



Sun StorageTek™ Availability Suite 4.0 ソフトウェアインストール および構成マニュアル

Sun Microsystems, Inc.
www.sun.com

Part No. 819-6358-10
2006 年 6 月, Revision A

コメントの送付: <http://www.sun.com/hwdocs/feedback>

Copyright 2006 Sun Microsystems, Inc., 4150 Network Circle, Santa Clara, California 95054, U.S.A. All rights reserved.

米国 Sun Microsystems, Inc. (以下、米国 Sun Microsystems 社とします)は、本書に記述されている技術に関する知的所有権を有しています。これら知的所有権には、<http://www.sun.com/patents>に掲載されているひとつまたは複数の米国特許、および米国ならびにその他の国におけるひとつまたは複数の特許または出願中の特許が含まれています。

本書およびそれに付属する製品は著作権法により保護されており、その使用、複製、頒布および逆コンパイルを制限するライセンスのもとにおいて頒布されます。サン・マイクロシステムズ株式会社による事前の許可なく、本製品および本書のいかなる部分も、いかなる方法によっても複製することが禁じられます。

本製品のフォント技術を含む第三者のソフトウェアは、著作権法により保護されており、提供者からライセンスを受けているものです。

本製品のの一部は、カリフォルニア大学からライセンスされている Berkeley BSD システムに基づいていることがあります。UNIX は、X/Open Company Limited が独占的にライセンスしている米国ならびに他の国における登録商標です。

本製品は、株式会社モリサワからライセンス供与されたリュウミン L-KL (Ryumin-Light) および中ゴシック BBB (GothicBBB-Medium) のフォント・データを含んでいます。

本製品に含まれる HG 明朝 L と HG ゴシック B は、株式会社リコーがリョービマジクス株式会社からライセンス供与されたタイプフェースマスタをもとに作成されたものです。平成明朝体 W3 は、株式会社リコーが財団法人日本規格協会 文字フォント開発・普及センターからライセンス供与されたタイプフェースマスタをもとに作成されたものです。また、HG 明朝 L と HG ゴシック B の補助漢字部分は、平成明朝体 W3 の補助漢字を使用しています。なお、フォントとして無断複製することは禁止されています。

Sun、Sun Microsystems、Java、AnswerBook2、docs.sun.com、Sun StorageTek は、米国およびその他の国における米国 Sun Microsystems 社の商標もしくは登録商標です。サンロゴマークおよび Solaris は、米国 Sun Microsystems 社の登録商標です。

すべての SPARC 商標は、米国 SPARC International, Inc. のライセンスを受けて使用している同社の米国およびその他の国における商標または登録商標です。SPARC 商標が付いた製品は、米国 Sun Microsystems 社が開発したアーキテクチャーに基づくものです。

OPENLOOK、OpenBoot、JLE は、サン・マイクロシステムズ株式会社の登録商標です。

ATOK は、株式会社ジャストシステムの登録商標です。ATOK8 は、株式会社ジャストシステムの著作物であり、ATOK8 にかかる著作権その他の権利は、すべて株式会社ジャストシステムに帰属します。ATOK Server/ATOK12 は、株式会社ジャストシステムの著作物であり、ATOK Server/ATOK12 にかかる著作権その他の権利は、株式会社ジャストシステムおよび各権利者に帰属します。

本書で参照されている製品やサービスに関しては、該当する会社または組織に直接お問い合わせください。

OPEN LOOK および Sun™ Graphical User Interface は、米国 Sun Microsystems 社が自社のユーザーおよびライセンス実施権者向けに開発しました。米国 Sun Microsystems 社は、コンピュータ産業用のビジュアルまたはグラフィカル・ユーザーインターフェースの概念の研究開発における米国 Xerox 社の先駆者としての成果を認めるものです。米国 Sun Microsystems 社は米国 Xerox 社から Xerox Graphical User Interface の非独占的ライセンスを取得しており、このライセンスは米国 Sun Microsystems 社のライセンス実施権者にも適用されます。

U.S. Government Rights—Commercial use. Government users are subject to the Sun Microsystems, Inc. standard license agreement and applicable provisions of the FAR and its supplements.

