



Guide d'administration du système de Sun Cluster 2.2

901 San Antonio Road
Palo Alto, , CA 94303-4900
Etats-Unis 650 960-1300 Fax 650 969-9131

Référence 806-5341-10
Juin 2000

Copyright Copyright 2000 Sun Microsystems, Inc. 901 San Antonio Road, Palo Alto, California 94303-4900 U.S.A. All rights reserved.

Ce produit ou document est protégé par copyright et distribué sous des licences qui en restreignent l'utilisation, la copie, la distribution et la décompilation. Aucune partie de ce produit ou document ne peut être reproduite sous aucune forme, par quelque moyen que ce soit, sans l'autorisation préalable et écrite de Sun et de ses bailleurs de licence, s'il y en a. Les logiciels de tierce partie, dont la technologie relative aux polices, sont protégés par copyright et licenciés par des fournisseurs de Sun.

Certaines parties de ce produit peuvent provenir des systèmes Berkeley BSD, licenciés par l'Université de Californie (University of California). UNIX est une marque déposée aux Etats-Unis et dans d'autres pays et licenciée exclusivement par X/Open Company, Ltd. La notice suivante est applicable à Netscape Communicator™: (c) Copyright 1995 Netscape Communications Corporation Tous droits réservés. Sun, Sun Microsystems, le logo Sun, AnswerBook2, docs.sun.com, et Solaris sont des marques de fabrique, des marques déposées, ou des marques de service, de Sun Microsystems, Inc. aux Etats-Unis et dans d'autres pays. Toutes les marques SPARC, utilisées sous licence, sont des marques de commerce ou des marques déposées appartenant à SPARC International, Inc. aux Etats-Unis et dans d'autres pays. Les produits portant des marques SPARC sont fondés sur une architecture développée par Sun Microsystems, Inc.

L'interface d'utilisation graphique OPEN LOOK et Sun™ a été développée par Sun Microsystems, Inc. pour ses utilisateurs et licenciés. Sun reconnaît les efforts novateurs de recherche entrepris par Xerox pour développer le concept des interfaces d'utilisation graphiques ou visuelles destinées à l'industrie de l'informatique. Sun détient une licence non exclusive de Xerox lui permettant d'exploiter l'interface d'utilisation graphique Xerox. Cette licence couvre aussi les licenciés de Sun qui mettent en oeuvre des interfaces graphiques OPEN LOOK en se conformant aux licences écrites de Sun.

CETTE PUBLICATION EST FOURNIE "EN L'ETAT" ET AUCUNE GARANTIE, EXPRESSE OU IMPLICITE, N'EST ACCORDEE, Y COMPRIS DES GARANTIES CONCERNANT LA VALEUR MARCHANDE, L'APTITUDE DE LA PUBLICATION A REPENDRE A UNE UTILISATION PARTICULIERE, OU LE FAIT QU'ELLE NE SOIT PAS CONTREFAISANTE DE PRODUIT DE TIERS. CE DENI DE GARANTIE NE S'APPLIQUERAIT PAS, DANS LA MESURE OU IL SERAIT TENU JURIDIQUEMENT NUL ET NON AVENU.



Sommaire

Préface 15

1. Préparation en vue de l'administration de Sun Cluster 21

Enregistrement des informations de partitionnement d'un disque (Solstice DiskSuite) 21

Enregistrement et restauration des informations VTOC (Solstice DiskSuite) 23

Enregistrement des informations de configuration de périphérique 24

Désignation et numérotation des instances 24

 Réinitialisations de reconfiguration 26

Connexion sur le serveur en qualité de `root` (superutilisateur) 27

2. Outils d'administration système de Sun Cluster 29

Utilitaires de surveillance 29

 Surveillance de la configuration avec `hastat(1M)` 30

 Consultation des fichiers de messages 32

 Utilitaires de service de données haute disponibilité 33

Système d'aide en ligne 33

 Éléments de la barre de menus de la fenêtre d'aide 36

 Boutons de la fenêtre Aide 38

Panneau de commandes de grappe 38

▼ Comment démarrer le PCG 39

Eléments du PCG	39
Emplacements des fichiers de configuration du PCG	40
Console de grappe	41
▼ Comment démarrer la console de grappe	41
Barre de menus de la fenêtre commune	42
Fichiers de configuration utilisés par la console de grappe	43
A propos du fichier <code>clusters</code>	43
A propos du fichier <code>serialports</code>	43
Utilisation de Sun Cluster Manager	44
Exécution de SCM en tant qu'application	45
▼ Comment exécuter SCM en tant qu'application	45
Exécution de SCM en tant qu'applet	46
▼ Comment configurer la trousse d'outils JDK	47
▼ Comment télécharger HotJava	48
▼ Comment exécuter l'applet SCM dans un navigateur HotJava à partir d'un noeud de grappe	48
▼ Comment exécuter l'applet SCM dans un navigateur HotJava à partir du poste de travail administratif	49
▼ Comment exécuter l'applet SCM dans un navigateur Netscape à partir d'un noeud de grappe	50
▼ Comment installer un serveur Web pour SCM	51
Accès à l'aide en ligne de SCM	52
3. Modification de la configuration de Sun Cluster	53
Ajout et suppression de noeuds de grappe	54
▼ Comment ajouter un noeud de grappe	54
▼ Comment supprimer un noeud de grappe	57
Modification du nom d'un noeud de grappe	57
▼ Comment modifier le nom d'un noeud de grappe	58
Modification des interfaces réseau privées	58

▼ Comment modifier les interfaces réseau privées	59
Impression de la configuration des noeuds	59
▼ Comment imprimer la configuration de la grappe	59
Ajout et suppression d'hôtes logiques	60
▼ Comment ajouter un hôte logique à la grappe	61
▼ Comment supprimer un hôte logique de la grappe	63
Modification de l'adresse IP de l'hôte logique	65
▼ Comment modifier l'adresse IP d'un hôte logique	66
Reconfiguration de grappe forcée	66
▼ Comment forcer une reconfiguration de grappe	66
Configuration des services de données de Sun Cluster	67
▼ Comment configurer un service de données Sun Cluster	67
Suppression de la configuration des services de données Sun Cluster	68
▼ Comment supprimer la configuration des services de données Sun Cluster	69
Ajout de services de données Sun Cluster	70
▼ Comment ajouter un service de données à une grappe à deux noeuds avec une BCG partagée	70
Modification du fichier <code>clusters</code>	72
▼ Comment modifier le fichier <code>clusters</code>	73
▼ Comment créer la table <code>clusters</code>	73
Modification du fichier <code>serialports</code>	74
▼ Comment modifier le fichier <code>serialports</code>	74
▼ Comment créer la table <code>serialports</code>	75
Modification des informations TC/PSS	75
▼ Comment modifier les informations CT/PSS	76
Modification du périphérique de quorum	77
▼ Comment modifier le périphérique de quorum	78
Configuration des délais de transition de la grappe	80

- ▼ Comment modifier les délais de la grappe 81
- 4. Administration générale de Sun Cluster 83**
- Démarrage de la grappe et des noeuds de grappe 83
 - ▼ Comment démarrer la grappe 84
- Arrêt de la grappe et des noeuds de grappe 86
 - ▼ Comment arrêter Sun Cluster sur un noeud de grappe 87
 - ▼ Comment arrêter Sun Cluster sur tous les noeuds 88
 - ▼ Comment arrêter un noeud Sun Cluster 89
 - ▼ Comment supprimer un verrouillage de noeud après une erreur grave de grappe 89
 - Arrêt du moniteur d'appartenance pendant l'exécution d'instances RDBMS 91
- Commutation des hôtes logiques 92
- Désactivation de la commutation automatique 93
- Activation du mode maintenance des hôtes logiques 93
- Récupération à partir des partitions de grappe 94
 - Partitions de double contrôle (VxVM seulement) 95
 - ▼ Comment sélectionner une nouvelle grappe 96
- Maintenance du système de fichiers `/var` 97
 - ▼ Comment réparer un système de fichiers `/var` saturé 97
- Administration du temps dans les configurations de Sun Cluster 98
- Remplacement d'un noeud défectueux 99
 - ▼ Comment remplacer un noeud défectueux 99
- Remplacement d'un concentrateur de terminaux défectueux 103
- Administration de la base de données de configuration de grappe 104
 - ▼ Comment vérifier la cohérence générale de la BCG 104
 - ▼ Comment sauvegarder la BCG 105
 - ▼ Comment restaurer la BCG 105
 - ▼ Comment activer et désactiver le quorum BCG 106

- ▼ Comment purifier la BCG 107
 - Dépannage de la BCG 107
- Réservation de disques partagés (VxVM) 109
- ▼ Comment réserver des périphériques partagés (VxVM) 109
- 5. Récupération après une panne de courant 111**
 - Récupération après une panne de courant généralisée 111
 - Récupération après une panne de courant partielle 112
 - Panne d'un seul noeud 112
 - Panne d'une unité d'expansion de disques multihôtes 113
 - Panne d'un serveur et d'une unité d'expansion de disques multihôtes 113
 - Mise sous tension du système 113
- 6. Administration des interfaces réseau 115**
 - Aperçu de la gestion de réseau public 115
 - Surveillance des défaillances et reprise par le service PNM 116
 - Groupes de sauvegarde 117
 - Mises à jour du fichier `nsswitch.conf` 118
 - Configuration et administration de la gestion de réseau public 119
 - ▼ Comment configurer le service PNM 119
 - ▼ Comment reconfigurer le service PNM 123
 - ▼ Comment vérifier l'état des groupes de sauvegarde 124
 - Paramètres configurables du service PNM 125
 - Résolution des erreurs du service PNM 126
 - Ajout et suppression d'interfaces réseau 127
 - Ajout d'une interface réseau 127
 - ▼ Comment ajouter une interface réseau 128
 - Suppression d'une interface réseau 130
 - ▼ Comment supprimer une interface réseau 130

- Administration de l'Agent de gestion de commutation 133
 - ▼ Comment ajouter des commutateurs et des cartes SCI 133
 - Dépannage du logiciel SCI 134
 - ▼ Comment vérifier la connectivité entre les noeuds 134
 - ▼ Comment vérifier la configuration des interfaces SCI 136
- 7. **Administration des composants serveurs 139**
 - Remplacement de cartes système 139
 - Ajout de modules de carte 140
 - ▼ Comment ajouter des modules de carte 140
 - Remplacement de cartes SBUS 141
 - ▼ Comment remplacer les cartes SBus 141
- 8. **Administration du concentrateur de terminaux 143**
 - Connexion à la console Sun Cluster 143
 - ▼ Comment se connecter à la console Sun Cluster 144
 - Réinitialisation d'une connexion au concentrateur de terminaux 146
 - ▼ Comment réinitialiser une connexion au concentrateur de terminaux 146
 - Entrée dans la mémoire PROM OpenBoot d'un serveur Sun Cluster 148
 - ▼ Comment entrer dans la mémoire PROM OpenBoot 148
 - Dépannage du concentrateur de terminaux 149
 - Erreurs d'accès à la configuration des ports 149
 - ▼ Comment corriger une erreur d'accès à la configuration d'un port 149
 - Interruptions aléatoires des connexions au concentrateur de terminaux 150
 - ▼ Comment établir un routage par défaut 151
 - Modification des informations CT/PSS 153
 - ▼ Comment modifier les informations relatives à l'hôte 154
 - ▼ Comment spécifier un numéro de port pour un PSS ou un CT 154
 - ▼ Comment modifier la configuration d'un CT 155

- ▼ Comment changer le mot de passe d'un PSS ou d'un CT 155
- 9. Utilisation de médiateurs dans une configuration à deux chaînes 157**
 - Aperçu des médiateurs 157
 - Médiateurs or 158
 - Configuration des médiateurs 159
 - Défaillances traitées à l'aide de médiateurs 160
 - Panne sur un serveur unique 160
 - Défaillance d'une seule chaîne 161
 - Défaillance d'un hôte et d'une chaîne 162
 - Administration des médiateurs 163
 - ▼ Comment ajouter des hôtes médiateurs 164
 - ▼ Comment vérifier l'état des données de médiateur 165
 - ▼ Comment corriger des données de médiateur erronées 165
 - Gestion des défaillances sans reprise automatique 166
 - Messages d'erreur relatifs aux médiateurs 168
- 10. Administration de disques locaux Sun Cluster 169**
 - Restauration d'un disque d'amorçage local à partir d'une copie de sauvegarde 170
 - ▼ Comment restaurer un disque d'amorçage local à partir d'une copie de sauvegarde (Solstice DiskSuite) 170
 - ▼ Comment restaurer un disque d'amorçage local à partir d'une copie de sauvegarde (VxVM) 172
 - Remplacement d'un disque local qui ne sert pas à l'amorçage 173
 - ▼ Comment remplacer un disque local qui ne sert pas à l'amorçage 174
- 11. Administration des tableaux SPARCstorage Array 177**
 - Reprise après une panne de courant 177
 - ▼ Comment effectuer une reprise après une panne de courant (Solstice DiskSuite) 178
 - ▼ Comment effectuer une reprise après une panne de courant (VxVM) 180
 - Réparation d'une connexion SPARCstorage Array 182

- ▼ Comment réparer une connexion SPARCstorage Array 183
- Ajout d'un tableau SPARCstorage Array 184
- ▼ Comment ajouter un tableau SPARCstorage Array 184
- Administration des plateaux d'un tableau SPARCstorage Array 185
- ▼ Comment mettre hors service un plateau de tableau SPARCstorage Array(Solstice DiskSuite) 186
- ▼ Comment mettre hors service un plateau de tableau SPARCstorage Array(VxVM) 188
- ▼ Comment remettre en service un plateau de tableau SPARCstorage Array(Solstice DiskSuite) 190
- ▼ Comment remettre en service un plateau de tableau SPARCstorage Array(VxVM) 192
- Remplacement d'un contrôleur du tableau SPARCstorage Array et modification du nom universel 193
- ▼ Comment modifier le nom universel d'un tableau SPARCstorage Array au moyen d'un système de maintenance 194
- ▼ Comment modifier le nom universel d'un tableau SPARCstorage Array 198
- Administration des disques SPARCstorage Array 203
 - Ajout d'un disque de tableau SPARCstorage Array 204
 - ▼ Comment ajouter un disque de tableau SPARCstorage Array (Solstice DiskSuite) 204
 - ▼ Comment ajouter un disque de tableau SPARCstorage Array (VxVM) 211
 - Remplacement d'un disque de tableau SPARCstorage Array 215
 - ▼ Comment remplacer un disque de tableau SPARCstorage Array(Solstice DiskSuite) 216
 - ▼ Comment remplacer un disque de tableau SPARCstorage Array(VxVM) 225
- Administration de la mémoire vive non-volatile (NVRAM) d'un tableau SPARCstorage Array 228
 - Activation et désactivation de la mémoire NVRAM 229
 - ▼ Comment activer et désactiver la mémoire NVRAM 229

Vidage et purge de la mémoire NVRAM 231

▼ Comment vider et purger la mémoire NVRAM 232

12. Administration des disques Sun StorEdge MultiPack et Sun StorEdge D1000 235

Reprise après une panne de courant 235

▼ Comment effectuer une reprise après une panne de courant (Solstice DiskSuite) 236

▼ Comment effectuer une reprise après une panne de courant (VxVM) 237

Administration des disques Sun StorEdge MultiPack et Sun StorEdge D1000 240

Réparation d'une connexion Sun StorEdge MultiPack ou Sun StorEdge D1000 240

▼ Comment réparer une connexion Sun StorEdge MultiPack ou Sun StorEdge D1000 241

Ajout d'un disque Sun StorEdge MultiPack ou Sun StorEdge D1000 241

▼ Comment ajouter un disque Sun StorEdge MultiPack ou Sun StorEdge D1000 242

Administration des disques Sun StorEdge MultiPack et Sun StorEdge D1000 243

Ajout de disques Sun StorEdge MultiPack ou Sun StorEdge D1000 243

▼ Comment ajouter un disque Sun StorEdge MultiPack ou Sun StorEdge D1000 244

Remplacement des disques Sun StorEdge MultiPack ou Sun StorEdge D1000 246

▼ Comment remplacer un disque Sun StorEdge MultiPack ou Sun StorEdge D1000 (Solstice DiskSuite) 246

▼ Comment remplacer un disque Sun StorEdge MultiPack ou Sun StorEdge D1000 (VxVM) 250

Remplacement des boîtiers Sun StorEdge MultiPack ou Sun StorEdge D1000 255

▼ Comment remplacer un boîtier Sun StorEdge MultiPack ou Sun StorEdge D1000 (VxVM) 255

13. Administration des disques Sun StorEdge A3x00 et Sun StorEdge A1000 263

Sources d'alimentation 263

- Ajout de boîtiers Sun StorEdge A3x00 ou Sun StorEdge A1000 264
 - ▼ Comment ajouter un boîtier de disque 264
- Administration des disques Sun StorEdge A3x00 ou Sun StorEdge A1000 265
 - Ajout d'un disque Sun StorEdge A3x00 ou Sun StorEdge A1000 265
 - Remplacement d'un disque Sun StorEdge A3x00 ou Sun StorEdge A1000 266
- 14. Administration des disques Sun StorEdge A5000 267**
 - Reprise après une panne de courant 267
 - ▼ Comment effectuer une reprise après une panne de courant (Solstice DiskSuite) 268
 - ▼ Comment effectuer une reprise après une panne de courant (VxVM) 269
 - Administration des disques Sun StorEdge A5000 272
 - Réparation d'une connexion Sun StorEdge A5000 272
 - ▼ Comment réparer une connexion Sun StorEdge A5000 273
 - Administration des disques Sun StorEdge A5000 273
 - Ajout ou remplacement de disques Sun StorEdge A5000 273
 - ▼ Comment ajouter un disque Sun StorEdge A5000 (Solstice DiskSuite) 274
 - ▼ Comment ajouter un disque Sun StorEdge A5000 (VxVM) 274
 - ▼ Comment remplacer un disque Sun StorEdge A5000 (Solstice DiskSuite) 275
 - ▼ Comment remplacer un disque Sun StorEdge A5000 (VxVM) 278
- A. Administration des gestionnaires de volumes 281**
 - Utilisation de Solstice DiskSuite dans l'environnement Sun Cluster 281
 - Administration de métapériphériques et d'ensembles de disques 282
 - Ajout d'un disque à un ensemble 283
 - ▼ Comment ajouter un disque à un ensemble (Solstice DiskSuite) 283
 - Retrait d'un disque d'un ensemble de disques 284
 - ▼ Comment retirer un disque d'un ensemble de disques (Solstice DiskSuite) 284

	Administration de métapériphériques multihôtes	284
	Administration de métapériphériques locaux	289
	Actions de métapériphérique destructives	289
	Utilisation de VxVM dans l'environnement Sun Cluster	289
	Aperçu de l'administration d'objets VxVM	290
	Administration des disques	291
	▼ Comment initialiser et configurer un disque (VxVM)	291
	Administration des groupes de disques	292
	Administration des objets VxVM	296
	Administration de miroirs locaux	299
	Sauvegarde de données multihôtes à l'aide de Solstice Backup	299
B.	Détection des défaillances de Sun Cluster	303
	Aperçu de la détection de défaillances	304
	Mécanisme de pulsation : moniteur d'appartenance à la grappe	304
	Vérification de validité du noeud qui effectue la vérification	305
	Surveillance du réseau public (PNM)	306
	Vérification des défaillances de Sun Cluster	307
	Vérification des défaillances propres à un service de données	308
	Vérification des défaillances de Sun Cluster HA pour NFS	309
	Vérification des défaillances SGBD-HD	310
	Vérification des défaillances de Sun Cluster HA pour Netscape	312
	Vérification des défaillances de Sun Cluster HA pour Lotus	314
	Vérification des défaillances de Sun Cluster HA pour Tivoli	315
	Vérification des défaillances de Sun Cluster HA pour SAP	315
C.	Utilisation des méthodes de gestion SNMP de Sun Cluster	317
	Base d'informations de l'agent SNMP de grappe et de gestion de grappe	318
	Base d'informations de gestion de grappe	319
	Attributs du <code>clustersTable</code>	320

Attributs du <code>clusterNodesTable</code>	320
Attributs du <code>switchesTable</code>	321
Attributs du <code>portsTable</code>	322
Attributs du <code>lhostTable</code>	323
Attributs du <code>dsTable</code>	324
Attributs de <code>dsinstTable</code>	324
Utilisation du démon SNMP de grappe et du démon super moniteur	325
Trappes SNMP	326
Modification du fichier <code>snmpd.conf</code>	330
▼ Comment modifier le fichier <code>snmpd.conf</code>	331
Configuration du port de l'agent SNMP de grappe	332
▼ Comment configurer le port de l'agent SNMP de grappe	332
Utilisation de l'agent SNMP avec SunNet Manager	333
▼ Comment utiliser l'agent SNMP avec SunNet Manager pour surveiller les grappes	334
▼ Comment reconfigurer <code>smond</code> pour surveiller une autre grappe	334
Dépannage de SNMP	335
D. Glossaire	337

Préface

Sun[™] Cluster 2.2 est un produit logiciel prenant en charge des configurations particulières fondées sur deux à quatre serveurs de noeuds. Ce logiciel est compatible avec les environnements Solaris 2.6, Solaris 7 et Solaris 8. Lorsqu'ils sont convenablement configurés, le matériel et le logiciel assurent ensemble des services de données à disponibilité élevée et un accès parallèle aux bases de données. Sun Cluster utilise les fonctions d'écriture miroir et de groupe de disques assurées par un gestionnaire de volumes. Sun Cluster prend en charge Solstice DiskSuite et VERITAS Volume Manager (VxVM). La fonction de grappe VxVM est prise en charge avec le service de données Oracle Parallel Server.

Ce manuel décrit les procédures de configuration du matériel et d'installation, de configuration et d'administration du logiciel Sun Cluster. Ce manuel doit être utilisé de pair avec les manuels énumérés à la section "Documentation connexe", page 16.

A qui ce guide est-il destiné ?

Ce guide s'adresse aux représentants et administrateurs de système Sun auxquels incombe l'installation et la maintenance des configurations de Sun Cluster 2.2. Les instructions et les sujets abordés sont complexes et destinés à des techniciens chevronnés.

Les instructions décrites dans ce manuel supposent que le lecteur possède une connaissance approfondie d'au moins un des gestionnaires de volumes utilisés avec Sun Cluster.

Les administrateurs de système connaissant le système UNIX[®] trouveront ce guide utile pour l'administration des configurations Sun Cluster 2.2.

Remarque - Il est déconseillé aux administrateurs moins expérimentés de tenter d'installer, de configurer ou d'administrer les configurations Sun Cluster 2.2.

Structure de ce guide

Ce guide est divisé en sections, couvrant chacune un sujet important relatif à l'administration du système. Chaque section est divisée en chapitres contenant un aperçu et des renseignements sur les tâches qui y sont décrites.

La plupart des renseignements contenus dans l'aperçu sont repris plus en détail dans les premiers chapitres de chaque section, et les chapitres subséquents contiennent des instructions détaillées sur l'exécution des tâches d'administration du système.

Documentation connexe

Les documents énumérés dans le Tableau P-1 contiennent des renseignements utiles pour l'administrateur système ou le fournisseur de services. Vous devez également conserver à portée de main les manuels d'installation et d'entretien du matériel.

TABLEAU P-1 Liste des documents connexes

Gamme de produits	Titre	Numéro de référence
Sun Cluster	<i>Sun Cluster 2.2 Software Installation Guide</i>	806-5342
	<i>Sun Cluster 2.2 API Developer's Guide</i>	806-5344
	<i>Sun Cluster 2.2 Error Messages Manual</i>	805-4242
	<i>Sun Cluster 2.2 Release Notes</i>	806-5345
	<i>Sun Cluster 2.2 Hardware Site Preparation, Planning, and Installation Guide</i>	806-5346
	<i>Sun Cluster 2.2 Hardware Service Manual</i>	806-5347

TABLEAU P-1 Liste des documents connexes (suite)

Gamme de produits	Titre	Numéro de référence
Solstice DiskSuite	<i>Installation et Notes sur le produit Solstice DiskSuite 4.2</i>	805-5960
	<i>Guide de l'utilisateur de Solstice DiskSuite 4.2</i>	805-5961
	<i>Guide de référence Solstice DiskSuite 4.2</i>	805-5962
	<i>Guide de référence Solstice DiskSuite 4.2.1</i>	806-3204
	<i>Guide de l'utilisateur de Solstice DiskSuite 4.2.1</i>	806-3205
	<i>Installation et Notes sur Solstice DiskSuite 4.2.1</i>	806-3206
VERITAS Volume Manager	<i>Veritas Volume Manager for Solaris</i>	100-001125A (VERITAS)

Règles typographiques

Le tableau qui suit énumère les règles typographiques utilisées dans ce guide.

TABLEAU P-2 Règles typographiques

Caractère ou symbole	Signification	Exemple
Courier	Nom des commandes, des fichiers et des répertoires ; affichage à l'écran.	Modifiez le fichier <code>.login</code> . Utilisez <code>ls -a</code> pour afficher la liste de tous les fichiers. <code>machine_name% Vous avez du courrier.</code>
gras	Ce que vous tapez, par contraste avec ce qui est affiché à l'écran.	<code>nom_de_l'ordinateur% su</code> <code>Mot de passe :</code>
<i>italique</i>	Paramètre substituable de ligne de commande : remplacez-la par un nom ou une valeur. Titres de publication, nouveaux mots ou termes mis en évidence.	Pour supprimer un fichier, tapez <code>rm <i>nom_du_fichier</i></code> .

Invites du shell dans des exemples de commandes

Le tableau suivant indique l'invite système par défaut et l'invite de root (superutilisateur) pour les shells C, Bourne et Korn.

TABLEAU P-3 Invites du shell

Shell	Invite
Invite du shell C	<code>nom_de_l'ordinateur%</code>
Invite de superutilisateur du shell C	<code>nom_de_l'ordinateur#</code>

TABLEAU P-3 Invites du shell (suite)

Shell	Invite
Invite des shells Bourne et Korn	\$
Invite de superutilisateur des shells Bourne et Korn	#

Comment obtenir de l'aide

Si vous n'arrivez pas à installer ou à utiliser le logiciel Sun Cluster, adressez-vous à votre fournisseur de services et communiquez-lui les renseignements suivants :

- Votre nom et votre adresse de courrier électronique (s'il y a lieu).
- Le nom, l'adresse et le numéro de téléphone de votre société.
- Les numéros de modèle et de série de vos systèmes.
- Le numéro de version du système d'exploitation (Solaris 2.6, par exemple).
- Le numéro de version de Sun Cluster (Sun Cluster 2.2, par exemple).

Utilisez les commandes suivantes pour obtenir des informations sur votre système :

TABLEAU P-4 Comment obtenir de l'aide

<code>prtconf -vprtconf</code>	Affiche la taille de la mémoire du système et fournit des renseignements sur les périphériques.
<code>psrinfo -vpsrinfo</code>	Affiche des informations sur les processeurs.
<code>showrev --p</code>	Indique les fichiers correctifs installés.
<code>prtdiag -v</code>	Affiche des données de diagnostic sur le système.

Ayez également à portée de la main le contenu du fichier `/var/adm/messages`.

Préparation en vue de l'administration de Sun Cluster

Ce chapitre décrit les procédures de préparation en vue de l'administration d'une configuration Sun Cluster. Certaines des procédures décrites varient en fonction du logiciel de gestion de volumes utilisé (Solstice DiskSuite ou VERITAS Volume Manager). Lorsqu'une procédure est tributaire du gestionnaire de volumes, son nom inclut le nom de ce gestionnaire. Ce chapitre comprend les sections suivantes :

- "Enregistrement des informations de partitionnement d'un disque (Solstice DiskSuite)", page 21
- "Enregistrement et restauration des informations VTOC (Solstice DiskSuite)", page 23
- "Enregistrement des informations de configuration de périphérique", page 24
- "Désignation et numérotation des instances", page 24
- "Connexion sur le serveur en qualité de root (superutilisateur)", page 27

Enregistrement des informations de partitionnement d'un disque (Solstice DiskSuite)

Conservez les informations de partitionnement des disques sur tous les noeuds et disques multihôtes dans la configuration Sun Cluster. Mettez à jour ces données chaque fois que de nouveaux disques sont ajoutés aux ensembles ou que vous repartitionnez des disques, car vous en aurez besoin lors du remplacement de disques.

Les informations de partitionnement des disques locaux ne sont pas aussi essentielles, car les disques locaux présents sur tous les noeuds Sun Cluster devraient avoir été partitionnés de la même façon. Lorsqu'un disque local tombe en panne, vous pouvez en général obtenir les informations de partitionnement à partir d'un autre noeud Sun Cluster.

Lorsque vous remplacez un disque multihôtes, le disque de remplacement doit être partitionné de la même façon que le disque défectueux. Selon le type de panne, il est possible que ces informations ne soient pas disponibles lorsque le remplacement est effectué. Il est donc particulièrement important de conserver un exemplaire des informations de partitionnement lorsque vous utilisez plusieurs types de partitionnement pour vos ensembles de disques.

Remarque - Bien que VxVM n'impose pas cette opération, il peut être très utile d'enregistrer ces données.

Une méthode d'enregistrement de ces informations est illustrée dans le script de l'exemple qui suit. Avant d'exécuter ce type de script, vous devez configurer le logiciel Sun Cluster. Dans cet exemple, les fichiers contenant les informations de la table des matières de volume (VTOC) sont écrits dans le répertoire local `/etc/opt/SUNWcluster/vtoc` au moyen de la commande `prtvtoc(1M)`.

```
#!/bin/sh
DIR=/etc/opt/SUNWcluster/vtoc
mkdir -p $DIR
cd /dev/rdisk
for i in *s7
do prtvtoc $i >$DIR/$i || rm $DIR/$i
done
```

Chaque disque d'un ensemble de disques de Solstice DiskSuite doit comporter une tranche 7. Cette tranche contient les répliques de base de données d'état des métapériphériques.

Lorsqu'un disque local contient également une tranche 7 valide, les informations VTOC sont également enregistrées par le script d'exemple. Toutefois, cette opération ne devrait pas intervenir dans le cas du disque d'initialisation, car celui-ci ne comporte pas habituellement de tranche 7 valide.

Remarque - Avant d'exécuter le script, assurez-vous qu'aucun disque n'appartient à un autre noeud Sun Cluster. Pour que ce script fonctionne, il faut que les hôtes logiques soient en mode de maintenance ou qu'ils appartiennent à l'hôte local, ou encore que Sun Cluster soit arrêté.

Enregistrement et restauration des informations VTOC (Solstice DiskSuite)

Lorsque vous enregistrez les données VTOC pour tous les disques multihôtes, vous pouvez ensuite utiliser ces informations lorsque vous remplacerez un disque. Le script de l'exemple qui suit utilise les informations VTOC enregistrées par le script ci-dessous pour partitionner le disque de remplacement de la même façon que le disque défectueux. En situation d'exploitation réelle, il suffit de remplacer `c1t0d0s7` et `c1t0d1s7` par le nom de chaque disque. Vous pouvez également spécifier plusieurs disques en entrant leur nom respectif un à la suite de l'autre, en les séparant par un espace.

```
#!/bin/sh
DIR=/etc/opt/SUNWcluster/vtoc
cd /dev/rdisk
for i in c1t0d0s7 c1t0d1s7
do fmthard -s $DIR/$i $i
done
```

Remarque - Le disque de remplacement doit avoir la même capacité et la même géométrie (et en général être du même modèle et du même fabricant) que le disque défectueux. Sinon, la table des matières de volume d'origine risque de ne pas convenir au nouveau disque.

Si vous n'avez pas enregistré ces informations de VTOC mais que vous avez créé des copies miroirs des tranches pour chaque disque individuel (mêmes informations VTOC des deux côtés du miroir, par exemple), vous pouvez copier les données VTOC sur le disque de remplacement à partir de l'autre disque sous-miroir. Pour que cette procédure fonctionne, il faut que le disque de remplacement soit en mode de maintenance ou qu'il appartienne au même hôte que le disque défectueux, ou encore que Sun Cluster soit arrêté. Cette procédure est illustrée dans l'exemple suivant.

```
#!/bin/sh
cd /dev/rdisk
OTHER_MIRROR_DISK=c2t0d0s7
REPLACEMENT_DISK=c1t0d0s7
prtvtoC $OTHER_MIRROR_DISK | fmthard -s - $REPLACEMENT_DISK
```

Si vous n'avez pas enregistré les données VTOC ni créé de copies miroirs pour chaque disque individuel, vous pouvez examiner la taille des composants indiquée par la commande `metaset(1M)` et décompiler les informations VTOC. Compte tenu

de la complexité des calculs effectués par cette procédure, celle-ci ne doit être exécutée que par un représentant spécialement formé à cette fin.

Enregistrement des informations de configuration de périphérique

Enregistrez les informations `/etc/path_to_inst` et `/etc/name_to_major` sur un support amovible (disquette ou unité de bande de sauvegarde).

Le fichier `path_to_inst(4)` contient les numéros d'unités mineurs pour chacun des disques contenus dans une unité d'expansion de disque multihôtes. Vous devez utiliser ces données lorsqu'un disque d'initialisation d'un noeud Sun Cluster tombe en panne et qu'il faut le remplacer.

- **Solstice DiskSuite** – Dans les configurations n'utilisant pas le pilote d'ID de disque (DID), le fichier `/etc/name_to_major` contient les numéros de périphériques majeurs pour les disques multihôtes. Solstice DiskSuite, par exemple, utilise les numéros majeurs qui ne changent pas d'une installation du système d'exploitation Solaris à une autre. Cela n'est valable que dans le cas des grappes mises à niveau de la version HA 1.3 à la version Sun Cluster 2.2. Pour de plus amples renseignements à ce sujet, reportez-vous à l'annexe de Solstice DiskSuite dans le manuel *Sun Cluster 2.2 Software Installation Guide*.
- **VxVM** – Pour éviter les messages d'erreur "Identificateur de fichier non valide" sur le client lors d'une reprise NFS, assurez-vous que le pilote `vxio` utilise les mêmes numéros majeurs de pseudo-périphériques sur tous les noeuds de la grappe. Vous trouverez ce numéro dans le fichier `/etc/name_to_major` une fois l'installation achevée. Pour de plus amples renseignements, reportez-vous aux chapitres sur Sun Cluster HA for NFS et sur la configuration de VxVM dans le manuel intitulé *Sun Cluster 2.2 Software Installation Guide*.

Désignation et numérotation des instances

Des noms d'instances sont parfois indiqués dans les messages d'erreur. Le nom d'instance désigne des périphériques système comme `ssd20` ou `hme5`.

Pour connaître les liens existant entre un nom d'instance et un nom physique, examinez la sortie `/var/adm/messages` ou `dmesg(1M)` :

```
ssd20 at SUNW,pln0:
ssd20 is /io-unit@f,e0200000/sbi@0,0/SUNW,soc@3,0/SUNW,pln@a0000800,20183777 \
/ssd@4,0
le5 at lebuffer5: SBus3 slot 0 0x60000 SBus level 4 sparc ipl 7
le5 is /io-unit@f,e3200000/sbi@0,0/lebuffer@0,40000/le@0,60000
```

Lorsqu'un nom d'instance est attribué à un périphérique, il y reste lié.

Les numéros d'instances sont codés dans un numéro mineur de périphérique. Pour que les numéros d'instances soient conservés entre les réinitialisations, le système les enregistre dans le fichier `/etc/path_to_inst`. Ce fichier n'est lu qu'au moment de l'initialisation et est mis à jour au moyen des commandes `add_drv(1M)` et `drvconfig(1M)`. Pour de plus amples renseignements à ce sujet, voir la page de manuel `path_to_inst(4)`.

Lorsque vous installez l'environnement d'exploitation Solaris sur un noeud, les numéros d'instances peuvent être modifiés si des éléments matériels ont été ajoutés ou retirés depuis la dernière installation de Solaris. Il faut donc être très prudent lors de l'ajout et du retrait, sur les noeuds Sun Cluster, de périphériques comme les cartes SBus ou FC/OM. Il est important de conserver la même configuration sur les périphériques existants pour éviter toute confusion du système lors de la réinitialisation ou après une réinstallation ou une reconfiguration.

Des problèmes de numéro d'instance peuvent également survenir. Prenons par exemple une configuration Sun Cluster constituée de trois tableaux SPARCstorage[™] avec cartes SBus de canal de fibres optiques (FC/S) installées dans les fentes SBus 1, 2 et 4 sur chacun des noeuds. Les contrôleurs portent les numéros `c1`, `c2` et `c3`. Si l'administrateur du système ajoute un autre tableau SPARCstorage à cette configuration (carte FC/S dans la fente SBus 3), le numéro du contrôleur correspondant sera `c4`. Si Solaris est réinstallé sur un des noeuds, les numéros de contrôleur `c3` et `c4` désigneront des tableaux SPARCstorage différents. L'autre noeud Sun Cluster désignera toujours les tableaux SPARCstorage portant les numéros d'instances originaux. Solstice DiskSuite ne communiquera pas avec les disques connectés aux contrôleurs `c3` et `c4`.

D'autres problèmes peuvent surgir sur le plan de la numérotation des instances associées aux connexions Ethernet. Ainsi, chacun des noeuds de Sun Cluster comporte trois cartes SBus Ethernet, installées dans les fentes 1, 2 et 3 et portant les numéros d'instances `hme1`, `hme2` et `hme3`. Si la carte du milieu (`hme2`) est retirée et que Solaris est réinstallé, la troisième carte SBus, auparavant `hme3`, portera désormais le nom `hme2`.

Réinitialisations de reconfiguration

Au cours de certaines procédures administratives décrites dans ce manuel, vous devez effectuer une réinitialisation de reconfiguration en exécutant la commande `OpenBoot™ PROM boot -r` ou en créant le fichier `/reconfigure` sur le noeud, puis en réinitialisant le système.

Remarque - Il n'est pas nécessaire d'effectuer cette réinitialisation pour ajouter des disques à une unité d'expansion de disque multihôtes existante.

N'effectuez pas de réinitialisation de reconfiguration de Solaris si un ou plusieurs éléments matériels (en particulier une unité d'expansion de disque multihôtes ou un disque) est hors tension ou défectueux. Dans de tels cas, la réinitialisation entraîne la suppression des inodes dans l'entrée `/devices` et des liens symboliques dans les entrées `/dev/dsk` et `/dev/rdsk` associées aux périphériques de disque. Ces disques deviennent alors inaccessibles pour Solaris, et ce jusqu'à une reconfiguration ultérieure. Lors de la réinitialisation de reconfiguration suivante toutefois, il peut arriver que les numéros mineurs initiaux du contrôleur ne soient pas restaurés et que le gestionnaire de volumes rejette les disques. Après restauration de la numérotation d'origine, le gestionnaire de volumes peut accéder aux objets associés à ces numéros.

Si tous les éléments matériels sont fonctionnels, vous pouvez effectuer une réinitialisation de reconfiguration sans inquiétude lors de l'ajout d'un contrôleur de disque à un noeud. Vous devez ajouter ces contrôleurs de façon symétrique aux deux noeuds (un déséquilibre temporaire est cependant accepté pendant la mise à niveau des noeuds). De même, si tous les éléments matériels sont fonctionnels, vous pouvez effectuer en toute sécurité une réinitialisation de reconfiguration pour supprimer certains de ces éléments.

Remarque - Sous Sun StorEdge A3000, dans le cas d'une panne touchant un seul contrôleur, vous devriez remplacer le contrôleur défectueux le plus rapidement possible. Les autres tâches d'administration exigeant normalement une initialisation à l'aide de la commande `boot --r` (ajout d'un nouveau périphérique SCSI, par exemple) ne doivent être exécutées qu'après le remplacement et la remise en ligne du contrôleur défectueux, lorsque tous les numéros d'unités logiques ont été remis à l'état en vigueur avant la défaillance. Pour de plus amples renseignements, reportez-vous à la documentation de Sun StorEdge A3000.

Connexion sur le serveur en qualité de root (superutilisateur)

Pour ouvrir une session sur les noeuds Sun Cluster en tant que root (superutilisateur) par l'intermédiaire d'un terminal autre que la console, vous devez modifier le fichier `/etc/default/login` et mettre en commentaire la ligne suivante :

```
CONSOLE=/dev/console
```

La ligne ci-dessus permet d'ouvrir une session avec privilèges de root (superutilisateur) au moyen des programmes `rlogin(1)` et `telnet(1)`, entre autres.

Outils d'administration système de Sun Cluster

Le présent chapitre couvre les sujets suivants :

- “Utilitaires de surveillance”, page 29
- “Système d'aide en ligne”, page 33
- “Panneau de commandes de grappe”, page 38
- “Console de grappe”, page 41
- “Utilisation de Sun Cluster Manager”, page 44

Trois interfaces utilisateur graphiques (IUG) facilitent l'administration du logiciel Sun Cluster :

Panneau de commandes de grappe – Démarre la console de grappe et d'autres outils d'administration du système.

Console de grappe – Permet l'exécution simultanée de commandes sur différents noeuds de la grappe et l'administration simplifiée de cette dernière.

Sun Cluster Manager – Surveille l'état courant de tous les noeuds de la grappe par l'intermédiaire d'un navigateur HotJava.

L'aide en ligne permet d'obtenir toutes les informations relatives à ces IUG. Des utilitaires vous permettent également de surveiller le logiciel Sun Cluster.

Utilitaires de surveillance

Vous pouvez recourir à l'utilitaire `hastat(1M)` de Sun Cluster ainsi qu'aux fichiers `/var/adm/messages` pour surveiller les configurations de Sun Cluster ou encore à

l'interface utilisateur graphique de Sun Cluster Manager, qui affiche l'état des principaux composants et sous-composants des grappes. Pour de plus amples renseignements sur Sun Cluster Manager, reportez-vous à la section "Utilisation de Sun Cluster Manager", page 44. Sun Cluster fournit également un agent SNMP (Simple Network Management Protocol) qui peut surveiller simultanément jusqu'à 32 grappes. Voir Annexe C.

Si vous utilisez Solstice DiskSuite, vous pouvez également employer les utilitaires `metastat(1M)`, `metadb(1M)`, `metatool(1M)`, `medstat(1M)` et `mdlogd(1M)` pour surveiller l'état des ensembles de disques de votre système. Le démon de journalisation SNMP de Solstice DiskSuite, `mdlogd(1M)`, génère une trappe SNMP générique lorsque Solstice DiskSuite consigne un message dans le fichier `syslog`. Vous pouvez configurer `mdlogd(1M)` pour qu'il n'envoie une trappe qu'à la consignation de certains messages ; pour ce faire, vous spécifiez une expression régulière dans le fichier de configuration `mdlogd.cf(4)`. La trappe est envoyée à l'hôte d'administration spécifié dans le fichier de configuration. Cet hôte d'administration doit exécuter une application de gestion de réseau comme Solstice SunNet Manager . `mdlogd(1M)` vous évite d'exécuter la commande `metastat(1M)` à intervalles réguliers ou d'analyser les données de sortie de `syslog` pour y retrouver des erreurs ou des avertissements provenant de Solstice DiskSuite. Pour de plus amples renseignements, consultez la page de manuel `mdlogd(1M)`.

Si vous exécutez VxVM, vous pouvez employer les utilitaires `vxprint`, `vxstat`, `vxtrace`, `vxnotify` et `vxva`. Pour obtenir des informations au sujet de ces utilitaires, reportez-vous à la documentation sur le logiciel de gestion des volumes.

Remarque - Pour obtenir des informations sur le dépannage et la réparation des composants défectueux, consultez la documentation fournie.

Surveillance de la configuration avec `hastat(1M)`

Le programme `hastat(1M)` affiche l'état courant de la configuration, c'est-à-dire les données sur les hôtes, les hôtes logiques, les réseaux privés et publics, les services de données, les disques locaux et les ensembles de disques, ainsi que les messages d'erreur les plus récents. Le programme `hastat(1M)` extrait les messages d'erreur liés à Sun Cluster dans le fichier `/var/adm/messages` et émet les quelques derniers messages de chaque hôte si l'option `-m` est spécifiée. La liste récente des messages d'erreur étant un extrait filtré des messages de consignation, le contexte de certains messages peut être perdu. Consultez le fichier `/var/adm/messages` pour obtenir la liste entière des messages. Les pages suivantes présentent un exemple des données de sortie obtenues avec la commande `hastat(1M)` :

```
# hastat -m 10
ETAT ET CONFIGURATION HAUTE DISPONIBILITE
```

(suite)

(suite)

```
-----  
LISTE DE LA GRAPPE  
  hôte-physique1 est un membre de la grappe  
  hôte-physique2 est un membre de la grappe  
ETAT DE LA : Stable  
TEMPS DE FONCTIONNEMENT DES NOEUDS DE LA GRAPPE  
  temps de fonctionnement de hôte-physique1 : StableEtat de la configuration sur hôte-  
physique2 :  
StableTEMPS DE FONCTIONNEMENT DES NOEUDS DE LA GRAPPE  
temps de fonctionnement de hôte-physique1 : 12:47pm  
en fonction depuis 12 jour(s), 21:11, 1 utilisateur, charge moyenne : 0,21, 0,15, 0,14  
temps de  
fonctionnement de hôte-  
physique2 : 12:46pm en fonction depuis 12 jour(s), 3:15, 3 utilisateurs, charge  
moyenne : 0.40, 0.20, 0.16
```

```
HÔTES LOGIQUES MAÎTRISES PAR LES MEMBRES DE LA GRAPPE  
Hôtes logiques maîtrisés sur hôte-physique1 :  
  hôte-hd1  
Hôtes de l'hôte de consignation pour lesquels hôte-physique1 sert de noeud de relève :  
  hôte-hd2  
Hôtes logiques maîtrisés sur hôte-physique2 :  
  hôte-hd2  
Hôtes de l'hôte de consignation pour lesquels hôte-physique2 sert de noeud de relève :  
  hôte-hd1  
HÔTES LOGIQUES AVEC ETAT MAINTENANCE  
  Aucun  
ETAT DES RESEAUX PRIVES DE LA GRAPPE  
  Etat des interconnexions sur hôte-physique1 :  
    interconnexion0 : sélectionnée  
    interconnexion1 : active  
  Etat des réseaux privés sur hôte-physique1 :  
    Vers hôte-physique1 - ACTIF  
    Vers hôte-physique2 - ACTIF  
  Etat des interconnexions sur hôte-physique2 :  
    interconnexion0 : sélectionnée  
    interconnexion1 : active  
  Etat des réseaux privés sur hôte-physique2 :  
    Vers hôte-physique1 - ACTIF  
    Vers hôte-physique2 - ACTIF  
ETAT DES RESEAUX PUBLICS DE LA GRAPPE  
Etat du réseau public sur hôte-physique1 :  
grp_releve adaptateur_réel état temps_reprise adaptateur_actif  
nafo0 le0 OK JAMAIS le0  
Etat du réseau public sur hôte-physique2 :  
bkggrp r_adp status fo_time live_adp  
nafo0 le0 OK JAMAIS le0
```

```

ETAT DES SERVICES S'EXECUTANT SUR LES HÔTES LOGIQUES DANS LA GRAPPE
Etat des services de données inscrits
q :                               Inactif
p :                               Inactif
nfs :                             Actif
oracle :                           Actif
dns :                              Actif
nshttp :                           Inactif
nslldap :                          Actif
Etat des services de données exécutés sur hôte-physique1
Service de données NFS HD :
Dans l'hôte logique hôte-hd1 : OK
Etat des services de données exécutés sur hôte-physique2
Service de données NFS HD :
Sur l'hôte logique hôte-hd2 : OK
Service de données ``oracle`` :
Etat des bases de données sur hôte-physique2 :
FICHIERSC22 - en cours d'exécution ;
Aucune méthode d'état avec le service de données ``dns``
MESSAGES D'ERREUR RECENTS PROVENANT DE LA GRAPPE
Messages d'erreur récents sur hôte-physique1
...
Messages d'erreur récents sur hôte-physique2
...

```

Consultation des fichiers de messages

En plus de transmettre les messages à la console, le logiciel Sun Cluster les écrit dans le fichier `/var/adm/messages`. Voici un exemple de transmission de messages après la défaillance d'un disque.

```

... 1 juin 16:15:26 hôte1 unix :
AVERTISSEMENT : /io-unit@f,e1200000/sbi@0.0/SUNW,pln@a0000000,741022/ssd@3,4(ssd49):
1 juin 16:15:26 hôte1 unix : Erreur de la commande 'write(I)' Err 1 juin 16:15:27 hôte1
unix : ou Niveau :
Fatal 1 juin 16:15:27 hôte1 unix : Bloc 144004 demandé, Bloc d'erreur : 715559 1 juin
16:15:27 hôte1 unix :
Clé capteur : Erreur de support 1 juin 16:15:27 hôte1 unix : Fournisseur
'CONNER' : 1 juin 16:15:27 hôte1 unix :
ASC=0x10(erreur d'ID CRC ou ECC),ASCQ=0x0,FRU=0x15 ...

```

Remarque - Les messages d'erreur de Solaris et de Sun Cluster étant stockés dans le fichier `/var/adm/messages`, le répertoire `/var` peut se saturer. Reportez-vous à la section "Maintenance du système de fichiers `/var`", page 97 pour corriger ce problème.

Utilitaires de service de données haute disponibilité

Sun Cluster comporte des utilitaires permettant de configurer et de gérer la prestation des services de données à haute disponibilité. Les utilitaires sont décrits dans les pages manuel correspondantes. En voici la liste :

- `cconsole(1)` - Démarre l'IUG de la console de grappe.
- `ccp(1)` -- Démarre l'IUG du panneau de commandes de grappe.
- `ctelnet(1)` - Ouvre une session telnet.
- `crlogin(1)` - Ouvre une session rlogin.
- `chosts(1)` - Développe le nom de grappe en une liste des hôtes appartenant à la grappe.
- `cports(1)` - Développe un nom d'hôte sous forme de triplet *hôte, noeud, port*. Utilisé par `cconsole(1)` pour identifier les consoles de port série des hôtes nommés au moyen du serveur de terminal retourné dans les triplets.
- `scconf(1M)` - Crée ou modifie les informations de configuration.

Système d'aide en ligne

Chaque outil d'administration de Sun Cluster fournit de l'aide en ligne détaillée. Pour accéder à cette aide, démarrez un des outils d'administration à partir du poste de travail administratif et sélectionnez l'option Aide de la barre de menus.

Vous pouvez également double-cliquer sur l'icône d'aide du panneau de commandes de grappe.

Les rubriques d'aide traitent en détail des outils d'administration et de certaines tâches administratives. Le Chapitre 4 contient d'autres directives détaillées sur l'exécution de tâches particulières.

Figure 2-1 illustre un exemple de fenêtre d'aide du panneau de commandes de grappe. Le texte affiché porte sur une rubrique particulière. La première fois que vous utilisez l'aide d'un utilitaire, la rubrique de démarrage, ou principale, s'affiche.

Par la suite, la fenêtre d'aide affiche la dernière rubrique consultée. Les rubriques connexes apparaissent sous la forme de liens hypertexte soulignés et colorés.

Cliquez une fois sur un lien hypertexte pour afficher le texte de la rubrique correspondante. Le système d'aide en ligne comporte également un historique automatique des rubriques consultées. Pour afficher cet historique, choisissez l'option Historique des rubriques du menu Affichage.

La fenêtre d'aide est dotée d'une zone de texte défilante, d'une barre de menus et de plusieurs boutons. Les sections suivantes traitent de ces éléments.

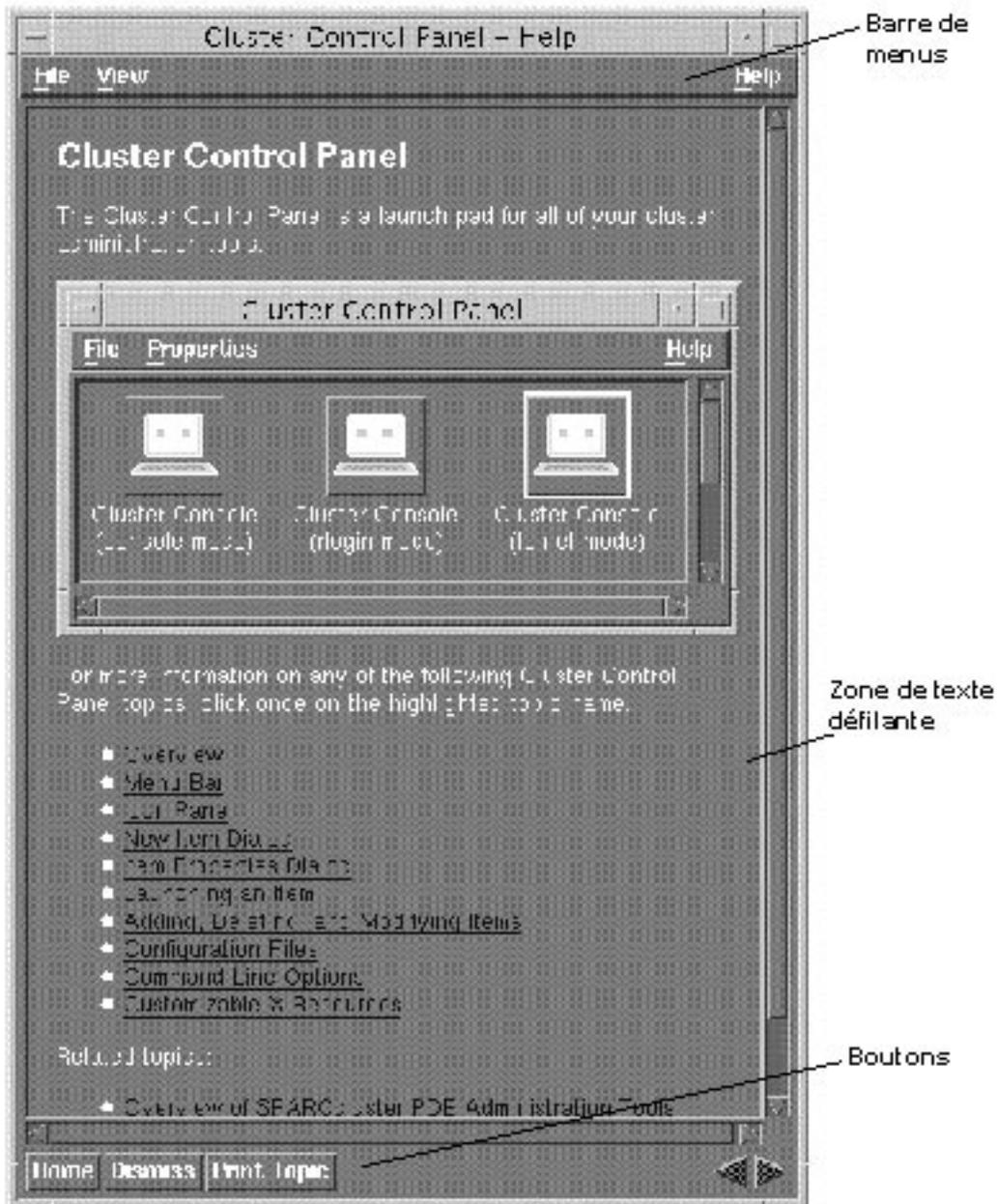


Figure 2-1 Exemple de page de démarrage de la fenêtre d'aide du panneau de commandes de grappe

Pour accéder aux différents menus déroulants de l'aide en ligne, procédez comme suit :

- Cliquez sur le nom du menu.

- Appuyez sur la touche mnémonique, c'est-à-dire la lettre soulignée d'un menu ou d'une option de menu (dans ce cas-ci, le menu déroulant doit être visible).
- Utilisez un accélérateur, ou séquence de touches, situé à droite de l'élément du menu.

Vous pouvez personnaliser les touches mnémoniques et les accélérateurs. Consultez l'aide en ligne pour de plus amples renseignements à ce sujet.

Les tableaux de cette section présentent les éléments des différents menus, décrivent les fonctions du menu traité et en indiquent les accélérateurs respectifs (combinaisons de touches).

Eléments de la barre de menus de la fenêtre d'aide

La fenêtre d'aide comprend les options des menus Fichier, Affichage et Aide. Cliquez sur un menu pour en afficher le contenu.

Menu Fichier

Le menu Fichier comporte les éléments suivants :

TABLEAU 2-1 Eléments du menu Fichier

Élément	Fonction	Accélérateur
Imprimer une rubrique	Imprime le contenu de la rubrique sélectionnée dans la zone de texte défilante de la fenêtre d'aide.	Alt + R
Fermer	Ferme la fenêtre Aide.	Alt + D

Menu Affichage

Le menu Affichage contient les éléments suivants :

TABLEAU 2-2 Eléments du menu Affichage

Elément	Fonction	Accélérateur
Rubrique précédente	Affiche la rubrique d'aide précédente (s'il y a lieu).	Alt + P
Rubrique suivante	Affiche la rubrique d'aide suivante (s'il y a lieu).	Alt + N
Rubrique de démarrage	Affiche la rubrique de démarrage (principale).	Alt + O
Historique des rubriques...	Affiche la boîte de dialogue Historique des rubriques permettant de sélectionner facilement une rubrique d'aide déjà consultée. Les rubriques de la liste déroulante s'affichent selon l'ordre de leur consultation dans le chemin d'accès courant. Le titre de la rubrique sélectionnée est mis en évidence.	Alt + I

Pour afficher la boîte de dialogue, sélectionnez Affichage, puis → Historique des rubriques... (Figure 2-2).

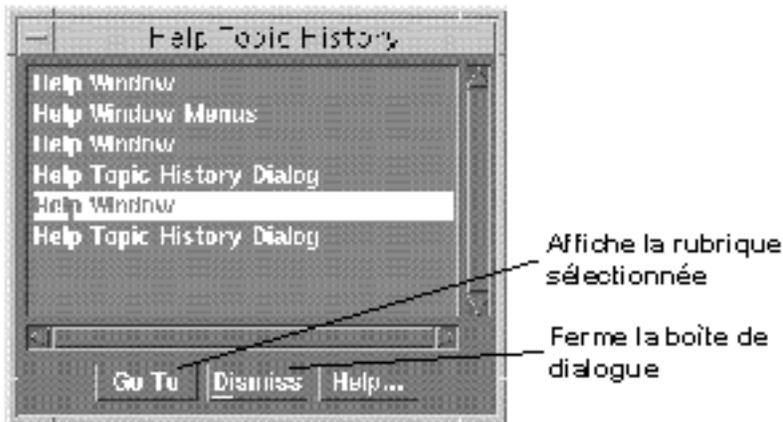


Figure 2-2 Historique des rubriques d'aide de la fenêtre d'aide

Menu Aide

Le menu Aide contient les éléments suivants :

TABLEAU 2-3 Eléments du menu Aide

Elément	Fonction
Aide sur l'aide...	Décrit la fenêtre d'aide et explique comment l'utiliser.
A propos de...	Affiche la fenêtre A propos de, qui contient des informations sur l'application, par exemple son numéro de version.

Boutons de la fenêtre Aide

Le tableau suivant dresse la liste des boutons de la fenêtre Aide et décrit leurs fonctions.

TABLEAU 2-4 Eléments du menu Aide

Bouton	Fonction
Accueil	Affiche la rubrique de démarrage de l'application.
Fermer	Ferme la fenêtre Aide.
Imprimer une rubrique	Imprime la rubrique sélectionnée sur l'imprimante par défaut.
	Sélectionne la rubrique d'aide précédente de la liste. Cliquez de façon répétée sur la flèche gauche pour afficher successivement les rubriques précédentes. L'historique de la fenêtre Aide se "souvient" automatiquement des rubriques consultées.
	Sélectionne la rubrique d'aide suivante de la liste.

Panneau de commandes de grappe

Le panneau de commandes de grappe (PCG) est une IUG vous permettant de démarrer la console de grappe et tout autre outil d'administration du système. Le PCG comporte des icônes représentant ces outils.

▼ Comment démarrer le PCG

Après avoir installé le logiciel client Sun Cluster sur le poste de travail administratif, utilisez la procédure ci-dessous pour exécuter une application à partir du PCG.

1. **En tant que root (superutilisateur), ajoutez le répertoire des outils Sun Cluster `/opt/SUNWcluster/bin` au chemin d'accès du poste de travail administratif.**

Remarque - Dans le cas des plates-formes E10000, vous devez d'abord ouvrir une session dans le processeur de services système (PSS) et vous connecter avec la commande `netcon`. Une fois connecté, appuyez sur les touches Maj-@ pour déverrouiller la console et obtenir l'accès en écriture. Passez ensuite à l'étape 2.

2. **Affichez le PCG à partir d'une fenêtre Shell de votre poste de travail.**

Entrez le nom de la grappe à surveiller :

```
# ccp nom_grappe
```

Remarque - Si les outils de Sun Cluster ne sont pas installés à l'emplacement par défaut `/opt/SUNWcluster`, vous devez modifier la variable d'environnement `$CLUSTER_HOME` en lui affectant le nom de l'emplacement où ils se trouvent.

Eléments du PCG

Le PCG (illustré dans la figure ci-dessous) comporte une barre de menus et un volet d'icônes regroupant tous les outils activés. Vous pouvez ajouter, supprimer ou modifier les outils du panneau de commandes à l'aide de la barre de menus.

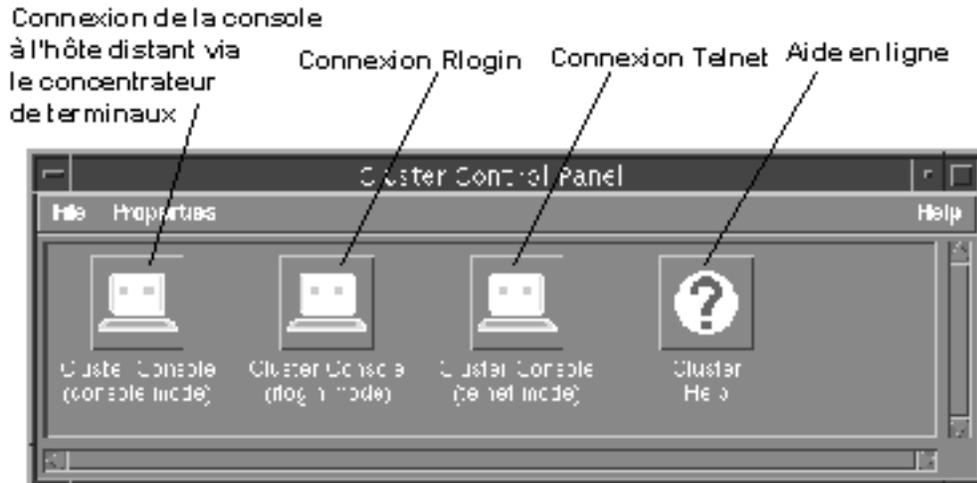


Figure 2-3 Exemple de Panneau de commandes de grappes

Les menus Fichier et Propriétés permettent :

- D'ajouter un nouvel élément.
- De supprimer un élément.
- De modifier un élément.

L'aide en ligne contient des informations détaillées sur le PCG.

La "Console de grappe", page 41 porte sur les programmes associés à ces outils et sur leur utilisation. L'"Utilisation de Sun Cluster Manager", page 44 porte sur l'utilisation des configurations de grappes à l'aide du navigateur HotJava.

Emplacements des fichiers de configuration du PCG

Le PCG stocke les propriétés et toutes informations connexes dans des fichiers de configuration regroupés dans un même répertoire. Par défaut, ce répertoire est /opt/SUNWcluster/etc/ccp.

Remarque - Vous devez être un root (superutilisateur) pour avoir le droit d'écrire dans ce répertoire par défaut. Seul un root (superutilisateur) peut ajouter, supprimer ou modifier les éléments du PCG contenus dans ce répertoire de configuration.

Cependant, vous pouvez créer votre propre répertoire de configuration et définir son emplacement à l'aide de la variable d'environnement `$CCP_CONFIG_DIR`. La variable `$CCP_CONFIG_DIR` spécifie le répertoire de configuration où se trouvent les fichiers de configuration contenant les propriétés des éléments. En l'absence de chemin d'accès défini, l'emplacement standard `/opt/SUNWcluster/etc/ccp` est utilisé. Pour définir votre propre répertoire de configuration, créez un nouveau répertoire et attribuez à la variable d'environnement `$CCP_CONFIG_DIR` le chemin d'accès complet du nouveau répertoire.

Il n'est pas nécessaire de modifier manuellement ces fichiers, le PCG s'en charge lui-même lorsque vous créez, modifiez ou supprimez un élément.

Console de grappe

L'IUG de la console de grappe (CG) permet l'exécution de commandes sur plusieurs noeuds à la fois, ce qui facilite l'administration des grappes. La console de grappe affiche une fenêtre de terminal pour chaque noeud de grappe, en plus d'une petite fenêtre commune servant à contrôler simultanément l'ensemble des fenêtres.

Différents types de sessions à distance vous permettent d'établir la connexion à la console de l'hôte ou de vous connecter à distance à l'aide de `rlogin` ou de `telnet`. La ligne de commande spécifie les hôtes, et la boîte de dialogue Sélectionner les hôtes permet de les ajouter ou de les supprimer après le démarrage du programme. Seule la ligne de commande permet de spécifier le type de session utilisée. Il est impossible d'opter pour un autre type après l'ouverture d'une session.

La fenêtre commune permet d'émettre des commandes vers plusieurs hôtes à la fois, alors que la fenêtre de terminal n'autorise que les commandes vers un seul hôte. Les fenêtres de terminal utilisent l'émulation de terminal VT100.

Vous pouvez également désactiver tous les hôtes dans le menu Hôtes, à l'exception de celui auquel vous souhaitez accéder, puis émettre les commandes à partir du champ de texte de la fenêtre commune.

▼ Comment démarrer la console de grappe

Démarrez la console de grappe avec le PCG (voir "Panneau de commandes de grappe", page 38) ou avec la ligne de commande dans une fenêtre Shell. Si un paramètre facultatif est employé, une fenêtre de terminal est créée pour chaque hôte de la grappe ou pour chaque hôte spécifié.

1. La commande `cconsole` permet d'établir l'accès à la console distante :

```
% cconsole [nom_groupe | nom_hôte...]
```

2. La commande `ctelnet` permet d'établir une connexion `telnet(1)` à partir de la console :

```
% ctelnet [nom_groupe | nom_hôte...]
```

3. La commande `crlogin`, suivie de votre nom d'utilisateur, permet d'établir une connexion `rlogin(1)` à partir de la console :

```
% crlogin -l nom d'utilisateur [  
nom_groupe | nom_hôte...]
```

Les trois commandes précédentes acceptent également les arguments standard de ligne de commande X/Motif. La fenêtre Console apparaît après le démarrage de la console de grappe.

Consultez l'aide en ligne pour obtenir des informations détaillées sur la console de grappe.

Barre de menus de la fenêtre commune

La fenêtre commune (voir la figure ci-dessous) est la principale fenêtre servant à commander tous les noeuds. Cette fenêtre apparaît toujours lorsque vous démarrez la console de grappe.



Figure 2-4 Barre de menus de la fenêtre commune de la console de grappe

Cette fenêtre comporte une barre avec trois menus et un champ de texte servant à l'entrée de commandes. Le menu Hôtes permet d'ouvrir la boîte de dialogue Sélectionner pour :

- Ajouter un hôte

- Ajouter tous les hôtes dans une grappe
- Retirer un hôte

Le menu Options permet d'activer ou de désactiver la liaison entre la fenêtre commune et les fenêtres de terminal.

Fichiers de configuration utilisés par la console de grappe

La console de grappe utilise deux fichiers de configuration, à savoir : `clusters` et `serialports`. Il peut s'agir de fichiers du répertoire `/etc` ou de bases de données NIS/NIS+. L'environnement NIS+ permet à différents postes de travail administratifs d'utiliser en même temps la console de grappe. Consultez le guide d'administration système NIS/NIS+ pour obtenir des informations détaillées au sujet de ces environnements.

A propos du fichier `clusters`

Le fichier `clusters` établit la correspondance entre le nom d'une grappe et la liste des noms d'hôtes contenant cette grappe. A chaque ligne du fichier correspond une grappe, comme dans l'exemple suivant :

```
planètes      Mercure Vénus Terre Mars
vin           zinfandel merlot chardonnay riesling
```

Les trois types de sessions de la console de grappe (`cconsole`, `ctelnet` et `crlogin`) utilisent le fichier `clusters` pour établir la correspondance entre les noms de grappes et les noms d'hôtes dans la ligne de commande et la boîte de dialogue Sélectionner les hôtes. Pour de plus amples renseignements, voir la "Modification du fichier `clusters`", page 72.

A propos du fichier `serialports`

Le fichier `serialports` fait correspondre un nom d'hôte au concentrateur de terminaux et à son port série auquel se connecte l'hôte. Chaque ligne de ce fichier de base de données spécifie un port série de l'hôte.

Voici des exemples d'entrée du fichier de base de données `serialports` pour Sun Enterprise 10000 :

Mercure	<code>nom_processeur_service_système</code>	23
Vénus	<code>nom_processeur_service_système</code>	23
Terre	<code>nom_processeur_service_système</code>	23
Mars	<code>nom_processeur_service_système</code>	23

Voici des exemples d'entrée du fichier de base de données `serialports` pour tous les autres noeuds :

Mercure	<code>planètes-tc</code>	5002
Vénus	<code>planètes-tc</code>	5003
Terre	<code>planètes-tc</code>	5004
mars	<code>planètes-tc</code>	5005

La version `cconsole` du programme utilise exclusivement le fichier `serialports` pour déterminer le concentrateur de terminaux et le port auxquels se connectent les hôtes et les grappes, ces derniers ayant été spécifiés avec la ligne de commande ou la boîte de dialogue Sélectionner les hôtes.

Dans l'exemple précédent, le noeud `Mercure` se connecte au port 2 de `planètes-tc` Port 2, alors que le noeud `Vénus` se connecte au port 3 de `planètes-tc`. Le port 1 sert uniquement à l'administration du concentrateur de terminaux.

Pour de plus amples renseignements, voir la "Modification du fichier `serialports`", page 74.

Utilisation de Sun Cluster Manager

Sun Cluster Manager (SCM) est l'outil de gestion de grappes de Sun Cluster 2.2. Le Sun Cluster Manager (SCM) sert d'interface unique à un grand nombre de fonctions de surveillance par ligne de commande de Sun Cluster. L'IUG de SCM transmet des informations sur les éléments suivants :

- Alarmes SCM
- Messages `syslog` sur chaque noeud de grappe

- Ressources de grappe, y compris : services de données parallèles, hôtes logiques, services HD inscrits, noeuds de grappe et gestionnaires de volumes.

Le gestionnaire SCM se compose de deux éléments principaux : un logiciel serveur SCM et une interface utilisateur graphique (IUG). Le serveur SCM s'exécute sur chaque noeud de la grappe. L'IUG du SCM peut s'exécuter comme une application ou un applet. Si elle est exécutée comme un applet, SCM s'exécute dans un navigateur compatible avec la trousse d'outils Java Development (JDK™), telle HotJava™ ou Netscape. N'importe quel ordinateur peut utiliser le navigateur, y compris les noeuds de grappe. Pour obtenir des informations à jour sur les versions JDK prises en charge, consultez votre fournisseur de services ou la dernière version de *Sun Cluster 2.2 Release Notes*.

Les informations des sections suivantes vous indiquent comment configurer et exécuter SCM.

Exécution de SCM en tant qu'application

Pour exécuter SCM en tant qu'application, procédez comme suit.

▼ Comment exécuter SCM en tant qu'application

1. Sur le poste de travail administratif, installez le SCM (SUNWscmgr) à partir du CD du produit Sun Cluster 2.2.

2. Installez la dernière version du fichier correctif SCM sur tous les noeuds de grappe et sur le poste de travail administratif.

Pour obtenir des informations à jour sur les fichiers correctifs et leurs numéros, consultez le *Sun Cluster 2.2 Release Notes*, adressez-vous à votre fournisseur de service ou visitez le site Web des fichiers correctifs de Sun, <http://sunsolve.sun.com>.

3. Exécutez l'application SCM.

Pour démarrer l'application, exécutez la commande suivante depuis n'importe quel noeud de grappe, *noeud_grappe* devant être un membre de la grappe.

```
# /opt/SUNWcluster/bin/scmgr noeud_grappe
```

Une fois l'application SCM démarrée, vous pouvez accéder à l'aide en ligne pour plus d'informations sur la navigation dans les menus, les tâches et les références.

