des plex. Pour créer un plex, vous devez identifier des sous-disques et les associer au plex que vous voulez créer.

Pour créer un plex à partir de sous-disques existants, entrez la commande suivante :

```
# vxmake plex nom_plex sd=nom_sous-disque,...
```

Par exemple, la commande suivante permet de créer le plex concaténé `vol01-02` à l'aide des deux sous-disques `disk02-01` et `disk02-02` :

```
# vxmake plex vol01-02 sd=disk02-01,disk02-02
```

Création d'un plex réparti par bandes

Pour créer un plex réparti par bandes, vous devez spécifier des attributs supplémentaires. Par exemple, la commande suivante permet de créer le plex réparti par bandes `p1-01` avec une largeur de bande de 32 secteurs et 2 colonnes :

```
# vxmake plex p1-01 layout=stripe stwidth=32 ncolumn=2 \  
sd=disk01-01,disk02-01
```

Si vous voulez utiliser un plex pour créer un volume, vous devez l'associer à ce volume (Pour plus d'informations, voir « Association d'un plex » à la page 195).

Association d'un plex

Un plex devient un actif pour un volume lorsqu'il est associé à ce volume. Pour associer un plex à un volume existant, entrez la commande suivante :

```
# vxplex att nom_volume nom_plex
```

Par exemple, la commande suivante permet d'associer le plex `vol01-02` au volume `vol01` :

```
# vxplex att vol01 vol01-02
```

Si le volume n'a pas été créé, vous pouvez associer un plex (ou plusieurs plex) au volume à créer lors de la création du volume à l'aide de la commande suivante :

```
# vxmake -U usetype vol nom_volume plex=nom_plex1, nom_plex2...
```

Par exemple, la commande suivante permet de créer le volume `home` de type `fsgen` mis en miroir et de lui associer les deux plex existants `home-1` et `home-2` :

```
# vxmake -Ufsgen vol home plex=home-1,home-2
```

Remarque Vous pouvez aussi utiliser la commande suivante sur un volume existant pour ajouter et associer un plex : `# vxassist mirror nom_volume`

Dissociation et suppression d'un plex

Lorsqu'un plex est inutile, vous pouvez le supprimer. Vous devrez peut-être supprimer un plex dans les cas suivants :

- ◆ pour libérer de l'espace disque ;
- ◆ pour réduire le nombre de miroirs dans un volume afin de pouvoir augmenter la longueur d'un autre miroir et de son volume associé. Lorsque vous supprimez des plex et des sous-disques, vous pouvez ajouter l'espace ainsi libéré aux autres volumes ;
- ◆ pour supprimer un miroir temporaire qui a été créé pour sauvegarder un volume et qui est devenu inutile ;
- ◆ pour modifier la structure d'un plex.

Attention Pour enregistrer les données à supprimer d'un plex, vous devez connaître la configuration de ce plex. Les paramètres de cette configuration (taille d'unité de bande et ordre des sous-disques) sont très importants pour garantir que le nouveau plex contiendra les mêmes données. Avant de supprimer un plex, vous devez enregistrer sa configuration. Pour plus d'informations, voir « Affichage d'informations sur les plex » à la page 198.

Pour dissocier et supprimer un plex du volume associé, entrez la commande suivante :

```
# vxplex -o rm dis nom_plex
```

Pour dissocier et supprimer le plex vol01-02, entrez la commande suivante :

```
# vxplex -o rm dis vol01-02
```

Cette commande supprime le plex vol01-02 et tous les sous-disques associés.

Vous pouvez commencer par dissocier le plex et les sous-disques, puis les supprimer au moyen des commandes :

```
# vxplex dis nom_plex  
# vxedit -r rm nom_plex
```

L'utilisation conjointe de ces commandes produit les mêmes résultats que `vxplex -o rm dis`.

Sauvegarde à l'aide de la mise en miroir

Si un volume est mis en miroir, vous pouvez y exécuter la sauvegarde en mettant l'un de ses miroirs hors ligne pendant une durée déterminée. Cette procédure rend inutile l'espace disque supplémentaire requis uniquement pour l'opération de sauvegarde. Cependant, elle supprime également la redondance du volume pendant toute la durée de la sauvegarde.

Remarque Les informations contenues dans cette section ne s'appliquent pas aux volumes RAID 5.

Pour effectuer la sauvegarde d'un volume mis en miroir d'un système actif, procédez comme suit :

1. Arrêtez brièvement les activités des utilisateurs (facultatif) afin d'améliorer la cohérence de la sauvegarde.
2. Dissociez l'un des miroirs de volume (vol01-01, dans cet exemple) :

```
# vxplex dis vol01-01
```
3. Créez un nouveau volume temporaire à l'aide du plex dissocié :

```
# vxmake -U gen vol tempvol plex=vol01-01
```
4. Démarrez le volume temporaire :

```
# vxvol start tempvol
```
5. Exécutez les procédures de sauvegarde appropriées à l'aide du volume temporaire.
6. Arrêtez le volume temporaire :

```
# vxvol stop tempvol
```
7. Dissociez le plex de sauvegarde de son volume temporaire :

```
# vxplex dis vol01-01
```
8. Réassociez le plex de sauvegarde à son volume d'origine pour rétablir la redondance du volume :

```
# vxplex att vol01 vol01-01
```
9. Supprimez le volume temporaire :

```
# vxedit rm tempvol
```

Opérations de plex

Pour plus d'informations sur une autre méthode de sauvegarde en ligne à l'aide de la commande `vxassist`, voir « Exécution d'une sauvegarde en ligne » à la page 208.

Affichage d'informations sur les plex

L'affichage de la liste des plex permet d'identifier les plex libres pour la construction de volumes. L'utilisation de la commande `vxprint` avec l'option plex (`-p`) permet d'afficher des informations sur tous les plex.

Pour afficher des informations détaillées sur tous les plex du système, entrez la commande suivante :

```
# vxprint -lp
```

Pour afficher des informations détaillées sur un plex spécifique, entrez la commande suivante :

```
# vxprint -l nom_plex
```

L'option `-t` imprime une seule ligne d'informations sur le plex. Pour afficher la liste des plex libres, entrez la commande suivante :

```
# vxprint -pt
```

Modification des attributs d'un plex

Attention Exercez la plus grande précaution lorsque vous modifiez les attributs des plex, et ne le faites que si cela est nécessaire.

La commande `vxedit` modifie les attributs des plex et des autres objets Volume Manager. Pour modifier les attributs des plex, entrez la commande suivante :

```
# vxedit set field=value ... nom_plex ...
```

Les commandes Volume Manager utilisent le champ `comment`, ainsi que les champs `putil` et `tutil`, une fois les plex créés. Les attributs `putil` sont conservés au redémarrage ; les champs `tutil` sont temporaires et les valeurs indiquées ne sont pas conservées au redémarrage.

Les champs `putil` et `tutil` comprennent trois fonctions et sont numérotés en fonction de ces dernières. Vous pouvez les modifier si nécessaire. Volume Manager utilise les champs d'utilitaire marqués `putil0` et `tutil0`. Les autres produits VERITAS utilisent les champs marqués `putil1` et `tutil1`. Les champs `putil2` et `tutil2` sont des champs utilisateur. Le tableau 4 indique les fonctions des champs `putil` et `tutil`.

Tableau 4. Champs putil[n] et tutil[n]

Champ	Description des champs d'utilitaire
putil0	Réservé aux commandes Volume Manager. Ce champ est conservé au redémarrage.
putil1	Réservé aux utilitaires de haut niveau tels que l'interface utilisateur graphique Volume Manager. Ce champ est conservé au redémarrage.
putil2	Réservé à l'administrateur système ou aux applications spécifiques du site. Ce champ est conservé au redémarrage.
tutil0	Réservé aux commandes Volume Manager. Ce champ est effacé au redémarrage.
tutil1	Réservé aux utilitaires de haut niveau tels que l'interface utilisateur graphique Volume Manager. Ce champ est effacé au redémarrage.
tutil2	Réservé à l'administrateur système ou aux applications spécifiques du site. Ce champ est effacé au redémarrage.

La commande :

```
# vxedit set comment="mon plex" tutil2="u" user="admin" vol01-02
```

utilise vxedit pour définir les attributs suivants :

- ◆ la valeur mon plex est entrée dans le champ comment (identifiant l'utilisation du plex) ;
- ◆ la valeur u est entrée dans le champ tutil2 pour indiquer que le sous-disque est en cours d'utilisation ;
- ◆ la valeur admin est choisie comme ID d'utilisateur.

Pour éviter d'associer un plex particulier à un volume, entrez une chaîne non nulle dans le champ putil0, comme dans l'exemple ci-dessous :

```
# vxedit set putil0="DO-NOT-USE" vol01-02
```

Modification de l'état d'un plex : détachement et attachement d'un plex

Une fois que vous avez créé et mis en ligne un volume (ENABLED), Volume Manager peut déconnecter temporairement ses plex. Cette fonction est utile, par exemple, lorsque vous devez réparer le composant matériel sur lequel se trouve le plex ou lorsque vous ne pouvez pas démarrer un volume et que vous devez choisir manuellement un plex source pour réactiver ce volume.

Pour corriger une défaillance de disque ou du système, vous devez mettre un volume hors ligne, puis attacher et détacher ses plex. Les deux commandes utilisées pour corriger une défaillance de disque sont vxmend et vxplex.

Opérations de plex

Pour mettre un plex hors ligne afin de pouvoir exécuter les opérations de réparation ou de maintenance sur le disque physique contenant les sous-disques de ce plex, entrez la commande suivante :

```
# vxmend off nom_plex ..
```

En cas de défaillance de la tête de disque, mettez hors ligne tous les plex dont les sous-disques associés sont sur le disque affecté. Par exemple, si des sous-disques des plex vol101-02 et vol102-02 se trouvent sur un lecteur à réparer, entrez la commande suivante :

```
# vxmend off vol101-02 vol102-02
```

Cette commande met vol101-02 et vol102-02 dans l'état OFFLINE ; ils conservent cet état tant que vous n'apportez pas de nouvelle modification.

Détachement d'un plex

Pour détacher temporairement un plex d'un volume mis en miroir, entrez la commande suivante :

```
# vxplex det nom_plex
```

Par exemple, la commande suivante permet de détacher temporairement le plex vol101-02 et de le placer en mode de maintenance :

```
# vxplex det vol101-02
```

Cette commande détache temporairement le plex, mais conserve son association à son volume. Cependant, le plex n'est pas utilisé pour les E/S. S'il est détaché à l'aide de la commande précédente, il sera récupéré au redémarrage du système. L'état du plex est STALE : ainsi, si une commande vxvol start est exécutée sur le volume approprié (par exemple, au redémarrage du système), le contenu de ce plex est récupéré et mis dans l'état ACTIVE.

Lorsque le plex est prêt à redevenir une partie active de son volume, procédez comme suit :

- ◆ Si le volume n'est pas dans l'état ENABLED, démarrez-le en entrant la commande suivante :

```
# vxvol start nom_volume
```

Si vous ne pouvez pas le démarrer, mettez l'un des plex dans l'état CLEAN en entrant la commande suivante :

```
# vxmend fix clean nom_plex
```

Démarrez ensuite le volume.

- ◆ Si l'état de noyau du plex n'est pas ENABLED, entrez la commande suivante :

```
# vxplex att nom_volume nom_plex ...
```

À l'instar de ce qui se passe lorsque vous changez l'état d'un plex de OFFLINE à ACTIVE, cette commande récupère le contenu des plex, puis leur attribue l'état ACTIVE.

Attachement d'un plex

Une fois que le disque est réparé ou remplacé, et qu'il est prêt à être utilisé, vous devez remettre les plex en ligne (l'état du plex est ACTIVE).

Si l'état du volume est ENABLED, entrez la commande suivante :

```
# vxplex att nom_volume nom_plex ...
```

Par exemple, la commande suivante s'applique au plex vol01-02 sur le volume vol01 :

```
# vxplex att vol01 vol01-02
```

Cette commande lance la récupération du contenu du plex et, à l'issue de la réactivation, elle attribue au plex l'état ACTIVE.

Si le volume n'est pas en cours d'utilisation (son état n'est pas ENABLED), entrez la commande suivante :

```
# vxmend on nom_plex
```

Par exemple, la commande suivante s'applique au plex vol01-02 :

```
# vxmend on vol01-02
```

Dans ce cas, l'état de vol01-02 est STALE. Au démarrage suivant du volume, les données du plex sont réactivées à partir de l'autre plex et elles sont incorporées dans le volume dont l'état est ACTIVE.

S'il est nécessaire de modifier manuellement l'état d'un plex, voir « Récupération d'un volume » à la page 192. Pour plus d'informations sur ces commandes, voir les pages man vxmake(1M) et vxmend(1M).

Déplacement d'un plex

Le déplacement d'un plex copie le contenu des données du plex d'origine sur un nouveau plex. La commande suivante permet de déplacer les données d'un plex sur un autre :

```
# vxplex mv plex_d'origine nouveau_plex
```

Pour réussir un déplacement, veillez à respecter les critères suivants :

- ◆ l'ancien plex doit être une partie active d'un volume actif (ENABLED) ;
- ◆ la taille du nouveau plex doit être au moins égale, voire supérieure, à celle de l'ancien plex ;
- ◆ Le nouveau plex ne doit pas être associé à un autre volume.

Opérations de sous-disque

La taille du plex a plusieurs implications :

- ◆ Si le nouveau plex est plus petit ou plus clairsemé que le plex d'origine, il sera créé une copie incomplète des données de ce dernier. S'il s'agit du résultat voulu, vous devez utiliser l'option `-o force`.
- ◆ Si le nouveau plex est plus long ou moins clairsemé que le plex d'origine, les données de ce dernier seront copiées sur le nouveau plex. Les zones qui n'existaient pas sur le plex d'origine, mais qui sont représentées sur le nouveau plex, sont remplies à partir des autres plex complets associés au même volume.
- ◆ Si le nouveau plex est plus long que le volume, la zone restante au-delà de la taille du volume n'est pas initialisée et reste inutilisée.

Copie d'un plex

Cette opération copie le contenu d'un volume sur un plex spécifié. Le volume à copier ne peut pas être activé. Le plex ne doit pas être associé à un autre volume. Pour copier un plex, entrez la commande suivante :

```
# vxplex cp nom_volume nouveau_plex
```

Lorsque la copie est terminée, *nouveau_plex* n'est pas associé au volume *nom_volume* spécifié. Le plex contient une copie complète des données du volume. Le plex en cours de copie doit avoir une taille égale ou supérieure à celle du volume. Si sa taille est inférieure à celle du volume, une copie incomplète des données est créée. Pour la même raison, *nouveau_plex* ne doit pas être clairsemé.

Opérations de sous-disque

Les sous-disques sont des blocs de construction de bas niveau dans une configuration Volume Manager et il servent à construire les volumes. Les sections suivantes décrivent les tâches que vous pouvez exécuter sur des sous-disques.

Création d'un sous-disque

Remarque Les sous-disques sont automatiquement créés lorsque vous utilisez vxassist ou Storage Administrator pour créer des volumes.

Vous pouvez utiliser la commande vxmake pour créer des objets Volume Manager, tels que les sous-disques. Lorsque vous créez un sous-disque, vous devez spécifier son nom et sa longueur, son point de départ (décalage) dans le disque et le nom d'accès logique au disque.

Pour créer un sous-disque, entrez la commande suivante :

```
# vxmake sd nom_sous-disque disque, décalage, longueur
```

Par exemple, la commande suivante permet de créer le sous-disques `disk02-01` qui commence au début du disque `disk02` et dont la longueur est de 8000 secteurs :

```
# vxmake sd disk02-01 disk02,0,8000
```

Par défaut, les commandes Volume Manager utilisent les tailles en secteurs. Vous pouvez ajouter un suffixe (par exemple, `k`, `m` ou `g`) pour changer l'unité de taille.

Si vous voulez utiliser le nouveau sous-disque pour créer un volume, vous devez associer le sous-disque à un plex (voir « Association d'un sous-disque » à la page 204). Vous devez suivre la même procédure afin de créer les sous-disques pour toutes les structures de plex (concaténés, répartis par bandes, RAID 5).

Suppression d'un sous-disque

Pour supprimer un sous-disque, entrez la commande suivante :

```
# vxedit rm nom_sous-disque
```

Par exemple, vous pouvez utiliser la commande suivante pour supprimer le sous-disque `disk02-01` :

```
# vxedit rm disk02-01
```

Affichage d'informations sur les sous-disques

La commande `vxprint` affiche des informations sur les objets Volume Manager. Pour afficher des informations générales sur tous les sous-disques, entrez la commande suivante :

```
# vxprint -st
```

L'option `-s` spécifie des informations sur les sous-disques. L'option `-t` affiche un enregistrement de sortie d'une seule ligne qui dépend du type de l'objet à lister.

La commande suivante vous permet d'afficher des informations complètes sur un sous-disque spécifique :

```
# vxprint -l nom_sous-disque
```

Par exemple, la commande suivante permet d'obtenir toutes les informations sur le sous-disque `disk02-01` :

```
# vxprint -l disk02-01
```

Opérations de sous-disque

Cette commande affiche les informations suivantes :

```
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
```

Association d'un sous-disque

L'association d'un sous-disque à un plex place la quantité d'espace disque définie par le sous-disque à un décalage spécifique dans le plex. La zone remplie par le sous-disque ne doit pas être entièrement occupée par une partie d'un autre sous-disque. Les sous-disques peuvent être associés aux plex de différentes façons, selon l'état général de la configuration.

Si vous avez déjà créé tous les sous-disques nécessaires pour un plex donné, vous pouvez associer des sous-disques à la création des plex en entrant la commande suivante :

```
# vxmake plex nom_plex sd=nom_sous-disque,...
```

Par exemple, la commande suivante crée le plex `home-1` et lui associe les sous-disques `disk02-01`, `disk02-00` et `disk02-02`.

```
# vxmake plex home-1 sd=disk02-01,disk02-00,disk02-02
```

Les sous-disques sont associés dans l'ordre indiqué, à partir du décalage 0. L'espace disque défini comme `disk02-01` occupe la première position, l'espace disque `disk02-00` la seconde, et `disk02-02` la troisième. Si vous utilisez ce type de commande, il n'est pas nécessaire de spécifier les différentes commandes nécessaires à la création du plex, puis d'associer à ce dernier chaque sous-disque. Cette méthode d'association de sous-disques est pratique pendant la configuration initiale.

Vous pouvez aussi associer les sous-disques à un plex existant. La commande suivante permet d'associer un ou plusieurs sous-disques à un plex existant :

```
# vxsd assoc nom_plex nom_sous-disque [nom_sous-disque2
nom_sous-disque3 ...]
```

Par exemple, vous pouvez utiliser la commande suivante pour associer les sous-disques `disk02-01`, `disk02-00` et `disk02-02` au plex `home-1` :

```
# vxsd assoc home-1 disk02-01 disk02-00 disk02-01
```

Si le plex n'est pas vide, les nouveaux sous-disques sont ajoutés après les sous-disques déjà associés au plex, sauf si vous spécifiez l'option `-1` avec la commande. Cette option associe les sous-disques à un décalage spécifique dans le plex.

L'option `-1` est nécessaire si vous avez créé un plex clairsemé (c'est-à-dire, un plex dont les sous-disques sont séparés par un espace) pour un volume particulier et que vous voulez convertir ce plex en un plex complet. Pour cela, vous devez créer un sous-disque dont la taille correspond exactement à celle de l'espace dans le plex clairsemé. Vous devez ensuite associer le sous-disque au plex en spécifiant le décalage du début de l'espace dans le plex. Pour cela, entrez la commande suivante :

```
# vxsd -1 décalage assoc nom_plex_clairsemé taille_exacte_sous-disque
```

Remarque La taille du sous-disque doit être correcte, car Volume Manager n'autorise pas le chevauchement de l'espace défini par deux sous-disques avec un seul plex.

Pour les sous-disques répartis par bandes, vous pouvez spécifier un numéro et un décalage de colonne :

```
# vxsd -1 n°_colonne/décalage assoc nom_plex nom_sous-disque ...
```

Si un seul nombre est spécifié avec l'option `-1` pour les plex répartis par bandes, il est interprété comme étant un numéro de colonne et le sous-disque est associé à la fin de cette colonne.

Association d'un sous-disque de journal

Les *sous-disques de journal* sont ajoutés à un plex qui est destiné à faire partie d'un volume à l'aide du journal des zones modifiées. Ce journal est activé pour un volume lorsque ce dernier est mis en miroir et a au moins un sous-disque de journal.

Pour plus d'informations sur le journal des zones modifiées, voir « Journal des zones modifiées (DRL) » à la page 73. Les sous-disques de journal sont ignorés dans les stratégies de plex courantes et ils servent uniquement à contenir le journal des zones modifiées.

Remarque Vous ne pouvez associer qu'un seul sous-disque de journal à un plex. Vous devez placer le sous-disque de journal (qui a enregistré de nombreuses écritures) sur un disque peu utilisé, faute de quoi les performances du système risquent d'être diminuées.

Vous pouvez utiliser la commande suivante pour ajouter un sous-disque de journal à un volume existant :

```
# vxassist addlog nom_volume disque
```

Cette commande crée automatiquement un sous-disque de journal dans un plex de journal pour le volume spécifié.

Opérations de sous-disque

Pour ajouter un sous-disque de journal à un plex existant, entrez la commande suivante :

```
# vxsd aslog plex sous-disque
```

où *sous-disque* représente le nom du sous-disque de journal. Le plex doit être associé à un volume mis en miroir pour que le journal des zones modifiées soit actif.

Par exemple, vous pouvez utiliser la commande suivante pour associer le sous-disque `disk02-01` au plex `vol101-02` (qui est déjà associé au volume `vol101`) :

```
# vxsd aslog vol101-02 disk02-01
```

Dissociation d'un sous-disque

Pour supprimer la connexion établie entre un sous-disque et son plex, il est nécessaire que le sous-disque soit *dissocié* du plex. Cela se produit lorsque le sous-disque est supprimé ou utilisé dans un autre plex. Pour dissocier un sous-disque, entrez la commande suivante :

```
# vxsd dis nom_sous-disque
```

Par exemple, vous pouvez dissocier le sous-disque `disk02-01` du plex auquel il est actuellement associé en entrant la commande suivante :

```
# vxsd dis disk02-01
```

Remarque Vous pouvez également entrer la commande suivante pour supprimer des sous-disques :

```
# vxsd -orm dis nom_sous-disque
```

Modification des attributs d'un sous-disque

Attention Exercez la plus grande précaution lorsque vous modifiez les attributs des plex, et ne le faites que si cela est nécessaire.

La commande `vxedit` modifie les attributs des sous-disques en d'autres objets Volume Manager. Pour modifier les informations sur un sous-disque, entrez la commande suivante :

```
# vxedit set champ=valeur ... nom_sous-disque
```

Par exemple, la commande suivante permet de modifier le champ `comment` du sous-disque `disk02-01` :

```
# vxedit set comment="nouveau_commentaire" disk02-01
```

La commande `vxedit` permet de modifier les champs de sous-disques suivants :

- ◆ `nom`
- ◆ champs `putil[n]`
- ◆ champs `tutil[n]`
- ◆ `len` (uniquement si le sous-disque est dissocié)
- ◆ `comment`

Remarque Si vous entrez des données dans le champ `putil0`, le sous-disque ne pourra pas être utilisé en tant que partie d'un plex, si ce n'est déjà le cas.

Déplacement d'un sous-disque

Le déplacement d'un sous-disque copie le contenu de son espace disque sur un autre sous-disque. Si le sous-disque en cours de déplacement est associé à un plex, ses données seront copiées sur le nouveau sous-disque. L'ancien sous-disque est dissocié du plex et le nouveau sous-disque est associé au plex. L'association s'effectue sur le même décalage dans le plex que dans le sous-disque source. Pour déplacer un sous-disque, entrez la commande suivante :

```
# vxsd mv ancien_nom_sous-disque nouveau_nom_sous-disque
```

Pour réussir le déplacement d'un sous-disque, vous devez respecter les conditions suivantes :

- ◆ les sous-disques inclus dans l'opération doivent avoir la même taille ;
- ◆ le sous-disque en cours de déplacement doit appartenir à un plex actif sur un volume actif (ENABLED) ;
- ◆ le nouveau sous-disque ne doit pas être associé à un autre plex.

Division d'un sous-disque

Cette opération divise un sous-disque existant en deux sous-disques. Pour diviser un sous-disque, entrez la commande suivante :

```
# vxsd -s taille split nom_sous-disque nouveau_sd1 nouveau_sd2
```

où :

- ◆ *nom_sous-disque* représente le nom du sous-disque d'origine.
- ◆ *nouveau_sd1* représente le nom du premier des deux sous-disques à créer.
- ◆ *nouveau_sd2* représente le nom du second sous-disque à créer.

Exécution d'une sauvegarde en ligne

L'option `-s` est nécessaire pour spécifier la taille du *premier* sous-disques à créer. Le second sous-disques occupe l'espace restant utilisé par le sous-disque d'origine.

Si le sous-disque d'origine est associé à un plex avant l'exécution de l'opération, les deux sous-disques résultants sont associés au même plex à l'issue de la division.

Pour diviser le sous-disque d'origine en plus de deux sous-disques, répétez la commande précédente autant de fois que nécessaire sur les sous-disques résultants.

Jonction de sous-disques

Cette opération combine en un seul sous-disque au moins deux sous-disques existants. Les sous-disques à joindre doivent être contigus sur le même disque. Si les sous-disques sélectionnés sont associés, ils doivent l'être au même plex et ils doivent être également contigus dans ce plex. Pour joindre des sous-disques, entrez la commande suivante :

```
# vxsd join sous-disque1 sous-disque2 nouveau_sous-disque
```

Exécution d'une sauvegarde en ligne

Volume Manager permet d'effectuer des sauvegardes de prise de vue des périphériques de volume. Pour cela, vous pouvez utiliser `vxassist` et d'autres commandes. Les sauvegardes peuvent être effectuées de plusieurs façons, en fonction des spécifications sur l'intégrité du contenu du volume. Ces procédures reposent sur la même spécification de départ : un plex d'une taille suffisante pour contenir tout le contenu du volume. Sa taille peut être supérieure à celle nécessaire ; en revanche, si vous utilisez un plex trop petit, la copie sera incomplète.

L'approche recommandée pour la sauvegarde du volume consiste à utiliser la commande `vxassist`, facile à utiliser. Les opérations `vxassist snapstart`, `snapwait` et `snapshot` permettent d'exécuter la sauvegarde en ligne des données en garantissant une interruption minimale des activités.

La procédure `vxassist snapshot` est constituée de deux étapes :

1. l'exécution de `vxassist snapstart` pour créer un miroir de prise de vue ;
2. l'exécution de `vxassist snapshot` pour créer un volume de prise de vue.

Remarque Vous pouvez utiliser la commande `vxassist` pour créer une prise de vue d'un volume RAID 5 pour la sauvegarde des volumes en suivant l'approche recommandée dans cette section.

L'opération `vxassist snapstart` crée un plex de sauvegarde accessible en écriture seule, qui est attaché au volume et synchronisé avec ce dernier. Lorsque la synchronisation est terminée, le plex de sauvegarde est prêt pour être utilisé en tant que

miroir de prise de vue. La fin de la procédure de mise à jour est indiquée par le changement d'état du nouveau miroir de prise de vue (son état devient SNAPDONE). Vous pouvez suivre le changement à l'aide de la tâche `vxassist snapwait task`, qui attend que l'état d'au moins un des miroirs devienne SNAPDONE. Si la procédure d'attachement échoue, le miroir de prise de vue est supprimé et l'espace qu'il occupait est libéré.

Une fois synchronisé, le miroir de prise de vue continue d'être mis à jour jusqu'à ce qu'il soit détaché. Vous pouvez alors choisir le moment voulu pour créer un volume de prise de vue qui constituera l'image du volume existant. Vous pouvez également demander aux utilisateurs de ne pas utiliser le système pendant la courte durée requise pour effectuer la prise de vue (généralement moins d'une minute). La durée nécessaire pour créer le miroir de prise de vue est assez longue par rapport à celle requise pour créer le volume de prise de vue.

Pour terminer la procédure de sauvegarde en ligne, exécutez la commande `vxassist snapshot` sur un volume dont l'état du miroir est SNAPDONE. Cette tâche détache la prise de vue terminée (qui devient un miroir normal), crée un nouveau volume normal et attache le miroir de prise de vue au volume de prise de vue. La prise de vue devient alors un miroir opérationnel normal et son état devient ACTIVE.

Si la procédure de prise de vue est interrompue, le miroir de prise de vue est automatiquement supprimé au démarrage du volume.

Procédez comme suit pour exécuter une sauvegarde `vxassist` complète :

1. Créez un miroir de prise de vue pour un volume à l'aide de la commande suivante :

```
# vxassist snapstart nom_volume
```
2. À l'issue de l'étape `snapstart`, lorsque l'état du miroir est SNAPDONE, choisissez le moment voulu pour terminer la tâche de prise de vue. Informez les utilisateurs de la prise de vue en cours et demandez-leur d'enregistrer leurs fichiers et de ne pas utiliser le système pendant la courte durée de l'opération.
3. Créez un volume de prise de vue qui reflète le volume d'origine à l'aide de la commande suivante :

```
# vxassist snapshot nom_volume nom_volume_temp
```
4. Utilisez `fsck` (ou un autre utilitaire adapté à l'application exécutée sur le volume) pour nettoyer le contenu du volume temporaire. Par exemple, vous pouvez utiliser la commande suivante :

```
# fsck -y /dev/vx/rdisk/nom_volume_temp
```

Exécution d'une sauvegarde en ligne

5. Copiez le volume temporaire sur une bande ou sur un autre support de sauvegarde approprié.
6. Supprimez le nouveau volume en entrant la commande suivante :

```
# vxedit -rf rm nom_volume_temp
```

Resynchronisation rapide des miroirs

La fonction de resynchronisation rapide des miroirs (FMR, *Fast Mirror Resynchronization*) exécute une resynchronisation rapide et efficace des miroirs obsolètes en augmentant l'efficacité de la prise de vue de VxVM pour garantir une meilleure prise en charge des opérations telles que la sauvegarde et la prise de décision.

Activation de la resynchronisation rapide des miroirs

Lorsque vous créez un nouveau volume au moyen de `vxassist`, vous pouvez spécifier un attribut pour activer ou désactiver la resynchronisation rapide des miroirs. Vous pouvez utiliser les deux mots-clé `fmr` et `fastresync` comme attributs afin de spécifier que la resynchronisation rapide des miroirs doit (ou non) être utilisée sur un volume.

Pour créer un volume lorsque la resynchronisation rapide des miroirs est activée, utilisez la commande `vxassist make` comme suit :

```
# vxassist make nom_volume taille fmr=on
```

Par défaut, la resynchronisation rapide des miroirs est désactivée (`off`), mais vous pouvez modifier cette valeur par défaut dans le fichier par défaut `vxassist`.

Vous pouvez aussi utiliser la commande `vxvol` pour activer (`on`) ou désactiver (`off`) la resynchronisation rapide des miroirs. Pour pouvoir utiliser cette fonction, elle doit être activée au moment de la prise de vue et elle doit le rester jusqu'à la fin de l'opération `snapback`. Si vous la désactivez, vous libérez toutes les mappes de suivi pour le volume spécifié. Les rattachements successifs n'utiliseront pas la fonction FMR, mais ils effectueront une resynchronisation complète du volume. Cela se produit même si vous activez par la suite la resynchronisation rapide des miroirs.

Pour activer la resynchronisation rapide des miroirs, entrez la commande suivante :

```
# vxvol set fmr=on nom_volume
```

Pour désactiver la resynchronisation rapide des miroirs, entrez la commande suivante :

```
# vxvol set fmr=off nom_volume
```

Fusion d'un volume de prise de vue

Vous pouvez fusionner avec le volume d'origine une copie de prise de vue de ce volume. Le plex de prise de vue est détaché du volume de prise de vue et il est attaché au volume d'origine. Le volume de prise de vue est supprimé. Cette opération resynchronise les données du volume de façon à garantir la cohérence des plex.

Pour fusionner une prise de vue avec son volume d'origine, entrez la commande suivante :

```
# vxassist snapback volume-replica
```

où *volume-replica* représente la copie de prise de vue du volume.

Par défaut, les données du plex d'origine sont utilisées pour le volume fusionné. Pour utiliser à la place la copie des données à partir du volume réplica, entrez la commande suivante :

```
# vxassist -o resyncfromreplica snapback volume-replica
```

Dissociation d'un volume de prise de vue

Vous pouvez supprimer définitivement le lien entre une prise de vue et son volume d'origine afin que le volume de prise de vue devienne indépendant.

Pour dissocier une prise de vue de son volume d'origine, entrez la commande suivante :

```
# vxassist snapclear volume-replica
```

où *volume-replica* est la copie de prise de vue du volume.

Affichage des informations sur le volume de prise de vue

La commande `vxassist snapprint` affiche les associations entre les volumes d'origine à leurs réplicas respectifs (copies de prise de vue).

La syntaxe de l'option `snapprint` est la suivante :

```
# vxassist snapprint [nom_volume]
```

Cette commande affiche les informations suivantes :

```
V NOM TYPEUTIL LONGUEUR RP NOM VOLUME LONGUEUR RRPLEXID
v vol fsgen 2048 rp vol-05 SNAP1-vol 3040 rp vol-04
```

Si vous spécifiez un volume, la commande affiche soit les informations propres à ce volume, soit un message d'erreur si aucune mappe FMR n'est activée pour ce volume. Dans le cas contraire, la commande affiche des informations sur tous les volumes du disk group.

Mise en miroir d'un volume sur un disque VM

Cette opération crée une ou plusieurs copies des volumes sur un autre disque. La création des copies miroirs des volumes vous permet de protéger votre système contre le risque de perte de données lié à une défaillance de disque. Vous pouvez utiliser cette opération sur le disque racine pour créer une seconde copie des informations de démarrage sur un autre disque. Vous pourrez ainsi démarrer le système, même si le disque racine est endommagé.

Remarque Cette opération ne met en miroir que des volumes concaténés. Les volumes qui sont déjà en miroir ou qui contiennent des sous-disques appartenant à plusieurs disques sont ignorés.

Pour mettre en miroir les volumes sur un disque, vous devez vous assurer que l'espace du disque cible est égal ou supérieur à celui du disque d'origine, puis procédez comme suit :

1. Sélectionnez l'option de menu 6 (Mettre en miroir des volumes sur un disque) dans le menu principal de vxdiskadm.
2. À l'invite suivante, entrez le nom du disque que vous voulez mettre en miroir :

```
Mettre en miroir des volumes sur un disque
Menu : VolumeManager/Disk/Mirror
```

```
Cette opération permet de mettre en miroir les volumes sur
un disque. Ces volumes peuvent être mis en miroir sur un
autre disque ou dans de l'espace disque disponible.
Les volumes ne seront pas mis en miroir s'ils le sont déjà
ou s'ils comprennent plusieurs sous-disques.
```

```
La mise en miroir des volumes à partir du disque
d'initialisation crée un disque qui peut être utilisé comme
disque de démarrage de remplacement.
```

```
Entrez le nom du disque [<disk>,list,q,?] disk02
```

3. À l'invite suivante, entrez le nom du disque cible (sa taille doit être supérieure ou égale à celle du disque d'origine) :

```
Vous pouvez mettre en miroir les volumes du disque
disk02 dans n'importe quel espace de disque disponible ou
dans un disque spécifique. Dans ce cas, sélectionnez le nom
de ce disque. Pour effectuer la mise en miroir dans de
l'espace de disque disponible, sélectionnez "any".
```

```
Enter destination disk [<disk>,list,q,?] (default: any) disk01
```

Remarque Veillez à toujours spécifier le disque de destination lorsque vous créez un autre disque racine, faute de quoi Volume Manager sélectionne le disque qui sera l'autre disque d'initialisation. Cependant, il n'est pas toujours possible de démarrer le système à partir de ce disque.