本書は、「現状のまま」をベースとして提供され、商品性、特定目的への適合性または第三者の権利の非侵害の黙示の保証を含みそれに限定されない、明示的であるか黙示的であるかを問わない、なんらの保証も行われぬものとします。

本書には、技術的な誤りまたは誤植の可能性があります。また、本書に記載された情報には、定期的に変更が行われ、かかる変更は本書の最新版に反映されます。さらに、米国サンまたは日本サンは、本書に記載された製品またはプログラムを、予告なく改良または変更することがあります。

本製品が、外国為替および外国貿易管理法(外為法)に定められる戦略物資等(貨物または役務)に該当する場合、本製品を輸出または日本国外へ持ち出す際には、サン・マイクロシステムズ株式会社の事前の書面による承諾を得ることのほか、外為法および関連法規に基づく輸出手続き、また場合によっては、米国商務省または米国所轄官庁の許可を得ることが必要です。

原典: Sun StorageTek Availability Suite 4.0 Software Installation and Configuration Guide
Part No: 819-6147-10
Revision A



目次

はじめに vii

1. Availability Suite ソフトウェアのアップグレード、インストール、およびアンインストール 1

Availability Suite ソフトウェアのアップグレード 2

- ▼ AVS 3.2 からアップグレードする 2

Availability Suite ソフトウェアのインストール 3

- ▼ AVS 4.0 をインストールする 3

Availability Suite ソフトウェアのアンインストール 4

- ▼ AVS 3.2 をアンインストールする 4

2. 初期構成手順 5

初期構成手順の概要 6

システムファイルの構成 7

- ▼ /etc/hosts ファイルを編集する 7

IP スタック (IPv4 および IPv6) の構成 8

- ▼ IPv6 アドレスを設定する 8

- ▼ /etc/services ファイルを編集する 11

- ▼ /etc/nsswitch.conf ファイルを編集する 12

設定の変更 13

- ▼ ビットマップの操作モードの設定 13

ボリュームセット数の上限の引き上げ	14
ストレージボリューム数の上限の引き上げ	14
dscfsgadm を使用した構成データベースの初期化およびサービスの開始	15
dscfsgadm ユーティリティー	15
使用法	15
オプション	15
構成データベースの初期化とサービスの開始	16
サービスの使用可能または使用不可への切り替え	16
ビットマップボリュームの使用	18
ビットマップボリュームの推奨場所	18
ビットマップボリュームのサイズの要件	18
ボリュームセットファイルの使用	20
構成情報のバックアップ	22
▼ 構成情報をバックアップする	22
3. 遠隔ミラーソフトウェアの構成	23
複製	24
同期複製	24
非同期複製	25
整合グループ	26
遠隔複製の計画	27
業務上の要件	27
アプリケーションの書き込み負荷	27
ネットワークの特性	28
非同期キューの設定	28
ディスクまたはメモリーのキュー	28
キューのサイズ	29
▼ キューのサイズを変更する	30
▼ 現在のキューのサイズ、長さ、および hwm を表示する	30

データベースの非同期キューの適切なサイズの設定	33
非同期キューのフラッシュスレッドの構成	34
ネットワークの調整	36
TCP バッファのサイズ	36
TCP バッファサイズの表示および調整	37
TCP バッファおよび値を表示するためのネットワークの調整	37
▼ すべての TCP バッファを表示する	37
▼ バッファ名ごとに設定を表示する	37
▼ ソケットのバッファサイズを表示する	38
▼ 起動スクリプトのバッファサイズを設定して確認する	38
遠隔ミラーによる TCP/IP ポートの使用	39
デフォルトの TCP 待機ポート	39
遠隔ミラーとファイアウォールの使用	40
遠隔ミラーソフトウェアとポイントインタイムコピーソフトウェアの併用	40
遠隔複製の構成	41
A. 用語集	43
索引	47

はじめに

『Sun StorageTek Availability Suite 4.0 ソフトウェアインストールおよび構成マニュアル』では、このソフトウェアを効率的にインストール、設定、および使用するための情報を記載します。

マニュアルの構成

このマニュアルは、次の章で構成されています。

- 第1章では、Availability Suite ソフトウェアのアップグレード、インストール、およびアンインストールに関する情報について説明します。
- 第2章では、Sun StorageTek™ Availability Suite ソフトウェアをはじめて使用する際に、あらかじめ必要な初期構成手順について説明します。
- 第3章では、遠隔ミラーソフトウェアの構成に関する問題について説明します。
- 用語集では、このマニュアルで使用される用語を定義します。