Pour afficher la fenêtre Aide depuis l'application SCM, sélectionnez Sommaire de l'aide dans le menu Aide. Vous pouvez également cliquer sur l'icône d'aide dans la barre d'outils au-dessus de l'icône du dossier.

Reportez-vous à la page `scmgr(1M)` pour plus d'informations sur l'exécution de SCM.

Exécution de SCM en tant qu'applet

Pour exécuter SCM en tant qu'applet, vous devez effectuer les tâches suivantes, qui sont décrites en détail dans les procédures ci-après.

1. Sur tous les noeuds de la grappe, installez la dernière version du fichier correctif SCM depuis SunSolve. Pour obtenir des informations à jour sur les fichiers correctifs et leurs numéros, consultez le *Sun Cluster 2.2 Release Notes*, adressez-vous à votre fournisseur de service ou visitez le site Web des fichiers correctifs de Sun, <http://sunsolve.sun.com>.
2. Installez un navigateur sur le poste de travail administratif. Les navigateurs pris en charge par Sun Cluster 2.2 4/00 sont HotJava (version 1.1.4 ou supérieure) et Netscape (version 4.5 ou supérieure).
3. Si vous avez installé le navigateur HotJava, installez la Trousse d'outils JDK sur le poste de travail administratif.
4. Installez et configurez un serveur Web sur tous les noeuds de grappe.
5. Démarrez l'applet SCM en entrant l'URL voulue dans le navigateur. Assurez-vous que l'hôte indiqué dans l'URL est bien membre de la grappe.

Remarque - Si vous optez pour le navigateur HotJava livré avec le système d'exploitation Solaris 2.6 ou 2.7, vous pouvez éprouver certaines difficultés avec les menus. Par exemple, l'option du menu sélectionnée peut parfois demeurer visible dans le navigateur après la sélection. Consultez les notes de version de *Sun Cluster 2.2 Release Notes* pour plus d'informations sur les problèmes liés à SCM. Solaris 8 ne prend pas en charge les navigateurs HotJava. Pour exécuter SCM avec Solaris 8, vous devez par conséquent utiliser un autre navigateur, par exemple Netscape. Voir "Comment exécuter l'applet SCM dans un navigateur Netscape à partir d'un noeud de grappe", page 50.

Remarque - Si vous utilisez le navigateur HotJava avec SCM, vous devez disposer d'au moins 40 méga-octets de zone de swap. Si vous vous apercevez que l'espace dans la zone de swap diminue, le fait de redémarrer le navigateur HotJava peut aider.

Pour exécuter SCM, vous devez posséder les versions adéquates de HotJava et avoir installé les modules JDK (`SUNWjvrt` et `SUNWjvjit`) sur le système sur lequel se trouve le navigateur HotJava. Vérifiez les numéros de vos versions par rapport à ceux

indiqués dans le tableau suivant. Consultez également la dernière version de *Sun Cluster 2.2 Release Notes* qui peut contenir des mises à jour de ces informations :

TABLEAU 2-5 Configuration requise pour SCM : JDK et HotJava

Environnement d'exploitation	Version de la Trousse d'outils JDK (Java Developer Kit)	Version HotJava	Version Netscape
Solaris 2.6	1.1.6 ou supérieure	1.1.4 ou supérieure	4.5 ou supérieure
Solaris 7	1.1.6 ou supérieure	1.1.4 ou supérieure	4.5 ou supérieure
Solaris 8	1.2 ou supérieure	Non pris en charge	4.5 ou supérieure

Deux options sont possibles. Vous pouvez :

- Exécuter le navigateur HotJava sur un noeud de grappe. Si vous choisissez cette option, vous devrez relancer HotJava sur un noeud différent si le noeud sur lequel il tournait tombe en panne.
- Installer un serveur Web pour l'exécution de SCM sur chaque noeud de la grappe. Si vous choisissez cette option, vous devez entrer l'URL d'un autre noeud dans le navigateur HotJava si la boîte de dialogue *Lost Connection* s'affiche.

Reportez-vous à la procédure correspondant à l'option choisie.

▼ Comment configurer la trousse d'outils JDK

1. Pour déterminer la version Java actuelle, entrez la commande suivante à partir de l'invite de console du serveur de la grappe :

```
# java -version
```

2. Si nécessaire, téléchargez une version plus récente de la Trousse d'outils JDK.

Si le système affiche une version de Java inférieure à la 1.1.6, suivez les instructions ci-après pour télécharger la version 1.1.6 de JDK (ou une version supérieure) depuis l'URL suivante :

<http://www.sun.com/solaris/java>

▼ Comment télécharger HotJava

1. Depuis l'ordinateur qui exécute le navigateur HotJava, sélectionnez **A propos de HotJava** dans le menu d'aide.

Si le navigateur affiche une version inférieure à la 1.1.4, ou si vous n'avez pas de navigateur HotJava, suivez les instructions ci-après pour télécharger la version 1.1.4 de JDK (ou une version supérieure) depuis l'URL suivante :

<http://java.sun.com/products/hotjava/index.html>

▼ Comment exécuter l'applet SCM dans un navigateur HotJava à partir d'un noeud de grappe

1. **Exécutez le navigateur HotJava sur un noeud de la grappe.**
Le navigateur HotJava se trouve dans le répertoire `/usr/dt/bin`.
2. **Affichez à distance ce navigateur sur un poste de travail X Windows.**
3. **Définissez les préférences de sécurité des applets dans le navigateur HotJava :**
 - a. Dans le menu **Editer**, choisissez **Préférences**, puis **Sécurité applet**.
 - b. Cliquez sur **Sécurité moyenne** comme paramètre par défaut pour les applets non signés.
4. Lorsque vous êtes prêt à commencer la surveillance de la grappe à l'aide de SCM, entrez l'URL appropriée. Par exemple :

```
file:/opt/SUNWcluster/scmgr/index.html
```

5. Cliquez sur **OK** dans chaque boîte de dialogue vous demandant d'autoriser l'accès à certains fichiers, ports et autres.

Remarque - Le téléchargement et l'exécution de l'applet par HotJava peuvent exiger un certain temps. Aucune information d'état n'apparaîtra pendant ce délai.

L'aide en ligne contient des informations détaillées sur la structure des menus et l'exécution des tâches et fournit également les références à leur sujet.

▼ Comment exécuter l'applet SCM dans un navigateur HotJava à partir du poste de travail administratif

1. **Lancez le navigateur HotJava sur un noeud de la grappe.**
Le navigateur HotJava se trouve dans le répertoire `/usr/dt/bin`.
2. **Configurez et démarrez un navigateur Web sur tous les noeuds de la grappe.**
Pour plus de détails, consultez la rubrique "Comment installer un serveur Web pour SCM", page 51.
3. **Définissez les préférences de sécurité pour l'applet dans le navigateur HotJava :**
 - a. **Dans le menu Editer, choisissez Préférences, puis Sécurité applet.**
 - b. **Cliquez sur Sécurité moyenne comme paramètre par défaut pour les applets non signés.**
4. **Lorsque vous êtes prêt à surveiller la grappe avec SCM, entrez l'URL appropriée.**

```
http://noeud_grappe/link_to_scm/index.html
```

5. **Cliquez sur OK dans chaque boîte de dialogue vous demandant d'autoriser l'accès à certains fichiers, ports et autres, à partir du poste de travail d'affichage distant vers le noeud de grappe où a été démarré le navigateur.**

Remarque - Le téléchargement et l'exécution de l'applet par HotJava peuvent exiger un certain temps. Aucune information d'état n'apparaîtra pendant ce délai.

L'aide en ligne contient des informations détaillées sur la structure des menus et l'exécution des tâches et fournit également les références à leur sujet.

▼ Comment exécuter l'applet SCM dans un navigateur Netscape à partir d'un noeud de grappe

1. **Installez Netscape sur les noeuds de la grappe.**
2. **Installez SCM et le fichier correctif SCM requis sur les noeuds de la grappe.**

Pour installer SCM, lancez la commande `scinstall(1M)`. `scinstall(1M)` installe le module SCM (`SUNWscmgr`) faisant partie du module serveur. Pour vous procurer le fichier correctif SCM, adressez-vous à votre ingénieur technico-commercial ou visitez le site Web SunSolve :

<http://sunsolve.sun.com/>

3. **Ajoutez les lignes suivantes au fichier `preferences.js`, le cas échéant.**

Ce fichier se trouve dans le répertoire `$HOME/.netscape`. Si les préférences ne figurent pas déjà dans le fichier, ajoutez les lignes suivantes :

```
user_pref('security.lower_java_network_security_by_trusting_proxies', true);
user_pref('signed.applets.codebase_principal_support', true);
```

4. **Sur un noeud de la grappe, définissez votre variable d'environnement `DISPLAY` de façon à afficher à distance le navigateur Netscape sur votre station de travail X Windows, puis exécutez le navigateur Netscape sur ce noeud.**
5. **Lorsque vous êtes prêt à surveiller la grappe avec SCM, entrez l'URL appropriée.**

```
file:/opt/SUNWcluster/scmgr/index.html
```

6. **Cliquez sur Attribuer dans chaque boîte de dialogue vous demandant d'autoriser l'accès à certains fichiers, ports et autres, du noeud de grappe où a été démarré le navigateur, à partir du poste de travail d'affichage distant.**
L'aide en ligne contient des informations détaillées sur la structure des menus et l'exécution des tâches et fournit également les références à leur sujet.

▼ Comment installer un serveur Web pour SCM

Si vous le souhaitez, vous pouvez installer un serveur Web sur les noeuds de grappe pour exécution avec SCM.

Remarque - Si vous exécutez le service Sun Cluster HA for Netscape HTTPet un serveur HTTP sur SCM, vous devez configurer chaque serveur HTTP avec un port d'écoute différent. Cette procédure s'avère nécessaire pour éliminer les risques de conflit au niveau des ports qu'ils utilisent.

1. **Installez un serveur Web sur tous les noeuds de la grappe.**
2. **Suivez la procédure de configuration du serveur Web pour vous assurer que le fichier `index.html` de SCM est accessible aux clients.**

L'applet client de SCM se trouve dans le fichier `index.html` dans le répertoire `/opt/SUNWcluster/scmgr`. Vous pouvez par exemple accéder au `document_root` de votre serveur HTTP et y créer un lien vers le répertoire `/opt/SUNWcluster/scmgr`.
3. **Démarrez votre navigateur HotJava à partir de votre poste de travail.**
4. **Définissez les préférences de sécurité pour l'applet dans le navigateur HotJava :**
 - a. **Dans le menu Editer, choisissez Préférences, puis Sécurité applet.**
 - b. **Cliquez sur Sécurité moyenne comme paramètre par défaut pour les applets non signés.**
5. **Lorsque vous êtes prêt à commencer la surveillance de la grappe à l'aide de SCM, entrez l'URL appropriée.**

Par exemple, si vous avez créé un lien à partir du répertoire `document_root` de votre serveur Web vers le répertoire `/opt/SUNWcluster/scmgr`, entrez l'URL suivante :

`http://cluster_node/scmgr/index.html`

6. **Cliquez sur OK dans chaque boîte de dialogue vous demandant d'autoriser l'accès à certains fichiers, ports et autres, à partir du noeud de la grappe où a été démarré le navigateur.**

Remarque - Le téléchargement et l'exécution de l'applet par HotJava peuvent exiger un certain temps. Aucune information d'état n'apparaîtra pendant ce délai.

L'aide en ligne contient des informations détaillées sur la structure des menus et l'exécution des tâches et fournit également les références à leur sujet.

Accès à l'aide en ligne de SCM

SCM comporte une aide en ligne qui contient des informations sur la structure des menus et l'exécution des tâches et fournit également des références à leur sujet. Cette aide est disponible quel que soit le mode d'exécution de SCM, en tant qu'application ou applet.

Pour afficher la fenêtre d'aide de SCM, sélectionnez Sommaire de l'aide dans le menu d'aide. Vous pouvez également cliquer sur l'icône d'aide (point d'interrogation) dans la barre d'outils au-dessus du dossier.

Au besoin, vous pouvez utiliser l'aide en ligne dans un autre navigateur en tapant l'URL suivante :

```
file:/opt/SUNWcluster/scmgr/help/locale/en/main.howtotopics.html
```

Par exemple, si vous avez créé un lien à partir du répertoire `document_root` de votre serveur Web vers le répertoire `/opt/SUNWcluster/scmgr`, entrez l'URL suivante :

```
http://clusternode/scmgr/help/locale/en/main.howtotopics.html
```

Lorsque vous avez fini de consulter l'aide en ligne, fermez le navigateur HotJava. Si vous sélectionnez l'aide de nouveau, un nouveau navigateur s'affiche et charge l'aide.

Modification de la configuration de Sun Cluster

Ce chapitre traite des sujets suivants :

- “Ajout et suppression de noeuds de grappe”, page 54
- “Modification du nom d’un noeud de grappe”, page 57
- “Modification des interfaces réseau privées”, page 58
- “Impression de la configuration des noeuds”, page 59
- “Ajout et suppression d’hôtes logiques”, page 60
- “Modification de l’adresse IP de l’hôte logique”, page 65
- “Reconfiguration de grappe forcée ”, page 66
- “Configuration des services de données de Sun Cluster”, page 67
- “Suppression de la configuration des services de données Sun Cluster”, page 68
- “Ajout de services de données Sun Cluster”, page 70
- “Modification du fichier `clusters`”, page 72
- “Modification du fichier `serialports`”, page 74
- “Modification des informations TC/PSS”, page 75
- “Modification du périphérique de quorum”, page 77
- “Configuration des délais de transition de la grappe”, page 80

Ajout et suppression de noeuds de grappe

Vous devez reconfigurer le logiciel Sun Cluster lorsque vous ajoutez ou retirez des noeuds de grappe. A la première installation de la grappe, vous avez spécifié le nombre de noeuds “actifs” et “potentiels” de la grappe à l’aide de la commande `scinstall(1M)`. La procédure présentée dans cette section permet d’ajouter des noeuds “potentiels” et de retirer des noeuds “actifs”.

Pour ajouter des noeuds qui ne sont pas déjà identifiés comme des noeuds potentiels, vous devez arrêter et reconfigurer la grappe entière.

▼ Comment ajouter un noeud de grappe

N’utilisez cette procédure qu’avec les noeuds que la première installation a identifiés comme étant “potentiels”.

1. Exécutez la commande `scinstall(1M)` pour installer Sun Cluster sur le noeud que vous ajoutez.

Utilisez les procédures décrites dans le *Sun Cluster 2.2 Software Installation Guide*, mais tenez compte des considérations suivantes lorsque vous répondez aux invites de la commande `scinstall(1M)` :

- Le nombre total de noeuds actifs doit inclure celui que vous ajoutez.
- Il n’y a pas d’invite relative aux informations partagées de la base de données de configuration de grappe (BCG), puisque la nouvelle grappe doit comporter plus de deux noeuds.
- (VxVM avec périphériques à connexion directe seulement) A l’invite du port de verrouillage de noeud, indiquez le périphérique et le port de verrouillage de noeud désignés.
- (VxVM seulement) Ne sélectionnez pas de périphérique de quorum à l’invite. Sélectionnez plutôt le mode `complex` puis l’option `N`. Vous exécuterez la commande `scconf -q` ultérieurement pour configurer le périphérique de quorum.
- (VxVM seulement) Sélectionnez l’option `Ask` à l’invite de sélection d’un comportement de partitionnement de grappe.

2. (Interface cohérente échelonnable [ICE] seulement) Mettez à jour le fichier modèle `sm_config` pour vérifier les informations sur le nouveau noeud.

Vous pouvez omettre cette étape avec les configurations Ethernet.

Le fichier `sm_config` doit contenir les noeuds spécifiés comme étant “potentiels” à la première installation, en plus de leurs noms d’hôtes mis en commentaire par les caractères `_%`. Supprimez la mise en commentaire d’un noeud pour l’activer. Assurez-vous que les informations de configuration du fichier correspondent à la disposition physique du noeud.

3. (ICE seulement) Exécutez la commande `sm_config`.

4. (VxVM seulement) Configurez le groupe de disques racine.

Pour plus de détails, consultez l’annexe VxVM dans le *Sun Cluster 2.2 Software Installation Guide*.

5. (SDS seulement) Configurez les ensembles de disques de Solstice DiskSuite.

Pour plus de détails, consultez l’annexe Solstice DiskSuite dans le *Sun Cluster 2.2 Software Installation Guide*.

6. Si un périphérique est connecté directement à chaque noeud, configurez l’indicateur de disque à connexion directe sur le nouveau noeud.

Pour configurer cet indicateur correctement dans les fichiers `cdb` de chaque noeud, exécutez la commande suivante sur tous les noeuds de la grappe. Dans cet exemple, la grappe porte le nom `grappe-sc` :

```
# scconf grappe-sc +D
```

7. (VxVM seulement) Sélectionnez un Périphérique de quorum commun.

Si VxVM assure la gestion des volumes et qu’un périphérique à connexion directe est connecté à chaque noeud, exécutez la commande suivante sur tous les noeuds et sélectionnez un périphérique de quorum commun.

```
# scconf grappe-sc -q -D
```

S’il n’y a pas de disque à connexion directe connecté à chaque noeud, exécutez la commande suivante pour chaque paire de noeuds partageant un périphérique de quorum avec le nouveau noeud.

```
# scconf -q
```

8. (VxVM seulement) Configurez le port de verrouillage de noeud sur le nouveau noeud.

Si vous venez d’installer un disque à connexion directe, configurez le port de verrouillage de noeud sur tous les noeuds.

Si la grappe contient déjà un disque à connexion directe, exécutez la commande suivante uniquement sur le nouveau noeud. Dans cet exemple, la grappe et le concentrateur de terminaux portent respectivement les noms `grappe-sc` et `ct-grappe`.

```
# # sconfig grappe-sc -t ct_grappe -l numéro_port
```

9. Arrêtez la grappe.

10. Exécutez la commande `sconfig -A` sur tous les noeuds pour mettre à jour le nombre de noeuds actifs.

Pour de plus amples renseignements, consultez la page de manuel `sconfig(1M)`. Dans le présent exemple, la grappe porte le nom `grappe-sc`, et le nouveau total de noeuds actifs est maintenant égal à trois.

```
# sconfig grappe-sc -A 3
```

11. (VxVM seulement) Retirez la BCG partagée si elle existe, celle-ci n'étant requise qu'avec les grappes à deux noeuds.

Exécutez la commande suivante sur tous les noeuds.

```
# sconfig grappe-sc -S none
```

12. A l'aide du protocole `ftp` (en mode binaire), copiez le fichier `cdb` d'un noeud existant vers le nouveau noeud.

Le chemin d'accès du fichier `cdb` est habituellement `/etc/opt/SUNWclus/conf/nom_grappe.cdb`.

13. Réinitialisez les nouveaux noeuds.

14. Démarrez le noeud.

Exécutez la commande suivante à partir d'un noeud quelconque.

```
# scadmin startcluster phys-hahost grappe-sc
```

Exécutez ensuite la commande suivante sur tous les autres noeuds.

```
# scadmin startnode
```

▼ Comment supprimer un noeud de grappe

La commande `sconf(1M)` vous permet de supprimer des noeuds en décrémentant le nombre de noeuds actifs spécifié lors de l'installation du logiciel Sun Cluster avec la commande `scinstall(1M)`. Cette procédure exige que vous exécutiez la commande `sconf(1M)` sur tous les noeuds de la grappe.

1. **Dans le cas d'une configuration HD, commutez tous les hôtes logiques dont la maîtrise est assurée actuellement par le noeud à supprimer.**

Passez directement à l'étape suivante dans le cas des configurations de bases de données parallèles.

```
# # haswitch phys-hahost3 hahost1
```

2. **Exécutez la commande `sconf -A` pour exclure le noeud.**

Exécutez la commande `sconf(1M)` sur tous les noeuds de grappe. Pour de plus amples renseignements, consultez la page de manuel `sconf(1M)`.

Remarque - Dans cette commande, le numéro indiqué ne représente pas un numéro de noeud. Ce numéro correspond au nombre total de noeuds de grappe qui seront actifs une fois l'opération `sconf` terminée. `sconf` supprime toujours de la grappe le noeud ayant le numéro de noeud le plus élevé. Il n'existe pas de procédure permettant de supprimer par exemple le numéro de noeud 2 dans une grappe à trois noeuds.

Dans cet exemple, le nom de la grappe est `grappe-sc` et le nombre total de noeuds actifs après l'opération `sconf` est de deux.

```
# sconf grappe-sc -A 2
```

Modification du nom d'un noeud de grappe

La commande `sconf(1M)` vous permet de modifier le nom des noeuds de grappe. Pour de plus amples renseignements, consultez la page de manuel `sconf(1M)`.

▼ Comment modifier le nom d'un noeud de grappe

1. Repérez les noms des noeuds de grappe actuels.

Vous pouvez exécuter la commande `scconf -p` sur n'importe quel noeud actif de la grappe.

```
# scconf nom_groupe -p
Configuration courante de la grappe nom_groupe :
  Hôtes de la grappe : phys-hahost1 phys-hahost2 phys-hahost3
  Interfaces réseau privées de
    phys-hahost1 : be0 bel
    phys-hahost2 : be0 bel
    phys-hahost3 : hme0 hme1
```

2. Exécutez la commande `scconf -h` sur tous les noeuds de la grappe.

Exécutez la commande `scconf(1M)` sur tous les noeuds. Pour de plus amples renseignements, consultez la page de manuel `scconf(1M)`.

```
# scconf -h nom_groupe nom_hôte0 [...nom_hôte3]
```

Spécifiez les nouveaux noms de noeuds en respectant l'ordre affiché par la commande `scconf -p`. Par exemple, exécutez la commande suivante sur tous les noeuds pour remplacer le nom de `phys-hahost3` par `phys_os_hahost1`.

```
# # scconf -h grappe-sc phys-hahost1 phys-hahost2 phys-hahost1
```

Modification des interfaces réseau privées

La commande `scconf(1M)` vous permet de modifier les interfaces réseau privées des noeuds de la grappe. Pour de plus amples renseignements, consultez la page de manuel `scconf(1M)`.

▼ Comment modifier les interfaces réseau privées

1. Exécutez la commande `scconf(1M)` sur tous les noeuds de la grappe.

Par exemple :

```
# scconf planètes -i Mercure scid0 scid1
# scconf planètes -i Vénus   scid0 scid1
# scconf planètes -i Pluton  scid0 scid1
# scconf planètes -i Jupiter  scid0 scid1
```

Après l'exécution de ces commandes, les quatre noeuds `Mercure`, `Vénus`, `Pluton` et `Jupiter` utiliseront les interfaces `scid0` et `scid1`.



Attention - N'utilisez pas la commande `ifconfig(1M)` lorsque la grappe est active. Cette commande provoque un comportement imprévisible du système en cours d'exécution.

Impression de la configuration des noeuds

La commande `scconf(1M)` permet d'imprimer les informations relatives à la configuration des noeuds. Pour de plus amples renseignements, consultez la page de manuel `scconf(1M)`.

▼ Comment imprimer la configuration de la grappe

1. Exécutez la commande `scconf(1M)` sur n'importe quel noeud actif de la grappe.

Par exemple :

```
# scconf planètes -p
```

Le système renvoie un message semblable à celui ci-dessous. (Selon le type d'interconnexion privée que vous utilisez, le message peut comporter la mention hme au lieu de scid.)

```
Configuration actuelle de la grappe planètes :
Hôtes de la grappe : Mercure Vénus Pluton Jupiter

Interfaces réseau privées de
  Mercure : scid0 scid1
  Vénus : scid0 scid1
  Pluton : scid2 scid3
  Jupiter : scid2 scid3
```

Ajout et suppression d'hôtes logiques

Les hôtes logiques sont les objets qui échouent lorsqu'un noeud échoue lui-même. Chaque hôte logique se compose d'un ou de plusieurs groupes de disques, d'une adresse IP relocalisable et d'un nom d'hôte logique. Les hôtes logiques ne sont utilisés qu'avec les configurations de services de données HD. Une configuration de bases de données parallèle ne contient pas d'hôtes logiques.

Pour ajouter ou supprimer des hôtes logiques, mettez à jour les informations de votre hôte logique et reconfigurez la grappe. Lors de la première configuration de la grappe, vous fournissez les informations sur la configuration de votre hôte logique à la commande `scinstall(1M)`. Une fois la grappe activée, vous pouvez modifier ces informations de deux façons :

- Exécutez à nouveau la commande `scinstall(1M)`. La commande `scinstall(1M)` fournit une interface à base de menus à la commande `scconf(1M)` et constitue la meilleure manière de modifier la configuration de votre hôte logique. Vous devez exécuter `scinstall(1M)` en tant que `root` (superutilisateur).
- Exécutez la commande `scconf(1M)`. Si vous optez pour cette commande, consultez la page de manuel `scconf(1M)` pour en connaître les options ou obtenir des informations à son sujet. Pour configurer un hôte logique avec plus d'un groupe de disques, exécutez la commande `scconf(1M)` prévue à cet effet.

▼ Comment ajouter un hôte logique à la grappe

Pour ajouter un hôte logique, vous devez fournir les informations suivantes :

- Noms des contrôleurs de réseau public principal des noeuds de la grappe.
- Utilisation de la grappe en tant que serveur de sous-réseau public secondaire.
- Initialisation du service de gestion de réseau public (PNM) sur le noeud de grappe sélectionné (vous ne devez réinitialiser le service PNM que si vous avez ajouté un contrôleur de réseau ou encore si vous avez modifié la configuration des contrôleurs au moment d'ajouter un nouvel hôte logique).
- Nom du nouvel hôte logique.
- Nom du maître par défaut du nouvel hôte logique.
- Nom du groupe de disques inclus dans l'hôte logique.
- Activation de la reprise automatique avec le nouvel hôte logique (une reprise automatique se produit lorsque l'hôte logique échoue et qu'il est remplacé par un noeud de relève. Le maître par défaut de l'hôte défaillant reprend la maîtrise, dès que ce dernier redevient actif dans la grappe). Voir "Désactivation de la commutation automatique", page 93 pour de plus amples renseignements.
- Nom du groupe de disques du nouvel hôte logique.

Ayez ces informations sous la main avant d'ajouter un hôte logique. Il faut noter que vous devez avoir déjà configuré le groupe de disques pour que le nouvel hôte logique puisse l'utiliser. Consultez l'annexe du *Sun Cluster 2.2 Software Installation Guide* traitant de votre gestionnaire de volumes (GV) pour en savoir davantage sur le sujet.

La procédure ci-dessous vous permet d'ajouter un hôte logique à une grappe.

1. Exécutez la commande `scinstall(1M)` et sélectionnez l'option **Modifier du menu principal**.

```
# scinstall
Nom de grappe par défaut de planètes
Remarque : La grappe planète est en cours d'exécution.
            Les installations et les désinstallations sont suspendues
            pendant le fonctionnement de la grappe.

            <<Appuyez sur Entrée pour continuer>>
            Vérification de l'état du module installé
            .....
            ===== Menu principal =====
            1) Modifier - Modification de la configuration de la grappe ou
            du service de données.
            2) Vérifier - Vérification des modules installés.
            3) Lister - Liste des ensembles de modules installés.
            4) Quitter - Fermeture du programme.
            5) Aide - Ecran d'aide du menu.
```

(suite)

(suite)

```
Sélectionner une option : [5] : 1
```

2. Sélectionnez l'option Hôtes logiques du menu Modifier.

```
===== Menu Modifier =====  
Options du menu :  
1) Hôtes logiques      -  
   Modification de la configuration des hôtes logiques.  
2) NAFO                - Réinitialisation de la configuration NAFO.  
  
3) Fermer              -  
   Fermeture du menu et retour au menu principal.  
4) Quitter             - Fermeture du programme.  
5) Aide                - Affichage de l'écran d'aide.  
Sélectionner une option : [5] 1
```

Cette option affiche le menu Configuration des hôtes logiques.

3. Sélectionnez l'option Ajouter du menu Configuration des hôtes logiques.

```
===== Configuration des hôtes logiques =====  
1) Ajouter      - Ajout d'un hôte logique à la grappe.  
2) Supprimer   - Suppression d'un hôte logique de la grappe.  
3) Lister      - Liste des hôtes logiques de la grappe.  
4) Fermer      - Retour au menu précédent.  
5) Quitter     - Fermeture du programme.  
Sélectionner une option : 1
```

Répondez aux questions concernant le nouvel hôte logique.

4. Entrez les informations demandées lorsque vous y êtes invité.

Le menu Configuration des hôtes logiques apparaît à la fin de la partie `scinstall(1M)` de la procédure.

```

Quel est le contrôleur de réseau public principal de l''phys-
hahost1'' ?
Quel est le contrôleur de réseau public principal de l''phys-
hahost2'' ?
La grappe est-elle utilisée comme serveur de sous-
réseaux publics secondaires
(oui/non)[non] ? Réinitialiser la configuration NAFO de ''phys-
hahost1''
avec un contrôleur par groupe (oui/non) ?
Quel est le nom du nouvel hôte logique ? hahost1
Quel est le nom du maître par défaut de ''hahost1'' ? phys-hahost1
Activation de la reprise automatique pour ''hptehdl'' (oui/
non) [non] ?
Quel est le nom du groupe de disques de
l'hôte logique ''hahost1'' [hahost1] ?
Permission d'ajouter l'hôte logique ''hahost1'' maintenant (oui/
non) [oui] ?
/etc/opt/SUNWcluster/conf/ha.cdb
Vérification de l'état du noeud...

```

5. Créez un nouveau système de fichiers d'administration HD et mettez à jour le fichier `/etc/opt/SUNWcluster/conf/hanfs/vfstab`.**logicalhost**.

Lorsque vous ajoutez un nouvel hôte logique, vous devez configurer un système de fichiers dans un groupe de disques de l'hôte logique afin de stocker des informations d'ordre administratif. La marche à suivre pour configurer le systèmes de fichiers d'administration HD varie en fonction du gestionnaire de volumes que vous utilisez. Les étapes nécessaires à cet effet font l'objet des annexes du *Sun Cluster 2.2 Software Installation Guide*.

Remarque - N'utilisez pas d'alias de nom d'hôte avec les hôtes logiques. Les clients NFS où sont montés les systèmes de fichiers Sun Cluster utilisant des alias de nom d'hôte peuvent éprouver des problèmes de récupération à la suite d'un verrouillage `statd`.

▼ Comment supprimer un hôte logique de la grappe

Pour supprimer un hôte logique de la configuration de la grappe, celle-ci doit être active, et l'hôte logique ne doit pas avoir de services de données inscrits.

1. Fermez toutes les applications de service de données en cours d'exécution dans l'hôte logique à supprimer.

```
# haresg -n service_données
```

2. Supprimez l'inscription du service de données.

```
# # haresg -u service_données
```

3. Supprimez l'hôte logique de la grappe.

Exécutez la commande `scinstall(1M)` comme le décrit le *Sun Cluster 2.2 Software Installation Guide* et sélectionnez l'option Modifier du menu principal.

```
# scinstall
Nom de grappe par défaut de planètes
Remarque : La grappe planète est en cours d'exécution.
           Les installations et les désinstallations sont suspendues pendant
           le fonctionnement de la grappe.
           <<Appuyez sur Entrée pour continuer>>
Vérification de l'état du module installé
.....
===== Menu principal =====
1) Modifier -
   Modification de la configuration de la grappe ou du service de données.
2) Vérifier - Vérification des modules installés.
3) Lister   - Liste des ensembles de modules installés.
4) Quitter  - Fermeture du programme.
5) Aide     - Ecran d'aide du menu.
Sélectionner une option : [5] : 1
```

4. Sélectionnez l'option Hôtes logiques du menu Modifier.

```
===== Menu Modifier =====
Options du menu :
1) Hôtes logiques -
   Modification de la configuration des hôtes logiques.
2) NAFO           -
   Réinitialisation de la configuration NAFO.
3) Fermer         -
   Fermeture du menu et retour au menu principal.
4) Quitter        - Fermeture du programme.
5) Aide           - Affichage de l'écran d'aide.
Sélectionner une option : [5] 1
```

Cette option affiche le menu Configuration des hôtes logiques.

5. Sélectionnez l'option Supprimer du menu Configuration des hôtes logiques.

```
==== Configuration des hôtes logiques ====
1) Ajouter      - Ajout d'un hôte logique à la grappe.
2) Supprimer    - Suppression d'un hôte logique de la grappe.
3) Lister       - Liste des hôtes logiques de la grappe.
4) Fermer       - Retour au menu précédent.
5) Quitter      - Fermeture du programme.
Sélectionner une option : 2
```

Cette option permet d'afficher la liste des hôtes logiques configurés.

6. Entrez le nom de l'hôte logique à supprimer de la liste des hôtes logiques configurés.

```
La liste des hôtes logiques inclut :
      hahost1
      hahost2
Lequel voulez-vous supprimer ? hahost1
```

La procédure est maintenant terminée, et le menu Configuration des hôtes logiques apparaît.

7. En tant que root (superutilisateur), supprimez le fichier /etc/opt/SUNWcluster/conf/hanfs/vfstab. *hôte_logique* créé au moment de l'ajout de l'hôte logique à la configuration de la grappe.

Modification de l'adresse IP de l'hôte logique

Pour modifier l'adresse IP d'un hôte logique, supprimez celui-ci puis ajoutez-le avec sa nouvelle adresse, en suivant la procédure de l'“Ajout et suppression d'hôtes logiques”, page 60 ou celle de la présente section.

Pour de plus amples renseignements, consultez la page de manuel `scconf(1M)`.

▼ Comment modifier l'adresse IP d'un hôte logique

Cette procédure ne doit être appliquée qu'à un seul noeud membre de la grappe.

1. **Pour supprimer l'entrée de l'hôte logique des fichiers de configuration, exécutez la commande suivante sur tous les noeuds :**

```
# scconf nom_grappe -L hôte_logique -r
```

2. **Pour créer une nouvelle entrée d'hôte logique avec le même nom d'hôte logique mais avec une nouvelle adresse IP, exécutez la commande suivante sur tous les noeuds de la grappe :**

```
# scconf nom_grappe -L hôte_logique -n liste_noeuds -g groupe_disque -i interfaces_et_IP
```

Reconfiguration de grappe forcée

Pour forcer la reconfiguration de grappe, exécutez la commande `haswitch(1M)` ou modifiez l'appartenance à la grappe avec la commande `scconf(1M)`.

▼ Comment forcer une reconfiguration de grappe

1. **Pour forcer la reconfiguration de grappe, exécutez la commande `haswitch(1M)` sur n'importe quel noeud membre de la grappe. Par exemple :**

```
# haswitch -r
```

Pour de plus amples renseignements, consultez la page de manuel `haswitch(1M)`.

Configuration des services de données de Sun Cluster

Cette section présente les procédures permettant de configurer les services de données de Sun Cluster. La configuration de ces services de données avec les hôtes logiques fait partie de l'installation de la grappe. Une fois cette installation effectuée, la configuration des hôtes logiques et des services de données demeure possible. Pour de plus amples renseignements sur un service de données particulier, reportez-vous au chapitre correspondant du *Sun Cluster 2.2 Software Installation Guide*.

Remarque - Toutes les commandes dont traite la présente section peuvent être exécutées sur n'importe quel noeud membre de la grappe, même sur un noeud ne pouvant maîtriser les hôtes logiques spécifiés ou qui est incapable d'exécuter les services de données spécifiés. Vous pouvez exécuter les commandes même s'il n'y a qu'un seul noeud appartenant à la grappe.



Attention - Les commandes décrites dans cette section mettent à jour la BCG, même en l'absence de quorum. L'arrêt des noeuds et leur remise en fonction selon une séquence incorrecte peut donc provoquer la perte des mises à jour de la BCG. Par conséquent, le dernier noeud à quitter la grappe doit y être ramené en premier avec la commande `scadmin startcluster`. Pour de plus amples renseignements sur la BCG, consultez le *Sun Cluster 2.2 Software Installation Guide*.

▼ Comment configurer un service de données Sun Cluster

1. Vérifiez si les tâches suivantes ont été complétées.

- Les hôtes logiques exécutant les services de données sont configurés. Pour de plus amples renseignements sur la configuration d'un hôte logique, consultez l'"Ajout et suppression d'hôtes logiques", page 60.
- Tous les groupes de disques, volumes logiques et systèmes de fichiers nécessaires sont configurés. Pour de plus amples renseignements à ce sujet, consultez le *Sun Cluster 2.2 Software Installation Guide*.
- Le système de fichiers administratif HD et le fichier `vfstab.hôte_logique` ont été configurés. Cette procédure varie en fonction du gestionnaire de volumes que vous utilisez. Consultez l'annexe du *Sun Cluster 2.2 Software Installation Guide* portant sur la configuration de votre gestionnaire de volumes.

2. Inscrivez le service de données.

Inscrivez chaque service de données Sun Cluster associé à son hôte logique.

```
# hareg -s -r service_données  
[-h hôte_logique]
```

On suppose ici que le service de données a déjà été installé et que ses méthodes sont disponibles.

L'option `-h` ajoutée à la commande `hareg -r` permet de configurer le service de données uniquement sur les hôtes logiques spécifiés par l'argument `hôte_logique`. L'omission de l'option `h` permet de configurer le service de données dans tous les hôtes logiques existants. Pour de plus amples renseignements, consultez la page de manuel `hareg(1M)`

Remarque - S'il faut associer le service de données à tous les hôtes logiques créés après son inscription, exécutez la commande `sconf -s` sur tous les noeuds de grappe pour étendre l'ensemble des hôtes logiques associés au service de données.

3. Configurez le service de données à l'aide de la commande interactive `hadsconfig(1M)` et de la documentation du service de données.

```
# hadsconfig
```

4. Démarrez le service de données.

```
# hareg -y service_données
```

Suppression de la configuration des services de données Sun Cluster

Cette procédure permet de supprimer la configuration des services de données Sun Cluster. Pour de plus amples renseignements sur un service de données particulier, reportez-vous au chapitre correspondant du *Sun Cluster 2.2 Software Installation Guide*.

▼ Comment supprimer la configuration des services de données Sun Cluster

1. **Fermez toutes les applications de service de données dont la configuration doit être supprimée.**

Fermez les différentes applications de service de données selon la procédure habituelle de chacune.

2. **Si le service de données est un système de gestion de base de données (SGBD), arrêtez tous les moniteurs de défaillance.**

3. **Arrêtez le service de données dans tous les hôtes logiques.**

```
# hareg -n service_données
```

4. **Supprimez l'inscription du service de données.**

```
# hareg -u service_données
```

Remarque - Si la commande `hareg -u` échoue, la base de données de configuration de grappe (BCG) peut devenir incohérente. Dans ce cas, exécutez la commande `scconf nom_grappe -R service_données` sur tous les noeuds de grappe pour forcer la suppression du service de données de la BCG.

5. **(Facultatif) Retirez les hôtes logiques de la configuration de la grappe.**

Vous ne pouvez supprimer un hôte logique de la configuration de la grappe qu'en dissociant tous les services de données de cet hôte.

Les méthodes suivantes permettent de supprimer un hôte logique.

Exécutez la commande `scconf(1M)` suivante sur un noeud membre de la grappe :

```
# scconf nom_grappe -L hôte_logique -r
```

Vous pouvez aussi lancer la commande `scinstall(1M)` comme indiqué dans "Ajout et suppression d'hôtes logiques", page 60. Si vous utilisez la commande `scinstall(1M)`, il n'est pas nécessaire de procéder à la reconfiguration de la grappe, comme l'indique l'étape 6, page 69.

6. **Reconfigurez la grappe avec la commande `haswitch(1M)`.**

```
# haswitch -r
```

Vous avez le choix de supprimer ou de renommer les fichiers `vfstab.hôte_logique` et `dfstab.hôte_logique` associés à l'hôte logique que vous avez supprimé, et de récupérer ensuite l'espace occupé par les volumes et les systèmes de fichiers de ce dernier. La suppression avec la commande `scconf(1M)` ne modifie pas ces fichiers.

Ajout de services de données Sun Cluster

Vous pouvez ajouter un service de données à une grappe existante à l'aide de la commande `scinstall(1M)`. Pour de plus amples renseignements, reportez-vous au Chapitre 3 du *Sun Cluster 2.2 Software Installation Guide*.

Pour ajouter un service de données à une grappe à deux noeuds avec une BCG partagée, des opérations supplémentaires sont nécessaires. Procédez comme suit pour ajouter un service de données à une grappe de ce type.

▼ Comment ajouter un service de données à une grappe à deux noeuds avec une BCG partagée

1. Supprimez le partage de la BCG.

Vous devez reconfigurer la grappe pour supprimer le partage de la BCG avant d'ajouter un nouveau service de données. Exécutez la commande suivante sur les deux noeuds, en tant que root (superutilisateur) lorsque les deux noeuds sont dans la grappe :

```
phys-hahost1# /opt/SUNWcluster/bin/scconf nom_grappe -S none  
phys-hahost2# /opt/SUNWcluster/bin/scconf nom_grappe -S none
```

Remarque - Vous devez supprimer le partage de la BCG. Si vous tentez d'ajouter un service de données alors que la BCG est partagée, seul le fichier local `ccd.database` sera mis à jour, et non le fichier BCG partagé, ce qui provoque l'échec de l'inscription du nouveau service de données.

2. Ajoutez les nouveaux services de données à l'aide des commandes suivantes.

Exécutez toutes les commandes en tant que root (superutilisateur). Dans ces exemples, les noms de noeud sont `phys-hahost1` et `phys-hahost2`.

a. Arrêtez la grappe sur le premier noeud.

```
phys-hahost1# scadmin stopnode
```

b. Utilisez la commande `scinstall(1M)` pour ajouter le nouveau module de service de données sur le premier noeud.

Pour de plus amples renseignements, consultez le Chapitre 3 du *Sun Cluster 2.2 Software Installation Guide*. Cette opération met automatiquement à jour le fichier BCG local.

```
phys-hahost1#scinstall
```

c. Arrêtez la grappe sur le deuxième noeud.

Remarque - Les clients ne pourront plus accéder aux services de données existants après l'arrêt de la grappe sur le deuxième noeud. L'accès est rétabli lorsque vous redémarrez la grappe sur le premier noeud.

```
phys-hahost2# scadmin stopnode
```

d. Redémarrez la grappe sur le premier noeud.

```
phys-hahost2# scadmin startcluster phys-hahost1 nom_grappe
```

e. Utilisez la commande `scinstall(1M)` pour ajouter le nouveau module de service de données sur le deuxième noeud. Pour de plus amples renseignements, consultez le Chapitre 3 du *Sun Cluster 2.2 Software Installation Guide*. Cette opération met automatiquement à jour le fichier BCG local.

```
phys-hahost2#scinstall
```

f. Ajoutez le deuxième noeud à la grappe.

```
phys-hahost2# scadmin startnode
```

3. Rétablissez la BCG partagée.

Exécutez la commande `scconf(1M)` sur les deux noeuds en tant que root (superutilisateur). Pour de plus amples renseignements, consultez la page de manuel `scconf(1M)`.

```
phys-hahost1# /opt/SUNWcluster/bin/scconf nom_groupe -S ccdvol  
phys-hahost2# /opt/SUNWcluster/bin/scconf nom_groupe -S ccdvol
```

Exécutez ensuite la commande `confccdssa(1M)` sur un seul noeud en tant que root (superutilisateur). Pour de plus amples renseignements, consultez la page de manuel `confccdssa(1M)`.

```
phys-hahost1# /opt/SUNWcluster/bin/confccdssa
```

Modification du fichier `clusters`

Le fichier `/etc/clusters` contient les informations relatives aux grappes connues dans le domaine de nomenclature local. Ce fichier, qui fait correspondre un nom de grappe à la liste des noms d'hôtes dans la grappe, peut être un mappage NIS ou NIS+, ou être créé localement dans le répertoire `/etc`.

Les opérations suivantes nécessitent la mise à jour du fichier `/etc/clusters` :

- Modification d'un nom d'hôte quelconque
- Ajout de grappes à Sun Cluster

Pour de plus amples renseignements sur le mappage NIS ou NIS+, consultez le *Guide d'administration NIS/NIS+*. Le *Sun Cluster 2.2 Software Installation Guide* contient

des informations sur la création du fichier `/etc/clusters`. La modification des fichiers NIS/NIS+ doit s'effectuer au niveau du serveur NIS/NIS+.

▼ Comment modifier le fichier `clusters`

1. **Modifiez le fichier `/etc/clusters` de façon à ajouter le nom de la grappe et les noms d'hôtes physiques de tous les noeuds.**

Par exemple, entrez la commande ci-dessous pour créer la grappe `hacluster` composée du noeud 0 `phys-hahost1`, du noeud 1 `phys-hahost2`, du noeud 2 `phys-hahost3` et du noeud 3 `phys-hahost4` et ajoutez cette entrée au fichier :

```
# Sun Enterprise Cluster nodes
grappehd phys-hahost1 phys-hahost2 phys-hahost3 phys-hahost4
```

Les fichiers `/etc/clusters` doivent être identiques sur tous les noeuds. Apportez les mêmes modifications au fichier `/etc/clusters` dans chaque noeud de la grappe.

▼ Comment créer la table `clusters`

1. **Dans un environnement NIS+, vous devez créer une table `clusters`. Les entrées de cette table sont identiques à celles du fichier `/etc/clusters`.**

Par exemple, utilisez la commande suivante pour créer une table `clusters` dans le domaine appelé `mon_domaine` d'un environnement NIS+ :

```
# nistbladm -c key-value key=SI value= clusters.mon_domaine.
```

Remarque - Le point (.) à la fin de la commande `nistbladm` est obligatoire.

Modification du fichier `serialports`

Le fichier `serialports` fait correspondre un nom d'hôte au concentrateur de terminaux et à son port série où est connectée la console de l'hôte. Ce fichier peut être un mappage NIS ou NIS+, ou être créé localement dans le répertoire `/etc`.

Les opérations suivantes nécessitent la mise à jour du fichier `serialports` :

- Modification de nom(s) d'hôte(s)
- Modification du nom du concentrateur de terminaux
- Modification du numéro de port de l'hôte dans le concentrateur de terminaux
- Ajout d'hôtes au concentrateur de terminaux
- Ajout de noeuds de grappe

Le *Sun Cluster 2.2 Software Installation Guide* contient des informations sur la création du fichier `/etc/serialports`. Pour de plus amples renseignements sur le mappage NIS ou NIS+, consultez le *Guide d'administration NIS/NIS+*.

▼ Comment modifier le fichier `serialports`

1. En tant que `root` (superutilisateur), créez un fichier `serialports` dans le répertoire `/etc`.
2. Dans le cas d'un système Sun Enterprise[™] 10000, entrez la ligne `nom_hôte nom_pss 23` dans le fichier `serialports`. Pour tout autre système matériel, entrez la ligne `nom_hôte concentrateur_terminaux port_série` dans le fichier `serialports`.

Avec un système Sun Enterprise 10000 :

```
# # Sun Enterprise Cluster nodes
phys-hahost1 nom_pss 23
phys-hahost2 nom_pss 23
phys-hahost3 nom_pss 23
phys-hahost4 nom_pss 23
```

Avec tout autre système matériel :

```
# Sun Enterprise Cluster nodes
phys-hahost1 ct_grappehd 5002
phys-hahost2 ct_grappehd 5003
phys-hahost3 ct_grappehd 5004
phys-hahost4 ct_grappehd 5005
```

▼ Comment créer la table `serialports`

1. Dans un environnement NIS+, vous devez créer une table `serialports`. Les entrées de cette table sont identiques à celles du fichier `/etc/serialports`. Utilisez la commande suivante pour créer une table `serialports` dans un domaine appelé `mon_domaine` d'un environnement NIS+ :

```
# nistbladm -c key-value key=SI value=clusters.mon_domaine.
```

Remarque - Le point (.) à la fin de la commande `nistbladm` est obligatoire.

Modification des informations TC/PSS

Lorsque vous installez le logiciel Sun Cluster, vous devez fournir des informations sur le concentrateur de terminaux (CT) ou sur un processeur de services système (PSS), qui sont stockées dans la base de données de configuration de grappe (BCG).

Ces informations servent dans les cas suivants :

- Interruption forcée de noeuds bloqués (protection contre les défaillances).
- Mise en oeuvre d'un verrou de grappe empêchant l'intégration des noeuds partitionnés dans la grappe.

Ces deux mécanismes assurent le maintien de l'intégrité des données dans le cas de grappes à quatre noeuds reliées directement à des périphériques de stockage.

Exécutez la commande `scconf(1M)` pour modifier les informations sur le CT ou le PSS associées à un noeud particulier, comme le décrivent les procédures ci-dessous.

Pour plus d'informations sur l'installation et la configuration du CT ou du PSS, consultez le chapitre traitant du concentrateur de terminaux dans le *Sun Cluster 2.2 Hardware Site Preparation, Planning, and Installation Guide*.

▼ Comment modifier les informations CT/PSS

Pour modifier les informations CT ou PSS, exécutez la commande `scconf(1M)` sur tous les noeuds de la grappe. Pour chacun de ces noeuds, entrez les nouvelles informations pertinentes. Les exemples suivants illustrent la syntaxe de la commande `scconf(1M)` pour chaque type de modification des informations.

1. **Type d'architecture de noeud et adresse IP** – Entrez les informations suivantes : **nom de grappe, nom d'hôte, nouveau type d'architecture et nouvelle adresse IP.**

```
# scconf nom_grappe -H nom_hôte -d E10000 -t nouvelle_adresse_ip
```

Remarque - Il est possible de connecter de multiples hôtes au même CT ; l'option `-H` n'influence que les informations associées à l'hôte spécifié sur la ligne de commande.

2. **Mot de passe d'un CT ou d'un PSS** – Entrez les informations suivantes : **nom de grappe, adresse IP et nouveau mot de passe.**

```
# scconf nom_grappe -t adresse_ip -P  
adresse_ip (129.34.123.51) Mot de passe :
```

3. **Numéro de port d'une console PSS** – Entrez les informations suivantes : **nom de grappe, nom d'hôte et nouveau numéro de port.**
 - Si vous utilisez un concentrateur de terminaux, spécifiez un numéro de port de CT libre de 1 à N.
 - Si vous utilisez un PSS, spécifiez la valeur -1.

```
# scconf nom_grappe -H nom_hôte -p nouveau_numéro_port
```

4. Nom de CT ou adresse IP – Entrez les informations suivantes : nom de grappe, nom d’hôte et nouveau nom de CT ou nouvelle adresse IP.

```
# sccnf nom_grappe -H nom_hôte -t nouveau_nom_ct|nouvelle_adresse_ip
```

Pour de plus amples renseignements sur la modification des informations CT ou PSS, consultez la page de manuel `sccnf(1M)` et Chapitre 8.

Modification du périphérique de quorum

Seules les configurations VxVMont recours aux périphériques de quorum. Les configurations Solstice DiskSuite ne les utilisent pas.

La commande `sccnf -q` permet d’associer un périphérique de quorum à un disque ou un contrôleur. Cette option s’avère utile lorsque le périphérique de quorum doit être réparé. Pour de plus amples renseignements, consultez la page de manuel `sccnf(1M)`.

Remarque - Si le périphérique de quorum est un disque, la commande `sccnf -q` doit être utilisée à chaque modification de l’adresse du disque (dont le format est `cxydzs2`), même si le numéro de série du disque demeure le même. Une telle modification de l’adresse du disque ne peut se produire que lorsque la fente SBus du contrôleur de disque change.



Attention - N’exécutez pas la commande `sccnf` avec l’option `-q` pour modifier la topologie du périphérique de quorum pendant que la grappe fonctionne. Vous ne pouvez ni ajouter ni supprimer un périphérique de quorum entre deux noeuds de grappe quelconques. Plus particulièrement : vous ne pouvez pas ajouter un périphérique de quorum s’il n’y en avait pas auparavant, et vous ne pouvez pas spécifier qu’il n’y a “pas de périphérique de quorum” s’il en existe déjà un. Cependant, vous pouvez modifier un périphérique de quorum (par exemple, en transférant son affectation d’un disque à un autre) dans une grappe en cours d’exécution avec l’option de commande `sccnf -q`.

▼ Comment modifier le périphérique de quorum

1. **Avant de réparer le périphérique de quorum, vous pouvez l'associer à un autre périphérique en exécutant la commande `scconf -q` sur tous les noeuds de grappe.**

Par exemple, pour modifier le périphérique de quorum de la grappe `grappehd` pour les noeuds `phys-hahost1` et `phys-hahost2`, exécutez la commande `scconf(1M)` de la façon suivante.

```
# scconf haclust -q phys-hahost1 phys-hahost2
Sélectionnez un périphérique de quorum pour les noeuds 0 (phys-
hahost1) et 1
(phys-hahost2).
Tapez le numéro correspondant au choix voulu.
Par exemple : 1<CR>
1) DISQUE : c2t2d0s2:01943825
2) DISQUE : c2t3d0s2:09064321
3) DISQUE : c2t4d0s2:02171369
4) DISQUE : c2t5d0s2:02149886
5) DISQUE : c2t8d0s2:09062992
6) DISQUE : c2t9d0s2:02166472
7) DISQUE : c3t2d0s2:02183692
8) DISQUE : c3t3d0s2:02183488
9) DISQUE : c3t4d0s2:02160277
10) DISQUE : c3t5d0s2:02166396
11) DISQUE : c3t8d0s2:02164352
12) DISQUE : c3t9d0s2:02164312
Périphérique de quorum : 12
```

L'option `-q` sonde la liste des périphériques connectés à chaque noeud et affiche la liste des périphériques que se partagent les deux noeuds. Il est alors possible de sélectionner le périphérique de quorum à partir de la liste affichée.

Pour activer le sondage des périphériques connectés aux hôtes distants, la commande modifie le fichier local `.rhosts` de manière à activer les permissions `rsh(1)`. Les permissions sont supprimées après l'exécution de la commande.

Remarque - Un tel comportement ne se produit que si la commande est exécutée simultanément sur tous les noeuds. Pour préserver l'accès distant à la racine, utilisez l'option `-m`.

2. **Sélectionnez un contrôleur SSA ou un disque de la liste comme périphérique de quorum.**

Si vous choisissez un contrôleur SSA, la liste des disques dans ce contrôleur apparaît.

3. Si vous avez choisi un contrôleur SSA à Etape 2, page 78, vous pouvez sélectionner un disque de ce contrôleur comme périphérique de quorum.

Si vous n'avez pas sélectionné de disque à cette étape, le contrôleur SSA sélectionné à l'étape précédente demeure le périphérique de quorum.

L'option `-q` vérifie également si le périphérique de quorum réserve un noeud, en raison d'un autre noeud n'appartenant pas à la grappe. Dans ce cas, l'option `-q` libère la réservation de l'ancien périphérique de quorum et réserve le nouveau périphérique de quorum.

Remarque - Tous les noeuds spécifiés doivent être initialisés pour que la commande `scconf -q` puisse s'exécuter correctement. Si l'un des noeuds n'est pas initialisé, la commande affiche la liste de tous les périphériques du noeud local. Assurez-vous de sélectionner un périphérique partagé comme périphérique de quorum.

Si vous connaissez déjà le nom du périphérique devant servir de périphérique de quorum, utilisez l'option `-m` pour spécifier le nouveau périphérique.

```
# scconf nom_grappe -q -m périphérique_quorum nom_hôte1 nom_hôte2
```

Le périphérique de quorum peut être le nom universel (World Wide Name ou WWN) d'un contrôleur SSA (WWN), un identificateur de disque ayant le format `WWN.no_série_disque` pour les disques dans les contrôleurs SSA, ou encore un identificateur de disque ayant le format `adresse_disque :no_série_disque` pour les disques autres que SSA. Le paramètre adresse-disque doit respecter le format `cxydzs2`. Exécutez la commande `finddevices(1M)` si vous désirez obtenir les numéros de série des disques SSA et non SSA.

Si la grappe possède plus de deux noeuds partageant le même périphérique de quorum, vous pouvez utiliser les options `-q -D` pour spécifier un nouveau périphérique de quorum commun.

```
# scconf nom_grappe -q -D
```

Etant donné que tous les hôtes de la grappe partagent le même périphérique, il n'est pas nécessaire de spécifier une liste d'hôtes.

Cette option interactive sonde la liste des périphériques connectés à chaque hôte et affiche la liste des périphériques partagés. Sélectionnez le périphérique de quorum dans cette liste.

Remarque - Tous les hôtes actifs définis dans la grappe doivent être initialisés pour que la commande `scconf -q -D` puisse s'exécuter correctement. Si l'un des hôtes n'est pas initialisé, la commande sonde et affiche la liste de tous les périphériques de l'hôte local. Assurez-vous de sélectionner un périphérique partagé comme périphérique de quorum.

L'option `-q -D` vérifie également si le périphérique de quorum réserve un noeud, en raison de la présence d'un autre noeud qui n'est pas membre de la grappe. Dans ce cas, l'option libère la réservation de l'ancien périphérique de quorum et réserve le nouveau périphérique de quorum.

L'exécution simultanée de cette commande sur tous les noeuds par l'intermédiaire des interfaces IUG `cconsole` et `crlogin` modifie le fichier local `.rhosts` de manière à activer les permissions `rsh(1)`. Cela permet le sondage des périphérique connectés aux hôtes distants. Les permissions sont supprimées après l'exécution de la commande.

Vous pouvez ajouter l'option `-m` pour empêcher l'accès distant à la racine. L'option `m` configure le périphérique de quorum et constitue le dernier argument de la commande pour les noeuds spécifiés.

```
# scconf nom_grappe -q -D -m périphérique_quorum
```

Le périphérique de quorum est un identificateur de disque dont le format est `cxydzs2 :no_série_disque`. Exécutez la commande `finddevices(1M)` pour obtenir les numéros de série des disques.

Configuration des délais de transition de la grappe

Sun Cluster comporte des délais de transition progressive de la grappe. Cette fonction permet de réserver et de libérer les hôtes logiques de la structure HD selon les modifications apportées à l'appartenance de la grappe. Vous pouvez modifier ces délais pour gérer efficacement les configurations comportant un grand nombre de services de données sur chaque noeud. L'établissement de valeurs fixes de délai avec un grand nombre de configurations ne constitue pas une solution pratique, à moins que leurs valeurs par défaut soient très élevées.

Deux facteurs principaux influencent le choix de la valeur du délai :

- le nombre d'hôtes logiques par noeud de grappe
- le nombre de services de données dans un hôte logique

Il n'est pas facile de déterminer la meilleure valeur pour une installation particulière. Seule la méthode par essais et erreurs permet d'établir une valeur adéquate. Vous pouvez toutefois vous servir des messages de la console de grappe, relatifs au début et à la fin de chaque étape de la transition de la grappe, à titre d'indication générale. Ces messages vous donnent une bonne idée de la durée d'exécution de chaque étape.

Les délais doivent tenir compte des scénarios les plus pessimistes. Lorsque vous configurez ces délais, posez-vous la question suivante : "Quel est le plus grand nombre d'hôtes logiques qu'un noeud de grappe doit pouvoir maîtriser à tout moment ?"

Par exemple, dans une configuration N+1, le noeud de relève peut maîtriser potentiellement tous les hôtes logiques des autres noeuds de la grappe. Dans ce cas, les délais doivent être suffisamment longs pour couvrir la durée nécessaire à la maîtrise de tous les hôtes logiques configurés de la grappe.

▼ Comment modifier les délais de la grappe

1. Pour régler les délais de reconfiguration de la grappe, exécutez la commande

```
scconf -T.
```

Par exemple, exécutez la commande suivante sur chaque noeud de la grappe pour régler la valeur configurable de délai d'étape de transition à 500 secondes :

```
# scconf nom_grappe -T 500
```

La valeur par défaut des étapes est de 720 secondes. Exécutez la commande `ssconf -p` pour afficher les valeurs actuelles de délai.

Pour chaque étape de reconfiguration, le nombre de services de données configurés sur un hôte logique a une influence sur la durée nécessaire à la maîtrise de cet hôte. S'il n'y a pas assez de temps pour maîtriser un hôte logique — c'est-à-dire si la valeur du paramètre `loghost_timeout` est trop faible — un message semblable à celui-ci apparaît sur la console :

```
ID[SUNWcluster.ccd.cdd.5001] : erreur cmd prise d'instantané =  
délai à la commande /opt/SUNWcluster/bin/loghost_sync.
```

Dans cet exemple, la structure de la grappe “tente” de ramener le système à un état cohérent par la libération de l’hôte logique. Si elle n’y parvient pas, le noeud peut être rejeté de la grappe pour prévenir toute incohérence.

2. **Exécutez la commande `scconf -l` pour régler le paramètre `loghost_timeout`.** La valeur par défaut est 180 secondes.

Remarque - Les délais des étapes de reconfiguration ne peuvent pas être inférieurs à la valeur `loghost_timeout`. Autrement, il se produit une erreur, et le fichier de configuration de la grappe n’est pas modifié. Pour vérifier que cette exigence est bien respectée, exécutez la commande `scconf -T` ou `scconf -l`. Un avertissement apparaît si l’un ou l’autre de ces délais est réglé à 100 secondes ou moins.

Administration générale de Sun Cluster

Le présent chapitre fournit des directives sur les sujets suivants.

- “Démarrage de la grappe et des noeuds de grappe”, page 83
- “Arrêt de la grappe et des noeuds de grappe”, page 86
- “Commutation des hôtes logiques”, page 92
- “Désactivation de la commutation automatique”, page 93
- “Activation du mode maintenance des hôtes logiques”, page 93
- “Récupération à partir des partitions de grappe”, page 94
- “Maintenance du système de fichiers `/var`”, page 97
- “Administration du temps dans les configurations de Sun Cluster”, page 98
- “Remplacement d’un noeud défectueux”, page 99
- “Remplacement d’un concentrateur de terminaux défectueux”, page 103
- “Administration de la base de données de configuration de grappe”, page 104
- “Réservation de disques partagés (VxVM)”, page 109

Démarrage de la grappe et des noeuds de grappe

La commande `scadmin startcluster` attribue le statut de premier membre à un noeud de la grappe. Celui-ci devient le noeud 0 de la grappe. L’exécution de la seule commande `scadmin startnode` démarre les autres noeuds de Sun Cluster. Cette commande démarre les programmes nécessaires à la synchronisation multi-noeuds et coordonne l’intégration des autres noeuds au premier noeud (si Sun Cluster est déjà

en cours d'exécution sur celui-ci). Vous pouvez supprimer des noeuds de la grappe en y exécutant la commande `scadmin` avec l'option `stopnode`.

Attribuez le statut de premier membre de grappe au noeud local. Pour que la commande `scadmin startcluster` fonctionne correctement, le noeud local doit être un noeud configuré de la grappe. Cette commande *doit* avoir été exécutée avec succès avant que d'autres noeuds puissent s'intégrer à la grappe. Si, pour une raison quelconque, le noeud local abandonne la procédure pendant l'intégration des autres noeuds à la grappe, il peut en résulter une altération de la BCG. Dans un tel cas, restaurez la BCG en suivant la procédure de la section "Comment restaurer la BCG", page 105.

Pour faire du noeud local un noeud configuré de la grappe, voir "Ajout et suppression de noeuds de grappe", page 54.

▼ Comment démarrer la grappe

Il est important qu'aucun autre noeud n'exécute le logiciel de grappes à ce moment. Si le noeud local détecte un autre noeud de grappe actif, il abandonne.

1. Démarrez le premier noeud de la grappe avec la commande `scadmin(1M)`.

```
# scadmin startcluster noeud_local nom_grappe
```

L'option `startcluster` ne peut fonctionner si `noeud_local` ne correspond pas au nom du noeud sur lequel la commande est exécutée. Pour de plus amples renseignements, consultez la page de manuel `scadmin(1M)`.

Par exemple :

```
phys-hahost1# scadmin startcluster phys-hahost1 grappehd
Le noeud spécifié est phys-hahost1
La grappe spécifiée est grappehd
===== AVERTISSEMENT =====
=          Création d'une nouvelle grappe          =
=====
Vous tentez de démarrer le noeud de grappe "phys-hahost1"
comme le seul noeud actif dans une nouvelle grappe. Il est
important qu'aucun autre noeud de grappe ne soit actif à ce
moment. Si ce noeud reçoit des informations des autres noeuds
de la grappe, il abandonne. Les autres noeuds ne peuvent intégrer
la grappe qu'à la fin de l'exécution de la commande. La présence
de plus d'une grappe active peut provoquer l'altération des données.
Voulez-vous continuer ? [o,n,?] y
```

Si vous recevez le message d'erreur `reconfig.4013`, c'est qu'il y a déjà un noeud dans la grappe ou qu'un autre noeud est cours d'arrêt. Exécutez la commande `get_node_status(1M)` sur le noeud qui pourrait être actif pour en connaître l'état.

2. Ajoutez tous les autres noeuds à la grappe.

Exécutez la commande ci-dessous sur tous les autres noeuds, séquentiellement.

```
# scadmin startnode
```

Si vous recevez le message d'erreur `reconfig.4015` ci-après, il se peut qu'il n'y ait pas de grappe existante. Redémarrez la grappe avec la commande `scadmin startcluster noeud_local`.

```
SUNWcluster.clustd.reconf.4015
''Abandon--il n'y a pas de grappe existante ou intacte à laquelle intégrer les noeuds.''
```

Ce message peut également provenir d'une défaillance de partition ou de noeud. (Par exemple, un troisième noeud tente de s'intégrer à une grappe à deux noeuds lorsque l'un d'eux échoue.) Dans un tel cas, attendez la fin des défaillances. S'il y a lieu, corrigez les problèmes et tentez ensuite de réintégrer le noeud à la grappe. S'il manque un des modules requis, la commande échoue et la console affiche un message semblable à celui-ci :

```
Nom de grappe par défaut de haclust
Erreur : module SC 'SUNWccm' non installé !
Abandon du démarrage de la grappe.
```

Pour des informations au sujet de l'installation des modules Sun Cluster, consultez le *Sun Cluster 2.2 Software Installation Guide*.

Arrêt de la grappe et des noeuds de grappe

Pour mettre un noeud dans n'importe quel mode à l'exception du mode multi-utilisateurs, ou pour arrêter ou réinitialiser un noeud, vous devez arrêter le moniteur d'appartenance de Sun Cluster. Utilisez alors la méthode d'administration préférée de votre site pour assurer la maintenance subséquente du noeud.

Pour arrêter la grappe, vous devez arrêter également le moniteur d'appartenance sur tous les noeuds de la grappe. Pour ce faire, exécutez simultanément la commande `scadmin stopnode` sur tous les noeuds.

- Vous ne pouvez arrêter le moniteur d'appartenance que lorsque le noeud local Sun Cluster ne possède aucun hôte logique.
- Pour arrêter la moniteur d'appartenance sur un noeud, commutez le ou les hôtes logiques vers un autre noeud avec la commande `haswitch(1M)`, puis exécutez la commande suivante :

```
phys-hahost1# haswitch hôte_destinationhôte_logique
phys-hahost1# scadmin stopnode
```

Si le noeud possède un hôte logique lorsque la commande `scadmin stopnode` est exécutée, il en perd la propriété au profit d'un autre noeud en mesure de maîtriser cet hôte avant l'arrêt du moniteur d'appartenance. Si le maître de relève de l'hôte logique est inactif, la commande `scadmin stopnode` interrompt les services de données en plus d'arrêter le moniteur d'appartenance.

Une fois la commande `scadmin stopnode` exécutée, Sun Cluster demeure arrêté jusqu'à l'exécution de la commande `scadmin startnode`, même après plusieurs réinitialisations du système.

La commande `scadmin stopnode` retire le noeud de la grappe. Si aucune autre défaillance ne se produit simultanément, vous pouvez arrêter autant de noeuds que vous le désirez, sans perdre le quorum au niveau des noeuds restants. (L'absence de quorum provoque l'arrêt de toute la grappe.)

Si vous arrêtez un noeud pour réparer un disque, vous devez également préparer le disque d'initialisation ou de données en suivant les procédures relatives aux disques d'initialisation du Chapitre 10, ou celles qui portent sur les disques de données dans la documentation fournie avec votre gestionnaire de volumes.

Vous devrez peut-être arrêter un ou plusieurs noeuds de Sun Cluster pour procéder à la maintenance matérielle, comme l'ajout ou la suppression de cartes SBUS. Les

sections suivantes décrivent la procédure requise pour arrêter un seul noeud ou la grappe en entier.

Remarque - Dans une grappe comportant deux noeuds ou davantage et un système de stockage directement connecté, un problème peut se produire si le dernier noeud de la grappe subit une erreur grave ou quitte la grappe de façon inhabituelle (sans exécuter la transition `stopnode`). Dans un tel cas, tous les noeuds sont retirés de la grappe et celle-ci n'existe plus, mais comme le dernier noeud s'est retiré de manière inhabituelle, il tient toujours le verrouillage de noeud. Lors d'un appel ultérieur de la commande `scadmin startcluster`, celle-ci ne pourra obtenir le verrouillage de noeud. Pour résoudre ce problème, effacez manuellement le verrouillage de noeud avant de redémarrer la grappe, à l'aide de la procédure "Comment supprimer un verrouillage de noeud après une erreur grave de grappe", page 89.

▼ Comment arrêter Sun Cluster sur un noeud de grappe

1. **S'il n'est pas nécessaire que les données demeurent disponibles, placez les hôtes logiques (groupes de disques) en mode de maintenance.**

```
phys-hahost2# haswitch -m hôte_logique
```

Pour de plus amples renseignements, consultez la page de manuel `haswitch(1M)`.

Remarque - L'arrêt d'un noeud Sun Cluster peut être obtenu avec la commande `halt(1M)` ; il s'ensuit une récupération des services de l'hôte logique sur le noeud de relève. Cependant, le résultat de la commande `halt(1M)` peut créer une confusion au niveau du noeud. L'exécution de la commande `haswitch(1M)` constitue une méthode de commutation de la propriété des hôtes logiques plus fiable.

2. **Arrêtez Sun Cluster sur un noeud sans arrêter les services en cours d'exécution sur les autres noeuds de la grappe.**

```
phys-hahost1# scadmin stopnode
```

Remarque - Lorsque vous arrêtez un noeud, le message d'erreur suivant peut s'afficher : `in.rdiscd[517] : setsockopt (IP_DROP_MEMBERSHIP) : Impossible d'attribuer l'adresse demandée`. Cette erreur est due à un problème de synchronisation entre le démon `in.rdiscd` et le module IP. Elle est sans gravité et peut être ignorée.

3. Arrêtez le noeud.

```
phys-hahost1# halt
```

Le noeud peut maintenant être réparé.

▼ Comment arrêter Sun Cluster sur tous les noeuds

L'arrêt de tous les noeuds d'une configuration Sun Cluster peut s'avérer nécessaire lorsque l'environnement présente des conditions dangereuses comme une panne du système de refroidissement ou un orage violent.

1. Arrêtez simultanément le moniteur d'appartenance sur tous les noeuds avec la commande `scadmin(1M)`.

Exécutez cette commande sur la console de chaque noeud de la grappe. Laissez chaque noeud quitter la grappe et attendez que les noeuds restants se soient reconfigurés complètement avant d'exécuter la commande sur le noeud suivant

```
phys-hahost1# scadmin stopnode
...
```

2. Arrêtez tous les noeuds avec la commande `halt(1M)`.

```
phys-hahost1# halt
...
```

▼ Comment arrêter un noeud Sun Cluster

1. **Arrêtez un noeud Sun Cluster quelconque avec la commande** `halt(1M)` **ou** `uadmin(1M)`.

Si le moniteur d'appartenance est en cours d'exécution à l'arrêt d'un noeud, il y a de fortes chances que ce dernier provoque un "dépassement du délai imparti". Le message suivant apparaît alors :

```
panic[cpu9]/thread=0x50f939e0 : Dépassement du délai imparti - unité
```

Une telle situation peut être évitée en arrêtant le moniteur d'appartenance avant le noeud. Pour de plus amples renseignements à ce sujet, voyez la procédure "Comment arrêter Sun Cluster sur tous les noeuds", page 88.

▼ Comment supprimer un verrouillage de noeud après une erreur grave de grappe

Dans une grappe comportant deux noeuds ou davantage et un système de stockage directement connecté, un problème peut se produire si le dernier noeud de la grappe subit une erreur grave ou quitte la grappe de façon inhabituelle (sans exécuter la transition `stopnode`). Dans un tel cas, tous les noeuds sont retirés de la grappe et celle-ci n'existe plus, mais comme le dernier noeud s'est retiré de manière inhabituelle, il tient toujours le verrouillage de noeud. Lors d'un appel ultérieur de la commande `scadmin startcluster`, celle-ci ne pourra obtenir le verrouillage de noeud.