4. À l'invite suivante, appuyez sur la touche de retour pour créer le miroir :

```
L'opération demandée consiste à mettre en miroir tous les
volumes du disque disk02 du disk group rootdg dans de
l'espace disque disponible sur le disque disk01.
```

```
REMARQUE : cette opération peut prendre un certain temps.
```

```
Continuer l'opération ? [y,n,q,?] (default: y)
```

vxdiskadm affiche l'état de l'opération de mise en miroir :

```
Mirror volume voltest-bk00 ...
```

```
La mise en miroir du disque disk01 est terminée.
```

5. À l'invite suivante, indiquez si vous voulez mettre en miroir les volumes d'un autre disque (**y**) ou revenir au menu principal de vxdiskadm (**n**) :

```
Mettre en miroir des volumes sur un autre disque ?
[y,n,q,?] (default: n)
```

Déplacement d'un volume à partir d'un disque VM

Avant de déplacer ou de supprimer un disque, vous pouvez déplacer les données de ce disque vers d'autres disques du système. Pour cela, assurez-vous que les disques cibles disposent d'un espace suffisant, puis procédez comme suit :

1. Sélectionnez l'option de menu 7 (Déplacer les volumes d'un disque) dans le menu principal de vxdiskadm.
2. À l'invite suivante, entrez le nom du disque dont vous voulez déplacer les volumes :

```
Déplacer les volumes d'un disque
Menu : VolumeManager/Disk/Evacuate
```

```
Utilisez cette opération du menu pour déplacer n'importe quel
volume qui utilise un disque vers d'autres disques.
Utilisez ce menu immédiatement avant de supprimer un disque
de manière permanente ou pour le remplacer. Vous pouvez
```

Exécution d'une sauvegarde en ligne

spécifier la liste des disques vers lesquels déplacer les volumes ou déplacer ces derniers vers n'importe quel espace disponible dans le même disk group.

REMARQUE : si les volumes sont supprimés d'un disque, sans que celui-ci le soit, ils peuvent être déplacés vers le disque par la suite. Par exemple, l'exécution de deux opérations de déplacement consécutives peut déplacer les volumes du second disque vers le premier.

Entrez le nom du disque [`<disk>,list,q,?`] `disk01`

Après l'affichage suivant, vous pouvez éventuellement spécifier la liste des disques vers lesquels déplacer le ou les volumes.

Vous pouvez spécifier maintenant la liste des disques vers lesquels effectuer le déplacement. Spécifiez la liste des noms d'accès logiques aux disques (p. ex., `disk01`) sur une seule ligne, séparés par des espaces. Si vous n'entrez pas de noms d'accès logique aux disques, les volumes seront déplacés dans n'importe quel espace disponible du disk group.

À l'invite suivante, appuyez sur la touche de retour pour déplacer les volumes :

L'opération demandée consiste à placer tous les volumes du disque `disk01` dans le groupe `rootdg`.

REMARQUE : cette opération peut prendre un certain temps.

Continuer l'opération ? [`y,n,q,?`] (default: `y`)

Au fur et à mesure que les volumes sont déplacés du disque, `vxdiskadm` affiche l'état de l'opération :

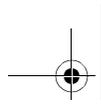
Déplacer le volume `voltest` ...
Déplacer le volume `voltest-bk00` ...

Lorsque tous les volumes ont été déplacés, `vxdiskadm` affiche le message de réussite suivant :

Le déplacement du contenu du disque `disk01` est terminé.

3. À l'invite suivante, indiquez si vous voulez déplacer les volumes d'un autre disque (y) ou revenir au menu principal de vxdiskadm (n) :

```
Déplacer les volumes d'un autre disque ?  
[y,n,q,?] (default: n)
```



Exécution d'une sauvegarde en ligne



Fonction de cluster de Volume Manager

6

Introduction

Ce chapitre présente la fonction de cluster qui accompagne VERITAS Volume Manager (VxVM). Volume Manager comprend une fonction de cluster facultative, qui permet d'utiliser VxVM dans un environnement cluster. Cette fonctionnalité fait l'objet d'une licence séparée.

Ce chapitre traite des sujets suivants :

- ◆ Présentation de la fonction de cluster
- ◆ Disques dans les clusters VxVM
- ◆ Journal des zones modifiées et environnements cluster
- ◆ Mise à niveau de la fonction de cluster de Volume Manager
- ◆ Utilitaires et démons Volume Manager relatifs au cluster
- ◆ Gestion des clusters et resynchronisation rapide des miroirs

Pour plus d'informations sur les messages d'erreur relatifs aux clusters, reportez-vous au chapitre traitant des messages d'erreur dans le *Guide de référence de VERITAS Volume Manager*.

Présentation de la fonction de cluster

La fonction de cluster de Volume Manager permet à plusieurs hôtes d'accéder à un jeu de disques donné contrôlé par Volume Manager (*disques VM*) et de le gérer simultanément. Un *cluster* est un ensemble d'hôtes qui partagent un jeu de disques ; chaque hôte est un *nœud* dans le cluster. Les nœuds sont connectés sur un réseau. En cas de défaillance de l'un de ces nœuds, les autres peuvent continuer d'accéder aux disques. La fonction de cluster de Volume Manager présente la même vue logique des configurations de disque (y compris les modifications) sur tous les nœuds.

Remarque Lorsque les clusters sont pris en charge, Volume Manager peut gérer un maximum de quatre nœuds par cluster.

Présentation de la fonction de cluster

Les sections suivantes contiennent davantage d'informations sur la fonction de cluster de Volume Manager.

Objets Volume Manager partagés

Lorsque la fonction de cluster est activée, tous les nœuds d'un cluster donné peuvent partager les objets Volume Manager.

La fonction de cluster de Volume Manager permet d'utiliser deux types de disk groups :

- ◆ *Les disk groups privés*, qui appartiennent à un seul nœud. Un disk group privé est uniquement importé par un système. Les disques d'un disk group privé peuvent être accessibles physiquement à partir d'un ou de plusieurs systèmes, mais l'accès réel est limité à un seul système.
- ◆ *Les disk groups partageables par le cluster*, qui sont partagés par tous les nœuds. Un disk group partageable par le cluster (ou *partagé*) est importé par tous les nœuds du cluster. Les disques d'un disk group partageable par le cluster doivent être accessibles physiquement à partir de tous les systèmes susceptibles de joindre le cluster.

Dans un cluster Volume Manager, les disk groups sont le plus souvent partagés. Cependant, le disk group racine (`rootdg`) est toujours un disk group privé.

Les disques d'un disk group partagé sont accessibles à partir de tous les nœuds d'un cluster, ce qui permet aux applications installées sur plusieurs nœuds de cluster d'accéder simultanément au même disque. Plusieurs nœuds du cluster peuvent accéder simultanément au volume d'un disk group partagé, en fonction des descriptions de licence et de mode d'activation des disk groups.

Un disk group partagé doit être activé sur un nœud pour que les E/S d'applications puissent accéder à ses volumes à partir de ce nœud. Le mode d'activation de ce disk group détermine la capacité des applications à lire ou à écrire sur les volumes. Les modes d'activation valides pour un disk group partagé sont *Écriture seule*, *Écriture partagée*, *Lecture seule*, *Lecture partagée* et *Désactivé* (ou inactif), comme indiqué dans le tableau 5, « Modes d'activation pour le disk group partagé ».

Tableau 5. Modes d'activation pour le disk group partagé

Écriture seule	Le nœud peut accéder au disk group en écriture seule. Aucun autre nœud ne peut activer le dg pour l'accès en écriture.
Écriture partagée	Le nœud peut accéder au disk group en écriture.
Lecture seule	Le nœud peut accéder au disk group en lecture et refuse l'accès en écriture à tous les autres nœuds du cluster. Il ne peut pas accéder au disk group en écriture. Les tentatives d'activation d'un disk group pour l'un des modes d'écriture sur les autres nœuds échoueront.
Lecture partagée	Le nœud peut accéder au disk group en lecture. Il ne peut pas y accéder en écriture, mais les autres nœuds peuvent avoir un accès en écriture.

Tableau 5. Modes d'activation pour le disk group partagé (suite)

Désactivé	Le nœud ne peut pas accéder au disk group en lecture ni en écriture. Les opérations d'interrogation sur le disk group sont autorisées.
------------------	--

Remarque L'activation des disk groups était une nouvelle fonction de VxVM 3.0. Pour maintenir la compatibilité avec les versions précédentes, les modes d'activation sont par défaut transparents aux utilitaires Volume Manager. Les disk groups partagés sont automatiquement activés en mode Écriture partagée.

L'utilisation des clusters à des fins spéciales, par exemple pour les applications HA et les sauvegardes sur un système différent de l'hôte, peut activer des disk groups pour contrôler explicitement les fonctions d'E/S des volumes à partir des différents nœuds du cluster. L'utilisation des modes d'activation est décrite à la section « Activation des disk groups » à la page 226.

Remarques :

- ◆ Les nouvelles fonctions de Volume Manager 3.0 (*volumes mis en miroir répartis par bande*, *moniteur de tâches* et *restructuration en ligne*) sont disponibles pour les disk groups privés, mais pas encore pour les disk groups partagés.
- ◆ La fonction de cluster de Volume Manager gère uniquement l'accès aux périphériques en mode raw. Elle ne prend pas en charge les volumes partagés avec les systèmes de fichiers.
- ◆ La fonction de cluster de Volume Manager ne gère pas actuellement les volumes RAID 5 des disk groups partageables par le cluster. En revanche, vous pouvez utiliser les volumes RAID 5 dans des disk groups privés attachés aux nœuds spécifiques d'un cluster.
- ◆ Si un disk group contenant des objets non pris en charge est importé en tant que disk group partagé, exportez-le. Réorganisez les volumes inclus afin d'obtenir des configurations prises en charge, puis réimportez le disk group en tant que disk group partagé.

Utilisation de la fonction de cluster de Volume Manager

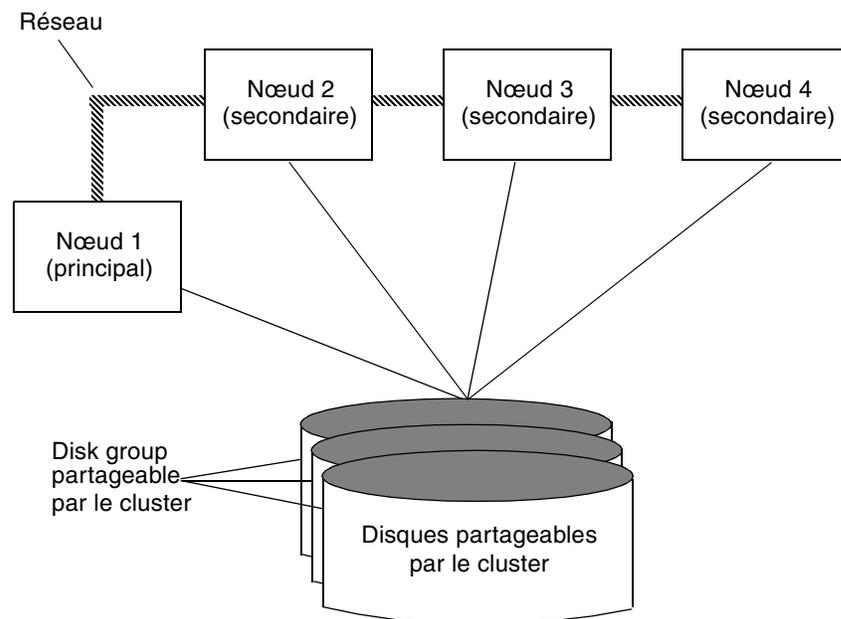
La fonction de cluster de Volume Manager fonctionne avec un *gestionnaire de clusters* externe, c'est-à-dire, un démon qui informe VxVM des modifications apportées aux composants du cluster. Chaque nœud démarre indépendamment et a ses propres copies du système d'exploitation, de VxVM avec prise en charge du cluster et du gestionnaire de clusters. Lorsqu'un nœud *joint* un cluster, il peut accéder aux disques partagés. Lorsqu'il *quitte* un cluster, il ne peut plus y accéder. Pour joindre un nœud à un cluster, l'administrateur système doit démarrer le gestionnaire de clusters sur ce nœud.

La figure 24, « Exemple de cluster à 4 nœuds », à la page 220 illustre une structure de cluster simple. Tous les nœuds sont connectés par un réseau. Ils sont ensuite connectés à un disk group partageable par le cluster. Tous les nœuds sont identiques pour le

Présentation de la fonction de cluster

gestionnaire de clusters. Cependant, la fonction de cluster de Volume Manager nécessite que l'un des nœuds soit le *nœud principal* ; les autres nœuds sont les *nœuds secondaires*. Le nœud principal assure la coordination de certaines des activités de Volume Manager. Le logiciel VxVM détermine le nœud principal (il peut s'agir de n'importe quel nœud) ; ce rôle change uniquement si le nœud principal quitte le cluster. Dans ce cas, l'un des nœuds secondaires devient le nouveau nœud principal. Dans la figure 24, « Exemple de cluster à 4 nœuds », le nœud 1 est le nœud principal, et les nœuds 2, 3 et 4 sont les nœuds secondaires.

Figure 24. Exemple de cluster à 4 nœuds



L'administrateur système désigne un disk group comme partageable par le cluster à l'aide de l'utilitaire `vxchg` (Pour plus d'informations, voir « `vxchg` » à la page 234). Lorsqu'un disk group est importé comme étant partageable par le cluster pour un nœud, les en-têtes de disque sont marqués par l'ID du cluster. Lorsque d'autres nœuds rejoignent le cluster, ils reconnaissent que le disk group est partageable par le cluster et ils l'importent. L'administrateur système peut importer ou exporter un disk group partagé à tout moment ; cette opération est effectuée de façon distribuée sur tous les nœuds.

Chaque disque physique est marqué par un ID de disque unique. Lorsque le cluster démarre sur le nœud principal, il importe tous les disk groups partagés (sauf s'ils ont le jeu d'attributs `noautoimport`). Lorsqu'un nœud secondaire tente de rejoindre le cluster, le

nœud principal lui envoie la liste des ID de disque qu'il a importés ; le nœud secondaire vérifie alors qu'il peut accéder à tous ces disques. S'il ne peut pas accéder à l'un des disques importés de la liste, il abandonne sa tentative de joindre le cluster. S'il peut accéder à tous les disques de la liste, il importe le même jeu de disk groups partagés que le nœud principal et il joint le cluster. Lorsqu'un nœud quitte le cluster, il exporte tous ses disk groups partagés et importés, qui restent cependant importés sur les autres nœuds.

La reconfiguration vers un disk group partagé s'effectue avec la coopération de tous les nœuds. Les modifications apportées à la configuration du disk group sont reflétées simultanément sur tous les nœuds et sont identiques. Ces modifications sont de nature atomique et se produisent simultanément sur tous les nœuds ou pas du tout.

Tous les membres du cluster ont simultanément accès en lecture et en écriture à n'importe quel disk group partageable par le cluster en fonction du mode d'activation. (Voir « Activation des disk groups » à la page 226.) L'accès par les nœuds actifs du cluster n'est pas affecté par la défaillance d'un autre nœud. Les données d'un disk group partageable par le cluster sont disponibles tant qu'au moins l'un des nœuds est actif sur le cluster. Quel que soit le nœud qui accède au disk group partageable par le cluster, la configuration du disk group semble être identique. Les applications exécutées sur chacun des nœuds peuvent accéder simultanément aux données des disques VM.

Remarque VxVM n'assure aucune protection contre l'écriture simultanée sur les volumes partagés par plusieurs nœuds. Le système considère que les contrôles de cohérence sont effectués au niveau de l'application (par exemple, au moyen d'un gestionnaire de verrouillage distribué).

Configuration et initialisation

Pour que les nœuds puissent joindre un nouveau cluster pour la première fois, l'administrateur système doit d'abord fournir certaines informations pendant la configuration du gestionnaire de clusters. Ces informations sont normalement stockées dans un type spécifique de base de données de configuration du gestionnaire de clusters. Leur contenu et format particulier dépendent des caractéristiques du gestionnaire de clusters. VxVM requiert les informations suivantes :

- ◆ ID de cluster ;
- ◆ ID de nœuds ;
- ◆ adresses réseau des nœuds ;
- ◆ adresses de port.

Lorsqu'un nœud joint le cluster, ces informations sont automatiquement chargées dans VxVM sur ce nœud au moment de son démarrage.

Présentation de la fonction de cluster

L'initialisation des nœuds est activée par le démarrage du gestionnaire de clusters, qui amène les divers composants du cluster (par exemple, VxVM avec prise en charge de cluster, le gestionnaire de clusters et un gestionnaire de verrouillage distribué) sur le nœud. À l'issue de la procédure, les applications peuvent être démarrées. L'administrateur système lance la procédure de démarrage du gestionnaire de clusters sur chacun des nœuds à joindre au cluster.

Pour VxVM dans un environnement cluster, l'initialisation consiste à charger les informations de configuration du cluster et à joindre les nœuds dans le cluster. Le premier nœud qui joint le cluster devient le nœud principal et les autres nœuds (secondaires) rejoignent le nœud principal. Si deux nœuds joignent le cluster en même temps, le logiciel VxVM choisit le nœud principal. Une fois qu'un nœud donné a joint le cluster, il a accès aux disques partagés.

Reconfiguration d'un cluster

Toute modification de l'état d'un cluster (lorsqu'un nœud quitte ou joint le cluster) est suivie d'une reconfiguration du cluster. Le gestionnaire de clusters de chaque nœud surveille les autres nœuds du cluster et informe VxVM de tout changement de la composition du cluster. VxVM prend ensuite les mesures appropriées.

La reconfiguration d'un cluster interrompt les E/S vers les disques partagés. Celles-ci reprennent à l'issue de la reconfiguration. Les applications peuvent alors sembler brièvement figées.

Si d'autres opérations (par exemple, récupérations ou opérations Volume Manager) sont en cours, les reconfigurations de cluster peuvent être retardées jusqu'à ce qu'elles soient terminées. Les reconfigurations de volume (décrites dans la suite de ce chapitre) ne se produisent pas en même temps que les reconfigurations de cluster. Selon les circonstances, vous pouvez suspendre une opération et la relancer par la suite. Généralement, la reconfiguration d'un cluster a priorité sur les autres opérations. Cependant, si la reconfiguration d'un volume en est au stade de la validation, elle sera terminée en premier.

Si vous utilisez le gestionnaire de clusters SunCluster™, l'utilitaire `vxclust` coordonne les reconfigurations de cluster et assure la communication entre VxVM et SunCluster™. L'interaction entre SunCluster™ et `vxclust` assure que chaque étape de la reconfiguration du cluster est effectuée dans l'ordre établi.

Si vous utilisez l'environnement de gestion de cluster VCS, les reconfigurations de cluster sont entièrement gérées dans le noyau VxVM.

Pour plus d'informations sur la reconfiguration d'un cluster, voir « `vxclust` » à la page 231 et « `vxclustadm` » à la page 232.

Reconfiguration d'un volume

La *reconfiguration d'un volume* comprend la procédure de création, de modification et de suppression des objets Volume Manager dans la configuration (par exemple, disk groups, volumes, miroirs, etc.). Dans un cluster, l'exécution de cette procédure nécessite la coopération de tous les nœuds. La reconfiguration d'un volume est distribuée sur tous les nœuds ; les configurations identiques sont modifiées simultanément sur tous les nœuds.

Remarque La reconfiguration d'un volume est lancée et coordonnée par le nœud principal ; par conséquent, l'administrateur système doit exécuter les utilitaires qui nécessitent la modification des objets Volume Manager sur le nœud principal.

Les démons `vxconfigd` jouent un rôle actif dans la reconfiguration d'un volume. Pour réussir la reconfiguration, vous devez exécuter `vxconfigd` sur tous les nœuds.

L'utilitaire installé sur le nœud principal contacte son démon `vxconfigd` local, qui exécute des vérifications locales afin de s'assurer que les modifications demandées sont raisonnables. Par exemple, la création d'un nouveau disk group échouera s'il en existe déjà un du même nom. `vxconfigd` du nœud principal envoie alors des messages contenant les détails des modifications aux démons `vxconfigd` installés sur tous les autres nœuds du cluster. Les démons `vxconfigd` situés sur chaque nœud secondaire effectuent alors leur propre vérification. Par exemple, un nœud secondaire vérifie qu'il ne possède pas un disk group privé du même nom que celui en cours de création ; si l'opération inclut un nouveau disque, tous les nœuds vérifient qu'ils peuvent y accéder. Lorsque les démons `vxconfigd` installés sur tous les nœuds reconnaissent que le changement proposé est raisonnable, ils en avertissent leur noyau, qui valide ou abandonne alors la transaction. Pour pouvoir valider la transaction, les noyaux doivent d'abord s'assurer qu'aucune opération d'E/S n'est en cours. Le nœud principal doit lancer la reconfiguration et coordonner la validation de la transaction.

Si le démon `vxconfigd` installé sur l'un des nœuds quitte le cluster pendant une reconfiguration, les autres nœuds en sont informés et l'opération échoue. Si un nœud quitte le cluster, l'opération échoue, sauf si le nœud principal l'a déjà validée. Si le nœud principal quitte le cluster, le nouveau nœud principal (qui était jusqu'à présent un nœud secondaire) termine ou abandonne l'opération, suivant que le nœud principal précédent lui a envoyé une notification l'informant de la réussite de l'opération. Si le nouveau nœud principal n'a pas reçu cette notification, les nœuds secondaires ne l'ont pas également reçue.

Si un nœud tente de rejoindre le cluster pendant que la reconfiguration d'un volume est en cours, le résultat de l'opération dépend du niveau atteint par la reconfiguration. Si le noyau n'est pas déjà actif, la reconfiguration du volume est interrompue et elle redémarrera une fois que le nœud aura joint le cluster. Dans le cas contraire, la reconfiguration se termine avant que le nœud ne puisse rejoindre le cluster.

Présentation de la fonction de cluster

Si une erreur se produit (par exemple, la vérification d'un nœud secondaire échoue ou un nœud quitte le cluster), elle est retournée à l'utilitaire et un message est envoyé à la console du nœud principal afin d'identifier le nœud sur lequel l'erreur est survenue.

Arrêt d'un nœud

L'administrateur système peut arrêter le cluster sur un nœud donné en appelant la procédure d'arrêt du gestionnaire de clusters sur ce nœud. Cette procédure arrête les composants du cluster lorsque ses applications sont arrêtées. VxVM gère l'*arrêt normal des nœuds* ; c'est-à-dire, le nœud peut quitter le cluster lorsque tous les accès aux volumes partagés ont cessé. L'hôte reste opérationnel, mais les applications de cluster ne peuvent pas être exécutées.

La fonction de cluster de Volume Manager gère des informations d'état globales pour chaque volume. VxVM peut ainsi déterminer avec précision les volumes à récupérer en cas d'arrêt anormal d'un nœud. Lorsqu'un nœud quitte le cluster à cause d'un arrêt anormal ou d'une autre façon non correcte, VxVM détermine les volumes sur lesquels l'écriture n'est pas terminée et le nœud principal les resynchronise. Si le journal des zones modifiées (DRL, *Dirty Region Logging*) est actif sur l'un de ces volumes, il sera utilisé.

L'arrêt normal des nœuds doit être utilisé après ou pendant une procédure destinée à arrêter toutes les applications de cluster. Suivant les caractéristiques de l'application en cluster et sa procédure d'arrêt, l'arrêt du nœud peut prendre un certain temps (de quelques minutes à plusieurs heures). Par exemple, de nombreuses applications utilisent le concept de « vidage », selon lequel elles n'acceptent aucune tâche nouvelle, mais exécutent tous les travaux en cours avant de quitter. Cette procédure peut prendre un certain temps si par exemple, une transaction de longue durée est active.

Lorsque vous appelez la procédure d'arrêt de VxVM, ce dernier vérifie tous les volumes sur tous les disk groups partagés du nœud en cours d'arrêt, puis il continue ou abandonne l'arrêt :

- ◆ Si tous les volumes des disk groups partagés sont fermés, VxVM en interdit l'accès aux applications. Le programme n'exécute pas de resynchronisation car les nœuds ont détecté que ces volumes sont fermés sur le nœud qui sort du cluster.
- ◆ Si aucun volume d'un disk group partagé n'est ouvert, la procédure d'arrêt VxVM échoue. Vous pouvez la relancer jusqu'à ce qu'elle réussisse. Cette opération n'inclut pas la vérification des délais — elle est conçue comme un service qui vérifie que les applications en clusters ne sont plus actives.

Remarque Lorsque le nœud est arrêté, il quitte le cluster. Vous devez attendre qu'il rejoigne de nouveau le cluster pour accéder aux volumes partagés.

Puisque l'arrêt peut prendre un certain temps, d'autres reconfigurations peuvent se produire au cours de la procédure d'arrêt. Généralement, la tentative d'arrêt est suspendue jusqu'à ce que l'autre reconfiguration soit terminée. Cependant, si l'arrêt est déjà trop avancé, il sera exécuté en premier.

Abandon d'un nœud

Il se peut qu'un nœud ne quitte pas le cluster normalement à cause de la défaillance de l'hôte ou parce qu'un composant du cluster a forcé le nœud à quitter le cluster d'urgence. La reconfiguration de cluster qui s'ensuit appelle l'abandon de VxVM. Cette fonction tente d'arrêter en même temps tous les accès aux volumes partagés, bien que la procédure doive d'abord attendre la fin des opérations d'E/S sur le disque.

Les opérations d'E/S qui n'ont pas encore commencé échouent et les volumes partagés sont supprimés. Les applications qui accédaient aux volumes partagés échoueront également.

Après l'abandon ou l'arrêt anormal d'un nœud, les volumes partagés doivent être récupérés (soit par un nœud restant, soit par un redémarrage successif du cluster) car il existe probablement des miroirs non synchronisés.

Arrêt d'un cluster

Lorsque tous les nœuds quittent un cluster, vous devez déterminer au démarrage suivant du cluster si les volumes partagés doivent être récupérés. Si tous les nœuds ont quitté le cluster normalement, la récupération est inutile. Elle est également inutile si le dernier nœud a quitté le cluster normalement et que la resynchronisation résultant de l'abandon anormal des nœuds du cluster est terminée. Toutefois, vous devez effectuer une récupération si le dernier nœud n'a pas quitté le cluster normalement ou si la resynchronisation a échoué.

Disques dans les clusters VxVM

Les nœuds d'un cluster doivent toujours reconnaître l'état d'un disque. Notamment, si l'un des nœuds ne peut pas écrire sur un disque donné, tous les nœuds doivent cesser d'accéder à ce disque avant que les résultats de l'opération d'écriture ne soient renvoyés au programme appelant. Par conséquent, si un nœud ne peut pas contacter un disque, il doit d'abord contacter un autre nœud afin de vérifier l'état du disque. En cas de défaillance du disque, les nœuds ne pourront pas y accéder et ils pourront se détacher du disque. Si la défaillance n'est pas due au disque, mais plutôt aux chemins d'accès à partir de certains nœuds, les nœuds ne pourront pas reconnaître l'état du disque. Une stratégie doit résoudre ce type d'incohérence.

Stratégies de détachement des disques

Les stratégies suivantes (définies pour un disk group) permettent de résoudre l'incohérence décrite ci-dessus. Vous pouvez les activer à l'aide de la commande `vxedit (1M)`.

Selon la stratégie de connectivité globale pour le ou les disk groups partagés, le détachement est effectué au niveau du cluster (globalement), si un nœud du cluster indique une défaillance d'un ou de plusieurs disques. Il s'agit de la politique par défaut.

Selon la stratégie de connectivité locale, en cas de défaillance d'un ou de plusieurs disques, les problèmes sont limités aux nœuds spécifiques qui ont indiqué cette défaillance. Le système tente de communiquer avec tous les nœuds du cluster pour confirmer que le ou les disques peuvent être utilisés. Si les tous nœuds indiquent un problème avec le ou les disques, un détachement se produit au niveau du cluster.

Activation des disk groups

L'activation des disk groups contrôle les fonctions d'E/S des volumes à partir des différents nœuds du cluster. Vous ne pouvez pas activer un disk group sur un nœud donné s'il est activé dans un mode conflictuel sur un autre nœud du cluster.

Le tableau 6, « Modes d'activation autorisés et conflictuels », récapitule les modes d'activation autorisés et conflictuels pour les disk groups partagés :

Tableau 6. Modes d'activation autorisés et conflictuels

Disk group activé dans le cluster en....	Tentative d'activer le disk group sur un autre nœud en....			
	Écriture seule	Écriture partagée	Lecture seule	Lecture partagée
Écriture seule	Échec	Échec	Échec	Réussite
Écriture partagée	Échec	Réussite	Échec	Réussite
Lecture seule	Échec	Échec	Réussite	Réussite
Lecture partagée	Réussite	Réussite	Réussite	Réussite

Pour mettre les modes d'activation sous le contrôle de l'utilisateur, le fichier `/etc/default/vxdg` doit être par défaut et il doit contenir la ligne suivante :

```
default_activation_mode=activation-mode
```

où le mode d'activation est : *off* (désactivé), *shared-writing* (écriture partagée), *shared-read* (lecture partagée), *read-only* (lecture seule) ou *exclusive-write* (écriture seule).

Le disk group partagé qui est créé ou importé, sera activé dans le mode spécifié. Lorsqu'un nœud joint le cluster, tous les disk groups partagés sont activés dans le mode spécifié.

Remarques :

- ◆ Lorsque vous lancez l'activation à l'aide du fichier par défaut, il est recommandé d'utiliser un fichier par défaut identique sur tous les nœuds du cluster, faute de quoi les résultats de l'activation seront imprévisibles.
- ◆ Si le nœud d'activation par défaut n'est pas dans un mode désactivé, l'activation suivant la jonction d'un cluster ou, la création ou l'importation d'un disk group, risque d'échouer si un autre nœud du cluster a activé le disk group dans un mode conflictuel.

Journal des zones modifiées et environnements cluster

Le *journal des zones modifiées* (DRL, *Dirty Region Logging*) est une propriété facultative des volumes qui assure la récupération rapide des volumes mis en mémoire après une défaillance du système. Le journal des zones modifiées est pris en charge dans les disk groups partageables par le cluster. Cette section présente un aperçu rapide du journal des zones modifiées et décrit son comportement dans un environnement cluster.

La fonction DRL assure le suivi des zones modifiées en enregistrant les E/S d'un volume mis en miroir et utilise ces informations pour récupérer uniquement les portions appropriées du volume. Elle divise un volume de façon logique, en un ensemble de zones consécutives, et gère un journal des zones modifiées dans lequel un bit d'état représente chaque zone du volume. Les *sous-disques de journal* permettent de stocker le journal des zones modifiées d'un volume dont la fonction DRL est activée. Un tel volume comprend au moins un sous-disque de journal associé à l'un de ses plex.

Avant d'écrire des données dans le volume, les zones écrites sont marquées comme modifiées dans le journal. Si une écriture modifie une zone du journal qui était auparavant intacte, le journal est écrit en mode synchrone sur le disque avant que l'opération d'écriture ne soit effectuée. La zone du journal redevient intacte lorsque les données sont écrites sur le miroir. Au redémarrage du système, Volume Manager récupère uniquement les zones du volume qui sont marquées comme modifiées dans le journal des zones modifiées.

Dans un environnement cluster, l'implémentation de la fonction DRL de Volume Manager diffère légèrement de la méthode normale. Les sections suivantes soulignent certaines de ces différences et décrivent différents aspects de l'implémentation de l'environnement cluster.

Format et taille du journal

À l'instar de VxVM sans cluster, le journal des zones modifiées en clusters existe dans un sous-disque de journal d'un volume mis en miroir.

Un journal des zones modifiées VxVM a un mappage de récupération et une seule mappe active. En revanche, un journal des zones modifiées en clusters a un mappage de récupération et plusieurs mappes actives (une pour chaque nœud du cluster). Contrairement à VxVM, la fonction de cluster met le mappage de récupération au début du journal.

Généralement, le journal des zones modifiées en clusters est plus volumineux que le journal des zones modifiées VxVM car il doit contenir les mappes actives pour tous les nœuds du cluster et un mappage de récupération. La taille de chaque mappe du journal des zones modifiées est d'un ou de plusieurs blocs. `vxassist` alloue automatiquement un journal des zones modifiées de taille suffisante.

La taille du journal dépend de celle du volume et du nombre de nœuds. Le journal doit avoir une taille suffisante pour contenir toutes les mappes (une mappe par nœud, plus un mappage de récupération). La longueur de chaque mappe doit être d'un bloc pour tous les 2 Go du volume. Un journal de 3 blocs (un bloc par mappe) devrait suffire pour un volume de 2 Go dans un cluster à 2 nœuds (il s'agit de la taille minimale du journal). Un volume de 4 Go dans un cluster à 4 nœuds nécessite un journal de 10 blocs, etc.

Lorsque vous ajoutez des nœuds à un cluster existant, vous devez détacher et supprimer les journaux des zones modifiées existants (à l'aide de `vxplex -o rm dis`), puis les recréer (à l'aide de `vxassist addlog`). Cette opération augmente la taille des journaux pour qu'ils puissent contenir les mappes pour les nœuds supplémentaires.

Compatibilité

Sauf pour l'ajout d'un numéro magique spécifique au cluster, les en-têtes du journal des zones modifiées dans un environnement cluster sont identiques à ceux de leurs homologues sans cluster.

Vous pouvez importer un disk group VxVM (et ses volumes) sous la forme d'un disk group partagé dans un environnement cluster et inversement. Cependant, le journal des zones modifiées du disk group importé peut être considéré comme non valide et il se peut qu'une récupération complète soit effectuée.

Si un disk group partagé est importé par un système VxVM sans gestion de cluster, VxVM considère que les journaux des volumes partagés ne sont pas valides et il effectue une récupération complète des volumes. À l'issue de cette récupération, Volume Manager utilise la fonction DRL du cluster.

La fonction de cluster de Volume Manager peut effectuer une récupération DRL sur un volume VxVM non partagé. Cependant, si un volume VxVM est déplacé sur un système VxVM avec gestion de cluster et qu'il est importé en tant que volume partagé, le journal

des zones modifiées ne sera probablement pas suffisant pour contenir tous les nœuds du cluster. La fonction de cluster marquera donc le journal comme non valide et effectuera tout de même une récupération complète. De même, le déplacement d'un volume DRL d'un cluster à 2 nœuds vers un cluster à 4 nœuds risque de créer un journal trop petit, que la fonction de cluster gère en effectuant une récupération complète des volumes. Dans les deux cas, l'administrateur système doit allouer un nouveau journal de taille suffisante.

Fonction DRL dans un environnement cluster

En cas de défaillance d'un ou de plusieurs nœuds d'un cluster, la fonction DRL tente de gérer la récupération de tous les volumes qui étaient utilisés par ces nœuds au moment du problème. Au démarrage initial du cluster, toutes les mappes actives sont incorporées à la mappe de récupération ; cela se produit pendant l'opération `volume start`.

Les nœuds défectueux (qui ont laissé le cluster dans un état instable) ne peuvent pas rejoindre le cluster tant que leurs mappes DRL actives n'ont pas été incorporées dans les mappes de récupération sur tous les volumes affectés. Les utilitaires de récupération comparent les mappes actives d'un nœud défectueux à la mappe de récupération et ils effectuent toutes les mises à jour nécessaires pour permettre au nœud de rejoindre le cluster et de reprendre les opérations d'E/S sur le volume (qui remplace la mappe active). Pendant ce temps, les autres nœuds peuvent poursuivre les opérations d'E/S.

Le noyau VxVM assure le suivi des nœuds défectueux. Si plusieurs récupérations de nœuds sont en cours sur un cluster à un moment donné, leurs récupérations et mises à jours des mappes de récupération respectives peuvent se concurrencer. Le noyau VxVM assure donc le suivi des modifications dans l'état de récupération DRL et évite les collisions des opérations d'E/S.

Le nœud principal assure un suivi « volatile » des mises à jour de la mappe de récupération DRL pour chaque volume et empêche plusieurs utilitaires de modifier simultanément la mappe de récupération.

Mise à niveau de la fonction de cluster de Volume Manager

La fonction de mise à niveau progressive permet à un administrateur de mettre à niveau la version de Volume Manager exécutée sur un cluster sans arrêter le cluster entier. Pour installer la nouvelle version de Volume Manager exécutée sur un cluster, l'administrateur système peut extraire un nœud du cluster, le mettre à niveau, puis joindre de nouveau le nœud au cluster. Il effectue cette opération pour chaque nœud du cluster.

À partir de la version 3.1, à chaque version de VxVM correspond un numéro de *version de protocole de cluster*. Ce numéro est différent du numéro de version. La version du protocole de cluster est stocké dans le fichier `/etc/vx/volboot`. Dans une nouvelle installation de

Utilitaires et démons Volume Manager relatifs au cluster

VxVM, le fichier `volboot` ne se trouve pas dans le répertoire `/etc/vx`. `vxctl init` crée ce fichier et attribue à la version du protocole de cluster le numéro le plus élevé autorisé.