UNIX コマンド

このマニュアルには、システムの停止、システムの起動、およびデバイスの構成などに使用する基本的な UNIX® コマンドと操作手順に関する説明は含まれていない可能性があります。これらについては、以下を参照してください。

- 使用しているシステムに付属のソフトウェアマニュアル
- 下記にある Solaris™ オペレーティングシステム (Solaris OS) のマニュアル

<http://docs.sun.com>

シェルプロンプトについて

シェル	プロンプト
UNIX の C シェル	<i>machine_name%</i>
UNIX の Bourne シェルと Korn シェル	\$
スーパーユーザー (シェルの種類を問わない)	#

書体と記号について

書体または記号*	意味	例
AaBbCc123	コマンド名、ファイル名、ディレクトリ名、画面上のコンピュータ出力、コード例。	.login ファイルを編集します。 ls -a を実行します。 % You have mail.
AaBbCc123	ユーザーが入力する文字を、画面上のコンピュータ出力と区別して表します。	% su Password:
<i>AaBbCc123</i>	コマンド行の可変部分。実際の名前や値と置き換えてください。	rm <i>filename</i> と入力します。
『 』	参照する書名を示します。	『Solaris ユーザーマニュアル』
「 」	参照する章、節、または、強調する語を示します。	第 6 章「データの管理」を参照。 この操作ができるのは「スーパーユーザー」だけです。
\	枠で囲まれたコード例で、テキストがページ行幅を超える場合に、継続を示します。	% grep `^#define` \ XV_VERSION_STRING '

* 使用しているブラウザにより、これらの設定と異なって表示される場合があります。

関連マニュアル

用途	タイトル	Part No.
マニュアルページ	sndradm iiadm dsstat kstat svadm dscfgadm	該当なし
システム管理	『Sun StorageTek Availability Suite 4.0 ポイントインタイムコピーソフトウェア 管理マニュアル』	819-6368
	『Sun StorageTek Availability Suite 4.0 遠隔ミラーソフトウェア管理マニ ュアル』	819-6363
統合	『Sun Cluster および Sun StorageTek Availability Suite 4.0 ソフトウェア統合 マニュアル』	819-6373
障害追跡	『Sun StorageTek Availability Suite 4.0 ソフトウェア障害追跡マニュアル』	819-6378
最新情報	『Sun StorageTek Availability Suite 4.0 ソフトウェアご使用にあたって』	819-6383

Sun のオンラインマニュアル

ローカライズ版を含む Sun の各種マニュアルは、次の URL から表示、印刷、または購入できます。

<http://www.sun.com/documentation>

Sun 以外の Web サイト

このマニュアルで紹介する Sun 以外の Web サイトが使用可能かどうかについては、Sun は責任を負いません。このようなサイトやリソース上、またはこれらを経由して利用できるコンテンツ、広告、製品、またはその他の資料についても、Sun は保証しておらず、法的責任を負いません。また、このようなサイトやリソース上、またはこれらを経由して利用できるコンテンツ、商品、サービスの使用や、それらへの依存に関連して発生した実際の損害や損失、またはその申し立てについても、Sun は一切の責任を負いません。

Sun の技術サポート

このマニュアルに記載されていない技術的な問い合わせについては、次の URL にアクセスしてください。

<http://www.sun.com/service/contacting>

コメントをお寄せください

マニュアルの品質改善のため、お客様からのご意見およびご要望をお待ちしております。コメントは下記よりお送りください。

<http://www.sun.com/hwdocs/feedback>

ご意見をお寄せいただく際には、下記のタイトルと Part No. を記載してください。

『Sun StorageTek Availability Suite 4.0 ソフトウェアインストールおよび構成マニュアル』、Part No. 819-6358-10

第1章

Availability Suite ソフトウェアの アップグレード、インストール、お よびアンインストール

この章では、Solaris 10 オペレーティング環境およびそれ以降の更新リリースでの、Availability Suite (AVS) ソフトウェアのアップグレード、インストール、およびアンインストールに関する情報について説明します。