Pour contourner ce problème, effacez manuellement le verrouillage de noeud avant de redémarrer la grappe. Utilisez la procédure suivante pour effacer manuellement le verrouillage de noeud et redémarrer la grappe, après fin anormale de celle-ci.

1. **En tant que root (superutilisateur), affichez la configuration de la grappe.**

```
# scconf nom_grappe -p
```

Repérez cette ligne dans la sortie :

```
nom_grappe Locking TC/SSP, port : A.B.C.D, E
```

- Si *E* est un nombre positif, le verrouillage de noeud est réglé sur le concentrateur de terminaux *A.B.C.D* et sur le port *E*. Passez à Étape 2, page 90.

- Si *E* a la valeur -1, le verrouillage est défini à un SSP. Passez à Etape 3, page 90.

2. Dans le cas d'un verrouillage de noeud sur un concentrateur de terminaux (CT), procédez comme suit.

- a. Etablissez une connexion telnet au concentrateur de terminaux *nom-ct*.**

```
$ telnet nom_ct
Essai de 192.9.75.51... Connecté à nom_ct. Le caractère d'échappement est '^]'.
```

Appuyez sur Entrée pour continuer.

- b. Précisez `cli` (interface de ligne de commande).**

```
Entrez le nom de port ou le numéro d'Annexe : cli
```

- c. Connectez-vous comme root (superutilisateur).**

- d. Exécutez la commande `admin`.**

```
annex# admin
```

- e. Réinitialisez le port *E*.**

```
admin : reset E
```

- f. Mettez fin à la connexion telnet.**

```
annex# hangup
```

- g. Passez à Etape 4, page 91.**

3. Dans le cas d'un verrouillage de noeud sur un processeur de services système (PSS), suivez les étapes ci-dessous.

- a. Connectez-vous au PSS.**

```
$ telnet nom_pss
```

b. Ouvrez une session en tant qu'utilisateur `pss`.

c. Affichez les informations concernant le fichier `nom_grappe.lock` à l'aide de la commande suivante. (Ce fichier est un lien symbolique à `/proc/csh.pid`.)

```
$ ls -l /var/tmp/nom_grappe.lock
```

d. Recherchez le processus `csh.pid`.

```
$ ps -ef | grep csh.pid
```

e. Si le processus `csh.pid` figure dans le résultat `ps -ef`, interrompez ce processus à l'aide de la commande suivante.

```
$ kill -9 csh.pid
```

f. Supprimez le fichier `nom_grappe.lock`.

```
$ rm -f /var/tmp/nom_grappe.lock
```

g. Quittez le PSS.

4. Redémarrez la grappe.

```
$ scadmin startcluster
```

Arrêt du moniteur d'appartenance pendant l'exécution d'instances RDBMS

Les instances de serveur de base de données ne peuvent s'exécuter sur un noeud que si vous avez appelé l'option `startnode` et que le noeud a été intégré correctement à la grappe. Toutes les instances de base de données doivent être arrêtées avant d'appeler l'option `stopnode`.

Remarque - Si vous utilisez Oracle7 Parallel Server, Oracle8 Parallel Server ou Informix XPS, consultez la documentation du produit correspondant pour connaître les détails de la procédure d'arrêt.

Si vous utilisez la commande `stopnode` pendant que l'instance Oracle7 ou Oracle8 est en cours d'exécution sur le noeud, `stopnode` bloque, et le message suivant apparaît sur la console :

```
ID[vxclust] : arrêt : attente de la fin de l'exécution des applications
```

Il faut arrêter l'instance Oracle7 ou Oracle8 pour que la commande `stopnode` s'exécute correctement.

L'exécution de la commande `stopnode` pendant que l'instance Informix-Online XPS est exécutée sur le noeud provoque le blocage de la base de données et la rend inutilisable.

Commutation des hôtes logiques

La commande `haswitch(1M)` permet de commuter les hôtes logiques spécifiés (ainsi que les groupes de disques, services de données et les adresses IP logiques connexes) vers le noeud spécifié par l'hôte cible. Par exemple, la commande suivante commute les hôtes logiques `hahost1` et `hahost2` vers `phys-hahost1`, qui en devient le maître.

```
# haswitch # haswitch hahost_phys1 hahost_phys2
```

Si l'hôte logique possède plus d'un service de données configuré, vous ne pouvez commuter qu'un seul service ou sous-ensemble de services. Vous ne pouvez que commuter *tous* les services de données vers l'hôte logique.



Attention - S'il se produit une relève ou une commutation quand le système de fichiers de l'hôte logique est occupé, ce dernier n'est pris en relève que partiellement ; certains des disques du groupe demeurent sur l'hôte physique cible initial. Ne tentez pas d'effectuer une commutation si le système de fichiers d'un hôte logique est occupé. En outre, n'accédez localement au système de fichiers d'aucun hôte, car le verrouillage de fichiers ne fonctionne pas correctement s'il y a à la fois verrouillage NFS et verrouillage local.

Remarque - L'hôte cible et le maître actuel de l'hôte logique *doivent* appartenir à la grappe pour que la commande s'exécute correctement. Sinon, elle échoue.

Désactivation de la commutation automatique

Dans le cas des grappes de services de données HD, vous pouvez configurer la commutation automatique pour l'éventualité suivante : un noeud échoue, les hôtes logiques dont il est le maître sont commutés vers un autre noeud, et le noeud défectueux est ramené ensuite dans la grappe. Le maître par défaut de ces hôtes logiques en reprend automatiquement possession, à moins que vous ne les ayez configurés pour qu'ils demeurent sous la maîtrise de l'hôte vers lequel ils ont été commutés.

Pour empêcher la commutation automatique d'un hôte logique vers son maître par défaut, utilisez l'option `-m` de la commande `scconf(1M)`. Pour de plus amples renseignements, consultez la page de manuel `scconf(1M)`.

Remarque - Pour désactiver la commutation automatique d'un hôte logique, il suffit d'exécuter la commande `scconf(1M)` sur un seul noeud actif appartenant à la grappe.

```
# scconf nom_grappe -L hôte_logique -n noeud1,noeud2 -g dg1 -i qe0,qe0,logaddr1 -m
```

Activation du mode maintenance des hôtes logiques

Le mode maintenance s'avère utile avec certaines tâches administratives au niveau des systèmes de fichiers et des groupes de disques. Pour activer le mode maintenance des groupes de disques d'un hôte logique, utilisez l'option `-m` de la commande `haswitch(1M)`.

Remarque - Contrairement aux autres types de propriété d'un hôte logique, le mode maintenance demeure activé après la réinitialisation des noeuds.

Dans l'exemple suivant, la commande active le mode maintenance de l'hôte logique `hahost1`.

```
phys-hahost2# haswitch -m hahost1
```

Cette commande interrompt les services de données associés à `hahost1` sur le noeud Sun Cluster qui est le propriétaire actuel du groupe de disques, et arrête également les programmes de surveillance des défaillances associés à `hahost1` sur tous les noeuds de Sun Cluster. La commande exécute le démontage (`umount (1M)`) de tous les systèmes de fichiers de Sun Cluster présents dans l'hôte logique. La propriété du groupe de disques associée à cet hôte est libérée.

Il est possible d'exécuter cette commande sur n'importe quel hôte, peu importe le propriétaire actuel de l'hôte logique et du groupe de disques.

Pour désactiver le mode maintenance d'un hôte logique, effectuez une commutation spécifiant l'hôte physique qui deviendra propriétaire du groupe de disques. Dans l'exemple ci-dessous, la commande désactive le mode maintenance de `hahost1` :

```
phys-hahost1# haswitch phys-hahost1 hahost1
```

Récupération à partir des partitions de grappe

Les tentatives des sous-ensembles de membres de grappe de demeurer actifs dans cette grappe peuvent provoquer des défaillances multiples (y compris le partitionnement du réseau). Normalement, ces sous-ensembles ont perdu, en tout ou en partie, leur capacité de communiquer entre eux. Dans ce cas, le logiciel tente de réduire le nombre de grappes valides à une seule. Pour y parvenir, il provoque l'abandon d'une partie ou de l'ensemble des noeuds. Voyons sur quels critères le logiciel fonde ses décisions à cet égard.

Le critère de quorum est un sous-ensemble comportant au moins la moitié des membres de l'ensemble des noeuds de grappe *original* (et non les seuls noeuds configurés). Si le sous-ensemble n'atteint pas le critère de quorum, les noeuds de ce sous-ensemble provoquent eux-mêmes leur abandon, et un message d'erreur

reconfig.4014 apparaît. La présence d'une partition au niveau du réseau ou d'une défaillance simultanée de plus de la moitié des noeuds de la grappe peuvent être à l'origine du non-respect du critère de quorum.

Remarque - Les grappes valides ne contiennent que des noeuds capables de communiquer entre eux sur des réseaux privés.

Prenons l'exemple d'une grappe à quatre noeuds qui se partitionne en deux sous-ensembles : on retrouve un seul noeud dans le premier sous-ensemble, alors que le second en comporte trois. Les deux sous-ensembles tentent d'atteindre le quorum requis. Comme le premier sous-ensemble ne possède qu'un seul noeud (sur les quatre d'origine), il ne respecte pas le critère de quorum. Par conséquent, le noeud du premier sous-ensemble s'arrête. Quant au second noeud, il possède trois des quatre noeuds originaux. Le quorum est atteint, et ce sous-ensemble demeure actif.

Prenons un autre exemple, celui d'une grappe à deux noeuds avec périphérique de quorum. Si une telle configuration comporte une partition, alors le critère de quorum est respecté avec la présence d'un noeud et du périphérique de quorum, et la grappe demeure active.

Partitions de double contrôle (VxVM seulement)

Une partition de double contrôle survient lorsqu'un sous-ensemble contient exactement la moitié des membres de la grappe. (Il n'y a pas de partition de double contrôle dans le cas d'une grappe à deux noeuds avec périphérique de quorum.) Au cours de la première installation de Sun Cluster, vous deviez décider du type de récupération privilégiée en cas de partition de double contrôle. Vous deviez choisir entre Demander et Sélectionner. Si vous avez opté pour Demander, le système vous demande de sélectionner les noeuds qui doivent demeurer actifs lorsque se produit une partition de double contrôle. Avec l'option `select`, le système sélectionne automatiquement les membres de la grappe qui demeurent actifs.

Si vous aviez choisi la politique de sélection automatique pour traiter les partitions de double contrôle, vous deviez choisir à nouveau entre les options ID de noeud le plus bas et ID de noeud le plus élevé. Si vous avez sélectionné l'option ID de noeud le plus bas, le sous-ensemble contenant le noeud dont l'ID est le plus bas devient la nouvelle grappe. Si vous avez sélectionné l'option ID de noeud le plus élevé, le sous-ensemble contenant le noeud dont l'ID est le plus élevé devient la nouvelle grappe. Pour de plus amples renseignements, consultez la section portant sur les procédures d'installation du *Sun Cluster 2.2 Software Installation Guide*.

Quelle que soit l'option choisie, vous devez arrêter manuellement les noeuds dans tous les autres sous-ensembles.

Si vous n'avez pas sélectionné une politique de sélection automatique ou si le système vous demande des précisions lorsque survient la partition, le message d'erreur suivant apparaît :

```
SUNWcluster.clustd.reconf.3010
``*** EXECUTER LA COMMANDE ABORTPARTITION OU CONTINUEPARTITION *** Grappe proposée :
xxx Noeuds inatteignables : yyy''
```

De plus, un message semblable à celui-ci apparaît toutes les dix secondes sur la console :

```
*** EXECUTER LA COMMANDE ISSUE ABORTPARTITION OU CONTINUEPARTITION ***
Si les noeuds inatteignables se sont formés en grappe, exécutez ABORTPARTITION.
(scadmin abortpartition <noeud_local> <nom_grappe>)
Vous pouvez autoriser la formation de la grappe proposée avec la commande CONTINUEPARTITION.
(scadmin continuepartition <noeud_local> <nom_grappe>)
Partition de grappe proposée : 0 Noeuds inatteignables : 1
```

Si vous n'avez pas choisi une sélection automatique, effectuez la procédure suivante pour sélectionner une nouvelle grappe.

Remarque - Pour redémarrer la grappe après une défaillance de double contrôle, vous devez attendre que le noeud arrêté soit complètement réactivé (ce délai provient de la reconfiguration ou de la réinitialisation du noeud) avant de le ramener dans la grappe avec la commande `scadmin startnode`.

▼ Comment sélectionner une nouvelle grappe

1. **Décidez du sous-ensemble qui formera la nouvelle grappe. Exécutez la commande suivante sur un noeud du sous-ensemble qui doit être abandonné.**

```
# scadmin abortpartition
```

Lorsque vous exécutez la commande `abortpartition` sur un noeud, le moniteur d'appartenance à une grappe (MAG) reproduit cette commande sur l'ensemble des noeuds de la partition concernée. En conséquence, *tous* les noeuds de la partition recevant la commande sont abandonnés. Au besoin, procédez à l'abandon manuel des noeuds que le MAG n'a pas réussi à contacter. Pour ce faire, exécutez la commande `scadmin abortpartition` sur les noeuds toujours actifs.

2. **Exécutez la commande suivante sur un noeud du sous-ensemble qui doit demeurer actif :**

```
# scadmin continuepartition
```

Remarque - Si la nouvelle grappe subit elle aussi une défaillance, un processus de reconfiguration supplémentaire s'enclenche. En tous temps, il n'y a qu'une seule grappe active.

Maintenance du système de fichiers / var

L'enregistrement des messages d'erreur des logiciels Solaris et Sun Cluster s'effectue dans le fichier `/var/adm/messages` et il existe un risque de saturation du système `/var`. La saturation du système de fichiers `/var` pendant l'exécution du noeud ne change rien à l'état de ce noeud. Par contre, il se peut que vous ne puissiez plus vous y connecter. Si le noeud tombe en panne, Sun Cluster ne pourra pas démarrer, et aucune connexion ne sera possible. Dans ce cas, vous devez réinitialiser le système en mode mono-utilisateur (commande `boot -s`).

Si le noeud signale que le système de fichier `/var` est saturé et qu'il continue d'exécuter les services Sun Cluster, effectuez les étapes de la procédure ci-dessous.

▼ Comment réparer un système de fichiers `/var` saturé

Dans cet exemple, `phys-hahost1` comporte un système de fichier `/var` saturé.

1. Effectuez une commutation.

Supprimez tous les hôtes logiques du noeud d'où provient le problème.

```
phys-hahost2# haswitch phys-hahost2 hahost1 hahost2
```

2. Supprimez l'appartenance du noeud à la grappe.

S'il y a une connexion active avec `phys-hahost1`, exécutez la commande suivante :

```
phys-hahost1 scadmin stopnode
```

S'il n'y a pas de connexion active avec `phys-hahost1`, arrêtez le noeud.

3. Réinitialisez le noeud en mode mono-utilisateur.

```
(0) ok boot -s
INIT : MODE MONO-UTILISATEUR
Appuyer sur Ctrl-D pour démarrer normalement
(ou donner le mot de passe de root (superutilisateur) pour
la maintenance système) :
mot_passe_superutilisateur
Activation du mode maintenance système
Sun Microsystems Inc. SunOS 5.6 générique août 1997
```

4. Procédez comme à l'habitude pour effacer le contenu du système de fichiers saturé.

5. Une fois le système de fichiers vidé, passez en mode multi-utilisateurs.

```
# exit
```

6. Exécutez la commande `scadmin startnode` pour réintégrer le noeud à la configuration.

```
# scadmin startnode
```

Administration du temps dans les configurations de Sun Cluster

Nous vous suggérons d'utiliser le protocole NTP (Network Time Protocol) fourni avec l'environnement d'exploitation Solaris pour préserver la synchronisation temporelle entre les noeuds de grappe.



Attention - Un administrateur ne peut régler l'heure des noeuds d'une configuration Sun Cluster. Il ne faut jamais utiliser les commandes `date(1)`, `rdate(1M)` ou `xntpdate(1M)` pour régler l'heure.

Dans l'environnement Sun Cluster, les noeuds de grappe peuvent fonctionner en tant que clients NTP. Pour utiliser le protocole NTP, un serveur NTP doit être installé et configuré hors de la grappe ; il est impossible de configurer les noeuds de grappe en tant que serveurs NTP. Consultez la page de manuel `xntpd(1M)` pour obtenir des informations sur les clients et serveurs NTP.

Si vous utilisez des noeuds de grappe en tant que clients NTP, assurez-vous qu'il n'y a pas d'entrées `crontab(1)` appelant la commande `ntpdate(1M)`. Il est plus prudent d'exécuter la commande `xntpd(1M)` sur les clients. De cette façon, on conserve la synchronisation des horloges sans compenser les écarts par des corrections importantes.

Remplacement d'un noeud défectueux

Effectuez les étapes ci-dessous lorsqu'un noeud éprouve une panne de matériel et qu'il doit être remplacé.

Remarque - Dans cette procédure, on suppose que le disque racine du noeud défectueux est toujours fonctionnel et qu'il peut encore servir. Si le disque racine n'est pas mis en miroir, communiquez avec un représentant local du service à la clientèle de Sun Enterprise ou avec le fournisseur de service agréé de votre région.

▼ Comment remplacer un noeud défectueux

Si le noeud défectueux n'est pas fonctionnel, passez à Etape 5, page 100.

1. **Si vous utilisez une configuration de base de données parallèle, arrêtez la base de données.**

Remarque - Consultez la documentation relative à vos services de données. La commande `scadmin stopnode` ferme automatiquement toutes les applications HD.

2. **Ouvrez la fenêtre de terminal avec la console de grappe.**

3. En tant que root (superutilisateur), exécutez la commande ci-dessous dans la fenêtre de terminal.

Cette commande supprime le noeud de la grappe, ferme le logiciel Sun Cluster et désactive le gestionnaire de volumes sur ce noeud.

```
# scadmin stopnode
```

4. Arrêtez le système d'exploitation du noeud.

Consultez le Guide d'administration Solaris à ce sujet.

5. Mettez le noeud hors tension.

Pour de plus amples renseignements, consultez le manuel d'entretien de l'équipement concerné.



Attention - Ne touchez pas aux câbles du noeud défectueux pour l'instant.

6. Retirez le disque d'initialisation du noeud défectueux.

Pour de plus amples renseignements, consultez le manuel d'entretien de l'équipement concerné.

7. Placez le disque d'initialisation au même emplacement dans le nouveau noeud.

L'adresse d'accès du disque racine doit demeurer la même. Pour de plus amples renseignements, consultez le manuel d'entretien de l'équipement concerné.

Remarque - Assurez-vous que l'adresse IP du nouveau noeud est la même que celle du système défectueux. Il peut s'avérer nécessaire de modifier les serveurs d'initialisation ou les serveurs `arp` pour rétablir la correspondance entre l'adresse IP et la nouvelle adresse Ethernet. Pour de plus amples renseignements à ce sujet, consultez le *Guide d'installation et de configuration NIS+ et DNS*.

8. Mettez le nouveau noeud sous tension.

Pour de plus amples renseignements, consultez le manuel d'entretien de l'équipement concerné.

9. Si le noeud s'initialise automatiquement, arrêtez le système d'exploitation et accédez au moniteur de mémoire morte programmable (PROM) OpenBoot.

Pour de plus amples renseignements, consultez la page de manuel `shutdown(1M)`.

10. Assurez-vous que chaque ID des initiateurs SCSI est correctement configurée.

Reportez-vous au Chapitre 4 du *Sun Cluster 2.2 Hardware Site Preparation, Planning, and Installation Guide* dans laquelle vous trouverez la procédure détaillée de configuration des ID des initiateurs SCSI.

11. Mettez le nouveau noeud hors tension.

Pour de plus amples renseignements, consultez le manuel d'entretien de l'équipement concerné.

12. Dans le noeud survivant qui partage les disques multihôtes avec le noeud défectueux, détachez tous les disques dans une unité d'expansion de disque attachée au noeud défectueux.

Pour de plus amples renseignements, consultez le manuel d'entretien de l'équipement concerné.

13. Mettez l'unité d'expansion de disque hors tension.

Pour de plus amples renseignements, consultez le manuel d'entretien de l'équipement concerné.

Remarque - Pendant que vous remplacez le noeud défectueux, des messages semblables à ceux-ci peuvent apparaître sur la console du système. Ne tenez pas compte de ces messages, puisqu'ils ne signalent pas nécessairement un problème.

```
Nov  3 17:44:00 updb10a unix : AVERTISSEMENT : /sbus@1f,0/SUNW,fas@0,8800000/
sd@2,0 (sd17) :
3 nov 17:44:00 updb10a unix : Echec du transport SCSI :
motif : 'incomplet' : réessayer \ commande
3 nov 17:44:03 updb10a unix : AVERTISSEMENT : /sbus@1f,0/SUNW,fas@0,8800000/
sd@2,0 (sd17) :
3 nov 17:44:03 updb10a unix : le disque ne réagit pas à la sélection
```

14. Débranchez le câble SCSI du noeud défectueux et connectez-le sur la fente correspondante du nouveau noeud.

Pour de plus amples renseignements, consultez le manuel d'entretien de l'équipement concerné.

15. Mettez l'unité d'expansion de disque sous tension.

Pour de plus amples renseignements, consultez le manuel d'entretien de l'équipement concerné.

16. Attachez de nouveau tous les disques détachés à Etape 12, page 101.

Pour de plus amples renseignements, consultez le manuel d'entretien de l'équipement concerné.

17. Attendez la fin de la récupération sur tous les volumes de l'unité d'expansion de disque avant de détacher l'unité correspondante.

Votre gestionnaire de volumes vous permet de déterminer le moment où s'est produite la récupération des volumes.

18. Répétez les Etape 12, page 101 à Etape 17, page 102 pour chacune des unités d'expansion de disque restantes.

19. Mettez le nouveau noeud (le noeud remplacé) sous tension.

Pour de plus amples renseignements, consultez le manuel d'entretien de l'équipement concerné.

20. Réinitialisez le noeud et attendez que le système redevienne actif.

```
<#0> boot
```

21. Déterminez l'adresse Ethernet du nouveau noeud (le noeud remplacé).

```
# /usr/sbin/arp nodename
```

22. Déterminez l'ID du nouveau noeud.

En procédant par élimination, déterminez le noeud ne faisant pas partie de la grappe. Les ID de noeuds sont énumérées par ordre croissant à partir du noeud 0.

```
# get_node_status
sc : inclus dans la grappe en cours d'exécution
ID de noeud : 0
appartenance : 0
interconnexion0 : inconnue
interconnexion1 : inconnue
type_gv : vxvm
vm_on_node : maître
gv : actif
b_données : inactive
```

23. Signalez la nouvelle adresse Ethernet (du nouveau noeud) au système de grappes en exécutant la commande suivante sur tous les noeuds de la grappe :

```
# scconf nom_grappe -N id_noeud adresse_ethernet_hôte
```

Toujours selon l'exemple de Etape 22, page 102, l'ID du noeud est 1 :

```
# scconf nom_grappe -N 1 adresse_ethernet_hôte
```

24. Démarrez le nouveau noeud.

```
# scadmin startnode
```

25. Si vous utilisez une configuration de base de données parallèle, redémarrez la base de données.

Remarque - Consultez la documentation relative à vos services de données. Les commandes `scadmin startcluster` et `scadmin startnode` démarrent automatiquement toutes les applications HA.

Remplacement d'un concentrateur de terminaux défectueux

Il n'est pas nécessaire que le concentrateur de terminaux soit fonctionnel pour que la grappe demeure active. Une défaillance du concentrateur n'a aucune incidence sur le fonctionnement de la grappe.

Vous pouvez remplacer un concentrateur de terminaux défectueux sans influencer le fonctionnement de la grappe. Si le nom, l'adresse IP et le mot de passe du concentrateur de terminaux de remplacement sont identiques à ceux de l'original, il n'est pas nécessaire d'exécuter les commandes sur les noeuds. Il suffit de brancher le nouveau concentrateur de terminaux pour qu'il fonctionne normalement.

Par contre, si le nom, l'adresse IP ou le mot de passe du nouveau concentrateur de terminaux ne sont pas les mêmes, exécutez la commande `scconf(1M)` comme le décrit la "Modification des informations TC/PSS", page 75 pour modifier ces

données dans la base de données de grappe. Le fonctionnement de la grappe en cours d'exécution n'en sera pas affecté.

Administration de la base de données de configuration de grappe

La commande `ccdadm(1M)` permet de gérer la base de données de configuration de grappe (BCG). Pour de plus amples renseignements, consultez la page de manuel `ccdadm(1M)`.

Remarque - En tant que `root` (superutilisateur), vous pouvez exécuter la commande `ccdadm(1M)` à partir de n'importe quel noeud actif. Cette commande met à jour tous les noeuds de la grappe.

Il est conseillé de contrôler point par point la BCG en ajoutant l'option `-c` (points de contrôle) à la commande `ccdadm(1M)` après chaque mise à jour de la configuration de grappe. La structure de Sun Cluster fait un usage intensif de la BCG pour stocker les données de configuration liées aux hôtes logiques et aux services de données HD. La BCG sert également à stocker les données de configuration de l'adaptateur réseau utilisées pour la gestion de réseau privé (GRP). Dès que la configuration HD ou GRP de la grappe est modifiée, nous vous suggérons fortement d'archiver un instantané valide de la BCG à jour avec l'option `-c` à titre d'assurance contre tout problème pouvant résulter d'une défaillance ultérieure. Il n'y a pas de raison de se soustraire à une telle pratique sûre. Après tout, même les administrateurs de bases de données ou de systèmes doivent procéder régulièrement à une sauvegarde des données pour se prémunir contre les crises majeures issues de circonstances imprévisibles.

▼ Comment vérifier la cohérence générale de la BCG

1. Utilisez l'option `-v` chaque fois que vous suspectez un problème avec la BCG dynamique.

Cette option permet de comparer l'enregistrement de cohérence de chaque exemplaire de la BCG de tous les noeuds de la grappe pour que vous puissiez vérifier que la base de données demeure cohérente dans l'ensemble des noeuds de la grappe. La fonction d'interrogation de la BCG est désactivée tout au long de la procédure de vérification.

```
# ccdadm nom_grappe -v
```

▼ Comment sauvegarder la BCG

1. Exécutez la commande précédente avec l'option `-c` une fois par semaine ou lorsque vous faites une sauvegarde de la BCG.

Cette option crée une copie de sauvegarde de la BCG dynamique. Cette copie peut servir par la suite à récupérer la BCG dynamique avec l'option `-r`. Voir la "Comment restaurer la BCG", page 105 pour de plus amples renseignements à ce sujet.

Remarque - Lorsque vous sauvegardez la BCG, activez le mode de maintenance de tous les hôtes logiques avant d'exécuter la commande `ccdadm -c`. Pour récupérer la BCG, les hôtes logiques doivent être en mode maintenance. Par conséquent, la présence d'un fichier de sauvegarde identique à l'état restauré de la BCG évite d'exposer inutilement le système aux erreurs ou aux défaillances.

```
# ccdadm nom_grappe -c nom_fichier_points_contrôle
```

Dans cette commande, *nom_fichier_points_contrôle* est le nom du fichier de sauvegarde.

▼ Comment restaurer la BCG

Exécutez la commande `ccdadm(1M)` suivie de l'option `-r` chaque fois que le contenu de la BCG est altéré. Cette option rejette la copie actuelle de la BCG dynamique et restaure cette dernière avec le contenu du fichier de récupération que vous spécifiez. Exécutez cette commande pour initialiser ou restaurer la BCG dynamique lorsque l'algorithme de reconfiguration `ccdd(1M)` est incapable de choisir une copie valide de la BCG au moment du redémarrage de la grappe. La BCG est alors identifiée comme étant valide.

1. Désactivez au besoin le quorum.

Voir la section "Comment activer et désactiver le quorum BCG", page 106 pour de plus amples renseignements à ce sujet.

```
# ccdadm nom_grappe -q off
```

2. Activez le mode maintenance des hôtes logiques.

```
# haswitch -m hôtes_logiques
```

3. Restaurez la BCG.

Dans la commande suivante, *nom_fichier_récupération* est le nom du fichier que vous récupérez.

```
# ccdadm nom_grappe -r nom_fichier_récupération
```

4. Si nécessaire, réactivez le quorum BCG.

```
# ccdadm nom_grappe -q on
```

5. Remettez les hôtes logiques en ligne.

Par exemple :

```
# haswitch hôte-physique1 hôte_logique1  
# haswitch hôte-physique2 hôte_logique2
```

▼ Comment activer et désactiver le quorum BCG

1. Habituellement, le logiciel de grappes requiert un quorum avant de mettre la BCG à jour. L'option `-c` vous permet de passer outre cette restriction et de mettre la BCG à jour avec n'importe quel nombre de noeuds.

Utilisez cette option pour activer ou désactiver le quorum au moment de mettre à jour ou de restaurer la BCG dynamique. L'indicateur `quorum` est un commutateur à deux valeurs : `actif` (activation du quorum) et `inactif` (désactivation du quorum). Par défaut, le quorum est activé.

Par exemple, si la grappe comporte trois noeuds physiques, vous avez besoin d'au moins deux noeuds pour mettre à jour la BCG. S'il y a eu une défaillance au niveau du matériel, vous ne pouvez réactiver qu'un seul noeud. Le logiciel de grappes ne vous permet donc pas de mettre la BCG à jour. Cependant, si vous exécutez la commande `ccdadm -q`, vous pouvez désactiver le contrôle logiciel et mettre la BCG à jour.

```
# ccdadm nom_grappe -q on/off
```

▼ Comment purifier la BCG

1. L'option `-p` vous permet de purifier le fichier de la BCG (c'est-à-dire de vérifier son contenu et la syntaxe utilisée). Utilisez cette option s'il y a des erreurs syntaxiques dans le fichier de la base de données de configuration de grappe.

```
# ccdadm -p nom_fichier_BCG
```

L'option `-p` signale toute erreur de format dans le fichier ciblé et enregistre une version corrigée dans le fichier `nom_fichier:pure`. Vous pouvez alors récupérer ce fichier "purifié" en tant que nouvelle BCG. Voir "Comment restaurer la BCG", page 105 pour de plus amples renseignements à ce sujet.

Dépannage de la BCG

Le système consigne les erreurs de la BCG dans le fichier `/var/opt/SUNWcluster/ccd/ccd.log`. Les messages d'erreur critique sont également transmis à la console de grappe. Il est rare que le système subisse une panne majeure, mais dans ce cas, le logiciel crée un fichier noyau dans `/var/opt/SUNWcluster/ccd`.

Voici un exemple de fichier `ccd.log`.

```
lpc204# cat ccd.log
16 avr 14:54:05 lpc204 ID[SUNWcluster.ccd.cdd.1005] : (info) démarrage
de la transition 'START' avec délai de 10000
16 avr 14:54:05 lpc204 ID[SUNWcluster.ccd.cdd.1005] : (info) transition 'START'
terminée avec état 0
16 avr 14:54:06 lpc204 ID[SUNWcluster.ccd.cdd.1005] : (info) démarrage
de la transition 'STEP1' avec délai de 20000
16 avr 14:54:06 lpc204 ID[SUNWcluster.ccd.cdd.1000] : (info)
ID_noeud = 0 Actif = 0 No_généré = 0 Date =
14 fév 10h30m00 1997 Récupération = 4
16 avr 14:54:06 lpc204 ID[SUNWcluster.ccd.cdd.1002] : (info) démarrage de la
reconfiguration de la BCG choisie à partir de ID_noeud = 0
16 avr 14:54:06 lpc204 ID[SUNWcluster.ccd.cdd.1004] :
(info) la BCG d'initialisation est cohérente
16 avr 14:54:06 lpc204 ID[SUNWcluster.ccd.cdd.1001] : (info) Activation
du noeud en tant que grappe à un noeud
après exécution de scadmin startcluster ; test de quorum BCG omis
16 avr 14:54:06 lpc204 ID[SUNWcluster.ccd.cdd.1005] :
(info) transition 'STEP1' terminée avec état 0
```

Le tableau ci-dessous dresse la liste des messages d'erreur courants et fournit des solutions aux différents problèmes. Le document *Sun Cluster 2.2 Error Messages Manual* contient la liste exhaustive de ces messages d'erreur.

TABLEAU 4-1 Messages d'erreur courants de la base de données de configuration de grappe

Numéro de message	Explication	Action
4200	Impossible d'ouvrir le fichier	Restaurez la BCG avec la commande <code>ccdadm -r</code> .
4302	Fichier non trouvé	Restaurez la BCG avec la commande <code>ccdadm -r</code> .
4307	BCG d'initialisation incohérente	Supprimez et réinstallez le logiciel Sun Cluster.
4402	Erreur d'inscription du serveur d'appel de procédure à distance (RPC)	Vérifiez l'état de votre réseau public (problème de réseautage).
4403	Echec de création de client RPC	Vérifiez l'état de votre réseau public (problème de réseautage).
5000	Erreur d'exécution système	Le script de synchronisation comporte une erreur. Vérifiez les permissions relatives au script.
5300	BCG non valide, doit être récupérée	Restaurez la BCG avec la commande <code>ccdadm -r</code> .
5304	Erreur d'exécution de la commande de prise d'instantané	Le script de synchronisation exécuté comporte des arguments erronés. Vérifiez la syntaxe du script.
5306	Pointeur de grappe nul	Ce message signale que la grappe n'existe pas (<code>ccdadm grappe</code>). Vérifiez l'exactitude du nom de grappe.

Réservation de disques partagés (VxVM)

La liste des disques mise à jour par le gestionnaire de volumes contient les différents périphériques assurant la protection contre les défaillances. Si un système ne comporte pas de groupe de disques, il n'y a pas de périphériques de protection contre les défaillances (il n'y a effectivement pas de données à protéger). Cependant, lorsque l'on importe de nouveaux groupes de disques alors qu'un ou plusieurs noeuds ne font pas partie de la grappe, il faut signaler à la grappe qu'un autre ensemble de périphériques a besoin de protection contre les défaillances.

▼ Comment réserver des périphériques partagés (VxVM)

1. **Lorsque l'on importe de nouveaux groupes de disques partagés alors qu'un ou plusieurs noeuds ne font pas partie de la grappe, il faut signaler à la grappe qu'un autre ensemble de périphériques a besoin de protection contre les défaillances. Pour ce faire, exécutez la commande `scadmin resdisk` à partir d'un noeud capable d'accéder au(x) nouveau(x) groupe(s) de disques.**

```
# scadmin resdisks
```

Cette commande réserve tous les périphériques connectés à un noeud, si aucun autre noeud n'appartient à la grappe (le noeud doit être capable de se connecter au même ensemble de périphériques). Autrement dit, les réservations ne sont influencées que si un et un seul noeud, parmi tous les noeuds connectés directement aux périphériques, appartient à la grappe. Si cette condition n'est pas respectée, la commande `scadmin resdisks` ne produit aucun résultat. Cette commande échoue également lorsque la reconfiguration de grappe est en cours. L'arrêt de ce noeud unique ou l'intégration d'autres noeuds connectés directement aux périphériques partagés provoque la libération automatique des réservations de périphériques partagés.

Remarque - L'exécution de la commande `scadmin resdisks` est inutile si l'on importe les groupes de disques partagés lorsque tous les noeuds sont présents dans la grappe. Si tous les noeuds appartiennent à la grappe, le recours aux réservations et à la protection contre les défaillances est inutile.

Par contre, l'exportation d'un groupe de disques partagés ne libère pas les réservations des périphériques partagés du groupe de disques exportés. Il n'y a pas de libération de ces réservations tant que le noeud d'où elles proviennent

n'est pas arrêté ou que l'autre noeud, avec qui il partage les périphériques, n'a pas intégré la grappe.

Pour activer et utiliser immédiatement l'ensemble de disques appartenant au groupe de disques exportés, exécutez successivement les deux commandes ci-dessous sur tous les noeuds de la grappe, après avoir exporté le groupe de disques partagés :

```
# scadmin reldisks  
# scadmin resdisks
```

La première commande libère les réservations sur tous les périphériques partagés. La seconde rétablit effectivement les réservations en fonction de l'ensemble de groupes de disques importé et exclut automatiquement l'ensemble de disques associé aux groupes de disques exportés.

Récupération après une panne de courant

Le présent chapitre décrit différents scénarios où survient une panne de courant et la marche à suivre pour rétablir le fonctionnement normal du système. La liste ci-dessous présente les sujets abordés dans le chapitre.

- “Récupération après une panne de courant généralisée”, page 111
- “Récupération après une panne de courant partielle”, page 112
- “Mise sous tension du système”, page 113

La gestion des configurations de Sun Cluster comprend le traitement de pannes comme la perte d'alimentation. Une panne de courant peut provoquer l'arrêt d'une configuration de Sun Cluster entière ou de certains de ses composants. Le comportement des noeuds de Sun Cluster varie en fonction du composant touché par la panne de courant. Les sections suivantes décrivent différents scénarios types et le comportement qui en résulte.

Récupération après une panne de courant généralisée

Dans le cas d'une configuration Sun Cluster alimentée par une seule source d'alimentation, une panne de courant affecte tous les noeuds Sun Cluster et leurs unités d'expansion de disques multihôtes. Lorsque tous les noeuds ne sont plus alimentés, la configuration toute entière échoue.

Dans le cas d'une panne généralisée, le fonctionnement des composants matériels de la grappe peut être rétabli de deux manières.

- Un noeud Sun Cluster se réinitialise avant le concentrateur de terminaux. Les erreurs signalées au moment de la réinitialisation du noeud sont stockées dans le fichier `/var/adm/messages` ou dans le journal d'erreurs identifié dans le fichier `/etc/syslog.conf`.
- Un noeud Sun Cluster se réinitialise avant l'unité d'expansion de disques multihôtes. Les disques associés ne seront pas accessibles. Au moins un noeud doit être réinitialisé après le rétablissement du fonctionnement de l'unité d'expansion de disques multihôtes. Une fois les noeuds réactivés, exécutez la commande `hasstat(1M)` et utilisez votre gestionnaire de volumes pour localiser toute erreur provoquée par la panne de courant.

Récupération après une panne de courant partielle

Si l'alimentation des noeuds de Sun Cluster et des unités d'expansion de disques multihôtes provient de sources distinctes, une panne de courant peut toucher un ou plusieurs composants. Divers scénarios peuvent se présenter. Voici les plus courants :

- Il y a rupture de l'alimentation d'un noeud Sun Cluster et seul le noeud tombe en panne.
- Il y a rupture de l'alimentation d'une unité d'expansion de disques multihôtes, et seule l'unité d'expansion tombe en panne.
- Il y a rupture de l'alimentation d'un noeud Sun Cluster et au moins une unité d'expansion de disques multihôtes tombe en panne.
- Il y a rupture de l'alimentation d'un noeud Sun Cluster et le noeud, au moins une des unités d'expansion de disques multihôtes et le concentrateur de terminaux tombent en panne.

Panne d'un seul noeud

Si les noeuds et les unités d'expansion de disques multihôtes possèdent leur propre alimentation et que la panne de courant ne touche qu'un seul noeud, les autres noeuds détectent la défaillance et amorcent une commutation.

Lorsque l'alimentation du noeud est rétablie, il s'initialise. Vous devez le réintégrer à la grappe avec la commande `scadmin startnode`. Effectuez ensuite une commutation manuelle avec la commande `haswitch(1M)`, ce qui rétablit la propriété par défaut de l'hôte logique.

Panne d'une unité d'expansion de disques multihôtes

S'il y a rupture de l'alimentation d'une des unités d'expansion de disques multihôtes, le gestionnaire de volumes détecte des erreurs sur les disques touchés par la panne et prend les mesures nécessaires pour leur attribuer un état d'erreur. Le moniteur de défaillances de Sun Cluster ne peut détecter cette panne à cause de la mise en miroir des disques. Aucune commutation ou relève ne se produit.

Lorsque l'alimentation de l'unité d'expansion de disques multihôtes est rétablie, exécutez la procédure documentée aux Chapitre 11 et Chapitre 12.

Panne d'un serveur et d'une unité d'expansion de disques multihôtes

S'il y a rupture de l'alimentation d'un des noeuds de Sun Cluster et d'une unité d'expansion de disques multihôtes, un noeud secondaire prend immédiatement la relève.

Lorsque le courant est rétabli, vous devez réinitialiser le noeud et le réintégrer à la configuration avec la commande `scadmin startnode`, après quoi vous pouvez commencer la surveillance des activités. Dans le cas d'une configuration avec commutation manuelle, exécutez la commande `haswitch(1M)` pour redonner manuellement la propriété de l'ensemble de disques au noeud ayant subi la panne de courant. Pour de plus amples renseignements, consultez la "Commutation des hôtes logiques", page 92.

Après avoir redonné la propriété de l'ensemble de disques au maître par défaut, vous devez remettre en service tous les disques multihôtes ayant signalés des erreurs. Les directives présentées aux chapitres concernant l'unité d'expansion de disque expliquent comment remettre les disques multihôtes en service.

Remarque - Il est possible que le noeud se réinitialise avant l'unité d'expansion de disques multihôtes. Les disques associés ne seront donc pas accessibles. Réinitialisez le noeud après la remise en service de l'unité d'expansion de disques multihôtes.

Mise sous tension du système

La mise sous tension des armoires, des noeuds et des disques d'initialisation du système varie en fonction du type d'armoire utilisé et du mode d'alimentation c.a. des noeuds.

Les tableaux de disques ne possédant pas leur propre source de courant alternatif sont alimentés lors de la mise sous tension de l'armoire du système.

Pour connaître les procédures de mise sous tension des Sun StorEdge MultiPack, consultez le *Sun StorEdge MultiPack Service Manual*.

L'alimentation en courant alternatif du concentrateur de terminaux provenant de l'armoire du système n'est établie que lorsque l'armoire est mise sous tension. Autrement, il faut mettre manuellement le concentrateur de terminaux sous tension.

Administration des interfaces réseau

Le présent chapitre fournit une description du service de gestion de réseau public (PNM) de Sun Cluster, ainsi que des directives permettant d'ajouter ou de remplacer des composants d'interface réseau. La liste ci-dessous présente les sujets abordés dans le chapitre.

- “Aperçu de la gestion de réseau public”, page 115
- “Configuration et administration de la gestion de réseau public”, page 119
- “Résolution des erreurs du service PNM”, page 126
- “Ajout et suppression d'interfaces réseau”, page 127
- “Administration de l'Agent de gestion de commutation”, page 133

Aperçu de la gestion de réseau public

Le service PNM de Sun Cluster a recours à la surveillance des défaillances et à la reprise pour empêcher la perte de disponibilité des noeuds provoquée par une seule panne d'adaptateur réseau ou de câble. Le service PNM effectue la surveillance des défaillances au niveau du noeud local ou dans la grappe toute entière et vérifie l'état des noeuds, des adaptateurs réseau, des câbles et du trafic sur le réseau. La fonction de reprise du service PNM utilise des ensembles d'adaptateurs réseau, appelés *groupes de sauvegarde*, pour assurer la redondance des connexions entre un noeud de grappe et le réseau public. La surveillance des défaillances et la reprise travaillent de concert pour garantir la disponibilité des services.

Vous devez activer le service PNM si votre système est configuré avec des services de données HD ; ces services de données dépendent de la surveillance des défaillances par le service PNM. Lorsqu'un service de données HD éprouve un problème de disponibilité, il interroge le service PNM par l'intermédiaire de la

structure de la grappe pour savoir si le problème est lié aux connexions du réseau public. Si tel est le cas, les services de données attendent que le service de gestion PNM ait résolu le problème. Si le réseau public n'est pas à l'origine du problème, les services de données font appel à leur propre mécanisme de reprise.

Le module de gestion PNM, `SUNWpnm`, est installé lors de la première installation du logiciel Sun Cluster. Les commandes associées au service PNM sont :

- `pnmset (1M)` – Configure le service PNM avant ou après la configuration de la grappe et vérifie la conformité de la configuration existante du service.
- `pnmstat (1M)` – Vérifie l'état du réseau et des adaptateurs.
- `pnmconf (1M)` – Affiche la configuration et l'état des interfaces réseau du service PNM.
- `pnmrtop (1M)` – Affiche le nom du groupe de sauvegarde ou le pseudonyme de l'adaptateur réseau (p. ex., `nafo1`) associé au nom réel de l'adaptateur réseau (p. ex., `hme2`) spécifié dans la commande.
- `pnmptor (1M)` – Affiche le nom réel de l'adaptateur réseau (p. ex., `hme2`) associé au pseudonyme d'adaptateur réseau ou au nom du groupe de sauvegarde (p. ex., `nafo1`) spécifié dans la commande.
- `pnmmd (1M)` – Démon PNM.

Pour de plus amples renseignements, consultez les pages de manuel pertinentes.

Surveillance des défaillances et reprise par le service PNM

Le service PNM surveille l'état du réseau public et des adaptateurs réseau associés à chaque noeud de la grappe et signale les cas suspects ou les erreurs. Lorsqu'il détecte l'absence de réponse d'un adaptateur réseau *principal* (celui par qui transitent les données en provenance ou à destination du noeud), le service PNM commute le service de réseau de ce noeud vers un adaptateur réseau en état de marche dans le groupe de sauvegarde. Il effectue ensuite certaines vérifications pour déterminer si l'adaptateur réseau ou le réseau lui-même est à l'origine de la défaillance.

Si l'adaptateur est défectueux, le service PNM envoie des messages d'erreur au fichier `syslog(3)` ; le gestionnaire de grappe détecte à son tour ces messages et les affiche par l'intermédiaire d'une IUG. Une fois réparé, l'adaptateur est testé puis réintégré dans le groupe de sauvegarde lors de la prochaine reconfiguration de la grappe. Si c'est le groupe d'adaptateurs réseau de sauvegarde tout entier qui tombe en panne, la structure Sun Cluster commande une reprise du noeud pour maintenir la disponibilité. Si une erreur se produit mais qu'elle dépasse le cadre de contrôle du service PNM, par exemple lorsqu'un sous-réseau entier tombe en panne, la reprise et la reconfiguration s'effectuent normalement.

Le service PNM fonctionne en deux modes de surveillance : gestion de grappe active et gestion de grappe inactive. Il est en mode de gestion de grappe active lorsque la

grappe est fonctionnelle. Dans ce cas, il utilise la base de données de configuration de grappe (BCG) pour surveiller l'état du réseau (pour de plus amples renseignements sur la BCG, consultez le chapitre d'aperçu du *Sun Cluster 2.2 Software Installation Guide*). Le service PNM utilise la BCG pour faire la distinction entre une panne de réseau public et une panne d'adaptateur réseau local. Voir Annexe B pour obtenir d'autres informations sur la reprise des hôtes logiques provoqué par une défaillance du réseau public.

Le service PNM passe en mode de gestion de grappe inactive lorsque la grappe n'est pas fonctionnelle. Dans ce mode, il ne peut se servir de la BCG et est donc incapable de faire la distinction entre une panne d'adaptateur réseau et une défaillance du réseau lui-même. En mode de gestion de grappe inactive, le service PNM ne détecte que les problèmes au niveau de la connexion au réseau local.

Le service PNM comporte une commande de surveillance, `pnmstat(1M)`, vous permettant de vérifier l'état du réseau public et des adaptateurs réseau. Pour de plus amples renseignements, consultez la page de manuel pertinente.

Groupes de sauvegarde

Les groupes de sauvegarde sont des ensembles d'adaptateurs réseau assurant la redondance des connexions entre un noeud de grappe unique et le réseau public. La commande `scinstall(1M)` permet de configurer ces groupes de sauvegarde lors de l'installation initiale ; par la suite, vous devez utiliser la commande `pnmset(1M)`. Le service PNM vous permet de configurer autant d'adaptateurs réseau redondants que vous le voulez dans un seul hôte.

Pour configurer les groupes de sauvegarde la première fois, exécutez la commande `pnmset(1M)` en tant que `root` (superutilisateur) avant de démarrer la grappe. La commande fonctionne comme un script interactif permettant de configurer les groupes de sauvegarde et d'en vérifier l'état. Elle sélectionne également un adaptateur réseau qui servira d'adaptateur principal ou actif. La commande `pnmset(1M)` attribue un nom aux groupes de sauvegarde `nafon`, où `n` est un nombre entier que vous assignez vous-même. Cette commande enregistre les informations sur les différents groupes de sauvegarde dans le fichier `/etc/pnmconfig`.

Pour modifier une configuration existante du service PNM sur un noeud de grappe, vous devez supprimer le noeud de la grappe et exécuter ensuite la commande `pnmset(1M)`. Le service PNM assure la surveillance des modifications apportées à l'appartenance au groupe de sauvegarde et les intègre à ce dernier de façon dynamique.

Remarque - Le fichier `/etc/pnmconfig` n'est pas supprimé même si on désinstalle le module `SUNWpnm`, par exemple, pour mettre un élément logiciel du système à niveau. Autrement dit, les informations sur l'appartenance au groupe de sauvegarde sont conservées pendant les mises à niveau des logiciels, et vous n'avez pas à exécuter de nouveau l'utilitaire `pnmset(1M)`, à moins que vous ne vouliez modifier l'appartenance au groupe de sauvegarde.

Mises à jour du fichier `nsswitch.conf`

Au moment de configurer le service PNM avec un adaptateur réseau de sauvegarde, le fichier `/etc/nsswitch.conf` doit comporter une des entrées suivantes relatives à l'entrée `masques_réseau`.

TABLEAU 6-1 Choix de services de nom pour le fichier `/etc/nsswitch.conf` File

Service de nom utilisé	Entrée <code>masques_réseau</code>
aucun	<code>masques_réseau : files</code>
nis	<code>masques_réseau : fichiers [INTROUVABLE=entrée] nis</code>
nisplus	<code>masques_réseau : fichiers [INTROUVABLE=entrée] nisplus</code>

Les paramètres précédents bloquent la consultation du paramètre `masques_réseau` dans une table de recherche NIS/NIS+. Ce blocage est essentiel si l'adaptateur défectueux est le réseau public principal et qu'il ne peut fournir les informations demandées. Si le réglage de l'entrée `masques_réseau` précédent est incorrect, la reprise vers l'adaptateur de sauvegarde échoue.



Attention - Les modifications apportées précédemment ont pour effet d'utiliser les fichiers locaux (`/etc/masques_réseau` et `/etc/groupes`) en tant que tables de recherche. Les services NIS/NIS+ ne peuvent servir que lorsque les fichiers locaux ne sont pas disponibles. Par conséquent, il faut constamment mettre à jour ces fichiers avec leurs versions NIS/NIS+. S'ils ne sont pas mis à jour, les noeuds de grappe ne peuvent accéder aux valeurs attendues qu'ils contiennent.

Configuration et administration de la gestion de réseau public

Cette section contient les procédures permettant de configurer le service de gestion de réseau public (PNM) et les groupes de sauvegarde.

▼ Comment configurer le service PNM

Voici les étapes générales permettant de configurer le service PNM :

- Configuration des composants matériels de noeud afin d'autoriser de multiples adaptateurs réseau par noeud et par sous-réseau.
- Installation des modules Sun Cluster et PNM (s'il y a lieu).
- Démarrage de la grappe.
- Vérification des interfaces réseau par défaut.
- Etablissement des groupes de sauvegarde du service PNM à l'aide de la commande `pnmset(1M)`.
- Vérification de la configuration du service PNM

Les étapes suivantes permettent de configurer le service PNM.

1. **Configurez les composants matériels de noeud de sorte qu'il y ait de multiples adaptateurs réseau sur un noeud unique et qu'ils utilisent le même sous-réseau.**
Consultez la documentation de Sun Cluster sur les composants matériels de noeud pour configurer les adaptateurs réseau.
2. **Si vous n'avez pas déjà installé les modules de noeud Sun Cluster, faites-le avec la commande `scinstall(1M)`.**
La commande `scinstall(1M)` fonctionne de manière interactive pour installer les modules sélectionnés. Le module PNM, `SUNWpnm`, fait partie de l'ensemble de modules de noeud. Le *Sun Cluster 2.2 Software Installation Guide* contient la procédure détaillée pour l'installation de la grappe.
3. **Si ce n'est déjà fait, inscrivez l'interface réseau par défaut dans chaque noeud.**
Vous devez inscrire une interface réseau par défaut par noeud dans la base de données d'interfaces associée à chaque noeud et vérifier que l'interface est étalonnée et qu'elle fonctionne correctement.
 - a. **Créez une base de données d'interfaces dans chaque noeud et inscrivez les interfaces principales de réseau public.**

Créez un fichier servant de base de données d'interface dans le répertoire / etc de chaque noeud. Nommez le fichier nom_hôte. *interface*, où *interface* désigne le type d'interface, par exemple qfe, hme, etc. Ajoutez ensuite une ligne contenant le nom d'hôte de ce noeud. Par exemple, dans le noeud phys-hahost1 où se trouve l'interface par défaut qfe-1, créez un fichier / etc/phys-hahost1.qfe1 contenant la ligne suivante :

```
phys-hahost1
```

b. Dans le fichier /etc/hosts de chaque noeud, associez une adresse IP au nom de l'interface principale de réseau public.

Dans l'exemple suivant, le nom d'hôte physique principal est phys-hahost1 :

```
129.146.75.200 phys-hahost1-qfe1
```

Si votre système utilise un mécanisme de nomenclature autre que / etc/ hosts, reportez-vous à la section appropriée du *Guide d'administration des communications de données et TCP/IP* pour savoir comment parvenir au même résultat.

4. Créez des groupes de sauvegarde du service PNM avec la commande

`pnmset (1M)`.

Exécutez la commande de script interactif `pnmset (1M)` pour configurer les groupes de sauvegarde.



Attention - Si vous avez déjà configuré les hôtes logiques et les services de données, vous devez interrompre les services de données HD avant de modifier l'appartenance au groupe de sauvegarde avec `pnmset (1M)`. Le défaut d'interrompre les services de données avant d'exécuter la commande `pnmset (1M)` peut entraîner de graves problèmes et des pannes majeures au niveau des services de données.

a. Exécutez la commande `pnmset (1M)`.

```
phys-hahost1# /opt/SUNWpnm/bin/pnmset
```

b. Entrez le nombre total de groupes de sauvegarde que vous souhaitez configurer.

Ce nombre correspond habituellement à celui des sous-réseaux publics.

```
Le message suivant vous demande de configurer la gestion de réseau public.  
Voulez-vous continuer ... [o/n] : o
```

(suite)

```
Nombre de groupes de sauvegarde NAFO sur l'hôte [1] : 2
```

c. Attribuez les numéros aux groupes de sauvegarde.

Lorsque vous y êtes invité, entrez un nombre entier compris entre 0 et 255. La commande `pnmset(1M)` ajoute ce nombre à la chaîne `nafo` pour former le nom du groupe de sauvegarde.

```
Entrez le numéro du groupe de sauvegarde [0] : 0
```

d. Attribuez des adaptateurs aux groupes de sauvegarde.

```
Veillez entrer tous les adaptateurs réseau sous nafo0 :  
qe0 qe1  
...
```

Attribuez ensuite les numéros de groupe de sauvegarde et les adaptateurs pour les autres groupes de sauvegarde de la configuration.

e. Donnez le temps à la commande `pnmset(1M)` de tester la configuration des adaptateurs.

La commande `pnmset(1M)` teste la conformité de la configuration des adaptateurs. Dans cet exemple, le groupe de sauvegarde contient un adaptateur réseau actif et deux adaptateurs réseau redondants.

```
Le test suivant évalue la conformité de la configuration NAFO...  
test de duplication des noms réussi  
  
Vérification de nafo0... < 20 secondes  
qe0 est actif  
adresse distante = 192.168.142.1  
test de nafo0 réussi  
  
Vérification de nafo1... < 20 secondes  
qe3 est actif  
adresse distante = 192.168.143.1  
test de qe4 attendez...  
test de qe2 attendez...
```

(suite)

(suite)

```
test de nafol réussi  
phys-hahost1#
```

Une fois la configuration vérifiée, le démon PNM `pnmd(1M)` enregistre automatiquement les modifications apportées à la configuration et amorce la surveillance des interfaces.

Remarque - Un seul adaptateur dans un groupe de sauvegarde doit être étalonné et posséder une entrée dans le fichier `/etc/nom_hôte.adaptateur`. N'attribuez pas d'adresse IP aux adaptateurs réseau de relève ; elles ne doivent pas être étalonnées.

Remarque - Le service PNM utilise la diffusion `ping(1M)` pour surveiller les réseaux ; ces derniers utilisant la diffusion de paquets ICMP (Internet Control Message Protocol) pour communiquer avec les autres hôtes distants. Certains routeurs ne transmettent pas les paquets ICMP ; par conséquent, le comportement de détection de défaillances s'en trouve touché. Voir les *Sun Cluster 2.2 Release Notes* pour contourner ce problème.

5. Démarrez la grappe avec la commande `scadmin(1M)`.

Exécutez ensuite les commandes suivantes sur un noeud :

```
# scadmin startcluster nom_hôte_physique sc-cluster
```

Ajoutez ensuite tous les autres noeuds à la grappe en exécutant la commande suivante à partir de ceux-ci :

```
# scadmin startnode
```

6. Vérifiez la configuration du service PNM avec la commande `pnmstat(1M)`.

```
phys-hahost1# /opt/SUNWpnm/bin/pnmstat -l
grp_sauvegarde adaptateur_réel état temps_reprise adaptateur_actif
nafo0 hme0 OK JAMAIS hme0
phys-hahost1#
```

La configuration initiale du service PNM est maintenant terminée.

▼ Comment reconfigurer le service PNM

Utilisez cette procédure pour modifier une configuration existante par l'ajout ou la suppression d'adaptateurs réseau. Les étapes suivantes permettent d'administrer un noeud à la fois tout en maintenant la disponibilité des services Sun Clustertout au long de la procédure.

1. Interrompez l'exécution du logiciel Sun Cluster sur le noeud à reconfigurer.

```
phys-hahost1# scadmin stopnode
```

2. Ajoutez ou supprimez des adaptateurs réseau.

Effectuez les procédures de la "Ajout et suppression d'interfaces réseau", page 127.

3. Exécutez la commande `pnmset (1M)` pour reconfigurer les groupes de sauvegarde.

Exécutez la commande `pnmset (1M)` pour reconfigurer les groupes de sauvegarde, comme l'indique Etape 4, page 120 de la procédure "Comment configurer le service PNM", page 119.

```
phys-hahost# pnmset
```

4. Redémarrez le logiciel Sun Cluster sur le noeud.

A partir du poste de travail administratif, redémarrez le noeud avec la commande suivante :

```
phys-hahost1# scadmin startnode
```

5. Effectuez de nouveau les étapes 1 à 4 pour chaque noeud à reconfigurer.

▼ Comment vérifier l'état des groupes de sauvegarde

Les commandes `pnmptor(1M)` et `pnmrtop(1M)` permettent de vérifier l'état des groupes de sauvegarde locaux ; par contre, la commande `pnmstat(1M)` vérifie l'état des groupes de sauvegarde locaux et distants.

1. Exécutez la commande `pnmptor(1M)` pour déterminer à quel groupe de sauvegarde appartient un adaptateur réseau.

La commande `pnmptor(1M)` établit la correspondance entre le pseudonyme d'adaptateur réseau spécifié et un nom d'adaptateur réseau réel. Dans l'exemple suivant, les données de sortie du système montrent que le pseudonyme d'adaptateur réseau `nafo0` est associé à l'adaptateur réseau actif `hme2` :

```
phys-hahost1# pnmptor nafo0
hme2
```

1. Exécutez la commande `pnmrtop(1M)` pour identifier l'adaptateur réseau actif associé à un groupe de sauvegarde donné.

Dans l'exemple ci-dessous, les données de sortie du système montrent que l'adaptateur réseau `hme1` appartient au groupe de sauvegarde `nafo0` :

```
phys-hahost1# pnmrtop hme1
nafo0
```

1. Exécutez la commande `pnmstat(1M)` pour connaître l'état d'un groupe de sauvegarde.

Utilisez l'option `-c` pour déterminer l'état d'un groupe de sauvegarde sur l'hôte local :

```
phys-hahost1# pnmstat -c nafo0
OK
JAMAIS
hme2
```

La syntaxe de la commande suivante permet de déterminer l'état d'un groupe de sauvegarde sur un hôte distant :

```
phys-hahost1# pnmstat -sh hôte_distant -c nafol
OK
JAMAIS
qel
```

Remarque - Les options `-s` et `-h` doivent être utilisées ensemble. L'option `-s` oblige `pnmstat(1M)` à communiquer par l'interconnexion privée. Si vous omettez l'option `-s`, `pnmstat(1M)` interroge le système par l'interconnexion publique. L'`hôte_distant` et l'hôte sur lequel vous exécutez la commande `pnmstat(1M)` doivent appartenir tous deux à la grappe.

Lorsque vous vérifiez l'hôte local ou l'hôte distant, la commande `pnmstat(1M)` renvoie des données sur leur état, leur historique et l'adaptateur réseau actif. Pour de plus amples renseignements, consultez la page de manuel pertinente.

Paramètres configurables du service PNM

Le tableau suivant contient la liste des paramètres du service PNM que l'utilisateur peut configurer. Après avoir installé le service PNM et avant d'afficher la grappe, configurez ces paramètres en modifiant manuellement le fichier de configuration file `/opt/SUNWcluster/conf/TEMPLATE.cdb` sur l'ensemble des noeuds de la grappe. Modifiez le fichier sur un noeud et copiez-le dans tous les autres noeuds, ou encore utilisez la console de grappe pour modifier simultanément le fichier dans tous les noeuds. Vous pouvez afficher la configuration actuelle du service PNM avec `pnmstat -t`. Pour de plus amples renseignements, consultez la page de manuel `pnmstat(1M)`.

TABLEAU 6-2 Paramètres configurables du service PNM

<code>pnmd.inactive_time</code>	Intervalle entre les sondages de défaillances, exprimé en secondes. La valeur par défaut est 5 secondes.
<code>pnmd.ping_timeout</code>	Durée avant le délai du sondage de défaillances, 4 secondes par défaut.
<code>pnmd.repeat_test</code>	Nombre maximum de tentatives de sondage par le service PNM avant le constat d'un problème, 3 sondages par défaut.
<code>pnmd.slow_network</code>	Temps d'attente entre la phase d'écoute et la phase de sondage actif des défaillances, 2 secondes par défaut. (Augmentez cette valeur si votre réseau est lent et qu'il force le service PNM à déclencher des processus de relève parasites).

Résolution des erreurs du service PNM

Voici les messages d'erreur les plus courants que renvoie le service PNM.

```
Échec du service PNM sur l'ordinateur distant
```

Ce message d'erreur indique que le démon PNM n'a pas démarré. Redémarrez-le avec la commande ci-dessous. *ID-noeud* *ID-noeud* représente la valeur retournée par la commande `/opt/SUNWcluster/bin/get_node_status`.

```
# /opt/SUNWpnm/bin/pnmd -s -c nom-grappe -l ID-noeud
```

```
PNM non démarré
```

Ce message indique qu'il n'y a pas de groupe de sauvegarde configuré. Exécutez la commande `pnmset(1M)` pour créer des groupes de sauvegarde.

```
Pas de nafoXX
```

Ce message signale que vous avez spécifié un nom de groupe de sauvegarde non admis. Exécutez la commande `pnmrtop(1M)` pour identifier le nom des groupes de sauvegarde associés à un adaptateur réseau donné. Exécutez de nouveau la commande avec un nom de groupe de sauvegarde valide.

Erreur de configuration du service PNM

Ce message signale que le démon PNM n'a pas pu configurer un adaptateur ou qu'il y a une erreur de formatage dans le fichier de configuration `/etc/pnmconfig`. Vérifiez les messages syslog et prenez les mesures spécifiées par Sun Cluster Manager. Pour de plus amples renseignements sur Sun Cluster Manager, reportez-vous au Chapitre 2.

Erreur de programme

Ce message indique que le démon PNM n'a pas pu exécuter un appel système. Vérifiez les messages syslog et prenez les mesures spécifiées par Sun Cluster Manager. Pour de plus amples renseignements sur Sun Cluster Manager, reportez-vous au Chapitre 2.

Ajout et suppression d'interfaces réseau

Les procédures décrites dans cette section permettent d'ajouter ou de supprimer des cartes d'interface réseau public dans une configuration de grappe.

Pour ajouter ou supprimer une interface réseau commandant un hôte logique, vous devez modifier la configuration de tous les hôtes logiques utilisant cette interface. Pour ce faire, vous devez supprimer entièrement l'hôte logique de la grappe, puis le rajouter avec les modifications nécessaires. Les commandes `scconf(1M)` et `scinstall(1M)` permettent de reconfigurer un hôte logique. Les exemples de cette section portent sur la commande `scconf(1M)`. Reportez-vous à "Ajout et suppression d'hôtes logiques", page 60 qui contient les étapes pour procéder à la configuration de l'hôte logique avec la commande `scinstall(1M)`.

Ajout d'une interface réseau

Pour ajouter une interface réseau, vous devez supprimer et refaire la configuration de tous les hôtes logiques qui lui sont associés. Il est à noter que l'accès à l'ensemble

des services de données est interrompu pour un court laps de temps pendant la procédure.

▼ Comment ajouter une interface réseau

Effectuez les étapes suivantes sur chaque noeud qui recevra une nouvelle carte d'interface réseau.

1. Interrompez l'exécution du logiciel de grappes.

```
phys-hahost# scadmin stopnode
```

2. Ajoutez la nouvelle carte d'interface en suivant les directives fournies avec celle-ci.

3. Configurez la nouvelle interface réseau sur chaque noeud.

Cette étape n'est nécessaire que si la nouvelle interface fait partie d'un hôte logique. Sautez cette étape si la configuration de votre système ne comporte pas d'hôtes logiques.

```
phys-hahost# pnmset
```

Dans le cas des interfaces Ethernet, créez un nouveau fichier `/etc/hostname.if` pour chaque nouvelle interface sur chaque noeud, puis exécutez la commande `ifconfig(1M)` comme vous le feriez avec un environnement ne comportant pas de grappe.

Remarque - Lorsque vous configurez un ensemble d'interfaces réseau dont se serviront les divers hôtes logiques d'une grappe, vous devez connecter toutes les interfaces au même sous-réseau.

4. Démarrez le logiciel de grappes.

Si tous les noeuds ont été arrêtés, exécutez la commande `scadmin startcluster` sur le noeud 0 puis la commande `scadmin startnode` sur tous les autres noeuds. Si le logiciel de grappes est toujours en cours d'exécution sur au moins un des noeuds, exécutez la commande `scadmin startnode` sur les noeuds restants.

```
phys-hahost# scadmin startnode
```

Si les nouvelles interfaces sont ajoutées à des groupes de sauvegarde déjà existants, la procédure est terminée.

Si vous avez modifié la configuration des groupes de sauvegarde, vous devez rétablir le fonctionnement normal de la grappe et reconfigurer chaque hôte logique qui utilise le nouvel ensemble de contrôleurs réseau. Vous devez ensuite supprimer puis reconfigurer chaque hôte logique. Vous devez donc exécuter la commande `scconf -p` pour imprimer la configuration actuelle avant de passer à ces étapes. Vous pouvez exécuter la commande `scconf -p` sur n'importe quel noeud actif de la grappe. Il n'est pas nécessaire de l'exécuter sur tous les noeuds de la grappe.

Pour supprimer la configuration et reconfigurer l'hôte logique, exécutez la commande `scconf(1M)` tel qu'illustré dans ces exemples, ou la commande `scinstall(1M)`, comme le décrit "Ajout et suppression de noeuds de grappe", page 54.

5. **Avisez les utilisateurs qu'ils ne pourront pas utiliser les services de données des hôtes logiques touchés pendant un court laps de temps.**
6. **Faites une copie de sauvegarde des fichiers `/etc/opt/SUNWcluster/conf/ccd.database` de chaque noeud au cas où vous auriez à restaurer la configuration initiale.**
7. **Désactivez les services de données.**

```
phys-hahost# hareg -n service_données
```

8. **Supprimez l'inscription des services de données.**

```
phys-hahost# hareg -u service_données
```

9. **Supprimez l'hôte logique de la grappe.**

Exécutez cette commande sur un des noeuds actifs appartenant à la grappe. Il n'est pas nécessaire de l'exécuter sur tous les noeuds.

```
phys-hahost# scconf nom_grappe -L hôte_logique -r
```

10. **Reconfigurez l'hôte logique pour qu'il inclue la nouvelle interface.**

Exécutez cette commande sur un des noeuds actifs appartenant à la grappe. Il n'est pas nécessaire de l'exécuter sur tous les noeuds.

```
phys-hahost# scconf nom_grappe -L hôte_logique -n nodelist -g dglst -i logaddrinfodglst
```

Le champ *logaddrinfo* définit le nom de la nouvelle interface. Consultez la liste produite par la commande `scconf -p` pour reconstruire chaque hôte logique.

11. Inscrivez les services de données.

```
phys-hahost# hareg [-s] -r service_données
```

12. Activez les services de données.

```
phys-hahost# hareg -y service_données
```

13. Vérifiez l'accès aux services de données.

14. Avisez les utilisateurs que les services de données sont maintenant disponibles.

La procédure d'ajout d'une interface réseau est terminée.

Suppression d'une interface réseau

La procédure suivante permet de supprimer une interface de réseau public d'une grappe.

- Si vous utilisez une configuration OPS, la suppression d'une interface réseau ne requiert aucune intervention au niveau de la grappe. Par contre, suivez la procédure ci-dessous pour supprimer les adaptateurs réseau des noeuds de grappe.
- Les configurations HD requièrent cette procédure pour supprimer et refaire la configuration de tout hôte logique utilisant l'interface réseau à supprimer. Il est alors nécessaire d'interrompre l'accès à l'ensemble des services de données pour un court laps de temps pendant la procédure.

▼ Comment supprimer une interface réseau

Pendant que tous les noeuds sont actifs dans la grappe, exécutez les étapes suivantes sur un seul noeud.

1. **Identifiez les hôtes logiques qui doivent être reconfigurés de manière à exclure l'interface réseau.**

La configuration de tous ces hôtes logiques doit être supprimée puis refaite. Exécutez la commande `sccconf -p` pour imprimer la liste des hôtes logiques de la configuration actuelle ; enregistrez cette liste pour pouvoir l'utiliser ultérieurement. Il n'est pas nécessaire d'exécuter la commande `sccconf -p` sur tous les noeuds de la grappe. Il suffit de l'exécuter sur n'importe quel noeud actif appartenant à la grappe.

2. **Exécutez la commande `pnmset (1M)` pour afficher la configuration PNM actuelle.**
3. **Supprimez au besoin le contrôleur d'un groupe de sauvegarde.**
Si le contrôleur à supprimer fait partie d'un groupe de sauvegarde, supprimez le contrôleur de tous les hôtes logiques, puis exécutez la commande `pnmset (1M)` pour supprimer le contrôleur du groupe de sauvegarde.
4. **Avisez les utilisateurs qu'ils n'auront pas accès aux services de données des hôtes logiques touchés durant une courte période.**
5. **Désactivez les services de données.**

```
phys-hahost# hares -n service_données
```

6. **Supprimez l'inscription des services de données.**

```
phys-hahost# hares -u service_données
```

7. **Supprimez l'hôte logique de la grappe.**

Remarque - Pour supprimer et refaire la configuration de l'hôte logique (Etape 7, page 131 et Etape 8, page 131), exécutez la commande `sccconf(1M)` tel que décrit, ou exécutez la commande `scinstall(1M)` comme l'indique "Ajout et suppression de noeuds de grappe", page 54.

Vous pouvez exécuter cette commande sur n'importe quel noeud actif appartenant à la grappe. Il n'est pas nécessaire de l'exécuter sur tous les noeuds.

```
phys-hahost# sccconf nom_grappe -L hôte_logique -r
```

8. **Reconfigurez l'hôte logique pour qu'il inclue la nouvelle interface.**
Vous pouvez exécuter cette commande sur n'importe quel noeud actif appartenant à la grappe. Il n'est pas nécessaire de l'exécuter sur tous les noeuds.

```
phys-hahost# scconf nom_groupe -L hôte_logique -n nodelist -g dglst -i logaddrinfo
```

Le champ *logaddrinfo* définit le nom de la nouvelle interface. Consultez la liste produite par la commande `scconf -p` pour reconstruire chaque hôte logique.