Une nouvelle version de VxVM prend en charge deux versions du protocole de cluster. La version la plus basse correspond à la version VxVM existante. Elle comprend un ensemble fixe de fonctions et de protocoles de communication. La version la plus élevée correspond à une nouvelle version de VxVM qui comprend un nouvel ensemble de ces fonctions. Si la nouvelle version de VxVM ne contient aucune modification de fonction ou de protocole, la version du protocole de cluster ne change pas, par exemple, en cas de changements mineurs ou de correction de bogues. Dans ce cas, il est inutile d'exécuter `vxctl upgrade`.

Pendant une mise à jour progressive, vous devez fermer tous les nœuds et installer la dernière version de VxVM avec la dernière version du protocole de cluster. Tous les nœuds qui ont la nouvelle version de VxVM continuent à utiliser la version de niveau inférieur. Un nœud secondaire sur lequel est installée la nouvelle version du protocole de cluster essaie de joindre le cluster. Si la nouvelle version du protocole de cluster n'est pas utilisée sur le nœud principal, celui-ci rejette la jonction et fournit la version courante du protocole de cluster au nœud secondaire. Le nœud secondaire retente la jonction avec la version du protocole de cluster fournie par le nœud secondaire. Si la jonction échoue à ce stade, la version du protocole de cluster installée sur le nœud principal ne correspond pas à celles gérées par le nœud secondaire en cours de jonction. Dans ce cas, l'administrateur système doit mettre à niveau le cluster vers la dernière version prise en charge pour le protocole de cluster, en passant par chaque version intermédiaire de VxVM.

Tous les nœuds mis à niveau vers la dernière version du protocole de cluster et les nouvelles fonctions sont disponibles.

Lorsque la nouvelle version est installée sur tous les nœuds, vous devez exécuter la commande `vxctl upgrade` sur le nœud principal pour activer la version la plus élevée du protocole de cluster.

Utilitaires et démons Volume Manager relatifs au cluster

Les utilitaires et/ou démons suivants ont été créés ou modifiés pour être utilisés avec Volume Manager dans un environnement cluster :

- ◆ `vxclust`
- ◆ `vxclustadm`
- ◆ `vxconfigd`
- ◆ `vxdbg`
- ◆ `vxdisk`
- ◆ `vxrecover`

- ◆ vxdctl
- ◆ vxstat

Les sections suivantes contiennent des informations sur l'utilisation de chacun de ces utilitaires dans un environnement cluster. Pour plus de détails sur ces utilitaires, reportez-vous aux pages man appropriées.

vxclust

À chaque reconfiguration de cluster, chaque nœud du cluster exécute l'utilitaire `vxclust` à chacune des étapes. Les fonctions du gestionnaire de clusters assurent que la même opération est exécutée sur tous les nœuds à la fois. Sur tous les nœuds, une étape donnée ne commence que lorsque la précédente est terminée. À chaque étape de la reconfiguration, `vxclust` détermine l'action suivante de la fonction de cluster de Volume Manager. Après avoir informé VxVM de son action suivante, `vxclust` attend le résultat de l'opération (réussite, échec ou relance) et communique cette information au gestionnaire de clusters.

Remarques :

- | |
|--|
| <ul style="list-style-type: none">◆ Les commandes Volume Manager nécessitent normalement des privilèges de super-utilisateur.◆ <code>vxclust</code> fonctionne avec SunCluster™ en tant que gestionnaire de clusters. |
|--|

Si un nœud ne répond pas à une requête `vxclust` pendant le délai spécifié, il est abandonné. `vxclust` décide alors de relancer la reconfiguration ou de l'abandonner, suivant les circonstances. Si la reconfiguration est due à une erreur irréparable locale, `vxclust` abandonne l'opération. Si un nœud ne peut pas exécuter une opération car un autre nœud a quitté le cluster, le nœud restant dépasse le délai. Dans ce cas, `vxclust` requiert une reconfiguration en tenant compte qu'un autre nœud quittera le cluster. Dans le cas contraire, `vxclust` force le nœud local à quitter le cluster.

Si une étape de reconfiguration échoue, `vxclust` renvoie une erreur au gestionnaire de clusters. Celui-ci peut décider d'abandonner le nœud, qui quitte immédiatement le cluster. Les E/S en cours vers le disque partagé échouent et l'accès aux disques partagés est arrêté.

`vxclust` décide des actions à prendre lorsqu'il est informé des modifications du cluster. Si un nouveau nœud principal est nécessaire (à cause de la défaillance du nœud principal précédent), `vxclust` détermine le nœud qui deviendra le nouveau nœud principal.

vxclustadm

Remarque vxclustadm ne fonctionne qu'avec VCS.

La commande vxclustadm active et désactive la fonction de cluster de VxVM sur un nœud du cluster. Elle est appelée à partir des scripts en ligne et hors ligne pendant le démarrage et l'arrêt du cluster VCS.

L'option startnode transmet les informations de configuration du cluster au noyau VxVM. En réponse à cette commande, le noyau et le démon de configuration, vxconfigd, effectuent l'initialisation.

L'option stopnode arrête la fonction de cluster sur un nœud. Elle attend que toutes les opérations d'E/S en cours se terminent et que les applications ferment tous les volumes partagés. L'option abortnode abandonne l'activité de gestion de cluster sur un nœud.

Il s'agit d'un arrêt d'urgence qui abandonne toutes les E/S en cours sur les volumes partagés.

L'option nodestate détermine l'état d'un nœud dans le cluster.

Pour plus d'informations, voir la page man vxclustadm(1M).

vxconfigd

Le démon de configuration de Volume Manager, vxconfigd, maintient les configurations des objets VxVM. vxconfigd reçoit les instructions relatives au cluster que l'utilitaire vxclust lui envoie sous SunCluster ou le noyau lors de l'exécution de VCS. Une copie séparée de vxconfigd se trouve sur chaque nœud ; ces copies communiquent entre elles par l'intermédiaire des fonctions de réseau. Pour chaque nœud d'un cluster, les utilitaires Volume Manager communiquent avec l'utilitaire vxconfigd exécuté sur ce nœud spécifique ; les utilitaires n'essaient pas de se connecter aux démons vxconfigd des autres nœuds. Pendant le démarrage du cluster, vxclust (SunCluster) ou le noyau (VCS) indique à vxconfigd de démarrer l'opération de cluster et lui indique son rôle (s'il est un nœud principal ou secondaire).

Lorsqu'un nœud est initialisé pour une opération de cluster, vxconfigd est notifié que le nœud est sur le point de rejoindre le cluster et il reçoit les informations suivantes (à partir de la base de données de configuration du gestionnaire de clusters) :

- ◆ l'ID de cluster ;
- ◆ les ID de nœud ;
- ◆ l'ID du nœud principal ;
- ◆ le rôle du nœud ;
- ◆ l'adresse réseau de chaque nœud vxconfigd.

Sur le nœud principal, `vxconfigd` définit la configuration partagée (il importe les disk groups partagés) et informe `vxclust` (SunCluster) ou le noyau (VCS) lorsque les nœuds secondaires peuvent joindre le cluster.

Sur les nœuds secondaires, `vxconfigd` est notifié du moment où le nœud secondaire peut joindre le cluster. Lorsque le nœud secondaire joint le cluster, `vxconfigd` et le noyau Volume Manager communiquent avec leurs homologues sur le nœud principal afin de définir la configuration partagée.

Lorsqu'un nœud quitte le cluster, le démon `vxconfigd` notifie le noyau sur tous les autres nœuds. Le nœud principal exécute alors toutes les opérations de nettoyage nécessaires. Si le nœud principal quitte le cluster, les noyaux choisissent un nouveau nœud principal et les démons `vxconfigd` de tous les nœuds sont notifiés de ce choix.

`vxconfigd` participe également à la reconfiguration d'un volume. Voir « Reconfiguration d'un volume » à la page 223 pour plus d'informations sur le rôle de `vxconfigd` dans la reconfiguration d'un volume.

Récupération `vxconfigd`

Vous pouvez arrêter/redémarrer à tout moment le démon `vxconfigd` de Volume Manager. Tant que `vxconfigd` est arrêté, les reconfigurations de volume sont interdites et les autres nœuds ne peuvent pas joindre le cluster ; ces opérations seront possibles dès le redémarrage de `vxconfigd`. Dans le cluster, les démons `vxconfigd` sur les nœuds secondaires sont toujours connectés au démon `vxconfigd` du nœud principal. Il est donc déconseillé d'arrêter le démon `vxconfigd` sur un nœud en cluster.

Si `vxconfigd` est arrêté pour une raison quelconque, le système effectue différentes actions suivant le nœud sur lequel le démon est arrêté :

- ◆ Si `vxconfigd` est arrêté sur le ou les nœuds secondaires, le nœud principal n'effectue aucune action. Lorsque `vxconfigd` redémarre sur le nœud secondaire, son démon `vxconfigd` tente de se reconnecter au démon du nœud principal et d'obtenir de nouveau les informations sur la configuration partagée. (Le noyau peut toujours voir la configuration partagée et l'accès aux disques partagés n'est pas affecté.) Tant que le démon `vxconfigd` du nœud secondaire n'a pas rejoint le nœud principal, il ne reçoit que très peu d'informations sur la configuration partagée et les tentatives d'afficher ou de modifier cette configuration risquent d'échouer. En particulier, si les disk groups partagés sont listés (à l'aide de `vxdg list`), ils sont marqués comme désactivés (`disabled`) ; lorsque la jonction réussit, ils sont marqués comme activés (`enabled`).
- ◆ Si `vxconfigd` est arrêté sur le nœud principal, le démon `vxconfigd` du ou des nœuds secondaires tente régulièrement de rejoindre le nœud principal. Cette opération réussit uniquement lorsque `vxconfigd` est redémarré sur le nœud

Utilitaires et démons Volume Manager relatifs au cluster

principal. Dans ce cas, les informations que le démon `vxconfigd` du nœud secondaire a reçues sur la configuration partagée ne sont pas perdues et les affichages de configuration sont exacts.

- ◆ Si le démon `vxconfigd` est arrêté à la fois sur le nœud principal et sur le ou les nœuds secondaires, celui-ci doit attendre le redémarrage de `vxconfigd` sur les deux nœuds et l'établissement d'une reconnexion entre ces derniers pour afficher des informations de configuration exactes.

Lorsque le démon `vxclust` (SunCluster) ou le noyau (VCS) détecte que `vxconfigd` est arrêté sur un nœud, `vxconfigd` redémarre.

Remarque Avec VxVM, l'option `-r reset` sur `vxconfigd` redémarre `vxconfigd` et crée tous les états. Cette option n'est pas disponible lorsqu'un nœud se trouve dans le cluster, car elle causerait la perte de ses informations ; si vous utilisez cette option dans ce cas spécifique, `vxconfigd` ne démarre pas.

vxdbg

L'utilitaire `vxdbg` gère les disk groups Volume Manager. Vous pouvez l'utiliser pour spécifier qu'un disk group est partageable par le cluster. L'option `-s` sur `vxdbg` permet d'initialiser ou d'importer un disk group à l'état « partagé ».

Si vous avez exécuté le logiciel de cluster pour configurer le cluster, vous pouvez entrer la commande suivante pour créer un disk group partagé :

```
vxdbg -s init diskgroup [medianame=]nom_d'accès
```

où *diskgroup* représente le nom du disk group ; *medianame*, le nom d'administration choisi pour le disque ; et *nom_d'accès*, le nom d'accès au disque (ou le nom du périphérique).

La commande `vxdbg -s import` vous permet d'importer les disk groups en tant que disk groups partagés. Si les disk groups ont été configurés avant l'exécution du logiciel de cluster, vous pouvez utiliser la commande suivante pour importer les disk groups dans la structure de cluster :

```
vxdbg -s import diskgroup
```

où *diskgroup* représente le nom ou l'ID du disk group. Lors des redémarrages suivants du cluster, le disk group sera automatiquement importé en tant que disk group partagé. Vous devrez peut-être l'exporter (à l'aide de la commande `vxdbg deport diskgroup`) avant d'appeler cette commande.

Pour convertir un disk group partagé en disk group privé, exportez-le au moyen de `vxdbg deport`, puis importez-le en utilisant la commande `vxdbg import diskgroup`.

Remarque Le système ne peut pas déterminer si un disque est partagé. Pour protéger l'intégrité des données lorsque des disques accessibles par plusieurs systèmes sont utilisés, l'administrateur système doit veiller à utiliser la désignation correcte pour ajouter un disque à un disk group. S'il essaie d'ajouter un disque qui n'est pas physiquement partagé à un disk group qui l'est, Volume Manager autorise cette opération sur le nœud où le disque est accessible, à condition qu'il soit unique dans le cluster. Cependant, les autres nœuds ne pourront pas joindre le cluster. De plus, si l'administrateur essaie d'ajouter simultanément le même disque à des disk groups différents sur deux nœuds, les résultats seront imprévisibles. Il faut donc veiller à gérer toutes les configurations sur un seul nœud.

`vxldg` comprend une option `force (-f)`, qui permet de forcer l'importation d'un disk group ou l'ajout d'un disque à un disk group.

Remarque L'option `force (-f)` doit être utilisée avec précaution, et uniquement si l'administrateur système connaît les conséquences éventuelles.

Au redémarrage d'un cluster, VxVM peut refuser d'importer automatiquement un disk group pour l'une des raisons suivantes :

- ◆ Un disque de ce disk group n'est plus accessible à cause d'erreurs matérielles. Dans ce cas, l'administrateur système peut réimporter le disk group en utilisant l'option `force` comme suit :


```
# vxldg -s -f import diskgroup
```
- ◆ Certains nœuds auxquels des disques du disk group sont attachés ne se trouvent pas actuellement dans le cluster et le disk group ne peut donc pas accéder à tous ses disques. Dans ce cas, l'importation forcée n'est pas sûre et vous ne devez pas l'exécuter (vous risquez d'obtenir des miroirs incohérents).

Si VxVM n'ajoute pas un disque à un disk group existant (parce que ce disque n'est pas attaché aux mêmes nœuds que les autres disques du disk group), l'administrateur système peut forcer son ajout comme suit :

```
# vxldg -f adddisk -g diskgroup [medianame=]nom_d'accès
```

`vxldg` permet également de lister les disk groups partagés. La commande suivante affiche une ligne d'information pour chaque disk group :

```
# vxldg list
```

Cette commande affiche la liste suivante :

NOM	ÉTAT	ID
rootdg	enabled	774215886.1025.teal
group2	enabled,shared	774575420.1170.teal
group1	enabled,shared	774222028.1090.teal

Utilitaires et démons Volume Manager relatifs au cluster

Les disk groups partagés sont signalés par l'indicateur `shared` (partagé).

La commande suivante affiche une ligne d'information pour chaque disk group partagé :

```
# vxdg -s list
```

Cette commande affiche la liste suivante :

NOM	ÉTAT	ID
group2	enabled, shared	774575420.1170.teal
group1	enabled, shared	774222028.1090.teal

La commande suivante affiche des informations sur un disk group spécifique (elle indique notamment s'il est partagé) :

```
# vxdg list diskgroup
```

où *diskgroup* représente le nom du disk group.

La commande `vxdg list group1` affiche les informations suivantes sur le nœud principal (pour le disk group `group1`) :

```
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
```

La valeur `shared` est attribuée au champ `flags`. La même commande affichera des informations légèrement différentes sur un nœud secondaire.

vxdisk

L'utilitaire `vxdisk` gère les disques Volume Manager. Il permet de déterminer si un disque est inclus dans un disk group partageable par le cluster :

```
# vxdisk list nom_accès
```

où *nom_accès* représente le nom d'accès au disque (ou le nom du périphérique).

Cette commande affiche les informations suivantes (pour le périphérique `c1t0d0s2`) :

```

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

```

La valeur `cvm` (le nom du cluster) est attribuée au champ `clusterid` et le champ `flags` inclut une entrée pour `shared`. Lorsqu'un nœud n'est pas joint, le champ `flags` contient l'indicateur `autoimport` à la place de l'indicateur `imported`.

vxrecover

L'utilitaire `vxrecover` récupère les plex et les volumes après le remplacement d'un disque.

Lorsqu'un nœud quitte le cluster, il peut laisser certains miroirs dans un état incohérent. L'utilitaire `vxrecover` exécute une récupération sur tous les volumes ayant cet état. L'option `-c` indique à `vxrecover` d'effectuer la récupération pour tous les volumes des `disk groups` partageables par le cluster. Le cas échéant, `vxclust` appelle automatiquement `vxrecover -c`.

Remarque Lorsque l'utilitaire `vxrecover` est actif, il se peut que les performances du système diminuent.

vxdctl

L'utilitaire `vxdctl` gère certains aspects du démon de configuration de volume, `vxconfigd`. L'option `-c` permet de demander des informations sur le cluster. Vous pouvez utiliser `vxdctl` comme suit pour déterminer si `vxconfigd` est activé et/ou en cours d'exécution :

```
# vxdctl -c mode
```

Suivant les circonstances, vous pouvez obtenir un affichage du type suivant :

```
mode: enabled: cluster active - MASTER
mode: enabled: cluster active - SLAVE
mode: enabled: cluster inactive
mode: enabled: cluster active - role not set
```

Remarque Si `vxconfigd` est désactivé, aucune information sur le cluster n'est affichée.

Pour une description complète de `vxdctl`, voir la page man `vxdctl(1M)`.

`vxdctl` indique la version du protocole de cluster et la plage du protocole de cluster. Une fois que tous les nœuds du cluster ont été mis à jour avec le nouveau protocole de cluster, le cluster entier est mis à niveau au moyen de la commande suivante :

```
# vxdctl upgrade
```

La commande `vxdctl protocolversion` permet de vérifier la version existante du protocole de cluster. Par exemple :

```
# vxdctl protocolversion
Cluster running at protocol 10
```

La commande `vxdctl protocolrange` indique les versions maximale et minimale de protocole de cluster prises en charge par la version actuelle de VxVM. Par exemple,

```
# vxdctl protocolrange
minprotoversion: 10, maxprotoversion: 20
```

La commande `vxdctl list` indique la version du protocole de cluster exécutée sur un nœud. Elle peut par exemple contenir les informations suivantes :

```
Volboot file
version: 3/1
seqno: 0.19
cluster protocol version: 20
hostid: giga
entries:
```

La commande `vxctl support` affiche les versions maximale et minimale de protocole prises en charge par le nœud ainsi que la version de protocole actuelle. Elle peut par exemple contenir les informations suivantes :

```
Support information:
  vold_vrsn: 11
  dg_minimum: 60
  dg_maximum: 70
  kernel: 10
  protocol_minimum: 10
  protocol_maximum: 20
  protocol_current: 20
```

vxstat

`vxstat` renvoie des statistiques sur les objets spécifiés. Dans un environnement cluster, `vxstat` réunit des statistiques provenant de tous les nœuds du cluster. Ces statistiques indiquent l'utilisation totale, par tous les nœuds, pour les objets demandés. Si un objet local est spécifié, la commande renvoie son utilisation locale.

`vxstat` permet au programme appelant de spécifier un sous-ensemble de nœuds facultatif :

```
# vxstat -g diskgroup -n noeud[,noeud...]
```

où *noeud* représente un nombre entier. Si vous entrez une liste de nœuds séparés par des virgules, `vxstat` affiche la somme des statistiques pour les nœuds de la liste.

Dans l'exemple suivant, `vxstat` doit obtenir des statistiques pour le nœud 2, volume `vol1` :

```
# vxstat -g group1 -n 2 vol1
```

Cette commande peut afficher les informations suivantes :

TYPE	NOM	OPÉRATIONS		BLOCS		DURÉE MOYENNE (ms)	
		LECTURE	ÉCRITURE	LECTURE	ÉCRITURE	LECTURE	ÉCRITURE
vol	vol1	2421	0	600000	0	99.0	0.0

`vxstat` peut également obtenir et afficher des statistiques pour tout le cluster, comme suit :

```
# vxstat -b
```

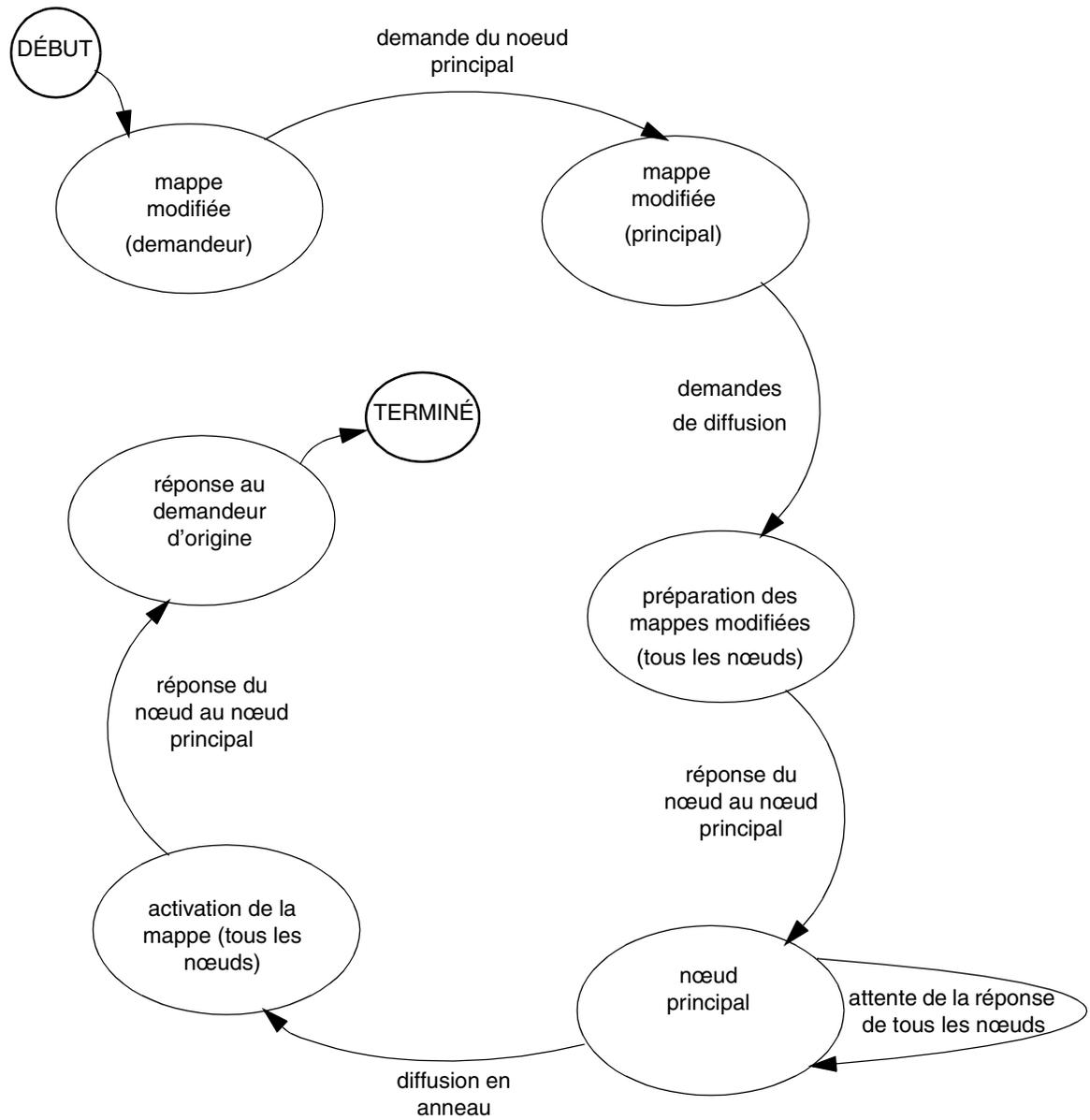
Les statistiques sur tous les nœuds sont ajoutées les unes aux autres. Par exemple, si le nœud 1 a effectué 100 E/S et le nœud 2, 200 E/S, `vxstat -b` renvoie 300.

Gestion des clusters et resynchronisation rapide des miroirs

La resynchronisation rapide des miroirs (FMR, *Fast Mirror Resynchronization*) est prise en charge pour les volumes partagés. Les mappes de mise à jour (mappes FMR) sont distribuées dans le cluster. L'existence de mappes « volatiles » est moins probable dans un environnement cluster, car seul un nœud du cluster doit être actif pour que les mappes FMR soient disponibles. La défaillance d'un seul nœud n'entraîne pas la perte de mappes FMR.

Les mappes sont mises à jour sur le nœud principal. Lorsque le nœud principal distribue les mises à jour à tous les nœuds, celles-ci s'appliquent soit à tous les nœuds, soit à aucun des nœuds. Le nœud principal orchestre un engagement en deux phases pour appliquer les mises à jour éventuelles. Voir la figure 25, « Clustérisation des bitmaps », à la page 241.

Figure 25. Clustérisation des bitmaps





Gestion des clusters et resynchronisation rapide des miroirs



Récupération

7

Introduction

VERITAS Volume Manager protège les systèmes contre des défaillances de disque et vous aide à effectuer la récupération après ces défaillances. Ce chapitre décrit les procédures de récupération et contient des informations qui vont aider à éviter les pertes de données ou les accès au système dus aux défaillances de disque. Il décrit également les états possibles des plex et des volumes.

Pour plus d'informations sur la protection du système, voir « Initialisation de Volume Manager » à la page 49 et « Configuration du système » à la page 52.

Ce chapitre traite des sujets suivants :

- ◆ Procédure de démarrage UNIX
- ◆ Configurations root (/), swap et usr possibles
- ◆ Défaillances et procédures de récupération
- ◆ Reconstruction automatique et défaillances du disque d'initialisation
- ◆ Rajout et remplacement de disques d'initialisation
- ◆ Rattachement des disques
- ◆ Récupération de réinstallation
- ◆ États des plex et des volumes
- ◆ Récupération des volumes RAID 5
- ◆ Opérations RAID 5 diverses

Procédure de démarrage UNIX

Le système Sun SPARC utilise l'invite de commande `boot` sauf si l'indicateur `autoboot` a été défini dans la zone de stockage permanente utilisée par les microprogrammes. Les ordinateurs ayant une PROM plus ancienne utilisent une invite différente de celles des nouvelles versions V2 et V3 de la PROM. Ces nouvelles versions sont également appelées OpenBoot PROM (OBP). La commande `boot` a une syntaxe différente pour ces deux types de PROM :

```
# ok boot [noms OBP] [nom_fichier] [indicateurs_démarrage]
```

noms OBP spécifie les désignations de PROM Open Boot. Par exemple, sur les systèmes Desktop SPARC, la désignation

```
# /sbus/esp@0,800000/sd@3,0:a
```

indique le disque SCSI (`sd`) à la cible 3, numéro d'unité logique 0 sur le bus SCSI, avec la carte de contrôleur `esp` connectée à l'emplacement d'extension 0.

Remarque Volume Manager vous permet d'utiliser les noms d'alias pour les disques de démarrage. Ces alias peuvent être les noms fournis par Volume Manager (par exemple, `vx-rootdisk` ou `vx-disk01`) ou par le système d'exploitation (par exemple, `disk1`). Vous pouvez afficher la liste des périphériques amorçables possibles en entrant la commande suivante à l'invite OK de OpenBoot : `devalias`.

Le paramètre *nom_fichier* représente le nom du programme autonome pour le programme `boot`. Par défaut, le démarrage s'effectue avec `/kernel/unix` à partir de la partition `root`. Vous pouvez spécifier un autre programme (par exemple, `/stand/diag`) sur la ligne de commande. Certaines versions des microprogrammes permettent d'enregistrer le nom de fichier par défaut dans la zone de stockage non volatile du système.

Le programme `boot` interprète l'indicateur `-a` comme une requête de type « demandez-moi » et vous invite à entrer le nom du programme autonome à démarrer. L'indicateur `-a` est ensuite transmis au programme autonome.

Remarque Un système exécutant Volume Manager avec le contrôle racine de Volume Manager n'utilise pas pour démarrer les paramètres par défaut représentés par l'indicateur `-a`. Voir « Copie de `/etc/system` disponible » à la page 253 pour obtenir les réponses correctes pour `boot -a`.

Le programme `boot` n'interprète pas les indicateurs. Il transmet tous les indicateurs de démarrage au fichier identifié par *nom_fichier*. Démarrez le système. Voir les pages `man kernel (1)` et `kadb (1M)` pour plus d'informations sur les options disponibles avec le programme autonome par défaut, `/kernel/unix`.

Démarrage après une défaillance

Si le disque d'initialisation est mis en miroir, vous pouvez utiliser le disque d'initialisation de remplacement pour initialiser le système en cas de défaillance du disque d'initialisation principal. Pour démarrer le système après la défaillance du disque d'initialisation principal, procédez comme suit :

1. Recherchez les disques VM avec alias qui utilisent la commande `devalias` à l'invite de commande OpenBoot.

Les disques qui sont des miroirs appropriés du disque racine sont listés sous le nom : `vx nom_logique`, où `nom_logique` représente le nom d'accès logique au disque contenant le système de fichiers racine candidat.

2. Entrez la commande suivante :

```
# ok boot nom_alias
```

où `nom_alias` représente le nom d'alias du disque sélectionné.

Si un disque sélectionné contient un miroir racine stale (non valide), `vxconfigd` affiche une erreur indiquant que le miroir est inutilisable ainsi que tous les disques amorçables de remplacement qui ne sont pas stale.

Configurations root (/), swap et usr possibles

Pendant l'installation, différentes configurations sont possibles pour les systèmes de fichiers racine, swap et `usr` :

- ◆ `usr` est un sous-répertoire du répertoire racine (`root`) et aucune partition séparée ne lui est allouée. Dans ce cas, `usr` devient une partie du volume `rootvol` lorsque le disque racine est encapsulé et placé sous le contrôle de Volume Manager.
- ◆ `usr` se trouve sur une partition séparée du disque racine. Dans ce cas, un volume séparé est créé pour la partition `usr`. `vxmirror` met en miroir le volume `usr` sur le disque de destination.
- ◆ `usr` se trouve sur une partition séparée qui n'est pas située sur le disque racine. Dans ce cas, un volume est créé pour la partition `usr` uniquement si ce disque est encapsulé par Volume Manager. L'encapsulation du disque racine et la présence de miroirs sur le volume racine sont inutiles si la partition `usr` devient inaccessible pour une raison ou une autre. Il est recommandé d'encapsuler aussi bien le disque qui contient la partition `usr` que le disque `root`, et de disposer de miroirs pour les volumes `usr`, `rootvol` et `swapvol` afin d'assurer la disponibilité maximale du système.

Le volume `rootvol` doit exister sur le disk group `rootdg`. Voir « Limitations des volumes au démarrage » à la page 81 pour plus d'informations sur les limitations des volumes `rootvol` et `usr`.

Configurations root (/), swap et usr possibles

Volume Manager vous permet de placer des partitions `swap` sur n'importe quel disque ; une zone de `swap` initiale est inutile pendant les premières phases du démarrage. Par défaut, l'installation de Volume Manager choisit la partition 0 du disque racine sélectionné comme partition `root`, et la partition 1 comme partition `swap`. Cependant, vous pouvez choisir la partition `swap` sur une partition qui ne se trouve pas sur le disque racine. Dans ce cas, le programme vous recommande d'encapsuler ce disque et de créer des miroirs pour le volume `swap`, faute de quoi l'endommagement de la partition `swap` risque de provoquer une défaillance du système. Vous pourrez peut-être démarrer le système, mais la présence de miroirs pour le volume `swapvol` empêche les défaillances du système.

Réparation des systèmes de fichiers root (/) ou /usr sur les volumes

Si le système de fichiers racine (/) ou /usr devient inutilisable, il est préférable de démarrer à partir d'un système de fichiers racine monté en réseau ou d'une sauvegarde valide. Celle-ci doit contenir toutes les partitions du système de fichiers appropriées sur le disque racine. Vous devez également disposer d'une copie imprimée de la table de partition du disque racine avant que ce dernier ne soit encapsulé.

Cette tâche est plus complexe lorsque le système de fichiers `root` ou /usr est défini sur un volume mis en miroir. Les modifications de la partition sous-jacentes à l'un des miroirs peuvent provoquer des dommages lorsque Volume Manager démarre par la suite et suppose que les miroirs sont correctement synchronisés.

Il existe deux solutions à ce problème :

- ◆ La solution la plus simple consiste à monter un plex du système de fichiers racine ou /usr, le réparer, le démonter et utiliser `dd` pour copier le plex fixe sur tous les autres plex. Cependant, cette procédure est sujette aux erreurs.
- ◆ L'autre solution consiste à restaurer le système à partir d'une bande de sauvegarde valide. Cette procédure est décrite ci-dessous. Elle ne nécessite pas que le système d'exploitation soit installé à partir du CD-ROM de base.

La procédure décrite ci-dessous est une méthode de récupération simple, efficace et fiable lorsque le disque racine et son miroir sont endommagés.

Récupération d'un disque racine (et du miroir racine) Volume Manager à partir d'une bande de sauvegarde

Cette procédure suppose ce qui suit :

- ◆ Vous disposez de la sauvegarde complète actuelle de tous les systèmes de fichiers contenus dans le disque racine Volume Manager d'origine.
- ◆ Vous avez installé un nouveau disque d'initialisation pour remplacer le disque d'initialisation d'origine qui a échoué, s'il a été physiquement endommagé.

Cette procédure nécessite la réinstallation du disque racine Volume Manager. Pour empêcher la perte de données sur les disques non inclus dans la réinstallation, vous ne devez utiliser que le disque racine Volume Manager dans la procédure de réinstallation.

Plusieurs des options d'installation automatiques accèdent aux disques différents du disque racine sans demander confirmation à l'administrateur. Par conséquent, il est recommandé de déconnecter tous les autres disques (contenant des volumes) du système avant de lancer cette procédure. Cette opération assure que les autres disques ne seront pas affectés par la réinstallation. Reconnectez ces disques à l'issue de la procédure.

Récupération d'un disque racine

La procédure ci-dessous suppose que le (nouveau) disque d'initialisation est `c0t0d0` et que vous devez récupérer à la fois les systèmes de fichiers `/` et `/usr` (respectivement `s0` et `s6`).

1. Démarrez le système d'exploitation à partir du CD-ROM.
2. Utilisez la commande `format` pour créer sur le (nouveau) disque d'initialisation (`c0t0d0`) des partitions identiques qui contiendront les systèmes de fichiers précédemment placés sur le disque d'initialisation d'origine.

Remarque Vous pouvez créer un maximum de sept partitions car vous devez restaurer les systèmes de fichiers. De plus, puisque vous réencapsulez provisoirement ce disque, vous ne pouvez disposer que de cinq partitions au maximum, étant donné que deux partitions sont requises pour les zones publiques et privées sur le disque.

3. Montez `/dev/rdisk/c0t0d0s0` sur `/a/root`. Restaurez le système de fichiers racine à partir de la bande. Installez un périphérique bootblock sur `/a/root` au moyen de `installboot`.
4. Montez `/dev/rdisk/c0t0d0s6` sur `/a/usr`. Restaurez le système de fichiers `/usr` à partir de la bande.
5. Modifiez le système de fichiers racine restauré comme suit :
 - ◆ `touch /a/root/etc/vx/reconfig.d/state.d/install-db`
 - ◆ modifiez `/a/root/etc/system` en supprimant les deux lignes suivantes :

```
rootdev:/pseudo/vxio@0:0
set vxio:vol_rootdev_is_volume=1
```

Configurations root (/), swap et usr possibles

- ◆ modifiez `/a/root/etc/vfstab` en remplaçant les entrées de périphérique de volume Volume Manager `/dev/vx/dsk` par les périphériques de disque standard `/dev/dsk/c0t0d0s0` et `/dev/dsk/c0t0d0s6`.
6. Redémarrez le système à partir du (nouveau) disque d'initialisation. Ainsi, celui-ci considère que Volume Manager *n'est pas* installé.

L'étape suivante de la procédure dépend de la présence ou de l'absence de miroirs racines dans l'ancien `rootdg` :

- ◆ Si l'ancien `rootdg` contient d'autres disques qui *ne sont pas* utilisés comme miroirs racines, passez à l'étape 7.
 - ◆ Si l'ancien `rootdg` contient uniquement des miroirs racines, passez à l'étape 8.
7. Si l'ancien `rootdg` contient d'autres disques qui *ne sont pas* utilisés comme miroirs racines, procédez comme suit afin de passer dans l'ancien `rootdg` (moins le disque d'initialisation que Volume Manager considérera comme un disque qui a échoué) et configurez le nouveau disque d'initialisation.