注 – この章の説明は、Solaris オペレーティング環境 (OE) に含まれていないバージョンの Availability Suite にのみ適用されます。

この章では、次の事項について説明します。

- 2 ページの「Availability Suite ソフトウェアのアップグレード」
- 3 ページの「Availability Suite ソフトウェアのインストール」
- 4 ページの「Availability Suite ソフトウェアのアンインストール」

Availability Suite ソフトウェアのアップグレード

AVS 4.0 へのアップグレードは、Solaris 8 または 9 で動作する AVS 3.2 からのみサポートされています。AVS 3.2 は Solaris 10 では動作しませんが、AVS 4.0 は Solaris 10 でのみ動作します。このため、システムを AVS 3.2 から AVS 4.0 へアップグレードする場合は、最初に Solaris オペレーティング環境を Solaris 10 へアップグレードする必要があります。

Solaris OE を適切に Solaris 10 へアップグレードしてから新しい AVS パッケージを Solaris 10 にインストールすると、AVS 3.2 の制御下にあったボリュームがすべて、自動的に使用可能になります。

▼ AVS 3.2 からアップグレードする

AVS 3.2 からアップグレードするには、次の手順に従ってください。

1. 念のため、`dscfg` データベースに格納されている情報を、遠隔地にあるファイルに保存します。

```
# dscfg -l > remote-node:/backup/database-file
```

このバックアップファイルには、AVS 3.2 の制御下にあるボリュームのリストが含まれています。これを使用すると、次の手順 3 に示す Solaris OE のアップグレードが失敗した場合に、AVS 3.2 セットを手動で再構築できます。たとえば、Solaris のアップグレード後にコントローラ番号が変更されると、`raw` スライスを使用する AVS セットを、更新後のコントローラ番号で再構成する必要が生じる場合があります。

2. 4 ページの「Availability Suite ソフトウェアのアンインストール」の手順に従って、AVS 3.2 パッケージを削除します。
3. オペレーティング環境を Solaris 10 OE 以降にアップグレードします。
Solaris 8 が動作しているシステムでは、中間の Solaris 9 へアップグレードすることなく、直接 Solaris 10 にアップグレードできます。システムに Solaris 10 OE 以降をクリーンインストールする場合は、アップグレードにならないことに注意してください。
4. 3 ページの「Availability Suite ソフトウェアのインストール」の手順に従って、新しい AVS パッケージをインストールします。

Availability Suite ソフトウェアのインストール

この節では、AVS 4.0 ソフトウェアのインストールに関する情報について説明します。

▼ AVS 4.0 をインストールする

AVS 4.0 をインストールするには、次の手順に従ってください。

1. Sun Cluster OE での AVS の実行を計画している場合は、AVS をインストールする前に Sun Cluster OE をインストールすることをお勧めします。すでに AVS がインストールされているシステムに Sun Cluster OE をインストールする場合は、AVS をアンインストールする必要はありません。
2. Sun Cluster OE のインストール後、`dscfgadm` を実行して Sun Cluster 固有の構成情報の場所を選択します。
3. `pkgadd(1M)` を使用して、新しい AVS パッケージを次の順序でインストールします。

```
SUNWscmr
SUNWscmu
SUNWspsvr
SUNWspsvu
SUNWiir
SUNWiiu
SUNWrdr
SUNWrdcu
```
4. はじめて AVS を使用する場合は、まず第 2 章、5 ページの「初期構成手順」に記載されている手順を実行します。

注 – Availability Suite ソフトウェアをインストールすると、ルートディレクトリに `reconfiguration` ファイルが作成されます。ただし、Availability Suite ソフトウェアを使用するために Solaris を再起動 (`boot -r`) する必要はありません。

Availability Suite ソフトウェアのアンインストール

この節では、AVS 3.2 ソフトウェアのアンインストールに関する情報について説明します。

▼ AVS 3.2 をアンインストールする

AVS 3.2 をアンインストールするには、次の手順に従ってください。

1. AVS 対応ボリュームに対する、すべてのアプリケーションからの書き込みを休止します。
2. `pkgrm(1M)` を使用して、AVS パッケージを次の順序でアンインストールします。

```
SUNWrdcu  
SUNWrdcr  
SUNWiiu  
SUNWiir  
SUNWspsvu  
SUNWspsvr  
SUNWscmu  
SUNWscmr
```