9. Si le contrôleur à supprimer faisait partie d'un groupe de sauvegarde, exécutez de nouveau la commande `pnmset (1M)`.

Exécutez de nouveau la commande `pnmset (1M)` pour exclure le contrôleur à supprimer.

10. (Facultatif) Si vous supprimez l'adaptateur réseau des noeuds, exécutez les étapes suivantes sur chaque noeud touché :

a. Interrompez l'exécution du logiciel de grappes.

```
phys-hahost# scadmin stopnode
```

b. Arrêtez le noeud et enlevez la carte d'interface.

c. Initialisez le noeud.

d. Exécutez les tâches d'administration habituelles dans le système Solaris pour supprimer une interface réseau (suppression du fichier *nom_hôte.if*, mise à jour de */etc/hosts*, etc).

e. Redémarrez le logiciel de grappes. Si tous les noeuds ont été arrêtés, démarrez le premier noeud avec la commande `scadmin startcluster`. Si le logiciel de grappes est en cours d'exécution sur au moins un des noeuds, redémarrez les autres noeuds.

```
phys-hahost# scadmin startnode
```

11. Inscrivez les services de données.

```
phys-hahost# hareg -r service_données
```

12. Activez les services de données.

```
phys-hahost# hareg -y service_données
```

13. Vérifiez l'accès aux services de données.

14. Avisez les utilisateurs que les services de données sont maintenant disponibles.

Administration de l'Agent de gestion de commutation

L'Agent de gestion de commutation (AGC) est un module de grappe assurant la gestion des canaux de communication au niveau de l'interconnexion privée de la grappe. Il surveille l'interconnexion privée et appelle une procédure de reprise vers un réseau de sauvegarde lorsqu'il détecte une défaillance.

Prenez note des restrictions suivantes avant de débiter la procédure.

- Les noeuds SC2000/SS1000 n'acceptent pas plus d'une carte SCI sur une carte système. S'il y en a davantage, il peut se produire des réinitialisations parasites au niveau des liens de l'interconnexion SCI.
- Un SBus de noeud E10000 doit comporter d'autres cartes en plus de la carte SCI.
- N'installez pas d'adaptateurs SCI et d'autres adaptateurs d'hôte A3000 sur le même SBus d'une configuration Sun StorEdge A3000.

Voir également l'Annexe B dans le *Sun Cluster 2.2 Hardware Site Preparation, Planning, and Installation Guide*.

▼ Comment ajouter des commutateurs et des cartes SCI

Cette procédure permet d'ajouter des commutateurs et des cartes SCI aux noeuds de la grappe. Pour de plus amples renseignements, consultez la page de manuel `sm_config(1M)`.

1. Modifiez le fichier modèle `sm_config` pour tenir compte des changements apportés à la configuration.

Ce fichier modèle se trouve habituellement dans le répertoire `/opt/SUNWsm/bin/Examples`.

2. Configurez les cartes SCI SBus en exécutant la commande `sm_config(1M)` à partir d'un noeud de la grappe.

Exécutez une seconde fois la commande pour vous assurer que les ID et les adresses IP des noeuds SCI sont correctement attribuées aux noeuds de la grappe. Toute attribution erronée peut nuire aux communications entre les noeuds.

3. Réinitialisez les nouveaux noeuds.

Dépannage du logiciel SCI

Si un problème survient avec le logiciel SCI, vérifiez si les éléments suivants s'appliquent :

- Le fichier modèle `sm_config(1M)` correspond à la configuration matérielle (lien et commutateur SCI) et à la topologie de la grappe.
- La commande `sm_config(1M)` peut être exécutée avec succès à partir d'un noeud de la grappe.
- Un noeud reconfiguré a été réinitialisé après l'exécution de la commande `sm_config(1M)`.

Prenez également note des problèmes suivants et de leurs solutions :

- Avec certaines applications telles que Oracle Parallel Server (OPS), il est nécessaire de spécifier une quantité minimale de mémoire partagée anormalement élevée dans le fichier `/etc/system`. Si la valeur du champ `shmsys:shminfo_shmmin` dans le fichier `/etc/system` dépasse les 200 octets, la commande `sm_config(1M)` ne peut pas obtenir sa part de mémoire partagée parce que ses besoins à ce niveau sont inférieurs de quelques octets seulement par rapport à la quantité minimale que le système peut lui allouer. Cela se traduit par un échec de l'appel système effectué par la commande `sm_config(1M)`, et cette commande ne peut être exécutée.

Pour pallier ce problème, modifiez le fichier `/etc/system` et fixez la valeur de `shmsys:shminfo_shmmin` à moins de 200 octets. Réinitialisez ensuite l'ordinateur pour qu'il tienne compte des nouvelles valeurs.

- Si des avertissements `semsys` et des vidages d'image mémoire se produisent, vérifiez que les valeurs de sémaphore contenues dans les champs `semsys:seminfo_*` du fichier `/etc/system` correspondent bien aux limites physiques réelles de l'ordinateur.

Pour de plus amples renseignements sur les composants SCI, voir l'Annexe B dans le *Sun Cluster 2.2 Hardware Site Preparation, Planning, and Installation Guide*.

▼ Comment vérifier la connectivité entre les noeuds

Deux commandes permettent de vérifier la connectivité entre les noeuds : `get_ci_status(1M)` et `ping(1)`.

1. Exécutez la commande `get_ci_status(1M)` sur tous les noeuds de grappe.

Voici un exemple de données de sortie obtenues avec la commande `get_ci_status(1M)` :

```
# /opt/SUNWsma/bin/get_ci_status
sma : sci #0 : no_emplacement_sbus 1 ; id_adaptateur 8 (0x08) ;
adresse_ip 1 ; id_commutateur 0 ; id_port 0 ;
    Etat de l'adaptateur - ACTIF ; Etat du lien - ACTIF
sma : sci #1 : no_emplacement_sbus 2 ; id_adaptateur 12 (0x0c) ;
adresse_IP 17 ; id_commutateur 1 ; id_port 0 ;
    Etat de l'adaptateur - ACTIF ; Etat du lien - ACTIF
sma : id_commutateur 0sma : id_port 1 : nom_hôte = interconn2 ;
id_adaptateur = 72 ; actif | fonctionnel
sma : id_port 2 : nom_hôte = interconn3 ; id_adaptateur = 136 ;
actif | fonctionnel
sma : id_port 3 : nom_hôte = interconn4 ; id_adaptateur = 200 ; actif | fonctionnel
sma : id_commutateur 1
sma : id_port 1 : nom_hôte = interconn2 ; id_adaptateur = 76 ; actif | fonctionnel
sma : id_port 2 : nom_hôte = interconn3 ; id_adaptateur = 140 ; actif | fonctionnel
sma : id_port 3 : nom_hôte = interconn4 ; id_adaptateur = 204 ; actif | fonctionnel
#
```

Les quatre premières lignes indiquent l'état du noeud logique (interconn1 dans le cas présent). Ce dernier communique avec `id_commutateur 0` et `id_commutateur 1` (État du lien - ACTIF).

```
sma : sci #0 : no_emplacement_sbus 1 ; id_adaptateur 8 (0x08) ; adresse_ip 1 ;
id_commutateur 0 ; id_port 0 ;
    Etat de l'adaptateur - ACTIF ; Etat du lien - ACTIF
sma : sci #1 : no_emplacement_sbus 2 ; id_adaptateur 12 (0x0c) ; adresse_IP 17 ;
id_commutateur 1 ; id_port 0 ;
    Etat de l'adaptateur - ACTIF ; Etat du lien - ACTIF
```

Les données de sortie restantes indiquent l'état général des autres noeuds de la grappe. Tous les ports des deux commutateurs communiquent avec leurs noeuds respectifs. S'il y a un problème au niveau du matériel, l'état inactif apparaît plutôt que actif. Si le problème se situe au niveau logiciel, l'état non fonctionnel apparaît plutôt que fonctionnel.

```
sma : id_commutateur 0
sma : id_port 1 : nom_hôte = interconn2 ; id_adaptateur = 72 ; actif | fonctionnel
sma : id_port 2 : nom_hôte = interconn3 ; id_adaptateur = 136 ; actif | fonctionnel
sma : id_port 3 : nom_hôte = interconn4 ; id_adaptateur = 200 ; actif | fonctionnel
sma : id_commutateur 1
sma : id_port 1 : nom_hôte = interconn2 ; id_adaptateur = 76 ; actif | fonctionnel
sma : id_port 2 : nom_hôte = interconn3 ; id_adaptateur = 140 ; actif | fonctionnel
sma : id_port 3 : nom_hôte = interconn4 ; id_adaptateur = 204 ; actif | fonctionnel
```

(suite)

```
#
```

1. Exécutez la commande `ping(1)` sur toutes les adresses IP des noeuds distants.

Voici un exemple de données de sortie obtenues avec la commande `ping(1)`.

```
# ping adresse-IP
```

Les adresses IP se trouvent dans le fichier `/etc/sma.ip`. Prenez soin d'exécuter la commande `ping(1)` sur chaque noeud de la grappe.

La commande `ping(1)` renvoie le message "en activité", ce qui confirme que les deux extrémités communiquent sans problème. Autrement, un message d'erreur apparaît.

Par exemple,

```
# ping 204.152.65.2
204.152.65.2 en activité
```

▼ Comment vérifier la configuration des interfaces SCI

1. Exécutez la commande `ifconfig -a` pour vérifier si toutes les interfaces SCI sont actives et que les noeuds de grappe possèdent une adresse IP exacte.

Les 8 derniers bits de l'adresse IP doivent correspondre à la valeur du champ IP dans le fichier `/etc/sma.config`.

```
# ifconfig -a
lo0 : indicateurs=849 <ACTIF,REBOUCLAGE,EN COURS D'EXECUTION,MULTIDESTINATAIRE> mtu 8232
inet 127.0.0.1 masque_réseau ff000000
hme0 : indicateurs=863<ACTIF,DIFFUSION,
PAS DE BLOCS DE FIN,EN COURS D'EXECUTION,MULTIDESTINATAIRES> mtu 1500
inet 129.146.238.55 masque_réseau ffffffff00 diffusion 129.146.238.255
ether 8:0:20:7b:fa:0
scid0 : indicateurs=80c1<ACTIF,EN COURS D'EXECUTION,PAS D'ARP,PRIVE> mtu 16321
```

(suite)

(suite)

```
inet 204.152.65.1 masques_réseau ffffffff0  
scidl : indicateurs=80cl<ACTIF,EN COURS D'EXECUTION,PAS D'ARP,PRIVE> mtu 16321  
inet 204.152.65.17 masques_réseau ffffffff0
```


Administration des composants serveurs

Le présent chapitre décrit la procédure permettant d'ajouter ou de supprimer les composants d'un noeud Sun Cluster. La liste ci-dessous présente les sujets abordés dans le chapitre.

- “Remplacement de cartes système”, page 139
- “Ajout de modules de carte”, page 140
- “Remplacement de cartes SBus”, page 141

Remplacement de cartes système

Le composant Solstice DiskSuite de Sun Cluster tient compte de la numérotation des périphériques, et la réorganisation des cartes système peut le désorienter. Pour de plus amples renseignements sur les noms et les numéros d'instances, reportez-vous au Chapitre 1.

A la première initialisation du noeud, les entrées de l'unité d'expansion de disques multihôtes du répertoire `/dev` sont liées à la fente de la connexion.

Par exemple, lorsque le noeud est initialisé, la carte système 0 et la fente SBus 1 font partie de l'identité de l'unité d'expansion de disques multihôtes. Si l'on installe la carte système ou la carte SBus dans un nouvel emplacement, Solstice DiskSuite devient confus, parce que Solaris assigne de nouveaux numéros aux contrôleurs SBus après leur déplacement.

Remarque - Vous pouvez déplacer les cartes SBus entre les différentes fentes, tant que le type de carte SBus demeure le même.

La réorganisation des câbles à fibres optiques connectés aux unités d'expansion de disques multihôtes peut créer des problèmes. Après avoir interverti les cartes SBus, reconnectez les unités d'expansion de disques multihôtes à la même fente SBus qu'avant la substitution.

Ajout de modules de carte

L'ajout ou le remplacement de modules de carte comme les barrettes de mémoire SIMM et les UC nécessitent une intervention au niveau tant logiciel que matériel.

▼ Comment ajouter des modules de carte

1. **Interrompez l'exécution de Sun Cluster sur le noeud qui doit recevoir le module de carte.**

Dans cet exemple, `phys-hahost2` reçoit le module de carte en premier.

```
phys-hahost2 scadmin stopnode
```

2. **Arrêtez le noeud.**

```
phys-hahost2# halt
```

3. **Mettez le noeud hors tension.**

4. **Installez le module de carte en respectant les directives du guide d'installation pertinent.**

5. **Mettez le noeud sous tension.**

6. **Effectuez une réinitialisation de la reconfiguration.**

```
ok boot -r
```

7. **Démarrez le logiciel de grappes sur le noeud.**

```
# scadmin startnode
```

8. Répétez Etape 1, page 140 jusqu'à Etape 7, page 140 sur les autres noeuds de Sun Cluster requérant la même mise à niveau matérielle.
9. Commutez au besoin les hôtes logiques à leurs maîtres par défaut respectifs.
Une commutation automatique se produit lorsque le mode manuel n'est pas activé.

```
phys-hahost2# haswitch phys-hahost1# hahost1
```

Remplacement de cartes SBus

Pour remplacer les cartes SBus des noeuds Sun Cluster, commutez les services de données vers le noeud fonctionnel avant de procéder au remplacement physique de la carte. Commutez ensuite les hôtes logiques vers leur maître par défaut respectif selon la procédure prescrite.

▼ Comment remplacer les cartes SBus

1. Commutez la propriété des hôtes logiques à partir du noeud Sun Cluster dont la carte SBus doit être remplacée.

Par exemple, entrez la commande suivante si vous remplacez la carte sur l'hôte physique `phys-hahost2` :

```
hôte_physique1# haswitch phys_host1 hahost1 hahost2
```

2. Interrompez l'exécution de Sun Cluster sur le noeud touché.

Exécutez la commande `scadmin(1M)` avec l'option `stopnode` sur l'hôte où se trouve la carte SBus défectueuse.

```
phys-hahost2# scadmin stopnode
```

3. Arrêtez le noeud touché et mettez-le hors tension.
4. Procédez au remplacement du composant matériel.

Pour savoir comment remplacer la carte SBus, reportez-vous aux directives du manuel d'entretien du composant matériel approprié.

5. Mettez le noeud sous tension et démarrez le logiciel de grappes sur le noeud.

```
# scadmin startnode
```

Le noeud réintègre automatiquement la configuration de Sun Cluster.

6. Commutez au besoin les hôtes logiques vers leurs maîtres par défaut respectifs.

Une commutation automatique se produit lorsque le mode manuel n'est pas activé.

```
phys-hahost2# haswitch phys-hahost2# hahost2
```

Administration du concentrateur de terminaux

Ce chapitre décrit les instructions d'utilisation du concentrateur de terminaux pour l'administration des configurations Sun Cluster. Consultez également le Chapitre 5 du *Sun Cluster 2.2 Hardware Site Preparation, Planning, and Installation Guide*.

Les sujets abordés sont énumérés ci-dessous.

- “Connexion à la console Sun Cluster”, page 143
- “Réinitialisation d'une connexion au concentrateur de terminaux”, page 146
- “Entrée dans la mémoire PROM OpenBoot d'un serveur Sun Cluster”, page 148
- “Dépannage du concentrateur de terminaux”, page 149
- “Modification des informations CT/PSS”, page 153

Connexion à la console Sun Cluster

Vous pouvez vous acquitter des tâches administratives depuis une fenêtre reliée à n'importe quel noeud Sun Cluster. Le guide de planification et d'installation du noeud Sun Cluster et la documentation fournie avec le concentrateur de terminaux contiennent les procédures permettant de configurer le concentrateur de terminaux et de définir les paramètres de sécurité.

La procédure suivante explique comment créer des connexions à partir du poste de travail administratif dans une configuration Sun Cluster.

La taille de `shelltool(1)` étant variable et la connexion étant établie par l'intermédiaire d'une interface de console sérielle, le port de console n'est pas en mesure de déterminer la taille de la fenêtre `shelltool(1)` depuis laquelle la

connexion a été établie. Vous devez régler manuellement la taille de la fenêtre sur les noeuds pour toute application ayant besoin des données sur les valeurs des colonnes et des rangées.

▼ Comment se connecter à la console Sun Cluster

1. Ouvrez une fenêtre `shelltool(1)` dans le bureau d'un poste de travail.
2. Exécutez la commande `tput(1)` et prenez en note la taille de la fenêtre `shelltool(1)`.

Ces valeurs serviront pour l'Etape 6, page 145.

```
# tput lines
35
# tput cols
80
```

3. Tapez la commande suivante pour établir une connexion `telnet(1)` avec un des noeuds Sun Cluster par l'intermédiaire du concentrateur de terminaux.

```
# telnet terminal-concentrator-name 5002
Tentative avec 192.9.200.1 ...
Connecté à 192.9.200.1.
Caractère d'échappement '^]'
```

Remarque - Les numéros de port varient en fonction de la configuration. Habituellement, les ports 2 et 3 (5002 et 5003 dans les exemples) servent à la première grappe Solaris d'un site.

4. Ouvrez une autre fenêtre de `shelltool(1)` et tapez la commande suivante pour établir une connexion `telnet(1)` avec l'autre noeud.

```
# telnet terminal-concentrator-name 5003
Tentative avec 192.9.200.1 ...
Connecté à 192.9.200.1.
Caractère d'échappement '^]'.

```

Remarque - Le système vous demande d'entrer le mot de passe du port si vous avez défini les paramètres de sécurité en suivant les directives du guide de planification et d'installation du noeud Sun Cluster. Après avoir établi la connexion, entrez le nom de connexion et le mot de passe.

5. Connectez-vous au noeud.

```
Connexion à la console : root (superutilisateur)
Mot de passe : racine-Mot de passe :

```

6. Utilisez la commande `stty(1)` pour réinitialiser les valeurs des rangées et colonnes du terminal à celles établies à Etape 2, page 144.

```
# stty rows 35
# stty cols 80

```

7. Réglez la variable d'environnement `TERM` à la valeur appropriée correspondant au type de fenêtre utilisée à l'Etape 1, page 151.

Par exemple, si vous utilisez une fenêtre `xterm`, tapez :

```
# TERM=xterm ; export TERM (sh ou ksh)
ou
# setenv TERM xterm (csh)

```

Réinitialisation d'une connexion au concentrateur de terminaux

Cette section explique comment réinitialiser une connexion au concentrateur de terminaux.

Si un autre utilisateur est connecté au port de console d'un noeud Sun Cluster sur le concentrateur de terminaux, vous pouvez réinitialiser ce port pour rompre la connexion de cet utilisateur. Cette procédure s'avère utile lorsque vous devez effectuer immédiatement une tâche administrative.

Si vous ne pouvez pas vous connecter au concentrateur de terminaux, le message suivant s'affiche :

```
# telnet terminal-concentrator-name 5002
Tentative avec 192.9.200.1 ...
telnet : Impossible de se connecter à l'hôte distant : Connexions refusée
#
```

Si vous utilisez le sélecteur de port, un message signalant que le port est occupé peut s'afficher.

▼ Comment réinitialiser une connexion au concentrateur de terminaux

1. **Après avoir établi la connexion, appuyez une nouvelle fois sur la touche Entrée et sélectionnez l'interface de ligne de commande (cli) pour vous connecter au concentrateur de terminaux.**

L'invite `annex:` s'affiche.

```
# telnet nom-concentrateur-terminaux
...
Entrer le nom ou le numéro du port Annex : cli
...
annex :
```

2. **Tapez la commande `su` et le mot de passe.**

Par défaut, le mot de passe est l'adresse IP du concentrateur de terminaux.

```
annex# su
Mot de passe :
```

3. Sélectionnez le port à réinitialiser.

Dans cet exemple, il s'agit du port 2. Utilisez la commande `who` intégrée au concentrateur de terminaux pour afficher les connexions.

```
annex# who
Port  Quoi  Utilisateur  Emplacement  Quand  Inactif
Adresse
2  PSVR  ---  ---  ---  1:27  192.9.75.12
v1  CLI  ---  ---  ---  192.9.76.10
```

4. Réinitialisez le port.

Utilisez la commande `reset` intégrée au concentrateur de terminaux pour réinitialiser le port. Dans l'exemple suivant, la connexion au Port 2 est interrompue.

```
annex# admin reset 2
```

5. Déconnectez-vous du concentrateur de terminaux.

```
annex# hangup
```

6. Reconnectez-vous au port.

```
# telnet nom-concentrateur-terminaux 5002
```

Entrée dans la mémoire PROM OpenBoot d'un serveur Sun Cluster

Cette section décrit la procédure permettant d'entrer dans la mémoire PROM OpenBoot à partir du concentrateur de terminaux.

▼ Comment entrer dans la mémoire PROM OpenBoot

1. Connectez-vous au port.

```
# telnet nom-concentrateur-terminaux 5002
Tentative avec 192.9.200.1 ...
Connecté à 129.9.200.1.
Caractère d'échappement '^]'.

```

2. Interrompez au besoin l'exécution du logiciel de grappes avec la commande `scadmin stopnode`, puis arrêtez le système.

Arrêtez le système en douceur avec la commande `halt(1M)`.

```
# halt

```

Si vous ne pouvez arrêter le système avec la commande `halt(1M)`, passez au mode de commande `telnet(1)`. Le caractère d'échappement `telnet(1)` par défaut est `Ctrl-]`.

3. Envoyez une interruption au noeud.

```
telnet> send brk

```

4. Exécutez les commandes de la mémoire PROM OpenBoot.

Dépannage du concentrateur de terminaux

Cette section décrit les techniques de dépannage applicables au concentrateur de terminaux. Consultez également le Chapitre 5 du *Sun Cluster 2.2 Hardware Site Preparation, Planning, and Installation Guide*.

Erreurs d'accès à la configuration des ports

L'affichage du message `connect : Connexion refusée` au moment d'accéder à un port particulier du concentrateur de terminaux avec `telnet(1)` a deux causes possibles :

- Une autre personne utilise déjà le port.
- Le port est mal configuré et n'accepte pas les connexions au réseau.

▼ Comment corriger une erreur d'accès à la configuration d'un port

1. **Etablissez une connexion telnet avec le concentrateur de terminaux sans préciser de port pour le moment. Vous le spécifierez lorsque cela vous sera demandé.**

```
# telnet nom-concentrateur-terminaux
Tentative avec adresse_ip ..
Connecté à 192.9.200.1
Le caractère d'échappement est '^'.
[Appuyer au besoin sur ENTREE pour afficher les invites suivantes]
Rotaries définis :
    cli
- Entrer le nom ou le numéro du port Annex : 2.
```

L'affichage du message suivant indique que le port est utilisé.

```
Port(s) occupé(s), voulez-vous attendre ? (o/n) [o] :
```

L'affichage du message suivant indique que le port est mal configuré.

```
Erreur sur le port 2 : Autorisation refusée.
```

Si le port est utilisé, réinitialisez les connexions au concentrateur de terminaux selon les directives fournies à la section “Réinitialisation d’une connexion au concentrateur de terminaux”, page 146.

Si le port est mal configuré, procédez comme suit :

- a. **Sélectionnez l’interpréteur de ligne de commande (cli) et adoptez l’identité de root (superutilisateur) du concentrateur de terminaux.**

```
Entrer le nom ou le numéro du port Annex : cli
Interpréteur de ligne de commande Annex
* Copyright 1991 Xylogics, Inc.
annex : su
Mot de passe :
```

- b. **En qualité de root (superutilisateur) du concentrateur de terminaux, réinitialisez le mode du port.**

```
annex# admin
Administration Annex MICRO-XL-UX R7.0.1, 8 ports
admin : port 2
admin : set port mode slave
Réinitialiser au besoin le port approprié et le sous-
système Annex ou
réinitialiser Annex pour appliquer les changements.
admin : reset 2
admin :
```

Le port est maintenant correctement configuré.

Pour de plus amples informations sur les commandes d’administration du concentrateur de terminaux, consultez le *Guide de référence générale du concentrateur de terminaux Sun*.

Interruptions aléatoires des connexions au concentrateur de terminaux

Les connexions au concentrateur de terminaux établies par un routeur peuvent subir des interruptions intermittentes. Elles peuvent s’activer pendant un certain temps, puis se désactiver à nouveau. A ce moment, les nouvelles tentatives de connexion au

concentrateur de terminaux dépassent le délai qui leur est imparti. Le concentrateur de terminaux ne montre aucun signe de réinitialisation. Par la suite, un routage requis peut être rétabli, puis disparaître à nouveau. Un dépassement de capacité de la table de routage du concentrateur de terminaux et une perte de la connexion au réseau sont à l'origine de ce problème.

Les connexions établies à partir d'un hôte résidant sur le même réseau que le concentrateur de terminaux ne subissent pas ce genre de défaillance.

Pour corriger la situation, il faut établir un routage par défaut à l'intérieur du concentrateur de terminaux et désactiver le mode `routed`. Vous devez désactiver le mode `routed` pour empêcher la perte du routage par défaut. La procédure ci-dessous vous indique comment faire. Pour de plus amples informations, consultez la documentation fournie avec le concentrateur de terminaux.

Le fichier `config.annex`, créé dans le système de fichiers de la mémoire EEPROM du concentrateur de terminaux, définit le routage à utiliser par défaut. Ce fichier `config.annex` vous permet également de définir des rotaries permettant de remplacer le numéro de port par un nom symbolique. Désactivez le mode `routed` avec la commande `set` du concentrateur de terminaux.

▼ Comment établir un routage par défaut

1. Établissez une connexion `shelltool(1)` avec le concentrateur de terminaux.

```
# telnet nom-concentrateur-terminaux
Tentative avec 192.9.200.2 ...
Connecté à xx-tc.
Caractère d'échappement '^]'.
Rotaries définis :
    cli

- Entrer le nom ou le numéro du port Annex : cli

Interpréteur de ligne de commande Annex
* Copyright 1991 Xylogics, Inc.
```

2. Exécutez la commande `su` et entrez le mot de passe d'administration.

Par défaut, le mot de passe est l'adresse IP du concentrateur de terminaux.

```
annex : su
Mot de passe : mot_passe-administration
```

3. Editez le fichier config.annex.

```
annex# edit config.annex
```

4. Tapez les renseignements mis en évidence dans l'exemple suivant, en remplaçant l'adresse IP par celle de votre routeur par défaut :

```
Ctrl-W : enregistrer et quitter Ctrl-X : quitter Ctrl-F : page suiv. Ctrl-
B : page préc.
%gateway
net default gateway 192.9.200.2 metric 1 active ^W
```

5. Désactivez le mode routed local.

```
annex# admin set annex routed n
Réinitialiser au besoin le port approprié du sous-
système Annex ou
réinitialiser Annex pour appliquer les modifications.
annex#
```

6. Réinitialisez le concentrateur de terminaux.

```
annex# boot
```

L'initialisation du concentrateur de terminaux prend quelques minutes. Pendant toute cette période, vous ne pouvez pas accéder à la console des noeuds Sun Cluster.

Modification des informations CT/PSS

L'installation de Sun Cluster 2.2 requiert des informations sur le concentrateur de terminaux (CT) ou un processeur de services système (PSS) (Sun Enterprise 10000 exclusivement). Ces informations sont stockées dans le fichier de configuration de grappe.

Elles servent dans les cas suivants :

- Interruption forcée de noeuds bloqués.
- Mise en oeuvre d'un verrou de grappe empêchant l'intégration des noeuds partitionnés dans la grappe.

Ces deux mécanismes assurent le maintien de l'intégrité des données dans le cas de grappes à quatre noeuds reliées directement à des périphériques de stockage.

Remarque - Si vous utilisez Solstice DiskSuite, les fonctions `tcmon` et `quorum` sont désactivées, et il n'est pas nécessaire de fournir des informations CT.

La commande `scconf(1m)` vous permet de modifier ces informations dans le fichier de configuration de grappe si, par exemple, des changements sont apportés à cette partie de la configuration matérielle de la grappe.

Pour de plus amples informations sur la modification des informations relatives au CT ou au PSS, consultez le Tableau 8-1 et la page de manuel `scconf(1M)`. Consultez également le Chapitre 5 du *Sun Cluster 2.2 Hardware Site Preparation, Planning, and Installation Guide*.

Remarque - Vous devez exécuter ces commandes sur tous les noeuds de la grappe.

TABLEAU 8-1 Modification des informations relatives à l'hôte pour tous les hôtes associés à un CT ou un PSS

Action	Commande
Remplacer l'adresse IP ou le nom d'un CT	<code>scconf(1m) -t -i nouvelle-adresse-ip ancienne-adresse-ip nom-CT</code>
Changer de mot de passe	<code>scconf(1m) -t -P ancienne-adresse-ip nom-CT</code>
Changer le numéro de port utilisé pour le verrou de grappe (CT exclusivement)	<code>scconf(1m) -t -l nouveau-port ancienne-adresse-ip nom-CT</code>

TABLEAU 8-1 Modification des informations relatives à l'hôte pour tous les hôtes associés à un CT ou un PSS (suite)

▼ Comment modifier les informations relatives à l'hôte

1. Exécutez la commande `scconf -H` pour modifier les informations relatives à un hôte particulier. Par exemple, pour modifier le type d'architecture d'un hôte donné et spécifier la nouvelle adresse IP de son PSS (ou de son CT), utilisez la commande suivante sur tous les noeuds de grappe, où `-d` spécifie la nouvelle architecture (Sun Enterprise 10000) associée à l'hôte, et `-t` spécifie la nouvelle adresse IP ou le nouveau nom d'hôte (`foo-ssp`) du PSS (ou du CT) connecté à l'hôte :

```
# scconf nom_grappe -H foo -d E10000 -t foo-ssp
```

▼ Comment spécifier un numéro de port pour un PSS ou un CT

1. Exécutez la commande `scconf -p` sur tous les noeuds de la grappe pour spécifier un numéro de port pour le PSS (ou le CT) sur la console de l'hôte.

```
# scconf nom_grappe -H nom_hôte -p numéro-port
```

Par exemple :

```
# scconf nom_grappe -H foo -p 10
```

Il est possible de connecter plusieurs hôtes au même CT ; l'option `-H` n'influence que les informations relatives à un hôte particulier.

▼ Comment modifier la configuration d'un CT

1. Exécutez la commande `scconf -t` sur tous les noeuds de la grappe pour modifier la configuration d'un CT particulier du système. Par exemple, pour modifier l'adresse IP du CT, exécutez la commande suivante, où `-i` et `-l` spécifient respectivement la nouvelle adresse IP (129.34.123.52) du concentrateur de terminaux (ou du PSS) concerné, et le nouveau port (8) servant de verrou de protection contre les défaillances :

```
# scconf nom_grappe -t foo-tc -i 129.34.123.52 -l -8
```

Si un concentrateur de terminaux est en utilisation, il faut spécifier un numéro de port de CT inutilisé allant de 2 à n , où n représente le nombre de ports du CT. Si un PSS est en utilisation, il faut spécifier la valeur `-1`.

▼ Comment changer le mot de passe d'un PSS ou d'un CT

1. Exécutez la commande `scconf -P` sur tous les noeuds de la grappe pour spécifier le nouveau mot de passe du PSS (ou du CT).

```
# scconf nom_grappe -t foo-ssp -P
foo-ssp(129.34.123.51) Mot de passe :*****
```

Remarque - Si vous avez modifié le mot de passe d'utilisateur sur le PSS ou le CT, vous devez également signaler ce changement au logiciel Sun Cluster en effectuant cette procédure à partir de chaque noeud de la grappe. Cette procédure s'avère nécessaire pour assurer le bon fonctionnement de la protection contre les défaillances lorsqu'on doit forcer l'interruption d'un noeud défectueux en exécutant la commande "send break" à partir du PSS ou du CT.

Utilisation de médiateurs dans une configuration à deux chaînes

Ce chapitre décrit la fonction de Solstice DiskSuite grâce à laquelle Sun Cluster peut exploiter des services de données à haute disponibilité au moyen de deux chaînes de disques seulement. Les sujets abordés sont énumérés ci-dessous. Pour de plus amples informations sur les concepts et fonctions de Solstice DiskSuite, consultez la documentation de Solstice DiskSuite.

- “Aperçu des médiateurs”, page 157
- “Configuration des médiateurs”, page 159
- “Défaillances traitées à l’aide de médiateurs”, page 160
- “Administration des médiateurs”, page 163

Aperçu des médiateurs

Avec Sun Cluster, une chaîne double, c’est-à-dire une configuration à deux chaînes seulement, doit pouvoir rester fonctionnelle, sans intervention de l’utilisateur, lorsqu’un noeud ou une chaîne d’unités tombe en panne.

Dans une configuration à deux chaînes, les répliques de base de données d’état des métapériphériques sont toujours disposées de telle sorte que la moitié exactement des répliques se trouvent sur une chaîne, et l’autre moitié sur la seconde chaîne. Un quorum (la moitié + 1 ou plus) de répliques est nécessaire pour que les données les plus récentes soient présentées. Lorsque, dans une configuration à deux chaînes, l’une des chaînes est indisponible, un quorum de répliques ne sera pas disponible.

Un médiateur est un hôte (noeud) où sont stockées les données de médiateur. Les données de médiateur fournissent des informations sur l’emplacement des autres

médiateurs et contiennent un compteur de validations identique à celui stocké dans les répliques de base de données. Ce compteur est utilisé pour confirmer la synchronisation des données de médiateur avec les données des répliques de base de données. Les données de médiateur sont vérifiées individuellement avant d'être utilisées.

Solstice DiskSuite nécessite un quorum de répliques (la moitié + 1) pour identifier les moments pendant lesquels les conditions d'exploitation sont "sûres". L'intégrité des données est ainsi assurée. Dans une configuration à deux chaînes, il arrive qu'une seule chaîne soit accessible. Dans un tel cas, il est impossible d'obtenir un quorum de répliques. Si des médiateurs sont utilisés et qu'un quorum de médiateurs est présent, les données de médiateur permettent parfois de déterminer si les données de la chaîne accessible sont à jour et qu'elles peuvent être utilisées sans aucun risque.

L'utilisation de médiateurs permet à Sun Cluster de veiller à ce que la plupart des données les plus récentes soient présentées en cas de défaillance d'une seule chaîne dans une configuration à deux chaînes.

Médiateurs or

Dans certains scénarios de défaillance de la configuration à deux chaînes, le concept d'un médiateur or a été mis en oeuvre afin d'éviter toute intervention inutile de la part de l'utilisateur. Si exactement la moitié des répliques de base de données sont accessibles et qu'un événement entraîne la mise à jour des hôtes médiateurs, deux tentatives de mise à jour des médiateurs sont effectuées. Lors de la première mise à jour, il y a tentative de modifier le nombre de validations et de définir un médiateur non or. La deuxième mise à jour n'a lieu que si, au cours de la première phase, tous les hôtes médiateurs ont été contactés avec succès et que le nombre de répliques accessibles (et dont le nombre de validations a été augmenté) correspond exactement à la moitié du nombre total de répliques. Si toutes les conditions sont satisfaites, la deuxième mise à jour attribue l'état or aux médiateurs. Cet état permet l'exécution de la relève de l'hôte or sans intervention de l'utilisateur. Si l'état or n'est pas attribué, les données passent en mode de lecture seule, et l'utilisateur doit intervenir pour assurer la relève ou la reprise. Pour que l'utilisateur puisse lancer une relève ou une reprise, il faut qu'exactly la moitié des répliques soient accessibles.

L'état or n'est stocké que dans la mémoire vive (RAM) non rémanente. Une fois la relève effectuée, les données de médiateur sont de nouveau mises à jour. Si l'un des hôtes médiateurs ne peut pas être mis à jour, l'état or est annulé. Comme l'état est en mémoire RAM seulement, la réinitialisation d'un hôte médiateur entraîne l'annulation de l'état or. L'état par défaut pour les médiateurs est non or.

Configuration des médiateurs

Figure 9-1 illustre un système Sun Cluster à configuration à deux chaînes et médiateurs sur deux noeuds Sun Cluster.

La grappe ne contient toujours que deux hôtes médiateurs, quel que soit le nombre de noeuds. Les hôtes médiateurs sont les mêmes pour tous les ensembles de disques utilisant des médiateurs dans une grappe donnée, même lorsqu'un hôte médiateur n'est pas membre de l'ensemble de serveurs pouvant maîtriser l'ensemble de disques.

Pour simplifier la présentation, les configurations illustrées ici n'utilisent qu'un seul ensemble de disques et une configuration symétrique. Dans ces scénarios, le nombre d'ensembles de disques n'a pas d'importance. Lorsque son état est stable, l'ensemble de disques est sous la maîtrise de `phys-hahost1`.

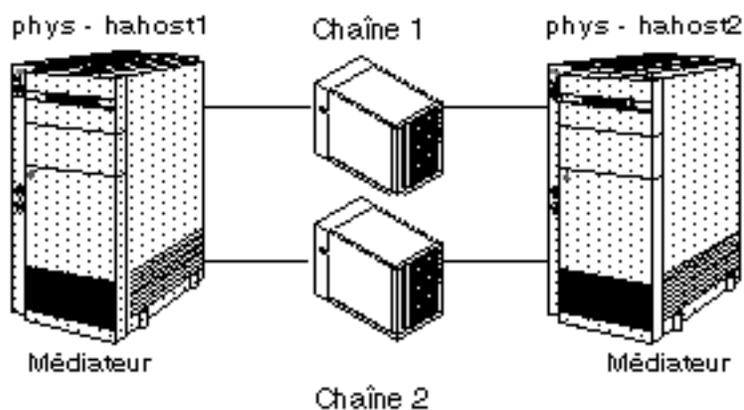


Figure 9-1 Sun Cluster Système en état stable avec médiateurs

En temps normal, les médiateurs ne sont pas utilisés lorsque la moitié + 1 des répliques de base de données sont accessibles. Lorsqu'exactly la moitié des répliques sont accessibles, le compteur de validations du médiateur peut être utilisé pour déterminer si la moitié accessible est la plus récente. Pour que le bon nombre de validations soit utilisé, il faut que les deux médiateurs soient accessibles ou encore que le médiateur soit de type `or`. La moitié + 1 des médiateurs constitue un *quorum de médiateurs*. Le quorum de médiateurs n'est pas tributaire du quorum de répliques.

Défaillances traitées à l'aide de médiateurs

Avec les médiateurs, il est possible d'effectuer une reprise lors de pannes simples et de certaines pannes doubles. Comme Sun Cluster n'assure la reprise automatique qu'en cas de défaillance simple, c'est ce type de panne qui sera traité en détails dans les paragraphes qui suivent. Des scénarios de pannes doubles sont également présentés, mais seuls les processus de reprise généraux sont décrits.

Figure 9-1 illustre une configuration à deux chaînes dont l'état est stable. Il faut noter que des médiateurs sont établis sur les deux noeuds Sun Cluster. Par conséquent, les deux noeuds doivent fonctionner pour qu'un quorum de médiateurs existe et que des médiateurs puissent être utilisés. Si l'un des noeuds Sun Cluster tombe en panne, il y a alors quorum de répliques. Lorsqu'une relève de l'ensemble de disques est nécessaire, celle-ci est effectuée sans l'aide des médiateurs.

Les sections qui suivent présentent divers scénarios de défaillance et décrivent le rôle des médiateurs dans la reprise.

Panne sur un serveur unique

Figure 9-2 illustre une situation de défaillance d'un noeud Sun Cluster. Dans ce cas, le logiciel médiateur n'est pas utilisé, car un quorum de répliques est disponible. Le noeud `phys-hahost2` Sun Cluster deviendra le maître de l'ensemble de disques auparavant sous la maîtrise de `phys-hahost1`.

Le processus de reprise dans ce scénario est identique à celui exécuté lors de la défaillance d'un noeud Sun Cluster dans une configuration à plus de deux chaînes de disques. L'administrateur ne doit intervenir que s'il faut commuter l'ensemble de disques après que `phys-hahost1` eut réintégré la grappe. Pour de plus amples informations sur la procédure de commutation, consultez la page de manuel `haswitch(1M)`.

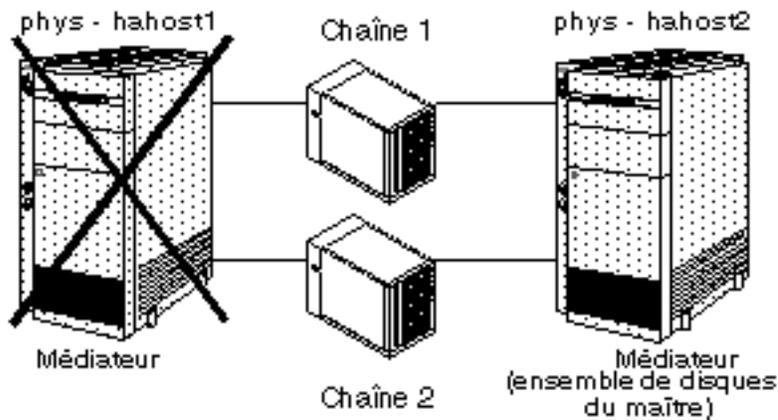


Figure 9-2 Défaillance d'un serveur Sun Cluster unique avec médiateurs

Défaillance d'une seule chaîne

Figure 9-3 illustre la défaillance d'une seule chaîne dans une configuration auparavant stable (voir Figure 9-1). Lorsque la chaîne 1 tombe en panne, les hôtes médiateurs de `phys-hahost1` et `phys-hahost2` sont mis à jour de façon à refléter cet événement, et le système continue de fonctionner, comme suit :

- Aucune relève n'est effectuée.
- Le noeud `phys-hahost1` Sun Cluster reste propriétaire de l'ensemble de disques.
- Comme la chaîne 1 a subi une défaillance, elle doit être resynchronisée par la chaîne 2. Pour de plus amples informations sur le processus de resynchronisation, consultez le *Guide de l'utilisateur de Solstice DiskSuite* et la page de manuel `metareplace(1M)`.

Le compteur de validations est incrémenté, et les médiateurs or sont conservés tels quels.

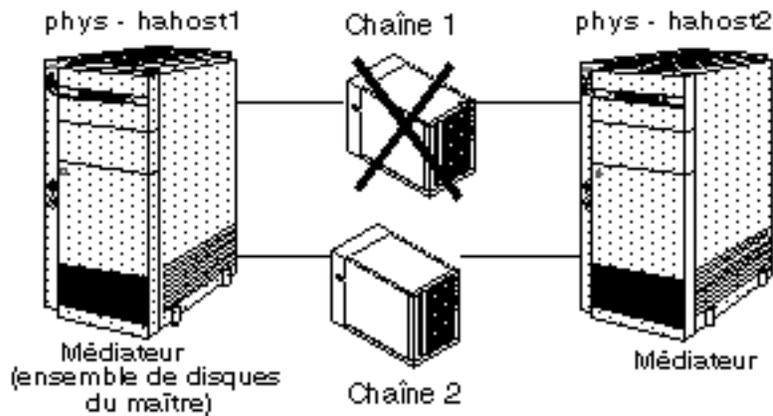


Figure 9-3 Défaillance d'une chaîne unique avec médiateurs

Dans ce scénario, l'administration nécessaire est la même que dans le cas d'une défaillance de chaîne unique dans une configuration à trois chaînes ou plus. Pour plus de détails sur ces procédures, reportez-vous au chapitre pertinent sur l'administration des unités d'expansion de disque.

Défaillance d'un hôte et d'une chaîne

Figure 9-4 illustre un cas de défaillance double où les deux chaînes 1 et `phys-hahost2` tombent en panne. Si la chaîne est tombée en panne en premier, suivie de l'hôte, le médiateur de `phys-hahost1` peut être `or`. Dans un tel cas, les conditions sont les suivantes :

- L'hôte `phys-hahost1` contient un médiateur `or`.
- La moitié des médiateurs sont disponibles.
- La moitié des répliques sont accessibles.
- Le compteur de validations affiché par le médiateur de `phys-hahost1` est identique au compteur de validations dans les répliques de la chaîne 2.

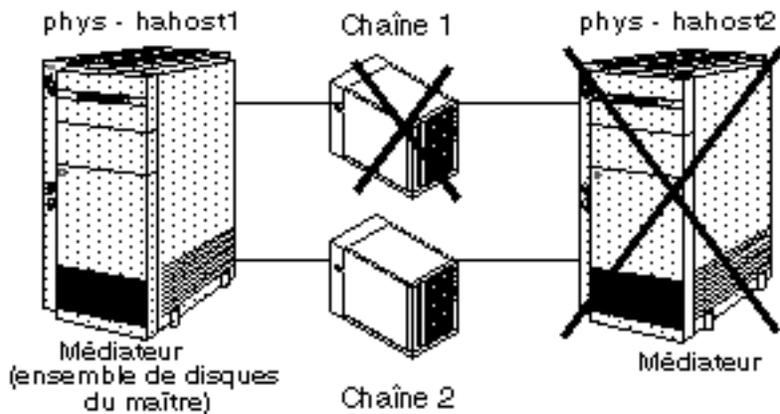


Figure 9-4 Défaillance multiple – Un serveur et une chaîne

Ce type de panne entraîne une reprise automatique par Sun Cluster. Si `phys-hahost2` était maître de l'ensemble de disques, c'est `phys-hahost1` qui devient maître de cet ensemble. Dans le cas contraire, `phys-hahost1` reste le maître de l'ensemble de disques. Après réparation de la chaîne 1, les données de cette chaîne doivent être resynchronisées avec celles de la chaîne 2. Pour de plus amples informations sur le processus de resynchronisation, consultez le *Guide de l'utilisateur de Solstice DiskSuite* et la page de manuel `metareplace(1M)`.



Attention - Bien qu'il soit possible d'effectuer une reprise dans un tel scénario, vous devez alors veiller à restaurer les composants défectueux immédiatement, puisqu'une troisième défaillance rendra la grappe indisponible.

S'il n'y a pas de médiateur or sur `phys-hahost1`, il n'y a pas de reprise automatique par Sun Cluster, et une intervention de l'administrateur est alors nécessaire. Dans ce cas, Sun Cluster génère un message d'erreur et l'hôte logique passe en mode de maintenance (lecture seule). Si une telle situation ou une autre panne multiple survient, cherchez assistance auprès de votre fournisseur de services.

Administration des médiateurs

Les hôtes médiateurs sont administrés au moyen des commandes `medstat(1M)` et `metaset(1M)`. Utilisez ces commandes pour ajouter ou supprimer des hôtes médiateurs, et pour vérifier et réparer les données de médiateur. Pour plus de détails, voir les pages de manuel `medstat(1M)`, `metaset(1M)` et `mediator(7)`.

▼ Comment ajouter des hôtes médiateurs

Effectuez cette procédure après avoir installé et configuré Solstice DiskSuite.

1. Lancez le logiciel de grappe sur tous les noeuds.

Sur le premier noeud :

```
# scadmin startcluster
```

Sur les autres noeuds :

```
# scadmin startnode
```

2. Identifiez le nom du lien privé pour chaque noeud.

Utilisez la commande `grep(1)` pour identifier le lien privé contenu dans le fichier `nom_grappe.cdb`.

```
hahost1# grep ``^cluster.node.0.hostname`` \  
/etc/opt/SUNWcluster/conf/nom_grappe.cdb  
cluster.node.0.hostname : hahost0  
phys-hahost1# grep ``cluster.node.0.hahost0`` \  
/etc/opt/SUNWcluster/conf/nom_grappe.cdb | grep 204  
204.152.65.33  
hahost1# grep ``^cluster.node.1.hostname`` \  
/etc/opt/SUNWcluster/conf/nom_grappe.cdb  
cluster.node.1.hostname : hahost1  
hahost1# grep ``cluster.node.1.hahost1`` \  
/etc/opt/SUNWcluster/conf/nom_grappe.cdb | grep 204  
204.152.65.34
```

Dans cet exemple, 204.152.65.33 est le lien privé pour hahost0 et 204.152.65.34 le lien privé pour hahost1.

3. Configurez les médiateurs à l'aide de la commande `metaset(1M)`.

Ajoutez chaque hôte avec connectivité à l'ensemble de disques comme médiateur pour cet ensemble. Exécutez chaque commande sur l'hôte qui est actuellement maître de l'ensemble de disques. Vous pouvez utiliser la commande `hastat(1M)` pour déterminer le maître actuel de l'ensemble de disques. Les informations produites par `hastat(1M)` pour l'hôte logique identifient le maître de l'ensemble de disques.

```
hahost1# metaset -s ensemble_disquesA -a -m hahost0,204.152.65.33
hahost1# metaset -s ensemble_disquesA -a -m hahost1,204.152.65.34
hahost1# metaset -s ensemble_disquesB -a -m hahost0,204.152.65.33
hahost1# metaset -s ensemble_disquesB -a -m hahost1,204.152.65.34
hahost1# metaset -s ensemble_disquesC -a -m hahost0,204.152.65.33
hahost1# metaset -s ensemble_disquesC -a -m hahost1,204.152.65.34
```

La commande `metaset(1M)` traite le lien privé comme un alias.

▼ Comment vérifier l'état des données de médiateur

1. Exécutez la commande `medstat(1M)`.

```
phys-hahost1# medstat -s ensemble_disques
```

Consultez la page de manuel `medstat(1M)` pour savoir comment interpréter la sortie. Si la sortie indique que les données de médiateur pour l'un ou l'autre des hôtes médiateurs d'un ensemble de disques donné sont erronées, procédez comme expliqué maintenant pour remédier au problème.

▼ Comment corriger des données de médiateur erronées

Remarque - La commande `medstat(1M)` permet de vérifier l'état des médiateurs. Effectuez cette procédure si la commande `medstat(1M)` signale un hôte médiateur défectueux.

1. **Supprimez les hôtes médiateurs défailants de tous les ensembles de disques touchés.**

Connectez-vous au noeud Sun Clusterpropriétaire de l'ensemble de disques touché et tapez :

```
phys-hahost1# metaset -s ensemble_disques -d -m hôte_mediateur_défectueux
```

2. **Restaurez l'hôte médiateur et ses alias :**

```
phys-hahost1# metaset -s ensemble_disques -a -m
hôte_médiateur_défectueux, alias_hôte_physique, ...
```

Remarque - Les liens privés doivent être attribués comme alias d'hôte médiateur. Spécifiez d'abord l'adresse IP de l'hôte physique et ensuite le lien privé HA sur la ligne de commande `metaset (1M)`. Consultez la page de manuel `mediator(7)` pour plus de détails sur l'utilisation de la commande `metaset (1M)`.

Gestion des défaillances sans reprise automatique

Dans certains cas de défaillancedouble, il ne peut y avoir de reprise automatique par Sun Cluster. Ces scénarios sont les suivants :

- Panne d'un noeud et d'une chaîne dans une configuration à deux chaînes en l'absence de médiateur or sur le noeud fonctionnel. Ce scénario est décrit plus en détails dans la "Défaillance d'un hôte et d'une chaîne", page 162.
- Données de médiateur erronées, non valides ou inexistantes sur l'un des noeuds ou les deux et sur l'une des chaînes lors d'une défaillance d'une configuration à deux chaînes. La tentative suivante d'acquisition des hôtes logiques échouera.
- Panne d'une chaîne dans une configuration à deux chaînes alors que le nombre de répliques intactes sur la chaîne fonctionnelle ne représente pas au moins la moitié du total des répliques de l'ensemble de disques défaillant. Lorsque DiskSuite tente de nouveau de mettre à jour ces répliques, une erreur système grave se produit.
- Une défaillance sans reprise automatique s'est produite, et il y a eu tentative de désactiver l'état de maintenance du ou des hôtes logiques affectés avant la fin de l'exécution des procédures de reprise manuelle.

Il est très important de vérifier régulièrement l'état des ensembles de disques, des répliques et des médiateurs. La commande `medstat(1M)` est utile à cette fin. Les données de médiateur, les répliques et les disques erronés doivent toujours être réparés sur-le-champ pour éviter toute complication dans les cas de pannes multiples.

Quand une défaillance de ce type se produit, l'une des séries suivantes de messages d'erreur est consignée :

```
ERREUR : metaset -s sortie de <ensemble_disques> -f -t avec code 66
ERREUR : base de données non valide pour ensemble <ensemble_disques>
AVIS : ensemble <ensemble_disques> libéré
ERREUR : metaset -s sortie de <ensemble_disques> -f -t avec code 2
```

(suite)

(suite)

```
ERREUR : données étiquetées pour ensemble <ensemble_disques>
AVIS : ensemble <ensemble_disques> libéré
ERREUR : metaset -s sortie de <ensemble_disques> -f -t avec code 3
ERREUR : seulement 50 % des répliques et 50 % des hôtes médiateurs
        sont disponibles pour <ensemble_disques>
AVIS : ensemble
<ensemble_disques> libéré
```

Les messages suivants finissent également par être affichés :

```
ERREUR : impossible de devenir propriétaire des hôtes logiques <hôte>, passage
        au mode de maintenance
ERREUR : l'état d'un hôte logique en mode de maintenance ne peut être modifié
        que par intervention manuelle de l'administrateur
ERREUR : l'administrateur doit trouver l'origine du problème et le corriger et,
        au besoin, utiliser la commande haswitch pour désactiver l'état de
maintenance des hôtes logiques
```

Il faut noter que dans le cas d'une défaillance double de ce type, les objectifs de haute disponibilité sont sacrifiés au profit du maintien de l'intégrité des données. Il est possible que les données ne soient pas disponibles pendant un certain temps. En outre, il n'est pas possible de garantir complètement la récupération ou l'intégrité des données.

Dans un tel cas, vous devez communiquer sur-le-champ avec votre fournisseur de services. Toute tentative de reprise manuelle pour ce type de panne double ne doit être effectuée que par un représentant autorisé. Des efforts bien planifiés et concertés sont nécessaires pour assurer la récupération des données. Ne faites rien avant l'arrivée du représentant.

Votre fournisseur examinera les messages consignés, évaluera le problème et effectuera si possible la réparation des éléments matériels endommagés. Votre fournisseur pourra ensuite tenter d'accéder aux données à l'aide de certaines des options `metaset(1M)` spéciales décrites à la page de manuel `mediator(7)`. Ces options doivent toujours être utilisées avec la plus grande prudence afin d'éviter la récupération des mauvaises données.



Attention - Ne tentez jamais d'alterner l'accès entre les deux chaînes. Cela ne ferait qu'aggraver la situation.

Avant de restaurer l'accès client aux données, exécutez toujours toutes les procédures de validation disponibles sur l'ensemble de données en entier ou sur les données touchées par les transactions effectuées récemment sur cet ensemble.

Avant d'exécuter la commande `haswitch(1M)` pour désactiver le mode de maintenance des hôtes logiques, veillez à libérer la propriété de l'ensemble de disques associé.

Messages d'erreur relatifs aux médiateurs

Les messages de consignation système ou de console signalent un problème de médiateurs ou de données de médiateur. Utilisez la procédure "Comment corriger des données de médiateur erronées", page 165 pour traiter les problèmes.

```
Attention : medstat indique des données de médiateur erronées sur l'hôte %s
           pour les ensembles de disques %s

Attention : medstat a détecté une erreur fatale dans les données de médiateur
           sur l'hôte %s de l'ensemble de disques %s !
Attention : échec de medstat pour l'ensemble de disques %s
```

Administration de disques locaux Sun Cluster

Ce chapitre présente les instructions permettant d'administrer les disques locaux Sun Cluster. Certaines des procédures décrites ici dépendent du logiciel de gestion de volumes (Solstice DiskSuite ou VxVM). Lorsqu'une procédure est tributaire du gestionnaire de volumes, son titre inclut le nom de ce gestionnaire.

Ce chapitre présente les rubriques suivantes :

- "Restauration d'un disque d'amorçage local à partir d'une copie de sauvegarde", page 170
- "Remplacement d'un disque local qui ne sert pas à l'amorçage", page 173

L'administration de Sun Cluster comprend la surveillance de l'état de la configuration. Pour plus de détails sur les méthodes de surveillance, voir le Chapitre 2. Le processus de surveillance peut révéler des anomalies des disques locaux. Les sections qui suivent présentent des instructions qui permettent de corriger ces problèmes.

Pour les méthodes d'administration de disque multihôtes, voir le chapitre sur l'administration de l'unité d'expansion de disque en question. Consultez également la documentation du logiciel gestionnaire de volumes en cas de remplacement ou de réparation du matériel de la configuration Sun Cluster.

Restauration d'un disque d'amorçage local à partir d'une copie de sauvegarde

Dans certains cas, il faut remplacer le disque d'amorçage d'un noeud de grappe, par exemple si l'état du disque d'amorçage est inconnu en raison d'un problème logiciel, si la mise à niveau du système d'exploitation échoue ou si un problème matériel survient. Ayez recours aux méthodes suivantes pour restaurer le disque d'amorçage à un état connu, ou encore pour remplacer le disque.

Remarque - Ces procédures supposent que l'on dispose d'une copie de sauvegarde du disque d'amorçage.

▼ Comment restaurer un disque d'amorçage local à partir d'une copie de sauvegarde (Solstice DiskSuite)

Lorsque les hôtes physiques se trouvent dans la même grappe, cette procédure est effectuée sur l'hôte local, pendant qu'un autre hôte offre les services de données pour tous les hôtes. Dans cet exemple, nous utilisons deux hôtes physiques, `phys-hahost1` et `phys-hahost2`, et deux hôtes logiques, `hahost1` et `hahost2`.

Les étapes suivantes permettent de restaurer un disque d'amorçage à partir d'une copie de sauvegarde, dans une configuration Solstice DiskSuite.

- Retrait de l'hôte qui contient le disque d'amorçage des ensembles de disques.
- Restauration du disque d'amorçage à partir d'une copie de sauvegarde.
- Renouvellement ou création de répliques sur le disque restauré.
- Rajout de l'hôte aux ensembles de disques.
- Lancement de Sun Cluster sur cet hôte.
- Passage de l'hôte logique à son maître par défaut (si le mode manuel est configuré en fonction d'une commutation).

Les étapes expliquées maintenant décrivent en détails la restauration d'un disque d'amorçage à partir d'une copie de sauvegarde, dans une configuration Solstice DiskSuite. Dans cet exemple, `phys-hahost1` contient le disque à restaurer. Le disque d'amorçage n'est pas mis en miroir.

1. Arrêtez l'hôte à restaurer.

2. **Sur les autres hôtes de la grappe, utilisez la commande `metaset(1M)` afin de retirer des ensembles de disques l'hôte en cours de restauration.**

Dans cet exemple, la commande `metaset(1M)` est exécutée depuis l'autre hôte de la grappe, soit `phys-hahost2`.

```
phys-hahost2# metaset -s hahost1 -f -d -h phys-hahost1
phys-hahost2# metaset -s hahost2 -f -d -h phys-hahost1
```

3. **Restaurez le disque d'amorçage sur l'hôte en cours de restauration à partir du support de sauvegarde.**

Suivez la procédure de restauration des fichiers et systèmes de fichiers présentée dans le Guide de l'administrateur du système Solaris, afin de restaurer le système de fichiers du disque d'amorçage.

4. **Réinitialisez l'hôte en cours de restauration.**

5. **Retirez les anciennes répliques de DiskSuite puis réinitialisez.**

Si vous remplacez un disque défectueux, les anciennes répliques ne sont pas présentes. Si vous restaurez un disque, exécutez la commande `metadb(1M)` afin de déterminer si des anciennes répliques sont présentes. Si c'est le cas, supprimez-les.

Remarque - L'emplacement par défaut des répliques est la tranche 7. Toutefois, vous n'êtes pas tenu de placer les répliques dans la tranche 7.

```
phys-hahost1# metadb -d -f c0t3d0s7
phys-hahost1# reboot
```

6. **Créez de nouvelles répliques DiskSuite sur le disque restauré au moyen de la commande `metadb(1M)`.**

```
phys-hahost1# metadb -afc 3 c0t3d0s7
```

7. **Ajoutez l'hôte restauré au(x) ensemble(s) de disques, à partir de l'hôte frère.**

```
phys-hahost2# metaset -s hahost1 -a -h phys-hahost1
phys-hahost2# metaset -s hahost2 -a -h phys-hahost1
```

8. Lancez Sun Cluster sur l'hôte restauré.

```
phys-hahost1# scadmin startnode
```

9. Refaites passer les hôtes logiques au maître par défaut, au besoin.

Si le mode manuel n'est pas défini, une commutation automatique aura lieu.

```
phys-hahost1# haswitch phys-hahost1 hahost1
```

▼ Comment restaurer un disque d'amorçage local à partir d'une copie de sauvegarde (VxVM)

Lorsque les hôtes physiques se trouvent dans la même grappe, cette procédure est effectuée sur l'hôte local, pendant qu'un autre hôte fournit les services de données à tous les hôtes. Dans cet exemple, nous utilisons deux hôtes physiques, `phys-hahost1` et `phys-hahost2`, et deux hôtes logiques, `hahost1` et `hahost2`. Dans cet exemple, le disque d'amorçage n'est pas mis en miroir.

Les étapes suivantes permettent de restaurer un disque d'amorçage à partir d'une copie de sauvegarde, dans une configuration VxVM.

- Interruption de l'hôte à restaurer.
- Restauration du disque d'amorçage à partir d'une copie de sauvegarde.
- Lancement de Sun Cluster sur cet hôte.
- Passage de l'hôte logique à son maître par défaut (si le mode manuel est configuré en fonction d'une commutation).

Les étapes expliquées maintenant décrivent en détails la restauration d'un disque d'amorçage à partir d'une copie de sauvegarde, dans une configuration VxVM. Dans cet exemple, `phys-hahost1` contient le disque à restaurer.

1. Arrêtez l'hôte à restaurer.
2. Restaurez le disque d'amorçage sur l'hôte en cours de restauration à partir du support de sauvegarde.

Suivez la procédure de restauration des fichiers et systèmes de fichiers présentée dans le Guide de l'administrateur du système Solaris, afin de restaurer le système de fichiers du disque d'amorçage.

3. Réinitialisez l'hôte en cours de restauration.

Suite à la réinitialisation, l'hôte repère tous les périphériques.

Remarque - Si les disques sont réservés, il peut être nécessaire d'exécuter la commande `vxdctl enable` ultérieurement, une fois les réservations annulées.

4. Lancez Sun Cluster sur l'hôte local.

```
phys-hahost1# scadmin startnode
```

5. Refaites passer les hôtes logiques au maître par défaut, au besoin.

Si le mode manuel n'est pas défini, une commutation automatique aura lieu.

```
phys-hahost1# haswitch phys-hahost1 hahost1
```

Remplacement d'un disque local qui ne sert pas à l'amorçage

Cette section décrit le remplacement d'un disque local défectueux qui ne contient pas l'environnement d'exploitation Solaris.

En général, si un disque local ne servant pas à l'amorçage tombe en panne, vous pouvez effectuer une reprise à partir d'une copie de sauvegarde, afin de restaurer les données sur un nouveau disque.

Les procédures de restauration d'un disque d'amorçage local sont présentées aux sections "Comment restaurer un disque d'amorçage local à partir d'une copie de sauvegarde (Solstice DiskSuite)", page 170 et "Comment restaurer un disque d'amorçage local à partir d'une copie de sauvegarde (VxVM)", page 172.

Ces étapes visent à remplacer un disque local défectueux qui ne sert pas à l'amorçage.

- (Facultatif) Arrêt de Sun Cluster sur le noeud comportant le disque défectueux, puis arrêt de ce noeud.

- Remplacement du disque.
- Formatage et partitionnement du nouveau disque.
- Restauration des données à partir d'une copie de sauvegarde.
- Lancement de Sun Cluster sur cet hôte.
- Commutation de l'hôte logique à son maître par défaut (si le mode manuel est configuré en fonction d'une commutation).

▼ Comment remplacer un disque local qui ne sert pas à l'amorçage

Les étapes expliquées maintenant décrivent en détails le remplacement d'un disque local, défectueux, ne servant pas à l'amorçage. Dans cet exemple, `phys-hahost2` contient le disque défectueux.

1. (Facultatif) Arrêtez les services Sun Cluster sur le noeud qui comporte le disque défectueux, puis arrêtez le noeud.

Vous pouvez sauter cette étape si le noeud s'amorce à partir d'un disque SPARCstorage Array. Toutefois, si le disque à remplacer se trouve sur le même bus SCSI que le disque d'amorçage opérationnel, vous devez arrêter Sun Cluster ainsi que le noeud.

```
# scadmin stopnode
...
# halt
```

2. Remplacez le disque.

Suivez la procédure décrite dans le manuel d'entretien du noeud Sun Cluster.

3. Lancez le noeud en mode mono-utilisateur.

4. Exécutez la commande `format(1M)` ou `fmthard(1M)` pour partitionner le nouveau disque.

Veillez à ce que la partition du nouveau disque soit identique à celle du disque remplacé. (L'enregistrement des informations sur le format du disque est décrit au Chapitre 1.)

5. Exécutez la commande `newfs(1M)` pour les nouvelles tranches, afin de créer les systèmes de fichiers.

- 6. Exécutez la commande `mount (1M)` afin de monter les systèmes de fichiers appropriés.**

Spécifiez le périphérique et les points de montage de chaque système de fichiers.

- 7. Restaurez les données à partir d'une copie de sauvegarde.**

Pour mener à bien cette étape, suivez les instructions du Guide de l'administrateur du système Solaris.

- 8. Réinitialisez le noeud.**

- 9. Lancez Sun Cluster sur l'hôte local.**

```
phys-hahost1# scadmin startnode
```

- 10. Refaites passer les hôtes logiques au maître par défaut, au besoin.**

Si le mode manuel n'est pas défini, une commutation automatique aura lieu.

```
phys-hahost2# haswitch phys-hahost2 hahost2
```


Administration des tableaux SPARCstorage Array

Ce chapitre fournit les instructions nécessaires à l'administration des tableaux SPARCstorage 100, SPARCstorage Array200 avec plateau, SCSI différentiel et SPARCstorage Array 200 avec plateaux de disques RSM . Consultez également le Chapitre 3 du *Sun Cluster 2.2 Hardware Service Manual*.

Ce chapitre présente les rubriques suivantes :

- “Reprise après une panne de courant”, page 177
- “Réparation d’une connexion SPARCstorage Array”, page 182
- “Ajout d’un tableau SPARCstorage Array”, page 184
- “Administration des plateaux d’un tableau SPARCstorage Array”, page 185
- “Remplacement d’un contrôleur du tableau SPARCstorage Array et modification du nom universel”, page 193
- “Administration des disques SPARCstorage Array”, page 203
- “Administration de la mémoire vive non-volatile (NVRAM) d’un tableau SPARCstorage Array”, page 228

Consultez toujours le manuel d’entretien de SPARCstorage Array ainsi que la documentation du gestionnaire de volumes lorsque vous remplacez ou réparez le matériel SPARCstorage Array utilisé dans le cadre d’une configuration Sun Cluster.

Reprise après une panne de courant

Lorsqu’un tableau SPARCstorage Array cesse d’être alimenté, les opérations d’E/S génèrent des erreurs qui sont détectées par le logiciel de gestion des volumes. Les

erreurs ne sont pas signalées tant que des transactions d'E/S ne sont pas effectuées sur le disque. Un remplacement dynamique peut être effectué si les périphériques affectés sont configurés à cette fin.

La configuration doit normalement faire l'objet d'une surveillance pour que ces événements soient toujours détectés. Pour de plus amples informations sur la surveillance de la configuration, voir le Chapitre 2.

▼ Comment effectuer une reprise après une panne de courant (Solstice DiskSuite)

La liste qui suit énumère les étapes générales d'une reprise après panne de courant sur un tableau SPARCstorage Array dans une configuration Solstice DiskSuite :

- Identification des répliques défectueuses.
- Remise en service des répliques défectueuses.
- Identification des périphériques défectueux.
- Remise en service des périphériques défectueux.
- Resynchronisation des disques.

Les étapes expliquées maintenant décrivent en détail la reprise après panne de courant sur un tableau SPARCstorage Array dans une configuration Solstice DiskSuite.

1. **Une fois le courant rétabli, exécutez la commande `metadb(1M)` pour identifier les répliques contenant des erreurs.**

```
# metadb -s ensemble_disques
```

2. **Remettez en service les répliques.**

Après une panne de courant, toutes les répliques de base de données d'état des métapériphériques du châssis de tableau SPARCstorage Array affecté passent à l'état d'erreur. Comme la restauration des répliques de base de données d'état des métapériphériques n'est pas automatique, il est plus prudent de récupérer celles-ci dès que le tableau SPARCstorage Array redevient opérationnel. A défaut, une nouvelle panne risque de mettre hors service la plupart des répliques et de provoquer une panne du noyau. Ce comportement de Solstice DiskSuite se produit en général lorsque le nombre de répliques disponibles n'est pas suffisant.

Bien que ces répliques erronées soient restaurées lors de la relève suivante (commande `haswitch(1M)` ou `reboot(1M)`), il est préférable de les remettre en service manuellement en les supprimant puis en les ajoutant de nouveau.

Remarque - Veillez à rajouter le même nombre de répliques que celui qui a été supprimé sur chaque tranche. Vous pouvez supprimer plusieurs répliques simultanément à l'aide de la commande `metadb(1M)`. Si une tranche doit contenir plusieurs exemplaires des répliques, ajoutez ces dernières en exécutant une seule fois la commande `metadb(1M)`, accompagnée de l'indicateur `-c`.

3. Exécutez la commande `metastat(1M)` pour identifier les métapériphériques contenant des erreurs.

```
# metastat -s ensemble_disques
```

4. Remettez en service les métapériphériques défectueux à l'aide de la commande `metareplace(1M)`, laquelle permet la resynchronisation des disques.

```
# metareplace -s ensemble_disques -e miroir composant
```

L'option `-e` permet au composant (tranche) de passer à l'état `Disponible` et lance la resynchronisation.

Les composants remplacés par une réserve dynamique doivent être remplacés en dernier, à l'aide de la commande `metareplace(1M)`. Si la réserve dynamique est remplacée en premier, elle pourrait remplacer un autre sous-miroir défectueux dès qu'elle sera disponible.

Vous ne pouvez resynchroniser qu'un seul composant de sous-miroir (métapériphérique) à la fois. Si tous les composants d'un sous-miroir ont été touchés par la panne de courant, chaque composant doit être remplacé séparément. La resynchronisation d'un disque de 1,05 Go demande environ 10 minutes.

Si plusieurs ensembles de disques ont été touchés par la panne de courant, vous pouvez resynchroniser simultanément les sous-miroirs de chaque ensemble. Ouvrez une session distincte sur chaque hôte et restaurez son ensemble de disques en y exécutant la commande `metareplace(1M)`.

Remarque - Selon le nombre de sous-miroirs et de composants figurant dans ces sous-miroirs, la resynchronisation peut demander beaucoup de temps. Ainsi, pour un seul sous-miroir contenant 30 disques de 1,05 Go, elle peut durer environ cinq heures. Avec une configuration moins complexe composée de cinq sous-miroirs, la resynchronisation est achevée au bout de 50 minutes environ.

▼ Comment effectuer une reprise après une panne de courant (VxVM)

Les pannes de courant peuvent provoquer la déconnexion des unités de disque et donc des périphériques virtuels, qui ne sont alors plus disponibles. Dans un miroir toutefois, le volume reste actif car les périphériques virtuels restants du volume sont toujours disponibles. Il est possible de reconnecter les unités de disque et de remédier à la situation sans arrêter les noeuds de la grappe.

La liste qui suit énumère les étapes générales d'une reprise après panne de courant sur un tableau SPARCstorage Array utilisé dans une configuration VxVM:

- Identification des périphériques virtuels défectueux à l'aide des commandes `vxprint` et `vxdisk`.
- Correction du problème qui a causé la panne de courant.
- Création des entrées `/devices` et `/dev` à l'aide des commandes `drvconfig` et `disks`.
- Analyse de la configuration de disque actuelle.
- Reconnexion des disques qui contenaient des erreurs temporaires.
- Vérification de la présence d'autres erreurs.
- (Facultatif) Dans le cas de groupes de disques partagés, exécution de la commande `vxdg` pour chaque disque dont l'alimentation a été coupée.
- Lancement de la récupération de volume.