- a. Supprimez les fichiers à présent inutiles utilisés pour l'installation :

```
rm -r /etc/vx/reconfig.d/state.d/installdb
```

- b. Démarrez les démons d'E/S VxVM :

```
# vxiod set 10
```

- c. Démarrez le démon de configuration VxVM en mode désactivé :

```
# vxconfigd -m disable
```

- d. Initialisez le démon `vxconfigd` :

```
# vxdctl init
```

- e. Activez `vxconfigd` :

```
# vxdctl enable
```

La procédure ci-dessus doit amener l'ancien `disk group` racine moins le disque racine que VxVM considérera comme ayant échoué.

- a. Utilisez la commande `vxedit` (ou Volume Manager Storage Administrator) pour supprimer les anciens volumes de disque racine et le disque racine.
- b. Utilisez la commande `vxdiskadm` pour encapsuler le (nouveau) disque d'initialisation et initialiser les disques qui serviront de miroirs racines. Après le redémarrage requis, mettez en miroir le disque racine sur les miroirs racines.

- c. Si l'ancien `rootdg` contient uniquement des miroirs racines :

Exécutez la commande `vxinstall` pour encapsuler le (nouveau) disque d'initialisation et initialiser le ou les miroirs racines.

Après le redémarrage requis, mettez en miroir le disque racine sur les miroirs racines.

Sauvegarde et restauration du système de fichiers racine

Il est recommandé de sauvegarder le système de fichiers racine `root` afin de pouvoir le restaurer s'il est endommagé.

Si vous utilisez le système de fichiers `ufs`, vous pouvez sauvegarder le système de fichiers `root` en entrant la commande suivante :

```
# /usr/lib/fs/ufs/ufsdump [dump-options] /dev/vx/rdisk/rootvol
```

Vous pouvez ensuite restaurer le système de fichiers `root` après une défaillance en procédant comme suit :

1. Démarrez à partir d'un CD-ROM ou d'un système de fichiers racine monté en réseau, puis exécutez Volume Manager (voir « Réparation des systèmes de fichiers `root` (`/`) ou `/usr` sur les volumes » à la page 246).
2. Montez et restaurez le système de fichiers `root` en entrant les commandes suivantes :

```
# newfs /dev/vx/rdisk/rootvol
# mount /dev/vx/dsk/rootvol /mnt
# cd /mnt
# /usr/lib/fs/ufs/ufsrestore [restore-options]
```

Défaillances et procédures de récupération

Il existe différents types de défaillance qui peuvent empêcher le démarrage d'un système, mais vous pouvez appliquer la même procédure de base pour relancer ce dernier. Lorsque vous ne pouvez pas démarrer un système, tentez d'abord d'identifier la défaillance à partir des données affichées à l'écran, puis de corriger le problème (par exemple, mettez sous tension un périphérique qui avait été accidentellement mis hors tension). Si vous ne pouvez pas corriger le problème (par exemple, des erreurs de données sur le disque d'initialisation), démarrez le système à partir d'un disque d'initialisation de remplacement (contenant un miroir du volume `root`) pour pouvoir réparer les dommages ou remplacer le disque qui a échoué.

Cette section présente quelques défaillances possibles et fournit des solutions de dépannage.

Défaillances dans le partitionnement UNIX

Une fois chargé, le programme de démarrage essaie d'accéder au disque d'initialisation en utilisant des informations de partition UNIX normales. Si ces informations sont endommagées, le programme de démarrage échoue et affiche l'erreur suivante :

```
File just loaded does not appear to be executable
(Le fichier qui vient d'être chargé ne semble pas être exécutable)
```

Si ce message s'affiche pendant la tentative de démarrage, vous devez démarrer le système à partir d'un disque d'initialisation de remplacement. Pendant le démarrage, les pilotes de disque affichent souvent sur la console des erreurs relatives aux informations de partition UNIX non valides sur le disque qui a échoué. Ces messages sont du type suivant :

```
WARNING: unable to read label
WARNING: corrupt label_sdo
```

Ces messages indiquent que la défaillance est due à une partition de disque non valide. Vous pouvez tenter de rattacher le disque comme décrit dans « Rajout d'un disque d'initialisation qui a échoué » à la page 257. Cependant, si le rattachement échoue, vous devez remplacer le disque en suivant la procédure décrite dans « Remplacement d'un disque d'initialisation qui a échoué » à la page 259.

Défaillances lors de l'accès au périphérique de démarrage

Des messages du type suivant peuvent s'afficher au début de la procédure de démarrage, juste après l'initialisation du système :

```
SCSI device 0,0 is not responding
Can't open boot device
```

Ce message signifie que la PROM système n'a pas pu lire le programme de démarrage à partir du lecteur de démarrage. Les causes de ce problème sont en général les suivantes :

- ◆ Le disque d'initialisation n'est pas sous tension.
- ◆ Le bus SCSI n'a pas de terminaison.
- ◆ Un contrôleur est défectueux.
- ◆ Un disque défectueux verrouille le bus, interdit aux disques de s'identifier auprès du contrôleur et indique au contrôleur qu'aucun disque n'est attaché.

La première étape pour diagnostiquer ce problème consiste à vérifier soigneusement le bon fonctionnement général du bus SCSI. Si les disques sont mis hors tension ou que le bus n'a pas de terminaison, corrigez le problème et redémarrez le système. Si l'un des disques a échoué, supprimez-le du bus et remplacez-le.

Si vous ne constatez aucun problème matériel, le problème est probablement dû à des erreurs de données sur le disque d'initialisation. Pour corriger ce problème, essayez de démarrer le système à partir d'un disque d'initialisation de remplacement (contenant un miroir du volume racine). Si vous n'y parvenez pas, il subsiste un problème matériel. De même, si vous ne pouvez toujours pas démarrer le système après avoir remplacé le disque d'initialisation défectueux par un disque d'initialisation de remplacement, le problème est alors d'ordre matériel.

Défaillances dues à des entrées incorrectes dans `/etc/vfstab`

Lorsque le disque racine est encapsulé et placé sous le contrôle de Volume Manager, lors de la procédure normale d'encapsulation, des volumes sont créés pour toutes les partitions du disque. Volume Manager modifie `/etc/vfstab` pour utiliser les volumes correspondants à la place des partitions de disque. Vous devez modifier manuellement le fichier `/etc/vfstab` avec précaution. Les entrées les plus importantes sont celles correspondant à `/` et `/usr`. Le fichier `vfstab` qui existait avant l'installation de Volume Manager est enregistré dans `/etc/vfstab.prevm`.

Entrée / endommagée dans `/etc/vfstab`

Si l'entrée correspondant à `/` dans `/etc/vfstab` est perdue ou incorrecte, le système démarre en mode mono-utilisateur. Des messages du type suivant s'affiche :

```
File just loaded does not appear to be executable
```

Il est recommandé d'exécuter `fsck` à ce stade :

```
# fsck /dev/vx/rdisk/rootvol
```

À ce stade du démarrage, `/` n'est pas encore monté en mode lecture/écriture. Puisque l'entrée dans `/etc/vfstab` était incorrecte ou supprimée, montez manuellement `/` en mode lecture/écriture en entrant la commande suivante :

```
# mount -o remount /dev/vx/dsk/rootvol
```

Après avoir monté `/` en mode lecture/écriture, quittez le shell. Le système vous demande d'indiquer le niveau d'exécution. Pour le mode multi-utilisateur, indiquez le mode d'exécution 3 :

```
ENTER RUN LEVEL (0-6,s or S) : 3
```

Restaurez l'entrée dans `/etc/vfstab` pour `/` après le démarrage du système.

Défaillances et procédures de récupération

Entrée /usr endommagée dans /etc/vfstab

/etc/vfstab a une entrée pour /usr uniquement si /usr se trouve dans une partition de disque séparée. Après l'encapsulation du disque contenant la partition /usr, Volume Manager modifie l'entrée de /etc/vfstab pour utiliser le volume correspondant.

En cas de perte de l'entrée pour /usr à partir de /etc/vfstab, vous ne pouvez pas démarrer le système (même si vous disposez de miroirs du volume /usr). Dans ce cas, démarrez le système à partir du CD-ROM et restaurez /etc/vfstab. (Voir « Réparation des systèmes de fichiers root (/) ou /usr sur les volumes » à la page 246.)

Défaillances dues à un fichier /etc/system manquant ou endommagé

Remarque Vous ne devez pas modifier les entrées de /etc/system qui sont ajoutées par Volume Manager. Les entrées de Volume Manager sont toutes incluses entre *vxvm_START et *vxvm_END.

Il est recommandé d'effectuer une copie de /etc/system dans le système de fichiers racine avant de le modifier. Vous pouvez ensuite spécifier le fichier système enregistré dans le programme boot si les modifications apportées au nouveau fichier /etc/system sont incorrectes. Pour cela, démarrez le système au moyen de la commande `boot -a`. Lorsque le système vous demande d'entrer le nom du fichier système, entrez le chemin du fichier système enregistré.

Copie de /etc/system non disponible

Si le fichier /etc/system est endommagé et que la copie enregistrée du fichier système n'est pas disponible, vous ne pouvez pas démarrer le système lorsque la fonction de contrôle racine est activée. Vous pouvez démarrer le système sans cette fonction (c'est-à-dire, sans que `rootvol` corresponde à /) si /usr n'est pas un volume.

Démarrez le système sans la fonction de contrôle racine selon la procédure décrite dans « Réparation des systèmes de fichiers root (/) ou /usr sur les volumes » à la page 246 pour :

1. démarrer Volume Manager à partir du CD-ROM. Reportez-vous au *Guide d'installation de VERITAS Volume Manager* pour plus d'informations ;
2. exécuter la commande `fixmountroot` ;
3. créer et monter `/tmp/rootvol`.

Vous pouvez ensuite modifier le fichier `/tmp/rootvol/etc/system` et effectuer toutes les autres réparations nécessaires.

Une fois que vous avez démarré le système, entrez ce qui suit dans `/etc/system` :

```
* vxvm_START
rootdev:/pseudo/vxio@0:0
set vxio:vol_rootdev_is_volume=1
* vxvm_END
```

Vous devez aussi forcer le chargement de tous les pilotes requis pour les disques miroirs racines. Pour cela, modifiez le fichier `/etc/system` pour qu'il contienne une ligne de la forme suivante pour chacun des pilotes :

```
# forceload: drv/driver_name
```

Pour obtenir le nom de ces pilotes de disque, vous pouvez effectuer un long listing sur `/dev/dsk/périphérique_racine`. Par exemple, un pilote peut être nommé `io-unit`.

Copie de `/etc/system` disponible

Si le fichier `/etc/system` est endommagé et que la copie enregistrée du fichier `etc/system` est disponible, vous pouvez démarrer le système avec la fonction de contrôle racine.

Pour démarrer le système avec cette fonction, utilisez la commande et les réponses suivantes (lorsque vous appuyez sur la touche de retour, vous acceptez les valeurs par défaut pour toutes les invites *sauf* le nom de périphérique racine) :

```
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` non disponible et `/usr` est un volume

Si le fichier `etc/system` est endommagé ou perdu, qu'aucune copie de sauvegarde n'est disponible et que `/usr` est un volume, vous devez démarrer le système à partir du CD-ROM (selon la procédure décrite dans « Réparation des systèmes de fichiers root (/)

Défaillances et procédures de récupération

ou /usr sur les volumes » à la page 246). À l'issue de cette opération, montez le volume racine et modifiez son fichier `etc/system`. Créez les entrées suivantes dans le fichier `etc/system`:

```
* vxvm_START
rootdev:/pseudo/vxio@0:0
set vxio:vol_rootdev_is_volume=1
set vxio:vol_swapdev_is_volume=1
* vxvm_END
```

Vous devez aussi forcer le chargement de tous les pilotes requis pour les disques miroirs racines (comme décrit précédemment). Une fois ces modifications effectuées, redémarrez le système à partir de la même partition `root` que celle sur laquelle le fichier système a été restauré.

Défaillances dues à un démarrage à partir de plex inutilisables ou stale

Si un disque n'est pas disponible lorsque le système est exécuté, tous les miroirs des volumes qui se trouvent sur ce disque deviennent stale (non valides). Autrement dit, les données de ce disque manquent de cohérence par rapport aux autres miroirs de ce volume. Au cours du démarrage, le système accède à une seule copie du volume `root` (la copie située sur le disque d'initialisation) tant qu'il ne peut obtenir une configuration complète pour ce volume.

Si le plex de ce volume, qui servait au démarrage, est stale, vous devez redémarrer le système à partir d'un disque d'initialisation de remplacement qui contient des plex non stale. Ce problème peut se produire, par exemple, si vous démarrez le système à partir de l'un des disques rendus amorçables par Volume Manager alors que le disque d'initialisation d'origine est désactivé. Le système démarre normalement, mais les plex qui se trouvent sur le disque hors tension sont stale. Si le système redémarre à partir du disque d'initialisation d'origine alors que le disque est réactivé, il s'initialise en utilisant ce plex stale.

Un autre problème potentiel peut survenir si les erreurs présentes dans les en-têtes de Volume Manager sur le disque d'initialisation empêchent Volume Manager d'identifier correctement le disque. Dans ce cas, Volume Manager ne connaît pas le nom de ce disque. Cela pose un problème car les plex sont associés aux noms de disque, et les plex du disque non identifié sont alors inutilisables.

Un problème peut également se produire lorsque le disque racine enregistre une défaillance qui affecte le plex du volume racine. À la tentative de démarrage suivante, le système s'attend toujours à utiliser le plex racine défectueux pour le démarrage. Si le disque racine a été mis en miroir au moment de la défaillance, vous pouvez spécifier un disque d'initialisation de remplacement (avec un plex racine valide) pour le démarrage.

Dans l'une de ces situations, l'utilitaire `vxconfigd` de Volume Manager note qu'il configure le système dans le cadre du traitement `init` de la séquence de démarrage. `vxconfigd` affiche un message décrivant l'erreur et la solution possible, puis arrête le système. Par exemple, si le plex `rootvol-01` du volume racine `rootvol` sur le disque `rootdisk` est stale, `vxconfigd` peut afficher le message suivant :

```
vxvm:vxconfigd: Warning Plex rootvol-01 for root volume is stale
or unusable.
vxvm:vxconfigd: Erreur : Le disque d'initialisation du système n'a
pas de plex root valide
Démarrez le système depuis l'un des disques suivants :
Disque : disk01                Périphérique : c0t1d0s2
vxvm:vxconfigd: Erreur : Échec du démarrage du système
The system is down.
```

Ce message indique à l'administrateur que le disque d'initialisation de remplacement `disk01` contient une copie utilisable du plex racine et devrait être utilisé pour le démarrage. Lorsque ce message s'affiche, redémarrez le système à partir du disque d'initialisation de remplacement.

Une fois que le système démarre, vous devez déterminer la nature exacte du problème. Si les plex placés sur le disque d'initialisation étaient simplement stale, ils sont actualisés automatiquement lors de la mise en route du système. En revanche, s'il existait un problème avec la zone privée du disque ou que le disque a échoué, vous devez rajouter ou remplacer ce dernier.

Si les plex du disque d'initialisation ne sont pas disponibles, vous devez recevoir des messages des utilitaires Volume Manager pour décrire le problème. Une autre méthode permettant de déterminer le problème consiste à lister les disques au moyen de l'utilitaire `vxdisk`. Dans l'exemple ci-dessus, si le problème est dû à une défaillance dans la zone privée de `rootdisk` (par exemple une défaillance de support ou le remplacement accidentel de la zone privée de Volume Manager sur le disque), `vxdisk list` affiche les informations suivantes :

PÉRIPHÉRIQUE	TYPE	DISQUE	GROUPE	ÉTAT
-	-	rootdisk	rootdg	échec était: c0t3d0s2
c0t1d0s2	sliced	disk01	rootdg	ONLINE

Reconstruction automatique et défaillances du disque d'initialisation

Si le disque d'initialisation (racine) échoue alors qu'il est mis en miroir, la reconstruction automatique tente de remplacer le miroir racine qui a échoué par un nouveau miroir. Pour cela, la reconstruction automatique utilise un miroir restant du disque racine afin de créer un nouveau miroir soit sur un disque de remplacement, soit sur un disque dont l'espace libre est suffisant. Cela assure que le disque racine dispose toujours d'au moins deux miroirs utilisables pour le démarrage. Le démon de reconstruction automatique appelle également l'utilitaire `vxbootsetup`, qui configure le disque en utilisant le nouveau miroir comme disque amorçable.

La reconstruction automatique peut échouer pour un disque racine si le group `disk rootdg` ne contient pas suffisamment d'espace libre ou de remplacement pour contenir les volumes du disque racine qui a échoué. Les volumes `rootvol` et `swapvol` nécessitent un espace contigu sur le disque. Si le volume racine et les autres volumes du disque racine qui a échoué ne peuvent pas être reconstruits sur un nouveau disque identique, vous pouvez reconstruire chacun de ces volumes sur un disque différent.

Les miroirs des volumes `rootvol` et `swapvol` doivent être alignés sur le cylindre de telle sorte qu'ils ne peuvent être créés que sur des disques dont l'espace est suffisant pour permettre aux sous-disques de commencer et de se terminer dans les limites du cylindre. La reconstruction automatique échoue si ces disques ne sont pas disponibles.

Rajout et remplacement de disques d'initialisation

Volume Manager peut uniquement accéder aux données qui ne sont pas critiques pour le démarrage du système lorsque celui-ci est totalement opérationnel ; par conséquent, ces données ne doivent pas se trouver dans des zones spécifiques. Cependant, les données qui sont critiques pour le démarrage doivent se trouver dans des zones spécifiques des disques de démarrage pour que la processus de démarrage puisse les trouver.

Sur certains systèmes, les actions spécifiques exécutées par le contrôleur de disque dans le processus et le BIOS du système limitent l'emplacement de ces données critiques.

Lorsqu'un disque échoue, il existe deux solutions pour corriger le problème :

- ◆ Si les erreurs sont temporaires ou corrigibles, vous pouvez réutiliser le même disque ; cela s'appelle *rajouter* un disque. Dans certains cas, il suffit de reformater un disque qui a échoué ou d'effectuer une analyse de surface pour reconstruire les mappages de secteurs alternatifs afin de créer un disque réutilisable et un candidat au rajout.
- ◆ Si le disque a réellement échoué, vous devez le remplacer.

Les sections suivantes décrivent comment rajouter ou remplacer un disque d'initialisation qui a échoué.

Rajout d'un disque d'initialisation qui a échoué

La procédure de rajout d'un disque est la même que celle de remplacement, à la différence que le même disque physique est utilisé. Généralement, vous devez *détacher* le disque à rajouter. Autrement dit, Volume Manager a détecté la défaillance du disque et n'y accède plus.

Remarque Votre système peut utiliser un *nom de périphérique* ou un *chemin d'accès* différents de ceux donnés dans les exemples. Pour plus d'informations sur les noms de périphériques, voir « Présentation de Volume Manager » à la page 17.

Par exemple, un système comprend deux disques, `disk01` et `disk02`, qui sont normalement mappés dans la configuration système au cours du démarrage, respectivement en tant que disques `c0t0d0s2` et `c0t1d0s2`. `disk01` est détaché à la suite d'une défaillance. Cela peut être confirmé en listant les disques à l'aide de la commande suivante de l'utilitaire `vxdisk` :

```
# vxdisk list
```

`vxdisk` affiche la liste suivante (exemple) :

PÉRIPHÉRIQUE	TYPE	DISQUE	GROUPE	ÉTAT
<code>c0t0d0s2</code>	sliced	-	-	erreur
<code>c0t1d0s2</code>	sliced	<code>disk02</code>	<code>rootdg</code>	en ligne
-	-	<code>disk01</code>	<code>rootdg</code>	échec était: <code>c0b0t0d0s0</code>

Aucun périphérique n'est associé au disque `disk01` dont l'état est échec (failed) avec l'indication du périphérique dont il a été détaché. Il se peut également que le périphérique `c0b0t0d0s0` ne soit pas listé du tout. Cela se produit lorsque le disque échoue entièrement et que le contrôleur de disque ne le détecte pas sur le bus (pour les systèmes qui utilisent un bus).

Dans certains cas, la sortie de `vxdisk list` diffère ; par exemple, si le disque d'initialisation présente des défaillances irrécupérables associées à la table de partition UNIX. Il se peut qu'une partition racine manquante ne puisse pas être réparée et que la zone privée de Volume Manager ne contienne aucune erreur. La commande `vxdisk list` affiche la liste suivante (exemple) :

PÉRIPHÉRIQUE	TYPE	DISQUE	GROUPE	ÉTAT
<code>c0t0d0s2</code>	sliced	<code>disk01</code>	<code>rootdg</code>	en ligne
<code>c0t1d0s2</code>	sliced	<code>disk02</code>	<code>rootdg</code>	en ligne

Cependant, puisque les procédures décrites n'ont pas permis de corriger l'erreur, le disque est considéré comme ayant échoué. Dans ce cas, il doit être détaché manuellement de son périphérique. Pour cela, utilisez la fonction « Supprimer un disque pour le remplacer » de l'utilitaire `vxdiskadm` (pour plus d'informations sur `vxdiskadm`, voir la

Rajout et remplacement de disques d'initialisation

page man `vxdiskadm` (1M)). Une fois que le disque est détaché du périphérique, vous pouvez appliquer toutes les procédures spéciales permettant de corriger le problème (par exemple, reformater le périphérique).

Pour rajouter le disque, utilisez la fonction « Remplacer un disque qui a échoué ou a été supprimé » de l'utilitaire `vxdiskadm` pour remplacer le disque et sélectionnez le *même* périphérique pour le remplacement. En reprenant les exemples précédents, le disque `disk01` doit être remplacé par le périphérique `c0t0d0s2` ou `c0b0t0d0s2` (pour les systèmes qui utilisent un bus).

Si la reconstruction automatique est activée pendant la défaillance d'un disque racine mis en miroir, le système tente de créer un nouveau miroir et de supprimer les sous-disques qui ont échoué du disque d'initialisation. Si le rajout de disque réussit après la reconstruction automatique, le volume racine et/ou les autres volumes affectés par la défaillance du disque n'existent plus sur le disque rajouté. À ce stade, vous pouvez appeler `vxunrelloc` pour déplacer les sous-disques reconstruits automatiquement sur le nouveau disque remplacé.

Rétablissement des sous-disques encapsulés reconstruits sur un nouveau disque

Lorsqu'un disque d'initialisation est encapsulé, le système de fichiers racine et les autres zones système, comme les périphériques de swap sur le disque d'initialisation, sont convertis en volumes. La partie de la zone de swap utilisée par Volume Manager pour créer la zone privée est située au milieu du disque. Lorsqu'un disque est initialisé, et non encapsulé, pour devenir un disque Volume Manager, la zone privée est créée au début du disque.

En cas de défaillance d'un disque d'initialisation encapsulé, tous les sous-systèmes sont reconstruits automatiquement sur les autres disques. Lors de la procédure de reconstruction, le nom du disque d'origine et les décalages sont stockés dans les enregistrements de sous-disque. Lorsque le disque d'initialisation qui a échoué est remplacé par un autre disque qui a la même capacité de stockage, il est « initialisé » et rajouté au `disk group`. Vous pouvez exécuter `vxunrelloc` pour déplacer automatiquement tous les sous-disques sur le disque. Cependant, la différence de structure entre un disque initialisé et un disque encapsulé affecte la façon dont le décalage d'un disque est calculé pour chaque sous-disque reconstruit dans le disque d'origine. Vous devez utiliser l'option `-f` pour déplacer les sous-disques vers le disque, mais pas vers le décalage exact. Si la capacité du disque remplacé dépasse d'au moins 2 Mo celle du disque d'initialisation d'origine, le déplacement de tous les sous-disques vers les décalages exacts du disque peuvent réussir.

`vxunrelloc` rend le nouveau disque amorçable après avoir déplacé tous les sous-disques vers le disque.

Remarque Généralement, un périphérique de vidage est configuré pour constituer la partition swap du disque racine. Lorsqu'un sous-disque de swap est déplacé (par reconstruction automatique, reconstruction dans le disque d'origine ou manuellement) d'un disque à un autre, le périphérique de vidage doit être reconfiguré vers le nouveau disque.

Sur Solaris 2.6 : le périphérique de vidage est stocké dans la structure `dumpfile`. Pour vérifier le paramètre de ce périphérique, entrez la commande suivante :

```
# echo dumpfile+0x10/s | adb -k /dev/ksyms /dev/mem
```

Pour modifier ce paramètre, redémarrez le système afin de configurer le périphérique de vidage sur la première partition swap.

Sur Solaris 2.7 et les versions ultérieures :

Le périphérique de vidage peut être affiché et défini à l'aide de la commande `dumpadm`. Pour plus de détails, entrez :

```
# man dumpadm
```

Remplacement d'un disque d'initialisation qui a échoué

Pour remplacer un disque d'initialisation, vous devez d'abord démarrer le système à partir d'un disque d'initialisation de remplacement. Si le disque qui a échoué n'est pas détaché de son périphérique, vous devez le détacher manuellement à l'aide de la fonction de `vxdiskadm` « Supprimer un disque pour le remplacer ». Pour plus d'informations sur `vxdiskadm`, voir la page `man vxdiskadm (1M)`. Lorsque le disque est détaché, arrêtez le système et remplacez le matériel.

Le disque de remplacement doit offrir une capacité de stockage au moins égale à l'espace occupé sur le disque à remplacer. Il doit être suffisamment grand pour que la zone du disque réservée au stockage des sous-disques puisse contenir tous les sous-disques du disque d'origine à leurs décalages actuels. Pour déterminer la taille minimale d'un disque de remplacement, vous devez déterminer l'espace occupé sur le disque qui a échoué.

Pour déterminer la taille du disque de remplacement, entrez la commande suivante :

```
# vxprint -st -e 'sd_disk="nom_disque"'
```

À partir de la sortie obtenue, ajoutez les valeurs des colonnes `DÉCDISQ` et `LONG` pour le dernier sous-disque indiqué. Le total est exprimé en multiples de 512 octets. Divisez la somme par 2 pour obtenir le total en kilo-octets.

Remarque Généralement, les tailles de disque indiquées par les constructeurs ne reflètent pas la capacité utilisable. De plus, certains constructeurs expriment la capacité en millions d'octets et non en méga-octets, ce qui est différent.

Rattachement des disques

Lorsqu'un disque de remplacement a été trouvé, arrêtez normalement le système et remplacez le matériel voulu. Une fois le remplacement effectué, démarrez le système. Utilisez la fonction « Remplacer un disque qui a échoué ou a été supprimé » de `vxdiskadm` pour remplacer le disque qui a échoué par le nouveau périphérique que vous venez d'ajouter.

Rattachement des disques

Vous pouvez rattacher un disque s'il a totalement échoué et si la reconstruction automatique est impossible, ou si certains pilotes de disque ne sont pas chargés et ne peuvent pas l'être au démarrage de Volume Manager (l'état des disques devient alors FAILED (échoué)). Si le problème est corrigé, vous pouvez utiliser la commande `vxreattach` pour rattacher les disques sans plex marqués comme stale. Cependant, le rattachement doit se produire avant le démarrage des volumes du disque.

La commande `vxreattach` est appelée lors de la récupération des disques à partir des menus `vxdiskadm` et pendant la procédure de démarrage. Si possible, `vxreattach` rattache l'enregistrement du nom logique du disque qui a échoué au disque ayant le même nom de périphérique. Le disque est rattaché dans le même disk group et il conserve son nom d'accès logique au disque.

Après le rattachement, il n'est peut-être pas nécessaire d'effectuer une récupération. Le rattachement peut échouer si le problème à l'origine de la défaillance du disque (ou une autre cause) persiste.

La commande `vxreattach -c` vérifie si un rattachement est possible, mais elle n'exécute pas l'opération. En revanche, elle affiche le nom du disk group et le nom d'accès logique au disque auquel le disque peut être rattaché.

Pour plus d'informations sur la commande `vxreattach`, voir la page man `vxreattach (1M)`.

Récupération de réinstallation

La réinstallation est nécessaire si toutes les copies du disque racine (de démarrage) sont endommagées ou si certains fichiers critiques ont été perdus à la suite de l'endommagement du système de fichiers. Lorsqu'une telle défaillance se produit, vous devez réinstaller entièrement le système, car il n'existe actuellement aucune méthode permettant de restaurer le système de fichiers racine à partir d'une sauvegarde.

Si une défaillance de ce type se produit, essayez de conserver la plus grande partie possible de la configuration Volume Manager d'origine. Vous pouvez enregistrer n'importe quel volume qui n'est pas directement inclus dans la défaillance. Il n'est pas nécessaire de reconfigurer les volumes conservés.

Informations générales sur la réinstallation

Cette section décrit les procédures qui permettent de réinstaller Volume Manager et de conserver le plus possible la configuration d'origine après une défaillance.

La réinstallation du système détruit le contenu de tous les disques utilisés.

Toutes les informations relatives à Volume Manager sont supprimées pendant la réinstallation. Il s'agit des données situées dans les zones privées des disques supprimés, qui contiennent l'identificateur de disque et des copies de la configuration de Volume Manager. La suppression des informations rend le disque inutilisable en tant que disque Volume Manager.

Le disque racine système est toujours utilisé dans le cadre de la réinstallation, mais vous pouvez aussi utiliser d'autres disques. Si le disque racine était placé sous le contrôle de Volume Manager soit pendant l'installation de Volume Manager, soit par une encapsulation ultérieure, ce disque, ainsi que tous les volumes ou miroirs qu'il contient, seront perdus pendant la réinstallation. Tous les autres disques qui sont utilisés lors de la réinstallation ou qui sont supprimés et remplacés peuvent perdre des données de configuration Volume Manager (y compris les volumes et les miroirs).

Si un disque, y compris le disque racine, n'est pas mis sous le contrôle de Volume Manager avant la défaillance, les données de configuration de Volume Manager ne seront pas perdues à la réinstallation. Vous pouvez remplacer les autres disques en suivant les procédures décrites dans la section « Commandes de disque et de disk group » à la page 92.

Si le disque racine n'est pas mis sous le contrôle de Volume Manager, cela simplifie la procédure de récupération à l'issue de la réinstallation, mais cela rend aussi une réinstallation plus probable. Lorsque vous mettez le disque racine sous le contrôle de Volume Manager et que vous créez des miroirs du contenu du disque racine, vous pouvez éliminer de nombreux problèmes qui nécessitent une réinstallation du système.

Lorsque la réinstallation est nécessaire, seuls sont enregistrés les volumes qui se trouvent ou qui possèdent des copies sur les disques qui ne sont pas directement impliqués dans la défaillance et la réinstallation. Tous les volumes, aussi bien sur le disque racine que sur les autres disques impliqués dans la défaillance et/ou la réinstallation, sont perdus pendant la réinstallation. S'il existe des copies de sauvegarde de ces volumes, vous pouvez restaurer ces derniers après la réinstallation. Cependant, sur certains systèmes, les systèmes de fichiers `root`, `stand` et `usr` ne peuvent pas être restaurés à partir d'une sauvegarde.

Procédures de réinstallation et de reconfiguration

Procédez comme suit pour réinstaller le système et récupérer la configuration de Volume Manager. Les diverses étapes sont décrites en détail dans les sections suivantes :

1. Préparez le système pour l'installation.

Cette étape comprend le remplacement des disques ou des composants matériels défectueux ainsi que le détachement des disques qui ne sont pas inclus dans la réinstallation.

2. Installez le système d'exploitation.

Pour cela, réinstallez le système de base ainsi que les autres packages Volume Manager indépendants.

3. Installez Volume Manager.

Ajoutez le package Volume Manager, mais *sans* exécuter la commande `vxinstall`.

4. Récupérez la configuration de Volume Manager.

5. Nettoyez la configuration de Volume Manager.

Cette opération comprend la restauration des informations situées sur les volumes affectés par la défaillance ou la réinstallation, ainsi que la création des volumes système (`rootvol`, `swapvol`, `usr` et autres).

Préparation du système pour la réinstallation

Pour éviter la perte des données situées sur les disques qui ne sont pas inclus dans la réinstallation, vous devez utiliser uniquement le disque racine pour la réinstallation.

Remarque Plusieurs options d'installation *automatiques* permettent d'accéder aux autres disques que le disque racine sans demander confirmation à l'administrateur. Par conséquent, il est recommandé de déconnecter tous les autres disques qui contiennent des volumes avant d'installer le système d'exploitation.

La déconnexion des autres disques garantit qu'ils ne sont pas affectés par la réinstallation. Par exemple, si pour l'installation initiale du système d'exploitation, le système de fichiers `home` se trouvait sur le second disque, vous pouvez toujours le récupérer. La suppression du second disque assure que le système de fichiers `home` restera intact.

Réinstallation du système d'exploitation

Lorsque les disques qui ont échoué ou qui sont en cours d'échec ont été remplacés et que les disques qui ne sont pas inclus dans la réinstallation ont été détachés, réinstallez le système d'exploitation en suivant la procédure décrite dans la documentation du système d'exploitation. Installez le système d'exploitation avant Volume Manager.

Vous devez vous assurer qu'aucun accès n'est effectué sur les autres disques que le disque racine pendant l'installation du système d'exploitation. Si des données sont écrites sur un disque non racine, la configuration de Volume Manager risque d'être détruite sur ce disque.

Remarque Pendant la réinstallation, vous pouvez modifier l'ID (ou le nom) de l'hôte. Il est recommandé de conserver l'ID (ou le nom) de l'hôte existant ; en effet, les sections suivantes considèrent que vous n'avez pas modifié cet ID.

Réinstallation de Volume Manager

L'installation de Volume Manager s'effectue en deux étapes :

- ◆ chargement de Volume Manager à partir du CD-ROM ;
- ◆ initialisation de Volume Manager.

Pour réinstaller Volume Manager, suivez les instructions de chargement de Volume Manager (à partir du CD-ROM) dans le *Guide d'installation de VERITAS Volume Manager*.

Remarque Pour reconstruire la configuration de Volume Manager laissée sur les disques non racines, *n'initialisez pas* Volume Manager (à l'aide de `vxinstall`) après la réinstallation.

Sur certains systèmes, vous pouvez utiliser `vxserial` pour installer la clé de licence Volume Manager (voir la page `man vxserial (1M)`).

Récupération de la configuration de Volume Manager

Une fois que le package Volume Manager est chargé, récupérez la configuration de Volume Manager en procédant comme suit :

1. Arrêtez le système.
2. Rattachez les disques qui ont été supprimés du système.
3. Redémarrez le système.

Récupération de réinstallation

4. Lorsque le système est opérationnel, passez en mode mono-utilisateur en entrant la commande suivante :

```
# shutdown -g0 -is -y
```
5. À l'invite du système, entrez le mot de passe et appuyez sur la touche de retour pour continuer.
6. Supprimez les fichiers relatifs à l'installation qui ont été créés lorsque vous avez chargé Volume Manager et qui ne sont plus nécessaires. Utilisez la commande suivante :

```
# rm -rf /etc/vx/reconfig.d/state.d/install-db
```
7. Une fois les fichiers supprimés, démarrez certains démons d'E/S Volume Manager. Pour cela, entrez la commande suivante :

```
# vxiod set 10
```
8. Démarrez le démon de configuration Volume Manager, `vxconfigd`, en mode désactivé en entrant la commande suivante :

```
# vxconfigd -m disable
```
9. Initialisez le démon `vxconfigd` en entrant la commande suivante :

```
# vxdctl init
```
10. Initialisez le sous-système DMP en entrant la commande suivante :

```
# vxdctl initdmp
```
11. Activez `vxconfigd` en entrant la commande suivante :