第2章

初期構成手順

Sun StorageTek Availability Suite ソフトウェアをインストールしてから、はじめて使用する際には、まずポイントインタイムコピーソフトウェアおよび遠隔ミラーソフトウェアで使用する特定のファイルを構成する必要があります。この章では、必要な初期構成手順について説明します。

- 6 ページの「初期構成手順の概要」
- 7 ページの「システムファイルの構成」
- 13 ページの「設定の変更」
- 18 ページの「ビットマップボリュームの使用」

この章では、使用する情報に関する次の事項についても説明します。

- 20 ページの「ボリュームセットファイルの使用」
- 15 ページの「dscfgadm を使用した構成データベースの初期化およびサービスの開始」
- 22 ページの「構成情報のバックアップ」

初期構成手順の概要

表 2-1 に、必要な初期構成作業を示します。

表 2-1 Availability Suite ソフトウェアの初期構成の概要

作業	手順
1. 次のファイルを構成します。 <ul style="list-style-type: none">• /etc/hosts• IP スタック (IPv4 および IPv6)• (省略可能) /etc/services• /etc/nsswitch.conf• (省略可能) /usr/kernel/drv/rdc.conf	7 ページの「システムファイルの構成」
2. (省略可能) ソフトウェアで使用できるボリュームのデフォルト数を調整します。	13 ページの「設定の変更」
3. (省略可能) 非同期キューを調整します。	『Sun StorageTek Availability Suite 4.0 遠隔ミラーソフトウェア管理マニュアル』
4. ビットマップボリュームを選択します。	18 ページの「ビットマップボリュームの使用」
5. (省略可能) 遠隔ミラーボリュームの任意の構成ファイルを設定します。	20 ページの「ボリュームセットファイルの使用」

システムファイルの構成

この節では、Sun StorageTek 遠隔ミラーソフトウェアを正しく動作させるために、次のシステムファイルを編集および確認する方法について説明します。

- 7 ページの「/etc/hosts ファイルを編集する」
- 8 ページの「IPv6 アドレスを設定する」
- 11 ページの「/etc/services ファイルを編集する」
- 12 ページの「/etc/nsswitch.conf ファイルを編集する」
- 22 ページの「構成情報をバックアップする」

▼ /etc/hosts ファイルを編集する

この手順によって、遠隔ミラーソフトウェアを実行するマシンで、/etc/hosts ファイル内のホスト名が読み取られ認識されるようになります。

- 遠隔ミラーソフトウェアの使用を計画しているすべてのマシンの名前および IP アドレスを /etc/hosts ファイルに追加します。

遠隔ミラーソフトウェアをインストールして実行する各マシンで、このファイルを編集してください。

IP スタック (IPv4 および IPv6) の構成

インターネットプロトコルバージョン 6 (IPv6) トランスポートプロトコルを複製に使用する場合は、ホストで、遠隔ミラーソフトウェアを使用するインタフェースに IPv4 と IPv6 のスタックを並行して構成してください。IPv6 プロトコルを使用すると、アドレス指定能力が向上します。Solaris 10 OS を使用する場合は、『Solaris のシステム管理 (IP サービス)』を参照して、IPv6 に関する情報を確認してください。

IPv6 プロトコルを使用するには、IPv4 と IPv6 のインタフェースを同じ名前で定義してください。一次ホストと二次ホストは、両方のマシンで同じトランスポートプロトコルが使用されるように定義する必要があります。

▼ IPv6 アドレスを設定する

ここでは、ネットワークインタフェースで IPv6 アドレスを使用するための設定方法の例について説明します。この手順を使用して、遠隔ミラーホストの接続をテストしてください。以降の例では、次の構成情報を使用することを前提としています。

ネットワークインタフェース	hme1
一次ホストのインタフェース名	sndrpri
二次ホストのインタフェース名	sndrsec

1. テキストエディタを使用して、一次ホストと二次ホストの両方で `/etc/hostname6.hme1` ファイルを作成します。一次ホストでは、このファイルにインタフェース名 `sndrpri` を追加します。二次ホストでは、このファイルにインタフェース名 `sndrsec` を追加します。ファイルを保存して閉じます。

```
primary-host# more /etc/hostname6.hme1
sndrpri
secondary-host# more /etc/hostname6.hme1
sndrsec
```