Ces étapes expliquent la reprise après panne de courant sur un tableau SPARCstorage Array dans une configuration VxVM.

1. Exécutez la commande `vxprint` pour afficher les périphériques virtuels contenant des erreurs.

Au besoin, spécifiez un groupe de disques avec l'option `-g groupe_disques`.

2. Exécutez la commande `vxdisk` pour identifier les disques contenant des erreurs.

```
# vxdisk list
DEVICE      TYPE      DISK      GROUP      STATUS
...
-           -         c1t5d0    toi        failed was:c1t5d0s2
...
```

3. Corrigez l'erreur qui a causé le problème afin de rétablir l'alimentation électrique de tous les disques en panne.

Vérifiez que les disques tournent avant de poursuivre.

4. Entrez les commandes suivantes sur tous les noeuds de la grappe.

Dans certains cas, les disques doivent être de nouveau reconnus par les noeuds.

```
# drvconfig
# disks
```

5. Entrez les commandes suivantes sur tous les noeuds de la grappe.

VxVM doit lire de nouveau la configuration actuelle du disque.

```
# vxdctl enable
# vxdisk -a online
```

6. Entrez la commande suivante sur tous les noeuds de la grappe.

Remarque - Si vous utilisez la fonction grappe de VxVM (utilisée avec Oracle Parallel Server), entrez d'abord la commande sur le noeud maître, puis ensuite seulement sur les noeuds esclaves.

Les disques qui contenaient des erreurs temporaires sont alors reconnectés.

```
# vxreattach
```

7. Vérifiez la sortie de la commande `vxdisk` pour repérer toute autre erreur éventuelle.

```
# vxdisk list
```

S'il reste des erreurs, exécutez de nouveau la commande `vxreattach` selon la marche à suivre indiquée (voir Etape 6, page 181).

8. Fonction grappe VxVM (OPS) exclusivement : si vous aviez des groupes de disques partagés et si le support a été remplacé à partir du noeud maître,

exécutez une nouvelle fois la commande pour chaque disque ayant été déconnecté.

Le disque physique et le nom d'accès du gestionnaire de volumes de ce disque doivent être reconnectés.

```
# vxvg -g nom-groupe_disques -k adddisk nom_support=nom_accès
```

Les valeurs pour `nom_support` et `nom_accès` figurent à la fin de la sortie de la commande `vxdisk list`.

Exemple :

```
# vxvg -g toi -k adddisk c1t5d0=c1t5d0s2
# vxvg -g toi -k adddisk c1t5d1=c1t5d1s2
# vxvg -g toi -k adddisk c1t5d2=c1t5d2s2
# vxvg -g toi -k adddisk c1t5d3=c1t5d3s2
# vxvg -g toi -k adddisk c1t5d4=c1t5d4s2
```

Vous pouvez également reconnecter les disques à l'aide de la commande `vxdiskadm` ou de l'interface graphique utilisateur.

9. Lancez la restauration du volume à partir du noeud (ou du noeud maître dans le cas de groupes de disques partagés).

```
# vxrecover -bv [-g groupe_disques]
```

10. (Facultatif) Exécutez la commande `vxprint -g` pour visualiser les résultats.

Réparation d'une connexion SPARCstorage Array

L'interruption de la connexion entre un tableau SPARCstorage Array et l'un des hôtes est le plus souvent provoquée par une défaillance d'un câble à fibres optiques, une carte SBus FC/S ou un module FC/OM.

Dès que la défaillance est détectée sur un hôte, ce dernier commence à générer des messages d'erreur. Par la suite, tout accès au tableau SPARCstorage Array engendre

des erreurs supplémentaires. L'hôte se comporte comme si la panne de courant touchait le tableau SPARCstorage Array en entier.

Les opérations d'E/S effectuées depuis les autres noeuds de la grappe ne sont pas touchées par ce type de panne.

Pour établir le diagnostic du problème, examinez l'affichage du tableau SPARCstorage Array. Celui-ci doit indiquer la connexion (A ou B) ayant été interrompue. Testez la carte FC/S et les modules FC/OM conformément aux procédures décrites dans le manuel d'entretien du noeud Sun Cluster afin d'identifier le composant défectueux. Pour effectuer un dépannage du matériel, libérez un noeud Sun Cluster et le tableau SPARCstorage Array qui paraît en panne.

▼ Comment réparer une connexion SPARCstorage Array

1. Préparez le système Sun Cluster comme pour un remplacement de composants.

Selon l'origine du problème, utilisez l'une des procédures suivantes pour préparer le système Sun Cluster.

- Si la défaillance s'est produite sur la carte FC/S ou le module FC/OM d'une carte FC/S, consultez le Chapitre 7 pour savoir comment préparer le système Sun Cluster pour la mise hors tension.
- Si c'est le câble à fibres optiques qui est défectueux, le logiciel de gestion de volumes aura déjà détecté le problème et préparé le système en vue du remplacement du câble.
- Si la défaillance s'est produite sur le module FC/OM du tableau SPARCstorage Array, exécutez la procédure intitulée "Comment mettre hors service un plateau de tableau SPARCstorage Array (Solstice DiskSuite)", page 186 ou la procédure "Comment mettre hors service un plateau de tableau SPARCstorage Array (VxVM)", page 188 sur chacun des plateaux de SPARCstorage Array de façon à préparer le tableau SPARCstorage Array en entier.

2. Remplacez le composant défectueux.

En cas de défaillance du câble à fibres optiques, de la carte SBus FC/S ou du module FC/OM, consultez le manuel d'entretien du noeud Sun Cluster pour des instructions détaillées sur leur remplacement.

3. Corrigez les erreurs du logiciel de gestion de volumes.

Effectuez les procédures décrites à la section "Reprise après une panne de courant", page 177.

Ajout d'un tableau SPARCstorage Array

Vous pouvez ajouter un tableau SPARCstorage Array à une configuration Sun Cluster à tout moment.

Toutefois, avant d'ajouter un tableau SPARCstorage Array, vous devez passer en revue la configuration des groupes de disques de votre grappe. Pour connaître l'impact du tableau SPARCstorage Array sur cette configuration, reportez-vous aux rubriques concernant la planification de la configuration du manuel *Sun Cluster 2.2 Software Installation Guide*.

▼ Comment ajouter un tableau SPARCstorage Array

1. Arrêtez le noeud de grappe devant recevoir le nouveau tableau SPARCstorage Array.

Exécutez la procédure intitulée “Comment arrêter Sun Cluster sur un noeud de grappe”, page 87 pour arrêter le noeud.

2. Installez la carte FC/S (carte SBus de câble de fibres optiques) dans le noeud.

Installez la carte FC/S conformément aux directives du manuel d'entretien du matériel pour le noeud Sun Cluster.

Remarque - Installez la carte FC/S dans le premier emplacement SBus disponible à la suite des autres cartes du noeud. Cela permet de conserver la numérotation des contrôleurs si l'environnement d'exploitation Solaris est réinstallé. Pour de plus amples informations à ce sujet, voir “Désignation et numérotation des instances”, page 24.

3. Branchez les câbles au tableau SPARCstorage Array et à la carte FC/S.

Suivez les instructions du manuel d'entretien du noeud Sun Cluster.

4. Effectuez une réinitialisation de reconfiguration du noeud.

```
ok boot -r
```

5. Exécutez la commande `haswitch(1M)` pour que le noeud réinitialisé devienne propriétaire de tous les hôtes logiques pouvant être contrôlés par un noeud maître.

```
phys-hahost1# haswitch phys-hahost2 hahost1 hahost2
```

6. Reprenez la procédure Etape 1, page 184 jusqu'à Etape 4, page 184 sur les autres noeuds connectés au tableau SPARCstorage Array.
7. Au besoin, réattribuez au maître par défaut approprié la propriété des hôtes logiques.

```
phys-hahost1# haswitch phys-hahost2 hahost2
```

8. Ajoutez les disques du tableau SPARCstorage Array au(x) groupe(s) de disques sélectionné(s).

Suivez les instructions de la documentation du gestionnaire de volumes pour ajouter les disques au(x) groupe(s) de disques sélectionné(s). Consultez également le manuel *Sun Cluster 2.2 Software Installation Guide* pour plus de renseignements sur Solstice DiskSuite et VxVM.

9. (Configuration Solstice DiskSuite exclusivement) Après avoir ajouté les disques à l'ensemble de disques au moyen de la commande `metaset(1M)`, exécutez la commande `scadmin(1M)` pour réserver et activer le mode failfast sur les disques spécifiés.

```
phys-hahost1# scadmin reserve cN tXdYsZ
```

Administration des plateaux d'un tableau SPARCstorage Array

Cette section décrit les procédures d'administration des plateaux de tableaux SPARCstorage Array. Utilisez les procédures décrites dans le manuel d'entretien du matériel de votre noeud pour identifier le plateau associé au composant défectueux.

Pour éviter toute perte de données ainsi que les défaillances qui pourraient exiger le remplacement du châssis SPARCstorage Array en entier, configurez les miroirs de telle façon que chaque châssis ne contienne qu'un seul sous-miroir.

Remarque - Sun Cluster prend en charge plusieurs modèles de tableaux SPARCstorage Array. Les procédures décrites dans cette section ne s'appliquent qu'aux tableaux SPARCstorage Array 100.

▼ Comment mettre hors service un plateau de tableau SPARCstorage Array(Solstice DiskSuite)

Avant de supprimer un plateau d'un tableau SPARCstorage Array, vous devez interrompre toutes les opérations d'E/S en cours et arrêter tous les disques du plateau. Comme les disques démarrent automatiquement en présence de demandes d'E/S, il faut absolument interrompre ces dernières avant d'arrêter les unités de disque.

La liste qui suit énumère les étapes générales de la mise hors service d'un plateau de tableau SPARCstorage Array dans une configuration Solstice DiskSuite :

- Commutation des hôtes logiques sur un noeud de grappe.
- Interruption des E/S sur le plateau affecté.
- Identification des répliques, des réserves dynamiques et des sous-miroirs sur le plateau affecté.
- Au besoin, vidage de la mémoire vive non-volatile (NVRAM).
- Arrêt et retrait du plateau.

Pour traiter le tableau SPARCstorage Array en entier, exécutez les étapes suivantes sur chaque plateau.

Les étapes expliquées maintenant décrivent en détail la mise hors service d'un plateau de tableau SPARCstorage Array dans une configuration Solstice DiskSuite.

1. **Exécutez la commande `haswitch(1M)` pour que d'autres noeuds deviennent propriétaires des hôtes logiques affectés.**

```
phys-hahost1# haswitch phys-hahost1 hahost1 hahost2
```

Il est possible que le tableau SPARCstorage Array à supprimer contienne des disques inclus dans plusieurs hôtes logiques. Si c'est le cas, attribuez à un autre noeud de la grappe la propriété de tous les hôtes logiques contenant des disques utilisant ce plateau. La commande `luxadm(1M)` sera utilisée plus tard pour arrêter les disques. Dans l'exemple qui suit, la commande `haswitch(1M)` a commuté les hôtes logiques sur `phys-hahost1`, permettant ainsi à `phys-hahost2` d'exécuter les fonctions administratives.

2. **Exécutez la commande `metastat(1M)` sur tous les hôtes logiques affectés pour identifier tous les sous-miroirs contenant des tranches sur le plateau à supprimer.**

```
phys-hahost1# metastat -s ensemble_disques
```

3. **Interrompez les E/S aux sous-miroirs dont les composants (tranches) se trouvent sur le plateau affecté.**

Utilisez la commande `metaoffline(1M)` pour cette étape. Cela met le sous-miroir hors ligne. Vous pouvez utiliser la commande `metadetach(1M)` pour interrompre les E/S, mais cela prolonge la durée de la resynchronisation.

Lorsque les sous-miroirs d'un plateau sont mis hors ligne, les miroirs correspondants n'assurent qu'une mise en miroir unidirectionnelle (c'est-à-dire sans redondance des données). (Un miroir tridirectionnel ne présente pas ce type de problème.) Lorsque le miroir est remis en ligne, la resynchronisation est effectuée automatiquement.

Lorsque tous les sous-miroirs affectés sont hors ligne, les E/S sur le plateau sont interrompues.

4. **Exécutez la commande `metadb(1M)` pour identifier les répliques sur le plateau.**

Enregistrez la sortie `metadb(1M)` pour pouvoir l'utiliser lorsque vous remplacerez le plateau.

5. **Exécutez la commande `metahs(1M)` pour identifier les périphériques à réserve dynamique disponibles et les sous-miroirs correspondants.**

Enregistrez la sortie `metahs(1M)` pour pouvoir l'utiliser lorsque vous remplacerez le plateau.

6. **Si la mémoire vive non-volatile (NVRAM) est activée, videz-la en transférant les données qui s'y trouvent sur le contrôleur, le plateau ou le(s) disque(s) approprié(s).**

```
phys-hahost1# luxadm sync_cache chemin
```

Le message de confirmation affiché indique que la mémoire vive non-volatile est maintenant vide. Pour plus de détails sur le vidage de la mémoire NVRAM, reportez-vous à la section "Vidage et purge de la mémoire NVRAM", page 231.

7. **Arrêtez le plateau à l'aide de la commande `luxadm stop`.**

Lorsque le voyant de verrouillage du plateau s'éteint, retirez le plateau et effectuez les corrections nécessaires.

```
phys-hahost1# luxadm stop cl
```

▼ Comment mettre hors service un plateau de tableau SPARCstorage Array(VxVM)

Avant de supprimer un plateau d'un tableau SPARCstorage Array, vous devez interrompre toutes les opérations d'E/S en cours et arrêter tous les disques du plateau. Comme les disques démarrent automatiquement en présence de demandes d'E/S, il faut absolument interrompre ces dernières avant d'arrêter les unités de disque.

La liste qui suit énumère les étapes générales de la mise hors service d'un plateau de tableau SPARCstorage Array dans une configuration VxVM :

- Commutation des hôtes logiques sur un noeud de grappe.
- Identification des objets VxVM sur le plateau affecté.
- Interruption des E/S sur le plateau affecté.
- Au besoin, vidage de la mémoire vive non-volatile (NVRAM).
- Arrêt et retrait du plateau.

Pour traiter le tableau SPARCstorage Array en entier, exécutez les étapes suivantes sur chaque plateau.

Les étapes expliquées maintenant décrivent en détail la mise hors service d'un plateau de tableau SPARCstorage Array dans une configuration VxVM.

1. **Exécutez la commande `haswitch(1M)` pour que d'autres noeuds deviennent propriétaires des hôtes logiques affectés.**

```
phys-hahost1# haswitch phys-hahost1 hahost1 hahost2
```

Il est possible que le tableau SPARCstorage Array à supprimer contienne des disques inclus dans plusieurs hôtes logiques. Si c'est le cas, attribuez à un autre noeud de la grappe la propriété de tous les hôtes logiques contenant des disques utilisant ce plateau. La commande `luxadm(1M)` sera utilisée plus tard pour arrêter les disques. Dans cet exemple, la commande `haswitch(1M)` a commuté les hôtes logiques sur `phys-hahost1`, permettant ainsi à `phys-hahost1` d'exécuter les fonctions administratives.

2. **Identifiez tous les volumes et périphériques virtuels correspondants sur les disques du plateau à mettre hors service.**
 - a. **Trouvez, dans l'adresse physique du périphérique (`cNtNdN`), le numéro du contrôleur et le numéro de la cible.**

Dans l'adresse `c3t2d0`, par exemple, le numéro du contrôleur est 3 et la cible 2.
 - b. **Trouvez, dans la sortie produite par la commande `vxdisk list`, les périphériques VxVM du plateau affecté.**

Si la cible est 0 ou 1, identifiez tous les périphériques dont l'adresse physique commence par cNt0 et cNt1. Si la cible est 2 ou 3, identifiez tous les périphériques dont l'adresse physique commence par cNt2 et cNt3. Si la cible est 4 ou 5, identifiez tous les périphériques dont l'adresse physique commence par cNt4 et cNt5. Voici un exemple illustrant comment la commande `vxdisk` peut être utilisée pour obtenir les informations.

```
# vxdisk -g groupe_disques -q list | egrep c3t2\|c3t3 | nawk '{print $3}'
```

c. Identifiez tous les périphériques virtuels sur les périphériques ci-dessus, à l'aide de la version qui convient (csh, ksh, ou shell Bourne) de la commande suivante.

```
PLLIST='vxprint -ptq -g groupe_disques  
-e '(aslist.sd_dm_name in (''c3t2d0'', ''c3t3d0'', ''c3t3d1'')) && (pl_kstate=ENABLED)' | nawk '{print $2}'
```

Pour csh, la syntaxe est `set PLLIST`. Pour ksh, la syntaxe est `export PLLIST=`. Avec le shell Bourne, vous devez utiliser la commande `export PLLIST` après définition de la variable.

3. Après avoir défini la variable, interrompez les E/S sur les volumes dont les composants (sous-disques) se trouvent sur le plateau.

Veillez à ce que tous les volumes associés à ce plateau soient détachés (configurations en miroir ou RAID5) ou arrêtés (périphériques virtuels simples). Exécutez la commande suivante pour détacher un périphérique virtuel mis en miroir.

```
# vxplex det ${PLLIST}
```

Autre commande permettant de détacher un périphérique virtuel d'un plateau :

```
# vxplex -g groupe_disques -v volume det périphérique_virtuel
```

Pour interrompre les E/S sur des périphériques virtuels simples, démontez les systèmes de fichiers ou interrompez tout accès à la base de données.

Remarque - Les volumes mis en miroir seront quand même actifs, car l'autre moitié du miroir est encore disponible.

4. Si la mémoire vive non-volatile (NVRAM) est activée, videz-la en transférant les données qui s’y trouvent sur le contrôleur, le plateau ou le(s) disque(s) approprié(s). Sinon, passez à Etape 5, page 190.

```
# luxadm sync_cache chemin
```

Le message de confirmation affiché indique que la mémoire vive non-volatile est maintenant vide. Pour plus de détails sur le vidage de la mémoire NVRAM, reportez-vous à la section “Vidage et purge de la mémoire NVRAM”, page 231.

5. Pour retirer le plateau, exécutez la commande `luxadm stop` afin de l’arrêter.

Lorsque le voyant de verrouillage du plateau s’éteint, retirez le plateau et effectuez les corrections nécessaires.

```
# luxadm stop c1
```

▼ Comment remettre en service un plateau de tableau SPARCstorage Array(Solstice DiskSuite)

La liste qui suit énumère les étapes générales de remise en service d’un plateau de tableau SPARCstorage Array dans une configuration Solstice DiskSuite :

- Démarrage des unités de disque.
- Restauration des répliques, des sous-miroirs et des réserves dynamiques.
- Commutation des hôtes logiques sur leurs maîtres par défaut.

Si le tableau SPARCstorage Array a fait l’objet d’un entretien complet, exécutez les étapes suivantes sur chacun des plateaux.

Les étapes expliquées maintenant décrivent en détail la remise en service d’un plateau de tableau SPARCstorage Array dans une configuration Solstice DiskSuite.

1. Si le tableau SPARCstorage Array a été retiré, démarrez les disques du plateau SPARCstorage Array. Sinon, passez à Etape 3, page 191.

Après avoir effectué toutes les opérations voulues sur un plateau SPARCstorage Array, remettez-le en place dans le châssis. Les disques démarrent automatiquement. Si les disques ne démarrent pas, exécutez la commande `luxadm(1M) start` pour les démarrer manuellement. Il y a ensuite un court délai (quelques secondes) entre l’exécution de la commande et le démarrage des disques du tableau SPARCstorage Array. Dans l’exemple qui suit, `c1` est l’ID de contrôleur :

```
phys-hahost1# luxadm start c1
```

2. Ajoutez toutes les répliques de base de données d'état des métapériphériques qui ont été supprimées des disques du plateau.

Utilisez les informations enregistrées à l' Etape 4, page 187 de la procédure "Comment mettre hors service un plateau de tableau SPARCstorage Array(Solstice DiskSuite)", page 186 pour restaurer les répliques de base de données d'état des métapériphériques.

```
phys-hahost1# metadb -s hahost1 -a répliques-supprimées
```

Pour ajouter plusieurs répliques sur la même tranche, utilisez l'option `-c`.

3. Après démarrage des disques, remettez en ligne tous les sous-miroirs qui avaient été mis hors ligne.

Exécutez la commande `metaonline(1M)` qui convient pour les disques du plateau.

```
phys-hahost1# metaonline -s hahost1 d15 d35
phys-hahost1# metaonline -s hahost1 d24 d54
...
```

Lorsque la commande `metaonline(1M)` est exécutée, une resynchronisation optimisée met automatiquement à jour les sous-miroirs. Cette resynchronisation optimisée ne copie que les zones de disque modifiées lorsque le sous-miroir était hors ligne. Cette zone ne représente généralement qu'une petite partie de la capacité du sous-miroir.

Exécutez la commande `metaonline(1M)` autant de fois que nécessaire, jusqu'à ce que tous les sous-miroirs soient de nouveau en ligne.

Remarque - Si vous avez utilisé la commande `metadetach(1M)` pour détacher le sous-miroir plutôt que la commande `metaoffline(1M)`, vous devez synchroniser le miroir en entier avec la commande `metattach(1M)`. Cette opération dure environ 10 minutes par giga-octet de données.

4. Ajoutez de nouveau toutes les réserves dynamiques supprimées au moment de la mise hors service du tableau SPARCstorage Array.

Exécutez la commande `metahs(1M)` qui convient à votre configuration des réserves dynamiques. Utilisez les informations enregistrées à l' Etape 5, page 187 de la procédure "Comment mettre hors service un plateau de tableau SPARCstorage Array(Solstice DiskSuite)", page 186 pour remplacer les réserves dynamiques.

```
phys-hahost1# metahs -s hahost1 -a pool_réserve_dynamiques cNtXdYsZ
```

5. Au besoin, commutez chaque hôte logique sur son maître par défaut.

```
phys-hahost1# haswitch phys-hahost2 hahost2
```

▼ Comment remettre en service un plateau de tableau SPARCstorage Array(VxVM)

La liste qui suit énumère les étapes générales de remise en service d'un plateau de tableau SPARCstorage Array dans une configuration VxVM :

- Démarrage des unités de disque.
- Restauration des objets VxVM.
- Commutation des hôtes logiques sur leurs maîtres par défaut.

Si le tableau SPARCstorage Array a fait l'objet d'un entretien complet, exécutez les étapes suivantes sur chacun des plateaux.

Les étapes expliquées maintenant décrivent la remise en service d'un plateau de tableau SPARCstorage Array dans une configuration VxVM.

1. Si le tableau SPARCstorage Array a été retiré, démarrez les disques du plateau SPARCstorage Array. Sinon, passez à Etape 2, page 192.

Après avoir effectué toutes les opérations voulues sur un plateau SPARCstorage Array, remettez-le en place dans le châssis. Les disques démarrent automatiquement. Si les disques ne démarrent pas, exécutez la commande `luxadm(1M) start` pour les démarrer manuellement. Il y a ensuite un court délai (quelques secondes) entre l'exécution de la commande et le démarrage des disques du tableau SPARCstorage Array. Dans l'exemple qui suit, `c1` est l'ID de contrôleur :

```
phys-hahost1# luxadm start c1
```

2. Une fois que les disques tournent, surveillez le déroulement de la reprise de la gestion des volumes.

Les volumes du plateau précédemment affectés devraient être progressivement remis en ligne, et la restauration des données devrait débuter automatiquement au bout de quelques minutes. Au besoin, exécutez les commandes `vxreattach` et `vxrecover` pour rattacher les disques et effectuer la reprise sur erreur. Pour

de plus amples renseignements, reportez-vous aux pages correspondantes du manuel.

Remarque - Les sous-disques DRL détachés doivent être rattachés manuellement.

3. Au besoin, commutez chaque hôte logique sur son maître par défaut.

```
phys-hahost1# haswitch phys-hahost2 hahost2
```

Remplacement d'un contrôleur du tableau SPARCstorage Array et modification du nom universel

Le contrôleur SPARCstorage Array porte un identificateur unique appelé "nom universel" qui permet au logiciel Solaris de l'identifier comme tel. Par conséquent, lorsque des défaillances du tableau SPARCstorage Array exigent le remplacement du contrôleur ou du châssis contenant le contrôleur, des procédures spéciales doivent être appliquées.

Le nom universel est semblable à l'ID machine stocké dans l>IDPROM de l'hôte d'une machine SPARC. Les quatre derniers chiffres du nom universel du tableau SPARCstorage Array sont affichés à l'écran ACL du châssis. Le nom universel fait partie du chemin `/devices` associé au tableau SPARCstorage Array et à ses disques.

Si vous devez remplacer le contrôleur du tableau SPARCstorage Array ou le châssis en entier, les noeuds Sun Cluster reconnaîtront le nouveau nom universel lors de leur réinitialisation. Pour éviter que le nouveau nom universel ne confonde pas les couches supérieures du logiciel Sun Cluster, remplacez le nom universel du nouveau contrôleur par le nom universel de l'ancien contrôleur. (Cette opération est semblable à la permutation de l>IDPROM lors du remplacement d'une carte système dans une machine SPARC.)

Choisissez la procédure qui convient en fonction de la situation :

- Dans la procédure décrite à la section "Comment modifier le nom universel d'un tableau SPARCstorage Array au moyen d'un système de maintenance", page 194, un système de maintenance distinct est utilisé pour activer le contrôleur à remplacer sans arrêter les noeuds de la grappe.

- Si la défaillance ne touche pas le tableau SPARCstorage Array en entier ou si celui-ci n'est pas remplacé pour une autre raison, préparez le remplacement en exécutant, pour chaque plateau du tableau SPARCstorage Array, les étapes décrites à la section "Administration des plateaux d'un tableau SPARCstorage Array", page 185. Exécutez ensuite la procédure décrite à la section "Comment modifier le nom universel d'un tableau SPARCstorage Array", page 198.
- Si la défaillance touche la totalité du contrôleur SPARCstorage Array, le logiciel de gestion de volumes aura déjà effectué les préparatifs en vue de l'échange. Dans ce cas, vous pouvez utiliser la procédure décrite à la section "Comment modifier le nom universel d'un tableau SPARCstorage Array au moyen d'un système de maintenance", page 194.

▼ Comment modifier le nom universel d'un tableau SPARCstorage Array au moyen d'un système de maintenance

Cette procédure explique comment modifier un contrôleur de tableau SPARCstorage Array et son nom universel par celui du contrôleur défectueux. Elle permet de remplacer le contrôleur d'un tableau SPARCstorage Array sans arrêter les noeuds de la grappe.

Dans cette procédure, le "système de maintenance" utilisé peut être n'importe quelle architecture Sun Microsystems capable de prendre en charge un tableau SPARCstorage Array. La présence d'un système de maintenance permet d'exécuter cette procédure sans arrêter les noeuds de la grappe.

Ce système doit être chargé avec la même version de l'environnement d'exploitation Solaris que celle utilisée pour les noeuds de la grappe, et doit contenir tous les fichiers correctifs applicables. Il doit également être doté d'un lecteur de CD-ROM, d'une carte SBus de canal de fibres optiques (FC/S) et d'un module de canal de fibres optiques (FC/OM). Le système doit comporter les révisions FCODE et matérielles qui conviennent. Comme solution de rechange, vous pouvez initialiser le système de maintenance sur le réseau.

Remarque - Si aucun "système de maintenance" n'est disponible, utilisez à cette fin l'un des noeuds de la grappe en suivant les étapes de cette procédure.

La liste qui suit énumère les étapes générales de la modification du nom universel d'un tableau SPARCstorage Array à l'aide d'un système de maintenance :

- (Facultatif) Si le contrôleur est le périphérique de quorum, utilisation de la commande `scconf(1M)` pour sélectionner un nouveau périphérique de quorum.
- Obtention du nom universel du tableau précédent.

- Débranchement des câbles optiques et remplacement du contrôleur ou du tableau.
- Branchement du câble optique entre le système de maintenance et le nouveau contrôleur.
- Initialisation du système de maintenance avec système “mini-unix” à partir d’un CD Solaris.
- Téléchargement du nom universel d’origine.
- Réinitialisation du SSA.
- Arrêt du système de maintenance.
- Connexion du contrôleur SSA aux noeuds de grappe.
- Vérification du niveau du microprogramme du nouveau contrôleur à partir du noeud de grappe.
- (Facultatif) Au besoin, mise à niveau du microprogramme du nouveau contrôleur à partir du noeud de grappe.
- Mise en ligne du plateau SSA et réactivation de la gestion des volumes.

Les étapes expliquées maintenant décrivent en détail la modification du nom universel d’un tableau SPARCstorage Array au moyen d’un système de maintenance.

1. Si le contrôleur du tableau SPARCstorage Array défaillant est le contrôleur de quorum, sélectionnez un nouveau contrôleur de quorum à l’aide de la commande `scconf(1M)`.

Pour de plus amples renseignements, consultez la page de manuel `scconf(1M)`.

2. Trouvez le nom universel du tableau SPARCstorage Array endommagé.

Si le tableau SPARCstorage Array est hors tension, suivez les instructions ci-dessous pour obtenir ce nom.

Le nom universel est constitué de 12 chiffres hexadécimaux. Ces chiffres font partie du composant de chemin d’accès. Il s’agit des 12 derniers chiffres suivant les caractères `p1n@a0`, à l’exclusion de la virgule. Utilisez la commande `ls(1)` sur un noeud de grappe connecté au SSA pour identifier le nom universel actuel.

```
# ls -l /dev/rdisk/cNt0d0s0
...SUNW,p1n@a0000000,7412bf ...
```

Dans cet exemple, le nom universel du tableau SPARCstorage Array remplacé est `0000007412bf`. La variable `N` dans le nom de périphérique correspond au numéro de contrôleur pour le tableau SPARCstorage Array défaillant. La chaîne “`t0d0s0`” n’est utilisée ici qu’à titre d’exemple. Utilisez un nom de périphérique existant sur le tableau SPARCstorage Array ou encore `/dev/rdisk/cN*` pour désigner tous les périphériques.

Si le tableau SPARCstorage Array est opérationnel, vous pouvez obtenir le nom universel en exécutant la commande `luxadm(1M)`.

Lorsque vous exécutez la commande `luxadm(1M)` avec l'option `display` et spécifiez un contrôleur, tous les renseignements à propos du tableau SPARCstorage Array sont affichés. Le numéro de série retourné par la commande `luxadm(1M)` est le nom universel.

```
# /usr/sbin/luxadm display cN
```

3. Débranchez le câble optique du contrôleur du tableau SPARCstorage Array défectueux.

4. Remplacez ce contrôleur.

Effectuez cette opération conformément aux instructions fournies dans le manuel d'entretien du tableau SPARCstorage Array.

Si le tableau SPARCstorage Array n'est pas entièrement défaillant ou s'il est remplacé pour une raison autre qu'une défaillance du contrôleur, préparez le remplacement en exécutant, pour chaque plateau du tableau SPARCstorage Array, les étapes décrites à la section "Administration des plateaux d'un tableau SPARCstorage Array", page 185.

Si la défaillance touche la totalité du contrôleur SPARCstorage Array, le gestionnaire de volumes aura déjà effectué les préparatifs en vue de l'échange.

5. Branchez le câble optique entre le système de maintenance et le nouveau contrôleur.

6. Accédez à la PROM OpenBoot du système de maintenance et initialisez-la avec "mini-unix".

Effectuez cette opération à partir du CD de distribution (ou son équivalent réseau) pour mettre le système de maintenance en mode mono-utilisateur et pour obtenir une version en mémoire de la structure de périphérique contenant le nouveau nom universel SPARCstorage Array.

```
<#0> ok boot cdrom -s  
ou <#0> ok boot netgel -s
```

Utilisez "mini-unix" pour empêcher la modification permanente des données de périphérique.

7. Exécutez la commande `luxadm download` pour définir le nom universel.

```
# /usr/sbin/luxadm -s -w nom download cN
```

Nom est le nom universel à 12 chiffres du contrôleur remplacé et *N* le numéro du contrôleur à partir de *cNtXdX* dans le nom de périphérique. Vous devez avoir obtenu le nom universel à Etape 2, page 195.

Remarque - Les zéros de gauche font partie du nom universel et doivent être entrés pour obtenir un nom à 12 chiffres.



Attention - N'interrompez pas le téléchargement. Attendez que l'invite du shell soit affichée après l'exécution de la commande `luxadm(1M)`.

8. Lorsque l'invite est de nouveau affichée, réinitialisez le SSA.

La nouvelle adresse devrait apparaître dans la fenêtre du tableau SPARCstorage Array.

9. Arrêtez le système de maintenance.

10. Rattachez le contrôleur SPARCstorage Array aux noeuds de grappe.

11. Vérifiez le niveau de microprogramme SPARCstorage Array sur le noeud de grappe.

Exécutez la commande `luxadm(1M)` pour déterminer la version actuelle du microprogramme. Indiquez le numéro du contrôleur (*N* dans l'exemple) dans la commande `luxadm(1M)`.

```
# /usr/sbin/luxadm display cN
```

Remarque - Si le système Solaris détecte une version ancienne du microprogramme, il affiche un message sur la console et dans `/var/adm/messages` semblable à ce qui suit : `NOTICE: pln0: Old SSA firmware has been detected (Ver:3.11) : Expected (Ver:3.12) - Please upgrade`

12. (Facultatif) Procédez comme suit pour mettre à niveau le microprogramme de contrôleur.

- a. Téléchargez le microprogramme qui convient. Pour plus de détails, consultez le fichier `README` accompagnant le fichier correctif du microprogramme.

```
# /usr/sbin/ssaadm download -f chemin /ssafirmware cN
```

où *chemin* est le chemin d'accès au répertoire où se trouve le microprogramme et *N* le numéro du contrôleur. Exemple :

```
# /usr/sbin/ssaadm download -f /usr/lib/firmware/ssa/ssafirmware cN
```

b. Réinitialisez le tableau SPARCstorage Array en appuyant sur le bouton SYS OK de l'unité.

L'unité est réinitialisée après un court délai.

c. Vérifiez de nouveau le niveau du microprogramme (Etape 11, page 197). Si le niveau du microprogramme ou le nom universel est toujours erroné, répétez l'Etape 12, page 197 en utilisant un autre contrôleur.

13. Débutez la restauration du gestionnaire de volumes.

Reportez-vous à la section "Administration des plateaux d'un tableau SPARCstorage Array", page 185. Attendez que le tableau SPARCstorage Array soit en ligne pour tous les noeuds et que tous les noeuds puissent voir tous les disques.

▼ Comment modifier le nom universel d'un tableau SPARCstorage Array



Attention - Cette procédure ne fonctionne pas si le disque racine est encapsulé par VxVM ou si le disque d'initialisation de l'un des noeuds se trouve sur ce tableau SPARCstorage Array. Dans de tels cas, utilisez la procédure décrite à la section "Comment modifier le nom universel d'un tableau SPARCstorage Array au moyen d'un système de maintenance", page 194.

Remarque - S'il y a défaillance d'un contrôleur de quorum, vous devez sélectionner un nouveau contrôleur de quorum avant d'arrêter un noeud.

La liste qui suit énumère les étapes générales de la modification du nom universel d'un tableau SPARCstorage Array:

- (Facultatif) Si le contrôleur est le périphérique de quorum, utilisation de la commande `scconf(1M)` pour sélectionner un nouveau périphérique de quorum.

- Commutation de la propriété des hôtes logiques à partir du noeud faisant l'objet des réparations ou du contrôleur devant être remplacé.
- Obtention du nom universel du tableau précédent.
- Remplacement du contrôleur ou du tableau.
- Arrêt du logiciel Sun Cluster et du noeud qui n'est pas propriétaire des disques.
- Avec "mini-unix", réinitialisation du noeud qui n'est pas propriétaire des disques.
- Identification du numéro du contrôleur pour le nouveau tableau.
- Définition du nouveau nom universel et reconfiguration du tableau.
- Réinitialisation, au besoin, des autres noeuds de grappe.
- Restauration de la gestion des volumes.

Les étapes expliquées maintenant décrivent en détail la modification du nom universel d'un tableau SPARCstorage Array.

1. **Si le contrôleur du tableau SPARCstorage Array défaillant est le contrôleur de quorum, sélectionnez un nouveau contrôleur de quorum à l'aide de la commande `scconf(1M)`.**

Pour de plus amples informations, consultez la page de manuel `scconf(1M)`.

2. **Sur le noeud de grappe connecté au SSA en cours de réparation, arrêtez le logiciel Sun Cluster et le système.**

Exécutez la commande `scadmin(1M)` pour attribuer la propriété de tous les hôtes logiques aux autres noeuds de la grappe et pour arrêter le logiciel Sun Cluster. Exécutez ensuite la commande `halt(1M)` afin d'arrêter la machine.

Dans cet exemple, `phys-hahost2` est le noeud à partir duquel la procédure de réparation est effectuée.

```
phys-hahost2# scadmin stopnode
...
phys-hahost2# halt
```

3. **Trouvez le nom universel du tableau SPARCstorage Array endommagé.**

Si le tableau SPARCstorage Array est hors tension, suivez les instructions ci-dessous pour obtenir ce nom.

Le nom universel est constitué de 12 chiffres hexadécimaux. Ces chiffres font partie du composant de chemin de périphérique contenant les caractères `p1n@a0`. Il s'agit des 12 derniers chiffres suivant les caractères `p1n@a0`, à l'exclusion de la virgule. Utilisez la commande `ls(1)` sur un noeud de grappe connecté au SSA pour identifier le nom universel actuel.

```
phys-hahost1# ls -l /dev/rdisk/cN t0d0s0
...SUNW,pln@a0000000,7412bf ...
```

Dans cet exemple, le nom universel du tableau SPARCstorage Array remplacé est 0000007412bf. La variable *N* dans le nom de périphérique correspond au numéro de contrôleur pour le tableau SPARCstorage Array défaillant. La chaîne *t0d0s0* n'est utilisée ici qu'à titre d'exemple. Utilisez un nom de périphérique existant sur le tableau SPARCstorage Array ou encore */dev/rdisk/cN** pour désigner tous les périphériques.

Si le tableau SPARCstorage Array est opérationnel, vous pouvez obtenir le nom universel en exécutant la commande `luxadm(1M)`.

Lorsque vous exécutez la commande `luxadm(1M)` avec l'option `display` et spécifiez un contrôleur, tous les renseignements à propos du tableau SPARCstorage Array sont affichés. Le numéro de série retourné par la commande `luxadm(1M)` est le nom universel.

```
phys-hahost1# /usr/sbin/luxadm display cN
```

4. Remplacez le contrôleur ou le tableau SPARCstorage Array.

Effectuez cette opération conformément aux instructions fournies dans le manuel d'entretien du tableau SPARCstorage Array.

Si le tableau SPARCstorage Array n'est pas entièrement défaillant ou s'il est remplacé pour une raison autre qu'une défaillance du contrôleur, préparez le remplacement en exécutant, pour chaque plateau du tableau SPARCstorage Array, les étapes décrites à la section "Administration des plateaux d'un tableau SPARCstorage Array", page 185.

Si la défaillance affecte la totalité du contrôleur SPARCstorage Array, le gestionnaire de volumes aura déjà effectué les préparatifs en vue de ce remplacement.

5. Accédez à la PROM OpenBoot sur le noeud arrêté et initialisez-la avec "mini-unix".

Effectuez cette opération à partir du CD de distribution (ou son équivalent réseau) pour mettre l'hôte en mode mono-utilisateur et obtenir une version en mémoire de la structure de périphérique contenant le nouveau nom universel SPARCstorage Array.

```
<#0> ok boot cdrom -s
ou
<#0> ok boot netqel -s
```

(suite)

Utilisez “mini-unix” pour empêcher la modification permanente des données de périphérique sur le noeud de grappe.

6. Trouvez le numéro de contrôleur pour le nouveau tableau SPARCstorage Array.

Utilisez la commande `ls(1)` et les quatre chiffres affichés à l’écran ACL du nouveau tableau SPARCstorage Array pour identifier le numéro de contrôleur.

Dans cet exemple, les quatre chiffres affichés à l’écran ACL sont 143b. Il est à noter que le nom de périphérique `c*t0d0s0` utilise une correspondance par modèle pour le numéro de contrôleur, mais spécifie une tranche existante connue. Cela réduit le nombre de lignes générées dans la sortie.

```
# ls -l /dev/rdisk/c*t0d0s0 | grep -i 143b
lrwxrwxrwx  1 root    root          98 Mar 14 13:38 /dev/rdisk/
c3t0d0s0 ->
../../../../devices/iommu@f,e0000000/sbus@f,e0001000/SUNW,soc@3,0/
SUNW,pln@a0000000,
74143b/ssd@0,0:a,raw
```

Dans cet exemple, 3 (dans `/dev/rdisk/c3...`) est le numéro de contrôleur du nouveau tableau SPARCstorage Array sous “mini-unix”.

Remarque - Les caractères hexadécimaux dans l’affichage ACL comportent des majuscules (lettres A, C, E et F) et des minuscules (b et d). Dans l’exemple, `grep -i` permet de ne pas tenir compte de la casse dans la comparaison.

7. Exécutez la commande `luxadm download` pour définir le nom universel.

Utilisez le numéro de contrôleur identifié à l’Etape 6, page 201. La commande suivante, par exemple, remplace la valeur de nom universel actuelle par la valeur déterminée à Etape 3, page 199 (0000007412bf). Le contrôleur SPARCstorage Array est le contrôleur 3.

```
phys-hahost2# /usr/sbin/luxadm download -w 0000007412bf c3
```

Remarque - Les zéros de gauche font partie du nom universel et doivent être entrés pour obtenir un nom à 12 chiffres.



Attention - N'interrompez pas le téléchargement. Attendez que l'invite du shell soit affichée après l'exécution de la commande `luxadm(1M)`.

8. Réinitialisez le tableau SPARCstorage Array en appuyant sur le bouton SYS OK de l'unité.

L'unité est réinitialisée au bout d'un court délai et la communication est établie avec les noeuds Sun Cluster.

9. Annulez "mini-unix" et initialisez l'hôte selon la procédure normale.

Envoyez une interruption à la console et initialisez l'ordinateur.

10. Vérifiez le niveau de microprogramme de SPARCstorage Array sur le noeud de grappe.

Exécutez la commande `luxadm(1M)` pour déterminer la version actuelle du microprogramme. Indiquez le numéro du contrôleur (*N* dans l'exemple) dans la commande `luxadm(1M)`.

```
phys-hahost2# /usr/sbin/luxadm display cN
```

Remarque - Si le système Solaris détecte une version ancienne du microprogramme, il affiche un message sur la console et dans `/var/adm/` messages semblable à ce qui suit : NOTICE : pln0 : Old SSA firmware has been detected (Ver:3.11) : Expected (Ver:3.12) - Please upgrade

11. (Facultatif) Procédez comme suit pour mettre à niveau le microprogramme de contrôleur.

a. Téléchargez le microprogramme qui convient. Pour plus de détails, consultez le fichier README accompagnant le fichier correctif du microprogramme.

```
# /usr/sbin/ssaadm download -f chemin/ssafirmware cN
```

où *chemin* est le chemin d'accès au répertoire où se trouve le microprogramme et *N* le numéro du contrôleur. Exemple :

```
# /usr/sbin/ssaadm download -f /usr/lib/firmware/ssa/ssafirmware cN
```

- b. Réinitialisez le tableau SPARCstorage Array en appuyant sur le bouton **SYS OK** de l'unité.

L'unité est réinitialisée après un court délai.

- c. Vérifiez de nouveau le niveau de microprogramme (voir Etape 10, page 202). Si le niveau du microprogramme ou le nom universel est toujours erroné, reprenez Etape 11, page 202 en utilisant un autre contrôleur.

12. Lancez le noeud.

```
phys-hahost2# scadmin startnode
```

13. Refaites passer les hôtes logiques au maître par défaut, au besoin.

14. Achevez le remplacement en restaurant les composants du gestionnaire de volumes sur le tableau SPARCstorage Array réparé.

Cette procédure est décrite dans la section "Administration des plateaux d'un tableau SPARCstorage Array", page 185.

15. Au besoin, réinitialisez les autres noeuds de la grappe.

Vous pourriez avoir à réinitialiser les autres noeuds de la grappe s'ils ne peuvent pas reconnaître tous les disques du tableau SPARCstorage Array après remplacement. Dans un tel cas, utilisez la commande `scadmin stopnode` pour arrêter l'activité de Sun Cluster, puis réinitialisez le système. Au besoin, après la réinitialisation, commutez de nouveau les hôtes logiques sur leurs maîtres par défaut. Pour de plus amples renseignements, consultez la page de manuel `scadmin(1M)`.

Administration des disques SPARCstorage Array

Dans le cadre de l'administration standard de Sun Cluster, il vous incombe de surveiller l'état de la configuration. Pour de plus amples informations sur les méthodes de surveillance, voir le Chapitre 2. Durant le processus de surveillance, vous pouvez découvrir des problèmes avec les disques multihôtes. Les sections qui suivent présentent des instructions qui permettent de corriger ces problèmes.

Sun Cluster prend en charge les types de disques SSA suivants :

- Modèles 100
- Modèles 200 avec plateau SCSI différentiel
- Modèles 200 avec RSM (214 RSM)

Selon le type de disque que vous utilisez et les caractéristiques mécaniques et électriques du boîtier, l'ajout de disques peut exiger la préparation de tous les disques connectés à un contrôleur particulier, de tous les disques d'un plateau particulier ou des disques ajoutés seulement. Ainsi, pour les modèles de tableau SPARCstorage Array 200 avec plateau SCSI différentiel, vous devez préparer le contrôleur de tableau et le boîtier de disque. Pour les modèles SPARCstorage Array 200 avec RSM (214 RSM), seul le nouveau disque doit être préparé. Pour le modèle SPARCstorage Array 110, un seul plateau doit être préparé.

Si vous utilisez un tableau SPARCstorage Array100, suivez les étapes indiquées. Si vous utilisez un tableau SPARCstorage Array200 avec plateau SCSI différentiel, vous devez arrêter tous les disques attachés au contrôleur de tableau qui sera connecté au nouveau disque. Autrement dit, vous devez répéter toutes les étapes relatives au plateau pour tous les boîtiers de disque attachés au contrôleur de tableau qui sera connecté au nouveau disque. Si vous utilisez un tableau SPARCstorage Array 214 RSM, il n'est pas nécessaire d'exécuter les étapes relatives au plateau, puisque les disques individuels peuvent être installés sans affecter les autres disques.

Pour la description du boîtier de disque, reportez-vous au manuel d'entretien de l'unité d'expansion de disque multihôtes.

Ajout d'un disque de tableau SPARCstorage Array

Selon le type du boîtier de disque, l'ajout de disques multihôtes à un tableau SPARCstorage Array (SSA) peut nécessiter la mise hors ligne de tous les objets du gestionnaire de volumes dans le plateau ou le boîtier de disque affecté. Par ailleurs, il est possible que le plateau ou le boîtier de disque contienne des disques de plusieurs groupes de disques, auquel cas un seul noeud doit être propriétaire de tous les groupes de disques affectés.

▼ Comment ajouter un disque de tableau SPARCstorage Array (Solstice DiskSuite)

La liste qui suit énumère les étapes générales de l'ajout d'un disque multihôtes dans une configuration Solstice DiskSuite:

- Commutation des hôtes logiques sur un noeud de grappe.

- Identification du contrôleur pour le nouveau disque et d'un connecteur libre dans le plateau ou le boîtier.
- Dans le cas d'un tableau SPARCstorage Array modèle 100, préparation du boîtier de disque en vue du retrait d'un plateau.
- Dans le cas d'un tableau SPARCstorage Array modèle 200 avec plateaux différentiels Wide-SCSI, mise hors tension du contrôleur et de tous les disques attachés.
- Suppression de toutes les réserves dynamiques sur les unités affectées.
- Suppression de toutes les bases de données d'état des métapériphériques des unités affectées.
- Mise hors ligne de tous les métapériphériques contenant des unités affectées.
- Arrêt de toutes les unités affectées.
- Ajout du nouveau disque.
- Remise en service des unités affectées.
 - Démarrage de toutes les unités.
 - Remise en ligne de tous les métapériphériques affectés.
 - Rajout de toutes les réserves dynamiques supprimées.
 - Recréation de tous les métapériphériques supprimés.
- Exécution des tâches administratives pour la préparation du disque pour son utilisation par Sun Cluster.
 - Création des fichiers spéciaux `/devices` et des liens `/dev/dsk` et `/dev/rdisk`.
 - Exécution de la commande `scdidadm -r`.
 - Ajout du disque à l'ensemble de disques.
 - Formatage et partitionnement du disque, au besoin.
 - Exécution des tâches administratives relatives au gestionnaire de volumes.

Les étapes expliquées maintenant décrivent en détail l'ajout d'un nouveau disque multihôtes dans une configuration Solstice DiskSuite.

1. Rendez d'autres noeuds de la grappe propriétaires de l'hôte logique qui inclura le nouveau disque.

Commutez les hôtes logiques dont les disques se trouvent dans le plateau qui sera retiré.

```
phys-hahost1# haswitch phys-hahost1 hahost1 hahost2
```

2. Trouvez le numéro de contrôleur du plateau auquel le disque sera ajouté.

Des noms universels sont attribués aux tableaux SPARCstorage Array. Le nom universel à l'avant du tableau SPARCstorage Array figure également dans l'entrée `/devices`, laquelle est reliée par un pointeur à l'entrée `/dev` contenant le numéro du contrôleur. Exemple :

```
phys-hahost1# ls -l /dev/rdisk | grep -i WWN | tail -1
```

Si le nom universel à l'avant du tableau SPARCstorage Arrayest 36cc, la sortie suivante sera affichée et le numéro de contrôleur sera c2 :

```
phys-hahost1# ls -l /dev/rdisk | grep -i 36cc | tail -1
lrwxrwxrwx 1 root  root      94 Jun 25 22:39 c2t5d2s7 ->
../../../../devices/io-unit@f,e1200000/sbi@0,0/SUNW,soc@3,0/SUNW,
pln@a0000800,201836cc/ssd@5,2:h,raw
```

3. Exécutez la commande luxadm(1M) avec l'option display pour afficher les emplacements libres.

```
phys-hahost1# luxadm display c2
                               SPARCstorage Array Configuration
...

```

	TRAY 1	DEVICE STATUS TRAY 2	TRAY 3
slot			
1	Drive: 0,0	Drive: 2,0	Drive: 4,0
2	Drive: 0,1	Drive: 2,1	Drive: 4,1
3	NO SELECT	NO SELECT	NO SELECT
4	NO SELECT	NO SELECT	NO SELECT
5	NO SELECT	NO SELECT	NO SELECT
6	Drive: 1,0	Drive: 3,0	Drive: 5,0
7	Drive: 1,1	NO SELECT	NO SELECT
8	NO SELECT	NO SELECT	NO SELECT
9	NO SELECT	NO SELECT	NO SELECT
10	NO SELECT	NO SELECT	NO SELECT
...			

Les emplacements vides ont l'état AUCUN CHOIX. La sortie affichée ici provient d'un tableau SPARCstorage Array110 ; l'affichage sera légèrement différent si vous utilisez un autre modèle.

Identifiez le plateau auquel vous ajouterez le nouveau disque. Si vous pouvez ajouter le disque sans influencer les autres unités (avec le modèle SPARCstorage Array 214 RSM, par exemple), passez à Etape 11, page 208.

Dans les étapes suivantes de la procédure, le plateau 2 est utilisé à titre d'exemple. L'emplacement sélectionné pour le nouveau disque est le plateau 2, emplacement 7. Le nouveau disque porte la désignation c2t3d1.

4. Trouvez toutes les réserves dynamiques affectées par l'installation.

Pour connaître l'état et l'emplacement de toutes les réserves dynamiques, exécutez la commande metahs(1M) avec l'option -i sur chaque hôte logique.

```
phys-hahost1# metahs -s hahost1 -i
...
phys-hahost1# metahs -s hahost2 -i
...
```

Remarque - Enregistrez la liste des réserves dynamiques. Cette liste sera utilisée plus tard dans cette procédure de maintenance. Prenez en note la liste des réserves dynamiques et de leurs pools.

5. Exécutez la commande `metahs(1M)` avec l'option `-d` pour supprimer toutes les réserves dynamiques affectées.

Pour plus de détails, consultez la page de manuel `metahs(1M)`.

```
phys-hahost1# metahs -s hahost1 -d composants-pool-réserves dynamiques
phys-hahost1# metahs -s hahost2 -d composants-pool-réserves dynamiques
```

6. Trouvez toutes les répliques de base de données d'état des métapériphériques situées sur les disques affectés.

Exécutez la commande `metadb(1M)` sur chaque hôte logique pour trouver toutes les bases de données d'état des métapériphériques. Dirigez la sortie vers des fichiers temporaires.

```
phys-hahost1# metadb -s hahost1 > /usr/tmp/mddb1
phys-hahost1# metadb -s hahost2 > /usr/tmp/mddb2
```

La sortie de la commande `metadb(1M)` indique l'emplacement des répliques de base de données d'état des métapériphériques dans ce boîtier de disque. Enregistrez ces informations pour utilisation lors de la restauration des répliques.

7. Supprimez les répliques de base de données d'état des métapériphériques situées sur les disques affectés.

Notez le numéro et l'emplacement des répliques que vous supprimez. Les répliques doivent être restaurées plus tard au cours de la procédure.

```
phys-hahost1# metadb -s hahost1 -d répliques
phys-hahost1# metadb -s hahost2 -d répliques
```

8. Exécutez la commande `metastat(1M)` pour identifier tous les composants de métapériphérique sur les disques affectés.

Dirigez la sortie de la commande `metastat(1M)` vers un fichier temporaire pour pouvoir utiliser ces informations plus tard au cours de la suppression et de l'ajout des métapériphériques.

```
phys-hahost1# metastat -s hahost1 > /usr/tmp/replicalog1
phys-hahost1# metastat -s hahost2 > /usr/tmp/replicalog2
```

9. Mettez hors ligne tous les sous-miroirs contenant des disques affectés.

Utilisez les fichiers temporaires pour créer un script de mise hors ligne de tous les sous-miroirs affectés dans l'unité d'expansion de disque. S'il n'y a que quelques sous-miroirs, exécutez la commande `metaoffline(1M)` pour les mettre chacun hors ligne. Voici un script d'exemple.

```
#!/bin/sh
# metaoffline -s <ensemble_disques> <miroir> <sous-miroir>
metaoffline -s hahost1 d15 d35
metaoffline -s hahost2 d15 d35
...
```

10. Arrêtez les disques affectés.

Arrêtez les disques SPARCstorage Array du plateau au moyen de la commande `luxadm(1M)`.

```
phys-hahost1# luxadm stop -t 2 c2
```

11. Ajoutez le nouveau disque.

Ajoutez le disque conformément aux directives du manuel d'entretien du matériel pour votre unité d'expansion de disque multihôtes. Après avoir ajouté le disque :

- Si le boîtier de disque est un tableau SPARCstorage Array 214 RSM, passez à Etape 16, page 210. (Ce type de disque peut être ajouté sans impact sur les autres unités.)
- Pour tous les autres types de tableaux SPARCstorage Array, passez à Etape 12, page 209.

12. Vérifiez que tous les disques du plateau sont en rotation.

Les disques dans le plateau SPARCstorage Array démarrent normalement automatiquement, mais si le plateau n'a pas démarré dans les deux minutes, forcez le démarrage au moyen de la commande suivante :

```
phys-hahost1# luxadm start -t 2 c2
```

13. Remettez en ligne les sous-miroirs.

Modifiez le script que vous avez créé à l' Etape 9, page 208 pour remettre en ligne les sous-miroirs.

```
#!/bin/sh
# metaonline -s <ensemble_disques> <miroir> <sous_miroir>
metaonline -s hahost1 d15 d35
metaonline -s hahost2 d15 d35
...
```

14. Restaurez les réserves dynamiques supprimées à l'Etape 5, page 207.

```
phys-hahost1# metahs -s hahost1 -a composants-pool-réserves dynamiques
phys-hahost1# metahs -s hahost2 -a composants-pool-réserves dynamiques
```

15. Restaurez le nombre initial de répliques de base de données d'état des métapériphériques sur les périphériques du plateau.

Les répliques ont été supprimées à Etape 7, page 207.

```
phys-hahost1# metadb -s hahost1 -a répliques
phys-hahost1# metadb -s hahost2 -a répliques
```

- 16. Exécutez les commandes `drvconfig(1M)` et `disks(1M)` pour créer les nouvelles entrées dans `/devices`, `/dev/dsk` et `/dev/rdsk` pour tous les nouveaux disques.**

```
phys-hahost1# drvconfig
phys-hahost1# disks
```

- 17. Attribuez la propriété de l'hôte logique auquel le disque va être ajouté à l'autre noeud connecté au tableau SPARCstorage Array.**

Pour cette opération, on suppose que chaque disque est connecté à deux noeuds.

```
phys-hahost1# haswitch phys-hahost2 hahost2
```

- 18. Exécutez les commandes `drvconfig(1M)` et `disks(1M)` sur le noeud de grappe désormais propriétaire de l'ensemble de disques auquel le disque sera ajouté.**

```
phys-hahost2# drvconfig
phys-hahost2# disks
```

- 19. Exécutez la commande `scdidadm(1M)` pour initialiser le nouveau disque pour son utilisation par le pseudo-pilote DID.**

Vous devez exécuter la commande `scdidadm(1M)` sur le noeud 0 de la grappe. Pour plus de détails sur le pseudo-pilote DID, reportez-vous au *Sun Cluster 2.2 Software Installation Guide*.

```
phys-hahost2# scdidadm -r
```

- 20. Ajoutez le disque à un ensemble de disques.**

La syntaxe de la commande est la suivante, où *ensemble_disques* représente le nom de l'ensemble de disques contenant le disque défectueux et *unité* le nom DID du disque dans le format `dN` (pour les nouvelles installations de Sun Cluster) ou `cNtYdZ` (pour les installations mises à niveau à partir de HA 1.3) :

```
# metaset -s ensemble_disques -a unité
```



Attention - La commande `metaset(1M)` peut repartitionner ce disque automatiquement. Pour de plus amples informations, voir la documentation de Solstice DiskSuite.

21. Exécutez la commande `scadmin(1M)` pour réserver et activer le mode failfast sur le disque spécifié qui vient d'être ajouté à l'ensemble de disques.

```
phys-hahost2# scadmin reserve cNtXdYsZ
```

22. Exécutez les tâches administratives habituelles pour le nouveau disque.

Vous pouvez maintenant exécuter les étapes d'administration habituelles pour la mise en service du nouveau disque (partitionnement du disque, ajout de celui-ci à la configuration comme réserve dynamique ou configuration comme métapériphérique, entre autres). Pour de plus amples informations sur ces tâches, voir la documentation de Solstice DiskSuite.

23. Au besoin, retransférez les hôtes logiques à leurs maîtres par défaut.

▼ Comment ajouter un disque de tableau SPARCstorage Array (VxVM)

La liste qui suit énumère les étapes générales de l'ajout d'un disque multihôtes dans une configuration VxVM:

- Commutation des hôtes logiques sur un noeud de grappe.
- Identification du contrôleur pour le nouveau disque et d'un emplacement vide dans le plateau ou le boîtier.
- Dans le cas d'un tableau SPARCstorage Array modèle 100, préparation du boîtier de disque en vue du retrait d'un plateau.
- Dans le cas d'un tableau SPARCstorage Array modèle 200 avec plateaux différentiels Wide-SCSI, mise hors tension du contrôleur et de tous les disques attachés.
- Identification des objets VxVM sur le plateau affecté.
- Arrêt des E/S dans les volumes avec sous-disques sur le plateau affecté.
- Ajout du nouveau disque.
- Remise en service des unités affectées.

- Démarrage de toutes les unités.
- Remise en ligne de tous les objets VxVM affectés.
- Exécution des tâches administratives pour la préparation du disque pour son utilisation par Sun Cluster.
 - Création des fichiers spéciaux `/devices` et des liens `/dev/dsk` et `/dev/rdisk`.
 - Repérage du nouveau disque.
 - Mise du disque sous le contrôle de la machine virtuelle (VM).
 - Formatage et partitionnement du disque, au besoin.
 - Exécution des tâches administratives relatives au gestionnaire de volumes.

Les étapes expliquées maintenant décrivent en détail l'ajout d'un disque multihôtes dans une configuration VxVM.

1. Rendez un autre noeud de la grappe propriétaire de l'hôte logique qui inclura le nouveau disque.

Commutez les hôtes logiques dont les disques se trouvent dans le plateau qui sera retiré.

```
phys-hahost1# haswitch phys-hahost1 hahost1 hahost2
```

Remarque - Dans une configuration en miroir, il n'est pas nécessaire de commuter les hôtes logiques si le noeud n'est pas arrêté.

2. Trouvez le numéro de contrôleur du plateau auquel le disque sera ajouté.

Des noms universels sont attribués aux tableaux SPARCstorage Array. Le nom universel à l'avant du tableau SPARCstorage Array figure également dans l'entrée `/devices`, laquelle est reliée par un pointeur à l'entrée `/dev` contenant le numéro du contrôleur. Exemple :

```
phys-hahost1# ls -l /dev/rdisk | grep -i WWN | tail -1
```

Si le nom universel à l'avant du tableau SPARCstorage Array est `36cc`, la sortie suivante sera affichée et le numéro de contrôleur sera `c2` :

```
phys-hahost1# ls -l /dev/rdisk | grep -i 36cc | tail -1
lrwxrwxrwx 1 root  root  94 Jun 25 22:39 c2t5d2s7 ->
../../../../devices/io-unit@f,e1200000/sbi@0,0/SUNW,soc@3,0/SUNW,
pln@a0000800,201836cc/ssd@5,2:h,raw
phys-hahost1#
```

3. Exécutez la commande `luxadm(1M)` avec l'option `display` pour afficher les emplacements vides.

Si vous pouvez ajouter le disque sans influencer les autres unités, passez à Etape 11, page 215.

```
phys-hahost1# luxadm display c2
                    SPARCstorage Array Configuration
...
          TRAY 1          DEVICE STATUS          TRAY 3
          slot
1      Drive: 0,0          Drive: 2,0          Drive: 4,0
2      Drive: 0,1          Drive: 2,1          Drive: 4,1
3      NO SELECT          NO SELECT          NO SELECT
4      NO SELECT          NO SELECT          NO SELECT
5      NO SELECT          NO SELECT          NO SELECT
6      Drive: 1,0          Drive: 3,0          Drive: 5,0
7      Drive: 1,1          NO SELECT          NO SELECT
8      NO SELECT          NO SELECT          NO SELECT
9      NO SELECT          NO SELECT          NO SELECT
10     NO SELECT          NO SELECT          NO SELECT
...

```

Les emplacements vides ont l'état AUCUN CHOIX. La sortie affichée ici provient d'un tableau SPARCstorage Array110 ; l'affichage sera légèrement différent si vous utilisez un autre modèle.

Identifiez le plateau auquel vous ajouterez le nouveau disque.

Dans les étapes suivantes de la procédure, le plateau 2 est utilisé à titre d'exemple. L'emplacement sélectionné pour le nouveau disque est le plateau 2, emplacement 7. Le nouveau disque porte la désignation `c2t3d1`.

4. Identifiez tous les volumes et les périphériques virtuels sur les disques du plateau qui contiendra le nouveau disque.

a. Trouvez, dans l'adresse physique du périphérique (`cNtNdN`), le numéro du contrôleur et le numéro de la cible.

Dans cet exemple, le numéro du contrôleur est 2 et la cible 3.

b. Identifiez les périphériques dans la sortie de la commande `vxdisk list`.

L'exemple qui suit montre comment la commande `vxdisk` peut être utilisée pour obtenir les informations.

```
# vxdisk -g groupe_disques -q list | nawk '/^c2/ {print $3}'
```

Notez le nom du volume pour les disques figurant sur la sortie de la commande.

- c. **Identifiez tous les périphériques virtuels sur les périphériques ci-dessus, à l'aide de la version qui convient (csh, ksh, ou shell Bourne) de la commande suivante.**

```
PLLIST=`vxprint -ptq -g groupe_disques  
-e '(aslist.sd_dm_name in ('`c2t3d0`')) && (pl_kstate=ENABLED)' | nawk '{print $2}'`
```

Pour csh, la syntaxe est `set PLLIST`. Pour ksh, la syntaxe est `export PLLIST=`. Avec le shell Bourne, vous devez utiliser la commande `export PLLIST` après définition de la variable.

5. **Après avoir défini la variable, interrompez les E/S sur les volumes dont les composants (sous-disques) se trouvent sur le plateau.**

Veillez à ce que tous les volumes associés à ce plateau soient détachés (configurations en miroir ou RAID5) ou arrêtés (périphériques virtuels simples). Exécutez la commande suivante pour détacher un périphérique virtuel mis en miroir.

```
# vxplex -g groupe_disques det ${PLLIST}
```

Autre commande permettant de détacher un périphérique virtuel d'un plateau :

```
# vxplex -g groupe_disques -v volume det périphérique_virtuel
```

Pour interrompre les E/S sur des périphériques virtuels simples, démontez les systèmes de fichiers ou interrompez tout accès à la base de données.

Remarque - Les volumes mis en miroir seront quand même actifs, car l'autre moitié du miroir est encore disponible.

6. **Ajoutez le nouveau disque.**

Ajoutez le disque conformément aux directives du manuel d'entretien du matériel pour votre unité d'expansion de disque multihôtes.

7. **Vérifiez que tous les disques du plateau sont en rotation.**

Les disques du plateau SPARCstorage Array démarrent normalement automatiquement, mais si le plateau ne démarre pas dans les deux minutes, forcez le démarrage au moyen de la commande suivante :

```
phys-hahost1# luxadm start -t 2 c2
```

8. Exécutez les commandes `drvconfig(1M)` et `disks(1M)` pour créer les nouvelles entrées dans `/devices`, `/dev/dsk` et `/dev/rdisk` pour tous les nouveaux disques.

```
phys-hahost1# drvconfig
phys-hahost1# disks
```

9. Forcez le pilote `vxconfigd VxVM` à rechercher les nouveaux disques.

```
phys-hahost1# vxdctl enable
```

10. Mettez le nouveau disque sous le contrôle de la machine virtuelle (VM) à l'aide de la commande `vxdiskadd`.

11. Exécutez les tâches administratives habituelles pour le nouveau disque.

Vous pouvez maintenant exécuter les étapes d'administration habituelles pour la mise en service du nouveau disque (partitionnement du disque, ajout de celui-ci à la configuration comme réserve dynamique ou configuration comme périphérique virtuel, entre autres).

Ainsi prend fin la procédure d'ajout d'un disque multihôtes à un tableau SPARCstorage Array existant.

Remplacement d'un disque de tableau SPARCstorage Array

Cette section explique comment remplacer un disque multihôtes dans un tableau SPARCstorage Array (SSA) sans interrompre les services Sun Cluster (remplacement en ligne) lorsque le gestionnaire de volumes signale des problèmes tels que :

- Composants affichant l'état "Maintenance requise".
- Remplacement de réserve dynamique.
- Erreurs de disque intermittentes.

▼ Comment remplacer un disque de tableau SPARCstorage Array(Solstice DiskSuite)

La liste qui suit énumère les étapes générales du remplacement d'un disque multihôtes dans une configuration Solstice DiskSuite. Certaines de ces étapes ne s'appliquent que dans le cas de configurations utilisant les tableaux SPARCstorage Array modèles 100 ou 200 (avec plateau SCSI différentiel).

- Commutation des hôtes logiques sur un noeud de grappe.
- Identification du disque qui doit être remplacé.
- Identification du plateau contenant le disque à remplacer.
- (SSA100 et SSA 200 exclusivement) Déconnexion des sous-miroirs du plateau ou du boîtier de disque affecté.
- (SSA100 et SSA 200 exclusivement) Exécution de la commande `metaclear(1M)` sur les sous-miroirs déconnectés.
- (SSA100 et SSA 200 exclusivement) Suppression des réserves dynamiques disponibles dans le plateau affecté.
- Retrait du disque défectueux de l'ensemble de disques.
- (SSA100 et SSA 200 exclusivement) Suppression des répliques de base de données d'état des métapériphériques éventuellement affectées sur les disques du plateau affecté.
- (SSA100 et SSA 200 exclusivement) Production d'une liste de métapériphériques dans le plateau affecté.
- (SSA100 et SSA 200 exclusivement) Exécution de la commande `metaoffline(1M)` sur les sous-miroirs dans le plateau affecté ou les sous-miroirs exploitant les réserves dynamiques du plateau.
- (SSA100 et SSA 200 exclusivement) Vidage de la mémoire non-volatile NVRAM, si celle-ci est activée.
- Arrêt des disques et retrait du plateau ou du boîtier de disque.
- Remplacement de l'unité de disque.
- Exécution de la commande `sddidadm -R`.
- Ajout du nouveau disque à l'ensemble de disques.
- Réserve et activation du mode failfast sur le nouveau disque.
- Partitionnement du nouveau disque.
- (SSA100 et SSA 200 exclusivement) Exécution de la commande `metainit(1M)` pour initialiser les périphériques précédemment supprimés par la commande `metaclear(1M)`.
- (SSA100 et SSA 200 exclusivement) Remise en ligne des miroirs hors ligne avec la commande `metaonline(1M)` et resynchronisation.
- (SSA100 et SSA 200 exclusivement) Reconnexion des sous-miroirs précédemment déconnectés.

- (SSA100 et SSA 200 exclusivement) Remplacement des réserves dynamiques utilisées par les sous-miroirs venant d'être reconnectés.
- (SSA100 et SSA 200 exclusivement) Retour des réserves dynamiques supprimées à leurs pools d'origine.
- Exécution de la commande `metastat(1M)` pour vérifier si le problème a été corrigé.

Les étapes expliquées maintenant décrivent le remplacement d'un disque multihôtes défectueux dans une configuration Solstice DiskSuite.

1. Exécutez la commande `haswitch(1M)` pour que d'autres noeuds deviennent propriétaires des hôtes logiques affectés.

```
phys-hahost1# haswitch phys-hahost1 hahost1 hahost2
```

Il est possible que le plateau SPARCstorage Array contenant le disque défectueux contienne aussi des disques inclus dans plusieurs hôtes logiques. Si c'est le cas, attribuez à un autre noeud de la grappe la propriété de tous les hôtes logiques contenant des disques utilisant ce plateau.

2. Trouvez le disque devant être remplacé en examinant les sorties produites par les commandes `metastat(1M)` et `/var/adm/messages`.