```
# vxdctl enable
```

La configuration conservée sur les disques non inclus dans la réinstallation a maintenant été récupérée. Cependant, puisque le disque racine a été réinstallé, Volume Manager ne le considère pas comme un disque Volume Manager. La configuration des disques conservés n'inclut pas le disque racine lors de la configuration de Volume Manager.

Si le disque racine du système et les autres disques inclus dans la réinstallation n'étaient pas sous le contrôle de Volume Manager lors de la défaillance et de la réinstallation, la reconfiguration est terminée. Plusieurs méthodes permettent de remplacer un disque ; choisissez celle que vous préférez.

Si le disque racine (ou un autre disque) était inclus dans la réinstallation, les volumes ou les miroirs qu'il contient (ou les autres disques qui ne sont plus attachés au système) sont désormais inaccessibles. Si un volume ne contenait qu'un plex sur un disque qui a été réinstallé, supprimé ou remplacé, les données de ce volume seront perdues et vous devez les restaurer à partir de la sauvegarde.

De plus, le système de fichiers `root` du système, la zone `swap` (sur certains systèmes, la zone `stand`) et le système de fichiers `/usr` ne se trouvent plus sur les volumes. Pour corriger ces problèmes, suivez les instructions de la section « Nettoyage de la configuration » à la page 265

Vous pouvez démarrer la fonction de reconstruction automatique une fois que `vxdctl enable` réussit, mais l'administrateur doit avoir déterminé avec certitude que ses services, lorsqu'ils sont activés et qu'ils fonctionnent, n'interfèrent pas avec les autres procédures de reconfiguration. Il est recommandé de démarrer la reconstruction automatique à l'issue de l'étape de « reconfiguration finale ». Pour plus d'informations sur cette procédure, voir « Démarrage de la reconstruction automatique » à la page 271.

Nettoyage de la configuration

Les sections suivantes décrivent la procédure à suivre pour nettoyer la configuration du système à l'issue de la réinstallation de Volume Manager.

Les types de nettoyage suivants sont décrits :

- ◆ nettoyage du contrôle racine ;
- ◆ nettoyage des volumes ;
- ◆ nettoyage des disques.

Ces sections sont suivies des informations de reconfiguration suivantes :

- ◆ reconfiguration du contrôle racine ;
- ◆ reconfiguration finale.

Nettoyage du contrôle racine

Pour lancer le nettoyage de la configuration de Volume Manager, supprimez tous les volumes associés au contrôle racine. Cette opération est obligatoire si le disque racine (et les autres disques inclus dans la procédure de démarrage du système) était sous le contrôle de Volume Manager. Vous devez supprimer les volumes suivants :

- ◆ `rootvol`, qui contient le système de fichiers `root` ;
- ◆ `swapvol`, qui contient la zone `swap` ;
- ◆ (sur certains systèmes) `standvol`, qui contient le système de fichiers `ficstand` ;
- ◆ `usr`, qui contient le système de fichiers `/usr`.

Récupération de réinstallation

Pour supprimer le volume racine, utilisez la commande `vxedit` :

```
# vxedit -fr rm rootvol
```

Répétez cette commande, en substituant `swapvol` et `usr (standvol)` à `rootvol`, pour supprimer les volumes `swap`, `stand` et `usr`.

Nettoyage des volumes

Une fois que vous avez nettoyé le contrôle racine, vous devez déterminer les volumes à restaurer à partir de la sauvegarde. Il s'agit des volumes dont tous les miroirs (toutes les copies du volume) se trouvent sur des disques qui ont été réinstallés ou supprimés. Ces volumes ne sont pas valides et doivent être supprimés, puis recréés et restaurés à partir d'une sauvegarde. Si seuls certains miroirs d'un volume se trouvent sur des disques réinitialisés ou supprimés, vous devez les supprimer. Vous pourrez les ajouter de nouveau par la suite.

Pour restaurer les volumes, procédez comme suit :

1. Déterminez les disques VM qui ont été supprimés ou réinstallés en entrant la commande suivante :

```
# vxdisk list
```

Volume Manager affiche la liste des périphériques de disque du système et leur état. Par exemple, pour un système réinstallé de trois disques et un disque racine réinstallé, la commande `vxdisk list` affiche ce qui suit :

PÉRIPHÉRIQUE	TYPE	DISQUE	GROUPE	ÉTAT
c0t0d0s2	partitioned	-	-	erreur
c0t1d0s2	partitioned	disk02	rootdg	en ligne
c0t2d0s2	partitioned	disk03	rootdg	en ligne
-	-	disk01	rootdg	échec était:c0t0d0s2

Remarque Le système peut utiliser un nom de périphérique différent de ceux donnés dans les exemples. Pour plus d'informations sur les noms de périphériques, voir « Présentation de Volume Manager » à la page 17.

L'affichage indique que le périphérique racine réinstallé, `c0t0d0s2`, n'est pas associé à un disque VM et que son état est `erreur`. `disk02` et `disk03` n'ont pas été inclus dans la réinstallation ; ils sont reconnus par Volume Manager et associés à leur périphérique (`c0t1d0s2` et `c0t2d0s2`). L'ancien `disk01`, qui correspondait au disque VM associé au périphérique de disque remplacé, n'est plus associé au périphérique (`c0t0d0s2`).

Si d'autres disques (contenant des volumes ou des miroirs) ont été supprimés ou remplacés pendant la réinstallation, ces disques correspondront également à un périphérique de disque listé avec l'état *erreur* et à un disque VM listé comme n'étant pas associé à un périphérique.

2. Une fois que vous savez quels disques ont été supprimés ou remplacés, recherchez tous les miroirs sur les disques qui ont échoué. Entrez la commande suivante :

```
# vxprint -sF "%vname" -e'sd_disk = "disque"'
```

où *disque* représente le nom d'un disque dont l'état est échec. Veillez à placer le nom du disque entre guillemets dans la commande, faute de quoi elle renverra un message d'erreur. La commande `vxprint` affiche la liste des volumes qui ont des miroirs sur le disque qui a échoué. Répétez cette commande pour tous les disques dont l'état est échec.

3. Vérifiez l'état de chaque volume. Utilisez la commande suivante pour imprimer les informations sur le volume :

```
# vxprint -th nom_volume
```

où *nom_volume* représente le nom du volume à examiner. La commande `vxprint` affiche l'état du volume, ses plex et les portions des disques qui composent ces plex. Par exemple, le volume `v01` n'a qu'un seul plex sur le disque réinstallé `disk01`. La commande `vxprint -th v01` affiche les informations suivantes :

V	NOM	TYPEUTIL	KSTATE	ÉTAT	LONG	RDPOL
PLEX	PRÉF					
PL	NOM	VOLUME	KSTATE	ÉTAT	LONG	STRUCT
COL/LARG	MODE					
SD	NOM	PLEX	DISQUE	DÉCDISQ	LONG	[COL/] DÉC
PÉRIPH	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				

L'unique plex du volume est indiqué sur la ligne commençant par `pl`. Le champ *ÉTAT* spécifique au plex `v01-01` est *NODEVICE*. Le plex a de l'espace sur un disque qui a été remplacé, supprimé ou réinstallé. Il n'est plus valide et doit être supprimé.

Récupération de réinstallation

Puisque v01-01 était le seul plex du volume, le contenu de ce dernier est également irrécupérable, sauf si vous restaurez le volume à partir d'une sauvegarde. Vous devez également supprimer le volume. S'il existe une copie de sauvegarde du volume, vous pouvez restaurer ce dernier par la suite. Notez le nom du volume et sa longueur ; ces informations seront nécessaires pour la sauvegarde.

4. Pour supprimer le volume v01, utilisez la commande `vxedit` :

```
# vxedit -r rm v01
```

Il se peut que seule une partie d'un plex se trouve sur le disque qui a échoué. Si un plex réparti par bandes est associé au volume, ce dernier est divisé sur plusieurs disques. Par exemple, le volume v02 a un plex réparti par bandes sur trois disques, dont l'un est le disque réinstallé `disk01`. La commande `vxprint -th v02` affiche les informations suivantes :

V	NOM	TYPEUTIL	KSTATE	ÉTAT	LONG	RDPOL		
PLEXPRÉF	PL	NOM	VOLUME	KSTATE	ÉTAT	LONG	STRUCT	
COL/LARG	MODE	SD	NOM	PLEX	DISQUE	DÉCDISQ	LONG	[COL/] DÉC
PÉRIPH	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		
	clt2d0	ENA						
sd	disk01-05	v02-01	disk01	620544	10240	1/0		
	clt2d1	DIS						
sd	disk03-01	v02-01	disk03	620544	10240	2/0		
	clt2d2	ENA						

L'affichage indique trois disques sur lesquels le plex v02-01 est réparti (les lignes commençant par `sd` représentent les bandes). L'une des zones de bande est située sur un disque qui a échoué. Ce disque n'est plus valide et le plex v02-01 est nommé `NODEVICE`. Puisqu'il s'agit du seul plex du volume, celui-ci n'est pas valide et vous devez le supprimer. S'il existe une copie de v02 sur le support de sauvegarde, vous pouvez le restaurer par la suite. Notez le nom du volume ainsi que la longueur des volumes que vous voulez restaurer depuis une sauvegarde.

- Utilisez la commande `vxedit` pour supprimer le volume en procédant comme décrit précédemment.

Un volume avec un miroir sur un disque qui a échoué peut également avoir d'autres miroirs sur des disques qui sont toujours valides. Dans ce cas, il est inutile de restaurer le volume à partir de la sauvegarde, puisque les données sont toujours correctes sur les disques valides.

La sortie de la commande `vxprint -th` pour un volume avec un plex situé sur un disque qui a échoué (`disk01`) et un autre plex sur un disque valide (`disk02`) est du type suivant :

V	NOM	TYPEUTIL	KSTATE	ÉTAT	LONG	RDPOL		
PLEXPRÉF	PL	NOM	VOLUME	KSTATE	ÉTAT	LONG	STRUCT	
COL/LARG	MODE	SD	NOM	PLEX	DISQUE	DÉCDISQ	LONG	[COL/] DÉC
PÉRIPH	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		
c1t3d0	ENA							
pl	v03-02	v03	DISABLED	NODEVICE	30720	CONCAT		
-		RW						
sd	disk01-04	v0302	disk03	262144	30720	0		
c1t2d2	DIS							

Ce volume a deux plex, `v03-01` et `v03-02`. Vous pouvez utiliser le premier plex (`v03-01`) car il n'utilise pas d'espace sur le disque non valide. Le second plex (`v03-02`) utilise de l'espace sur le disque non valide `disk01` et son état est `NODEVICE`. Il doit être supprimé. Cependant, le volume a toujours un plex valide contenant des données correctes. Si vous devez mettre le volume en miroir, vous pourrez ajouter un autre plex par la suite. Notez le nom du volume pour créer un autre plex par la suite.

- Pour supprimer un plex non valide, vous devez dissocier le plex du volume, puis le supprimer à l'aide de la commande `vxp1ex`. Pour supprimer le plex `v03-02`, entrez la commande suivante :

```
# vxplex -o rm dis v03-02
```

- Une fois que tous les volumes ont été nettoyés, nettoyez la configuration du disque, comme décrit dans la section « Nettoyage des disques » à la page 270.

Récupération de réinstallation

Nettoyage des disques

Une fois que vous avez supprimé tous les volumes et les plex non valides, vous pouvez nettoyer la configuration de disque. Vous devez supprimer de la configuration tous les disques qui ont été supprimés, réinstallés ou remplacés (comme déterminé par la sortie de la commande `vxdisk list`).

Pour supprimer le disque, utilisez la commande `vxdbg`. Pour supprimer le disque qui a échoué `disk01`, entrez la commande suivante :

```
# vxdbg rmdisk disk01
```

Si la commande `vxdbg` renvoie un message d'erreur, cela indique qu'il existe des miroirs non valides. Répétez les procédures décrites dans la section « Nettoyage des volumes » à la page 266 autant de fois que nécessaire pour supprimer tous les volumes et les miroirs non valides.

Reconfiguration du contrôle racine

Une fois que vous avez supprimé tous les disques non valides, vous pouvez ajouter les disques de remplacement ou ajouter les disques réinstallés sous le contrôle de Volume Manager. Si le disque était à l'origine sous le contrôle de Volume Manager ou que vous voulez maintenant placer le disque racine sous le contrôle de Volume Manager, commencez par ajouter ce disque.

Pour ajouter le disque racine sous le contrôle de Volume Manager, utilisez Volume Manager Support Operations (`vxdiskadm`). Entrez la commande suivante :

```
# vxdiskadm
```

Dans le menu principal de `vxdiskadm`, sélectionnez l'option de menu 2 (Encapsuler un ou plusieurs disques). Suivez les instructions et encapsulez le disque racine pour le système.

À l'issue de l'encapsulation, redémarrez le système en mode multi-utilisateur.

Reconfiguration finale

Une fois que le disque racine est encapsulé, vous devez ajouter tous les autres disques qui ont été remplacés à l'aide de `vxdiskadm`. Si les disques ont été réinstallés en même temps que le système d'exploitation, vous devez les encapsuler ; dans le cas contraire, vous pouvez les ajouter.

Une fois que vous avez ajouté tous les disques au système, vous pouvez recréer les volumes qui ont été complètement supprimés lors du nettoyage de la configuration et restaurer leur contenu à partir de la sauvegarde. Pour recréer les volumes, vous pouvez utiliser `vxassist` ou l'interface utilisateur graphique.

Vous pouvez recréer les volumes v01 et v02 à l'aide de la commande `vxassist` suivante :

```
# vxassist make v01 24000
# vxassist make v02 30720 layout=stripe nstripe=3
```

Une fois les volumes créés, vous pouvez les restaurer à partir de la sauvegarde en utilisant les procédures normales de sauvegarde/restauration.

Si des plex ont été supprimés de certains volumes lors de leur nettoyage, vous pouvez recréer ces miroirs en suivant les instructions de mise en miroir d'un volume à l'aide de `vxassist`, comme décrit dans ce manuel. Pour remplacer le plex supprimé du volume v03 à l'aide de `vxassist`, entrez la commande suivante :

```
# vxassist mirror v03
```

Une fois que vous avez restauré les volumes et les plex perdus pendant la réinstallation, la récupération est terminée et le système devrait être configuré tel qu'il était avant la défaillance.

Démarrage de la reconstruction automatique

À ce stade, l'administrateur doit redémarrer le système ou lancer manuellement la reconstruction automatique (si nécessaire). Ces deux opérations activent le démon de reconstruction (ainsi que son processus `vxnotify`).

Pour démarrer la reconstruction automatique, entrez les commandes suivantes.

Démarrez le démon de surveillance. Un message électronique est alors envoyé à l'administrateur dès qu'un problème est détecté. Pour modifier l'adresse à laquelle les rapports de problème seront envoyés, remplacez l'argument de `vxrelocd` :

```
# nohup /usr/lib/vxvm/bin/vxrelocd root &
```

La commande suivante est également valide :

```
# nohup /usr/lib/vxvm/bin/vxrelocd root > /dev/null 2>&1 &
```

La commande suivante peut détecter si la reconstruction automatique a démarré :

```
# ps -ef | grep vxrelocd | grep -v grep
```

États des plex et des volumes

Les sections suivantes décrivent les états des plex et des volumes.

États des plex

Les états des plex indiquent si ceux-ci sont des copies complètes et cohérentes (miroirs) du contenu des volumes. Les utilitaires Volume Manager actualisent automatiquement l'état des plex. Cependant, vous pouvez modifier leur état si les modifications apportées au volume auquel ce plex est associé ne doivent pas être écrites. Par exemple, si un disque avec un plex spécifique enregistre une défaillance, vous pouvez le désactiver temporairement.

Remarque Un plex ne doit pas obligatoirement être associé à un volume. Il peut être créé à l'aide de la commande `vxmake plex`. Un tel plex peut ensuite être attaché à un volume.

Les utilitaires Volume Manager utilisent les états des plex pour :

- ◆ indiquer si le contenu d'un volume a été initialisé dans un état connu ;
- ◆ déterminer si un plex contient une copie valide (miroir) du contenu du volume ;
- ◆ vérifier si un plex était utilisé au cours d'une défaillance du système ;
- ◆ surveiller les opérations effectuées sur les plex.

Cette section présente en détail les états des plex ; elle s'adresse aux administrateurs qui souhaitent acquérir une connaissance approfondie de ces états.

Les plex associés à un volume peuvent avoir l'un des états suivants :

- ◆ EMPTY (vide)
- ◆ CLEAN (valide)
- ◆ ACTIVE (actif)
- ◆ STALE (non valide)
- ◆ OFFLINE (hors ligne)
- ◆ TEMP (temporaire)
- ◆ TEMPRM (supprimé temporairement)
- ◆ TEMPRMSD (sous-disque supprimé temporairement)
- ◆ IOFAIL (erreur d'E/S)

Le journal des zones modifiées (DRL, *Dirty Region Logging*) ou le plex de journal RAID 5 constitue un cas spécial car son état est toujours LOG (journal).

État des plex EMPTY

La création de volumes attribue l'état EMPTY à tous les plex associés au volume pour indiquer que ce plex n'est pas encore initialisé.

État des plex CLEAN

L'état d'un plex est CLEAN lorsqu'il contient une copie cohérente (miroir) du contenu du volume qu'une opération a désactivé. Par conséquent, lorsque tous les plex d'un volume sont nettoyés, aucune action n'est requise pour garantir que les plex sont identiques au démarrage de ce volume.

État des plex ACTIVE

L'état d'un plex peut être ACTIVE dans deux cas :

- ◆ lorsque, après le démarrage du volume, le plex participe pleinement aux E/S normales de volume (le contenu du plex change avec celui du volume) ;
- ◆ lorsque le volume a été arrêté par un arrêt anormal du système, à un moment où le plex était ACTIVE.

Dans le second cas, une défaillance du système peut entraîner une certaine incohérence dans le contenu du plex. Au démarrage d'un volume, Volume Manager exécute la récupération afin de garantir que les plex marqués ACTIVE auront exactement le même contenu.

Remarque Sur un système fonctionnant sans problème, l'état ACTIVE est affiché pour la plupart des plex de volume.

État des plex STALE

Si un plex ne contient pas le contenu complet et actuel d'un volume, son état devient STALE. De plus, si une erreur d'E/S se produit sur un plex, le noyau cesse d'utiliser et d'actualiser son contenu et une opération définit son état comme étant STALE.

L'opération `vplex att` récupère le contenu d'un plex STALE à partir d'un plex ACTIVE. Les opérations de copie atomique copient le contenu du volume sur les plex dont l'état est STALE. L'administrateur système peut forcer un plex à devenir STALE à l'aide d'une opération `vplex det`.

État des plex OFFLINE

La commande `vxmend off task` détache indéfiniment un plex d'un volume en lui attribuant l'état OFFLINE. Bien que le plex détaché reste associé au volume, les modifications apportées à ce dernier se sont pas reflétées sur le plex OFFLINE. En effet,

États des plex et des volumes

pour pouvoir être actualisé, ce plex doit être mis en ligne et rattaché à l'aide de la commande `vxplex att task`. Dans ce cas, l'état de ce plex devient STALE et son contenu est récupéré lors de l'opération `vxvol start` suivante.

État des plex TEMP

Lorsque l'état d'un plex devient TEMP, cela facilite certaines opérations de plex qui ne pourraient pas être exécutées de façon véritablement atomique. Par exemple, lors de l'attachement d'un plex à un volume activé, vous devez copier le contenu de ce dernier sur le plex pour qu'il puisse être considéré comme réellement attaché.

Un utilitaire définit l'état du plex comme TEMP au début d'une telle opération et lui attribue l'état approprié à la fin de l'opération. En cas de défaillance du système, l'état TEMP du plex indique que l'opération n'a pas été terminée. Une opération `vxvol start` ultérieure dissocie les plex dont l'état est TEMP.

État des plex TEMPRM

L'état TEMPRM d'un plex est similaire à l'état TEMP, sauf que le plex TEMPRM est supprimé à l'issue de l'opération. Certaines opérations de sous-disque nécessitent un plex temporaire. Par exemple, l'association d'un sous-disque à un plex nécessite que vous mettiez à jour le sous-disque avec le contenu du volume avant de réellement associer le sous-disque. Lors de cette mise à jour, le sous-disque doit être associé à un plex temporaire marqué TEMPRM jusqu'à la fin de l'opération et jusqu'à ce que le plex TEMPRM soit supprimé.

En cas de défaillance du système, l'état TEMPRM indique que l'opération n'a pas été terminée. Une opération ultérieure dissocie et supprime les plex TEMPRM.

État des plex TEMPRMSD

`vxassist` utilise l'état TEMPRMSD pour attacher de nouveaux plex. Si l'opération ne se termine pas, le plex et ses sous-disques sont supprimés.

État des plex IOFAIL

L'état IOFAIL du plex est associé à un enregistrement permanent de l'état. Dès que la défaillance d'un plex ACTIVE est détectée, `vxconfigd` lui attribue l'état IOFAIL pour l'exclure du processus de sélection lors de la récupération au démarrage du volume.

Cycle des états des plex

La modification des états des plex fait partie des opérations normales. Les changements de l'état d'un plex indiquent des anomalies que Volume Manager doit corriger. Au démarrage du système, les volumes démarrent automatiquement et la tâche `vxvol start` fait passer tous les plex de ACTIVE à CLEAN. Si tout se passe bien jusqu'à l'arrêt du système, l'opération d'arrêt du volume marque comme CLEAN tous les plex ACTIVE et le cycle continue. Si tous les plex sont CLEAN au démarrage (avant que `vxvol start` ne leur attribue l'état ACTIVE), cela indique un arrêt normal et optimise le démarrage.

État du noyau du plex

L'état du noyau du plex indique l'accessibilité du plex. Cet état est surveillé dans le pilote de volume et il permet d'attribuer à un plex un mode de fonctionnement hors ligne (DISABLED), de maintenance (DETACHED) ou en ligne (ENABLED).

Le noyau du plex peut avoir l'un des états suivants :

- ◆ DISABLED — Le plex est inaccessible.
- ◆ DETACHED — Les écritures sur le volume ne sont pas reflétées sur le plex. Une requête de lecture à partir du volume ne se reflète pas sur le plex. Les opérations de plex et les fonctions `ioctl` sont acceptées.
- ◆ ENABLED — Une demande d'écriture sur le volume est reflétée sur le plex. Une requête de lecture à partir du volume est exécutée depuis le plex.

Remarque Aucune intervention de l'utilisateur n'est requise pour définir ces états ; ils sont actualisés selon une procédure interne. Lorsqu'un système fonctionne correctement, tous les plex sont activés.

États des volumes

Les volumes peuvent avoir différents états, dont certains sont similaires à ceux des plex :

- ◆ CLEAN — Le volume ne démarre pas (l'état du noyau est DISABLED) et ses plex sont synchronisés.
- ◆ ACTIVE — Le volume démarre (l'état du noyau est actuellement ENABLED) ou il était utilisé (l'état du noyau était ENABLED) lorsque le système a été réinitialisé. Si l'état du volume est actuellement ENABLED, l'état de ses plex à tout moment est incertain (puisque le volume est en cours d'utilisation). Si l'état du volume est actuellement DISABLED, cela signifie qu'il est impossible de garantir la cohérence des plex, mais qu'elle le sera au démarrage du volume.

États des plex et des volumes

- ◆ **EMPTY** — Le contenu du volume n'est pas initialisé. L'état du noyau est toujours **DISABLED** lorsque le volume est **EMPTY**.
- ◆ **SYNC** — Soit le volume est en mode de récupération lecture-réécriture (l'état du noyau est actuellement **ENABLED**), soit il était en mode de récupération lecture-réécriture lorsque le système a redémarré (l'état du noyau est **DISABLED**). Avec la récupération lecture-réécriture, la cohérence des plex est récupérée en lisant les données à partir des blocs d'un plex et en les écrivant sur tous les autres plex en écriture. Si l'état du volume est **ENABLED**, cela signifie que les plex sont resynchronisés à l'aide d'une lecture-réécriture. Si l'état du volume est **DISABLED**, cela signifie que les plex étaient resynchronisés à l'aide d'une lecture-réécriture lorsque le système a redémarré et que par conséquent, ils doivent être synchronisés.
- ◆ **NEEDSYNC** — Le volume doit être resynchronisé lors de son prochain démarrage.

L'interprétation de ces indicateurs pendant le démarrage du volume est modifiée par le journal d'état permanent du volume (par exemple, l'indicateur **DIRTY/CLEAN**). Si l'indicateur **CLEAN** est activé, aucun volume avec l'état **ACTIVE** n'a été écrit par un processus ni même ouvert au redémarrage ; par conséquent, il peut être considéré comme **CLEAN**. Dans tous les cas, l'indicateur **CLEAN** est toujours défini lorsque le volume est marqué **CLEAN**.

États des volumes RAID 5

Les volumes RAID 5 ont leurs propres états :

- ◆ **CLEAN** — Le volume ne démarre pas (l'état du noyau est **DISABLED**) et sa parité est correcte. Les plex RAID 5 sont cohérents.
- ◆ **ACTIVE** — Le volume démarre (l'état du noyau est actuellement **ENABLED**) ou il était utilisé (l'état du noyau était **ENABLED**) lorsque le système a été redémarré. Si l'état du volume est actuellement **ENABLED**, l'état de ses plex RAID 5 à tout moment est incertain (puisque le volume est en cours d'utilisation). Si l'état du volume est actuellement **DISABLED**, la synchronisation de la parité ne peut pas être garantie.
- ◆ **EMPTY** — Le contenu du volume n'est pas initialisé. L'état du noyau est toujours **DISABLED** lorsque le volume est **EMPTY**.
- ◆ **SYNC** — Soit la parité du volume fait l'objet d'une resynchronisation (l'état du noyau est actuellement **ENABLED**), soit elle était en cours de resynchronisation lorsque le système a été redémarré (l'état du noyau est **DISABLED**).
- ◆ **NEEDSYNC** — Le volume doit être resynchronisé lors de son prochain démarrage.
- ◆ **REPLAY** — Le volume est dans un état temporaire lors de la relecture du journal. Celle-ci se produit lorsqu'il devient nécessaire d'utiliser la parité et les données enregistrées dans le journal.

État du noyau du volume

L'*état du noyau du volume* indique l'accessibilité du volume. Cet état permet d'attribuer à un volume un mode de fonctionnement hors ligne (DISABLED), de maintenance (DETACHED) ou en ligne (ENABLED).

Le noyau de volume peut avoir l'un des états suivants :

- ◆ DISABLED — Le volume est inaccessible.
- ◆ DETACHED — Il est impossible de lire ou d'écrire sur le volume, mais les opérations de périphérique de plex et les fonctions ioctl sont acceptées.
- ◆ ENABLED — Il est possible de lire ou d'écrire sur le volume.

Récupération des volumes RAID 5

Cette section décrit le fonctionnement et la récupération des volumes RAID 5. Pour plus d'informations sur les volumes RAID 5, voir « Présentation de Volume Manager » à la page 17.

Structure des volumes RAID 5

Un volume RAID 5 comprend un ou plusieurs plex, contenant chacun un ou plusieurs sous-disques. Contrairement aux volumes mis en miroir, les plex d'un volume RAID 5 ne servent pas tous à conserver une copie miroir des données du volume.

Un volume RAID 5 peut avoir deux types de plex :

- ◆ le *plex RAID 5* permet de conserver aussi bien les données que la parité pour le volume :
- ◆ les *plex de journal* conservent les fichiers journaux des données écrites sur le volume afin d'accélérer et d'optimiser la récupération.

Récupération des volumes RAID 5

Plex RAID 5

Les volumes RAID 5 conservent aussi bien les données que les informations de parité dans un seul plex RAID 5. Celui-ci comprend des sous-disques organisés en colonnes, selon le modèle de répartition des données par bandes sur plusieurs disques. L'affichage ci-dessous représente un plex RAID 5, ainsi que la sortie de vxprint qui lui est associée :

PL NOM	VOLUME	KSTATE	ÉTAT	LONG	STRUCT
COL/LARG	MODE				
SD NOM	PLEX	DISQUE	DÉCDISQ	LONG	[COL/] DÉC
PÉRIPH	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				

La ligne de plex indique que la structure du plex est RAID, avec trois colonnes et une unité de bande de 16 secteurs. Chaque ligne de sous-disque indique la colonne dans le plex et le décalage dans la colonne dans laquelle il se trouve.

Remarque Votre système peut utiliser un *nom de périphérique* différent de ceux donnés dans les exemples. Voir « Présentation de Volume Manager » à la page 17 pour plus d'informations sur les noms de périphériques.

Journaux RAID 5

Chaque volume RAID 5 a un plex RAID 5 contenant les données et la parité. Tous les autres plex associés au volume sont utilisés pour enregistrer les informations sur les données et la parité écrites sur le volume. Ces plex sont appelés *plex de journal RAID 5* ou *journaux RAID 5*.

Les journaux RAID 5 peuvent être des plex concaténés ou répartis par bandes, et chaque journal RAID 5 associé à un volume RAID 5 contient une copie complète des informations de journal associées au volume. Il est recommandé d'avoir au moins deux plex RAID 5 pour chaque volume RAID 5. Ces plex de journal doivent être situés sur des disques différents. L'existence de deux plex de journal RAID 5 pour chaque volume RAID 5 offre une protection contre la perte d'informations de journal due à la défaillance d'un seul disque.

Pour gérer l'accès simultané à l'ensemble de disques RAID 5, la taille du journal doit être plusieurs fois celle de la bande du plex RAID 5.

Vous pouvez distinguer un plex de journal RAID 5 d'un plex RAID 5 d'un volume RAID 5 en examinant la sortie `vxprint`. Le champ ÉTAT pour un plex de journal indique LOG. L'affichage suivants représente la sortie de la commande `vxprint` pour un volume RAID 5.

```

V  NOM          TYPEUTIL  KSTATE  ÉTAT    LONG    RDPOL
PLEXPRÉF
PL  NOM          VOLUME    KSTATE  ÉTAT    LONG    STRUCT
COL/LARG  MODE
SD  NOM          PLEX      DISQUE  DÉCDISQ LONG    [COL/]DÉC
PÉRIPH    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

```

Le volume RAID 5 (`r5vol`) peut être identifié en tant que tel car sa politique de lecture est RAID. Son plex RAID 5 (`r5vol-01`) est similaire à celui décrit précédemment. Il comprend deux journaux RAID 5 dans les plex `r5vol-11` et `r5vol-12` qui sont identifiés par le contenu du champ ÉTAT (LOG). Ils sont associés à un volume RAID 5 et leur structure n'est pas RAID.

Récupération des volumes RAID 5

Création de volumes RAID 5

Pour créer des volumes RAID 5, vous pouvez utiliser `vxassist` (recommandé) ou `vxmake`. Ces deux options sont décrites dans cette section.

Un volume RAID 5 a un plex RAID 5 qui comprend au moins deux sous-disques situés sur au moins deux disques physiques. Il ne peut y avoir qu'un seul plex RAID 5 par volume.

vxassist et volumes RAID 5

Vous pouvez créer un volume RAID 5 à l'aide de la commande `vxassist` suivante :

```
# vxassist make nom_volume longueur layout=raid5
```

Par exemple, pour créer un volume 10M RAID 5 nommé `volraid`, entrez :

```
# vxassist make volraid 10m layout=raid5
```

Cette opération crée un volume RAID 5 avec une taille d'unité de bande par défaut sur le nombre par défaut de disques.

Initialisation des volumes RAID 5

Vous devez initialiser un volume RAID 5 si vous l'avez créé au moyen de `vxmake` et que vous ne l'avez pas encore initialisé, ou si son état a été défini comme non initialisé.

Vous pouvez initialiser un volume RAID 5 avec `vxvol` en entrant l'une des commandes suivantes :

```
# vxvol init zero nom_volume
```

ou

```
# vxvol start nom_volume
```

La commande `vxvol init zero` écrit des zéros sur tous les plex de journal RAID 5 et sur la longueur totale du volume. Elle attribue ensuite l'état `ACTIVE` au volume.

`vxvol start` récupère la parité en appliquant XOR aux unités de bande de données correspondantes dans toutes les autres colonnes. Même si elle est plus lente qu'une opération `vxvol init zero`, l'opération `vxvol start` rend le volume RAID 5 immédiatement disponible.

Défaillances et volumes RAID 5

Il existe deux sortes de défaillances : les *défaillances du système* et les *défaillances de disque*. Dans le cas d'une défaillance du système, le système cesse immédiatement de fonctionner à la suite d'un arrêt de l'alimentation électrique ou d'un problème relatif au système d'exploitation. Dans le cas des défaillances de disque, les données situées sur certains disques deviennent inaccessibles à cause d'une défaillance du système (par exemple, une panne de tête, un problème électronique sur le disque ou une défaillance du contrôleur de disque).

Défaillances du système

Les volumes RAID 5 sont conçus pour rester disponibles, tout en occupant un espace disque minimum en cas de défaillance de disque. Cependant, les différents types de volume RAID 5 peuvent faire l'objet de perte de données en cas de défaillance du système. La perte de données peut se produire lorsque la défaillance du système désynchronise les données et la parité dans le volume RAID 5. La perte de synchronisation survient parce qu'il est impossible de déterminer l'état des écritures en cours au moment de la défaillance.

Si une perte de synchronisation se produit pendant l'accès à un volume RAID 5, celui-ci est décrit comme ayant une *parité stale*. La parité doit alors être reconstruite en lisant toutes les colonnes de non parité dans chaque bande, en recalculant la parité et en écrivant l'unité de bande de parité dans la bande. Cette opération doit être effectuée pour chaque bande du volume, ce qui peut prendre un certain temps.

Attention Pendant l'exécution de cette resynchronisation, toute défaillance de l'un des disques de l'ensemble de disques entraîne la perte des données du volume. Cela ne concerne que les volumes RAID 5 *sans* plex de journal.

La resynchronisation rend le système plus vulnérable aux pannes et peut également ralentir son fonctionnement en utilisant une partie de ses ressources.

Les journaux RAID 5 réduisent l'endommagement dû aux défaillances du système car ils conservent une copie des données écrites au moment de la panne. La resynchronisation consiste à lire les données et la parité à partir des journaux et de les écrire dans les zones appropriées du volume RAID 5. Cela contribue à réduire considérablement la durée de resynchronisation des données et de la parité. Cela signifie également que le volume ne devient jamais véritablement stale. La défaillance d'un seul disque n'entraîne pas la perte des données du volume car les données et la parité pour toutes les bandes du volume sont connues à tout moment.

Récupération des volumes RAID 5

Défaillance d'un disque

La défaillance d'un disque peut rendre ses données non disponibles. Pour un volume RAID 5, cela signifie qu'un sous-disque n'est plus disponible.

Ce problème peut survenir à la suite d'une erreur d'E/S incorrigible lors d'une écriture sur le disque. Cette erreur d'E/S peut détacher le sous-disque de l'ensemble de disques ou peut empêcher tout accès au disque lors du démarrage du système (par exemple, à la suite d'un problème de câblage ou de la mise hors tension d'un lecteur).

Dans ce cas, le sous-disque ne peut plus conserver de données et il est considéré comme étant à la fois STALE (non valide) et DETACHED (détaché). Si le disque sous-jacent devient disponible ou est remplacé, le sous-disque est toujours considéré comme stale et il n'est pas utilisé.

Si le système essaie de lire les données situées sur un sous-disque stale, ces données seront reconstruites à partir de toutes les autres unités de bande. Cette opération est appelée *lecture de reconstruction*. Elle dépasse une simple lecture et peut réduire les performances de lecture. Lorsque certains sous-disques d'un volume RAID 5 sont stale, celui-ci est considéré comme étant en *mode dégradé*.

Vous pouvez reconnaître un volume RAID 5 en mode dégradé à partir des informations données par `vxprint`, comme illustré dans l'affichage suivant :

Récupération des volumes RAID 5

V	NOM	TYPEUTIL	KSTATE	ÉTAT	LONG	RDPOL
PLEXPRÉF	PL NOM	VOLUME	KSTATE	ÉTAT	LONG	STRUCT
COL/LARG	MODE	PLEX	DISQUE	DÉCDISQ	LONG	[COL/] DÉC
SD NOM	PÉRIPH	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	-					

Le volume `r5vol` est en mode dégradé, comme l'indique le champ `ÉTAT` (`DEGRADED`). Le sous-disque qui a échoué est `disk01-00`, comme l'indiquent les indicateurs de la dernière colonne (le paramètre `d` indique que le sous-disque est détaché et le paramètre `S`, que son contenu est stale).

Il se peut qu'un disque contenant un journal RAID 5 présente une défaillance. Ce problème n'a aucune incidence directe sur le fonctionnement du volume. Cependant, la perte de tous les journaux RAID 5 sur un volume rend ce dernier vulnérable à une défaillance totale. Dans la sortie de `vxprint -ht`, la défaillance d'un plex de journal RAID 5 est indiquée par l'état du plex `BADLOG`. Cela est illustré par l'affichage suivant, dans lequel le plex de journal RAID 5 `r5vol-11` a échoué :

Récupération des volumes RAID 5

V	NOM	TYPEUTIL	KSTATE	ÉTAT	LONG	RDPOL
PLEXPRÉF	PL NOM	VOLUME	KSTATE	ÉTAT	LONG	STRUCT
COL/LARG	MODE	SD NOM	PLEX	DISQUE	DÉCDISQ	LONG
PÉRIPH	MODE					[COL/] DÉC
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					

Récupération RAID 5

Les types de récupération suivants sont généralement nécessaires pour les volumes RAID 5 :

- ◆ resynchronisation de parité ;
- ◆ récupération des sous-disques stale ;
- ◆ récupération des plex de journal.

Ces types de récupération sont décrits dans les sections suivantes. La resynchronisation de parité et la récupération des sous-disques stale s'exécutent normalement :

- ◆ au démarrage du volume RAID 5 ;
- ◆ peu après le démarrage du système ;
- ◆ en appelant la commande `vxrecover`.

Pour plus d'informations sur le démarrage des volumes RAID 5, voir « Démarrage des volumes RAID 5 » à la page 289.

Si la reconstruction automatique est activée lors d'une défaillance de disque, aucune intervention n'est requise de l'administrateur système, sauf s'il n'existe aucun espace disque approprié disponible pour la reconstruction. La reconstruction automatique est déclenchée par la défaillance et l'administrateur est informé de cette défaillance par courrier électronique.

La reconstruction automatique tente de reconstruire les sous-disques d'un plex RAID 5 qui a échoué. À l'issue de toute reconstruction, le démon de reconstruction automatique (`vxrelocd`) lance également une resynchronisation de parité.

Dans le cas d'un plex de journal RAID 5 qui a échoué, la reconstruction se produit uniquement si le plex de journal est mis en miroir ; `vxrelocd` lance alors une resynchronisation de miroir afin de recréer le plex de journal RAID 5. Si la reconstruction automatique est désactivée au moment de la défaillance, l'administrateur système devra peut-être lancer une resynchronisation ou une récupération.

Récupération de parité

Le plus souvent, un ensemble de disques RAID 5 n'a pas une parité stale. Celle-ci ne se produit qu'après la défaillance de tous les plex de journal RAID 5 du volume RAID 5, et uniquement en cas de défaillance du système. Même si un volume RAID 5 n'a pas une parité stale, ce problème est normalement corrigé au démarrage du volume.

Si un volume sans journal RAID 5 valide est démarré et que la procédure est arrêtée avant la resynchronisation du volume, il en résulte un volume actif avec une parité stale. Cela peut être confirmé en vérifiant l'état du volume affiché dans la sortie de la commande `vxprint -ht`, comme illustré dans l'affichage suivant.