2. 両方のマシンを停止してから再起動し、IPv6 を起動します。

```
# /etc/shutdown -y -i 6 -g 0
```

3. 両方のマシンを再起動したあと、hme1 インタフェースアドレスの IPv6 の `inet` アドレスを取得します。

4. 次の例では、アドレスは fe80::a00:20ff:febd:c33f/128 です。

```
# ifconfig -a
lo0: flags=1000849<UP,LOOPBACK,RUNNING,MULTICAST,IPv4> mtu 8232 index 2
    inet 127.0.0.1 netmask ff000000
hme0: flags=1000843<UP,BROADCAST,RUNNING,MULTICAST,IPv4> mtu 1500 index 3
    inet 192.9.200.125 netmask ffffffff broadcast 192.9.200.255
    ether 8:0:20:ae:85:fa
lo0: flags=2000849<UP,LOOPBACK,RUNNING,MULTICAST,IPv6> mtu 8252 index 2
    inet6 ::1/128
hme0: flags=2000841<UP,RUNNING,MULTICAST,IPv6> mtu 1500 index 3
    ether 8:0:20:ae:85:fa
    inet6 fe80::a00:20ff:feae:85fa/10
hme1: flags=2000841<UP,RUNNING,MULTICAST,IPv6> mtu 1500 index 4
    ether 8:0:20:bd:c3:3f
    inet6 fe80::a00:20ff:febd:c33f/128
```

5. /etc/inet/ipnodes ファイルを編集し、手順 4 で取得した inet アドレスを追加します。一次ホストのアドレスは sndrpri に、二次ホストのアドレスは sndrsec に割り当てます。アドレスの /128 の部分は入力しません。

注 - 遠隔ミラーソフトウェアを実行している各システムの /etc/inet/ipnodes ファイルに、各システムの IPv6 の inet アドレスおよび名前が指定されていることを確認してください。

6. ファイルを保存して閉じたあと、ファイルの内容を確認します。

次の例で、sndrsec は二次ホストのインタフェース名です。

```
primary-host# more /etc/inet/ipnodes
#
# Internet host table
#
::1                localhost
127.0.0.1          localhost
fe80::a00:20ff:febd:c33f          sndrpri
fe80::a00:20ff:fee1:195e          sndrsec
```

7. /etc/nsswitch.conf ファイルを編集して、ipnodes: に files が指定されていることを確認します。

ファイルで次の文字列を探し、ipnodes: の行がコメントになっていないことを確認してください。

```
# consult /etc "files" only if nis is down.
hosts: files nis [NOTFOUND=return] files
ipnodes: files
```

8. 遠隔ミラーソフトウェアを使用するすべてのマシンのホスト名および IPv6 の inet 一次アドレスを、各マシンの /etc/hosts ファイルに追加します。

遠隔ミラーソフトウェアをインストールして実行する各マシンで、このファイルを編集してください。

注 – この手順 (詳細は、7 ページの「/etc/hosts ファイルを編集する」を参照) を完了しないと、遠隔ミラーソフトウェアを使用可能にしたときに、次のエラーメッセージが表示されます。

```
sndradm: Error: neither sndrpri nor sndrsec is local
```

9. あるシステムからほかのシステムへの ping を行なって、これらのシステムが IPv6 プロトコルを使用していることを確認します。

一次ホストから ping を行うには、次のように入力します。

```
# ping -s sndrsec
PING sndrsec: 56 data bytes
64 bytes from sndrsec (fe80::a00:20ff:fee1:195e): icmp_seq=0. time=0. ms
64 bytes from sndrsec (fe80::a00:20ff:fee1:195e): icmp_seq=1. time=0. ms
64 bytes from sndrsec (fe80::a00:20ff:fee1:195e): icmp_seq=2. time=0. ms
```

二次ホストから ping を行うには、次のように入力します。

```
# ping -s sndrpri
PING sndrpri: 56 data bytes
64 bytes from sndrpri (fe80::a00:20ff:febd:c33f): icmp_seq=0. time=0. ms
64 bytes from sndrpri (fe80::a00:20ff:febd:c33f): icmp_seq=1. time=0. ms
64 bytes from sndrpri (fe80::a00:20ff:febd:c33f): icmp_seq=2. time=0. ms
```