Si la commande `metastat(1M)` signale qu'un périphérique est en état de maintenance ou que des composants ont été remplacés par des réserves dynamiques, vous devez trouver et remplacer le périphérique en question. Un exemple de sortie de la commande `metastat(1M)` est présenté ci-après. Dans cet exemple, le périphérique `c3t3d4s0` est en état de maintenance.

```
phys-hahost1# metastat -s hahost1
...
d50:Submirror of hahost1/d40
  State: Needs Maintenance
  Stripe 0:
    Device      Start Block      Dbase      State      Hot Spare
    c3t3d4s0    0                 No         Okay       c3t5d4s0
...
```

Vérifiez `/var/adm/messages` afin de déterminer quel type de problème a été détecté.

```
...
Jun 1 16:15:26 host1 unix: WARNING: /io-unit@f,e1200000/sbi@0.0/SUNW,pln@a0000000,741022/
ssd@3,4(ssd49):
Jun 1 16:15:26 host1 unix: Error for command 'write(I)' Err
Jun 1 16:15:27 host1 unix: or Level: Fatal
Jun 1 16:15:27 host1 unix: Requested Block 144004, Error Block: 715559
```

(suite)

(suite)

```
Jun 1 16:15:27 host1 unix: Sense Key: Media Error
Jun 1 16:15:27 host1 unix: Vendor 'CONNER':
Jun 1 16:15:27 host1 unix: ASC=0x10(ID CRC or ECC error),ASCQ=0x0,FRU=0x15
...
```

3. Déterminez l'emplacement du disque défaillant à l'aide de la commande luxadm(1M).

La commande luxadm(1M) énumère les plateaux et les disques qui leur sont associés. La sortie varie en fonction du modèle de tableau SPARCstorage Array. Un exemple de sortie pour un tableau SPARCstorage Array 100 est illustré ci-dessous. Le disque défectueux figure en caractères gras.

```
phys-hahost1# luxadm display c3
SPARCstorage Array Configuration
Controller path:
/devices/iommu@f,e0000000/sbus@f,e0001000/SUNW,soc@0,0/
SUNW,pln@a0000000,779a16:ctlr
DEVICE STATUS
TRAY1          TRAY2          TRAY3
Slot
1      Drive:0,0      Drive:2,0      Drive:4,0
2      Drive:0,1      Drive:2,1      Drive:4,1
3      Drive:0,2      Drive:2,2      Drive:4,2
4      Drive:0,3      Drive:2,3      Drive:4,3
5      Drive:0,4      Drive:2,4      Drive:4,4
6      Drive:1,0      Drive:3,0      Drive:5,0
7      Drive:1,1      Drive:3,1      Drive:5,1
8      Drive:1,2      Drive:3,2      Drive:5,2
9      Drive:1,3      Drive:3,3      Drive:5,3
10     Drive:1,4      Drive:3,4      Drive:5,4
CONTROLLER STATUS
Vendor:      SUN
Product ID:  SSA110
Product Rev:  1.0
Firmware Rev: 3.9
Serial Num:   000000741022
Accumulate performance Statistics: Enabled
```

4. Déconnectez tous les sous-miroirs avec composants sur le disque à remplacer.

Si vous détachez un sous-miroir dont un composant est défectueux, vous devez forcer l'opération au moyen de la commande metadetach -f. Dans l'exemple qui suit, la commande détache le sous-miroir d50 du métamiroir d40.

```
phys-hahost1# metadetach -s hahost1 -f d40 d50
```

5. Exécutez la commande `metaclear(1M)` pour supprimer les sous-miroirs détachés à l' Etape 4, page 218.

```
phys-hahost1# metaclear -s hahost1 -f d50
```

6. Avant de supprimer les répliques et les réserves dynamiques, notez l'emplacement (tranche), le nombre de répliques et les informations relatives aux réserves dynamiques (nom et liste des périphériques contenant les pools correspondants) afin de pouvoir répéter les actions dans l'ordre inverse, une fois le disque remplacé.

7. Supprimez toutes les réserves dynamiques ayant l'état `Disponible` qui se trouvent dans le plateau contenant le disque défectueux.

Cela inclut toutes les réserves dynamiques, quel que soit l'hôte logique auquel elles sont attribuées. Dans l'exemple qui suit, la commande `metahs(1M)` signale la présence de réserves dynamiques sur `hahost1`, mais indique également qu'il n'y en a aucune sur `hahost2`.

```
phys-hahost1# metahs -s hahost1 -i
hahost1:hsp000 2 hot spares
      c1t4d0s0          Available      2026080 blocks
      c3t2d5s0          Available      2026080 blocks
phys-hahost1# metahs -s hahost1 -d hsp000 c3t2d4s0
hahost1:hsp000:
      Hotspare is deleted
phys-hahost1# metahs -s hahost2 -i
phys-hahost1#
hahost1:hsp000 1 hot spare
      c3t2d5s0          Available      2026080 blocks
```

8. Exécutez la commande `metaset(1M)` afin de supprimer le disque défectueux de l'ensemble de disques.

La syntaxe de la commande est précisée ci-après. Dans cet exemple, `ensemble_disques` représente le nom de l'ensemble de disques contenant le disque défectueux et `unité` le nom DID du disque dans le format `dN` (pour les nouvelles installations de Sun Cluster) ou `cNtYdZ` (pour les installations mises à niveau à partir de la version HA 1.3).

```
# metaset -s ensemble_disques -d unité
```

Cette opération peut demander 15 minutes et même davantage, selon la taille de la configuration et le nombre de disques.

9. Supprimez les répliques de base de données d'état des métapériphériques présentes sur les disques du plateau affecté.

La commande `metadb(1M)` avec l'option `-s` signale des répliques dans l'ensemble de disques spécifié.

```
phys-hahost1# metadb -s hahost1
phys-hahost1# metadb -s hahost2
phys-hahost1# metadb -s hahost1 -d répliques-dans-plateau
phys-hahost1# metadb -s hahost2 -d répliques-dans-plateau
```

10. Trouvez les sous-miroirs utilisant des composants résidant dans le plateau affecté.

Une méthode possible consiste à utiliser la commande `metastat(1M)` pour créer des fichiers temporaires contenant les noms de tous les métapériphériques.

Exemple :

```
phys-hahost1# metastat -s hahost1 > /usr/tmp/hahost1.stat
phys-hahost1# metastat -s hahost2 > /usr/tmp/hahost2.stat
```

Recherchez les composants en question dans les fichiers temporaires (`c3t3dn` et `c3t2dn` dans cet exemple). Les informations contenues dans les fichiers temporaires se présentent de la manière suivante :

```
...
hahost1/d35: Submirror of hahost1/d15
  State: Okay
  Hot Spare pool: hahost1/hsp100
  Size: 2026080 blocks
  Stripe 0:
    Device      Start Block  Dbase   State   Hot Spare
    c3t3d3s0    0            No      Okay
hahost1/d54: Submirror of hahost1/d24
  State: Okay
  Hot Spare pool: hahost1/hsp106
  Size: 21168 blocks
  Stripe 0:
    Device      Start Block  Dbase   State   Hot Spare
    c3t3d3s6    0            No      Okay
```

(suite)

(suite)

```
...
```

11. Mettez hors ligne tous les autres sous-miroirs ayant des composants dans le plateau affecté.

En utilisant la sortie obtenue des fichiers temporaires à l'Étape 10, page 220, exécutez la commande `metaoffline(1M)` sur tous les sous-miroirs du plateau affecté.

```
phys-hahost1# metaoffline -s hahost1 d15 d35
phys-hahost1# metaoffline -s hahost1 d24 d54
...
```

Exécutez la commande `metaoffline(1M)` autant de fois que nécessaire pour mettre hors ligne tous les sous-miroirs. Cela force Solstice DiskSuite à cesser d'utiliser les composants de sous-miroir.

12. Si la mémoire non-volatile NVRAM est activée, videz-la en transférant ses données au contrôleur, au plateau, au disque ou aux disques qui conviennent.

```
phys-hahost1# luxadm sync_cache pathname
```

Le message de confirmation affiché indique que la mémoire non-volatile NVRAM a été vidée. Pour plus de détails sur le vidage de la mémoire NVRAM, reportez-vous à la section "Vidage et purge de la mémoire NVRAM", page 231.

13. Arrêtez tous les disques contenus dans les plateaux de tableau SPARCstorage Array affectés.

Utilisez la commande `luxadm stop` pour arrêter les disques. Pour plus de détails à ce sujet, consultez la page de manuel `luxadm(1M)`.

```
phys-hahost1# luxadm stop -t 2 c3
```



Attention - N'exécutez pas de commandes Solstice DiskSuite pendant l'arrêt d'un plateau du tableau SPARCstorage Array, car cela pourrait entraîner le démarrage d'un ou plusieurs disques du plateau.

14. Remplacez le disque.

Pour plus de détails sur cette procédure, reportez-vous aux manuels d'entretien du matériel de tableau SPARCstorage Array.

15. Mettez à jour la base de données de pilotes DID en précisant l'ID du nouveau périphérique.

Utilisez l'indicateur `-l` avec la commande `scdidadm(1M)` pour identifier le nom DID du périphérique de bas niveau de l'unité à remplacer. Mettez ensuite à jour la base de données des pilotes DID en utilisant l'indicateur `-R` avec la commande `scdidadm(1M)`. Consultez le *Sun Cluster 2.2 Software Installation Guide* pour des détails complémentaires sur le pseudo-pilote DID.

```
phys-hahost1# scdidadm -o name -l /dev/rdisk/c3t3d4
6 phys-hahost1:/dev/rdisk/c3t3d4 /dev/did/rdsk/d6
phys-hahost1# scdidadm -R d6
```

16. Vérifiez que tous les disques dans l'unité d'expansion de disque multihôtes sont en rotation.

Les disques dans l'unité d'expansion de disque multihôtes démarrent normalement automatiquement. Si le plateau n'a pas démarré au bout de deux minutes, forcez le démarrage au moyen de la commande suivante :

```
phys-hahost1# luxadm start -t 2 c3
```

17. Ajoutez le nouveau disque dans l'ensemble de disques avec la commande `metaset(1M)`.

Au cours de cette étape, les répliques supprimées du disque défectueux sont automatiquement ajoutées. La syntaxe de la commande est la suivante, où *ensemble_disques* représente le nom de l'ensemble de disques contenant le disque défectueux et *unité* le nom DID du disque dans le format `dN` (pour les nouvelles installations de Sun Cluster) ou `cNtYdZ` (pour les installations mises à niveau à partir de HA 1.3) :

```
# metaset -s ensemble_disques -a unité
```

- 18. (Facultatif) Si vous avez supprimé des répliques, appartenant à d'autres ensembles de disques, de disques contenus dans le même plateau que le disque défectueux, exécutez la commande `metadb(1M)` pour rétablir ces répliques.**

```
phys-hahost1# metadb -s hahost2 -a répliques-supprimées
```

Pour ajouter plusieurs répliques sur la même tranche, utilisez l'option `-c`.

- 19. Exécutez la commande `scadmin(1M)` pour réserver et activer le mode failfast sur le disque spécifié qui vient d'être ajouté à l'ensemble de disques.**

```
phys-hahost2# scadmin reserve c3t3d4
```

- 20. Exécutez la commande `format(1M)` ou `fmthard(1M)` pour repartitionner le nouveau disque.**

Veillez à partitionner le nouveau disque exactement de la même manière que l'ancien. (L'enregistrement des données de format du disque est recommandé au Chapitre 1.)

- 21. Exécutez la commande `metainit(1M)` pour réinitialiser les disques supprimés à l'Etape 5, page 219.**

```
phys-hahost1# metainit -s hahost1 d50
```

- 22. Remettez en ligne tous les sous-miroirs mis hors ligne à l' Etape 11, page 221.**

```
phys-hahost1# metaonline -s hahost1 d15 d35  
phys-hahost1# metaonline -s hahost1 d24 d54  
...
```

Exécutez la commande `metaonline(1M)` autant de fois que nécessaire pour mettre en ligne tous les sous-miroirs.

Lorsque les sous-miroirs sont remis en ligne, Solstice DiskSuite resynchronise automatiquement tous les sous-miroirs et met ainsi à jour toutes les données.

Remarque - Si la commande `metastat(1M)` était exécutée à cette étape-ci, l'affichage indiquerait que tous les métapériphériques contenant des composants résidant sur le plateau affecté seraient en cours de resynchronisation.

23. Attachez les sous-miroirs déconnectés à l'Etape 4, page 218.

Pour ce faire, exécutez la commande `metattach(1M)`. Pour plus de détails, consultez la page de manuel `metattach(1M)`.

```
phys-hahost1# metattach -s hahost1 d40 d50
```

24. Remplacez s'il y a lieu les réserves dynamiques utilisées dans les sous-miroirs attachés à l'Etape 23, page 224.

Si, avant d'être détaché, un sous-miroir a fait l'objet d'un remplacement dynamique, ce remplacement reste en vigueur lorsque le sous-miroir est rattaché. Au cours de cette étape, la réserve dynamique retourne à l'état `Disponible`.

```
phys-hahost1# metareplace -s hahost1 -e d40 c3t3d4s0
```

25. Restaurez toutes les réserves dynamiques supprimées à l'Etape 7, page 219.

Exécutez la commande `metahs(1M)` pour rajouter les réserves dynamiques. Pour plus de détails, consultez la page de manuel `metahs(1M)`.

```
phys-hahost1# metahs -s hahost1 -a hsp000 c3t2d5s0
```

26. Au besoin, commutez les hôtes logiques sur leurs maîtres par défaut.

```
phys-hahost1# haswitch phys-hahost2 hahost2
```

27. Vérifiez si le remplacement a remédié au problème.

```
phys-hahost1# metastat -s hahost1
```

▼ Comment remplacer un disque de tableau SPARCstorage Array(VxVM)

Dans une configuration VxVM, il est possible de remplacer un disque SPARCstorage Array sans arrêter le système, à condition que la configuration en question ait été mise en miroir.

Remarque - Si vous devez remplacer un disque dans un tableau SPARCstorage Array amorçable, ne supprimez pas les plateaux SSA contenant le disque d'initialisation des hôtes. Arrêtez plutôt l'hôte dont le disque d'initialisation se trouve sur ce plateau. Laissez le logiciel de grappe reconfigurer les noeuds restants pour qu'il y ait reprise avant la réparation du disque défaillant. Pour de plus amples informations, reportez-vous au *Guide de l'utilisateur des tableaux SPARCstorage*.

La liste qui suit énumère les étapes générales du remplacement d'un disque multihôtes dans un environnement VxVM utilisant des disques de tableau SPARCstorage Array100 :

- Identification de tous les volumes et des périphériques virtuels correspondants sur les disques du plateau contenant le disque défectueux.
- Identification du contrôleur et du numéro de cible du disque défectueux.
- Identification des périphériques du plateau au moyen de la commande `vxdisk list`.
- Identification de tous les périphériques virtuels sur le plateau affecté.
- Déconnexion des tous les périphériques virtuels sur le plateau affecté.
- Suppression du disque de son groupe de disques.
- Arrêt des disques du plateau.
- Remplacement de l'unité de disque.
- Démarrage des disques du plateau.
- Initialisation du disque de remplacement.
- Analyse de la configuration de disque actuelle.
- Ajout du disque de remplacement dans le groupe de disques.
- Resynchronisation des volumes.

Les étapes expliquées maintenant décrivent en détail le remplacement d'un disque multihôtes dans un environnement VxVM utilisant des disques de tableau SPARCstorage Array 100.

1. **Si le disque remplacé est un périphérique de quorum, utilisez la commande `scconf -q` pour faire d'un autre disque le périphérique de quorum.**

2. Identifiez tous les volumes et les périphériques virtuels correspondants sur les disques du plateau contenant le disque défectueux.

- a. Trouvez, dans l'adresse physique du périphérique ($cNtNdN$), le numéro du contrôleur et le numéro de la cible.

Dans l'adresse $c3t2d0$, par exemple, le numéro du contrôleur est 3 et la cible 2.

- b. Identifiez les périphériques dans la sortie de la commande `vxdisk list`.

Si la cible est 0 ou 1, identifiez tous les périphériques dont l'adresse physique commence par $cNt0$ et $cNt1$, où N est le numéro du contrôleur. Si la cible est 2 ou 3, identifiez tous les périphériques dont l'adresse physique commence par $cNt2$ et $cNt3$. Si la cible est 4 ou 5, identifiez tous les périphériques dont l'adresse physique commence par $cNt4$ et $cNt5$. Voici un exemple illustrant comment la commande `vxdisk` peut être utilisée pour obtenir les informations.

```
# vxdisk -g groupe_disques -q list | egrep c3t2\|c3t3 | nawk '{print $3}'
```

- c. Notez le nom du support du volume pour le disque défectueux (selon la sortie produite par la commande).

Vous en aurez besoin à l'Etape 10, page 227.

- d. Identifiez tous les périphériques virtuels sur les périphériques ci-dessus, à l'aide de la version qui convient (`cs`, `ks`, ou `shell Bourne`) de la commande suivante.

```
PLLIST=`vxprint -ptq -g groupe_disques  
-e '(aslist.sd_dm_name in ('c3t2d0','c3t3d0','c3t3d1')) && (pl_kstate=ENABLED)' | nawk '{print $2}'`
```

Pour `cs`, la syntaxe est `set PLLIST` Pour `ks`, la syntaxe est `export PLLIST=` Avec le `shell Bourne`, vous devez utiliser la commande `export PLLIST` après définition de la variable.

3. Après avoir défini la variable, interrompez les E/S sur les volumes dont les composants (sous-disques) se trouvent sur le plateau.

Veillez à ce que tous les volumes associés à ce plateau soient détachés (configurations en miroir ou RAID5) ou arrêtés (périphériques virtuels simples). Exécutez la commande suivante pour détacher un périphérique virtuel mis en miroir.

```
# vxplex det ${PLLIST}
```

Autre commande permettant de détacher un périphérique virtuel d'un plateau :

```
# vxplex -g groupe_disques -v volume det périphérique_virtuel
```

Pour interrompre les E/S sur des périphériques virtuels simples, démontez les systèmes de fichiers ou interrompez tout accès à la base de données.

Remarque - Les volumes mis en miroir seront quand même actifs, car l'autre moitié du miroir est encore disponible.

4. Supprimez le disque du groupe de disques.

```
# vxdg -g groupe_disques rmdisk disque
```

5. Arrêtez les disques du plateau.

```
# luxadm stop -t contrôleur_plateau
```

6. Remplacez le disque défectueux.

7. Démarrez les disques.

```
# luxadm start -t contrôleurplateau
```

8. Initialisez le disque de remplacement.

```
# vxdisksetup -i périphérique
```

9. Détectez de nouveau la configuration actuelle du disque.

Exécutez les commandes suivantes sur tous les noeuds de la grappe.

```
# vxdctl enable  
# vxdisk -a online
```

10. Ajoutez le nouveau disque au groupe de disques.

L'élément *nom_support_périphérique* correspond au nom du support de volume enregistré à l'Étape 2, page 226c.

```
# vxvg -g groupe_disques -k adddisk nom-support-périphérique=nom-périphérique
```

11. Resynchronisez les volumes.

```
# vxrecover -g groupe_disques -b -o
```

Administration de la mémoire vive non-volatile (NVRAM) d'un tableau SPARCstorage Array

La mémoire NVRAM prend en charge les opérations d'écriture rapide dans les tableaux SPARCstorage Array. En l'absence de cette mémoire, les demandes d'écriture synchrone en provenance d'un programme doivent être enregistrées sur disque, et un accusé de réception doit être reçu par le programme pour que la demande suivante puisse être soumise. La mémoire NVRAM met en cache les demandes d'écriture dans une mémoire non-volatile et vide périodiquement les données sur disque. Après le stockage des données en mémoire NVRAM, un accusé de réception est renvoyé au programme comme si les données avaient été écrites sur disque. Cela améliore les performances des applications qui font intervenir un grand nombre d'opérations d'écriture avec les tableaux SPARCstorage Array.

Les procédures décrites ici utilisent l'interface de ligne de commande. Cependant, dans les configurations Solstice DiskSuite, vous pouvez également utiliser l'interface graphique *metatool* pour gérer la mémoire NVRAM d'un disque, d'un plateau ou d'un contrôleur. Pour de plus amples informations sur Solstice DiskSuite, consultez la documentation correspondante.



Attention - Faites usage de cette fonctionnalité avec prudence. Il s'agit d'un outil puissant de gestion des tableaux SPARCstorage Array. Sauvegardez toujours vos données avant d'exécuter ces procédures.

Activation et désactivation de la mémoire NVRAM

L'écriture rapide peut être configurée :

- au niveau du contrôleur, pour tous les disques d'un tableau SPARCstorage Array ;
- au niveau du disque, pour l'écriture rapide sur un disque individuel ;
- au niveau du plateau, par l'intermédiaire de l'interface graphique (GUI) de Solstice DiskSuite.

Lorsque l'écriture rapide est activée, elle peut être enregistrée et conservée — lors des cycles sous tension — comme élément de la configuration SPARCstorage Array.

L'écriture rapide est automatiquement désactivée au niveau du contrôleur si la batterie de la mémoire NVRAM est déchargée, absente ou défectueuse.

Avant d'activer l'écriture rapide, vous devez interrompre toutes les E/S sur le contrôleur ou le disque. Vous devez plus particulièrement vous assurer que le propriétaire de l'ensemble de disques est libéré, car des flux d'E/S sont maintenus tant que ce propriétaire existe. La procédure qui suit permet d'interrompre toutes les E/S.

Exécutez la commande `luxadm(1M)` pour activer et désactiver la mémoire NVRAM. Pour des informations complètes sur cette commande, consultez la page de manuel `luxadm(1M)`.

Remarque - La mémoire NVRAM doit être désactivée si vous utilisez la fonction de grappe de VxVM (utilisée avec Oracle Parallel Server).

▼ Comment activer et désactiver la mémoire NVRAM

La liste qui suit énumère les étapes générales de l'activation et de la désactivation de la mémoire NVRAM :

- Création d'une copie de sauvegarde de toutes les données.
- Obtention des privilèges de root (`superutilisateur`).
- Identification du contrôleur ou du disque sur lequel la mémoire NVRAM doit être activée ou désactivée.
- Arrêt de toutes les E/S sur le périphérique.
- Activation ou désactivation de la mémoire NVRAM.
- Redémarrage du périphérique et resynchronisation des données.

Les étapes expliquées maintenant décrivent l'activation et la désactivation de la mémoire NVRAM.

1. Identifiez le contrôleur, le plateau ou le disque individuel dont la mémoire NVRAM doit être activée ou désactivée.

Vous pouvez tirer parti de la commande `luxadm(1M)` pour afficher des informations sur le contrôleur, le plateau ou le disque individuel spécifié. L’affichage reproduit ci-dessous, par exemple, identifie tous les disques du contrôleur `c2`.

```
phys-hahost1# luxadm display c2
                                SPARCstorage Array Configuration
Controller path:
/devices/iommu@f,e0000000/sbus@f,e0001000/SUNW,soc@0,0/
SUNW,pln@a0000000,779a16:ctlr
                                DEVICE STATUS
                                TRAY 2
TRAY 1                                TRAY 3
slot
1    Drive: 0,0                        Drive: 2,0                        Drive: 4,0
2    Drive: 0,1                        Drive: 2,1                        Drive: 4,1
3    NO SELECT                          NO SELECT                          NO SELECT
4    NO SELECT                          NO SELECT                          NO SELECT
5    NO SELECT                          NO SELECT                          NO SELECT
6    Drive: 1,0                        Drive: 3,0                        Drive: 5,0
7    Drive: 1,1                        NO SELECT                          NO SELECT
8    NO SELECT                          NO SELECT                          NO SELECT
9    NO SELECT                          NO SELECT                          NO SELECT
10   NO SELECT                          NO SELECT                          NO SELECT
                                CONTROLLER STATUS
...
```

2. Interrompez toutes les E/S sur le périphérique affecté.

Solstice DiskSuite :

- Dans le cas d’un contrôleur ou d’un plateau, voir les étapes qui s’appliquent dans la procédure intitulée “Comment mettre hors service un plateau de tableau SPARCstorage Array(Solstice DiskSuite)”, page 186.
- Dans le cas d’un disque, voir les étapes qui s’appliquent dans la procédure intitulée “Comment remplacer un disque de tableau SPARCstorage Array(Solstice DiskSuite)”, page 216.

VxVM :

- Dans le cas d’un contrôleur ou d’un plateau, voir les étapes qui s’appliquent dans la procédure intitulée “Comment mettre hors service un plateau de tableau SPARCstorage Array(VxVM)”, page 188.
- Dans le cas d’un disque, voir les étapes qui s’appliquent dans la procédure intitulée “Comment remplacer un disque de tableau SPARCstorage Array(VxVM)”, page 225.

3. Activez ou désactivez l’écriture rapide sur le contrôleur ou le disque.

Choisissez l'option qui convient parmi les trois disponibles pour la commande `luxadm(1M)` (activation de l'écriture rapide pour toutes les opérations d'écriture, pour les opérations d'écriture synchrone seulement ou désactivation de l'écriture rapide).

- `-e` active l'écriture rapide pour toutes les opérations d'écriture.
- `-c` n'active l'écriture rapide que pour les opérations d'écriture synchrone.
- `-d` désactive l'écriture rapide.

Dans l'exemple suivant, la configuration de la mémoire NVRAM est enregistrée lorsque la machine est remise sous tension, et l'écriture rapide est activée pour toutes les opérations d'écriture. Pour plus de détails sur ces options, consultez la page de manuel `luxadm(1M)`.

```
phys-hahost# luxadm fast_write -s -e chemin
```

Le message de confirmation affiché indique que l'écriture rapide a été activée.

4. Exécutez les étapes nécessaires pour que le composant fonctionne normalement sous Sun Cluster.

Solstice DiskSuite :

- Dans le cas d'un contrôleur ou d'un plateau, voir les étapes qui s'appliquent dans la procédure intitulée "Comment remettre en service un plateau de tableau SPARCstorage Array(Solstice DiskSuite)", page 190.
- Dans le cas d'un disque, voir les étapes qui s'appliquent dans la procédure intitulée "Comment remplacer un disque de tableau SPARCstorage Array(Solstice DiskSuite)", page 216.

VxVM :

- Dans le cas d'un contrôleur ou d'un plateau, voir les étapes qui s'appliquent dans la procédure intitulée "Comment remettre en service un plateau de tableau SPARCstorage Array(VxVM)", page 192.
- Dans le cas d'un disque, voir les étapes qui s'appliquent dans la procédure intitulée "Comment remplacer un disque de tableau SPARCstorage Array(VxVM)", page 225.

Vidage et purge de la mémoire NVRAM

La commande `luxadm sync_cache` permet de vider sur disque toutes les opérations d'écriture en attente dans la mémoire NVRAM. Si une erreur survient au cours du vidage, vous devez purger les données au moyen de la commande `luxadm purge`. La purge des données entraîne la "suppression" des opérations d'écriture en attente dans la mémoire NVRAM.



Attention - La purge des données d'écriture rapide doit être effectuée avec prudence, et uniquement en cas de panne d'une unité de disque, car elle peut entraîner la perte de données.

Si la batterie de la mémoire NVRAM est déchargée, absente ou défectueuse, la mémoire NVRAM n'est pas opérationnelle et les données qui s'y trouvent sont perdues.

▼ Comment vider et purger la mémoire NVRAM

La liste qui suit énumère les étapes générales de vidage ou de purge, sur disque, des opérations d'écriture en attente pour le contrôleur sélectionné (et tous les disques correspondants) ou des opérations d'écriture individuelles :

- Création d'une copie de sauvegarde de toutes les données.
- Obtention des privilèges de root (superutilisateur).
- Identification du contrôleur ou du disque sur lequel les opérations d'écriture doivent être vidées.
- Vidage ou purge de toutes les opérations d'écriture en attente.
- Arrêt de toutes les E/S sur le périphérique.
- Remise en service du périphérique sous Sun Cluster.

Les étapes expliquées maintenant décrivent le vidage ou la purge des données de la mémoire NVRAM.

1. Identifiez le contrôleur ou le disque individuel devant être vidé ou purgé.

Vous pouvez utiliser la commande `luxadm(1M)` pour afficher des informations relatives au contrôleur, au plateau ou au disque spécifié. L'affichage reproduit ci-dessous, par exemple, identifie tous les disques du contrôleur `c2`.

```
phys-hahost1# luxadm display c2
                               SPARCstorage Array Configuration
Controller path:
/devices/iommu@f,e0000000/sbus@f,e0001000/SUNW,soc@0,0/
SUNW,pln@a0000000,779a16:ctrlr
                               DEVICE STATUS
slot   TRAY 1                TRAY 2                TRAY 3
1      Drive: 0,0            Drive: 2,0            Drive: 4,0
2      Drive: 0,1            Drive: 2,1            Drive: 4,1
3      NO SELECT             NO SELECT             NO SELECT
4      NO SELECT             NO SELECT             NO SELECT
5      NO SELECT             NO SELECT             NO SELECT
6      Drive: 1,0            Drive: 3,0            Drive: 5,0
```

(suite)

(suite)

```
7      Drive: 1,1          NO SELECT          NO SELECT
8      NO SELECT          NO SELECT          NO SELECT
9      NO SELECT          NO SELECT          NO SELECT
10     NO SELECT          NO SELECT          NO SELECT
                                CONTROLLER STATUS
Vendor:      SUN
Product ID:  SSA110
Product Rev: 1.0
Firmware Rev: 3.9
Serial Num:  000000779A16
Accumulate Performance Statistics: Enabled
phys-hahost1#
```

2. Interrompez toutes les E/S sur le périphérique affecté.

Solstice DiskSuite :

- Dans le cas d'un contrôleur ou d'un plateau, voir les étapes qui s'appliquent dans la procédure intitulée "Comment mettre hors service un plateau de tableau SPARCstorage Array(Solstice DiskSuite)", page 186.
- Dans le cas d'un disque, voir les étapes qui s'appliquent dans la procédure intitulée "Comment remplacer un disque de tableau SPARCstorage Array(Solstice DiskSuite)", page 216.

VxVM :

- Dans le cas d'un contrôleur ou d'un plateau, voir les étapes qui s'appliquent dans la procédure intitulée "Comment mettre hors service un plateau de tableau SPARCstorage Array(VxVM)", page 188.
- Dans le cas d'un disque, voir les étapes qui s'appliquent dans la procédure intitulée "Comment remplacer un disque de tableau SPARCstorage Array(VxVM)", page 225.

3. Videz ou purgez le contenu de la mémoire NVRAM sur un contrôleur, un plateau ou un disque individuel.

Si vous pouvez accéder aux disques du tableau SPARCstorage Array, videz la mémoire NVRAM. Ne purgez la mémoire NVRAM que si vous ne pouvez plus accéder au tableau SPARCstorage Array ou au disque.

```
phys-hahost1# luxadm sync_cache chemin
ou
phys-hahost1# luxadm purge chemin
```

Le message de confirmation affiché indique que la mémoire NVRAM a été vidée ou purgée.

4. Exécutez les étapes nécessaires pour que le composant fonctionne normalement sous Sun Cluster.

Solstice DiskSuite :

- Dans le cas d'un contrôleur ou d'un plateau, voir les étapes qui s'appliquent dans la procédure intitulée "Comment remettre en service un plateau de tableau SPARCstorage Array(Solstice DiskSuite)", page 190.
- Dans le cas d'un disque, voir les étapes qui s'appliquent dans la procédure intitulée "Comment remplacer un disque de tableau SPARCstorage Array(Solstice DiskSuite)", page 216.

VxVM :

- Dans le cas d'un contrôleur ou d'un plateau, voir les étapes qui s'appliquent dans la procédure intitulée "Comment remettre en service un plateau de tableau SPARCstorage Array(VxVM)", page 192.
- Dans le cas d'un disque, voir les étapes qui s'appliquent dans la procédure intitulée "Comment remplacer un disque de tableau SPARCstorage Array(VxVM)", page 225.

Administration des disques Sun StorEdge MultiPack et Sun StorEdge D1000

Ce chapitre explique comment administrer les disques Sun StorEdge™ MultiPack et Sun StorEdge D1000. Certaines des procédures décrites ici dépendent du logiciel de gestion de volumes (Solstice DiskSuite ou VxVM). Le titre de ces procédures inclut le nom du gestionnaire de volumes.

- “Reprise après une panne de courant”, page 235
- “Administration des disques Sun StorEdge MultiPack et Sun StorEdge D1000”, page 240
- “Administration des disques Sun StorEdge MultiPack et Sun StorEdge D1000”, page 243

Utilisez le manuel d’entretien des disques Sun StorEdge MultiPack et Sun StorEdge D1000, ainsi que la documentation du logiciel de gestion de volumes, pour remplacer ou réparer le matériel du disque dans la configuration Sun Cluster.

Reprise après une panne de courant

Lorsque l’alimentation électrique d’un disque Sun StorEdge MultiPack ou Sun StorEdge D1000 est coupée, les opérations d’E/S produisent des erreurs qui sont détectées par le logiciel de gestion des volumes. Les erreurs ne sont pas signalées tant que des transactions d’E/S ne sont pas effectuées sur le disque.

Vous devez surveiller la configuration de ces événements à l’aide des commandes décrites au Chapitre 2.

▼ Comment effectuer une reprise après une panne de courant (Solstice DiskSuite)

La liste qui suit énumère les étapes générales de la récupération suite à une panne de courant d'un boîtier de disque dans un environnement Solstice DiskSuite:

- Identification des répliques défectueuses
- Remise en service des répliques défectueuses.
- Identification des périphériques défectueux.
- Remise en service des périphériques défectueux.
- Resynchronisation des disques.

Les étapes expliquées maintenant décrivent en détails la récupération suite à une panne de courant d'un boîtier de disque dans l'environnement Solstice DiskSuite.

1. **Une fois l'alimentation rétablie, exécutez la commande `metadb(1M)` pour identifier les répliques contenant des erreurs :**

```
# metadb -s ensemble_disques
```

2. **Remettez en service les répliques.**

Après la panne de courant, toutes les répliques de base de données d'état des métapériphériques qui se trouvent sur le châssis de boîtier du disque affecté affichent un état d'erreur. Etant donné que la récupération d'une réplique de la base de données d'état des métapériphériques n'est pas automatique, il est préférable d'effectuer la récupération immédiatement après la remise en service du boîtier de disque. A défaut, une nouvelle panne risque de mettre hors service la plupart des répliques et de provoquer une panne du noyau. Ce comportement de Solstice DiskSuite se produit en général lorsque le nombre de répliques disponibles n'est pas suffisant.

Bien que ces répliques erronées soient restaurées lors de la relève suivante (commande `haswitch(1M)` ou `reboot(1M)`), vous pourriez vouloir les remettre en service manuellement en les supprimant et en les ajoutant de nouveau.

Remarque - Veillez à rajouter le même nombre de répliques que celui qui a été supprimé sur chaque tranche. Vous pouvez supprimer plusieurs répliques simultanément à l'aide de la commande `metadb(1M)`. Si une tranche doit contenir plusieurs exemplaires des répliques, ajoutez ces dernières en exécutant une seule fois la commande `metadb(1M)`, accompagnée de l'indicateur `-c`.

3. **Exécutez la commande `metastat(1M)` pour identifier les métapériphériques contenant des erreurs.**

```
# metastat -s ensemble_disques
```

4. Remettez en service les métapériphériques contenant des erreurs au moyen de la commande `metareplace(1M)`, puis resynchronisez les disques.

```
# metareplace -s ensemble_disques-e miroir composant
```

L'option `-e` permet au composant (tranche) de passer à l'état de disponibilité et lance la resynchronisation.

Les composants ayant été remplacés par une réserve dynamique doivent être les derniers périphériques remplacés avec la commande `metareplace(1M)`. Si la réserve dynamique est remplacée en premier, elle pourrait remplacer un autre sous-miroir défectueux dès qu'elle sera disponible.

Vous ne pouvez resynchroniser qu'un seul composant de sous-miroir (métapériphérique) à la fois. Si tous les composants d'un sous-miroir ont été touchés par la panne de courant, chaque composant doit être remplacé séparément. La resynchronisation d'un disque de 1,05 Go demande environ 10 minutes.

Si les deux ensembles de disques d'une configuration symétrique ont été touchés par la panne de courant, vous pouvez resynchroniser en même temps les sous-miroirs touchés de chaque ensemble de disques. Connectez-vous séparément à chaque hôte afin de récupérer l'ensemble de disques de chacun d'eux en lançant la commande `metareplace(1M)` pour chacun.

Remarque - Selon le nombre de sous-miroirs et de composants figurant dans ces sous-miroirs, la resynchronisation peut demander beaucoup de temps. Un seul sous-miroir composé de 30 unités de 1,05 Go peut exiger environ cinq heures, tandis qu'une configuration comprenant des sous-miroirs à cinq composants peut n'exiger que 50 minutes.

▼ Comment effectuer une reprise après une panne de courant (VxVM)

Les pannes de courant peuvent provoquer la déconnexion des unités de disques et donc des périphériques virtuels, qui ne sont alors plus disponibles. Le volume demeure toutefois actif, car les périphériques virtuels qui restent dans un volume mis en miroir sont toujours disponibles. Il est possible de reconnecter les unités de disque et de remédier à la situation sans arrêter les noeuds de la grappe.

La liste qui suit énumère les étapes générales de la récupération suite à une panne de courant d'un boîtier de disque dans une configuration VxVM:

- Identification des périphériques virtuels défectueux à l'aide des commandes `vxprint` et `vxdisk`.
- Correction du problème qui a causé la panne de courant.
- Utilisation des commandes `drvconfig` et `disks` pour créer les entrées `/devices` et `/dev`.
- Analyse de la configuration de disque actuelle.
- Reconnexion des disques qui contenaient des erreurs temporaires.
- Vérification de la présence d'autres erreurs.
- (Facultatif) Dans le cas de groupes de disques partagés, exécution de la commande `vxdg` pour chaque disque dont l'alimentation a été coupée.
- Lancement de la récupération de volume.

Les étapes expliquées maintenant décrivent en détails la récupération suite à une panne de courant d'un boîtier de disque dans une configuration VxVM.

1. Exécutez la commande `vxprint` pour afficher les périphériques virtuels défectueux.

Vous pouvez également spécifier un groupe de disques avec l'option `-g groupe_disques`.

2. Exécutez la commande `vxdisk` pour identifier les disques contenant des erreurs.

```
# vxdisk list
DEVICE      TYPE      DISK      GROUP      STATUS
..
-           -         c1t5d0    toi        failed was:c1t5d0s2
...
```

3. Corrigez l'erreur qui a causé le problème afin de rétablir l'alimentation électrique de tous les disques en panne.

Vérifiez que les disques tournent avant de poursuivre.

4. Entrez les commandes suivantes sur tous les noeuds de la grappe.

Dans certains cas, les disques doivent être de nouveau reconnus par les noeuds.

```
# drvconfig
# disks
```

5. Entrez les commandes suivantes sur tous les noeuds de la grappe.

Le gestionnaire de volumes doit analyser de nouveau la configuration de disque actuelle.

```
# vxdctl enable
# vxdisk -a online
```

6. Entrez la commande suivante sur tous les noeuds de la grappe.

Remarque - Pour la fonction de grappe de VxVM (utilisée avec Oracle Parallel Server), entrez la commande d'abord sur le noeud maître, et ensuite seulement sur les noeuds restants.

Ainsi, vous reconnecterez les disques qui comportaient une anomalie temporaire et vous amorcerez leur récupération.

```
# vxreattach -r
```

7. Vérifiez la sortie de la commande `vxdisk` pour repérer toute autre erreur éventuelle.

```
# vxdisk list
```

8. Si le support a été remplacé, entrez la commande suivante depuis le noeud maître pour chaque disque ayant été déconnecté.

Le disque physique et le nom d'accès du gestionnaire de volumes de ce disque doivent être reconnectés.

```
# vxdg -g groupe_disques -k adddisk nom_support=nom_accès
```

Les valeurs pour `nom_support` et `nom_accès` figurent à la fin de la sortie de la commande `vxdisk list`.

Exemple :

```
# vxdg -g toi -k adddisk c1t5d0=c1t5d0s2
# vxdg -g toi -k adddisk c1t5d1=c1t5d1s2
# vxdg -g toi -k adddisk c1t5d2=c1t5d2s2
# vxdg -g toi -k adddisk c1t5d3=c1t5d3s2
# vxdg -g toi -k adddisk c1t5d4=c1t5d4s2
```

Vous pouvez également rattacher les disques à l'aide de la commande `vxdiskadm` ou de l'interface graphique.

9. Depuis le noeud, lancez la récupération du volume.

```
# vxrecover -bv [-g groupe_disques]
```

Si vous disposez de groupes de disques partagés, utilisez les options `-svc` de la commande `vxrecover`.

10. (Facultatif) Exécutez la commande `vxprint -g` pour afficher les modifications.

Administration des disques Sun StorEdge MultiPack et Sun StorEdge D1000

Cette section explique les procédures nécessaires à l'administration des composants Sun StorEdge MultiPack et Sun StorEdge D1000. Utilisez les procédures qui figurent dans le manuel du matériel du serveur pour identifier le composant défectueux.

Réparation d'une connexion Sun StorEdge MultiPack ou Sun StorEdge D1000

Lorsque la connexion entre un boîtier de disque et un noeud de la grappe est défectueuse, l'anomalie est sans doute imputable à un câble SCSI-2 ou une carte SBus défaillante.

Dans tous les cas, le noeud sur lequel la panne s'est produite se met à générer des erreurs lorsque la panne est découverte. Les accès subséquents au boîtier de disque produisent d'autres erreurs. Le noeud présente le même comportement que si l'alimentation du boîtier de disque avait été coupée.

Les opérations d'E/S effectuées depuis les autres noeuds de la grappe ne sont pas touchées par ce type de panne.

Pour diagnostiquer la panne, vous pouvez avoir recours aux procédures d'essai du module de carte qui figurent dans le manuel d'entretien du noeud Sun Cluster, afin de déterminer quel composant est défectueux. Vous devez normalement libérer un noeud et le boîtier de disque qui semble être en panne afin de dépanner le matériel.

▼ Comment réparer une connexion Sun StorEdge MultiPack ou Sun StorEdge D1000

1. Préparez le système Sun Cluster comme pour un remplacement de composants.

Selon la cause de la perte de connexion, préparez le noeud Sun Cluster en effectuant l'une des procédures suivantes.

- Si le composant défectueux est une carte SBus, voir le Chapitre 7, afin de préparer le noeud Sun Cluster en vue de la mise hors tension.
- Si le problème provient d'un câble SCSI-2 défectueux, le logiciel de gestion de volumes aura détecté l'anomalie et préparé le système en vue du remplacement du câble.

2. Remplacez le composant défectueux.

Si le câble SCSI-2 ou la carte SBus tombe en panne, reportez-vous au manuel d'entretien du noeud Sun Cluster pour les instructions sur le remplacement de ces éléments.

3. Corrigez les erreurs du logiciel de gestion de volumes.

Effectuez les procédures décrites à la section "Reprise après une panne de courant", page 235.

Ajout d'un disque Sun StorEdge MultiPack ou Sun StorEdge D1000

Vous pouvez ajouter des Sun StorEdge MultiPack ou Sun StorEdge D1000 à une configuration Sun Cluster à tout moment.

Vous devez passer en revue la configuration du groupe de disques dans votre configuration Sun Cluster avant d'ajouter un boîtier de disque. Les considérations exposées au Chapitre 2 du *Sun Cluster 2.2 Software Installation Guide* et à l'Annexe

A du présent manuel vous aideront à évaluer l'impact du boîtier de disque sur la configuration des groupes de disques.

▼ Comment ajouter un disque Sun StorEdge MultiPack ou Sun StorEdge D1000

1. Arrêtez un des noeuds de la grappe.

Utilisez la procédure décrite à la section “Arrêt de la grappe et des noeuds de grappe”, page 86 pour arrêter le noeud.

2. Au besoin, installez une carte SBus supplémentaire dans le noeud.

Suivez les instructions du manuel d'entretien du noeud Sun Cluster pour installer la carte SBus.

Remarque - Installez la carte SBus dans le premier emplacement SBus disponibles, après toutes les autres cartes du noeud. Ainsi, la numérotation du contrôleur sera préservée si l'environnement d'exploitation Solaris est réinstallé. Pour de plus amples informations à ce sujet, voir “Désignation et numérotation des instances”, page 24.

3. Branchez les câbles SCSI-2 au boîtier de disque.

Suivez les instructions du manuel d'entretien du noeud Sun Cluster.

4. Définissez l'ID de l'initiateur SCSI comme il convient.

Suivez les instructions du manuel d'entretien du noeud Sun Cluster.

5. Effectuez une réinitialisation de reconfiguration du noeud.

```
ok boot -r
```

6. Exécutez la commande `haswitch(1M)` afin que le noeud réinitialisé devienne propriétaire de tous les hôtes logiques pouvant faire l'objet d'une maîtrise.

```
phys-hahost1# haswitch phys-hahost2 hahost1 hahost2
```

7. Répétez Etape 1, page 242 jusqu'à Etape 5, page 242 pour les autres noeuds connectés à ce boîtier de disque.

- 8. Au besoin, réattribuez la propriété des hôtes logiques au maître par défaut approprié.**

Exemple :

```
phys-hahost1# haswitch phys-hahost2 hahost2
```

- 9. Ajoutez les disques des boîtiers de disque au groupe de disques sélectionné.**

Suivez les instructions de la documentation du gestionnaire de volumes pour ajouter les disques au(x) groupe(s) de disques sélectionné(s). Reportez-vous également aux annexes du *Sun Cluster 2.2 Software Installation Guide* pour plus de détails sur Solstice DiskSuite ou VxVM.

Administration des disques Sun StorEdge MultiPack et Sun StorEdge D1000

Dans le cadre de l'administration standard de Sun Cluster, il vous incombe de surveiller l'état de la configuration. Pour de plus amples informations sur les méthodes de surveillance, voir le Chapitre 2. Durant le processus de surveillance, vous pouvez découvrir des problèmes avec les disques multihôtes. Les procédures suivantes décrivent la correction de ces problèmes.

Sun Cluster prend en charge divers types de disque. Pour la description du boîtier de disque, reportez-vous au manuel d'entretien de l'unité d'expansion de disque multihôtes.

Ajout de disques Sun StorEdge MultiPack ou Sun StorEdge D1000

Dans une configuration symétrique, le boîtier de disque peut contenir des disques de plusieurs groupes de disques : un seul noeud doit être propriétaire de tous les groupes de disques visés.

▼ Comment ajouter un disque Sun StorEdge MultiPack ou Sun StorEdge D1000

La liste qui suit énumère les étapes générales d'ajout d'un disque Sun StorEdge MultiPack ou Sun StorEdge D1000:

- Identification du contrôleur de ce nouveau disque et d'un emplacement libre dans le boîtier de disque.
- Ajout du nouveau disque.
- Exécution des tâches administratives pour la préparation du disque pour son utilisation par Sun Cluster.
 - Création des fichiers spéciaux `/devices` et des liens `/dev/dsk` et `/dev/rdisk`.
 - Ajout du disque au groupe de disques.
 - Formatage et partitionnement du disque, au besoin.
 - Exécution des tâches administratives liées à la gestion de volumes.

Les étapes expliquées maintenant décrivent en détails l'ajout d'un disque Sun StorEdge MultiPack ou Sun StorEdge D1000.

1. Déterminez le numéro de contrôleur du boîtier de disque auquel le disque doit être ajouté.

Utilisez la commande `mount(1M)` ou `format(1M)` pour déterminer le numéro du contrôleur.

2. Repérez un emplacement de disque disponible et adapté dans le boîtier de disque en vue de l'ajout du disque.

Identifiez les emplacements libres en observant les DEL des unités de disque situées sur le devant du boîtier de disque ou en enlevant le couvercle gauche de l'unité. Les ID d'adresse qui correspondent aux emplacements figurent sur la partie médiane de la baie d'unité.

Dans les étapes suivantes, le plateau 2 est utilisé à titre d'exemple.

L'emplacement sélectionné pour le nouveau disque est le plateau 2, emplacement 7. Le nouveau disque porte la désignation `c2t3d1`.

3. Ajoutez le nouveau disque.

Suivez les instructions du manuel d'entretien du boîtier de disque afin d'ajouter le matériel (disque).

4. Exécutez les commandes `drvconfig(1M)` et `disks(1M)` pour créer les nouvelles entrées dans `/devices`, `/dev/dsk` et `/dev/rdisk` pour tous les nouveaux disques.

```
phys-hahost1# drvconfig
phys-hahost1# disks
```

5. **Transférez la propriété des hôtes logiques à l'autre noeud de la grappe auquel le disque est connecté.**

```
phys-hahost1# haswitch phys-hahost2 hahost1 hahost2
```

6. **Exécutez les commandes `drvconfig(1M)` et `disks(1M)` sur le noeud à présent propriétaire du groupe de disques auquel le disque va être ajouté.**

```
phys-hahost2# drvconfig
phys-hahost2# disks
```

7. **Ajoutez le disque à un groupe de disques à l'aide du logiciel de gestion de volumes.**

Pour Solstice DiskSuite, la syntaxe de la commande est la suivante, où *ensemble_disques* représente le nom de l'ensemble de disques contenant le disque défectueux et *unité* le nom DID du disque au format *dN* (pour les nouvelles installations de Sun Cluster), ou *cNtYdZ* (pour les installations ayant fait l'objet d'une mise à niveau à partir de HA 1.3) :

```
# metaset -s ensemble_disques -a unité
```

Pour VxVM, vous pouvez utiliser la ligne de commande ou d'interface graphique pour ajouter le disque au groupe de disques.



Attention - Si vous utilisez Solstice DiskSuite, la commande `metaset(1M)` peut repartitionner automatiquement ce disque. Pour de plus amples informations, voir la documentation de Solstice DiskSuite.

8. **(Configuration Solstice DiskSuite exclusivement) Une fois les disques ajoutés à l'ensemble de disques avec la commande `metaset(1M)`, exécutez la commande `scadmin(1M)` pour réserver et activer le mode failfast sur les disques spécifiés.**

```
phys-hahost1# scadmin reserve nom_unité
```

9. Exécutez les tâches administratives habituelles pour le nouveau disque.

Vous pouvez maintenant effectuer les étapes d'administration normales qu'on doit exécuter lorsqu'une nouvelle unité est mise en service. Pour de plus amples renseignements sur ces tâches, reportez-vous à la documentation du logiciel de gestion de volumes.

10. Au besoin, réattribuez les hôtes logiques à leurs maîtres par défaut.

Remplacement des disques Sun StorEdge MultiPack ou Sun StorEdge D1000

Cette section décrit le remplacement d'un disque multihôtes sans interruption des services de Sun Cluster (remplacement en ligne) lorsque le gestionnaire de volumes signale des problèmes tels que :

- Composant dans l'état `Maintenance exigée`.
- Remplacement de réserve dynamique.
- Erreurs de disque intermittentes.

Consultez la documentation du logiciel de gestion de volumes pour connaître les méthodes de remplacement hors ligne.

▼ Comment remplacer un disque Sun StorEdge MultiPack ou Sun StorEdge D1000(Solstice DiskSuite)

Utilisez la procédure qui suit si vous avez déterminé que des composants d'un disque sont dans l'état `Maintenance exigée`, si une réserve dynamique a remplacé un composant ou si un disque génère des erreurs intermittentes.

La liste qui suit énumère les étapes générales du remplacement d'un disque Sun StorEdge MultiPack ou Sun StorEdge D1000 dans une configuration Solstice DiskSuite.

- Identification du disque qui doit être remplacé.
- Identification de l'unité d'expansion de disque qui renferme le disque à remplacer.
- Retrait du disque défectueux de l'ensemble de disques.
- Arrêt du disque et ouverture du boîtier de disque.
- Remplacement de l'unité de disque.
- Exécution de la commande `scdidadm -R`.
- Ajout du nouveau disque à l'ensemble de disques.
- Mise en réserve et activation du mode `failfast` sur le disque.

- Partitionnement du nouveau disque.
- Exécution de la commande `metastat(1M)` pour vérifier si le problème a été corrigé.

Les étapes expliquées maintenant décrivent en détails le remplacement d'un disque Sun StorEdge MultiPack ou Sun StorEdge D1000 défectueux dans une configuration Solstice DiskSuite.

1. Appliquez la procédure sur l'hôte maître de l'ensemble de disques dans lequel se trouve le disque défectueux. Pour cela, vous devrez peut-être commuter sur l'ensemble de disques à l'aide de la commande `haswitch(1M)`.

2. Déterminez quel disque doit être remplacé.

Exécutez la commande `metastat(1M)` et produisez la sortie `/var/adm/messages`.

Si la commande `metastat(1M)` signale qu'un périphérique est en état de maintenance ou que des composants ont été remplacés par des réserves dynamiques, vous devez trouver et remplacer le périphérique en question. Un exemple de sortie de la commande `metastat(1M)` est présenté ci-après. Dans cet exemple, le périphérique `c3t3d4s0` est en état de maintenance :

```
phys-hahost1# metastat -s hahost1
...
d50:Submirror of hahost1/d40
  State: Needs Maintenance
  Stripe 0:
    Device      Start Block    Dbase    State    Hot Spare
    c3t3d4s0    0              No       Okay     c3t5d4s0
...
```

Vérifiez `/var/adm/messages` afin de déterminer quel type de problème a été détecté.

```
...
Jun 1 16:15:26 host1 unix:
WARNING: /io-unit@f,e1200000/sbi@0.0/SUNW,pln@a0000000,741022/ssd@3,4(ssd49):
Jun 1 16:15:26 host1 unix: Error for command 'write(I)'. Err
Jun 1 16:15:27 host1 unix: or Level: Fatal
Jun 1 16:15:27 host1 unix: Requested Block 144004, Error Block: 715559
Jun 1 16:15:27 host1 unix: Sense Key: Media Error
Jun 1 16:15:27 host1 unix: Vendor 'CONNER':
Jun 1 16:15:27 host1 unix: ASC=0x10(ID CRC or ECC error),ASCQ=0x0,FRU=0x15
...
```

3. Déterminez l'emplacement du disque défectueux.

Utilisez la commande `mount(1M)` ou `format(1M)` pour déterminer le numéro du contrôleur.

4. Si le disque défectueux contient des répliques, prenez note de la tranche et du numéro, puis supprimez les répliques.

Exécutez la commande `metadb(1M)` pour supprimer les répliques.

5. Déconnectez tous les sous-miroirs avec composants sur le disque à remplacer.

Si vous déconnectez un sous-miroir qui comporte un composant défectueux, vous devez commander la déconnexion au moyen de l'option `metadetach -f`.

L'exemple suivant déconnecte le sous-miroir `d50` du métamiroir `d40`.

```
phys-hahost1# metadetach -s hahost1 -f d40 d50
```

6. Exécutez la commande `metaclear(1M)` afin de vider les sous-miroirs déconnectés à Etape 5, page 248.

```
phys-hahost1# metaclear -s hahost1 -f d50
```

7. Si le disque défectueux contient des réserves dynamiques, prenez note du nom des périphériques et de la liste des périphériques qui contiennent des pools de réserves dynamiques, puis supprimez les réserves dynamiques.

Exécutez la commande `metahs(1M)` pour supprimer les réserves dynamiques.



Attention - Vous devez prendre note des informations avant de supprimer les objets, afin de pouvoir procéder dans l'ordre inverse des opérations suite au remplacement du disque.

8. Exécutez la commande `metaset(1M)` afin de supprimer le disque défectueux de l'ensemble de disques.

La syntaxe de la commande est la suivante, où *ensemble_disques* représente le nom de l'ensemble de disques contenant le disque défectueux et *unité* le nom DID du disque dans le format `dN` (pour les nouvelles installations de Sun Cluster) ou `cNt YdZ` (pour les installations mises à niveau à partir de HA 1.3) :

```
phys-hahost1# metaset -s ensemble_disques -d unité
```

Cette opération peut nécessiter 15 minutes ou plus, selon la taille de la configuration et le nombre de disques.

9. Remplacez le disque défectueux.

Pour plus de détails, reportez-vous aux manuels d'entretien du boîtier de disque.

10. Vérifiez si le nouveau disque tourne.

Le disque doit tourner automatiquement.

11. Mettez à jour la base de données de pilotes DID en précisant l'ID du nouveau périphérique.

Remarque - Si vous avez effectué une mise à niveau depuis HA 1.3, le système n'utilise pas le programme pilote DID. Si c'est le cas, sautez cette étape.

Utilisez l'indicateur `-l` de la commande `sccidadm(1M)` pour identifier le nom DID du périphérique de bas niveau de l'unité à remplacer. Mettez ensuite à jour la base de données des pilotes DID en utilisant l'indicateur `-R` avec la commande `sccidadm(1M)`. Consultez le *Sun Cluster 2.2 Software Installation Guide* pour des détails complémentaires sur le pseudo-pilote DID.

```
phys-hahost1# sccidadm -o name -l /dev/rdisk/c3t3d4
6 phys-hahost1:/dev/rdisk/c3t3d4 /dev/did/rdsk/d6
phys-hahost1# sccidadm -R d6
```

12. Rajoutez le nouveau disque dans l'ensemble de disques à l'aide de la commande `metaset(1M)`.

Cette étape rajoute automatiquement le nombre adéquat de répliques qui ont été supprimées du disque défectueux. La syntaxe de la commande est indiquée ci-après. Dans cet exemple, *ensemble_disques* représente le nom de l'ensemble de disques contenant le disque défectueux et *unité* le nom DID du disque dans le format `dN` (pour les nouvelles installations de Sun Cluster) ou `cNtYdZ` (pour les installations mises à niveau à partir de la version HA 1.3).

```
phys-hahost1# metaset -s ensemble_disques -a unité
```

Cette opération peut demander 15 minutes et même davantage, selon la taille de la configuration et le nombre de disques.

13. Exécutez la commande `scadmin(1M)` pour réserver et activer le mode failfast sur le disque spécifié qui vient d'être rajouté à l'ensemble de disques.

```
phys-hahost1# scadmin reserve c3t3d4
```

14. Exécutez la commande `format(1M)` ou `fmthard(1M)` pour repartitionner le nouveau disque.

Veillez à partitionner le nouveau disque exactement de la même manière que l'ancien. (L'enregistrement des données de format du disque est recommandé au Chapitre 1.)

15. Exécutez la commande `metainit(1M)` pour réinitialiser les disques supprimés à l'Étape 6, page 248.

```
phys-hahost1# metainit -s hahost1 d50
```

16. Attachez les sous-miroirs déconnectés à l'Étape 5, page 248.

Pour ce faire, exécutez la commande `metattach(1M)`. Pour plus de détails, consultez la page de manuel `metattach(1M)`.

```
phys-hahost1# metattach -s hahost1 d40 d50
```

17. Restaurez toutes les réserves dynamiques supprimées à l'Étape 7, page 248.

Exécutez la commande `metahs(1M)` pour rajouter les réserves dynamiques. Pour plus de détails, consultez la page de manuel `metahs(1M)`.

```
phys-hahost1# metahs -s hahost1 -a hsp000 c3t2d5s0
```

18. Vérifiez si le remplacement a remédié au problème.

```
phys-hahost1# metastat -s hahost1
```

▼ Comment remplacer un disque Sun StorEdge MultiPack ou Sun StorEdge D1000 (VxVM)

La liste qui suit énumère les étapes générales du remplacement d'un disque Sun StorEdge MultiPack ou Sun StorEdge D1000 dans une configuration VxVM.

- Retrait du disque défectueux du boîtier de disque au moyen de la commande `vxdiskadm`.
- Remplacement du disque défectueux.
- Remplacement du disque retiré auparavant au moyen de la commande `vxdiskadm`.

Remarque - Pour les systèmes n'utilisant pas des groupes de disques partagés, le noeud maître est le noeud qui a importé le groupe de disques.

1. **Si vous utilisez des groupes de disques partagés, déterminez le noeud maître et le noeud esclave en entrant la commande suivante sur tous les noeuds de la grappe :**

```
# vxctl -c mode
```

Remarque - Effectuez les étapes suivantes depuis le noeud maître.

2. **Déterminez si le disque en question a connu des pannes et s'il est en état NODEVICE.**
Si ce n'est pas le cas, passez à l'Etape 8, page 253.
3. **Exécutez l'utilitaire vxdiskadm puis entrez 4 (Retirer un disque en vue du remplacement).**
Cette option retire un disque physique tout en conservant son nom. L'utilitaire vous demande ensuite de spécifier le périphérique à remplacer.
4. **Entrez le nom du disque ou list.**
L'exemple suivant illustre le retrait du disque c2t8d0.

```
Enter disk name [<disk>,list,q,?] list
Disk group: rootdg
DM NAME      DEVICE      TYPE      PRIVLEN  PUBLLEN  STATE
...
Disk group: demo
DM NAME      DEVICE      TYPE      PRIVLEN  PUBLLEN  STATE
dm c1t2d0    c2t2d0s2    sliced    1519     4152640  -
dm c1t3d0    c2t3d0s2    sliced    1519     4152640  -
dm c1t4d0    c2t4d0s2    sliced    1519     4152640  -
dm c1t5d0    c2t5d0s2    sliced    1519     4152640  -
dm c1t8d0    c2t8d0s2    sliced    1519     4152640  -
dm c1t9d0    c2t9d0s2    sliced    1519     4152640  -
dm c2t2d0    c1t2d0s2    sliced    1519     4152640  -
dm c2t3d0    c1t3d0s2    sliced    1519     4152640  -
dm c2t4d0    c1t4d0s2    sliced    1519     4152640  -
dm c2t5d0    c1t5d0s2    sliced    1519     4152640  -
dm c2t8d0    c1t8d0s2    sliced    1519     4152640  -
dm c2t9d0    c1t9d0s2    sliced    1519     4152640  -
Enter disk name [<disk>,list,q,?] c2t8d0
The requested operation is to remove disk c2t8d0 from disk group
demo. The disk name will be kept, along with any volumes using
the disk, allowing replacement of the disk.
```

(suite)

(suite)

```
Select "Replace a failed or removed disk" from the main menu
when you wish to replace the disk.
```

5. Tapez **y** ou Entrée pour continuer.

```
Continue with operation? [y,n,q,?] (default: y) y
Removal of disk c2t8d0 completed successfully.
```

6. Tapez **q** pour quitter l'utilitaire.

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

7. Tapez **vxdisk list** et **vxprint** pour visualiser les changements.

Le disque exemple **c2t8d0** a été retiré.

```
# vxdisk list
.
c2t3d0s2    sliced    c1t3d0     demo       online shared
c2t4d0s2    sliced    c1t4d0     demo       online shared
c2t5d0s2    sliced    c1t5d0     demo       online shared
c2t8d0s2    sliced    c1t8d0     demo       online shared
c2t9d0s2    sliced    c1t9d0     demo       online shared
-           -         c2t8d0     demo       removed
# vxprint
.
dm c2t3d0    c1t3d0s2  -          4152640   -         -         -
dm c2t4d0    c1t4d0s2  -          4152640   -         -         -
dm c2t5d0    c1t5d0s2  -          4152640   -         -         -
dm c2t8d0    -          -          -          -         REMOVED  -         -
dm c2t9d0    c1t9d0s2  -          4152640   -         -         -
pl demo05-02 -          DISABLED  51200     -         REMOVED  -         -
sd c2t8d0-1  demo05-02 DISABLED  51200     0         REMOVED  -         -
.
.
.
```

8. Remplacez l'unité physique sans mettre hors tension aucun composant.

Pour plus de détails, reportez-vous à la documentation fournie avec le boîtier de disque.

Remarque - Pendant que vous remplacez l'unité, des messages semblables à ceux de l'exemple suivant peuvent apparaître sur la console système. Ne vous alarmez pas, car ces messages ne révèlent pas la présence d'un problème. Poursuivez plutôt le remplacement, conformément aux étapes décrites ci-après.

```
3 nov 17:44:00 updb10a unix: WARNING: /sbus@1f,0/SUNW,fas@0,8800000/sd@2,0 (sd17) :
3 nov 17:44:00 updb10a unix : SCSI transport failed: reason "incomplete": \
retrying command
3 Nov 17:44:03 updb10a unix: WARNING: /sbus@1f,0/SUNW,fas@0,8800000/sd@2,0 (sd17) :
3 nov 17:44:03 updb10a unix : disk not responding to selection
```

9. Exécutez l'utilitaire vxdiskadm puis tapez 5 (Remplacer un disque défectueux ou retiré).

10. Entrez le nom du disque.

Vous pouvez taper `list` pour afficher la liste des disques dont l'état est `REMOVED`.

Remarque - L'état apparent du disque peut être `NODEVICE` s'il a subi des pannes.

```
Select a removed or failed disk [<disk>,list,q,?] list
Disk group: rootdg DM NAME          DEVICE          TYPE          PRIVLEN  PUBLN
STATE ... Disk group: demo DM NAME          DEVICE          TYPE          PRIVLEN  PUBLN
STATE dm c2t8d0                    -              -              -          REMOVED
Select a removed or failed disk [<disk>,list,q,?] c2t8d0
```

L'utilitaire `vxdiskadm` détecte le nouveau périphérique et vous demande si celui-ci doit remplacer le périphérique retiré.

Remarque - Si d'autres disques inutilisés sont reliés au système, l'utilitaire `vxdiskadm` propose également ces disques à titre de choix valable.

11. Entrez le nom du périphérique ou, si l'utilitaire indique le périphérique par défaut, appuyez sur Entrée.

```
The following devices are available as replacements:
  c1t8d0s2
  You can choose one of these disks to replace c2t8d0.
  Choose "none" to initialize another disk to replace c2t8d0.
  Choose a device, or select "none"
  [<device>,none,q,?] (default: c1t8d0s2) <Return>
  The requested operation is to use the initialized device c1t8d0s2
  to replace the removed or failed disk c2t8d0 in disk group demo.
```

12. Tapez y ou Entrée pour confirmer votre volonté de faire de ce périphérique (dans l'exemple, c1t8d0s2) le disque de remplacement.

```
Continue with operation? [y,n,q,?] (default: y) <Entrée>
  Replacement of disk c2t8d0 in group demo with disk device
  c1t8d0s2 completed successfully.
```

13. Tapez n ou Entrée pour quitter l'utilitaire.

```
Replace another disk? [y,n,q,?] (default: n) <Entrée>
```

14. Tapez vxdisk list et vxprint pour visualiser les changements.

L'état du disque d'exemple, c2t8d0, n'est plus REMOVED.

```
# vxdisk list
...
c2t2d0s2    sliced    c1t2d0    demo      online shared
c2t3d0s2    sliced    c1t3d0    demo      online shared
c2t4d0s2    sliced    c1t4d0    demo      online shared
c2t5d0s2    sliced    c1t5d0    demo      online shared
c2t8d0s2    sliced    c1t8d0    demo      online shared
c2t9d0s2    sliced    c1t9d0    demo      online shared
# vxprint
...
dm c2t4d0    c1t4d0s2  -         4152640  -         -         -
dm c2t5d0    c1t5d0s2  -         4152640  -         -         -
dm c2t8d0    c1t8d0s2  -         4152640  -         -         -
dm c2t9d0    c1t9d0s2  -         4152640  -         -         -
```

(suite)

(suite)

```
...
```

Remplacement des boîtiers Sun StorEdge MultiPack ou Sun StorEdge D1000

Cette section explique comment remplacer un boîtier complet Sun StorEdge MultiPack ou Sun StorEdge D1000 exécutant VxVM.

▼ Comment remplacer un boîtier Sun StorEdge MultiPack ou Sun StorEdge D1000 (VxVM)

La liste qui suit énumère les étapes générales du remplacement d'un boîtier complet Sun StorEdge MultiPack ou Sun StorEdge D1000 défectueux dans une configuration VxVM:

- Retrait de tous les disques du boîtier de disque défectueux au moyen de la commande `vxdiskadm`
- Remplacement du boîtier de disque défectueux
- Remise en place de tous les disques précédemment retirés dans le nouveau boîtier de disque au moyen de la commande `vxdiskadm`

Remarque - Pour les systèmes n'utilisant pas des groupes de disques partagés, le noeud maître est le noeud qui a importé le groupe de disques.

1. Si vous utilisez des groupes de disques partagés, déterminez le noeud maître et le noeud esclave en entrant la commande suivante sur tous les noeuds de la grappe :

```
# vxctl -c mode
```

Remarque - Effectuez les étapes suivantes depuis le noeud maître.

2. Retirez tous les disques du boîtier de disque défectueux en exécutant l'utilitaire `vxdiskadm` et en tapant 4 (Retirer un disque en vue d'un remplacement).

Remarque - Cette option ne permet de retirer qu'un seul disque à la fois. Répétez cette procédure pour chaque disque.

3. Tapez la commande `list`.

Dans l'exemple suivant, on suppose que le boîtier de disque du contrôleur `c2` doit être remplacé. Selon le résultat de la commande `list`, les noms VxVM de ces disques sont `c2t2d0`, `c2t3d0`, `c2t4d0`, `c2t5d0`, `c2t8d0` et `c2t9d0`.

```
Remove a disk for replacement
Menu: VolumeManager/Disk/RemoveForReplace
  Use this menu operation to remove a physical disk from a disk
  group, while retaining the disk name. This changes the state
  for the disk name to a "removed" disk. If there are any
  initialized disks that are not part of a disk group, you will be
  given the option of using one of these disks as a replacement.
Enter disk name [<disk>,list,q,?] list
Disk group: rootdg
DM NAME          DEVICE          TYPE          PRIVLEN     PUBLLEN     STATE
...
Disk group: demo
DM NAME          DEVICE          TYPE          PRIVLEN     PUBLLEN     STATE
dm c1t2d0        c2t2d0s2        sliced        1519        4152640     -
dm c1t3d0        c2t3d0s2        sliced        1519        4152640     -
dm c1t4d0        c2t4d0s2        sliced        1519        4152640     -
dm c1t5d0        c2t5d0s2        sliced        1519        4152640     -
dm c1t8d0        c2t8d0s2        sliced        1519        4152640     -
dm c1t9d0        c2t9d0s2        sliced        1519        4152640     -
dm c2t2d0        c1t2d0s2        sliced        1519        4152640     -
dm c2t3d0        c1t3d0s2        sliced        1519        4152640     -
dm c2t4d0        c1t4d0s2        sliced        1519        4152640     -
dm c2t5d0        c1t5d0s2        sliced        1519        4152640     -
dm c2t8d0        c1t8d0s2        sliced        1519        4152640     -
dm c2t9d0        c1t9d0s2        sliced        1519        4152640     -
```

4. Tapez le nom du disque (dans cet exemple, `c2t2d0`).

```
Enter disk name [<disk>,list,q,?] c2t2d0
The following volumes will lose mirrors as a result of this
operation:
  demo-1
No data on these volumes will be lost.
The requested operation is to remove disk c2t2d0 from disk group
demo. The disk name will be kept, along with any volumes using
```

(suite)

(suite)

```
the disk, allowing replacement of the disk.  
Select "Replace a failed or removed disk" from the main menu  
when you wish to replace the disk.
```

5. Tapez **y** ou **Entrée** pour confirmer votre volonté de remplacer le disque.

```
Continue with operation? [y,n,q,?] (default: y) <Entrée>  
Removal of disk c2t2d0 completed successfully.
```

6. Tapez **y** pour continuer.

```
Remove another disk? [y,n,q,?] (default: n) y  
Remove a disk for replacement  
Menu: VolumeManager/Disk/RemoveForReplace  
Use this menu operation to remove a physical disk from a disk  
group, while retaining the disk name. This changes the state  
for the disk name to a "removed" disk. If there are any  
initialized disks that are not part of a disk group, you will be  
given the option of using one of these disks as a replacement.
```

7. Tapez le nom du disque suivant dans l'exemple, **c2t3d0**.

```
Enter disk name [<disk>,list,q,?] c2t3d0  
The following volumes will lose mirrors as a result of this  
operation:  
demo-2  
No data on these volumes will be lost.  
The following devices are available as replacements:  
c1t2d0  
You can choose one of these disks now, to replace c2t3d0.  
Select "none" if you do not wish to select a replacement disk.
```

8. Tapez **none** au besoin.

Remarque - Cette interrogation survient chaque fois que l'utilitaire repère un disque correct dans le système. Si aucun disque correct n'est trouvé, cette interrogation n'apparaît pas.

```
Choose a device, or select "none"
[<device>,none,q,?] (default: c1t2d0) none
```

9. Tapez y ou Entrée pour confirmer votre volonté de supprimer le disque.

```
The requested operation is to remove disk c2t3d0 from disk group
demo. The disk name will be kept, along with any volumes using
the disk, allowing replacement of the disk.
Select "Replace a failed or removed disk" from the main menu
when you wish to replace the disk.
Continue with operation? [y,n,q,?] (default: y) <Entrée>
Removal of disk c2t3d0 completed successfully.
```

10. Répétez l'Etape 6, page 257 jusqu'à l'Etape 9, page 258 pour chaque disque indiqué à l'Etape 3, page 256.

11. Mettez le boîtier de disque hors tension puis remplacez-le.

Pour de plus amples renseignements, reportez-vous à la documentation du boîtier de disque.

Remarque - Pendant que vous remplacez le boîtier de disque, des messages semblables à ceux de l'exemple suivant peuvent être affichés sur la console système. Ne vous alarmez pas, car ces messages ne révèlent pas la présence d'un problème. Poursuivez plutôt le remplacement, conformément aux étapes décrites dans la section suivante.

```
3 nov 17:44:00 updb10a unix: WARNING: /sbus@1f,0/SUNW,fas@0,8800000/sd@2,0 (sd17) :
3 nov 17:44:00 updb10a unix : SCSI transport failed: reason "incomplete": \
retrying command
3 nov 17:44:03 updb10a unix: WARNING: /sbus@1f,0/SUNW,fas@0,8800000/sd@2,0 (sd17) :
3 nov 17:44:03 updb10a unix : disk not responding to selection
```

12. Mettez le boîtier de disque sous tension.

Pour de plus amples renseignements, reportez-vous au manuel d'entretien du boîtier de disque.

13. Connectez tous les disques précédemment retirés en exécutant l'utilitaire `vxdiskadm` et en tapant 5 (Remplacer un disque défectueux ou retiré).