```

V  NOM          TYPEUTIL  KSTATE  ÉTAT      LONG      RDPOL
PLEXPRÉF
PL  NOM          VOLUME    KSTATE  ÉTAT      LONG      STRUCT
COL/LARG  MODE
SD  NOM          PLEX      DISQUE  DÉCDISQ  LONG      [COL/] DÉC
PÉRIPH    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

```

Récupération des volumes RAID 5

Ces informations indiquent que l'état du volume est NEEDSYNC, c'est-à-dire qu'il faut resynchroniser la parité. Cet état aurait pu aussi être SYNC et indiquer que le système avait tenté une synchronisation au démarrage et qu'une procédure de synchronisation devrait être exécutée. S'il n'existe aucune procédure de ce type ou si l'état du volume est NEEDSYNC, vous pouvez lancer manuellement une synchronisation en utilisant le mot-clé `resync` pour la commande `vxvol`. Par exemple, pour resynchroniser le volume RAID 5 de l'affichage précédent, entrez la commande suivante :

```
# vxvol resync r5vol1
```

Pour régénérer la parité, exécutez la commande `ioctl VOL_R5_RESYNC` sur le volume RAID 5. La resynchronisation commence au début du volume RAID 5 et resynchronise une zone égale au nombre de secteurs spécifié par l'option `-o iosize`. Si vous ne spécifiez pas cette option, le système utilise la taille d'E/S maximale par défaut. La resynchronisation passe alors à la zone suivante, et ainsi de suite jusqu'à ce que la longueur totale du volume RAID 5 ait été resynchronisée.

Plus le volume est important, plus la régénération de la parité sera longue. Il se peut que le système s'arrête ou tombe en panne avant la fin de l'opération. Dans le premier cas, la progression de la régénération de parité doit être conservée d'une réinitialisation à l'autre. Dans tous les autres cas, vous devez redémarrer l'opération depuis le début.

Pour éviter de redémarrer la procédure, la régénération de la parité est dotée d'un point de reprise. Cela signifie que le décalage jusqu'au niveau duquel la parité a été régénérée est enregistré dans la base de données de configuration. L'option `-o checkpoint=size` contrôle la fréquence d'enregistrement du point de reprise. Si elle n'est pas spécifiée, le système utilise la taille du point de reprise par défaut.

La sélection d'une taille trop petite pour le point de reprise peut prolonger la durée de la régénération de la parité car l'enregistrement du décalage du point de reprise nécessite une transaction. Lorsque le système a redémarré, un volume RAID 5 avec un décalage de point de reprise inférieur à la longueur du volume lance une resynchronisation de la parité au décalage du point de reprise.

Récupération des sous-disques

La récupération des sous-disques stale s'effectue généralement au démarrage du volume. Cependant, il se peut que la procédure de récupération échoue ou que le volume soit démarré avec une option empêchant la récupération des sous-disques. Il se peut également que le disque contenant le sous-disque ait été remplacé sans aucune opération de récupération. Dans tous les cas, vous pouvez exécuter une récupération des sous-disques en utilisant le mot-clé `recover` de la commande `vxvol`. Par exemple, pour récupérer le sous-disque stale du volume RAID 5 illustré à la figure 26 à la page 290, entrez la commande suivante :

```
# vxvol recover r5vol1 disk01-00
```

Il se peut que le volume RAID 5 contienne plusieurs sous-disques stale à traiter en même temps. Lorsque vous utilisez `vxvol recover` avec uniquement le nom du volume, vous sélectionnez les différents sous-disques stale :

```
# vxvol recover r5vol
```

Récupération des journaux après les défaillances

Les défaillances de disque risquent de détacher les plex de journal RAID 5. Vous pouvez rattacher ces journaux RAID 5 en utilisant le mot-clé `att` de la commande `vxplex`. Pour rattacher le plex de journal RAID 5 illustré à la figure 27 à la page 295, entrez la commande suivante :

```
# vxplex att r5vol r5vol-11
```

Opérations RAID 5 diverses

Plusieurs opérations permettent de manipuler les volumes RAID 5 et leurs objets associés. Ces opérations sont normalement effectuées par d'autres commandes telles que `vxassist` et `vxrecover` dans le cadre d'opérations de plus grande envergure, comme le déplacement du contenu des disques. Généralement, ces opérations de ligne de commande ne devraient pas être nécessaires pour les opérations plus limitées de Volume Manager.

Manipulation des journaux RAID 5

Les journaux RAID 5 sont représentés sous la forme de plex de volumes RAID 5 et ils sont manipulés à l'aide de la commande `vxplex`. Vous pouvez ajouter un journal RAID 5 à l'aide de `vxplex att` :

```
# vxplex att r5vol r5log
```

L'attachement ne peut se poursuivre que si la taille du nouveau journal est suffisamment grande pour contenir toutes les données de la bande. Si le volume RAID 5 contient déjà des journaux, la longueur du nouveau journal correspond à la longueur minimum de chaque journal. En effet, le nouveau journal est un miroir des anciens journaux.

Si le volume RAID 5 n'est pas activé, le nouveau journal est marqué `BADLOG` et il est activé au démarrage du volume. Cependant, le contenu du journal est ignoré.

Si le volume RAID 5 est activé et comprend d'autres journaux RAID 5 activés, le contenu du nouveau journal est synchronisé avec les autres journaux à l'aide des commandes `ioctl ATOMIC_COPY`.

Opérations RAID 5 diverses

Si aucun journal n'est actuellement activé dans le volume RAID 5, le nouveau journal est remis à zéro avant d'être activé.

Vous pouvez supprimer les plex de journal d'un volume en entrant la commande `vxplex dis` suivante :

```
# vxplex dis r5log3
```

Si la suppression du journal laisse moins de deux journaux valides sur le volume, un avertissement s'affiche et l'opération ne peut pas continuer. Vous devez forcer l'opération à l'aide de l'option `-o force`.

Manipulation des sous-disques RAID 5

À l'instar des autres sous-disques, les sous-disques du plex RAID 5 d'un volume RAID 5 sont manipulés à l'aide de la commande `vxsd`. Pour effectuer l'association, vous devez utiliser le mot-clé `assoc` de la même façon que pour les plex répartis par bande. Par exemple, pour ajouter des sous-disques à la fin de chaque colonne du volume RAID 5 utilisé dans l'affichage précédent, entrez la commande suivante :

```
# vxsd assoc r5vol-01 disk10-01:0 disk11-01:1 disk12-01:2
```

Si un sous-disque remplit un « espace » dans le plex (c'est-à-dire, une partie de l'espace d'adressage logique du volume est mappée par le sous-disque), le sous-disque est considéré comme stale. Si le volume RAID 5 est activé, l'opération d'association régénère les données situées sur le sous-disque à l'aide des commandes `ioctl VOL_R5_RECOVER`. Dans le cas contraire, il est marqué comme stale et il est récupéré au démarrage du volume.

Vous pouvez supprimer les sous-disques du plex RAID 5 à l'aide de `vxsd dis` :

```
# vxsd dis disk10-01
```

Attention Si le sous-disque mappe une partie de l'espace d'adressage du volume RAID 5, le volume passe en mode DÉGRADÉ. Dans ce cas, l'opération `dis` affiche un avertissement et vous devez la forcer à l'aide de l'option `-o force`. De même, si la suppression du sous-disque rend le volume RAID 5 inutilisable parce qu'un autre sous-disque de la même bande est inutilisable ou manquant et que le volume n'est ni vide ni avec l'état `DISABLED`, cette opération n'est pas autorisée.

Vous pouvez utiliser `vxsd mv` pour déplacer les sous-disques afin de modifier les disques occupés par un volume RAID 5. Par exemple, si vous devez déplacer le contenu de `disk03` et que `disk22` a suffisamment d'espace sur deux de ses parties, vous pouvez entrer la commande suivante :

```
# vxsd mv disk03-01 disk22-01 disk22-02
```

Cette commande est similaire à celle utilisée pour les plex répartis par bandes, mais le mécanisme sous-jacent ne l'est pas.

Déplacements des sous-disques RAID 5

Pour déplacer des sous-disques RAID 5, vous devez supprimer le sous-disque actuel du plex RAID 5 et vous devez le remplacer par les nouveaux sous-disques. Les nouveaux sous-disques sont marqués comme *stale*, puis récupérés à l'aide des opérations `VOL_R5_RECOVER`. La récupération s'effectue via `vxsd` ou (si le volume n'est pas actif) au démarrage du volume. *Cela signifie que le volume RAID 5 est dégradé pour toute la durée de l'opération.*

Une autre défaillance des bandes utilisées dans le déplacement rend le volume inutilisable. Il se peut aussi que le volume RAID 5 ne soit plus valide si sa parité devient *stale*.

Pour éviter ce type de situation, l'utilitaire `vxsd` empêche tout déplacement des sous-disque si :

- ◆ un sous-disque *stale* et le sous-disque en cours de déplacement occupent la même bande :
- ◆ le volume RAID 5 est arrêté anormalement (la parité est considérée comme *stale*) ;
- ◆ le volume RAID 5 est actif et n'a plus aucune zone de journal valide.

L'option `-o force` s'applique uniquement au troisième cas.

Vous pouvez également diviser et joindre les volumes RAID 5 à l'aide de `vxsd split` et de `vxsd join`. Ces opérations fonctionnent de la même façon que pour les volumes mis en miroir.

Remarque Les sous-disques RAID 5 sont déplacés de la même façon que les autres sous-disques, sans que la redondance ne soit dégradée.

Démarrage des volumes RAID 5

Au démarrage, le volume RAID 5 peut avoir divers états. Après l'arrêt normal du système, l'état du volume doit être `CLEAN` et ne doit pas nécessiter de récupération. Cependant, si le volume n'était pas fermé ou s'il n'était pas monté avant l'arrêt anormal du système, vous devrez peut-être effectuer une récupération au démarrage, pour le rendre disponible. Cette section décrit les actions que vous pouvez prendre dans certaines conditions.

Dans ces conditions normales, les volumes démarrent automatiquement après un redémarrage et la récupération est effectuée automatiquement ou à l'aide de la commande `vxrecover`.

Volumes RAID 5 impossibles à démarrer

Un volume RAID 5 est inutilisable si une partie du plex RAID 5 ne mappe pas la longueur du volume :

- ◆ Le plex RAID 5 ne peut pas être clairsemé par rapport à la longueur du volume RAID 5.
- ◆ Le plex RAID 5 ne mappe pas une zone dans laquelle deux sous-disques ont échoué dans une même bande soit parce qu'ils sont stales, soit parce qu'ils sont construits sur un disque qui a échoué.

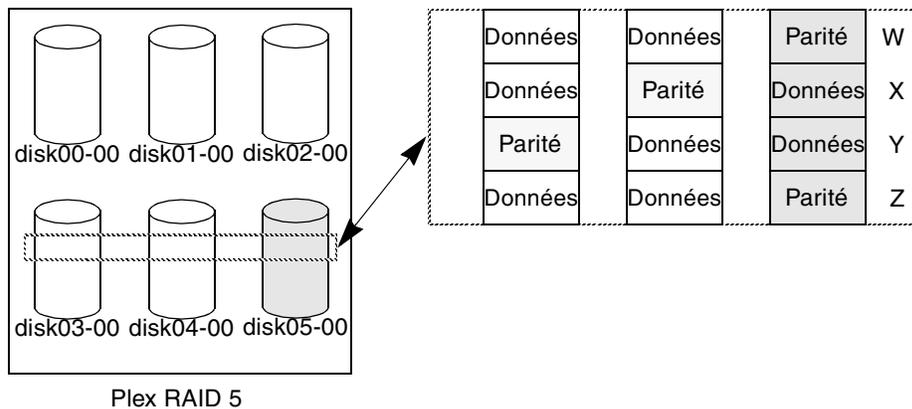
Dans ce cas, la commande `vxvol start` renvoie le message d'erreur suivant :

```
vxvm:vxvol: ERREUR : Le volume r5vol ne peut pas être démarré ;
le plex RAID 5 ne mappe pas la longueur entière du volume.
```

À ce stade, le contenu du volume RAID 5 est inutilisable.

Il existe une autre raison pour laquelle le volume RAID 5 ne peut pas être démarré : la parité est stale et un sous-disque devient détaché ou stale. Ce problème se produit lorsque l'unité de bande de parité n'est pas valide dans les bandes contenant le sous-disque qui a échoué (parce que la parité est stale) *et* lorsque l'unité de bande sur le sous-disque qui a échoué n'est également pas valide. Cette situation est illustrée à la figure 26, « Volume RAID 5 non valide », qui représente un volume RAID 5 qui n'est plus valide à cause d'une parité stale et d'un sous-disque qui a échoué.

Figure 26. Volume RAID 5 non valide



Cet exemple illustre quatre bandes dans l'ensemble de disques RAID 5. La parité est stale et le sous-disque `disk05-00` a échoué. Par conséquent, les bandes X et Y sont inutilisables puisque deux défaillances se sont produites dans ces bandes.

Cela correspond à deux défaillances dans une même bande et rend le volume inutilisable. Dans ce cas, la commande `vxvol start` affiche la sortie suivante :

```
vxvm:vxvol: ERREUR : Le volume r5vol ne peut pas être démarré ;
certains disques sont inutilisables et la parité est stale.
```

Vous pouvez éviter cette situation en utilisant *toujours* au moins deux plex de journal RAID 5 dans les volumes RAID 5. Avec les plex de journal RAID 5, la parité du volume ne peut plus devenir stale. Pour plus d'informations, voir « Défaillances du système » à la page 281.

Démarrage forcé de volumes RAID 5

Vous pouvez démarrer un volume même si les sous-disques sont marqués comme stale. Par exemple, un volume arrêté a une parité stale et aucun journal RAID 5 et un disque est détaché, puis rattaché.

Le sous-disque est considéré comme stale même si les données sont toujours actuelles (car le volume était utilisé lorsque le sous-disque n'était pas disponible) et le volume RAID 5 est considéré comme non valide. Pour éviter cette situation, veillez à toujours associer plusieurs journaux RAID 5 valides à l'ensemble de disques. Cependant, cette solution ne sera pas toujours possible.

Pour démarrer un volume RAID 5 avec des sous-disques stale, vous pouvez utiliser l'option `-f` avec la commande `vxvol start`. Tous les sous-disques stale sont ainsi marqués comme non stale. Ce marquage est effectué avant que l'opération `start` n'évalue la validité du volume RAID 5 et les conditions nécessaires à son démarrage. De plus, vous pouvez marquer des sous-disques individuels comme non stale en utilisant la commande `vxmend fix unstale sous-disque`.

Récupération au démarrage des volumes RAID 5

Plusieurs opérations peuvent se révéler nécessaires pour restaurer entièrement le contenu d'un volume RAID 5 et le rendre utilisable. Lorsque vous démarrez un volume, les plex de journal RAID 5 sont remis à zéro avant le démarrage proprement dit. Cela permet d'éviter que des données aléatoires ne soient interprétées comme une entrée de journal et endommagent le contenu du volume. De plus, vous devrez peut-être récupérer des sous-disques ou resynchroniser la parité (si les journaux RAID 5 ont échoué).

Les étapes suivantes sont effectuées au démarrage d'un volume RAID 5 :

1. Si le volume RAID 5 n'a pas été arrêté normalement, le système vérifie qu'il contient des plex de journal RAID 5 valides.
 - ◆ S'il existe des plex de journal valides, ils sont relus. Pour cela, le noyau du volume est placé dans l'état DETACHED et le volume, dans l'état REPLAY (relecture). Les plex de journal RAID 5 sont ensuite activés. Si le système peut lire les journaux et que la lecture réussit, passez à l'étape 2.
 - ◆ S'il n'existe pas de plex de journal valide, vous devez resynchroniser la parité. Pour cela, placez le noyau du volume dans l'état DETACHED et le volume dans l'état SYNC. Les plex de journal conservent l'état DISABLED.

Le volume n'est pas disponible pendant la resynchronisation de la parité parce que les défaillances de sous-disque qui se produisent pendant ce temps rendent le volume inutilisable. Pour remédier à ce problème, utilisez l'option `-o unsafe start` avec `vxvol`. S'il existe des sous-disques stale, le volume RAID 5 est inutilisable.

Attention L'option `-o unsafe start` est considérée comme dangereuse car elle peut rendre le contenu du volume inutilisable. Son utilisation est de ce fait déconseillée.

2. Les plex de journal existants sont remis à zéro et activés. Si tous les journaux échouent pendant cette opération, le démarrage est abandonné.
3. S'il n'existe pas de sous-disque stale ou que ceux présents sont récupérables, le noyau du volume est mis dans l'état ENABLED (activé) et le volume, dans l'état ACTIVE. Le volume est alors démarré.
4. Si certains sous-disques sont stale et doivent être récupérés, et qu'il existe des journaux valides, le volume doit être activé en mettant son noyau dans l'état ENABLED ; il pourra alors être utilisé pendant la récupération des sous-disques. Dans le cas contraire, le noyau du volume est mis dans l'état DETACHED et le volume ne sera pas disponible pendant la récupération des sous-disques.

Cela se produit lorsque le système s'arrête anormalement ou que le volume a été arrêté incorrectement pendant qu'il était actif : la parité devient alors stale et le volume inutilisable. Si cette situation ne convient pas, vous pouvez démarrer le volume à l'aide de l'option `-o unsafe start`.

Attention L'option `-o unsafe start` est considérée comme dangereuse, puisqu'elle peut rendre le contenu du volume inutilisable. Son utilisation est donc déconseillée.

5. Le volume est mis dans l'état RECOVER (récupération) et les sous-disques stale sont restaurés. Le sous-disque n'est plus marqué comme stale car les données situées sur les différents sous-disques sont à présent valides.
6. Si la récupération d'un sous-disque échoue et qu'aucun journal n'est valide, le démarrage du volume est abandonné car les sous-disques restent stale et qu'un arrêt anormal du système rend le volume RAID 5 inutilisable. Vous pouvez aussi corriger ce problème à l'aide de l'option `-o unsafe start`.

Attention L'option `-o unsafe start` est considérée comme dangereuse, puisqu'elle peut rendre le contenu du volume inutilisable. Son utilisation est donc déconseillée.

Si le volume a des journaux valides, les défaillances de récupération des sous-disques sont enregistrées, mais elles n'arrêtent pas la procédure de démarrage.

7. Une fois que tous les sous-disques ont été récupérés, le noyau du volume est mis dans l'état ENABLED et le volume est marqué comme ACTIVE. Il est alors démarré.

Modification des attributs d'un volume RAID 5

Vous pouvez modifier plusieurs attributs d'un volume RAID 5 tels que la longueur du volume et celle du journal RAID 5 à l'aide de la commande `vxvol set`. Pour modifier la longueur d'un volume RAID 5, entrez la commande suivante :

```
# vxvol set len=10240 r5vol
```

La longueur d'un volume ne peut pas dépasser la zone mappée (appelée la *zone contiguë*, ou *contiglen*) du plex RAID 5. Vous ne pouvez pas étendre la longueur afin de rendre le volume inutilisable. Si le volume RAID 5 est actif et que vous réduisez sa longueur, vous devez forcer l'opération au moyen de l'option `-o force usage`. Cela permet d'éviter la suppression de l'espace à partir des applications qui utilisent le volume.

Récupération au démarrage des volumes RAID 5

Vous pouvez aussi modifier la longueur des journaux RAID 5 en utilisant `vxvol set` avec la commande suivante :

```
# vxvol set loglen=2M r5vol
```

Veillez noter que les plex de journal RAID 5 ne sont valides que s'ils mappent entièrement la longueur du journal du volume RAID 5. Si l'augmentation de la longueur du journal rend l'un des journaux RAID 5 non valide, l'opération n'est pas autorisée. De plus, si le volume n'est pas actif et qu'il est modifié (il n'a pas été arrêté normalement), vous ne pouvez pas modifier la longueur du journal. Cela évite la perte du contenu des journaux (si la longueur du journal est réduite) ou l'introduction de données aléatoires dans les journaux (si la longueur est augmentée).

Écriture dans les ensembles de disques RAID 5

Cette section décrit la procédure d'écriture pour les ensembles de disques RAID 5.

Lecture-Modification-Écriture

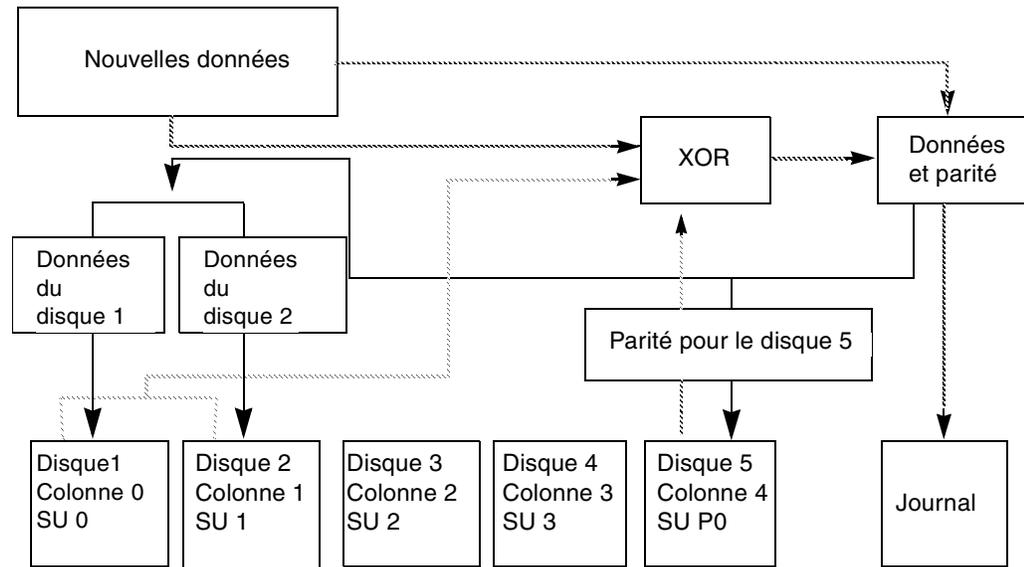
Lorsque vous écrivez dans un ensemble de disques RAID 5, vous pouvez effectuer les étapes suivantes pour chaque bande incluse dans les E/S :

1. Le système accède aux unités de bande de données à mettre à jour avec les nouvelles données écrites et les place dans des tampons internes. L'unité de bande de parité est lue dans des tampons internes.
2. La parité est mise à jour pour refléter le contenu de la nouvelle zone de données. Dans un premier temps, le contenu des anciennes données est soumis à une opération OR exclusive (XOR) avec la parité (suppression logique des anciennes données). Les nouvelles données sont alors soumises à l'opération XOR dans la parité (ajout logique des nouvelles données). Les nouvelles données et la nouvelle parité sont écrites dans un journal.
3. La nouvelle parité est écrite dans l'unité de bande de parité. Les nouvelles données sont écrites dans les unités de bande de données. Les unités de bande sont toutes écrites au cours d'une même opération.

Cette procédure est un cycle *lecture-modification-écriture*, qui est le type d'écriture par défaut pour RAID 5. Si un disque échoue, aussi bien les unités de bande de données que les unités de bande de parité ne sont plus disponibles. On dit alors que l'ensemble de disques fonctionne en mode *dégradé*.

La séquence lecture-modification-écriture est illustrée à la figure 27, « Lecture-Modification-Écriture ».

Figure 27. Lecture-Modification-Écriture



SU = Unité de bande
 = Étape 1 : Lit les données (à partir de l'unité de bande de parité P0 et les unités de bande de données 0 et 1).
 - - - - - = Étape 2 : Exécute des XOR entre les données et la parité pour calculer la nouvelle parité. Enregistre dans le journal les nouvelles données et parité.
 ————— = Étape 3 : Écrit la nouvelle parité (résultant de XOR) sur l'unité de bande de parité P0 et les nouvelles données sur les unités de bande de données 0 et 1.

Récupération au démarrage des volumes RAID 5

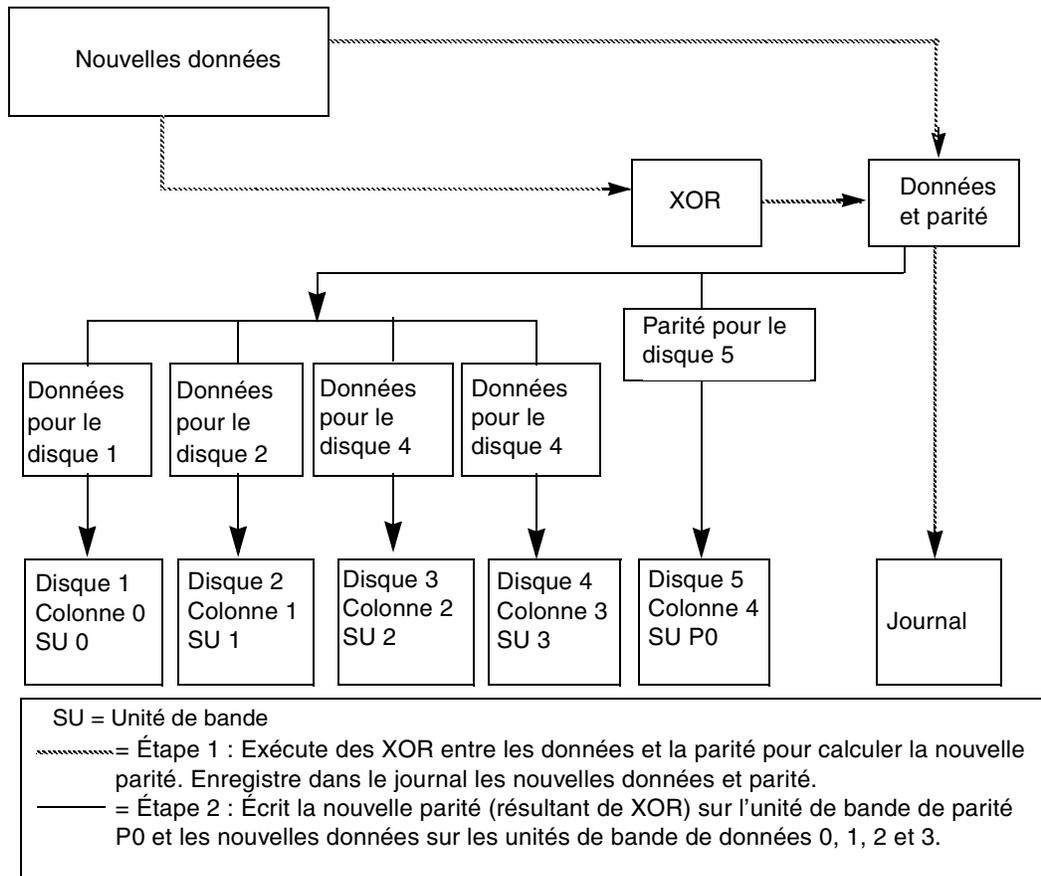
Écriture de bande complète

En cas d'écritures volumineuses (des écritures couvrant entièrement la bande de données), vous pouvez ne pas suivre la procédure de lecture-modification-écriture et utiliser à la place celle d'*écriture de bande complète*. Une écriture de bande complète est plus rapide qu'une opération lecture-modification-écriture parce qu'elle ne nécessite pas l'exécution de la lecture. L'élimination du cycle de lecture réduit la durée des E/S nécessaires à l'écriture sur le disque. Une procédure d'écriture de bande complète comprend les étapes suivantes :

1. L'opération XOR est appliquée sur toutes les nouvelles unités de bande de données, ce qui génère une nouvelle valeur de parité. Les nouvelles données et parité sont écrites dans un journal.
2. La nouvelle parité est écrite dans une unité de bande de parité. Les nouvelles données sont écrites dans les unités de bande de données. La bande entière est écrite au cours d'une seule opération.

La figure 28, « Écriture de bande complète » illustre une écriture de bande complète.

Figure 28. Écriture de bande complète



Récupération au démarrage des volumes RAID 5

Reconstruction-Écriture

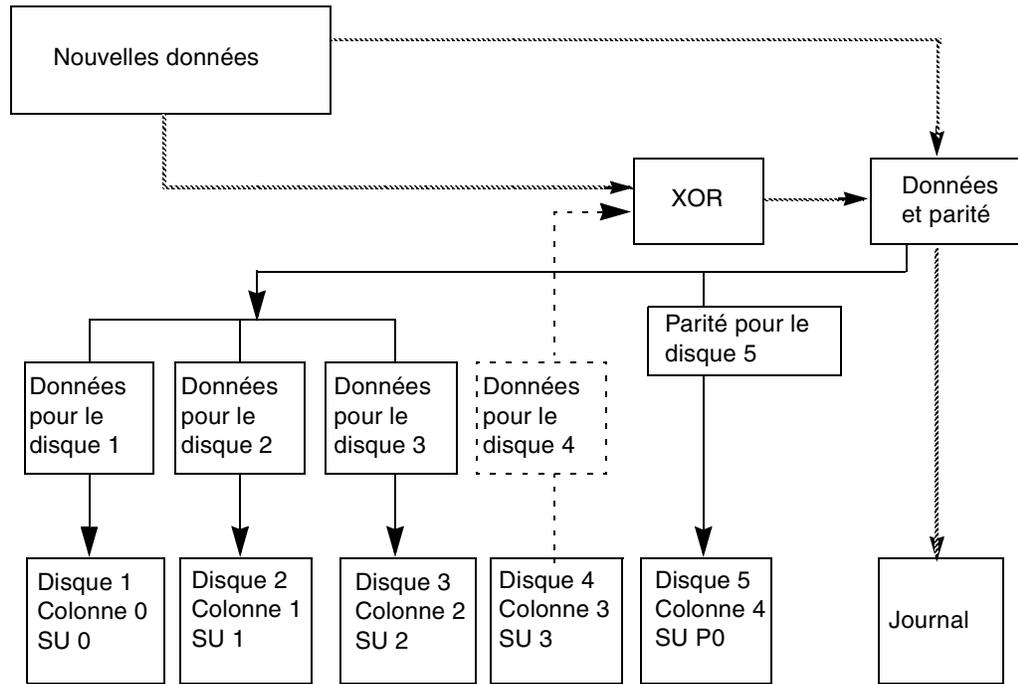
Lorsque au moins 50 % des disques de données sont écrits lors d'une même E/S, vous pouvez utiliser une *reconstruction-écriture*. Une reconstruction-écriture enregistre la durée d'E/S par XOR. L'opération XOR ne nécessite pas la lecture de la zone de parité, mais uniquement la lecture des données non affectées, qui correspondent à moins de la moitié des unités de bande.

Une reconstruction-écriture comprend les étapes suivantes :

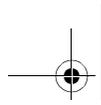
1. Les données non affectées sont lues à partir des unités de bande de données non modifiées.
2. L'opération XOR est effectuée à la fois sur les nouvelles et les anciennes données non affectées afin de générer une nouvelle unité de bande de parité. Les nouvelles données et la parité résultante sont enregistrées dans le journal.
3. La nouvelle parité est écrite dans l'unité de bande de parité. Les nouvelles données sont écrites dans les unités de bande de données. Toutes les unités de bande sont écrites lors d'une même opération.

La figure 29, « Reconstruction-Écriture » illustre une reconstruction-écriture. Cette opération est préférable dans ce cas à une procédure de lecture-modification-écriture parce qu'elle lit uniquement les disques de données nécessaires, au lieu des disques et du disque de parité.

Figure 29. Reconstruction-Écriture



SU = Unité de bande
 = Étape 1 : Écrit les données à partir de l'unité de bande de données non affectée 3.
 = Étape 2 : Exécute des XOR entre les nouvelles et les anciennes données non affectées. Enregistre dans le journal les nouvelles données et parité.
 ——— = Étape 3 : Écrit la nouvelle parité (résultant de XOR) sur l'unité de bande de parité P0 et les nouvelles données sur les unités de bande de données 0, 1 et 2.



Récupération au démarrage des volumes RAID 5



Contrôle des performances de VxVM

8

Introduction

La gestion logique des volumes est un outil qui permet d'améliorer les performances générales du système. Ce chapitre contient des recommandations pour la configuration et à la gestion des performances qui pourront vous aider à tirer parti des avantages de Volume Manager. Il contient des informations qui vous permettront de définir les priorités voulues pour les performances et il décrit les méthodes pour obtenir et utiliser les données appropriées.

Ce chapitre traite des sujets suivants :

- ◆ Recommandations pour les performances
- ◆ Contrôle des performances
- ◆ Réglage de Volume Manager

Recommandations pour les performances

Volume Manager rend la configuration du stockage plus souple, ce qui permet d'améliorer les performances du système. Il existe deux stratégies de base pour optimiser les performances :

- ◆ attribuer des données aux disques physiques pour répartir régulièrement la charge d'E/S entre les disques disponibles ;
- ◆ identifier les données les plus souvent accédées et augmenter la largeur de bande d'accès à ces données au moyen de la répartition des données par bandes sur plusieurs disques et de la mise en miroir.

Volume Manager assure également la redondance des données (via la mise en miroir et RAID 5) qui permet un accès continu aux données en cas de défaillance d'un disque.

Attribution des données

Lorsqu'un administrateur système décide de l'emplacement des systèmes de fichiers, il tente généralement d'équilibrer la charge entre les disques disponibles. L'efficacité de cette approche peut être limitée par le fait qu'il est extrêmement difficile d'anticiper les modèles d'utilisation future et qu'il est impossible de diviser des systèmes de fichiers entre plusieurs lecteurs. Par exemple, si un seul système de fichiers enregistre la plupart des accès aux disques, l'installation de ce système de fichiers sur un autre disque ne fait que déplacer le goulot d'étranglement vers ce disque.

Vous pouvez obtenir un meilleur niveau de granularité pour le placement des données car Volume Manager peut diviser des volumes entre plusieurs disques. Après avoir examiné les modèles d'accès réels, l'administrateur système peut affiner ses décisions de placement des systèmes de fichiers. Il peut reconfigurer en ligne les volumes après avoir défini ou modifié les modèles de performance, sans que cela n'affecte leur disponibilité.

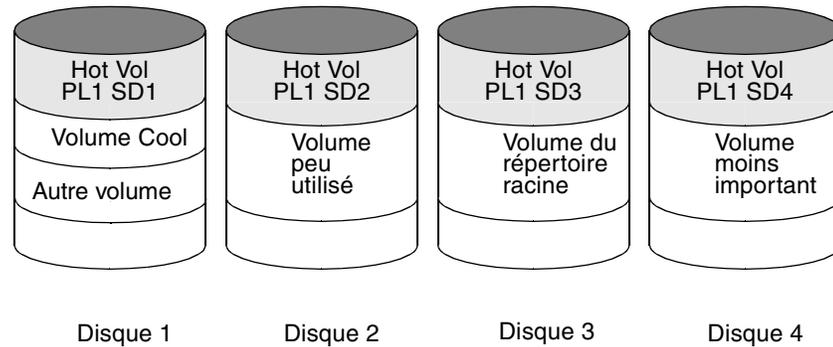
Répartition des données par bandes sur plusieurs disques

La répartition des données par bandes sur plusieurs disques permet de « partitionner » les données et de les stocker sur plusieurs périphériques afin d'améliorer les performances d'accès. Elle accroît la largeur de bande d'accès pour un plex. Les performances d'accès des plex répartis par bandes sont améliorées pour les opérations de lecture et d'écriture.

L'identification des volumes qui enregistrent le plus grand nombre d'accès (ils contiennent des systèmes de fichiers ou des bases de données) vous permet d'améliorer les performances du système. Vous pouvez augmenter la largeur de bande d'accès de ces données en répartissant par bandes les données à « trafic élevé » sur des portions de plusieurs disques.

La figure 30, « Utilisation de la répartition des données par bande sur plusieurs disques pour optimiser l'accès aux données », illustre l'exemple d'un volume (HOT VOL) identifié comme créant un goulot d'étranglement pour l'accès aux données. Ce volume est réparti par bandes sur quatre disques et l'espace restant sur ces disques peut être utilisé par les volumes moins sollicités.

Figure 30. Utilisation de la répartition des données par bande sur plusieurs disques pour optimiser l'accès aux données



Mise en miroir

La mise en miroir est une technique qui permet de stocker plusieurs copies des données sur un système. Lorsqu'elle est correctement appliquée, elle peut rendre les données constamment disponibles en assurant une protection contre les risques de perte de données dus aux défaillances des supports physiques. L'utilisation de la mise en miroir améliore les possibilités de récupération des données en cas de panne du système ou de défaillance d'un disque.

La mise en miroir permet aussi parfois d'améliorer les performances du système. La mise en miroir des données les plus sollicitées permet non seulement de protéger ces données contre tout risque de perte dû à la défaillance d'un disque, mais aussi d'améliorer les performances d'E/S. Cependant, contrairement à la répartition des données par bandes sur plusieurs disques, les performances obtenues à l'aide de la mise en miroir dépendent du rapport de lecture/écriture lors de l'accès aux disques. Si la charge de travail du système consiste surtout en des opérations d'écritures (par exemple, plus de 30 % d'écritures), la mise en miroir risque de réduire les performances.

Pour optimiser les performances pour différents types de volumes mis en miroir, Volume Manager gère les stratégies de lecture suivantes :

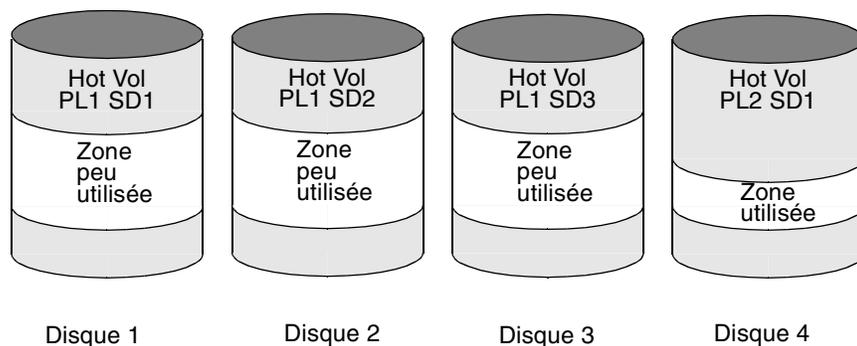
- ◆ la stratégie de *lecture alternative (round-robin)* – `round` –, dans laquelle les requêtes de lecture sur le volume sont traitées alternativement à partir de tous les plex du volume (c'est-à-dire, les plex sont traités à tour de rôle) ;
- ◆ la stratégie de lecture à partir d'un *plex de préférence* – `prefer` –, dans laquelle les requêtes de lecture sont traitées à partir d'un plex spécifique (en général, le plex dont les performances sont les meilleures), sauf si ce plex a échoué, auquel cas le système accède à un autre plex.

Recommandations pour les performances

- ◆ La stratégie de lecture par défaut – `select` –, qui sélectionne la stratégie de lecture appropriée pour la configuration. Par exemple, la lecture à partir d'un plex de préférence sera sélectionnée lorsqu'un seul plex réparti par bandes est associé au volume et la lecture alternative sera choisie dans la plupart des autres cas.

Dans l'exemple de configuration illustré à la figure 31, « Utilisation de la mise en miroir et de la répartition des données par bande sur plusieurs disques pour améliorer les performances », la stratégie de lecture du volume `Hot Vol` devrait être `prefer` pour le plex `PL1`. Ainsi, les lectures de `PL1` distribuent la charge entre plusieurs disques autrement peu sollicités, et non sur un seul disque.

Figure 31. Utilisation de la mise en miroir et de la répartition des données par bande sur plusieurs disques pour améliorer les performances



Vous pouvez attacher un maximum de 32 plex au même volume pour améliorer les performances lorsque la charge de travail est principalement composée d'opérations de lecture. Cependant, ce scénario se traduit par une réduction de l'utilisation réelle de l'espace disque. Vous pouvez aussi améliorer les performances en répartissant les données par bandes pour former un plex sur chaque moitié des disques disponibles. Lorsque cette méthode est possible, elle assure la meilleure configuration de Volume Manager sur un jeu de disques et permet d'optimiser les performances tout en assurant une fiabilité adéquate.

Mise en miroir et répartition des données par bandes sur plusieurs disques

L'utilisation conjointe de la mise en miroir et de la répartition des données par bande sur plusieurs disques combine les avantages de la répartition des données sur plusieurs disques et de la redondance des données.

Vous pouvez combiner la mise en miroir et la répartition des données par bande sur plusieurs disques pour améliorer considérablement les performances lorsqu'il existe plusieurs flux d'E/S. La répartition des données par bande sur plusieurs disques peut améliorer l'accès en série lorsque les E/S correspondent exactement à toutes les unités de bande d'une même bande. Elle accroît le débit, car les flux d'E/S peuvent s'effectuer simultanément sur des périphériques séparés.

Vous devrez peut-être utiliser parfois la mise en miroir pour les charges de travail composées principalement d'opérations d'écriture car elle permet le plus souvent d'assurer une protection contre tout risque de perte de données dûs à des défaillances de disque. Dans ce cas, vous pouvez combiner la mise en miroir et la répartition des données par bande sur plusieurs disques pour optimiser à la fois la disponibilité et les performances.

Répartition des données par bandes sur plusieurs disques et mise en miroir

L'utilisation conjointe de la répartition des données par bande sur plusieurs disques et de la mise en miroir combine les avantages inhérents aussi bien à la répartition des données sur plusieurs disques qu'à la redondance des données.

Vous pouvez combiner la répartition des données par bande sur plusieurs disques et la mise en miroir pour améliorer considérablement les performances lorsqu'il existe plusieurs flux d'E/S. La répartition des données par bande sur plusieurs disques peut améliorer l'accès en série lorsque les E/S correspondent exactement à toutes les unités de bande d'une même bande. Elle accroît le débit, car les flux d'E/S peuvent s'effectuer simultanément sur des périphériques séparés.

Vous devrez peut-être utiliser parfois la mise en miroir pour les charges de travail composées principalement d'opérations d'écriture car elle permet le plus souvent d'assurer une protection contre tout risque de perte de données dus à des défaillances de disque. Dans ce cas, vous pouvez combiner la mise en miroir et la répartition des données par bande sur plusieurs disques pour optimiser à la fois la disponibilité et les performances. Voir « Volumes répartis en couches » à la page 42.

Contrôle des performances

RAID 5

La structure RAID 5 offre la plupart des avantages de l'utilisation conjointe de la mise en miroir et de la répartition des données par bande sur plusieurs disques, mais nécessite moins d'espace disque. Les performances de lecture d'une structure RAID 5 sont similaires à celle de la répartition des données par bande sur plusieurs disques et la parité RAID 5 offre une redondance similaire à celle de la mise en miroir. Cependant, la structure RAID 5 présente plusieurs inconvénients, notamment une certaine lenteur d'écriture.

Remarque La structure RAID 5 n'est normalement pas considérée comme une solution pour l'amélioration des performances, sauf pour les rapports lecture-écriture élevés des modèles d'accès de l'application.

Contrôle des performances

Un administrateur système doit contrôler deux types de performances : les performances *physiques*, concernant le matériel, et celles *logiques*, concernant la gestion du logiciel et de ses opérations.

Types de performances

Les performances physiques sont caractérisées par l'équilibre et la concentration des E/S sur chaque unité de disque afin de réduire au maximum le temps de recherche. En fonction des résultats contrôlés, vous pouvez déplacer des sous-disques pour équilibrer les disques.

Les performances logiques concernent les opérations logicielles et leur gestion. En fonction de la surveillance, vous pouvez mettre en miroir certains volumes ou les répartir par bandes afin d'améliorer leurs performances. Vous pouvez améliorer les performances des volumes les plus importants aux dépens du débit général. Seul l'administrateur système peut déterminer la priorité du système et les compromis acceptables.

Généralement, vous pouvez optimiser les performances en mettant en miroir et en répartissant par bandes tous les volumes sur un nombre raisonnable de disques, et en effectuant une mise en miroir sur les contrôleurs, lorsque cela est possible. Cette opération tend à équilibrer la charge entre tous les disques. Cependant, elle complique également l'administration de Volume Manager. Si vous disposez de nombreux disques (plusieurs centaines, voire plusieurs milliers), vous pouvez les placer dans des groupes de 10 (via des disk groups), où chaque groupe permet de répartir par bandes et de mettre en miroir un jeu de volumes. Cette opération assure toujours de bonnes performances et facilite l'administration.

Obtention des données de performances

Volume Manager offre deux types d'informations qui facilitent le contrôle des performances : les statistiques et le suivi des E/S que vous pouvez obtenir respectivement à l'aide des utilitaires `vxstat` et `vxtrace`. Ce chapitre comprend une brève présentation de ces deux utilitaires.

Obtention des statistiques d'E/S (`vxstat`)

L'utilitaire `vxstat` accède aux informations d'activité sur les volumes, les plex, les sous-disques et les disques placés sous le contrôle de Volume Manager. Il indique les statistiques qui reflètent les niveaux d'activité des objets Volume Manager depuis le démarrage. Vous pouvez afficher à un moment donné les statistiques sur un objet Volume Manager spécifique ou sur tous les objets. Vous pouvez également spécifier un disk group, auquel cas le programme affiche uniquement les statistiques sur les objets de ce disk group. Si vous ne spécifiez aucun disk group, le programme utilise `rootdg` par défaut.

La quantité d'informations affichées dépend des options spécifiées pour `vxstat`. Pour plus d'informations sur les options disponibles, voir la page `man vxstat (1M)`.

Volume Manager enregistre les statistiques d'E/S suivantes :

- ◆ le nombre d'opérations ;
- ◆ le nombre de blocs transférés (une opération peut inclure plusieurs blocs) ;
- ◆ la durée moyenne de l'opération (qui reflète la durée totale dans l'interface de Volume Manager et ne peut pas servir aux comparaisons avec les autres programmes de statistiques).

Volume Manager enregistre les statistiques d'E/S précédentes pour les E/S logiques. Ces statistiques incluent pour chaque volume les lectures, les écritures, les copies atomiques, les lectures et les écritures vérifiées, ainsi que les lectures et les écritures de plex. Par conséquent, une écriture sur un volume à deux plex se traduit par cinq opérations minimum : une pour chaque plex, une pour chaque sous-disque et une pour le volume. De plus, une lecture répartie sur deux sous-disques indique au moins quatre lectures : une pour chaque sous-disque, une pour le plex et une pour le volume.

Volume Manager gère également d'autres données statistiques. Pour chaque plex, il indique les erreurs de lecture et d'écriture. Pour les volumes, il indique les erreurs de lecture et d'écriture corrigées outre les erreurs de lecture et d'écriture.

`vxstat` peut également remettre à zéro les informations statistiques. La commande `vxstat -r` permet d'effacer toutes les statistiques, aussi bien pour tous les objets que pour les seuls objets spécifiés. Si vous effectuez cette remise à zéro juste avant le lancement d'une opération, vous pourrez mesurer l'effet de cette dernière.

Contrôle des performances

L'exemple suivant illustre une commande `vxstat` :

TYP	NOM	OPÉRATIONS		BLOCS		DURÉE MOYENNE (ms)	
		LECTURE	ÉCRITURE	LECTURE	ÉCRITURE	LECTURE	ÉCRITURE
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

D'autres statistiques de volume sont disponibles sur les configurations RAID 5. Pour plus d'informations, voir la page man `vxstat` (1M).

Suivi des E/S (`vxtrace`)

La commande `vxtrace` assure les opérations de suivi sur les volumes. Elle imprime soit les erreurs d'E/S du noyau, soit les enregistrements de suivi des E/S dans la sortie standard ou elle écrit les enregistrements dans un fichier de format binaire. Le suivi peut être effectué sur des types d'objets d'E/S du noyau spécifiques ou bien sur des objets ou des périphériques spécifiés. Pour plus d'informations, voir la page man `vxtrace`(1M).

Utilisation des données de performances

Une fois collectées, les données de performances permettent de déterminer la configuration système optimale pour une utilisation efficace des ressources du système. Les sections suivantes décrivent l'utilisation possible de ces données.

Utilisation des statistiques d'E/S

L'examen des statistiques d'E/S peut suggérer une reconfiguration. Il existe deux statistiques principales : l'activité d'E/S d'un volume et l'activité d'E/S d'un disque.

Avant d'obtenir des statistiques, effacez (réinitialisez) toutes les statistiques existantes à l'aide de la commande `vxstat -r`. Lorsque vous effacez les statistiques, vous éliminez les différences entre les volumes ou les disques résultant de la création des volumes, et vous supprimez également les statistiques du démarrage (qui ne présentent normalement pas d'intérêt).

Une fois que vous avez effacé les statistiques, exécutez le système normalement. Pour mesurer l'effet d'une application ou d'une charge de travail spécifique, vous devez exécuter le système sur cette application ou cette charge de travail spécifique. Lorsque vous surveillez un système utilisé pour diverses activités, n'essayez pas d'utiliser de façon excessive une application donnée. Lorsque vous surveillez un système qui s'exécute « en temps partagé » avec de nombreux utilisateurs, efforcez-vous de collecter des statistiques pendant plusieurs heures d'utilisation normale par jour.

Pour afficher les statistiques sur les volumes, entrez la commande `vxstat` sans aucun argument. Les statistiques suivantes s'affichent :

TYP	NOM	OPÉRATIONS		BLOCS		DURÉE MOYENNE (ms)	
		LECTURE	ÉCRITURE	LECTURE	ÉCRITURE	LECTURE	ÉCRITURE
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

Cette sortie permet d'identifier les volumes qui enregistrent un grand nombre d'opérations ou un nombre excessif de lectures ou d'écritures.

Pour afficher les statistiques sur les disques, entrez la commande `vxstat -d`. Les statistiques suivantes s'affichent :

TYP	NOM	OPÉRATIONS		BLOCS		DURÉE MOYENNE (ms)	
		LECTURE	ÉCRITURE	LECTURE	ÉCRITURE	LECTURE	ÉCRITURE
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

Pour déplacer des volumes d'un disque à un autre ou pour déplacer le volume `archive` sur un autre disque, vous devez d'abord identifier le ou les disques sur lesquels il se trouve à l'aide de la commande suivante :