10. netstat(1M) コマンドを使用して、適切な IPv6 アドレスおよび IPv6 名がインタフェースに使用されていることを確認します。

sndrpri ホストおよび sndrsec ホストの両方で、このコマンドを実行します。次に例を示します。

```
# netstat -in
Name Mtu Net/Dest Address Ipkts Ierrs Opkts Oerrs Collis Queue
lo0 8232 127.0.0.0 127.0.0.1 3844 0 3844 0 0 0
hme0 1500 192.0.0.0 192.9.200.225 22007 0 1054 0
0 0

Name Mtu Net/Dest Address Ipkts Ierrs Opkts Oerrs Collis Queue
lo0 8252 ::1 ::1 3844 0
3844 0 0
hme1 1500 fe80::a00:20ff:febd:c33f fe80::a00:20ff:febd:c33f 43 0 65 0 0
```

```
# netstat -i
Name Mtu Net/Dest Address Ipkts Ierrs Opkts Oerrs Collis Queue
lo0 8232 loopback localhost 3844 0 3844 0 0 0
hme0 1500 arpanet rick1 22038 0
1067 0 0 0

Name Mtu Net/Dest Address Ipkts Ierrs
Opkts Oerrs Collis
lo0 8252 localhost localhost 3844 0 3844 0 0
hme1 1500 sndrpri sndrpri 43 0 65
0 0
```

▼ /etc/services ファイルを編集する

ポート 121 は、遠隔ミラーの rdc デーモンが使用するデフォルトのポートです。

```
# cat /etc/services
...
rdc 121/tcp # SNDR server daemon
...
```

ポート番号を変更する場合は、構成セット内のすべての遠隔ミラーホスト（つまり、一次ホストおよび二次ホスト、また 1 対多、多対 1、マルチホップ構成のすべてのホスト）で、ポート番号を変更する必要があります。

1. 遠隔ミラーソフトウェアを実行している各マシンで /etc/services ファイルを編集します。
2. すべてのホストを停止して再起動し、新しいポート番号を有効にします。

▼ /etc/nsswitch.conf ファイルを編集する

ファイルに `hosts:` および `services:` エントリが含まれている場合は、`files` が `nis`、`nisplus`、`ldap`、`dns` またはマシンで使用されるその他のサービスより前に指定されていることを確認してください。たとえば、NIS ネームサービスを使用しているシステムでは、ファイルに次の行を指定します。

```
hosts: files nis
services: files nis
```

- ホストおよびサービスのエントリが正しくない場合は、ファイルを編集して保存します。

IPv6 プロトコルを使用している場合、このファイルの変更については、8 ページの「IP スタック (IPv4 および IPv6) の構成」を参照してください。

設定の変更

このあとの節では、遠隔ミラーソフトウェアの設定を変更する方法について説明します。

- 13 ページの「ビットマップの操作モードの設定」
- 14 ページの「ボリュームセット数の上限の引き上げ」
- 14 ページの「ストレージボリューム数の上限の引き上げ」

注 - この節の手順に従ってファイルを編集したあとは、`shutdown` コマンドでサーバーを停止してから再起動して、変更内容を有効にしてください。64 を超えるボリュームセットを使用するように `rdc.conf` ファイルを編集する場合は、大きなスワップ空間などの十分なシステム資源があることを確認してください。

ビットマップの操作モードの設定

ディスクに保存されたビットマップは、`/usr/kernel/drv/rdc.conf` 内の `rdc_bitmap_mode` の設定によっては、システムで障害が発生しても保持することができます。デフォルトの設定は、1 (書き込み処理ごとにビットマップの書き込みを強制) です。

- `rdc.conf` ファイルを編集し、次に示すセクションを探します。ビットマップモードの値を編集し、ファイルを保存してから閉じます。

```
#
# rdc_bitmap_mode
# - Sets the mode of the RDC bitmap operation, acceptable values are:
#   0 - autodetect bitmap mode depending on the state of SDBC (default).
#   1 - force bitmap writes for every write operation, so an update resync
#       can be performed after a crash or reboot.
#   2 - only write the bitmap on shutdown, so a full resync is
#       required after a crash, but an update resync is required after
#       a reboot.
#
rdc_bitmap_mode=1;
```