Remarque - Cette option ne permet de remplacer qu'un seul disque à la fois. Répétez cette procédure pour chaque disque.

14. Exécutez la commande `list` pour afficher la liste des noms de disque dont l'état est à présent `REMOVED`.

```
Replace a failed or removed disk
Menu: VolumeManager/Disk/ReplaceDisk
Use this menu operation to specify a replacement disk for a disk
that you removed with the "Remove a disk for replacement" menu
operation, or that failed during use. You will be prompted for
a disk name to replace and a disk device to use as a replacement.
You can choose an uninitialized disk, in which case the disk will
be initialized, or you can choose a disk that you have already
initialized using the Add or initialize a disk menu operation.
Select a removed or failed disk [<disk>,list,q,?] list
Disk group: rootdg
DM NAME          DEVICE          TYPE          PRIVLEN  PUBLLEN  STATE
...
Disk group: demo
DM NAME          DEVICE          TYPE          PRIVLEN  PUBLLEN  STATE
```

(suite)

(suite)

dm c2t2d0	-	-	-	-	REMOVED
dm c2t3d0	-	-	-	-	REMOVED
dm c2t4d0	-	-	-	-	REMOVED
dm c2t5d0	-	-	-	-	REMOVED
dm c2t8d0	-	-	-	-	REMOVED
dm c2t9d0	-	-	-	-	REMOVED

15. Tapez le nom du disque (dans l'exemple, c2t2d0).

```
Select a removed or failed disk [<disk>,list,q,?] c2t2d0
The following devices are available as replacements:
c1t2d0s2 c1t3d0s2 c1t4d0s2 c1t5d0s2 c1t8d0s2 c1t9d0s2
```

L'utilitaire `vxdiskadm` détecte les nouveaux périphériques et vous demande si ceux-ci doivent remplacer les périphériques retirés.

16. Tapez le nom du périphérique de “remplacement” ou du “nouveau” périphérique ou, si l'utilitaire indique le périphérique par défaut, appuyez sur Entrée.

```
You can choose one of these disks to replace c2t2d0.
Choose "none" to initialize another disk to replace c2t2d0.

Choose a device, or select "none"
[<device>,none,q,?] (default: c1t2d0s2) <Entrée>
```

17. Tapez `y` ou Entrée pour confirmer votre volonté de faire de ce périphérique (dans l'exemple, c1t8d0s2) le disque de remplacement.

```
The requested operation is to use the initialized device c1t2d0s2
to replace the removed or failed disk c2t2d0 in disk group demo.
Continue with operation? [y,n,q,?] (default: y) <Entrée>
Replacement of disk c2t2d0 in group demo with disk device
c1t2d0s2 completed successfully.
```

18. Tapez y pour continuer.

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

Répétez l'Étape 15, page 260 jusqu'à l'Étape 18, page 261 pour chacun des noms de disque ayant l'état REMOVED/NODEVICE.

Administration des disques Sun StorEdge A3x00 et Sun StorEdge A1000

Ce chapitre explique les instructions nécessaires à l'administration des unités d'expansion et les disques Sun StorEdge A3x00 et Sun StorEdge A1000 utilisés dans votre configuration Sun Cluster. Les deux unités d'expansion prennent en charge le matériel RAID5. L'unité Sun StorEdge A3x00 comporte deux contrôleurs RAID5 et l'unité Sun StorEdge A1000 un seul. La plupart des procédures d'administration s'appliquent aux deux unités.

- “Sources d'alimentation”, page 263
- “Ajout de boîtiers Sun StorEdge A3x00 ou Sun StorEdge A1000”, page 264
- “Administration des disques Sun StorEdge A3x00 ou Sun StorEdge A1000”, page 265

Lorsque vous remplacez ou réparez le matériel des unités de disque dans une configuration Sun Cluster, reportez-vous toujours au manuel d'entretien des disques Sun StorEdge A3x00 ou Sun StorEdge A1000 et à la documentation accompagnant votre logiciel de gestion des volumes.

Sources d'alimentation

Les unités d'expansion Sun StorEdge A3x00 et Sun StorEdge A1000 comportent des séquenceurs d'alimentation redondants. Comme chacun des séquenceurs assure l'alimentation de la moitié des composants de l'unité d'expansion, le système reste disponible même si l'un d'eux tombe en panne. Le module de contrôleur dispose également de sources d'alimentation redondantes. Par conséquent, la disponibilité du système est assurée en cas de panne unique dans ce module.

Aucune procédure particulière n'est nécessaire sous Sun Cluster pour la reprise en cas de panne de la totalité de l'unité d'expansion. Pour remettre l'unité en service, il suffit de suivre les instructions fournies dans le manuel d'entretien de l'unité d'expansion.

Ajout de boîtiers Sun StorEdge A3x00 ou Sun StorEdge A1000

Cette section explique les procédures utilisées pour ajouter une unité d'expansion Sun StorEdge A3x00 ou Sun StorEdge A1000. Il n'est pas nécessaire d'arrêter Sun Cluster pour exécuter ces procédures.

Remarque - Pour la mise à niveau du microprogramme des unités d'expansion, consultez le manuel d'entretien matériel de votre système.

▼ Comment ajouter un boîtier de disque

1. **Commutez tous les hôtes logiques sur un des noeuds Sun Cluster de la grappe qui doit recevoir le nouveau boîtier de disque.**

```
phys-hahost1# haswitch phys-hahost2 hahost1 hahost2
```

2. **Arrêtez Sun Cluster sur le noeud qui ne contrôle plus aucun hôte logique.**

```
phys-hahost1# scadmin stopnode
```

3. **Ajoutez la carte d'interface de disque (UDWIS) au noeud Sun Cluster.**
Suivez les instructions accompagnant la carte d'interface UDWIS.

4. **Câblez la nouvelle carte UDWIS au boîtier de disque.**

5. **Réinitialisez le noeud contenant la nouvelle carte UDWIS pour appliquer la nouvelle configuration.**

```
phys-hahost1# boot -r
```

6. Démarrez Sun Cluster sur le noeud.

```
phys-hahost1# scadmin startnode
```

7. Commutez les hôtes logiques sur un autre noeud de la grappe.

```
phys-hahost1# haswitch phys-hahost1 hahost1 hahost2
```

8. Exécutez la procédure de l'Etape 2, page 264 à l' Etape 6, page 265 sur le noeud qui contrôlait précédemment les hôtes logiques.

9. Configurez les sous-systèmes de disque dans la grappe.

Configurez les disques conformément à la documentation de l'unité d'expansion et de VxVM.

10. Exécutez la commande `haswitch(1M)` sur les deux noeuds pour appliquer la nouvelle configuration à Sun Cluster.

```
# haswitch -r
```

Administration des disques Sun StorEdge A3x00 ou Sun StorEdge A1000

Les procédures d'administration des disques Sun StorEdge A3x00 ou Sun StorEdge A1000 dans une configuration Sun Cluster sont identiques à celles utilisées pour les noeuds qui ne sont pas en grappe. Reportez-vous à la documentation accompagnant votre unité d'expansion pour savoir comment ajouter, remplacer ou réparer des disques ou des composants de disques dans une unité d'expansion.

Ajout d'un disque Sun StorEdge A3x00 ou Sun StorEdge A1000

Vous pouvez ajouter des disques à une unité Sun StorEdge A3x00 ou Sun StorEdge A1000 lorsque le système fonctionne. Ne réinitialisez pas le système. En effet, vous

risqueriez alors de perdre des informations de configuration sur les nouveaux disques, de même que des données et des éléments de configuration d'unités logiques sur les disques existants.

Si un problème quelconque survient, comme par exemple des unités logiques préexistantes invisibles après l'ajout de disques et le redémarrage du système, retirez les disques que vous venez d'ajouter, redémarrez le système, puis ajoutez-les un à un.

Cela ne se produit pas lorsque les disques sont ajoutés à un groupe de disques "inutilisé" sur un système déjà lancé (le système peut ensuite être réinitialisé sans aucun problème). Cette procédure convient puisque les unités Sun StorEdge A3x00 et Sun StorEdge A1000 prennent toutes deux en charge les branchements dynamiques.

Remplacement d'un disque Sun StorEdge A3x00 ou Sun StorEdge A1000

Sun Cluster considère les disques Sun StorEdge A3x00 et Sun StorEdge A1000 comme des unités logiques et non comme des disques physiques. Par conséquent :

- Tant que l'unité logique est disponible, aucune intervention n'est nécessaire lors du remplacement d'un disque physique défectueux.
- Si l'unité logique n'est pas disponible (ou dans un état optimal) et qu'elle a été utilisée comme périphérique de quorum, vous devez exécuter la commande `scconf -q` pour désigner une autre unité logique (disque) comme périphérique de quorum avant de poursuivre la procédure de remplacement.

Pour de plus amples renseignements sur les unités logiques et les disques physiques, reportez-vous à la documentation accompagnant votre unité d'expansion. Consultez également les Chapitre 3 et 9 du *Sun Cluster 2.2 Hardware Service Manual*.

Il faut également savoir que certaines tâches administratives sur les disques Sun StorEdge A3x00 et Sun StorEdge A1000 nécessitent le remplacement des cartes UDWIS. Pour cette procédure, consultez le Chapitre 12 du *Sun Cluster 2.2 Hardware Service Manual*.

Administration des disques Sun StorEdge A5000

Ce chapitre décrit les instructions à suivre pour l'administration des disques Sun StorEdge A5000.

- “Reprise après une panne de courant”, page 267
- “Administration des disques Sun StorEdge A5000”, page 272
- “Administration des disques Sun StorEdge A5000”, page 273

Lorsque vous remplacez ou réparez le matériel des unités de disque dans une configuration Sun Cluster, consultez le manuel d'entretien des disques Sun StorEdge A5000 et la documentation accompagnant votre logiciel de gestion des volumes.

Reprise après une panne de courant

Lorsqu'un disque Sun StorEdge A5000 cesse d'être alimenté, les opérations d'E/S génèrent des erreurs que détecte votre gestionnaire de volumes. Les erreurs ne sont pas signalées tant que des transactions d'E/S ne sont pas effectuées sur le disque.

Vous devez surveiller la configuration de ces événements à l'aide des commandes décrites au Chapitre 2.

▼ Comment effectuer une reprise après une panne de courant (Solstice DiskSuite)

La liste qui suit énumère les étapes générales de la récupération suite à une panne de courant d'un boîtier de disque dans un environnement Solstice DiskSuite :

- Identification des répliques défectueuses
- Remise en service des répliques défectueuses
- Identification des périphériques défectueux
- Remise en service des périphériques défectueux
- Resynchronisation des disques

Les étapes expliquées maintenant décrivent en détails la récupération suite à une panne de courant d'un boîtier de disque dans l'environnement Solstice DiskSuite.

1. **Une fois l'alimentation rétablie, exécutez la commande `metadb(1M)` pour identifier les répliques contenant des erreurs :**

```
# metadb -s ensemble_disques
```

2. **Remettez en service les répliques.**

Après la panne de courant, toutes les répliques de base de données d'état des métapériphériques qui se trouvent sur le châssis de boîtier du disque affecté affichent un état d'erreur. Etant donné que la récupération d'une réplique de la base de données d'état des métapériphériques n'est pas automatique, il est préférable d'effectuer la récupération immédiatement après la remise en service du boîtier de disque. A défaut, une nouvelle panne risque de mettre hors service la plupart des répliques et de provoquer une panne du noyau. Ce comportement de Solstice DiskSuite se produit en général lorsque le nombre de répliques disponibles n'est pas suffisant.

Bien que ces répliques erronées soient restaurées lors de la relève suivante (commande `haswitch(1M)` ou `reboot(1M)`), vous pourriez vouloir les remettre en service manuellement en les supprimant et en les ajoutant de nouveau.

Remarque - Veillez à rajouter le même nombre de répliques que celui qui a été supprimé sur chaque tranche. Vous pouvez supprimer plusieurs répliques simultanément à l'aide de la commande `metadb(1M)`. Si une tranche doit contenir plusieurs exemplaires des répliques, ajoutez ces dernières en exécutant une seule fois la commande `metadb(1M)`, accompagnée de l'indicateur `-c`.

3. **Exécutez la commande `metastat(1M)` pour identifier les métapériphériques contenant des erreurs.**

```
# metastat -s ensemble_disques
```

4. Remettez en service les métapériphériques contenant des erreurs au moyen de la commande `metareplace(1M)`, puis resynchronisez les disques.

```
# metareplace -s ensemble_disques -e miroir composant
```

L'option `-e` permet au composant (tranche) de passer à l'état de disponibilité et lance la resynchronisation.

Les composants ayant été remplacés par une réserve dynamique doivent être les derniers périphériques remplacés avec la commande `metareplace(1M)`. Si la réserve dynamique est remplacée en premier, elle pourrait remplacer un autre sous-miroir défectueux dès qu'elle sera disponible.

Vous ne pouvez resynchroniser qu'un seul composant de sous-miroir (métapériphérique) à la fois. Si tous les composants d'un sous-miroir ont été touchés par la panne de courant, chaque composant doit être remplacé séparément. La resynchronisation d'un disque de 1,05 Go demande environ 10 minutes.

Si les deux ensembles de disques d'une configuration symétrique ont été touchés par la panne de courant, vous pouvez resynchroniser en même temps les sous-miroirs touchés de chaque ensemble de disques. Connectez-vous séparément à chaque hôte afin de récupérer l'ensemble de disques de chacun d'eux en lançant la commande `metareplace(1M)` pour chacun.

Remarque - Selon le nombre de sous-miroirs et de composants figurant dans ces sous-miroirs, la resynchronisation peut demander beaucoup de temps. Ainsi, pour un seul sous-miroir contenant 30 disques de 1,05 Go, elle peut durer environ cinq heures. Avec une configuration moins complexe composée de cinq sous-miroirs, la resynchronisation est achevée au bout de 50 minutes environ.

▼ Comment effectuer une reprise après une panne de courant (VxVM)

Les pannes de courant peuvent provoquer la déconnexion des unités de disques et donc des périphériques virtuels, qui ne sont alors plus disponibles. Le volume demeure toutefois actif, car les périphériques virtuels qui restent dans un volume mis en miroir sont toujours disponibles. Il est possible de reconnecter les unités de disque et de remédier à la situation sans arrêter les noeuds de la grappe.

La liste qui suit énumère les étapes générales de la récupération suite à une panne de courant d'un boîtier de disque dans une configuration VxVM :

- Identification des périphériques virtuels défectueux à l'aide des commandes `vxprint` et `vxdisk`.
- Correction du problème qui a causé la panne de courant.
- Utilisation des commandes `drvconfig` et `disks` pour créer les entrées `/devices` et `/dev`.
- Analyse de la configuration de disque actuelle.
- Reconnexion des disques qui contenaient des erreurs temporaires.
- Vérification de la présence d'autres erreurs.
- (Facultatif) Dans le cas de groupes de disques partagés, exécution de la commande `vxdg` pour chaque disque dont l'alimentation a été coupée.
- Lancement de la récupération de volume.

Les étapes expliquées maintenant décrivent en détails la récupération suite à une panne de courant d'un boîtier de disque dans une configuration VxVM.

1. Exécutez la commande `vxprint` pour afficher les périphériques virtuels défectueux.

Au besoin, spécifiez un groupe de disques avec l'option `-g groupe_disques`.

2. Exécutez la commande `vxdisk` pour identifier les disques contenant des erreurs.

```
# vxdisk list
DEVICE      TYPE      DISK      GROUP      STATUS
..
-           -         c1t5d0    toi        failed was:c1t5d0s2
...
```

3. Corrigez l'erreur qui a causé le problème afin de rétablir l'alimentation électrique de tous les disques en panne.

Vérifiez que les disques tournent avant de poursuivre.

4. Entrez les commandes suivantes sur tous les noeuds de la grappe.

Dans certains cas, les disques doivent être de nouveau reconnus par les noeuds.

```
# drvconfig
# disks
```

5. Entrez les commandes suivantes sur tous les noeuds de la grappe.

Le gestionnaire de volumes doit analyser de nouveau la configuration de disque actuelle.

```
# vxdctl enable
# vxdisk -a online
```

6. Entrez la commande suivante tout d'abord sur le noeud maître, et ensuite seulement sur tous les autres noeuds de la grappe.

Les disques contenant des erreurs temporaires sont alors rattachés.

```
# vxreattach
```

7. Vérifiez la sortie de la commande `vxdisk` pour repérer toute autre erreur éventuelle.

```
# vxdisk list
```

8. Si le support a été remplacé, entrez la commande suivante à partir du noeud maître pour chacun des disques déconnectés.

Le disque physique et le nom d'accès du gestionnaire de volumes de ce disque doivent être reconnectés.

```
# vxdg -g groupe_disques -k adddisk nom_support=nom_accès
```

Les valeurs pour `nom_support` et `nom_accès` figurent à la fin de la sortie de la commande `vxdisk list`.

Exemple :

```
# vxdg -g toi -k adddisk c1t5d0=c1t5d0s2
# vxdg -g toi -k adddisk c1t5d1=c1t5d1s2
# vxdg -g toi -k adddisk c1t5d2=c1t5d2s2
# vxdg -g toi -k adddisk c1t5d3=c1t5d3s2
# vxdg -g toi -k adddisk c1t5d4=c1t5d4s2
```

Vous pouvez également rattacher les disques à l'aide de la commande `vxdiskadm` ou de l'interface graphique.

9. Depuis le noeud, démarrez la récupération du volume.

Si vous utilisez des groupes de disques partagés, spécifiez les options `-svc` dans la commande `vxrecover`.

```
# vxrecover -bv [-g groupe_disques]
```

10. (Facultatif) Exécutez la commande `vxprint -g` pour afficher les modifications.

Administration des disques Sun StorEdge A5000

Ce chapitre décrit les procédures d'administration des composants Sun StorEdge A5000. Utilisez les procédures qui figurent dans le manuel du matériel du serveur pour identifier le composant défectueux.

Réparation d'une connexion Sun StorEdge A5000

L'interruption de la connexion entre un boîtier de disque et l'un des noeuds de grappe est habituellement provoquée par une défaillance d'un câble SCSI-2 ou d'une carte SBus.

Dans tous les cas, le noeud sur lequel la panne s'est produite se met à générer des erreurs lorsque la panne est découverte. Les accès subséquents au boîtier de disque produisent d'autres erreurs. Le noeud présente le même comportement que si l'alimentation du boîtier de disque avait été coupée. Les opérations d'E/S effectuées depuis les autres noeuds de la grappe ne sont pas touchées par ce type de panne.

Pour diagnostiquer la panne, vous pouvez avoir recours aux procédures d'essai du module de carte qui figurent dans le manuel d'entretien du noeud Sun Cluster, afin de déterminer quel composant est défectueux. Vous devez normalement libérer un noeud et le boîtier de disque qui semble être en panne afin de dépanner le matériel.

▼ Comment réparer une connexion Sun StorEdge A5000

1. Préparez le système Sun Cluster comme pour un remplacement de composants.

Selon la cause de la perte de connexion, préparez le noeud Sun Cluster en effectuant l'une des procédures suivantes.

- Si la défaillance s'est produite sur l'adaptateur d'hôte FC-100 SBus, consultez le Chapitre 7 pour savoir comment préparer le noeud Sun Cluster pour la mise hors tension.
- Si c'est le câble à fibres optiques FC-100 qui est défectueux, le logiciel de gestion de volumes aura déjà détecté le problème et préparé le système en vue du remplacement du câble.

2. Remplacez le composant défectueux.

Si la défaillance s'est produite sur le câble à fibres optiques FC-100 ou sur l'adaptateur d'hôte FC-100 SBus, consultez le *Manuel d'installation et d'entretien Sun StorEdge A5000* pour des instructions détaillées sur leur remplacement.

3. Corrigez les erreurs du logiciel de gestion de volumes.

Effectuez les procédures décrites à la section "Reprise après une panne de courant", page 267.

Ainsi prend fin la procédure de rétablissement d'une connexion.

Administration des disques Sun StorEdge A5000

Cette section explique comment ajouter et remplacer des disques Sun StorEdge A5000 dans une configuration Sun Cluster.

Ajout ou remplacement de disques Sun StorEdge A5000

Lorsque vous ajoutez ou remplacez des disques Sun StorEdge A5000, reportez-vous toujours à la documentation accompagnant votre système.

Remarque - Pour remplacer un disque A5000 défectueux sous contrôle VxVM, vous ne pouvez pas vous contenter de retirer le disque et de le remplacer par un autre. En effet, chaque disque porte un nom universel unique. Pour de plus amples renseignements sur les noms universels, voir “Remplacement d’un contrôleur du tableau SPARCstorage Array et modification du nom universel”, page 193.

▼ Comment ajouter un disque Sun StorEdge A5000 (Solstice DiskSuite)

1. Exécutez la commande `luxadm` pour insérer le nouveau disque.

Procédez à l’installation physique du ou des nouveau(x) disque(s) lorsque le système vous invite à le faire. Répétez cette procédure pour chaque noeud physiquement connecté au tableau.

```
# luxadm insert boitier.emplacement
```

2. Insérez le nouveau disque et appuyez sur Entrée.

3. Au besoin, utilisez la commande `Format` pour créer une étiquette et repartitionner le disque.

4. Exécutez la commande `scdidadm(1M)` pour reconnaître le nouveau disque et créer une instance DID correspondante.

N’exécutez cette commande qu’à partir du noeud 1. Pour plus de détails, consultez la page de manuel `scdidadm(1M)`.

```
# scdidadm -r -H node2,node3...
```

Ainsi prend fin la procédure d’ajout de disques.

▼ Comment ajouter un disque Sun StorEdge A5000 (VxVM)

1. Utilisez la commande `luxadm` pour préparer la boucle en vue de l’ajout d’un nouveau périphérique.

Procédez à l'installation physique du ou des nouveau(x) disque(s) quand le système vous invite à le faire.

```
# luxadm insert
```

2. Signalez à VxVM la présence du nouveau disque.

```
# vxdctl enable
```

3. Au moyen de la commande `vxdiskadm`, placez le(s) nouveau(x) disque(s) sous le contrôle de VxVM.

Tapez 1 (Ajout ou initialisation d'un ou plusieurs disques).

Ainsi prend fin la procédure d'ajout de disques.

▼ Comment remplacer un disque Sun StorEdge A5000 (Solstice DiskSuite)

1. Identifiez tous les métapériphériques ou toutes les applications utilisant le disque défectueux.

Si les métapériphériques ont été mis en miroir ou s'il s'agit de périphériques RAID5, le disque peut être remplacé sans qu'il soit nécessaire d'arrêter les métapériphériques. Sinon, il faut interrompre les E/S sur le disque au moyen des commandes appropriées. Pour démonter un système de fichiers sur une bande ou une concaténation, par exemple, vous devez utiliser la commande `umount (1M)`.

2. Conservez au besoin l'étiquette de disque.

Exemple :

```
# prtvtoc /dev/rdisk/c1t3d0s2 > /tmp/c1t3d0.vtoc
```

3. (Facultatif) Exécutez la commande `metareplace` pour remplacer des tranches de disque si ce dernier n'a pas fait l'objet d'un remplacement dynamique.

Exemple :

```
# metareplace d1 c1t3d0s2 c1t2d0s2
d1: device c1t3d0s2 is replaced with c1t2d0s2
```

4. Exécutez la commande `luxadm -F` pour supprimer le disque.

L'option `-F` doit être spécifiée, car Solstice DiskSuite ne met pas les disques hors ligne. Répétez cette commande pour tous les hôtes, s'il y en a plusieurs. Exemple :

```
# luxadm remove -F /dev/rdisk/c1t3d0s2
WARNING!!! Please ensure that no filesystems are mounted on these device(s).
All data on these devices should have been backed up.The list of devices which
will be removed is: 1: Box Name ``macsl'' rear slot 1 Please enter 'q' to Quit
or <Return> to Continue: stopping: Drive in ``macsl'' rear slot 1...Done
offlining: Drive in ``macsl'' rear slot 1...Done Hit <Return> after removing
the device(s).
```

Remarque - L'icône FPM de l'unité de disque à supprimer doit normalement clignoter, de même que la DEL ambre sous cette unité.

5. Retirez le disque et appuyez sur Entrée.

La sortie résultante doit normalement se présenter comme suit :

```
Hit <Return> after removing the device(s).
Drive in Box Name ``macsl'' rear slot 1
Removing Logical Nodes:
Removing c1t3d0s0 Removing c1t3d0s1 Removing c1t3d0s2 Removing c1t3d0s3
Removing c1t3d0s4 Removing c1t3d0s5 Removing c1t3d0s6 Removing c1t3d0s7
Removing c2t3d0s0 Removing c2t3d0s1 Removing c2t3d0s2 Removing c2t3d0s3
Removing c2t3d0s4 Removing c2t3d0s5 Removing c2t3d0s6 Removing c2t3d0s7
#
```

6. Répétez l'Etape 4, page 276 pour tous les noeuds si le tableau de disques est configuré en mode multihôtes.

7. Exécutez la commande `luxadm insert` pour insérer le nouveau disque.

Répétez cette procédure pour tous les noeuds. La sortie résultante doit normalement se présenter comme suit :

```
# luxadm insert macs1,r1
The list of devices which will be inserted is:
1: Box Name ``macs1`` rear slot 1
Please enter `q` to Quit or <Return> to Continue: Hit <Return> after inserting the device(s)
```

8. Insérez le disque et appuyez sur Entrée.

La sortie résultante doit normalement se présenter comme suit :

```
Hit <Return> after inserting the device(s). Drive in Box Name ``macs1`` rear slot 1
Logical Nodes under /dev/dsk and /dev/
rdsk : clt3d0s0 clt3d0s1 clt3d0s2 clt3d0s3 clt3d0s4 clt3d0s5
      clt3d0s6 clt3d0s7 c2t3d0s0 c2t3d0s1 c2t3d0s2 c2t3d0s3 c2t3d0s4 c2t3d0s5 c2t3d0s6 c2t3d0s7
#
```

Remarque - L'icône FPM de l'unité de disque remplacée doit normalement clignoter, de même que la DEL verte sous cette unité.

9. Exécutez la commande `scdidadm(1M)` pour mettre à jour les informations de pseudo-périphérique DID.

Exécutez la commande suivante sur tous les noeuds connectés au disque pour mettre à jour les informations d'ID du nouveau disque.

```
# scdidadm -R instance_DID
```

où *instance_DID* est le numéro d'instance du disque remplacé. Pour de plus amples informations à ce sujet, voir la page de manuel `scdidadm(1M)`.

10. Réinitialisez tous les noeuds connectés au nouveau disque.

Pour éviter toute interruption, exécutez la commande `haswitch(1M)` pour commuter la propriété des hôtes logiques pouvant être sous la maîtrise du noeud à réinitialiser. Exemple :

```
# haswitch phys-hahost2 hahost1 hahost2
```

11. Etiquetez le disque si nécessaire.

Exemple :

```
# cat /tmp/clt3d0.vtoc | fmthard -s - /dev/rdisk/clt3d0s2
fmthard: New volume table of contents now in place.
```

12. Au besoin, remplacez metadb.

Exemple :

```
# metadb -d clt3d0s0; metadb -a clt3d0s0
```

13. Activez les tranches du nouveau disque au moyen de la commande metareplace -e.

Exemple :

```
# metareplace -e d0 clt3d0s0
d0: device clt3d0s0 is enabled
```

Ainsi prend fin la procédure de remplacement de disque.

▼ Comment remplacer un disque Sun StorEdge A5000 (VxVM)

1. Identifiez tous les volumes ou applications utilisant le disque défectueux.

Si les volumes ont été mis en miroir ou s'il s'agit de périphériques RAID5, le disque peut être remplacé sans qu'il soit nécessaire d'arrêter le volume. Sinon, il faut interrompre les E/S sur le disque au moyen des commandes appropriées. Pour démonter un système de fichiers sur une bande ou une concaténation, par exemple, vous devez utiliser la commande `umount (1M)`.

2. Exécutez la commande `vxdiskadm` pour remplacer et mettre hors ligne une unité de disque.

Sous VxVM, exécutez ces commandes sur l'ordinateur contrôlant l'hôte logique propriétaire du groupe de disques.

Entrez **4** (Suppression du disque pour remplacement) et ensuite **11** (Désactivation [mise hors ligne] du disque).

Vous pouvez également utiliser l'interface graphique, si vous préférez.

3. Exécutez la commande `luxadm` pour supprimer le périphérique et les noeuds de périphérique.

Il s'agit d'une commande interactive qui vous invite à retirer le disque physique. Exécutez cette commande sur chacun des noeuds connectés au tableau. Exemple :

```
# luxadm remove_device -F /dev/rdisk/c2t20d0s2
```

4. Remplacez le disque physique et exécutez ensuite la commande `luxadm` pour insérer le nouveau disque.

Le nouveau périphérique et les noeuds de périphérique sont alors créés. Exécutez cette commande sur chacun des noeuds connectés au tableau. Exemple :

```
# luxadm insert_device ratbert,r4
```

5. Signalez la présence du nouveau disque au gestionnaire de volumes.

```
# vxdctl enable
```

6. Exécutez la commande `vxdiskadm` pour placer le nouveau disque sous le contrôle de VxVM.

Tapez 5 (Remplacement d'un disque défectueux ou supprimé).

7. (Facultatif) Le volume peut à présent être restauré, si nécessaire.

Ainsi prend fin la procédure de remplacement de disque.

Administration des gestionnaires de volumes

Cette annexe fournit des instructions relatives à l'administration des ensembles de disques et des métapériphériques Solstice DiskSuite, et à l'administration des objets VERITAS Volume Manager. Les méthodes décrites sont tributaires du logiciel de gestion des volumes utilisé.

- "Utilisation de Solstice DiskSuite dans l'environnement Sun Cluster", page 281
- "Utilisation de VxVM dans l'environnement Sun Cluster", page 289
- "Sauvegarde de données multihôtes à l'aide de Solstice Backup", page 299

Utilisation de Solstice DiskSuite dans l'environnement Sun Cluster

Cette section indique comment utiliser DiskSuite pour administrer :

- les ensembles de disques.
- les disques d'un ensemble.
- les métapériphériques multi-hôtes.
- les métapériphériques locaux.

Pour une description détaillée de l'administration des objets DiskSuite, reportez-vous à la documentation de Solstice DiskSuite.

Administration de métapériphériques et d'ensembles de disques

Les métapériphériques et les ensembles de disques se créent et s'administrent à l'aide des utilitaires de ligne de commande Solstice DiskSuite ou de l'interface utilisateur graphique DiskSuite Tool (`metatool(1M)`).

Lisez ce chapitre avant d'avoir recours à la documentation Solstice DiskSuite pour administrer les ensembles de disques et les métapériphériques d'une configuration Sun Cluster.

Les ensembles de disques sont des groupes de disques. La principale tâche d'administration à effectuer en rapport avec les ensembles de disques a trait à l'ajout et au retrait de disques.

Pour pouvoir utiliser un disque qui a été placé dans un ensemble de disques, vous devez configurer un métapériphérique à l'aide des tranches du disque. Un métapériphérique peut être une concaténation, un entrelacement, une image miroir ou un périphérique de consignation UFS (également appelé périphérique trans). Vous pouvez en outre créer des pools de réserve dynamiques qui renferment des tranches qui serviront de remplacement en cas de défaillance d'un métapériphérique.

Remarque - Les noms des métapériphériques débutent par la lettre `d` qui est suivie d'un nombre. Par défaut, dans une configuration Sun Cluster, on retrouve 128 métapériphériques compris entre 0 et 127. Chaque périphérique de consignation UFS que vous créez utilise au moins sept noms de métapériphériques. Par conséquent, dans le cas d'une configuration Sun Cluster d'envergure, les 128 noms de métapériphériques par défaut peuvent ne pas suffire. Pour des directives sur la façon de changer la quantité par défaut, reportez-vous à la documentation de Solstice DiskSuite. Les noms des pools de réserve dynamiques débutent par les lettres `hsp`, qui sont suivies d'un nombre. Vous pouvez avoir jusqu'à 1000 pools de réserve dynamiques, allant de `hsp000` à `hsp999`.

A propos des ensembles de disques

Cette section présente un aperçu général des ensembles de disques et de leurs relations avec les hôtes logiques, ainsi que des procédures d'ajout et de retrait de disques d'un ensemble associé à l'hôte logique.

Les hôtes logiques Sun Cluster ont pour maîtres des hôtes physiques. Seul l'hôte physique qui est actuellement le maître d'un hôte logique peut accéder à l'ensemble de disques de cet hôte. Lorsqu'un hôte physique est le maître de l'ensemble de disques d'un hôte logique, on dit qu'il est propriétaire de l'ensemble de disques. En général, Sun Cluster se charge de la propriété des ensembles de disques. Cependant, si l'hôte logique est en cours de maintenance, comme l'indique la commande `hastat(1M)`, vous pouvez avoir recours à la commande `metaset -t` de DiskSuite pour assumer manuellement la propriété de l'ensemble de disques. Avant de

remettre l'hôte logique en service, libérez la propriété de l'ensemble de disques à l'aide de la commande `metaset -r`.

Remarque - N'effectuez jamais l'administration de l'ensemble de disques à l'aide des options `-t` (prise en charge de la propriété) ou `-r` (libération de la propriété) de la commande `metaset(1M)` pendant que les hôtes logiques sont en marche. En effet, ces options sont utilisées par le logiciel Sun Cluster et leur emploi doit être coordonné entre les noeuds de grappe.

Ajout d'un disque à un ensemble

Si le disque à ajouter à un ensemble doit faire office de sous-miroir, vous devez disposer de deux disques sur deux unités d'expansion de disques multihôtes différentes afin de permettre l'écriture miroir. Toutefois, si le disque doit être utilisé comme réserve dynamique, vous pouvez ajouter un seul disque.

▼ Comment ajouter un disque à un ensemble (Solstice DiskSuite)

1. Vérifiez que le disque ne contient aucune donnée.

Cet aspect est important, car la table de partition sera réécrite et un espace sera alloué, sur le disque, pour la réplique de la base de données d'état des métapériphériques.

2. Insérez le disque dans l'unité d'expansion de disque multihôtes.

Suivez les instructions qui figurent dans la documentation du matériel de l'unité d'expansion de disque (ajout et retrait de disques).

3. Ajoutez le disque à un ensemble de disques.

La syntaxe de la commande est précisée ci-après. Dans cet exemple, *ensemble_disques* est le nom de l'ensemble auquel le disque doit être ajouté, et *unité* le nom DID du disque sous la forme *dN* (pour les nouvelles installations de Sun Cluster), ou *cNt YdZ* (pour les installations mises à niveau à partir de HA 1.3).

```
# metaset -s ensemble_disques -a unité
```

4. Après avoir ajouté les disques à l'ensemble de disques en utilisant la commande `metaset(1M)`, utilisez la commande `scadmin(1M)` afin de réserver et d'activer la fonction failfast sur les disques indiqués.

```
phys-hahost1# scadmin reserve nom_unité
```

Retrait d'un disque d'un ensemble de disques

Vous pouvez retirer un disque d'un ensemble de disques à n'importe quel moment, à condition qu'aucune des tranches de ce disque ne soient en cours d'utilisation dans les métapériphériques ou les pools de réserve dynamiques.

▼ Comment retirer un disque d'un ensemble de disques (Solstice DiskSuite)

1. Utilisez la commande `metastat(1M)` pour vous assurer qu'aucune des tranches n'est utilisée comme métapériphérique ou réserve dynamique.
2. Utilisez la commande `metaset(1M)` afin de retirer le disque voulu de l'ensemble.

La syntaxe de la commande est précisée ci-après. Dans cet exemple, *ensemble_disques* est le nom de l'ensemble de disques qui contient le disque (défectueux) qui doit être retiré et *unité* le nom du disque sous la forme *dN* (pour les nouvelles installations de Sun Cluster), ou *cNtYdZ* (pour les installations mises à jour à partir de HA 1.3).

```
# metaset -s ensemble_disques -d unité
```

Cette opération peut demander 15 minutes et même davantage, selon la taille de la configuration et le nombre de disques.

Administration de métapériphériques multihôtes

Les sections suivantes présentent les différences entre l'administration de métapériphériques en environnement Sun Clustermultihôte et en environnement à un seul hôte.

Sauf indication contraire dans les sections qui suivent, vous pouvez suivre les instructions qui figurent dans la documentation de Solstice DiskSuite.

Remarque - Les instructions qui figurent dans la documentation de Solstice DiskSuite ne concernent que les configurations à un seul hôte.

Les sections ci-après présentent les programmes de ligne de commande de Solstice DiskSuite à utiliser pour exécuter une tâche. Vous pouvez également vous servir de l'interface utilisateur graphique `metatool(1M)` pour toutes les tâches, à moins d'indications contraires. Utilisez l'option `-s` lorsque vous exécutez `metatool(1M)`, car vous pouvez ainsi indiquer le nom de l'ensemble de disques.

Gestion des métapériphériques

Pour la gestion courante des métapériphériques, vous devez surveiller en permanence leurs éventuelles anomalies de fonctionnement, comme décrit dans "Utilitaires de surveillance", page 29

Lorsque `hastat(1M)` signale qu'un ensemble de disques est défectueux, employez la commande `metastat(1M)` afin de déterminer quel métapériphérique est défectueux.

Vous devez utiliser l'option `-s` lorsque vous exécutez `metastat(1M)` ou `metatool(1M)`, afin de pouvoir indiquer le nom de l'ensemble de disques.

Remarque - Vous devez enregistrer les informations de configuration du métapériphérique lorsque vous modifiez la configuration. Utilisez la commande `metastat -p` pour produire un résultat similaire à celui du fichier `md.tab` puis enregistrez le résultat. Pour plus de détails sur l'enregistrement de données de partitionnement, reportez-vous à la section "Enregistrement des informations de partitionnement d'un disque (Solstice DiskSuite)", page 21.

Ajout d'un miroir à un ensemble de disques

Les métapériphériques mis en miroir peuvent être utilisés dans le cadre d'un système de fichiers UFS de consignment pour les applications Sun Cluster à haute disponibilité.

On peut configurer en métapériphérique les tranches inactives de disques dans un ensemble de disques à l'aide de la commande `metainit(1M)`.

Retrait d'un miroir d'un ensemble de disques

Sun Cluster peuvent utiliser des métapériphériques mis en miroir bruts pour le stockage des bases de données. Celles-ci ne sont pas mentionnées dans le fichier `dfstab.hôte logique` ou `vfstab` de chaque hôte logique, mais figurent dans les fichiers de configuration de base de données de Sun Cluster. Le miroir doit être retiré de ces fichiers, et le système de base de données Sun Cluster doit cesser d'utiliser ce miroir. Puis, le miroir peut être supprimé à l'aide de la commande `metaclear(1M)`.

Mise hors ligne de sous-miroirs

Pour les tableaux SPARCstorage, notez qu'avant de remplacer ou d'ajouter une unité de disques sur un plateau, tous les périphériques de ce plateau doivent être mis hors ligne.

Dans des configurations symétriques, la mise hors ligne de sous-miroirs en vue de la maintenance est une tâche complexe, car les disques de chacun des deux ensembles de disques doivent se trouver dans le même plateau, dans le tableau SPARCstorage. Vous devez mettre hors ligne les métapériphériques de chaque ensemble de disques avant de retirer le plateau.

Employez la commande `metaoffline(1M)` afin de mettre hors ligne tous les sous-miroirs de chaque disque du plateau.

Création de nouveaux métapériphériques

Une fois qu'un disque a été ajouté à un ensemble de disques, vous pouvez créer de nouveaux métapériphériques en utilisant `metainit(1M)` ou `metatool(1M)`. Si les nouveaux périphériques sont des éléments de réserve dynamiques, employez la commande `metahs(1M)` afin de placer ces éléments dans un pool de réserve dynamique.

Remplacement de composants défectueux

Pour remplacer le composant défectueux d'un métapériphérique, utilisez la commande `metareplace(1M)`.

Vous devez disposer d'une tranche (ou d'un disque) de remplacement. Il peut s'agir d'un périphérique existant qui n'est pas utilisé ou d'un nouveau périphérique que vous avez ajouté à l'ensemble de disques.

Vous pouvez aussi remettre en service les unités ayant subi des erreurs passagères (par exemple suite à une panne d'alimentation du châssis) à l'aide de la commande `metareplace -e` commande.

Suppression de métapériphériques

Avant de supprimer un métapériphérique, assurez-vous qu'aucun de ses composants n'est utilisé par Sun Cluster HA for NFS. Utilisez ensuite la commande `metaclear(1M)`, pour supprimer le métapériphérique.

Développement de métapériphériques

Pour développer un métapériphérique, vous devez disposer d'au moins deux tranches (de disque) situées dans des unités d'expansion de disques multihôtes

différentes. Chacune des deux nouvelles tranches doit être ajoutée à un sous-miroir différent à l'aide de la commande `metainit(1M)`. Employez ensuite la commande `growfs(1M)` pour développer le système de fichiers.



Attention - Lorsque la commande `growfs(1M)` est en cours d'exécution, les clients peuvent subir des interruptions de service.

Si une relève survient pendant que le système de fichiers se développe, le développement n'aura pas lieu. Vous devez relancer la commande `growfs(1M)` une fois la relève terminée.

Remarque - Le système de fichiers qui contient `/hôtelogique/statmon` ne peut être développé. Étant donné que le programme `statd(1M)` modifie ce répertoire, celui-ci serait paralysé pendant une longue période, le temps que le développement du système de fichiers ait lieu. Les répercussions sur le protocole de verrouillage des fichiers du réseau seraient alors imprévisibles. Ce problème ne concerne que les configurations utilisant Sun Cluster HA for NFS.

Gestion des pools de réserve dynamiques

Vous pouvez ajouter ou supprimer à tout moment des périphériques de réserve dynamique dans des pools de réserve dynamique, à condition qu'ils ne soient pas en cours d'utilisation. De plus, vous pouvez créer de nouveaux pools de réserve dynamiques, puis les associer à des sous-miroirs à l'aide de la commande `metahs(1M)`.

Gestion des journaux UFS

Tous les journaux UFS qui se trouvent sur les disques multihôtes sont mis en miroir. Lorsqu'un sous-miroir tombe en panne, le système indique que ce composant est défectueux. Vous pouvez réparer la panne à l'aide de la commande `metareplace(1M)` ou `metatool(1M)`.

Si le miroir entier qui contient le journal UFS tombe en panne, vous devez démonter le système de fichiers, sauvegarder toutes les données accessibles, corriger l'erreur, réparer le système de fichiers (à l'aide de la commande `fsck(1M)`), puis remonter le système de fichiers.

Ajout de consignation UFS à un hôte logique

Tous les systèmes de fichiers UFS figurant dans un hôte logique doivent être des systèmes UFS de consignation de sorte que la reprise ou les critères de dépassement

impartis de `haswitch(1M)` soient respectés. Ainsi, on favorise une commutation et une relève rapides.

Pour configurer le système de fichiers UFS de consignation, créez un périphérique trans à l'aide d'un périphérique de consignation mis en miroir et d'un système de fichiers maître UFS mis en miroir. Le périphérique de consignation et le périphérique maître UFS doivent être mis en miroir.

En général, la tranche 6 de chaque unité d'un ensemble de disques peut être utilisée comme journal UFS. Les tranches peuvent être employées pour les sous-miroirs des journaux UFS. Si la taille des tranches est inférieure à celle du journal voulu, vous pouvez regrouper plusieurs tranches. Habituellement, un méga-octet par groupe de 100 méga-octets convient aux journaux UFS, jusqu'à un maximum de 64 méga-octets. Idéalement, les tranches d'un journal doivent se trouver sur une unité différente de celle du périphérique maître UFS.

Remarque - Si vous devez découper le disque afin de libérer de l'espace pour les journaux UFS, conservez la tranche 7 existante, qui débute au cylindre 0 et qui contient au moins deux méga-octets. En effet, cet espace est réservé pour les répliques de base de données d'état des métapériphériques. Les champs *Etiquette* et *Indicateur* (dont le contenu est signalé par la commande `format(1M)`) doivent être réservés pour la tranche 7. La commande `metaset(1M)` définit correctement les champs *Etiquette* et *Indicateur* au moment de la configuration initiale.

Une fois que le périphérique trans a été configuré, créez le système de fichiers UFS, à l'aide de la commande `newfs(1M)`, sur le périphérique trans.

Une fois le processus `newfs` terminé, ajoutez le système de fichiers UFS au fichier `vfstab` pour l'hôte logique, en modifiant le fichier `/etc/opt/SUNWcluster/conf/hanfs/vfstab.hôte logique` pour mettre à jour les informations des systèmes de fichiers administratifs et multihôtes.

Assurez-vous que les fichiers `vfstab.hôte logique` de tous les noeuds de grappe contiennent les mêmes informations. Utilisez la fonction `cconsole(1)` pour modifier simultanément les fichiers `vfstab.hôte logique` de tous les noeuds de la grappe.

Voici un exemple de fichier `vfstab.hôte logique` sur lequel apparaissent le système de fichiers administratif et quatre autres systèmes de fichiers UFS :

#device	device	mount	FS	fsck	mount	mount	
#to mount	to fsck	point		type	pass	all	options#
/dev/md/hahost1/dsk/d11	/dev/md/hahost1/rdsk/d11	/hahost1	ufs	1	no	-	
/dev/md/hahost1/dsk/d1	/dev/md/hahost1/rdsk/d1	/hahost1/1	ufs	1	no	-	
/dev/md/hahost1/dsk/d2	/dev/md/hahost1/rdsk/d2	/hahost1/2	ufs	1	no	-	
/dev/md/hahost1/dsk/d3	/dev/md/hahost1/rdsk/d3	/hahost1/3	ufs	1	no	-	
/dev/md/hahost1/dsk/d4	/dev/md/hahost1/rdsk/d4	/hahost1/4	ufs	1	no	-	

Si le système de fichiers est destiné à être également utilisé par Sun Cluster HA for NFS, suivez la procédure de partage des systèmes de fichiers NFS tels que décrite dans le Chapitre 11 du *Sun Cluster 2.2 Software Installation Guide*.

Le nouveau système de fichier sera monté automatiquement à la prochaine reconfiguration du moniteur d'appartenance. Pour forcer la reconfiguration de l'appartenance, utilisez la commande suivante :

```
#haswitch -r
```

Administration de métapériphériques locaux

Les disques locaux peuvent être mis en miroir. Si un seul miroir tombe en panne, suivez les instructions de la documentation de Solstice DiskSuite pour remplacer le miroir défectueux et resynchroniser le disque de remplacement avec le bon disque.

Actions de métapériphérique destructives

Les actions de métapériphériques non prises en charge par les configurations Sun Clustersont :

- Création d'un miroir unidirectionnel dans un ensemble de disques.
- Création d'une configuration comportant trop peu de répliques de base de données d'état des métapériphériques sur les disques locaux.
- Modification de répliques de base de données d'état des métapériphériques sur des disques multihôtes, à moins que des instructions explicites à cet effet figurent dans le présent manuel ou une autre publication Sun Cluster.

Utilisation de VxVM dans l'environnement Sun Cluster

VERITAS Volume Manager (VxVM) et le dispositif VxVM du cluster sont des variations du même gestionnaire de volumes. Le dispositif VxVM du cluster est uniquement utilisé dans les configurations Oracle Parallel Server OPS. Cette section décrit l'utilisation de disques relevant de la responsabilité du gestionnaire de volumes en vue de l'administration :

- disques du gestionnaire de volumes.

- groupes de disques.
- sous-disques.
- périphériques virtuels.
- volumes.

Pour une description détaillée de l'administration de ces objets, reportez-vous à la section correspondante.

Aperçu de l'administration d'objets VxVM

Les objets gérés par le gestionnaire de volumes sont créés et administrés à l'aide des utilitaires de ligne de commande, ou de l'interface utilisateur graphique Visual Administrator.

Lisez bien ce chapitre avant de consulter la documentation VxVM concernant l'administration des objets gérés par un gestionnaire de volumes dans une configuration Sun Cluster. Les méthodes présentées ci-après ne représentent qu'un moyen d'effectuer ces tâches. Utilisez la méthode la mieux adaptée à votre configuration.

Ces objets ont généralement la relation suivante :

- Les disques sont mis sous la commande du gestionnaire de volumes et réunis en groupes de disques.
- Un ou plusieurs sous-disques (représentant chacun une certaine partie d'un disque) sont réunis afin de former des périphériques virtuels, ou miroirs.
- Un volume est composé d'un ou plusieurs périphériques virtuels.

Le groupe de disques par défaut est `rootdg` (le groupe de disques racine). Vous pouvez créer d'autres groupes de disques au besoin. Les principales tâches d'administration à effectuer en rapport avec les groupes de disques ont trait à l'ajout et au retrait de disques.

Avant d'utiliser un disque que vous avez placé dans un groupe de disques, vous devez configurer des disques et des sous-disques (sous la commande du gestionnaire de volumes) afin de constituer des périphériques virtuels, ou miroirs, à l'aide des tranches du disque physique. Un périphérique virtuel peut être une concaténation ou une tranche.

Avec VxVM, les applications accèdent aux volumes (créés sur les disques du gestionnaire de volumes) plutôt qu'aux tranches.

Les sections ci-après présentent les programmes de ligne de commande de VxVM à utiliser pour effectuer une tâche. Vous pouvez également vous servir de l'interface utilisateur graphique pour toutes les tâches, sauf indication contraire.

Remarque - Sur les noeuds qui font tourner des services de données Sun ClusterHA, n'exécutez jamais manuellement les options `vxdg import` ou `deport` sur un groupe de disques géré par Sun Cluster, à moins que l'hôte logique de ce groupe de disques ne soit en mode maintenance. Avant d'importer ou de déporter manuellement un groupe de disques, vous devez soit arrêter Sun Cluster sur tous les noeuds qui peuvent agir comme maîtres de ce groupe de disques (en lançant `scadmin stopnode` sur tous ces noeuds), soit utiliser la commande `haswitch -m` pour basculer en mode maintenance les hôtes logiques correspondants. Lorsque vous êtes prêts à redonner la commande du groupe de disques à Sun Cluster, la méthode la plus sûre consiste à déporter le groupe de disques avant de lancer `scadmin startnode` ou avant d'utiliser `haswitch(1M)` pour remettre l'hôte logique sous la commande de Sun Cluster.

Administration des disques

Avant d'être utilisé par VxVM, un disque doit être identifié, ou initialisé, comme disque géré par un gestionnaire de volumes. Un disque complètement initialisé peut être ajouté à un groupe de disques, ou encore servir à remplacer un disque défectueux ou à créer un nouveau groupe de disques.

▼ Comment initialiser et configurer un disque (VxVM)

1. Vérifiez que le disque ne contient aucune donnée.

Cet aspect est important, car les données figurant sur un disque sont détruites lorsque celui-ci est initialisé.

2. Insérez le périphérique de disque et installez-le dans le boîtier de disque en suivant les instructions figurant dans la documentation fournie avec le matériel.

3. Initialisez le disque et ajoutez-le à un groupe de disques.

Pour ce faire, vous pouvez utiliser les menus `vxdiskadm` ou l'interface utilisateur graphique. Vous pouvez également utiliser les utilitaires de ligne de commande `vxdisksetup` et `vxdg addisk` pour initialiser le disque et le placer dans un groupe de disques.

Mise hors ligne d'un disque

Il est parfois nécessaire de mettre un disque physique hors ligne. Si un disque est endommagé, vous devez le désactiver puis le retirer. Vous devez en outre désactiver un disque avant de déplacer le périphérique de disque physique à un autre endroit en vue de la connexion à un autre système.

Pour mettre un disque physique hors ligne, retirez d'abord celui-ci de son groupe de disques. Puis, mettez le disque hors ligne à l'aide de la commande `vxdisk(1M)`.

Retrait d'un disque

Vous pouvez retirer un disque dans le but de le placer dans un autre système, ou encore parce qu'il est en panne ou risque de l'être. Il est également possible de retirer les volumes qui ne sont plus utilisés.

Pour retirer un disque du groupe de disques, utilisez la commande `vxdg(1M)`. Pour faire en sorte qu'un disque ne soit plus géré par le gestionnaire de volumes en supprimant les partitions privées et publiques, utilisez la commande `vxdiskunsetup(1M)`. Pour plus de détails sur ces commandes, reportez-vous aux pages `vxdg(1M)` et `vxdiskunsetup(1M)` du manuel.

Administration des groupes de disques

Pour VxVM, il est préférable de créer et de garnir les groupes de disques à partir du noeud actif qui est le maître par défaut de ce groupe de disques. Dans une configuration N+1, chacun de ces noeuds maîtres par défaut partage la connexion au disque multihôtes avec un seul autre noeud de la grappe (le noeud de relève dynamique). En utilisant ces noeuds pour garnir les groupes de disques, vous ne courez pas le risque de générer des groupes incorrectement configurés.

Création d'un groupe de disques (VxVM)

Vous pouvez utiliser les menus `vxdiskadm` ou l'interface utilisateur graphique pour créer un nouveau groupe de disques. Vous pouvez aussi recourir à l'utilitaire de ligne de commande `vxdg init`.

Après les avoir créés et garnis, vous devez déporter chaque groupe de disques à l'aide de la commande `vxdg deport`. Ensuite, il convient d'importer chaque groupe sur le noeud de relève dynamique avec l'option `-t`. L'option `-t` est importante, dans la mesure où elle empêche l'importation de se poursuivre à la prochaine initialisation. Avant de poursuivre, il est nécessaire de créer tous les périphériques virtuels et les volumes de VxVM et d'activer les volumes.

Déplacer un disque vers un groupe de disques différent (VxVM)

Utilisez la procédure suivante pour déplacer un disque vers un groupe de disques différent.

▼ Comment déplacer un disque vers un groupe de disques différent (VxVM)

Pour déplacer un disque d'un groupe vers un autre, retirez-le du groupe où il se trouve pour le placer dans l'autre groupe.

Dans cet exemple, le disque physique `clt0d1` est déplacé du groupe de disques `acct` vers le groupe de disques `log_node1` à l'aide d'utilitaires de ligne commande.

1. Utilisez la commande `vxprint(1M)` pour déterminer si le disque est en cours d'utilisation.

```
# vxprint -g acct
TY NAME      ASSOC      KSTATE  LENGTH  PLOFFS  STATE    TUTILO  PUTILO
dg acct      acct       -        -        -        -        -        -
dm clt0d0    clt0d0s2  -        2050272 -        -        -        -
dm clt0d1    clt0d1s2  -        2050272 -        -        -        -
dm c2t0d0    c2t0d0s2  -        2050272 -        -        -        -
dm c2t0d1    c2t0d1s2  -        2050272 -        -        -        -
v newvol     gen       ENABLED  204800  -        ACTIVE   -        -
pl newvol-01 newvol    ENABLED  205632  -        ACTIVE   -        -
sd clt0d1-01 newvol-01 ENABLED  205632  0        -        -        -
pl newvol-02 newvol    ENABLED  205632  -        ACTIVE   -        -
sd c2t0d1-01 newvol-02 ENABLED  205632  0        -        -        -
v vol01     gen       ENABLED  1024000 -        ACTIVE   -        -
pl vol01-01 vol01     ENABLED  1024128 -        ACTIVE   -        -
sd clt0d0-01 vol01-01  ENABLED  1024128 0        -        -        -
pl vol01-02 vol01     ENABLED  1024128 -        ACTIVE   -        -
sd c2t0d0-01 vol01-02  ENABLED  1024128 0        -        -        -
```

2. Employez la commande `vxedit(1M)` afin de retirer le volume et de libérer le disque `clt0d1`.

Vous devez exécuter la commande `vxedit` à partir du noeud qui contrôle le groupe de disques partagé.

```
# vxedit -g acct -fr rm newvol
```

L'option `-f` force l'exécution de l'opération. L'option `-r` rend l'opération récursive.

3. Retirez le disque `c1t0d1` du groupe `acct`.

Vous devez lancer la commande `vxdg` depuis le noeud qui est le maître du groupe de disques partagé.

```
# vxdg -g acct rmdisk c1t0d1
```

4. Ajoutez le disque `c1t0d1` au groupe de disques `log_node1`.

```
# vxdg -g log_node1 adddisk c1t0d1
```



Attention - Cette méthode n'enregistre pas la configuration ni les données sur le disque.

Le groupe de disques `acct`, après le retrait du disque `c1t0d1`, est représenté ci-dessous.

```
# vxprint -g acct
TY NAME      ASSOC      KSTATE  LENGTH  PLOFFS  STATE  TUTILO  PUTILO
dg acct      acct      -       -       -       -       -       -
dm c1t0d0    c1t0d0s2  -       2050272 -       -       -       -
dm c2t0d0    c2t0d0s2  -       2050272 -       -       -       -
dm c2t0d1    c2t0d1s2  -       2050272 -       -       -       -
v vol01      gen        ENABLED 1024000 -       ACTIVE  -       -
pl vol01-01  vol01      ENABLED 1024128 -       ACTIVE  -       -
sd c1t0d0-01 vol01-01   ENABLED 1024128 0       -       -       -
pl vol01-02  vol01      ENABLED 1024128 -       ACTIVE  -       -
sd c2t0d0-01 vol01-02   ENABLED 1024128 0       -       -       -
```

Le groupe de disques `log_node1`, après le retrait du disque `c1t0d1`, est présenté ci-dessous.

```
# vxprint -g log_node1
TY NAME      ASSOC      KSTATE  LENGTH  PLOFFS  STATE  TUTILO  PUTILO
dg log_node1 log_node1  -       -       -       -       -       -
dm c1t0d1    c1t0d1s2  -       2050272 -       -       -       -
```

(suite)

(suite)

```
dm c1t3d0      c1t3d0s2  -      2050272  -      -      -      -
dm c2t3d0      c2t3d0s2  -      2050272  -      -      -      -
#
```

Pour modifier les autorisations ou la propriété des volumes, vous devez utiliser la commande `vxedit`.



Attention - N'utilisez pas les commandes `chmod` et `chgrp`. Les autorisations et la propriété définies par les commandes `chmod` et `chgrp` sont automatiquement remises à `root` (superutilisateur) durant la réinitialisation.

Voici un exemple des autorisations et de la propriété attribués aux volumes `vol01` et `vol02` dans le répertoire `/dev/vx/rdisk` avant une modification.

```
# ls -l
crw----- 1 root root nnn,nnnnn date time vol01
crw----- 1 root root nnn,nnnnn date time vol02
...
```

Voici un exemple de changement des autorisations et de la propriété du `vol01`.

```
vxedit -g nom_groupe set mode=755 user=oracle vol01
```

Une fois que la modification a été apportée, vous constaterez que les autorisations et la propriété ont changé.

```
# ls -l
crwxr-xr-x 1 oracle root nnn,nnnnn date time vol01
crw----- 1 root root nnn,nnnnn date time vol02
...
```

Administration des objets VxVM

Les volumes, ou disques virtuels, peuvent contenir des systèmes de fichiers ou des applications telles que des bases de données. Un volume peut comporter jusqu'à 32 périphériques virtuels, chacun d'entre eux renfermant un ou plusieurs sous-disques. Pour qu'un volume puisse être utilisé, au moins un périphérique virtuel comportant au moins un sous-disque correspondant doit y être associé. Précisons que tous les sous-disques d'un volume doivent appartenir au même groupe de disques.

Création de volumes et ajout de miroirs aux volumes

Utilisez l'interface graphique utilisateur ou l'utilitaire de ligne de commande `vxassist(1M)` pour créer des volumes dans chaque groupe de disques et pour créer un miroir associé à chaque volume.

La taille réelle d'un périphérique VxVM est légèrement inférieure à celle de l'unité de disque complète. VxVM réserve un petit espace à l'utilisation privée, appelé zone privée.

Remarque - On peut assigner le même nom à des volumes qui appartiennent à des groupes de disques différents.

Ajout de la consignation de zone modifiée

La consignation de zone modifiée (DRL) est une propriété facultative d'un volume qui sert à assurer une reprise rapide des volumes mis en miroir suite à une panne de système. La DRL assure le suivi des zones qui ont changé en raison d'opérations d'E/S sur un volume mis en miroir : elle se sert de ces informations pour ne récupérer que les parties requises du volume.

Création d'un fichier journal pour un volume existant

Les sous-disques de consignation servent à stocker le journal de zone modifiée d'un volume dont la DRL est activée. Un volume ayant une DRL dispose d'au moins un sous-disque de consignation. En outre, plusieurs sous-disques de consignation peuvent être utilisés afin de mettre en miroir le journal de zone modifiée. Chaque sous-disque de consignation est associé à l'un des périphériques virtuels du volume. Il ne peut y avoir qu'un seul sous-disque de consignation par périphérique virtuel. Si le périphérique virtuel ne contient qu'un sous-disque de consignation et aucun sous-disque de données, ce périphérique peut être appelé périphérique virtuel de consignation. Le sous-disque de consignation peut également être associé à un périphérique virtuel ordinaire qui contient des sous-disques de données ; dans un tel cas, le sous-disque de consignation risque de ne pas être disponible si le

périphérique virtuel doit être déconnecté en raison de la panne de l'un de ses sous-disques de données.

Utilisez l'interface graphique utilisateur ou l'utilitaire de ligne de commande `vxassist(1M)` pour créer un journal pour un volume existant.

Utilisation de la redistribution dynamique

La redistribution dynamique est la capacité d'un système à réagir automatiquement aux pannes d'E/S sur des objets redondants du gestionnaire de volumes (mis en miroir ou RAID5), ainsi qu'à rétablir la redondance et l'accès à ces objets. La redistribution dynamique n'est prise en charge que sous les configurations qui utilisent VxVM. VxVM détecte les pannes d'E/S sur les objets du gestionnaire de volumes et redistribue les sous-disques touchés aux disques désignés à titre de réserve ou d'espace libre dans le groupe de disques. Puis, VxVM reconstitue les objets qui existaient avant la panne, il rétablit leur redondance et les rend à nouveau accessibles.

Lorsqu'une panne partielle d'un disque a lieu (autrement dit, seuls certains sous-disques du disque sont touchés), les données redondantes qui figurent sur la partie défectueuse du disque sont redistribuées, et les volumes qui contiennent les parties intactes du disque demeurent accessibles.

Remarque - La redistribution dynamique n'est exécutée que pour les sous-disques redondants (mis en miroir ou RAID5) d'un disque défectueux. Les sous-disques non redondants d'un disque défectueux ne sont pas redistribués, mais le système vous avise de la panne.

Un disque de réserve doit être initialisé et placé dans un groupe de disques à titre de réserve pour qu'il puisse être utilisé à des fins de remplacement. Lorsqu'une panne survient, si aucun disque n'a été désigné comme réserve, VxVM utilise automatiquement l'espace libre, le cas échéant, du groupe de disques dans lequel la panne est survenue. Si l'espace du disque de réserve ne suffit pas, le système utilise cet espace de réserve ainsi que l'espace libre. Vous pouvez désigner un ou plusieurs disques, dans chaque groupe de disques, comme réserve pour la redistribution dynamique. Pour désigner les disques devant servir de réserve, utilisez la commande `vxedit(1M)`.

Utilisation des systèmes de fichiers VxFS

Vous pouvez configurer et spécifier les systèmes de fichier UFS ou VxFS d'un hôte logique se trouvant sur des volumes de type `fsgen`. Si un noeud de grappe est maître d'un hôte logique, les systèmes de fichiers de cet hôte logique qui sont associés aux groupes de disques sont montés sur les points de montage spécifiés du noeud maître.

Durant la reconfiguration d'un hôte logique, on doit vérifier les systèmes de fichiers à l'aide de la commande `fsck(1M)`. Bien que cette opération soit exécutée en mode parallèle non-interactif avec les systèmes de fichiers UFS, elle est susceptible d'avoir une incidence sur la durée globale d'exécution de la séquence de reconfiguration. La fonction de consignation des systèmes de fichiers UFS, SDS et VxFS réduit considérablement le temps d'exécution de la commande `fsck(1M)` avant le montage des systèmes de fichiers.

Lorsqu'on doit effectuer la commutation d'un service de données en plus de la récupération du volume, la récupération dure plus longtemps que la limite permise pour les étapes de reconfiguration. Ainsi le délai imparti est dépassé et l'opération est interrompue sur le noeud.

Par conséquent, lorsque vous configurez des volumes mis en miroir, ajoutez toujours un journal DRL afin de réduire le temps de récupération des volumes en cas de panne du système. Si des volumes mis en miroir sont utilisés dans l'environnement de grappe, la DRL doit être définie pour les volumes dépassant 500 méga-octets.

Faites appel à VxFS si vous utilisez des systèmes de fichiers volumineux (de plus de 500 méga-octets) pour les services de données HD. Dans la plupart des cas, VxFS n'est pas intégré à Sun Cluster et doit donc être acheté séparément de VERITAS.

Remarque - Il est possible de configurer des hôtes logiques disposant de très petits systèmes de fichiers mis en miroir ; cependant, en raison des risques de dépassement du délai imparti lié l'augmentation de la taille du système de fichiers, il est préférable d'utiliser la consignation de zones modifiées (DRL) ou des systèmes de fichiers VxFS.

Développement d'un système de fichier

Pour développer un volume entrelacé ou RAID5 contenant un système de fichiers, vous devez disposer d'espace libre sur le même nombre de disques que celui figurant actuellement dans le volume entrelacé ou RAID5. Par exemple, si vous disposez de quatre disques de 1 Go entrelacés les uns aux autres (constituant ainsi un système de fichiers de 4 Go) et que vous souhaitez ajouter 1 Go d'espace (afin d'avoir un système de 5 Go), vous devez installer quatre nouveaux disques offrant chacun au moins 0,25 Go d'espace libre. Autrement dit, vous ne pouvez pas ajouter un disque à un entrelacement de quatre disques.

L'interface utilisateur graphique de VxVM détermine le disque à utiliser pour le développement du système de fichiers. Pour sélectionner des disques particuliers en vue du développement du système de fichiers, employez plutôt l'interface de ligne de commande.

Les systèmes de fichiers UFS ne peuvent pas être réduits. La seule façon de "réduire" un système de fichiers consiste à recréer le volume, à lancer la commande `newfs` pour ce volume, puis à restaurer les données à partir de la copie de sauvegarde.

Administration de miroirs locaux

Les disques locaux peuvent être mis en miroir. Si un miroir tombe en panne, suivez les directives énoncées dans la documentation du gestionnaire de volumes pour remplacer le miroir défectueux et resynchroniser le disque de remplacement avec le bon disque.

Sauvegarde de données multihôtes à l'aide de Solstice Backup

Cette section comporte des conseils sur l'utilisation de Solstice Backup[™] pour la sauvegarde des systèmes de fichiers Sun Cluster.

Solstice Backup lance chaque copie du logiciel serveur sur un serveur unique. Solstice Backup s'attend à ce que la récupération des fichiers ait lieu à l'aide du même serveur physique que celui utilisé pour la sauvegarde.

Solstice Backup dispose de données considérables sur les ordinateurs physiques (noms et ID d'hôtes) correspondant au serveur et aux clients. Les informations de Solstice Backup relatives aux ordinateurs physiques sous-jacents sur lesquels les hôtes logiques sont configurés ont une incidence sur le type de stockage des index clients.

Ne placez pas la base de données `/nsr` de Solstice Backup sur les disques multihôtes. En effet, des conflits risquent de se produire si deux serveurs Solstice Backup tentent d'accéder à la même base de données `/nsr`.

Étant donné la façon dont Solstice Backup stocke les index clients, pour la sauvegarde d'un client particulier, n'utilisez pas des serveurs Solstice Backup différents à des jours différents. Faites en sorte qu'un hôte logique particulier ait toujours pour maître le même serveur physique lorsque des sauvegardes sont effectuées. Ainsi, les opérations de récupération s'exécuteront sans accroc.

Remarque - Par défaut, les systèmes Sun Cluster ne génèrent pas la liste complète des systèmes de fichiers pour la configuration de sauvegarde. Si la liste de sauvegarde contient le mot-clé "All", le fichier `/etc/vfstab` est passé en revue afin de déterminer quels systèmes de fichiers doivent être sauvegardés. Les fichiers Sun Cluster `vfstab` étant par défaut conservés dans `/etc/opt/SUNWcluster/conf/hanfs`, Solstice Backup ne les trouvera pas, à moins que vous ne précisiez explicitement les systèmes de fichiers Sun Cluster à sauvegarder. Lorsque vous testez vos procédures de sauvegarde, assurez-vous que tous les systèmes de fichiers Sun Cluster qui doivent être sauvegardés figurent dans la liste de systèmes de fichiers de Solstice Backup.

Quatre méthodes de configuration de Solstice Backup sont présentées ci-après. Choisissez celle qui convient à votre configuration de Sun Cluster. À cet égard, vous

pouvez tenir compte des temps de commutation. Une fois que vous avez choisi une méthode, continuez de l'utiliser afin que les opérations de récupération futures réussissent.

Voici une description des méthodes de configuration :

- Utilisez un serveur qui n'est pas un noeud de grappe et qui n'est pas à haute disponibilité, configuré comme serveur Solstice Backup.

Configurez un serveur supplémentaire différent des serveurs Sun Cluster, qui fera office de serveur Solstice Backup. Configurez les hôtes logiques en tant que clients du serveur. Pour obtenir des résultats optimaux, vérifiez toujours que les hôtes logiques sont configurés sur leurs maîtres par défaut respectifs avant d'effectuer la sauvegarde quotidienne. Cela peut exiger une commutation. Si les hôtes logiques ont pour maître des serveurs différents à des jours différents (peut-être suite à une relève), Solstice Backup risque d'être déconcerté s'il doit effectuer une récupération, en raison de la façon dont Solstice Backup stocke les index clients.

- Utilisez un serveur Sun Cluster configuré de façon à effectuer les sauvegardes locales.

Configurez l'un des serveurs Sun Cluster en vue de l'exécution des sauvegardes locales. Faites toujours passer les hôtes logiques au serveur Solstice Backup avant d'effectuer la sauvegarde quotidienne. Ainsi, si `phys-hahost1` et `phys-hahost2` sont les serveurs Sun Cluster et que `phys-hahost1` est le serveur Solstice Backup, faites toujours passer les hôtes logiques à `phys-hahost1` avant d'effectuer les sauvegardes. Une fois que les sauvegardes sont terminées, refaites passer l'hôte logique à son maître habituel, soit `phys-hahost2`.

- Utilisez les serveurs Sun Cluster configurés à titre de serveurs Solstice Backup.

Configurez chaque serveur Sun Cluster pour l'exécution des sauvegardes locales de l'hôte logique dont il est, par défaut, le maître. Assurez-vous toujours que les hôtes logiques sont configurés sur leurs maîtres par défaut respectifs avant d'effectuer la sauvegarde quotidienne. Cela peut exiger une commutation. Si les hôtes logiques ont pour maître des serveurs différents à des jours différents (peut-être suite à une relève), Solstice Backup risque d'être déconcerté s'il doit effectuer une récupération, en raison de la façon dont Solstice Backup stocke les index clients.

- Utilisez un serveur Sun Cluster configuré à titre de serveur Solstice Backup.

Configurez un serveur Sun Cluster afin qu'il sauvegarde son hôte logique localement et les frères et soeurs de celui-ci sur le réseau. Assurez-vous toujours que les hôtes logiques sont configurés sur leurs maîtres par défaut respectifs avant d'effectuer la sauvegarde quotidienne. Cela peut exiger une commutation. Si les hôtes logiques ont pour maître des serveurs différents à des jours différents (peut-être suite à une relève), Solstice Backup risque d'être déconcerté s'il doit effectuer une récupération, en raison de la façon dont Solstice Backup stocke les index clients.

Avec les quatre méthodes de sauvegarde présentées ci-dessus, un autre serveur peut être configuré afin d'effectuer temporairement les sauvegardes en cas de panne du serveur Solstice Backup désigné. Notez que vous ne pouvez ni utiliser le serveur Solstice Backup temporaire pour récupérer des fichiers sauvegardés par le serveur

Solstice Backup normal, ni récupérer des fichiers sauvegardés par le serveur temporaire depuis le serveur de sauvegarde normal.