```
# vxprint -tvh archive
```

Exemple d'affichage :

V	NOM	TYPEUTIL	KSTATE	ÉTAT	LONGUEUR	RDPOL	PLEXPRÉF
PL	NOM	VOLUME	KSTATE	ÉTAT	LONGUEUR	STRUCTURE	MODE
COL/LARG							
SD	NOM	PLEX	DÉCALPL	DÉCDISQ	LONGUEUR	[COL/]OFF	FLAGS
v	archive	fsgen	ENABLED	ACTIF	204800	SELECT	-
pl	archive-01	archive	ENABLED	ACTIF	204800	CONCAT	- RW
sd	disk03-03	archive-01	0	409600	204800	0	<i>périphérique</i>

Remarque Votre système peut utiliser un *nom de périphérique* ou un *chemin* différents de ceux donnés dans les exemples. Pour plus d'informations sur les noms de périphériques, voir « Présentation de Volume Manager » à la page 18.

Contrôle des performances

La liste des sous-disques associés indique que le volume `archive` se trouve sur le disque `disk03`. Pour déplacer le volume de `disk03`, entrez la commande suivante :

```
# vxassist move archive !disk03 disque_dest
```

où *disque_dest* représente le disque vers lequel vous voulez déplacer le volume. Il n'est pas nécessaire de spécifier un tel disque. Dans ce cas, le volume est déplacé vers un disque disponible dont l'espace est suffisant pour contenir le volume.

Par exemple, entrez la commande suivante pour déplacer le volume de `disk03` à `disk04` :

```
# vxassist move archive !disk03 disk04
```

Cette commande indique que le volume doit être réorganisé de façon à ce qu'il n'en reste aucune partie sur `disk03`.

Remarque L'interface utilisateur graphique permet de déplacer facilement des parties de volumes d'un disque à un autre et elle peut être préférable à l'utilisation de la ligne de commande.

Si deux volumes (autres que le volume racine) sont occupés, déplacez-les de façon à ce que chacun se trouve sur un disque différent.

Si un seul volume est particulièrement occupé (notamment si son nombre moyen de lectures ou d'écritures est exceptionnellement élevé), répartissez-le par bandes (ou divisez-le en plusieurs éléments, en plaçant chacun d'eux sur un disque différent). Si vous effectuez cette opération en ligne, la conversion d'un volume de façon à utiliser la répartition des données par bandes sur plusieurs disques nécessite un espace libre suffisant pour contenir une copie supplémentaire du volume. Si l'espace libre disponible est insuffisant, vous pouvez effectuer à la place une copie de sauvegarde. Pour effectuer une conversion vers la répartition des données par bandes sur plusieurs disques, vous devez créer un plex réparti par bandes du volume, puis supprimer l'ancien plex. Par exemple, pour répartir par bandes le volume `archive` sur les disques `disk02`, `disk03` et `disk04`, entrez les commandes suivantes :

```
# vxassist mirror archive layout=stripe disk02 disk03 disk04
# vxplex -o rm disarchive-01
```

Une fois que vous avez réorganisé des volumes particulièrement occupés, vérifiez les statistiques sur le disque. Si certains volumes ont été réorganisés, effacez d'abord les statistiques, puis collectez-les pendant une durée raisonnable.

Si certains disques semblent excessivement occupés (ou si leur durée de lecture ou d'écriture est particulièrement longue), vous pouvez reconfigurer certains volumes. Si un disque contient deux volumes relativement occupés, rapprochez-les de façon à réduire les temps de recherche sur le disque. Si le nombre de volumes relativement occupés est trop élevé sur un disque, déplacez-les vers un disque moins occupé.

Utilisez le suivi d'E/S (ou les statistiques de sous-disque) pour déterminer si l'activité des volumes est excessive dans des zones particulières du volume. Si vous pouvez identifier les zones actives, divisez les sous-disques du volume et déplacez ces zones vers un disque moins occupé.

Attention La répartition par bandes d'un volume ou sa division sur plusieurs disques augmente le risque qu'une défaillance de disque entraîne celle de ce volume. Par exemple, si vous répartissez pas bandes cinq volumes sur les mêmes cinq disques, la défaillance de l'un de ces disques nécessite la restauration des cinq volumes à partir d'une sauvegarde. Si chaque volume se trouvait sur un disque séparé, il suffirait de restaurer un seul volume. Utilisez la mise en miroir ou la structure RAID 5 pour réduire le risque qu'une seule défaillance de disque entraîne celle d'un plus grand nombre de volumes.

Généralement, les systèmes de fichiers et les bases de données utilisent différemment avec le temps l'espace alloué et les informations spécifiques à la position sur un volume sont souvent inutiles. Pour les bases de données, vous pouvez identifier l'espace utilisé par une table ou un index particulièrement occupé. Si cet espace peut être identifié, vous pouvez le déplacer vers des disques non occupés.

L'examen du rapport entre les lectures et les écritures permet d'identifier les volumes que vous pouvez mettre en miroir afin d'améliorer leurs performances. Si le rapport lecture/écriture est élevé, la mise en miroir peut augmenter les performances et la fiabilité. Le rapport lectures/écritures pour lequel la mise en miroir peut améliorer les performances dépend dans une large mesure des disques, du contrôleur de disque, si plusieurs contrôleurs peuvent être utilisés et de la vitesse du bus système. Si un volume particulièrement occupé a un rapport lectures/écritures élevé, il est probable que la mise en miroir améliore considérablement les performances de ce volume.

Utilisation du suivi d'E/S

Les statistiques d'E/S fournissent les données nécessaires à l'analyse des performances de base ; le suivi d'E/S sert à une analyse plus détaillée. Il permet d'obtenir plus précisément une trace d'événement pour une charge de travail spécifique. Il est ainsi possible d'identifier explicitement l'emplacement et la taille d'un point actif et de l'application qui en est la cause.

L'utilisation des données provenant des suivis d'E/S permet de simuler de réelles charges de travail sur les disques et d'assurer le suivi des résultats. L'administrateur système peut anticiper à l'aide de ces statistiques les limitations du système et planifier des ressources supplémentaires.

Réglage de Volume Manager

Cette section décrit les mécanismes de contrôle des ressources utilisés par Volume Manager. Il est parfois nécessaire de régler certaines valeurs afin d'obtenir les meilleures performances (selon le type de ressources système disponible).

Recommandations générales pour le réglage

Volume Manager est réglé pour la plupart des configurations, des petits systèmes aux grands serveurs. Lorsque le réglage permet d'augmenter les performances sur les grands systèmes au détriment d'une ressource importante (par exemple, la mémoire), Volume Manager est généralement réglé de façon à être exécuté sur la plus petite configuration prise en charge. Ces modifications de réglage doivent être effectuées avec précaution car elles risquent d'affecter les performances générales du système, voire de rendre Volume Manager inutilisable.

Divers mécanismes permettent de régler Volume Manager. Sur certains systèmes, vous pouvez régler plusieurs paramètres au moyen du fichier réglable global `/etc/system`. En revanche, le réglage d'autres valeurs ne peut se faire que dans l'interface de ligne de commande de Volume Manager.

Éléments réglables

Sur certains systèmes, vous devez utiliser la commande `idtune` pour modifier les éléments réglables. Pour plus d'informations, voir la page `man idtune(1M)`.

Sur d'autres systèmes, vous pouvez modifier les éléments réglables en ajoutant des lignes au fichier `/etc/system`, puis en redémarrant le système. Les éléments réglables modifiés prennent alors effet.

Par exemple, pour remplacer la valeur par défaut de l'élément réglable `vol_tuneme` par la valeur 5000, vous devez insérer la ligne suivante dans la section appropriée du fichier `/etc/system`:

```
set vxio:vol_tuneme=5000
```

Dans bien des cas, les éléments réglables se trouvent dans la structure `volinfo`, décrite dans la page `man vxio(7)`.

Les sections suivantes décrivent des éléments réglables spécifiques.

vol_maxvol

Cette valeur contrôle le nombre maximal de volumes que vous pouvez créer sur le système. Elle peut être comprise entre 1 et le nombre maximal de mineurs dans le système.

La valeur par défaut pour cet élément réglable correspond à la moitié de la valeur maximale des mineurs pour le système.

voliomem_maxpool_sz

Cet élément réglable permet d'éviter qu'une E/S n'utilise toute la mémoire du système.

VxVM alloue deux pools de `voliomem_maxpool_sz`, l'un pour la structure RAID 5 et l'autre pour les volumes mis en miroir.

Lorsqu'une écriture effectuée pour un volume RAID 5 dépasse la valeur de `volio_maxpoll_sz/10`, elle est divisée et exécutée par incréments de `volio_maxpoll_sz/10`.

Lorsqu'une écriture effectuée pour un volume mis en miroir dépasse la valeur de `volio_maxpoll_sz/2`, elle est divisée et exécutée par incréments de `volio_maxpoll_sz/2`.

vol_subdisk_num

Cet élément réglable permet de contrôler le nombre maximal de sous-disques que vous pouvez attacher à un même plex. Ce nombre ne fait en théorie l'objet d'aucune limite, mais, dans la pratique, il a été limité par défaut à 4 096. Vous pouvez modifier cette valeur par défaut, le cas échéant.

vol_maxioctl

Cette valeur contrôle la taille maximale des données qui peuvent être transmises à Volume Manager via un appel `ioctl`. L'augmentation de cette limite permet d'exécuter des opérations plus importantes. Généralement, il est déconseillé de réduire cette limite car certains utilitaires dépendent de l'exécution d'opérations d'une taille donnée et ils risquent d'échouer de façon inattendue s'ils envoient des requêtes `ioctl` excessives.

La valeur par défaut pour cet élément réglable est 32 768 octets (32 Ko).

vol_maxspecialio

Cet élément réglable contrôle la taille maximale d'une E/S envoyée par un appel `ioctl`. La requête `ioctl` en elle-même peut être minime, mais son exécution aura exigé une opération d'E/S volumineuse. Cet élément réglable limite la taille de ces E/S. Si nécessaire, une requête dépassant cette valeur peut échouer ou les E/S peuvent être interrompues et effectuées en mode synchrone.

La valeur par défaut pour cet élément réglable est 512 secteurs (256 Ko).

Réglage de Volume Manager

vol_maxio

Cette valeur contrôle la taille maximale des opérations d'E/S logiques que vous pouvez exécuter sans interrompre la requête. Les requêtes d'E/S physiques supérieures à cette valeur seront divisées et effectuées en mode synchrone. Les E/S physiques sont divisées en fonction des capacités du périphérique de disque et elles ne sont pas affectées par les modifications apportées à cette limite maximale de requête logique.

La valeur par défaut pour cet élément réglable est 512 secteurs (256 Ko).

L'augmentation de cette limite peut poser un problème si la taille d'une E/S oblige la procédure à occuper davantage de mémoire ou d'espace de mappage du noyau qu'il n'en existe, et donc à bloquer le système. La limite maximale pour `vol_maxio` est 20 % de la plus petite des valeurs entre la mémoire physique ou la mémoire virtuelle du noyau. Il est déconseillé de dépasser cette limite car cela risque d'entraîner un blocage du système.

Si certaines des bandes dépassent la valeur de `vol_maxio`, leurs E/S complètes sont divisées afin d'éviter des lectures-écritures complètes sur la bande. En effet, cette procédure risque de bloquer le débit d'E/S du volume pour les E/S séquentielles ou plus volumineuses.

Cet élément réglable devrait avoir au moins la valeur de la taille de la plus large des bandes. Cette recommandation s'applique aussi bien à la répartition des données par bandes sur plusieurs disques RAID 0 que RAID 5.

vol_maxkiocount

Cet élément réglable contrôle le nombre maximal d'E/S que Volume Manager peut exécuter en parallèle. Les autres E/S qui tentent d'utiliser un périphérique de volume seront mises en file d'attente jusqu'à ce que le nombre d'activités actuel tombe au-dessous de cette valeur.

La valeur par défaut pour cet élément réglable est 2 048.

Il faudrait exécuter en parallèle des requêtes d'E/S de 2 Ko pour atteindre la limite des E/S actives dans le noyau car les threads de processus ne peuvent souvent envoyer qu'une seule E/S à la fois. L'augmentation de cette limite ne devrait guère apporter d'amélioration, sauf pour les plus grands systèmes.

vol_default_iodelay

Cette valeur représente le nombre de tops d'horloge pendant lesquels les utilitaires vont effectuer une pause entre chaque envoi d'E/S s'ils doivent ralentir la vitesse d'émission des E/S, mais n'ont reçu aucune instruction spécifique quant au délai approprié. Les utilitaires qui exécutent des opérations telles que la resynchronisation des miroirs ou la reconstruction des colonnes RAID 5 utiliseront cette valeur.

La valeur par défaut pour cet élément réglable est 50 tops.

L'augmentation de cette valeur ralentira les opérations de récupération, et par conséquent, réduira son incidence sur le système pendant l'exécution des récupérations.

voldrl_min_regionsz

Avec le journal des zones modifiées (DRL, *Dirty Region Logging*), Volume Manager divise logiquement un volume en un ensemble de zones consécutives. L'élément réglable `voldrl_min_regionsz` spécifie le nombre minimum de secteurs pour une zone de volume DRL.

Le noyau de Volume Manager définit actuellement la valeur par défaut pour cet élément réglable à 1024 secteurs.

Des tailles de zones plus importantes tendent à améliorer le taux de réussite pour les zones. Cela améliore les performances d'écriture au détriment de la durée de récupération.

voldrl_max_dirty

Certains volumes, tels que ceux utilisés pour les journaux de relecture Oracle, sont écrits séquentiellement et n'utilisent pas ce nettoyage des bits DRL. Pour ces volumes, vous pouvez utiliser un *journal des zones modifiées séquentiel* afin de diminuer encore davantage le nombre de bits modifiés et d'accélérer la récupération. Le nombre de bits modifiés autorisés pour un journal des zones modifiées séquentiel sera limité par l'élément réglable `voldrl_max_dirty`. L'utilisation d'un journal des zones modifiées séquentiel sur des volumes écrits séquentiellement risque d'affecter gravement le débit des E/S.

voldrl_max_drtregs

Cet élément réglable spécifie le nombre maximal de zones modifiées qui peuvent exister sur le système à un moment donné. Cette valeur globale s'applique au système entier, quel que soit le nombre de ses volumes actifs.

La valeur par défaut pour cet élément réglable est 2 048.

Vous pouvez utiliser l'élément réglable `voldrl_max_drtregs` afin de régler la durée de récupération dans le pire des cas pour le système, à la suite d'une défaillance. Une valeur plus élevée peut améliorer les performances du système au détriment de la durée de récupération.

vol_maxparallelio

Cet élément réglable contrôle le nombre d'opérations d'E/S que le démon `vxconfigd(1M)` peut demander à partir du noyau dans un seul `VOL_VOLDIO_READ` par appel `ioctl VOL_VOLDIO_WRITE`.

La valeur par défaut pour cet élément réglable est 256 ; il est déconseillé de la modifier.

Réglage de Volume Manager

vol_mvr_maxround

Cette valeur contrôle la granularité de la stratégie de lecture alternative à partir des miroirs. Une lecture sera traitée par le même miroir que la dernière lecture si son décalage est compris dans le nombre de secteurs décrit par cet élément réglable de la dernière lecture.

La valeur par défaut pour cet élément réglable est 512 secteurs (256 Ko).

L'augmentation de cette valeur diminuera le nombre de passages aux autres miroirs lors de la lecture. Cette procédure est souhaitable si les E/S sont exécutées principalement en mode séquentiel, entrecoupées d'un nombre réduit de recherches. Pour de nombreuses lectures de volume distribuées de façon aléatoire, il est souvent préférable d'effectuer la lecture à partir d'autres miroirs.

voliot_iobuf_limit

Cette valeur définit la limite applicable à la quantité de mémoire qui sera utilisée pour stocker les tampons de suivi dans le noyau. Le noyau de Volume Manager utilise les tampons de suivi pour stocker les enregistrements d'événements de suivi. Au fur et à mesure que ces tampons doivent être stockés dans le noyau, la mémoire nécessaire est prise de ce pool.

L'augmentation de cette taille peut permettre l'exécution d'un suivi supplémentaire au détriment de l'utilisation de la mémoire système. Il est déconseillé d'attribuer à cette taille une valeur supérieure à celle possible.

La valeur par défaut pour cet élément réglable est 131 072 octets (128 Ko).

voliot_iobuf_max

Cette valeur contrôle la taille de tampon maximale applicable à un seul tampon de suivi. Les requêtes de tampon supérieures à cette taille seront automatiquement réduites à cette taille. Une requête de taille de tampon maximale à partir de l'interface de suivi se traduira (suivant les limites d'utilisation) par un tampon de cette taille.

La valeur par défaut de ce tampon est 65 536 octets (64 Ko).

L'augmentation de ce tampon permet d'extraire des suivis plus importants sans perte, dans le cas des volumes les plus utilisés. Vous devez veiller à ne pas augmenter cette valeur au-delà de celle définie pour l'élément réglable `voliot_iobuf_limit`.

voliot_iobuf_default

Cette valeur représente la taille par défaut pour la création d'un tampon de suivi lorsque vous ne spécifiez pas la taille voulue pour le tampon du noyau lors du suivi `ioctl`.

La valeur par défaut pour cet élément réglable est 8 192 octets (8 Ko).

Cette valeur doit être augmentée si les données de suivi sont souvent perdues à cause de la trop petite taille de ce tampon.

voliot_errbuf_default

Cet élément réglable indique la taille par défaut du tampon géré pour les événements de suivi des erreurs. Ce tampon est alloué au moment du chargement du pilote et sa taille ne peut pas être réglée pendant l'exécution de Volume Manager.

La valeur par défaut pour cet élément réglable est 16 384 octets (16 Ko).

L'augmentation de ce tampon peut permettre de conserver davantage d'événements d'erreur, au détriment toutefois de la mémoire système. La réduction de la taille du tampon peut empêcher la détection d'une erreur via le périphérique de suivi. Les applications dont les réponses dépendent du suivi des erreurs utilisent ce tampon.

voliot_max_open

Cette valeur contrôle le nombre maximal de canaux de suivi qui peuvent être ouverts simultanément. Les canaux de suivi sont des points d'entrée clone dans le pilote du périphérique de suivi. Chaque commande `vxt race` exécutée sur le système utilisera un seul canal de suivi.

Le nombre de canaux par défaut est 32. L'allocation de chaque canal occupe environ 20 octets, même s'il n'est pas utilisé.

vol_checkpt_default

Cet élément réglable contrôle l'intervalle auquel les utilitaires exécutant des récupérations ou des opérations de resynchronisation chargeront le décalage actuel dans le noyau, de sorte qu'une défaillance du système ne nécessitera pas une récupération complète, mais pourra se poursuivre à partir du dernier point de reprise atteint.

La valeur par défaut du point de reprise est 20 480 secteurs (10 Mo).

L'augmentation de cette taille réduit le délai d'exécution des points de reprise sur les opérations de récupération, au détriment de la récupération supplémentaire effectuée après une défaillance système survenue au cours d'une récupération.

volraid_rsrtransmax

Cet élément réglable RAID 5 contrôle le nombre maximal d'opérations de reconstruction non permanentes qui peuvent être exécutées en parallèle. Une opération de reconstruction non permanente s'effectue sur un volume RAID 5 non dégradé ; elle est donc imprévisible. Lorsque vous limitez le nombre d'opérations de ce type susceptibles de se

Réglage de Volume Manager

produire simultanément, vous éliminez tout risque de noyer le système sous différentes opérations de reconstruction simultanées, ce qui réduit le risque de saturation de la mémoire.

Le nombre par défaut de reconstructions non permanentes pouvant être exécutées en parallèle est 1.

L'augmentation de cette valeur peut améliorer les performances initiales du système lors de la première défaillance et avant le détachement d'un objet qui a échoué, mais elle risque de saturer la mémoire.

`vliomem_chunk_size`

La mémoire système est alloué à Volume Manager et libérée sur la base de cette granularité. Une granularité plus élevée réduit (dans une certaine mesure) la durée de l'allocation de mémoire en permettant à Volume Manager de conserver une quantité de mémoire plus importante.

La valeur par défaut pour cet élément réglable est 64 Ko.

Réglage pour les gros systèmes

Sur les petits systèmes (moins d'une centaine de lecteurs), le réglage est normalement inutile et Volume Manager devrait pouvoir utiliser des valeurs par défaut raisonnables pour tous les paramètres de configuration. Cependant, sur les gros systèmes, certaines configurations peuvent nécessiter un contrôle supplémentaire pour le réglage de ces paramètres, aussi bien en termes de capacité que de performances.

En règle générale, vous ne pouvez prendre qu'un nombre limité de décisions importantes lors de la configuration de Volume Manager sur un gros système. L'une d'elles consiste à déterminer la taille des disk groups et le nombre de copies de configuration à gérer pour chaque disk group. Une autre concerne la taille à attribuer à la zone privée pour tous les disques d'un disk group.

Les disk groups plus volumineux présentent l'avantage d'offrir un pool d'espace libre plus important dans lequel la commande `vxassist (1M)` peut effectuer une sélection et ils permettent également de créer des ensembles de disques plus volumineux. Cependant, les disk groups plus petits ne nécessitent pas une base de données de configuration aussi importante et par conséquent, ils peuvent avoir des zones privées plus petites. Les disk groups vraiment volumineux peuvent finalement utiliser toute la taille de la zone privée, avec pour résultat d'empêcher l'ajout d'autres objets de configuration dans ce disk group. À ce stade, vous devez soit diviser la configuration en plusieurs disk groups, soit augmenter les zones privées. Cette procédure implique la réinitialisation de chaque disque du disk group (elle peut même nécessiter une reconfiguration totale et une restauration à partir de la sauvegarde).

Il est recommandé aux utilisateurs de sous-systèmes d'ensembles de disques de créer un seul disk group pour chaque ensemble de disques, de façon à ce que le disk group puisse être déplacé physiquement en tant qu'unité entre les systèmes.

Nombre de copies de configuration pour un disk group

La sélection du nombre de copies de configuration pour un disk group est basée sur un compromis entre la redondance et les performances. En règle générale, moins il existe de copies de configuration dans un disk group, plus l'accès initial au disk group, le démarrage initial de `vxconfigd` (1M) et l'exécution des transactions sont rapides sur le disk group.

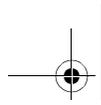
Attention Une redondance plus faible des copies de base de données présente un risque de perte de la base de données de configuration, qui peut à son tour entraîner la perte de tous ses objets et de toutes les données contenues dans le disk group.

La stratégie par défaut pour les copies de configuration dans le disk group consiste à allouer une copie de configuration à chaque contrôleur identifié dans le disk group ou à chaque cible contenant plusieurs disques adressables sur la même cible. Cette stratégie est suffisante au niveau de la redondance, mais dans certaines circonstances, cela peut se traduire par un nombre élevé de copies de configuration.

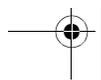
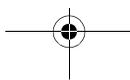
Dans ce cas, il est recommandé de limiter le nombre de copies de configuration à un minimum de 4. La sélection de l'emplacement de ces copies s'effectue comme auparavant, en fonction du contrôleur maximal ou de la répartition des cibles.

Le mécanisme permettant de définir le nombre de copies pour un disk group consiste à utiliser la commande `vxdbg init` pour une nouvelle configuration de groupe (pour plus d'informations, voir la page `man vxdbg` (1M)). De plus, vous pouvez modifier les copies d'un groupe existant au moyen de la commande `vxedit set` (pour plus d'informations, voir la page `man vxedit` (1M)). Par exemple, pour définir le disk group `foodg` afin qu'il contienne 5 copies, entrez la commande suivante :

```
# vxedit set nconfig=5 foodg
```



Réglage de Volume Manager



Glossaire

Abandon d'un nœud

Situation dans laquelle un nœud quitte un cluster (en cas d'urgence) sans essayer d'arrêter les opérations en cours.

Arrêt normal d'un nœud

Possibilité offerte à un nœud de quitter le cluster normalement lorsque tous les accès aux volumes partagés ont cessé.

Associer

Opération consistant à établir une relation entre les objets Volume Manager ; par exemple, un sous-disque qui a été créé et défini comme ayant un point de départ dans un plex est considéré comme associé à ce plex.

Attaché

État dans lequel un objet VxVM est à la fois associé à un autre objet et activé pour être utilisé.

Bande

Jeu d'unités de bande qui occupent les mêmes positions dans une série de colonnes.

Bande de données

Ce terme désigne la partie des données utilisable d'une bande ; elle correspond à la bande moins la zone de parité.

Base de données de configuration

Ensemble d'enregistrements contenant des informations détaillées sur les objets Volume Manager existants (par exemple des attributs de disque et de volume). On appelle copie de configuration une copie unique d'une base de données de configuration.

Bloc

Unité minimale de transfert des données vers un disque ou un ensemble de disques.

Chemin d'accès

Lorsqu'un disque est connecté à un hôte, le chemin d'accès au disque est représenté par l'adaptateur de bus hôte (HBA, *Host Bus Adapter*) sur l'hôte, le connecteur SCSI ou de câbles à fibres, et par le contrôleur sur le disque ou l'ensemble de disques. Ces composants constituent le chemin d'accès à un disque. En cas de défaillance de l'un de ces composants, la fonction DMP (*Dynamic Multipathing*) tente de transférer toutes les E/S pour ce disque vers les autres chemins d'accès (de remplacement).

Chemin d'accès activé

Chemin d'accès vers un disque disponible pour les E/S.

Chemin d'accès désactivé

Chemin d'accès à un disque non disponible pour les E/S. Un chemin peut être *désactivé* à la suite de défaillances matérielles véritables ou si l'utilisateur a employé la commande `vxddm adm disable` sur ce contrôleur.

Chemin d'accès multiple

Un disque connecté à un système est dit à chemins d'accès multiples lorsque plusieurs chemins d'accès physiques permettent d'y accéder. Tout logiciel résidant sur l'hôte (par exemple, le pilote DMP) qui masque cette caractéristique à l'utilisateur assure une fonctionnalité de chemin d'accès multiple.

Chemin d'accès principal

Dans les ensembles de disques de type actif/passif, un disque peut être lié à un contrôleur particulier sur l'ensemble de disques ou il peut appartenir à un contrôleur. L'accès au disque s'effectue alors à l'aide du chemin via ce contrôleur spécifique. Voir « *chemin d'accès* », « *chemin d'accès secondaire* ».

Chemin d'accès secondaire

Dans les ensembles de disques de type actif/passif, il s'agit des chemins d'accès à un disque différents du chemin d'accès principal. Généralement, l'accès à un disque s'effectue exclusivement via le chemin principal, tant qu'il n'échoue pas, auquel cas la propriété du disque est transférée à l'un des chemins secondaires. Voir « *chemin d'accès* », « *chemin d'accès principal* ».

Cluster

Ensemble d'hôtes partageant un jeu de disques.

Colonne

Ensemble d'un ou de plusieurs sous-disques dans un plex réparti par bandes. La répartition des données par bandes sur plusieurs disques est assurée par l'allocation alternée et régulière des données dans toutes les colonnes d'un plex.

Concaténation

Style de structure qui se caractérise par une disposition séquentielle et contiguë des sous-disques.

Configuration racine

Base de données de configuration pour le disk group racine. Elle est spéciale dans la mesure où elle contient toujours des enregistrements d'autres disk groups, qui servent exclusivement pour la sauvegarde. Elle contient également des enregistrements de disques qui définissent tous les périphériques de disque du système.

Contrôle racine

Possibilité de placer le système de fichiers `root` et le périphérique `swap` sous le contrôle de Volume Manager. Les volumes ainsi obtenus peuvent être mis en miroir afin de garantir la redondance et permettre la récupération en cas de défaillance d'un disque.

Contrôleur de disque

Le contrôleur (HBA) connecté à l'hôte *ou* à l'ensemble de disques et représenté par le système d'exploitation comme le nœud parent du disque est appelé contrôleur de disque par le sous-système à chemins d'accès multiples de Volume Manager. Par exemple, si un disque est représenté par le nom de périphérique :

```
/devices/sbus@1f,0/QLGC,isp@2,10000/sd@8,0:c
```

le contrôleur du disque `sd@8,0:c` est :

```
QLGC,isp@2,10000
```

Ce contrôleur (HBA) est connecté à l'hôte.

Détaché

État dans lequel un objet VxVM est associé à un autre objet, mais n'est pas activé pour être utilisé.

Disk group

Jeu de disques partageant une configuration commune. Une configuration de disk group est un ensemble d'enregistrements qui contiennent des informations détaillées sur des objets Volume Manager existants (par exemple, des attributs de disque et de volume) et

sur leurs relations. Chaque disk group a un nom attribué par l'administrateur et un ID unique défini en mode interne. Le disk group racine (`rootdg`) est un disk group privé particulier qui existe toujours.

Disk group partagé

Disk group dans lequel les disques sont partagés par plusieurs hôtes (aussi appelé *disk group partageable par les clusters*).

Disk group partageable par les clusters

Disk group dans lequel les disques sont partagés par plusieurs hôtes (aussi appelé *disk group partagé*).

Disk group privé

Disk group dans lequel l'accès aux disques s'effectue uniquement via un hôte spécifique.

Disk group racine

Disk group privé spécial qui existe toujours sur le système. Le disk group racine est nommé `rootdg`.

Disque

Ensemble de blocs de données en lecture/écriture qui sont indexées et dont l'accès est relativement rapide. Chaque disque est doté d'un identificateur universel unique.

Disque d'initialisation

Disque servant au démarrage. Ce disque peut être sous le contrôle de VxVM.

Disque physique

Périphérique de stockage sous-jacent, qui peut être sous le contrôle de Volume Manager.

Disque racine

Disque contenant le système de fichiers racine. Ce disque peut se trouver sous le contrôle de VxVM.

Disque VM

Disque placé sous le contrôle de Volume Manager, qui est attribué à un disk group. Les disques VM sont parfois appelés *disques Volume Manager* ou simplement *disques*. Dans l'interface utilisateur graphique, les disques VM sont représentés par une icône (un cylindre marqué D).

Disque VM partagé

Disque VM appartenant à un disk group partagé.

Dissocier

Opération de suppression d'un lien existant entre deux objets Volume Manager. Par exemple, la dissociation d'un sous-disque d'un plex supprime le sous-disque de ce plex et l'ajoute au pool d'espace libre.

Fichier volboot

Petit fichier servant à trouver des copies de la configuration racine. Ce fichier peut lister les disques qui contiennent des copies de configuration aux emplacements standard et il peut également contenir des pointeurs directs vers les emplacements des copies de la configuration. volboot est stocké dans un emplacement dépendant du système.

Encapsulation

Procédure assurant la conversion de partitions existant sur un disque spécifié en volumes. Si l'une de ces partitions contient des systèmes de fichiers, les entrées /etc/vfstab sont modifiées de façon à ce que les systèmes de fichiers soient montés à la place sur des volumes. L'encapsulation ne peut pas être appliquée à tous les systèmes.

Enregistrements du chemin d'accès au disque

Enregistrements de configuration permettant de spécifier le chemin d'accès à des disques particuliers. Chaque enregistrement du chemin d'accès au disque contient un nom, un type, voire des informations spécifiques au type, que Volume Manager utilise pour déterminer la façon d'accéder et de manipuler le disque défini par l'enregistrement du chemin d'accès au disque.

Enregistrement du nom logique du disque

Enregistrement de configuration qui identifie un disque spécifique par son ID et lui donne un nom logique (ou d'administration).

Ensemble de disques

Jeu de disques organisés de façon logique dans un objet. Ils offrent souvent des avantages spécifiques, comme la redondance ou l'amélioration des performances.

Ensemble de disques actif/actif

Ce type d'ensemble de disques à chemins d'accès multiples permet d'accéder à l'un de ses disques via tous les chemins d'accès en même temps, sans que cela entraîne une diminution des performances.

Ensemble de disques actif/passif

Ce type d'ensemble de disques à chemins d'accès multiples permet de désigner le chemin d'accès à un disque comme chemin d'accès principal et de l'utiliser pour accéder à ce disque à tout moment. L'utilisation d'un chemin différent du chemin actif désigné peut entraîner une forte diminution des performances de certains ensembles de disques. Voir « *chemin d'accès* », « *chemin d'accès principal* », « *chemin d'accès secondaire* ».

Espace libre

Zone d'un disque placé sous le contrôle de VxVM, qui n'est allouée à aucun sous-disque ni réservée à l'utilisation d'un autre objet Volume Manager.

Gestionnaire de clusters

Démon externe exécuté sur tous les nœuds d'un cluster. Les gestionnaires de clusters sur chaque nœud communiquent entre eux et informent VxVM des modifications apportées aux membres des clusters.

Gestionnaire de verrouillage distribué

Gestionnaire de verrouillage exécuté sur des systèmes différents, qui assure un accès cohérent aux ressources distribuées.

Hostid

Chaîne permettant d'identifier un hôte dans Volume Manager. La chaîne *hostid* d'un hôte est stockée dans son fichier *volboot* et elle permet de définir la propriété des disques et des disk groups.

ID de disk group

Identificateur unique permettant d'identifier un disk group.

ID de disque

Identificateur universel unique attribué à chaque disque et qui permet de l'identifier, même s'il est déplacé.

Jonction de nœud

Procédure par laquelle un nœud se joint à un cluster et peut accéder aux disques partagés.

Journal des événements de volume

Le périphérique de journal des événements de volume (*/dev/vx/event*) est l'interface par laquelle les événements de pilote de volume sont signalés aux utilitaires.

Journal des zones modifiées

Procédure par laquelle Volume Manager contrôle et enregistre dans un journal les modifications apportées à un plex. Le système conserve une bitmap des zones modifiées dans un sous-disque associé (le *sous-disque de journal*).

Logging d'état permanent

Type de logging garantissant que seuls les miroirs actifs sont utilisés pour la récupération et interdisant la sélection des miroirs qui ont échoués pour la récupération. Aussi appelé *logging de noyau*.

Miroir

Copie dupliquée d'un volume et des données qu'il contient (sous la forme d'un jeu ordonné de sous-disques). Chaque miroir est une copie du volume auquel le miroir est associé. Les termes *miroir* et *plex* sont synonymes.

Mise en miroir

Technique de structure qui met en miroir le contenu d'un volume sur plusieurs plex. Chaque plex duplique les données stockées sur le volume, mais les plex peuvent avoir des structures différentes.

Mode de lecture-réécriture

Mode de récupération dans lequel chaque opération de lecture récupère la cohérence des plex pour la zone qu'elle couvre. La cohérence des plex est récupérée en lisant les blocs qui appartiennent à un plex et en écrivant ces données sur tous les autres plex en écriture.

Nœud

L'un des hôtes d'un cluster.

Nœud de démarrage

Nœud sur lequel l'administrateur système exécute un utilitaire qui demande la modification des objets Volume Manager. Ce nœud démarre la reconfiguration d'un volume.

Nœud esclave

Nœud qui n'est pas désigné comme nœud principal.

Nœud maître

Nœud auquel un disque est attaché. Aussi appelé *propriétaire de disque*.

Nœud principal

Nœud désigné comme nœud « principal » par le logiciel. Il peut s'agir de n'importe quel nœud. Le nœud principal coordonne certaines opérations Volume Manager.

Nom d'accès logique au disque

Nom logique ou d'administration choisi pour le disque, comme `disk03`. Le terme *nom du disque* peut également désigner le nom d'accès logique au disque.

Nom d'accès physique au disque

Nom permettant d'accéder à un disque physique, par exemple `c0t0d0`. La syntaxe `c#t#d#s#` identifie le contrôleur, l'adresse cible, le disque et la partition. Le terme *nom de périphérique* peut également désigner le nom d'accès physique au disque.

Nom de périphérique

Nom de périphérique ou adresse permettant d'accéder à un disque physique, par exemple `c0t0d0`. La syntaxe `c#t#d#s#` identifie le contrôleur, l'adresse cible, le disque et la zone (sur certains systèmes, partition).

Numéro de série de l'ensemble de disques

Numéro de série de l'ensemble de disques. Il apparaît généralement sur le boîtier de l'ensemble de disques ou il peut être obtenu en envoyant une commande SCSI spécifique du fournisseur aux disques de l'ensemble. Le sous-système DMP utilise ce numéro pour identifier de façon unique un ensemble de disques.

Objet

Entité définie et reconnue en mode interne par Volume Manager. Les objets VxVM sont les suivants : volume, plex, sous-disque, disque et disk group. Il existe en fait deux types d'objets de disque — un pour l'aspect physique du disque et l'autre pour l'aspect logique.

Opération atomique

Opération qui soit réussit totalement, soit échoue en conservant pour toutes les données dans l'état qui était le leur avant le début de l'opération. Si l'opération réussit, tous ses aspects prennent effet en même temps et les états de changement intermédiaires sont invisibles. Si un aspect de l'opération échoue, celle-ci est abandonnée et aucun changement partiel n'est effectué.

Parité

Valeur calculée qui permet de reconstruire les données après une défaillance. La parité est également calculée en utilisant l'opérateur XOR (*OU exclusif*) avec les données pendant qu'elle sont écrites sur un volume RAID 5. La parité ainsi obtenue est alors écrite sur le volume. Si une partie d'un volume RAID 5 échoue, il est possible de recréer les données qu'elle contenait à partir des données restantes et des informations de parité.

Partition

Division standard d'un périphérique de disque physique, directement prise en charge par le système d'exploitation et les unités de disque.

Partition racine

Zone du disque contenant le système de fichiers racine.

Périphérique de configuration de volume

Le périphérique de configuration de volume (`/dev/vx/config`) est l'interface par laquelle sont effectuées toutes les modifications de la configuration du pilote de périphérique de volume.

Pilote de périphérique de volume

Pilote qui constitue le lecteur de disque virtuel entre l'application et le pilote de périphérique physique. L'accès au pilote de périphérique de volume s'effectue au moyen d'un nœud de périphérique de disque virtuel dont les nœuds de périphériques de caractère apparaissent dans `/dev/vx/rdisk` et dont les nœuds de périphériques de bloc apparaissent dans `/dev/vx/dsk`.

Plex

Copie dupliquée d'un volume et des données qu'il contient (sous la forme d'un ensemble ordonné de sous-disques). Chaque plex est une copie du volume auquel il est associé. Les termes *miroir* et *plex* sont synonymes.

Plex associé

Plex associé à un volume.

Plex clairsemé

Plex qui n'est pas aussi long que le volume ou qui comporte des trous (zones du plex sans sous-disque de sauvegarde).

Plex de journal

Plex permettant de stocker un journal RAID 5. Le terme *plex de journal* peut également renvoyer à un plex de journal des zones modifiées.

Plex dissocié

Plex dissocié d'un volume.

RAID

Un ensemble de disques RAID (*Redundant Array of Independent Disks*) est configuré avec une partie de la capacité de stockage combiné utilisée pour les informations dupliquées sur les données stockées dans cet ensemble de disques. Il permet de régénérer les données en cas de défaillance de disque.

Reconstruction automatique

Technique permettant de restaurer automatiquement la redondance et l'accès aux volumes mis en miroir et RAID 5 en cas de défaillance d'un disque. Elle implique la reconstruction des sous-disques affectés en disques spécifiés comme disques de remplacement et/ou espace libre dans le même disk group.

Répartition

Technique de structure permettant de répartir sur plusieurs disques physiques un volume (et son système de fichiers ou sa base de données) trop grand pour tenir sur un seul disque.

Répartition des données par bandes sur plusieurs disques

Technique de structure qui répartit les données sur plusieurs disques au moyen de bandes. Les données sont allouées alternativement aux bandes dans les sous-disques de chaque plex.

Secteur

Unité de taille variable suivant le système. Un secteur correspond normalement à 512 octets.

Sous-disque

Ensemble consécutif de blocs de disques contigus formant un segment de disque logique. Les sous-disques peuvent être associés aux plex pour former des volumes.

Sous-disque associé

Sous-disque associé à un plex.

Sous-disque de journal

Sous-disque permettant de stocker un journal des zones modifiées. Voir *Journal des zones modifiées*.

Sous-disque dissocié

Sous-disque dissocié d'un plex.

Sous-disque libre

Sous-disque non associé à un plex et dont le champ `put il [0]` est vide.

Système de fichiers

Ensemble de fichiers organisés dans une structure. Le système de fichiers UNIX est une structure hiérarchique comprenant des répertoires et des fichiers.

Système de fichiers racine

Système de fichiers initial monté au cours de la séquence de démarrage du noyau UNIX.

Taille de bande

Somme des tailles d'unités de bande comprenant une seule bande sur toutes les colonnes en cours de répartition.

Taille d'unité de bande

Taille de chaque unité de bande. L'unité de bande par défaut a une taille de 32 secteurs (16 Ko). La *taille d'unité de bande* est également appelée *largeur de bande*.

Transaction

Ensemble des modifications de la configuration qui réussissent ou qui échouent en tant que groupe, et non individuellement. Les transactions sont utilisées en mode interne pour assurer des configurations homogènes.

Unité de bande

Zones de taille égale allouées alternativement aux sous-disques (dans les colonnes) de chaque plex réparti par bandes. Dans un ensemble de disques, il s'agit d'un groupe de blocs logiques contigus qui existent sur chaque disque avant que les allocations ne soient effectuées à partir du disque suivant de l'ensemble. Une *unité de bande* est quelquefois appelé *élément de bande*.

Unité de bande de parité

Zone de stockage d'un volume RAID 5 contenant des informations de parité. Les données contenues dans l'unité de bande de parité peuvent servir à reconstruire les zones manquantes d'un volume RAID 5 à la suite d'erreurs d'E/S ou d'une défaillance de disque.

Volume

Disque virtuel représentant une plage adressable de blocs de disques utilisés par des applications telles que les systèmes de fichiers ou les bases de données. Un volume peut comprendre 1 à 32 plex.

Volume de swap

Volume VxVM configuré pour servir de zone de swap.

Volume partagé

Volume qui, appartenant à un disk group partagé, est ouvert sur plusieurs nœuds à la fois.

Volume racine

Volume VxVM contenant le système de fichiers racine, si un tel volume est désigné par la configuration du système.

vxconfigd

Démon de configuration de Volume Manager, qui modifie la configuration de VxVM. Ce démon doit être exécuté pour pouvoir effectuer des opérations VxVM.

Zone de swap

Zone de disque servant à conserver les copies des pages de mémoire permutées par la procédure de pagination du système.

Zone privée

Zone d'un disque physique servant à stocker des informations Volume Manager structurées et privées. La *zone privée* contient un en-tête de disque, une table des matières et une base de données de configuration. La table des matières mappe le contenu du disque. L'en-tête de disque contient un ID de disque. Les données de la zone privée sont toutes dupliquées afin de garantir une fiabilité supplémentaire.

Zone publique

Zone d'un disque physique gérée par Volume Manager, qui contient l'espace disponible et permet d'allouer les sous-disques.

A

- Accès au disque
 - nom 21, 142
- Accès logique au disque
 - nom 142
- Activation
 - accès à un disk group 174
 - d'un disk group 140
 - d'un disque 140
 - disque 166
- Affichage
 - configuration des volumes 192
 - informations sur les disques 108, 163
- Affichage de sous-disques
 - vxprint 203
- Affichage DMP 135
- Ajout
 - d'un disque 97, 138
 - d'un disque à un disk group 99
 - d'un disque pour une utilisation future 155
 - disque à un disk group 170
 - journal des zones modifiées 188
 - journal RAID 5 188
- Ajout de disques 93
 - format 94
- Association de miroirs
 - vxmake 195
- Association de plex
 - vxmake 195
- Association de sous-disques
 - vxmake 204
 - vxsd 204
- Association de sous-disques de journal 205
 - vxsd 206
- Attribution des données 302
- Augmentation de la taille du volume 183

B

- Basculement de chemin
 - DMP 83

C

- Chemin d'accès à un périphérique 142
- Chemins d'accès multiples
 - DMP 82
- Cluster
 - disques 225
 - environnement 217
 - fonctionnalité 217

- objets partagés 218
- Colonne de bande 30
- Colonnes, dans la répartition des données par bandes sur plusieurs disques 30
- Concaténation 28
- Concaténé
 - volumes, création 181
- Contrôle 80
- Contrôle racine
 - nettoyage 265
- Contrôleur 142
- Conversion
 - structure de stockage 63
- Copie de miroirs
 - vxplex 202
- Création
 - disk group 170
 - volume concaténé 181
 - volume RAID 5 183
 - volume réparti 181
 - volume réparti par bandes 182
 - volume sur un disque VM 182
- Création de disk groups
 - vxdbg 125
- Création de miroirs
 - vxmake 194
- Création de sous-disques 203
 - vxmake 203
- Cycle des états des plex 275

D

- Défaillances 249, 282, 286
 - disques 110, 282
 - procédures de récupération 249
 - système 281
- Défaillances de disques
 - récupération 243
- Démarrage
 - après défaillance 245
 - vxdiskadm 137
- Démarrage à partir des périphériques DMP
 - DMP 84
- Démarrage des volumes 289
- Démarrage forcé des volumes 291
- Démon d'E/S 52
- Démons 51
 - reconstruction automatique 69
 - Volume Manager 51
 - vxrelocd 113

- Déplacement
 - des volumes à partir d'un disque 140
 - des volumes d'un disque 213
- Déplacement de disk groups
 - vx dg 128
 - vx recover 129
- Déplacement de disques 109
- Déplacement de miroirs 201
 - vxplex 201
- Déplacement des sous-disques
 - RAID 5 289
 - vx sd 207
- Désactivation
 - d'un disk group 140, 173
 - d'un disque 140, 165
- devname 142
- Disk group partageable par le cluster 218
- Disk group privé 218
- Disk groups 21, 91
 - activation 140
 - ajout d'un disque 170
 - création 170
 - déplacement 128, 129
 - déplacement d'un système à un autre 128
 - déportation 129
 - désactivation 173
 - exportation 128, 173
 - importation 128, 129, 174
 - initialisation 170
 - par défaut 170
 - renommer 130
 - suppression 128
- disk groups 91
- Disque d'initialisation
 - défaillance 256
 - défaillance et reconstruction automatique 256
 - mise en miroir 105
 - rajout 257
 - remplacement 256, 259
- Disque racine 60, 80, 245, 256
 - mise en miroir 105
 - rajout 256
- Disques 142, 282
 - activation 166
 - affichage d'informations 163
 - ajout 93
 - ajout pour une utilisation future 155
 - dans un cluster 225
 - défaillance
 - reconstruction automatique 69
 - défaillances 110, 282
 - reconstruction automatique 110
 - récupération 243
 - déplacement 109
 - déplacement des volumes 213
 - désactivation 165
 - détachés 119
 - disque d'initialisation 105, 256
 - disque racine 60, 80, 105, 245, 256
 - disques de remplacement pour la reconstruction automatique 114, 158
 - disques VM 21
 - encapsulation 60, 80, 133, 138, 152, 245
 - initialisation 93
 - mise en miroir de volumes 212
 - physiques
 - activation 140
 - affichage d'informations 108
 - ajout 97, 138
 - ajout à un disk group 99
 - déplacement des volumes 140
 - désactivation 140
 - mise hors ligne 105
 - remplacement 139, 167
 - réservation 104
 - sous le contrôle de VxVM 97
 - suppression 107, 139
 - rajout 256
 - rattachement 260
 - réinitialisation 155
 - remplacement 121, 122, 256
 - suppression 164
 - VM
 - création d'un volume 182
 - volatiles 135
- Disques ayant échoué 110
- Disques VM 21
 - définition 21
- Dissociation de miroirs
 - vxplex 196
- Dissociation des miroirs
 - vxplex 187
- Division des sous-disques
 - vx sd 207
- DMP
 - chemins d'accès multiples 82

- démarrage à partir des périphériques
 - DMP 84
 - équilibrage de charge 84
 - mécanisme de basculement de chemin 83
- Données
 - conservation 60
 - redondance 35
- Données de performances 307
 - obtention 307
 - utilisation 308
- DRL 188
- E**
 - E/S
 - statistiques 308
 - obtention 307
 - suivi 308, 311
 - Écritures
 - bande complète 296
 - lecture-modification 294
 - reconstruction 298
 - Éléments réglables 312
 - Encapsulation 60, 80, 133, 138, 152, 245
 - Encapsulation des disques 138, 152
 - Encapsulations spéciales
 - vxdisk 133
 - Environnement cluster 87
 - Équilibrage de charge
 - DMP 84
 - États
 - plex 272
 - volume 275
 - États des plex
 - ACTIVE 273
 - CLEAN 273
 - EMPTY 272
 - IOFAIL 274
 - OFFLINE 273
 - STALE 273
 - TEMP 274
 - TEMPRM 274
 - États des volumes
 - ACTIVE 275, 276
 - CLEAN 275
 - EMPTY 275, 276
 - SYNC 275, 276
 - États du noyau du plex 275
 - DETACHED 275
 - DISABLED 275
 - ENABLED 275
 - États du noyau du volume 277
 - DETACHED 277
 - DISABLED 277
 - ENABLED 277
 - Exportation
 - d'un disk group 128, 140, 173
- F**
 - Fermeture
 - vxdiskadm 175
 - Fonction de cluster 87
- H**
 - Hôtes
 - multiples 217
- I**
 - Importation
 - d'un disk group 128
 - Importation d'un disk group 140, 174
 - Indicateur, autoboot 244
 - Informations 206
 - Informations sur les disques, affichage 163
 - Initialisation des disques 93
 - Interface de menu 136
 - Interface utilisateur graphique 45, 92
- J**
 - Jonction des sous-disques
 - vxsd 208
 - Journal
 - ajout 188
 - RAID 5 188
 - sous-disques 73, 205, 227
 - Journal des zones modifiées 188
 - dans un environnement cluster 227
 - recommandations 57
 - sous-disques de journal 205
 - Journal des zones modifiées (DRL) 73
 - Journalisation (logging) 41
 - Journaux 278, 287
- L**
 - Lecture
 - stratégies 303
 - Limitations des volumes
 - au démarrage 81
 - Limitations des volumes racine 81
 - Limitations des volumes swap 81

Liste des miroirs
vxprint 198

M

Manipulation des sous-disques 288

Mineurs

réservation 132

Miroirs 23, 24

affichage 198

création 194

dissociation 187, 196

mise hors ligne 200

récupération 120

sauvegarde 197

suppression 187, 196

Mise en miroir 35, 139, 303, 305

disque d'initialisation 105

disque racine 105

Recommandations 56

tous les volumes 187

volumes d'un disque 212

Mise hors ligne

d'un disque 105, 140, 165

Mode détérioré 282

Modification des attributs de volume 293

N

Nœud principal 220

Nœud secondaire 220

Nœuds 217

Nom

accès au disque 19

périphérique 142

Nom d'accès logique au disque 21, 90, 142

Nom de disque 90, 92

Nom de périphérique 90, 142

nopriv 133, 134

périphériques 135

O

Obtention des données de performances 307

OFFLINE 200

Opérations de disque VM 107

P

Parité 34, 37, 286

Partagés

objets 218

Partition 142

Partitions 19

UNIX 142

Performances 306

contrôle 306

gestion 301

optimisation 301

recommandations 301

types 306

périphériques de disques standard 90

Périphériques spéciaux

utilisation 133

Plage de protocole de cluster 238

Plex 23, 278

affichage 198

attachement 199, 201

copie 202

création 194

définition 23

déplacement 201

détachement 199

en tant que miroirs 24

liste 198

modification des informations 198

volume 24

Plex de journal 277

Plex RAID 5 277

Plex réparti par bandes 30

Plex, états 272

Point de reprise 286

Prise de vue 208, 209

RAID 5 208

Procédure, boot 244

Procédures de reconfiguration 262

putil 198

R

RAID 0 30

RAID 1 35

RAID 5 277, 278, 281, 282, 285, 286, 287, 288,
289, 290, 291, 293, 306

déplacement des sous-disques 289

journal 188

prise de vue 208

recommandations 59

récupération 284, 292

volumes

création 183

Rajout de disques 256

Rattachement de disques 260

Recommandations

journal des zones modifiées 57

- mise en miroir 56
 - mise en miroir et répartition des données par bandes sur plusieurs disques 58
 - RAID 5 59
 - Recommandations pour la configuration 301
 - Reconfiguration de cluster 222
 - Reconfiguration de volume 223
 - Reconstruction automatique 69, 110, 141, 158
 - désignation des disques de remplacement 70, 158
 - disque d'initialisation 256
 - modification de vxrelocd 112
 - suppression des remplacements 160
 - Reconstruction-lecture 282
 - Récupération 243, 292
 - journaux 286
 - procédures 249
 - volumes 192
 - volumes RAID 5 284, 292
 - Récupération de parité 285, 286
 - Réduction de la taille d'un volume 185
 - Réglage
 - Volume Manager 312
 - Réinitialisation des disques 155
 - Réinstallation 262, 263
 - Remplacement de disques 121, 139, 167, 256
 - vxdiskadm 121
 - Renommer les disk groups 130
 - Répartis par bandes
 - plex 30
 - Répartition 28
 - volumes
 - création 181
 - Répartition des données par bandes sur plusieurs disques 30, 302, 305
 - Répartition par bandes
 - volumes
 - création 182
 - Réservation de disques pour des utilisations spéciales 104
 - Resilvering 85
 - Restructuration en ligne 63
 - caractéristiques de la restructuration 68
 - principes de fonctionnement 64
 - récupération après défaillance 68
 - restructurations et longueur de volume 68
 - types de restructuration 65
 - Resynchronisation
 - bases de données Oracle 85
 - Resynchronisation des volumes 72
 - Resynchronisation rapide des miroirs (FMR) 74
 - rootdg 21, 91
 - renommer 130
- S**
- Sauvegarde d'un volume 191
 - Sauvegarde en ligne 208
 - Sauvegardes 61, 208
 - miroirs 197
 - vxassist 208
 - Sous-disques
 - affichage 203
 - association 203, 204
 - déplacement 207
 - dissociation 206
 - division 207
 - jonction 208
 - modification des informations 206
 - suppression 203
 - Sous-disques de journal 57, 73, 205, 227
 - association 205
 - Storage Administrator 45
 - Structure
 - symétrique gauche 39
 - Structure de stockage
 - conversion 63
 - Suivi des E/S 308
 - vxtrace 308
 - Suppression
 - d'un disque 139, 164
 - d'un disque physique 107
 - d'un volume 185
 - du journal des zones modifiées 188
 - Suppression de disk groups 128
 - vxdg 128
 - Suppression de disques
 - vxdg 102
 - Suppression de sous-disques
 - vxedit 203
 - Système
 - défaillances 281
- T**
- Taille
 - volumes
 - augmentation 183
 - réduction 185

tutil 198

U

Unités de bande 31

UNIX

partition 142

Utilisation de périphériques spéciaux 133

Utilisation des disk groups

vxassist 93

Utilisation des données de performances
308

Utilisation des statistiques d'E/S 308

Utilitaire format 94

Utilitaires de disk group 92

Utilitaires de la ligne de commande 92

V

Version de protocole de cluster 238

Volume

resynchronisation 72

Volume concaténé

création 181

Volume Manager 45, 63

contrôle 80

démons 51

interface utilisateur graphique 92

structures 44

système d'exploitation 43

Volume Manager Support Operations 91

Volume RAID 5

création 183

Volume réparti

création 181

Volume réparti par bandes

création 182

Volume, états 275

Volumes 24

affichage de la configuration 192

arrêt 189

augmentation de la taille 183

création

concaténé 181

définition 18, 24

démarrage 189

déplacement à partir d'un disque 213

mise en miroir 186

mise en miroir de tous les volumes

existants 187

mise en miroir sur un disque 212

nettoyage 266

plex 24

récupération 192

réduction de la taille 185

sauvegarde 191

stratégie de lecture 190

structure 277

suppression 185

Volumes inutilisables 290

Volumes racine

démarrage 81

Volumes RAID 5 292

vxassist 60, 93, 122, 186, 208, 209

création de volumes 179

growby 184

growto 184

sauvegarde 208

shrinkby 184

shrinkto 184

vxassist addlog 188

vxassist growby 184

vxassist growto 184

vxassist make 105, 182

vxassist snapshot 192

vxassist snapstart 191

vxclust 231

vxconfigd 51, 81, 232

vxctl 238

vxlg 93, 102, 125, 128, 130, 234

déplacement de disk groups 128

suppression de disk groups 128

vxlg rmdisk 107

vxdisk 93, 102, 133, 236

encapsulations spéciales 133

rm 102

vxdisk list 108

vxdisk rm 107

vxdiskadd 91, 92, 94, 97, 108, 123

vxdiskadm 80, 91, 92, 94, 102, 103, 121, 130

démarrage 137

remplacement de disques 121

vxedit 196, 198, 203, 206, 266

suppression de sous-disques 203

vxedit rename 104

vxedit set 104

vxinfo 121

vxiod 51, 52

vxmake 186, 194, 195, 202, 204

association de miroirs 195

association de sous-disques 204

- création de miroirs 194
 - création de sous-disques 203
 - vxmend 192, 199, 200, 201
 - vxmirror 245
 - vxplex 186, 187, 195, 196, 197, 199, 200, 201
 - copie de miroirs 202
 - déplacement de miroirs 201
 - dissociation des miroirs 187, 196
 - vxprint 119, 192, 198, 203
 - affichage de sous-disques 203
 - liste des miroirs 198
 - vxreattach 260
 - vxrecover 120, 237
 - déplacement de disk groups 129
 - vxrelocd 69, 113
 - modification 112
 - vxsd 204, 206, 207
 - association des sous-disques 204
 - association des sous-disques de journal 206
 - déplacement des sous-disques 207
 - division des sous-disques 207
 - jonction des sous-disques 208
 - VxSmartSync 85
 - vxstat 119, 239, 307, 309
 - vxtrace 307, 308
 - VxVM 45, 63
 - vxvol 122, 200
- Z**
- zone privée 91
 - zone publique 91