Détection des défaillances de Sun Cluster

Cette annexe décrit la détection des défaillances pour Sun Cluster et traite des thèmes suivants :

- “Aperçu de la détection de défaillances”, page 304
- “Surveillance du réseau public (PNM)”, page 306
- “Vérification des défaillances de Sun Cluster”, page 307
- “Vérification des défaillances propres à un service de données”, page 308

Cette section présente un aperçu de la détection de défaillances par Sun Cluster. Cette détection de défaillances englobe trois méthodes générales :

- Un mécanisme de pulsation
- Une surveillance des défaillances des réseaux
- Une surveillance des défaillances de services de données particuliers

La surveillance des défaillances effectue des vérifications de validité afin de s'assurer que la responsabilité du problème est imputée au noeud défectueux et non au noeud fonctionnel.

Certains renseignements présentés sont propres à cette version de Sun Cluster, et peuvent changer au fur et à mesure que le produit évolue. Les estimations de temps précisées pour la détection des diverses défaillances sont approximatives et ne visent qu'à donner des indications du comportement général de Sun Cluster. Ce document n'est pas un manuel relatif au fonctionnement logique interne de Sun Cluster ; il ne décrit pas non plus une interface de programmation.

Aperçu de la détection de défaillances

Tel que mentionné à la section d'architecture de base de Sun Cluster, lorsqu'un serveur tombe en panne, l'autre serveur prend la relève. Un point important reste à déterminer : comment le serveur sait-il que l'autre est en panne ?

Sun Cluster emploie trois méthodes de détection des défaillances.

- Pulsation et surveillance du lien AGC – Ces moniteurs se servent des liens privés. Pour Ethernet, il existe deux moniteurs : un moniteur de lien AGC et un moniteur d'appartenance à la grappe. Pour SCI, on compte trois moniteurs : un moniteur de lien AGC, un moniteur d'appartenance à la grappe et un moniteur de pulsation SCI de bas niveau.
- Surveillance des défaillances du réseau – Toutes les connexions des serveurs au réseau public sont vérifiées : si un serveur n'est pas en mesure de communiquer par l'entremise du réseau public en raison d'une anomalie matérielle ou logicielle, un autre serveur de l'ensemble de serveurs prend la relève.
- Vérification des défaillances propre à un service de données – Chaque service de données de Sun Cluster effectue la détection de défaillances qui lui est propre. Cette dernière méthode consiste à déterminer si le service de données effectue des tâches utiles et pas seulement à savoir si l'ordinateur et le système d'exploitation semblent fonctionner.

Dans le cas des deuxième et troisième méthodes, un serveur vérifie si l'autre serveur envoie une réponse. Après avoir détecté une anomalie apparente, le serveur qui effectue la surveillance réalise différentes vérifications de validité sur lui-même avant de prendre, de force, la relève de l'autre serveur. Ces vérifications de validité visent à s'assurer que le problème qui touche le serveur qui effectue la vérification n'est pas la cause de l'absence de réponse de l'autre serveur. Ces vérifications de validité sont réalisées par `hactl(1M)`, un sous-programme de bibliothèque qui fait partie de l'environnement de base de Sun Cluster. Ainsi, le code de détection de défaillances propre à un service de données n'a qu'à lancer la commande `hactl(1M)` pour effectuer les vérifications de validité du serveur qui effectue la vérification. (Pour plus de détails, consultez la page de manuel `hactl(1M)`.)

Mécanisme de pulsation : moniteur d'appartenance à la grappe

Sun Cluster utilise un mécanisme de pulsation. Le traitement des pulsations est assuré par un processus en temps réel à priorité élevée qui est fixé en mémoire ; ainsi, il n'est pas soumis à l'échange de pages. Ce processus est appelé *moniteur d'appartenance à la grappe*. Dans une liste `ps(1)`, son nom est `clustd`.

Chaque serveur envoie le message “Tout va bien”, ou une pulsation, sur les deux liens privés, environ toutes les deux secondes. De plus, chaque serveur est à l’écoute des messages de pulsation émis par les autres serveurs, sur les deux liens privés. La réception de la pulsation sur un des liens privés suffit pour indiquer qu’un autre serveur fonctionne. Un serveur détermine qu’un autre serveur est en panne s’il ne reçoit pas de message de pulsation provenant de ce serveur pendant une période suffisante, soit environ 12 secondes.

Dans la stratégie globale de détection de défaillances, le mécanisme de pulsation du moniteur d’appartenance à la grappe est le moyen de première intervention. En cas d’absence de pulsation, les pannes du matériel et les anomalies du système d’exploitation sont immédiatement détectées. Il est également possible de détecter les problèmes globaux du système d’exploitation, par exemple la disparition du contenu de tous les tampons de communication. Le mécanisme de pulsation est également la méthode de détection de défaillances la plus rapide de Sun Cluster. Etant donné que le moniteur d’appartenance à la grappe fonctionne en temps réel et qu’il est fixé en mémoire, un court délai d’absence de pulsation est acceptable. En revanche, pour les autres méthodes de détection de défaillances, Sun Cluster ne doit pas indiquer qu’un serveur est en panne si celui-ci est tout simplement lent. Pour ces méthodes, on définit des délais relativement longs, équivalents à plusieurs minutes et, dans certains cas, deux dépassements ou plus du délai accordé sont requis pour que Sun Cluster prenne la relève.

Puisque le moniteur d’appartenance à la grappe tourne en temps réel et est fixé en mémoire, il se peut, paradoxalement, que le moniteur d’appartenance fonctionne même si son serveur n’effectue aucune tâche utile relative aux services de données. D’où l’utilité de la surveillance des défaillances propre à un service de données, décrite dans “Vérification des défaillances propres à un service de données”, page 308.

Vérification de validité du noeud qui effectue la vérification

La vérification des défaillances du réseau et la vérification des défaillances propre à un service de données exige de chaque noeud qu’il vérifie si un autre noeud envoie une réponse. Avant de prendre la relève, le noeud qui effectue la surveillance réalise différentes vérifications de validité élémentaires sur lui-même. Ces vérifications visent à s’assurer que le problème n’est pas imputable au noeud qui effectue la surveillance. Il s’agit également de faire en sorte que la relève du serveur qui semble être défectueux permette réellement d’améliorer la situation. Si on ne procède pas aux vérifications de validité, des *relèves erronées* risquent de se produire. Autrement dit, un noeud en panne pourrait, à tort, indiquer qu’un autre noeud n’envoie pas de réponse et prendre la relève du serveur qui fonctionne bien.

Le noeud qui effectue la vérification effectue les vérifications de validité suivantes sur lui-même avant de prendre la relève d’un autre noeud :

- Le noeud qui effectue la vérification détermine sa propre capacité à utiliser le réseau public (voir “Surveillance du réseau public (PNM)”, page 306).
- Le noeud qui effectue la vérification détermine également si ses propres services de données HD répondent. Tous les services de données HD qui sont exécutés par le noeud effectuant la vérification sont également contrôlés. Si l’un d’eux ne répond pas, la relève est interdite, selon l’hypothèse que le noeud qui effectue la vérification ne pourra pas faire tourner les services d’un autre noeud s’il n’arrive pas à exécuter les siens. De plus, l’absence de réponse de la part des services de données HD du noeud qui effectue la vérification peut indiquer que celui-ci a un problème sous-jacent pouvant provoquer l’échec de vérification de l’autre noeud. Sun Cluster HA pour NFS offre un exemple de ce phénomène : pour verrouiller un fichier sur un autre noeud, les démons `lockd` et `statd` du noeud qui effectue la vérification doivent fonctionner. En vérifiant la réponse de ses démons `lockd` et `statd`, le noeud qui effectue la vérification peut déterminer que l’absence de réponse de la part de ses propres démons est la cause de l’absence de réponse de l’autre noeud.

Surveillance du réseau public (PNM)

L’élément PNM a deux fonctions principales :

- Surveiller l’état des adaptateurs configurés d’un noeud et signaler les pannes générales des adaptateurs ou du réseau.
- Effectuer une reprise transparente en faveur d’autres adaptateurs de relève d’un noeud en cas de panne de l’adaptateur principal.

La PNM est mise en oeuvre à titre de démon (`pnmd`) qui recueille périodiquement les statistiques du réseau sur l’ensemble des interfaces de réseau public d’un noeud. Si les résultats indiquent des anomalies, `pnmd` tente d’identifier une des situations suivantes :

- Le réseau est au repos.
- Le réseau est en panne.
- L’interface réseau est en panne.

La PNM envoie ensuite une commande `ping` multidestinataires. La PNM place les résultats de ses recherches dans le CCD et compare les résultats locaux aux résultats d’autres noeuds (qui sont également placés dans le CCD). Cette comparaison sert à déterminer si le réseau est en panne ou si l’interface réseau est défectueuse. Si la PNM établit que l’interface réseau est défectueuse et que des adaptateurs de secours sont configurés, elle assure la reprise pour l’adaptateur réseau.

Remarque - Le `ping` multidestinaires établi par la PNM pourrait ne pas être compris par certains composants matériels non-Sun présents dans la configuration. Ainsi, devez-vous connecter directement un dispositif de réseau Sun au réseau que vous surveillez.

Les résultats de la surveillance PNM sont utilisés par diverses entités. La composante de reprise de l'adaptateur réseau de la PNM se sert des résultats de la surveillance pour déterminer si le recours à un adaptateur de reprise est justifié. Par exemple, si le réseau est en panne, aucune reprise d'adaptateur n'est effectuée. Les moniteurs de défaillances associés aux services de données HD SC et la commande API `hactl` utilisent la fonction PNM pour déterminer la cause de la panne des services de données. Les informations produites par PNM servent à établir s'il convient de transférer le service de données ainsi que l'emplacement du service de données après le transfert.

Les messages `syslog` enregistrés par la fonction PNM suite à la détection de pannes d'adaptateur sont lus par le gestionnaire SC, qui les traduit en icônes graphiques affichées par l'entremise de l'interface utilisateur graphique.

Il est également possible de lancer les utilitaires PNM depuis la ligne de commande afin de déterminer l'état des composants réseau. Pour de plus amples renseignements, consultez les pages de manuel `pnmset(1M)`, `pnmstat(1M)`, `pnmptor(1M)`, `pnmrtop(1M)`, et `pnmmd(1M)`.

Vérification des défaillances de Sun Cluster

PNM vérifie l'état du réseau public et commande un passage aux connexions de secours au besoin. Toutefois, en cas d'impossibilité complète d'accéder au réseau public, PNM n'assure pas la reprise pour les services de données ou les hôtes logiques. En pareil cas, PNM signale l'anomalie, mais il incombe à un vérificateur des défaillances externe de prendre en charge la commutation entre les noeuds de secours.

Si vous utilisez VxVM comme gestionnaire de volumes, l'environnement Sun Cluster est responsable de la surveillance de chaque groupe de sauvegarde de reprise d'adaptateur réseau (NAFO) défini par hôte logique, ainsi que de la mise en branle d'une commutation vers un noeud de secours, lorsque l'une des situations suivantes survient :

- Aucun accès au réseau public n'est possible (aucun groupe de sauvegarde NAFO n'est disponible) et le noeud de secours dispose d'au moins un groupe NAFO.

- L'accès au réseau public est en partie interrompu — au moins un groupe de sauvegarde NAFO est encore actif lorsque plus d'un groupe (plusieurs sous-réseaux) est défini pour un hôte logique — et le noeud de secours dispose d'un nombre supérieur de groupes de sauvegarde NAFO valides et actifs.

Si aucune de ces conditions n'existe, Sun Cluster ne commande pas de commutation.

Si vous utilisez Solstice DiskSuite comme gestionnaire de volumes, la perte de l'accès au réseau public entraîne l'interruption du noeud déconnecté ; en outre, les hôtes logiques qui ont ce noeud pour maître sont transférés au noeud de secours.

L'environnement Sun Cluster ne surveille les réseaux publics que si la configuration comprend un hôte logique et qu'un service de données est "activé" et enregistré sur cet hôte logique. Seuls les groupes de sauvegarde NAFO utilisés par un hôte logique sont surveillés.

Vérification des défaillances propres à un service de données

La vérification des défaillances propre à un service de données découle du fait que, même si le noeud serveur et le système d'exploitation tournent, les logiciels ou le matériel peuvent être dans un état tel qu'aucune tâche utile ne peut être exécutée par les services de données. Dans l'environnement global, la panne totale du noeud ou du système d'exploitation est détectée par le mécanisme de pulsation du moniteur d'appartenance à la grappe. Toutefois, un noeud peut fonctionner suffisamment bien pour que le mécanisme de pulsation continue à être exécuté même si le service de données n'effectue aucune tâche utile.

En revanche, la vérification des défaillances propre à un service de données ne doit pas forcément détecter la panne d'un noeud ou l'arrêt de l'envoi, par celui-ci, de messages de pulsation à la grappe. On présume que le moniteur d'appartenance à la grappe détecte ces anomalies et que la vérification des défaillances des services de données ne contient aucun processus pour remédier à ces situations.

Une vérification des défaillances de service de données agit comme un client du service de données. Une vérification des défaillances qui tourne sur un ordinateur surveille le service de données exporté par cet ordinateur et, ce qui est plus important encore, le service de données exporté par un autre serveur. Un serveur en panne n'est pas suffisamment fiable pour détecter ses propres anomalies : ainsi, chaque serveur surveille un autre noeud en plus de se vérifier lui-même.

En plus de se comporter comme un client, la vérification des défaillances propre à un service de données se sert également, dans certains cas, des statistiques du service de données pour déterminer si des tâches utiles sont exécutées ou non. Une vérification peut en outre détecter la présence de certains processus cruciaux pour un service de données particulier.

En général, la vérification des défaillances réagit à l'absence de service en commandant à un serveur de prendre la relève d'un autre serveur. Dans certains cas, la vérification des défaillances tente d'abord de redémarrer le service de données sur l'ordinateur initial avant de commander la relève. Si plusieurs redémarrages ont lieu à l'intérieur d'une brève période, on considère que cet ordinateur a des problèmes graves. En pareil cas, la relève par un autre serveur a lieu immédiatement, sans qu'un autre redémarrage local ne soit tenté.

Vérification des défaillances de Sun Cluster HA pour NFS

Le serveur de surveillance effectue deux types de vérification périodique du service NFS d'un autre serveur.

1. Le serveur qui effectue la vérification envoie l'indication `NULL RPC` à tous les processus démons du noeud cible qui doivent assurer un service NFS ; ces démons sont `rpcbind`, `mountd`, `nfsd`, `lockd` et `statd`.
2. Le serveur qui effectue la vérification effectue un essai bout en bout : il tente de monter un système de fichiers NFS depuis l'autre noeud, puis de lire et d'écrire un fichier dans ce système de fichiers. Il réalise cet essai bout en bout pour chaque système de fichiers que l'autre noeud partage actuellement. Etant donné que le montage est coûteux, on y a recours moins souvent qu'aux autres vérifications.

Si une de ces vérifications repère une anomalie, le noeud qui effectue la vérification envisagera la possibilité de prendre la relève du noeud en utilisation. Toutefois, certaines conditions peuvent empêcher la relève d'avoir lieu immédiatement :

- Délai de grâce pour redémarrage local – Avant d'effectuer la relève, le noeud qui effectue la vérification attend pendant une courte période :
 - Afin que le noeud défectueux puisse prendre connaissance de son anomalie et remédier à la situation en redémarrant localement ses propres démons
 - Afin que le noeud défectueux puisse réduire sa charge de travail (dans le cas où il serait simplement surchargé)

Après cette attente, le vérificateur effectue une nouvelle vérification et il ne commande la relève que si une anomalie est signalée. En général, deux dépassements du délai imparti, en ce qui concerne la vérification, sont requis pour qu'une relève ait lieu, afin de tenir compte des serveurs lents.

- Réseaux publics multiples – Si l'autre noeud est relié à plusieurs réseaux publics, le noeud qui effectue la vérification vérifie au moins deux de ceux-ci.
- Verrous – Certains utilitaires de sauvegarde emploient la fonction `lockfs(1M)`, qui interdit différents types de mises à jour d'un système de fichiers, afin que la sauvegarde puisse prendre un instantané d'un système de fichiers demeurant inchangé. Malheureusement, dans l'environnement NFS, la commande

`lockfs(1M)` signale que le système de fichiers est inaccessible. Ainsi, le message suivant apparaît à l'intention des clients NFS : Le serveur NFS ne répond pas. Avant d'effectuer la relève, le noeud qui effectue la vérification interroge l'autre noeud afin de déterminer si le système de fichiers est en mode `lockfs` et, si c'est le cas, la relève est interdite. La relève est interdite parce que la commande `lockfs` est une composante normale de tout processus administratif de sauvegarde. Précisons que les utilitaires de sauvegarde n'utilisent pas tous `lockfs`. Certains permettent en effet au service NFS de continuer sans interruption.

- Démons – L'absence de réponse de la part des démons `lockd` et `statd` n'entraîne pas une relève. Les démons `lockd` et `statd` assurent, de concert, le verrouillage réseau des fichiers NFS. Si ces démons n'envoient pas de réponse, la situation est simplement consignée dans `syslog`, et aucune relève n'a lieu. `lockd` et `statd`, dans le cadre de leurs tâches normales, doivent effectuer des RPC des ordinateurs clients, de sorte qu'un client en panne ou partitionné puisse provoquer l'arrêt de `lockd` et `statd` pendant une période prolongée. Ainsi, un client défectueux peut faire croire que `lockd` et `statd` sont en panne sur le serveur. Par ailleurs, si une relève de la part du serveur qui effectue la vérification a lieu, ce serveur sera sans doute interrompu de la même façon par le client défectueux. Dans le modèle actuel, un client défectueux ne provoque pas de relève erronée.

Suite à l'exécution de ces essais propres à Sun Cluster HA pour NFS, le processus d'établissement de la pertinence d'une relève se poursuit par des appels de la commande `hactl(1M)` (voir "Vérification de validité du noeud qui effectue la vérification", page 305).

Le serveur qui effectue la vérification vérifie également son propre service NFS. La logique employée est semblable à celle des essais de l'autre serveur, mais au lieu de prendre la relève, des messages d'erreur sont consignés dans `syslog`, et on tente de redémarrer les démons dont les processus n'existent plus. Autrement dit, le redémarrage d'un processus démon n'est effectué que si ce processus est terminé ou en panne. Le redémarrage d'un processus démon n'est pas tenté si ce processus existe toujours mais ne répond pas, car il faudrait alors mettre un terme au démon sans savoir quelles structures de données il met à jour. En outre, aucun redémarrage n'est tenté si le dernier redémarrage local a eu lieu récemment (il y a moins d'une heure). L'autre serveur reçoit plutôt l'indication d'envisager une relève (pourvu que cet autre serveur réussisse ses vérifications de validité). Enfin, le démon `rpcbind` n'est jamais redémarré, car aucun moyen ne permet d'informer les processus qui étaient inscrits sous `rpcbind` qu'ils doivent se réinscrire.

Vérification des défaillances SGBD-HD

Les vérifications des défaillances de Sun Cluster HA pour Oracle, Sun Cluster HA pour Sybase et Sun Cluster HA pour Informix surveillent de façon semblable le serveur de base de données. Les vérifications des défaillances du SGBD-HD se configurent par le lancement d'un des utilitaires, soit `haoracle(1M)`,

hasybase(1M) soit hainformix(1M). (Pour une description détaillée des options de ces utilitaires, voir les pages de manuel en ligne.)

Une fois que les utilitaires ont été configurés et activés, deux processus sont démarrés sur le noeud local, et deux autres sont lancés sur le noeud à distance, simulant ainsi un accès client. Le vérificateur de défaillances distant est amorcé par le démon `ha_dbms_serv` et démarré lorsque `hareg -y dataservicename` est lancé.

Le module SGBD-HD se sert de deux méthodes pour déterminer si le service SGBD est disponible. Tout d'abord, SGBD-HD extrait des données du SGBD lui-même :

- Sous Oracle, le tableau `V$SYSSTAT` est consulté.
- Sous Sybase, les variables globales `@@io_busy`, `@@pack_received`, `@@pack_sent`, `@@total_read`, `@@total_write` et `@@connections` sont consultées.
- Sous Informix, le tableau `SYSPROFILE` est consulté.

Si les données extraites révèlent que des tâches sont exécutées pour des clients, aucune autre vérification du SGBD n'est effectuée. Ensuite, si les données du SGBD précisent qu'aucune tâche n'est réalisée, SGBD-HD présente une petite transaction d'essai au SGBD. Si tous les clients sont au repos, les données du SGBD indiquent qu'aucune tâche n'est exécutée. Autrement dit, la transaction d'essai fait la distinction entre une panne de la base de données et une situation de repos normale. Etant donné que la transaction d'essai n'est exécutée que si les données indiquent qu'aucune activité n'a lieu, elle n'impose aucune charge additionnelle si la base de données est active. La transaction d'essai consiste à :

- Créer un tableau portant le nom `HA_DBMS_REM` ou `HA_DBMS_LOC`
- Entrer des valeurs dans le tableau créé
- Mettre à jour la valeur introduite
- Effacer le tableau créé

SGBD-HD filtre minutieusement les codes d'erreur produits par le SGBD, à l'aide d'un tableau qui précise les codes qui doivent provoquer ou non une relève. Par exemple, dans le cas de Sun Cluster HA pour Oracle, une condition `table space full` (espace de tableau plein) ne provoque pas une relève, car un administrateur doit intervenir pour remédier à la situation. (Si une relève avait lieu, le nouveau serveur maître se buterait au même problème, soit `table space full`.)

En revanche, si un code d'erreur tel que `could not allocate Unix semaphore` (impossible d'allouer le sémaphore UNIX) survient, Sun Cluster HA pour Oracle tente de redémarrer ORACLE localement sur ce serveur. Si un redémarrage local a eu lieu récemment, l'autre ordinateur prend plutôt la relève (après avoir réussi ses propres vérifications de validité).

Vérification des défaillances de Sun Cluster HA pour Netscape

Les moniteurs de défaillances de tous les services de données Sun Cluster HA pour Netscape ont recours à une méthode commune de surveillance de l'instance du service. Ils utilisent tous la surveillance des défaillances à distance et locale.

Le moniteur de défaillances qui tourne sur le noeud qui est actuellement le maître de l'hôte logique sur lequel le service de données s'exécute est appelé le moniteur de défaillances local. Le moniteur de défaillances qui tourne sur un noeud pouvant être le maître de l'hôte logique est appelé un moniteur de défaillances distant.

Les moniteurs de défaillances de Sun Cluster HA pour Netscape effectuent périodiquement une opération de service de données simple avec le serveur. Si cette opération échoue ou dépasse le délai accordé, on considère qu'il y a une anomalie.

Si une vérification échoue, la vérification des défaillances locale tente de redémarrer localement le service de données. En général, cette mesure suffit pour rétablir le service de données. La vérification à distance conserve un dossier des échecs de la vérification, mais elle ne met en branle aucune mesure. Suite à deux échecs successifs de la vérification (ce qui indique que le redémarrage du service de données n'a pas permis de corriger le problème), la vérification à distance lance la commande `hactl(1M)` en mode "relève" afin d'amorcer la reprise de l'hôte logique. Certains services de données Netscape utilisent un algorithme de fenêtre coulissante des réussites et des échecs des vérifications : si un nombre prédéterminé d'échecs survient dans la fenêtre, la vérification prend des mesures.

Vous pouvez utiliser la commande `hadsconfig(1M)` pour ajuster les valeurs d'intervalle de vérification et de délai des moniteurs de défaillances de Sun Cluster HA pour Netscape. Si vous réduisez l'intervalle de vérification des défaillances, la détection des anomalies est plus rapide, mais vous risquez de provoquer des reprises erronées en raison de problèmes passagers. Par ailleurs, si vous diminuez la valeur du délai de vérification, la détection des anomalies se rapportant au service de données est plus rapide, mais vous risquez de provoquer des reprises erronées si le service de données est simplement occupé en raison d'une charge importante. Dans la plupart des cas, les valeurs par défaut de ces paramètres sont adéquates. Ces paramètres sont décrits à la page de manuel `hadsconfig(1M)` ainsi qu'aux sections de configuration du chapitre consacré à chaque service de données dans le *Sun Cluster 2.2 Software Installation Guide*.

Vérification des défaillances de Sun Cluster HA pour DNS

La vérification des défaillances de Sun Cluster HA pour DNS effectue une opération `nslookup` afin de déterminer l'état du serveur Sun Cluster HA pour DNS. Elle recherche le nom de domaine de l'hôte logique Sun Cluster HA pour DNS dans le serveur Sun Cluster HA pour DNS. Selon la configuration de votre fichier `/etc/resolv.conf`, `nslookup` peut contacter d'autres serveurs si le serveur Sun Cluster HA pour DNS principal est en panne. Ainsi, l'opération `nslookup` peut réussir

même si le serveur Sun Cluster HA pour DNS principal est en panne. Pour prévenir cette situation, la vérification des défaillances détermine si les réponses proviennent du serveur Sun Cluster HA pour DNS principal ou d'autres serveurs.

Vérification des défaillances de Sun Cluster HA pour Netscape HTTP

La vérification des défaillances de Sun Cluster HA pour Netscape HTTP vérifie l'état du serveur `http` en tentant de se relier à celui-ci, à l'adresse de l'hôte logique figurant sur le port configuré. Précisons que le moniteur de défaillances se sert du numéro du port spécifié pour `hadsconfig(1M)` lors de la configuration de l'instance du service `nshttp`.

Vérification des défaillances de Sun Cluster HA pour Netscape News

La vérification des défaillances de Sun Cluster HA pour Netscape News détermine l'état du serveur de nouvelles en se reliant à celui-ci, à l'adresse IP de l'hôte logique et au numéro de port `nntp`. Puis, elle tente de lancer la commande `date NNTP` sur le serveur de nouvelles et vérifie si la réponse du serveur lui parvient à l'intérieur du délai précisé pour la vérification.

Vérification des défaillances de Sun Cluster HA pour Netscape Mail ou du serveur de messagerie

La vérification des défaillances de Sun Cluster HA pour Netscape Mail ou du serveur de messagerie détermine l'état du serveur de courrier ou de messagerie en le vérifiant sur les trois ports de service pris en charge par le serveur, soit SMTP, IMAP et POP3 :

- SMTP (port 25) — Exécute un message "hello" SMTP sur le serveur puis lance la commande `quit`.
- IMAP (port 143) — Exécute la commande `CAPABILITY IMAP4` suivie de la commande `LOGOUT IMAP4`.
- POP3 (port 110) — Exécute la commande `quit`.

Pour tous ces essais, la vérification des défaillances prévoit de recevoir en réponse une chaîne du serveur, à l'intérieur du délai de vérification. Précisons que si la vérification détecte une anomalie sur l'un des trois ports de service ci-dessus, on estime que le serveur est en panne. Afin d'éviter les reprises erronées, la vérification des défaillances `nsmail` fait appel à un algorithme de fenêtre coulissante pour effectuer le suivi des échecs et des réussites de la vérification. Si le nombre d'échecs précisés dans la fenêtre coulissante est supérieur au nombre prédéterminé, la vérification à distance commande une relève.

Vérification des défaillances de Sun Cluster HA pour Netscape LDAP

La vérification locale de Sun Cluster HA pour Netscape LDAP peut effectuer un nombre variable de redémarrages locaux avant d'amorcer une reprise. Le mécanisme de redémarrage local se sert d'un algorithme de fenêtre coulissante. Ainsi, une reprise n'a lieu que si le nombre maximal de tentatives de cette fenêtre est atteint.

La vérification à distance de Sun Cluster HA pour Netscape LDAP utilise une connexion Telnet simple avec le port LDAP pour vérifier l'état du serveur. Le numéro du port LDAP est celui qui est précisé à l'installation à l'aide de `hadsconfig(1M)`.

La vérification locale :

- Vérifie le serveur en lançant un script de surveillance. Le script recherche le nom LDAP commun "moniteur". Le nom commun est défini par le serveur de répertoires et il n'est utilisé que pour la surveillance. La vérification se sert de l'utilitaire `ldapsearch` pour effectuer cette opération.
- Tente de redémarrer le serveur localement suite à la détection d'une anomalie du serveur.
- Lance la commande `hactl(1M)` en mode `abandon` après avoir décidé que le noeud local ne peut pas lancer en toute confiance le serveur de répertoires, tandis que la vérification à distance lance la commande `hactl(1M)` en mode `relève`. Si l'hôte logique peut avoir plusieurs maîtres, toutes les vérifications à distance exécutent l'opération de relève à l'unisson. Toutefois, après la relève, l'environnement sous-jacent s'assure qu'un noeud maître unique est choisi pour le serveur de répertoires.

Vérification des défaillances de Sun Cluster HA pour Lotus

La vérification des défaillances de Sun Cluster HA pour Lotus comporte deux éléments — une vérification locale qui tourne sur le noeud où les processus du serveur Lotus Domino s'exécutent actuellement, et une vérification à distance qui tourne sur tous les autres noeuds qui sont des maîtres possibles de l'hôte logique du serveur Lotus Domino.

Les deux vérifications se servent d'une connexion Telnet simple avec le port Lotus Domino afin de vérifier l'état du serveur Domino. Si une vérification n'arrive pas à se connecter, elle lance une reprise ou une relève en appelant la commande `hactl(1M)`.

La vérification des défaillances locale peut effectuer trois redémarrages locaux avant de lancer une reprise. Le mécanisme de redémarrage local utilise un algorithme de fenêtre de temps coulissante. Ainsi, une reprise n'a lieu que si le nombre maximal de tentatives de cette fenêtre est atteint.

Vérification des défaillances de Sun Cluster HA pour Tivoli

Sun Cluster HA pour Tivoli ne se sert que d'une vérification des défaillances locale. Celle-ci tourne sur le noeud où le répartiteur d'objets Tivoli, soit le démon `oserv`, s'exécute actuellement.

La vérification des défaillances se sert de la commande Tivoli `wping` pour vérifier l'état du démon `oserv` observé. Les situations suivantes peuvent faire échouer la commande `wping` du démon `oserv` :

- Le démon `oserv` surveillé ne tourne pas.
- Le démon `oserv` du serveur se termine pendant la surveillance d'un démon `oserv` client.
- Les rôles Tivoli adéquats (autorisation) n'ont pas été définis pour l'utilisateur administratif. Pour plus de détails sur Tivoli, voir le *Sun Cluster 2.2 Software Installation Guide*.

Si la vérification locale n'arrive pas à détecter le démon `oserv` à l'aide de la commande `wping`, elle lance une reprise en appelant la commande `hactl(1M)`. La vérification des défaillances effectue un redémarrage local avant de lancer une reprise.

Vérification des défaillances de Sun Cluster HA pour SAP

La vérification des défaillances de Sun Cluster HA pour SAP surveille la disponibilité de l'instance Centrale, particulièrement le serveur de messagerie, le serveur de mise en file d'attente et le répartiteur. La vérification ne surveille que le noeud local en vérifiant la présence de processus SAP cruciaux. Elle utilise également l'utilitaire SAP `lgtst` pour vérifier s'il est possible de joindre le serveur de messagerie.

Dès qu'une anomalie est détectée, par exemple lorsqu'un processus se termine prématurément ou lorsque la commande `lgtst` signale une erreur, la vérification des défaillances tentera d'abord de redémarrer SAP sur le noeud local un certain nombre de fois (ce nombre est défini à l'aide de la commande `hadsconfig(1M)`). Si le nombre de redémarrages défini par l'utilisateur a été épuisé, la vérification des défaillances lance une commutation en appelant la commande `hactl(1M)`, si cette instance a été configurée afin de permettre une reprise (également modifiable à l'aide de la commande `hadsconfig(1M)`). L'instance Centrale est arrêtée avant que la commutation n'ait lieu, puis elle est redémarrée sur le noeud distant une fois que la commutation a été effectuée.

Affichage des messages LOG_DB_WARNING pour la vérification SAP

Le paramètre Sun Cluster HA pour SAP LOG_DB_WARNING détermine si les messages d'avertissement doivent ou non être affichés lorsque le système de vérification de Sun Cluster HA pour SAP ne peut se connecter à la base de données. Lorsque LOG_DB_WARNING est défini sur `y` et que le système de vérification ne peut se connecter à la base de données, un message est créé au niveau avertissement de la fonction `local0`. Par défaut, le démon `syslogd(1M)` n'affiche pas ces messages sur `/dev/console` ou sur `/var/adm/messages`. Pour les visualiser, vous devez modifier le fichier `/etc/syslog.conf` afin d'afficher les messages de priorité `local0.warning`. Par exemple :

```
...
*.err;kern.notice;auth.notice;local0.warning /dev/console
*.err;kern.debug;daemon.notice;mail.crit;local0.warning /var/adm/messages
...
```

Après avoir modifié le fichier, vous devez relancer `syslogd(1M)`. Pour plus d'informations, consultez les pages `syslog.conf(1M)` et `syslogd(1M)` du manuel.

Utilisation des méthodes de gestion SNMP de Sun Cluster

Cette annexe indique comment utiliser SNMP pour surveiller une configuration Sun Cluster.

- “Base d’informations de l’agent SNMP de grappe et de gestion de grappe”, page 318
- “Base d’informations de gestion de grappe”, page 319
- “Utilisation du démon SNMP de grappe et du démon super moniteur”, page 325
- “Trappes SNMP”, page 326
- “Modification du fichier `snmpd.conf`”, page 330
- “Configuration du port de l’agent SNMP de grappe”, page 332
- “Utilisation de l’agent SNMP avec SunNet Manager”, page 333

Vous pouvez utiliser les méthodes de gestion SNMP suivantes pour surveiller les configurations de Sun Cluster :

- Agent SNMP Sun Cluster
- Gestionnaire de domaine
- Gestionnaire Enterprise
- Gestionnaire Sun Net
- HP OpenView conforme à SNMP

Base d'informations de l'agent SNMP de grappe et de gestion de grappe

Sun Cluster comprend un agent de protocole de gestion de réseau simple (SNMP), ainsi qu'une base d'informations de gestion (BIG), pour la grappe. Le nom du fichier de l'agent est `snmpd` (démon SNMP), tandis que le nom de la BIG est `sun.mib`.

L'agent SNMP de grappe est un agent proxy qui est en mesure de surveiller plusieurs grappes (maximum de 32) en même temps. Vous pouvez gérer un système Sun Cluster ordinaire depuis le poste de travail administratif ou le processeur des services système (SSP). En installant l'agent SNMP de grappe sur le poste de travail administratif ou le SSP, le trafic réseau est contrôlé, et la puissance de l'UC des noeuds n'est pas utilisée inutilement au transfert de paquets SNMP.

Le démon `snmpd` :

- Est un agent SNMP conforme à RFC 1157.
- Est dédié au support des extensions BIG Sun Cluster (SC) sous le groupe Enterprise de Sun Microsystems, Inc.
- Offre la grappe `sun.mib` en format ASCII.
- Supporte les opérations du protocole SNMP, notamment `GET-REQUEST`, `GETNEXT-REQUEST` et `TRAP`.
- Offre l'agent super moniteur `smond` pour la collecte de données.

Le démon super moniteur `smond` recueille les informations sur la configuration matérielle et les événements de grappe critiques en se reliant au démon `in.mond` pour chacun des noeuds faisant partie de la ou des grappes. Le démon `smond` transmet ensuite ces informations au démon SNMP (`snmpd`).

Remarque - Vous ne devez configurer qu'un seul démon `smond` pour recueillir les informations sur plusieurs grappes.

Le logiciel `SUNWcsnmp` contient les éléments suivants :

- fichiers binaires `/opt/SUNWcluster/bin/snmpd` et `/opt/SUNWcluster/bin/smond`
- fichier ASCII `/opt/SUNWcluster/etc/sun.mib`
- script `/opt/SUNWcluster/bin/init.snmpd` (commande `snmpd`)
- fichier `/var/opt/SUNWcluster/snmpd.conf` (configuration SNMP)
- fichier `/opt/SUNWcluster/etc/snmp.traps` (trappes SNMP)
- fichier `/opt/SUNWcluster/etc/sun-snmp.schema` (schéma SunNet Manager)
- script `/opt/SUNWcluster/bin/smond_conf` (configuration `smond`)

- script `/opt/SUNWcluster/bin/smond_ctl` (commande `smond`)
- Pages de manuel correspondantes

Pour plus d'informations sur les démons `snmpd` et `smond`, référez-vous aux pages du manuel correspondantes.

Base d'informations de gestion de grappe

La base d'informations de gestion (BIG) est un ensemble d'objets auxquels on peut accéder à l'aide d'un protocole de gestion de réseau. La définition des objets doit être générique et cohérente, afin que diverses plates-formes de gestion puissent la lire et l'analyser.

Lancez le démon `snmpd` sur le serveur de gestion, qui est sur la station de travail d'administration de la grappe, ou sur n'importe quel autre client. Cet agent fournit des informations (recueillies à l'aide de `smond`) pour tous les attributs SNMP définis dans la BIG de grappe. Ce fichier BIG est en général compilé afin de constituer un gestionnaire de réseau prenant en charge "SNMP", comme la console SunNet Manager. Voir "Modification du fichier `snmpd.conf`", page 330.

Le fichier `sun.mib` donne des renseignements sur les grappes dans les tableaux suivants :

- `clustersTable`
- `clusterNodesTable`
- `switchesTable`
- `portsTable`
- `lhostTable`
- `dsTable`
- `dsinstTable`

Remarque - Dans les tableaux mentionnés ci-dessus, l'heure précisée est l'heure locale du serveur SNMP (où le tableau se trouve). Ainsi, l'heure indique à quel moment un changement d'attribut est reporté sur le serveur.

Attributs du `clustersTable`

Le tableau des grappes contient des entrées pour toutes les grappes surveillées. Chaque entrée du tableau contient des attributs particuliers qui fournissent des informations sur la grappe. Pour les attributs du `clustersTable`, voir Tableau C-1.

TABLEAU C-1 Attributs du `clustersTable`

Nom d'attribut	Description
<code>clusterName</code>	Nom de la grappe.
<code>clusterDescr</code>	Description de la grappe.
<code>clusterVersion</code>	Version de la grappe.
<code>numNodes</code>	Nombre de noeuds de la grappe.
<code>nodeNames</code>	Noms de tous les noeuds de la grappe, séparés par des virgules.
<code>quorumDevices</code>	Noms de tous les périphériques de quorum de la grappe, séparés par des virgules.
<code>clusterLastUpdate</code>	La dernière fois qu'un des attributs de cette entrée a été modifié.

Attributs du `clusterNodesTable`

Le tableau des noeuds de grappe comprend les noeuds connus de toutes les grappes surveillées. Chaque entrée contient des renseignements particuliers sur le noeud. Pour les attributs du `tableauNoeudsGrappe`, voir Tableau C-2.

Remarque - Lorsque vous utilisez une référence croisée, l'attribut `belongsToCluster` fait office de référence principale entre ce tableau et le `clusterNodesTable`.

TABLEAU C-2 Attributs du `clusterNodesTable`

Nom d'attribut	Description
<code>nodeName</code>	Nom d'hôte du noeud.
<code>belongsToCluster</code>	Nom de la grappe (à laquelle ce noeud appartient).
<code>scState</code>	Etat du logiciel Sun Cluster de ce noeud (arrêté, interrompu, en transition, inclus, exclu ou inconnu). Une trappe propre à Enterprise signale un changement d'état.
<code>vmState</code>	Etat du logiciel gestionnaire de volumes de ce noeud. Une trappe propre à Enterprise signale un changement d'état.
<code>dbState</code>	Etat du logiciel de base de données de ce noeud (en panne, en marche ou inconnu). Une trappe propre à Enterprise signale un changement d'état.
<code>vmType</code>	Type du gestionnaire de volumes qui est utilisé actuellement sur ce noeud.
<code>vmonNod</code>	Mode du logiciel VxVM figurant sur ce noeud (maître, esclave ou inconnu). Une trappe propre à Enterprise signale un changement d'état. Cet attribut n'est pas valide pour les grappes disposant d'autres gestionnaires de volumes.
<code>nodeLastUpdate</code>	La dernière fois qu'un des attributs de cette entrée a été modifié.

Attributs du `switchesTable`

Le tableau des commutateurs comprend des entrées pour tous les commutateurs. Chaque entrée du tableau contient des renseignements sur un commutateur de la grappe. Pour les attributs du `switchesTable`, voir Tableau C-3.

TABLEAU C-3 Attributs du `switchesTable`

Nom d'attribut	Description
<code>switchName</code>	Nom du commutateur.
<code>numPorts</code>	Nombre de ports du commutateur.
<code>connectedNodes</code>	Noms de tous les noeuds actuellement connectés aux ports du commutateur.
<code>switchLastUpdate</code>	La dernière fois qu'un des attributs du commutateur de cette entrée a été modifié.

Attributs du `portsTable`

Le tableau des ports comprend des entrées pour tous les ports de commutateur. Chaque entrée du tableau contient des renseignements sur un port particulier d'un commutateur. Pour les attributs du `portsTable`, voir Tableau C-4.

Remarque - Lorsque vous utilisez une référence croisée, l'attribut `belongsToSwitch` fait office de référence principale entre ce tableau et le `switchesTable`.

TABLEAU C-4 Attributs du `portsTable`

Nom d'attribut	Description
<code>portId</code>	Identification ou numéro du port.
<code>belongsToSwitch</code>	Nom du commutateur (auquel le port appartient).
<code>connectedNode</code>	Nom du noeud (auquel le port est actuellement connecté).
<code>nodeAdapterId</code>	Identification de l'adaptateur (de la carte SCI) du noeud auquel le port est connecté.

TABLEAU C-4 Attributs du `portsTable` (suite)

Nom d'attribut	Description
<code>portStatus</code>	Etat du port (actif, inactif, etc.).
<code>portLastUpdate</code>	La dernière fois qu'un des attributs du port de cette entrée a été modifié.

Attributs du `lhostTable`

Le tableau des hôtes logiques comporte des entrées pour chacun des hôtes logiques configurés dans la grappe. Pour les attributs du `lhostTable`, voir Tableau C-5.

TABLEAU C-5 Attributs du `lhostTable`

Nom d'attribut	Description
<code>lhostName</code>	Nom de l'hôte logique.
<code>lhostMast</code>	Liste des noms des noeuds qui constituent l'hôte logique.
<code>lhostCurrMaster</code>	Nom du noeud qui agit actuellement comme maître de l'hôte logique.
<code>lhostDS</code>	Liste des services de données qui sont configurés pour s'exécuter sur l'hôte logique.
<code>lhostDG</code>	Groupes de disques configurés sur l'hôte logique.
<code>lhostLogicalIP</code>	Adresse IP logique associée à l'hôte logique.
<code>lhostStatus</code>	Etat actuel de l'hôte logique (en marche ou en panne).
<code>lhostLastUpdate</code>	La dernière fois qu'un des attributs de cette entrée a été modifié.

Attributs du dsTable

Le tableau des services de données comprend des entrées pour tous les services de données configurés pour tous les hôtes logiques des grappes surveillées. Chaque entrée du tableau contient des renseignements sur un service de données configuré sur un hôte logique. Pour les attributs du dsTable, voir Tableau C-6.

Remarque - Lorsque vous utilisez une référence croisée, l'attribut dsonLhost fait office de référence principale entre ce tableau et le lhostTable.

TABLEAU C-6 Attributs du dsTable

Nom d'attribut	Description
dsName	Nom du service de données.
dsOnLhost	Nom de l'hôte logique où le service de données est configuré.
dsReg	Cette valeur est 1 ou 0, selon que le service de données est enregistré et configuré pour s'exécuter (1) ou ne pas s'exécuter (0).
dsStatus	Etat actuel du service de données (activé, désactivé, inst en panne).
dsDep	Liste des autres services de données dont ce service dépend.
dsPkg	Nom du logiciel du service de données.
dsLastUpdate	La dernière fois qu'un des attributs de cette entrée a été modifié.

Attributs de dsinstTable

Le tableau d'instance de service de données comprend des entrées pour toutes les instances de service de données. Pour les attributs du dsinstTable, voir Tableau C-7.

Remarque - Lorsque vous utilisez une référence croisée, l'attribut `dsinstOfDS` peut faire office de référence principale entre ce tableau et le `dsTable`. De même, l'attribut `dsinstOnLhost` peut être utilisé comme référence principale entre ce tableau et le `lhostTable`.

TABLEAU C-7 Attributs du `dsinstTable`

Nom d'attribut	Description
<code>dsinstName</code>	Nom de l'instance du service de données.
<code>dsinstOfDS</code>	Nom du service de données dont cette instance est tirée.
<code>dsinstOnLhost</code>	Nom de l'hôte logique sur lequel cette instance de service de données tourne.
<code>dsinstStatus</code>	Etat de l'instance du service de données.
<code>dsinstLastUpdate</code>	La dernière fois qu'un des attributs de cette entrée a été modifié.

Utilisation du démon SNMP de grappe et du démon super moniteur

Le démon SNMP, fonctionne suivant les modalités suivantes :

- Le démon `smond` se connecte à `in.mond` sur tous les noeuds de grappe demandés.
- Le démon `smond` transmet les informations `config` et `syslog` recueillies au démon `snmpd`.
- Le démon `snmpd` remplit les tableaux BIG de la grappe (qui sont mis à la disposition des clients par l'entremise des opérations SNMP GET).
- Le démon `snmpd` envoie des trappes propres à Enterprise pour les événements de grappe cruciaux lorsqu'il est avisé par des données `smond syslog`.

Trappes SNMP

Les trappes SNMP sont des avis asynchrones générés par l'agent SNMP, qui indiquent un changement imprévu de l'état des objets surveillés.

Le logiciel génère des trappes propres à Sun Cluster pour les événements de grappe cruciaux. Ces événements sont indiqués dans les tableaux ci-après.

Tableau C-8 présente la liste des trappes Sun Cluster qui indiquent l'état du logiciel de grappe sur un noeud.

TABLEAU C-8 Trappes Sun Cluster indiquant l'état du logiciel sur un noeud

Numéro de la trappe	Nom de la trappe
0	sc:arrêté
1	sc:interrompu
2	sc:en_transition
3	sc:inclus
4	sc:exclu
5	sc:inconnu

Tableau C-9 présente la liste des trappes Sun Cluster qui indiquent l'état du gestionnaire de volumes sur un noeud.

TABLEAU C-9 Trappes Sun Cluster indiquant l'état du gestionnaire de volumes sur un

TABLEAU C-9 Trappes Sun Cluster indiquant l'état du gestionnaire de volumes sur un noeud *(suite)*

noeud

Numéro de la trappe	Nom de la trappe
10	gv:panne
11	gv:marche
12	gv:inconnu

Tableau C-10 présente la liste des trappes Sun Cluster qui indiquent l'état de la base de données sur un noeud.

TABLEAU C-10 Trappes Sun Cluster indiquant l'état de la base de données sur un noeud

Numéro de la trappe	Nom de la trappe
20	bd:panne
21	bd:marche
22	bd:inconnu

Tableau C-11 présente la liste des trappes de Sun Cluster indiquant la nature de la grappe VxVM (maître ou esclave) sur un noeud.

TABLEAU C-11 Trappes de Sun Cluster indiquant la nature de VxVM sur un noeud

Numéro de la trappe	Nom de la trappe
30	gv_sur_noeud:maître
31	gv_sur_noeud:esclave
32	gv_sur_noeud:inconnu

Tableau C-12 présente la liste des trappes Sun Cluster qui indiquent l'état d'un hôte logique.

TABLEAU C-12 Trappes Sun Cluster indiquant l'état d'un hôte logique

Numéro de la trappe	Nom de la trappe
40	hôte:abandon
41	hôte:donné
42	hôte:relève
43	hôte:pris
46	hôte:inconnu

Tableau C-13 présente la liste des trappes Sun Cluster qui indiquent l'état d'une instance de service de données.

TABLEAU C-13 Trappes Sun Cluster indiquant l'état d'une instance de service de données

Numéro de la trappe	Nom de la trappe
50	sd:démarré
51	sd:arrêté

TABLEAU C-13 Trappes Sun Cluster indiquant l'état d'une instance de service de données (suite)

Numéro de la trappe	Nom de la trappe
52	sd:en_transition
53	sd:panne-locale
54	sd:panne-distante
57	sd:inconnu

Tableau C-14 présente la liste des trappes Sun Cluster qui indiquent l'état du service de données NFS-HD.

TABLEAU C-14 Trappes Sun Cluster indiquant l'état de l'instance du service de données NFS-HD

Numéro de la trappe	Nom de la trappe
60	nfshd:démarrage
61	nfshd:arrêt
70	nfshd:inconnu

Tableau C-15 présente la liste des trappes Sun Cluster qui indiquent les erreurs SNMP.

TABLEAU C-15 Trappes Sun Cluster indiquant les erreurs SNMP

Numéro de la trappe	Nom de la trappe
100	ERREUR_SOCKET:noeud_manque_ressources_système
101	ERREUR_CONNEXION:noeud_manque_ressources_système

TABLEAU C-15 Trappes Sun Cluster indiquant les erreurs SNMP (suite)

Numéro de la trappe	Nom de la trappe
102	ERREUR_MONDFAUTIF: noeud_exécute_version_mond_fautive/ ancienne
103	ERREUR_MONDABSENT: mond_pas_installé_sur_noeud
104	ERREUR_AUCUN_MOND: mond_sur_noeud_répond_pas: noeud_peut_réinitialiser
105	ERREUR_DELAI_ACCORDE: dépassement_délai_tentative_connexion_noeuds_mond
106	ERREUR_ACCES: mond_noeud_inaccessible: anomalies_réseau??
107	ERREUR_LECTURE: noeud_manque_ressources_système
108	ERREUR_AUCUNE_REPONSE: noeud_manque_ressources_système
109	ERREUR_REPONSE_ERRONEE: message_bienvenue_imprévu_de_mond_noeud
110	ERREUR_ARRET: arrêt_mond_noeud
200	Fatale: sortie_démon_super_moniteur(smond)!

Pour les trappes 100 à 110, vérifiez le noeud défectueux puis corrigez le problème. Pour la trappe 200, voir “Dépannage de SNMP”, page 335.

Modification du fichier `snmpd.conf`

Le fichier `snmpd.conf` est utilisé pour les informations de configuration. Chaque entrée du fichier comprend un mot clé suivi d'une chaîne de paramètres. Les valeurs par défaut qui figurent dans ce fichier devraient convenir à vos besoins.

▼ Comment modifier le fichier `snmpd.conf`

1. Editez le fichier `/snmpd.conf/`.

Pour la description des mots clés, reportez-vous à la page de manuel `snmpd(7)`.

2. Après avoir modifié le fichier `snmpd.conf`, arrêtez les programmes `smond` et `snmpd`, puis redémarrez les scripts en tapant :

```
# /opt/SUNWcluster/bin/smond_ctl stop
# /opt/SUNWcluster/bin/init.snmpd stop
# /opt/SUNWcluster/bin/init.snmpd start
# /opt/SUNWcluster/bin/smond_ctl start
```

Voici un exemple de fichier `snmpd.conf`.

```
sysdescr          Sun SNMP Agent, SPARCstation 10, Company
                  Property Number 123456
syscontact       Coby Phelps
sysLocation      Room 123
#
system-group-read-community    public
system-group-write-community   private
#
read-community   all_public
write-community  all_private
#
trap             localhost
trap-community   SNMP-trap
#
#kernel-file     /vmunix
#
managers         lvs golden
```

Configuration du port de l'agent SNMP de grappe

Par défaut, l'agent SNMP de grappe vérifie si des demandes arrivent, sur le port 161 du protocole datagramme utilisateur (UDP), du gestionnaire SNMP, par exemple la console SunNet Manager. Vous pouvez changer ce port à l'aide de l'option `-p` des démons `snmpd` et `smond`.

Les démons `snmpd` et `smond` doivent tous deux être configurés sur le *même* port pour qu'ils fonctionnent correctement.



Attention - Si vous installez l'agent SNMP de grappe sur un SSP ou sur un poste de travail administratif qui tourne sous Solaris 2.6 ou des versions compatibles, configurez toujours les programmes `snmpd` et `smond` sur un port autre que le port UDP 161 par défaut.

Par exemple, avec le SSP, l'agent SNMP de grappe nuit à l'agent SNMP SSP, qui utilise également le port UDP 161. Cette obstruction risque d'entraîner la disparition des caractéristiques RAS du serveur Sun Enterprise 10000.

▼ Comment configurer le port de l'agent SNMP de grappe

Pour configurer l'agent SNMP de grappe sur un port différent du port 161 par défaut, effectuez les étapes ci-dessous.

1. **Modifiez le fichier** `/opt/SUNWcluster/bin/init.snmpd` **et changez la valeur de la variable** `CSNMP_PORT`, **qui est de 161, en faveur du nombre souhaité.**
2. **Modifiez le fichier** `/opt/SUNWcluster/bin/smond_ctl` **et changez la valeur de la variable** `CSNMP_PORT`, **qui est de 161, en faveur du nombre choisi à Étape 1, page 332.**
3. **Arrêtez puis redémarrez les démons** `snmpd` **et** `smond` **pour valider les modifications.**

```
# /opt/SUNWcluster/bin/smond_ctl stop
# /opt/SUNWcluster/bin/init.snmpd stop
```

(suite)

(suite)

```
# /opt/SUNWcluster/bin/smond_ctl start  
# /opt/SUNWcluster/bin/init.snmpd start
```

Remarque - Vous pouvez avoir à modifier les fichiers de configuration propres au gestionnaire SNMP afin que celui-ci soit au courant du nouveau numéro de port. Pour de plus amples renseignements, reportez-vous à la documentation du gestionnaire SNMP. Vous pouvez également configurer l'agent SNMP maître sur le poste de travail administratif afin de lancer l'agent proxy SNMP de grappe à titre de sous-agent sur un port autre que 161. Pour plus de détails sur la configuration de l'agent SNMP maître, reportez-vous au *Guide de l'utilisateur des agents Solstice Enterprise* ou à la page du manuel `snmpdx(1M)`.

Utilisation de l'agent SNMP avec SunNet Manager

SunNet Manager est intégré à l'agent SNMP de grappe SunNet Manager. Effectuez les procédures suivantes avant d'utiliser SunNet Manager pour surveiller les grappes.

Remarque - Ces procédures présument que vous utilisez le port UDP 161 pour SNMP. Si vous avez changé le numéro de port conformément à "Configuration du port de l'agent SNMP de grappe", page 332, vous devez lancer l'agent proxy SNMP de SunNet Manager, `na.snmp`, pour pouvoir utiliser l'autre port.

▼ Comment utiliser l'agent SNMP avec SunNet Manager pour surveiller les grappes

1. Copiez la BIG de la grappe, /opt/SUNWcluster/etc/sun.mib, dans le fichier /opt/SUNWconn/snm/agents/cluster.mib sur la console SunNet Manager.
2. Sur la console SunNet Manager exécutez mib2schema pour le fichier cluster.mib que vous avez copié :

```
# /opt/SUNWconn/snm/bin/mib2schema cluster.mib
```

3. Sur le poste de travail administratif Sun Cluster, modifiez le fichier snmpd.conf puis indiquez, dans la chaîne de paramètres du mot clé trap, le nom de la console SunNet Manager.
Pour de plus amples renseignements sur la modification du fichier snmpd.conf, reportez-vous à "Modification du fichier snmpd.conf", page 330.
4. Exécutez la commande smond_conf sur le poste de travail administratif Sun Cluster pour chaque grappe à surveiller. Par exemple :

```
# /opt/SUNWcluster/bin/smond_conf -h [nomgrappe ...]
```

5. Pour le proxy de grappe-snmplib, indiquez le nom de la console SunNet Manager.

Remarque - Pour surveiller les grappes, vous devez également surveiller le poste de travail administratif à l'aide de SunNet Manager.

▼ Comment reconfigurer smond pour surveiller une autre grappe

Vous pouvez reconfigurer le démon smond afin de surveiller une grappe différente.

1. Arrêtez le démon snmpd en tapant :

```
# /opt/SUNWcluster/bin/init.snmpd stop
```

2. Reconfigurez le démon smond en tapant :

```
# /opt/SUNWcluster/bin/smond_conf -h [nomgrappe ...]
```

3. Démarrez le démon snmpd en tapant :

```
# /opt/SUNWcluster/bin/init.snmpd start
```

4. Démarrez le démon smond en tapant :

```
# /opt/SUNWcluster/bin/smond_ctl start
```

Dépannage de SNMP

Si les tableaux de base d'informations de gestion (BIG) de la grappe ne contiennent pas votre application ou si vous recevez la trappe numéro 200, assurez-vous que les démons snmpd et smond tournent en tapant :

```
# ps -ef | grep snmpd  
# ps -ef | grep smond
```

Si les démons ne tournent pas, aucune sortie n'est affichée.

Si les démons ne tournent pas, tapez :

```
# /opt/SUNWcluster/bin/init.snmpd start  
# /opt/SUNWcluster/bin/smond_ctl start
```


Glossaire

Agent de gestion de commutation (AGC)	Composant logiciel assurant la gestion des sessions pour les liens et les commutateurs SCI et Ethernet.
Agent SNMP de grappe	L'agent SNMP (protocole de gestion de réseau simple) de grappe est utilisé pour surveiller simultanément plusieurs grappes (jusqu'à concurrence de 32).
Bande	Semblable à une concaténation, la bande se distingue toutefois par l'absence de chevauchement dans l'adressage des blocs de composants et un entrelacement sur les tranches (partitions) plutôt qu'une disposition séquentielle. La partition en bandes permet d'améliorer les performances. Lorsque des données sont en bandes sur plusieurs disques gérés par des contrôleurs différents, il est possible d'y accéder simultanément à partir de multiples contrôleurs.
Base de données d'état des métapériphériques	Informations conservées en mémoire rémanente (sur disque) pour assurer le maintien de l'état et de la configuration des métapériphériques.
Base de données de configuration de grappe (BCG)	Base de données répliquée à haute disponibilité pouvant servir au stockage des données des services de données HD et à divers autres usages de configuration de Sun Cluster.
Base de données parallèle	Image d'une base de données unique à laquelle plusieurs utilisateurs peuvent accéder simultanément à partir d'hôtes multiples.
Base de données des états DiskSuite	Base de données répliquée utilisée pour stocker la configuration et l'état des métapériphériques.

Commutation	Déplacement coordonné d'un hôte logique d'un serveur HD opérationnel à un autre. La commutation est effectuée par l'administrateur à l'aide de la commande <code>haswitch(1M)</code> .
Concaténation	Métapériphérique créé par mappage séquentiel de blocs sur plusieurs tranches physiques (partitions) d'un périphérique logique. Deux composants physiques ou plus peuvent être concaténés. Les tranches sont accessibles de façon séquentielle et ne sont pas entrelacées (comme c'est le cas dans une bande).
Concentrateur de terminaux	Périphérique utilisé pour permettre à un poste de travail administratif de communiquer de façon sécuritaire avec tous les noeuds de Sun Cluster.
Configuration symétrique	Configuration à deux noeuds où un serveur sert de relève dynamique pour l'autre.
Connexions par canal à fibres optiques	Connexions par fibre reliant les noeuds au SPARCstorage Arrays.
Consignation UFS	Enregistrement des mises à jour UFS dans un journal (le périphérique de consignation) avant l'application des mises à jour dans le système de fichiers UFS (le périphérique maître).
Défaillance partielle	Transfert en cas de défaillance dans un sous-ensemble d'hôtes logiques sous la maîtrise d'un hôte physique unique.
Détection de défaillances	Les programmes Sun Cluster détectent deux types de défaillances. Le premier type couvre les pannes de bas niveau comme les erreurs système graves et les défaillances matérielles (c'est-à-dire celles qui rendent le serveur complètement inopérant). Ces problèmes peuvent être détectés rapidement. Le deuxième type de défaillances sont celles affectant les services de données, et il faut parfois plus de temps pour les détecter.
Disques locaux	Disques attachés à un serveur HD mais ne faisant pas partie d'un ensemble de disques. Les disques locaux contiennent le module de distribution Solaris et le module de gestion de volumes de Sun Cluster. Les disques locaux ne doivent pas contenir de données exportées par le service de données Sun Cluster.
Disque multihôtes	Disque configuré de façon à être accessible à partir de multiples serveurs. Le logiciel Sun Cluster permet d'exporter les données contenues sur un disque multihôtes vers des clients réseau par l'intermédiaire d'un service de données à haute disponibilité.

Écriture miroir	Reproduction, sur des périphériques multiples (les sous-miroirs), de toutes les opérations d'écriture effectuées sur un périphérique logique unique (le miroir), accompagnée d'une distribution des opérations de lecture. Cette fonction assure la redondance des données en cas de défaillance.
Ensemble de disques	Voir «Groupe de disques».
Gestionnaire de verrouillage distribué (GVR)	Logiciel de verrouillage utilisé dans un environnement de disques partagés Oracle7 ou de serveur parallèle Oracle8 (OPS). Le GVR permet aux processus Oracle exécutés sur des noeuds différents de synchroniser l'accès à la base de données. Le GVR est conçu pour assurer une disponibilité élevée. Lorsqu'un processus ou un noeud tombe en panne, il n'est pas nécessaire d'arrêter et de redémarrer les noeuds restants. Une reconfiguration rapide du GVR est alors exécutée pour assurer la reprise.
Grappe	Groupe de deux à quatre noeuds configurés ensemble pour l'exécution d'un logiciel de base de données parallèle ou des services de données à haute disponibilité.
Groupe de disques	Dans une configuration HD, groupe bien défini de disques multihôtes utilisé comme une unité unique par deux serveurs. Il peut s'agir, par exemple, d'un ensemble de disques Solstice DiskSuite ou d'un groupe de disques VERITAS Volume Manager.
Groupe de relève	Utilisé par la reprise sur panne d'adaptateur réseau (NAFO). Ensemble d'adaptateurs réseau situés sur le même sous-réseau. Les adaptateurs d'un ensemble assurent leur sauvegarde mutuelle.
Hôte	Ordinateur pouvant faire partie d'une grappe Sun. Dans la documentation Sun Cluster, «hôte» et «noeud» sont synonymes.
Hôte frère	Un des serveurs physiques dans une configuration HD symétrique.
Hôte médiateur	Hôte fournissant un "troisième vote" décisif par exécution du démon <code>rpc.metamed(1M)</code> , et qui a été ajouté à un ensemble de disques.
Hôte multirésident	Hôte résident sur plusieurs réseaux publics.
Hôtes logiques	Ensemble de ressources utilisé comme une seule unité par des serveurs HD. Dans le produit actuel, les ressources comprennent un ensemble de noms d'hôtes réseau et leurs adresses IP respectives,

ainsi qu'un groupe de disques (ensemble de disques). Chaque hôte logique est sous la maîtrise d'un seul hôte physique à la fois.

Interconnexion de grappes	Interface réseau privée entre les noeuds d'une grappe.
Interface cohérente échelonnable	Interconnexion à haute vitesse utilisée comme interface de réseau privé.
Interface réseau logique	Dans l'architecture Internet, un hôte peut porter une ou plusieurs adresses IP. HD configure des interfaces réseau logiques supplémentaires pour mapper plusieurs interfaces réseau logiques à une seule interface réseau physique. Cela permet à une interface réseau physique unique de répondre à de multiples interfaces réseau logiques. Cela permet également à l'adresse IP de passer d'un serveur HD à un autre en cas de relève ou de commutation (commande <code>haswitch(1M)</code>) sans qu'une interface matérielle supplémentaire ne soit nécessaire.
Liens privés	Réseau privé reliant les noeuds utilisés pour envoyer et recevoir les pulsations entre membres d'un ensemble de serveurs.
Maître	Serveur possédant des droits d'accès exclusifs (en lecture et en écriture) à un ensemble de disques. L'hôte maître actuel de l'ensemble de disques exécute le service de données, et les adresses IP logiques sont mappées à son adresse Ethernet.
Maître par défaut	Noeud configuré pour maîtriser un groupe de disques lorsque les hôtes logiques sont configurés.
Maître potentiel	Hôte physique quelconque pouvant maîtriser un hôte logique particulier.
Médiateur	Dans une configuration à deux chaînes, le médiateur fournit un "troisième vote" décisif lorsqu'il faut déterminer si l'accès aux répliques de base de données d'état des métapériphériques doit être accordé ou refusé. Le médiateur n'est utilisé que si exactement la moitié des répliques de base de données d'état des métapériphériques est accessible.
Médiateur or	Dans les configurations Solstice DiskSuite, état intrinsèque d'un hôte médiateur en vigueur lorsque des conditions particulières existaient lors de la dernière mise à jour des données de médiateur. Cet état permet l'exécution des opérations de type «take» même si un quorum d'hôtes médiateurs n'est pas disponible.

Métapériphérique	Groupe de composants accessibles en tant que périphérique logique unique par concaténation, partition en bandes, écriture miroir ou consignation de périphérique physiques. Les métapériphériques sont parfois appelés «pseudo-périphériques».
Moniteur d'appartenance	Processus de surveillance des serveurs exécuté sur tous les serveurs HD. Le moniteur d'appartenance envoie des pulsations à ses hôtes frères et en reçoit. Le moniteur peut déclencher une relève si l'émission de pulsations cesse. Il effectue également le suivi des serveurs actifs.
Moniteur d'appartenance à une grappe (MAG)	Logiciel gérant une liste cohérente d'appartenance aux grappes afin d'éviter l'altération de la base de données et la transmission subséquente aux clients de données erronées ou incohérentes. Lorsque des noeuds sont inclus ou retirés d'une grappe - et que les données d'appartenance doivent être modifiées - le MAG traite ces noeuds de façon à coordonner la reconfiguration globale des divers services système.
Moniteur de défaillances	Démon et programmes utilisés pour analyser les diverses parties des services de données.
Noeud	Ordinateur pouvant faire partie d'une grappe Sun. Dans la documentation Sun Cluster, «noeud» et «hôte» sont synonymes.
Noeud de grappe	Ordinateur faisant partie d'une grappe Sun. Egalement appelé «hôte de grappe» ou «serveur de grappe».
Nom d'hôte logique	Nom attribué à une des interfaces réseau logiques. Le nom d'hôte logique est utilisé par les clients du réseau pour désigner l'emplacement des données et des services de données. Le nom d'hôte logique est le nom du chemin de l'hôte logique. Etant donné qu'un hôte peut se trouver sur plusieurs réseaux, il arrive qu'un hôte logique porte plusieurs noms d'hôte logique.
Nom d'hôte logique principal	Nom désignant un hôte logique sur le réseau public principal.
Nom d'hôte logique secondaire	Nom désignant un hôte logique sur un réseau public secondaire.
Nom d'hôte physique principal	Nom désignant un hôte physique sur le réseau public principal.
Nom d'hôte physique secondaire	Nom désignant un hôte physique sur un réseau public secondaire.
Périphérique de consignation UFS	Dans les configurations Solstice DiskSuite, composant d'un périphérique trans contenant le journal UFS.

Périphérique de quorum	Dans les configurations VxVM, le système vote par quorum majoritaire afin d'éviter le partitionnement du réseau. Comme deux noeuds ne peuvent pas dégager un quorum majoritaire, un périphérique de quorum est inclus lors du vote. Ce périphérique peut être un contrôleur ou un disque.
Périphérique directement attaché	Unité de stockage sur disque connectée physiquement à tous les noeuds de la grappe.
Périphérique maître UFS	Dans les configurations Solstice DiskSuite, composant d'un périphérique trans contenant le système de fichiers UFS.
Périphérique trans	Dans les configurations Solstice DiskSuite, pseudo-périphérique assurant la gestion du contenu d'un journal UFS.
Poste de travail administratif	Poste de travail extérieur à la grappe, ou un des noeuds de grappe utilisé pour exécuter le logiciel d'administration de grappe.
Pulsation	Message périodique que s'envoient mutuellement plusieurs moniteurs d'appartenance. L'absence d'une pulsation après le délai ou le nombre de tentatives spécifié peut entraîner une relève.
Quorum BCG	Ensemble des bases de données de configuration de grappe nécessaire pour la sélection d'une copie valide et cohérente de la base de données de configuration de grappe.
Quorum de grappes	Nombre des noeuds de grappe pouvant prendre une décision au nom de l'ensemble.
Quorum de médiateurs	Condition satisfaite lorsque la moitié + 1 des hôtes médiateurs sont accessibles.
Quorum de répliques	Concept de Solstice DiskSuite désignant les situations où LA MOITIÉ + 1 des répliques de base de données d'état des métapériphériques sont accessibles.
Quorum MAG	Voir «Quorum de grappes».
Reconfiguration de grappe	Processus ordonné à plusieurs étapes appelé chaque fois qu'une modification importante est apportée à l'état de la grappe (relève, commutation ou réinitialisation de l'hôte physique). Lorsqu'une grappe est reconfigurée, le logiciel Sun Cluster coordonne tous les hôtes physiques qui ont été lancés et sur lesquels des communications sont en cours. Ces hôtes identifient ensemble les

	hôtes physiques maîtres et les hôtes logiques sous la maîtrise de ceux-ci.
Relève	Déplacement automatique d'un hôte logique d'un serveur HD à un autre après détection d'une défaillance. Le serveur HD défaillant est alors contraint d'abandonner la maîtrise de l'hôte logique.
Réplique	Voir «Réplique de base de données d'état des métapériphériques».
Réplique de base de données d'état des métapériphériques	Copie de la base de données d'état. La conservation de plusieurs exemplaires de cette base de données permet d'éviter les pertes de données d'état et de configuration. Ces informations sont essentielles pour l'exécution de toutes les opérations de métapériphérique.
Réseau public principal	Nom désignant le premier réseau public.
Réseau public secondaire	Nom désignant le deuxième réseau public ou les réseaux publics suivants.
Serveur	Ordinateur pouvant faire partie d'une grappe Sun. Dans la documentation Sun Cluster, «noeud» et «hôte» sont synonymes.
Serveur actif	Dans une configuration Sun Cluster, noeud assurant des services de données à haute disponibilité.
Serveur de relève dynamique	Dans une configuration N+1, noeud connecté à tous les disques multihôtes de la grappe. Ce serveur est également le noeud administratif. Si un ou plusieurs des noeuds actifs sont défaillants, les services de données passent du noeud défectueux au serveur de relève dynamique. Cependant, il n'est pas nécessaire que le noeud +1 s'abstienne d'exécuter les services de données en mode de fonctionnement normal.
Service de données	Service réseau permettant aux clients du réseau d'accéder aux données sur disque en lecture et en écriture. Le NFS est un exemple de service de données. Un service de données peut comporter de multiples processus fonctionnant de concert.
Service de données à haute disponibilité	Service de données apparaissant comme toujours disponible même en cas de panne ponctuelle du matériel de serveur ou des composants logiciels.
Solstice DiskSuite	Produit logiciel assurant la fiabilité des données par partition en bandes, concaténation et mise en miroir de disques, consignation UFS, développement dynamique des métapériphériques et des

systèmes de fichiers, et répliques de base de données d'état des métapériphériques.

Sous-miroir	Métapériphérique faisant partie d'un miroir. Voir aussi «Ecriture miroir».
Sun Cluster	Logiciel et matériel permettant à plusieurs ordinateurs d'agir comme serveurs de données en mode lecture/écriture tout en assurant leur sauvegarde réciproque.
Système de fichiers administratif HD	Système de fichiers spécial créé sur chaque hôte logique lorsque Sun Cluster est installé pour la première fois. Ce système de fichiers est utilisé par Sun Cluster et par les services de données en couches pour stocker des copies de leurs données administratives.
Topologie de paire de grappes	Deux paires de noeuds Sun Cluster fonctionnant dans un cadre administratif à une grappe.
Topologie échelonnable	Voir <i>Topologie N à N</i> .
Topologie en anneau	Configuration selon laquelle un serveur principal et un serveur de sauvegarde sont spécifiés pour chaque ensemble de services de données.
Topologie N à N	Configuration selon laquelle tous les noeuds sont connectés directement à un ensemble de disques partagés.
Topologie N+1	Configuration fondée sur un nombre (N) de serveurs actifs et un (+1) serveur de relève dynamique. Les serveurs actifs assurent des services de données continus, et le serveur de relève dynamique effectue la reprise du traitement de ces services en cas de défaillance d'un ou de plusieurs serveurs actifs.
UFS	Acronyme désignant le système de fichier UNIX®
Unité d'expansion de disque	Enceinte de stockage physique contenant les disques multihôtes. Par exemple, SPARCstorage Arrays, Sun StorEdge Multipacks, Sun StorEdge A3000 et Sun StorEdge A5000.
Unité d'expansion de disque multihôtes	Voir <i>Unité d'expansion de disque</i> .
Verrouillage de noeud	Mécanisme utilisé dans des trappes à plus de deux noeuds. Dans la documentation Sun Cluster, «noeud» et «hôte» sont synonymes.