ボリュームセット数の上限の引き上げ

構成されるボリュームセットの数は、デフォルトで 64 です。64 を超えるボリュームセットを構成する場合は、遠隔ミラーソフトウェアが動作している各マシンで /usr/kernel/drv/rdc.conf ファイルの rdc_max_sets フィールドを編集してください。

- rdc.conf ファイルを編集し、次に示すセクションを探します。ボリュームセットの値を編集し、ファイルを保存してから閉じます。

たとえば、128 セットを使用するには、次のようにファイルを変更します。

```
#
# rdc_max_sets
# - Configure the maximum number of RDC sets that can be enabled on
# this host. The actual maximum number of sets that can be enabled
# will be the minimum of this value and nsc_max_devices (see
# nsctl.conf) at the time the rdc kernel module is loaded.
#
rdc_max_sets=128;
```

ストレージボリューム数の上限の引き上げ

Availability Suite ソフトウェアでは、ストレージボリューム数がデフォルトで 4096 に制限されています。ストレージボリュームドライバデバイス (ボリューム) 数のデフォルトの値は、nsctl.conf ファイルの nsc_max_devices で設定します。

このボリューム数は、遠隔ミラーソフトウェアとポイントインタイムコピーソフトウェアとの間で分割して使用されます。たとえば、ポイントインタイムコピーソフトウェアだけを使用する場合は、それぞれマスターボリューム、シャドウボリューム、およびビットマップボリュームで構成されるボリュームセットを 1365 セットまで使用できます。遠隔ミラーソフトウェアパッケージとポイントインタイムコピーソフトウェアパッケージを一緒に使用すると、ボリュームセット数は 2 つのパッケージ間で分割されます。

環境によっては、この上限の変更が役に立ちます。使用できるメモリー量にゆとりのあるサイトでは、必要であれば、上限を引き上げてより多数のストレージボリュームを使用できます。使用できるメモリー量に制約のあるサイトでは、上限を引き下げてシステム資源を解放することができます。

- nsctl.conf ファイルを編集し、nsc_max_devices フィールドを探します。値を編集し、ファイルを保存して閉じます。

dscfgadm を使用した構成データベースの初期化およびサービスの開始

Availability Suite ソフトウェアの起動および停止は、サービス管理機能 (Service Management Facility、SMF) サービスを介して実装されます。このサービスは、dscfgadm ユーティリティーを使用して管理できます。

```
# svcs | grep nws_
online      Mar_14    svc:/system/nws_scm:default
online      Mar_14    svc:/system/nws_sv:default
online      Mar_14    svc:/system/nws_ii:default
online      Mar_14    svc:/system/nws_rdc:default
online      Mar_14    svc:/system/nws_rdcsyncd:default
```

dscfgadm ユーティリティー

dscfgadm は、構成情報の場所を設定する機能や、Availability Suite サービスの使用可能および使用不可を切り替える機能を提供して、Availability Suite の構成サービスを制御します。

```
# dscfgadm [-x]
```

使用法

```
dscfgadm [-x ]
dscfgadm [-x ] -i
dscfgadm [-x ] -e [-r] [-p]
dscfgadm [-x ] -d [-r]
```

オプション

-i を指定すると、Availability Suite サービスに関する情報が表示されます。

-e を指定すると、デフォルトではすべての Availability Suite SMF サービスが使用可能になります。

-d を指定すると、デフォルトではすべての Availability Suite SMF サービスが使用不可になります。

-r を指定すると、遠隔ミラーソフトウェアが使用可能または使用不可になります。

-p を指定すると、ポイントインタイムコピーソフトウェアが使用可能になります。

-x を指定すると、詳細なデバッグ情報が表示されます。

構成データベースの初期化とサービスの開始

デフォルトでは、Availability Suite サービスは開始されず、Availability Suite 構成データベースもシステムに存在しません。dscfadm は、オプションを指定せずに、または -x のみを指定して実行すると、対話型モードで実行されます。このモードでは、Availability Suite ソフトウェアに必要なローカルの構成データベースを初期化することができます。また、初期化中に Availability Suite サービスを開始するかどうかを選択できます。

データベースの初期化中に Availability Suite SMF サービスの開始を選択しない場合は、あとで dscfadm -e コマンドを使用して SMF サービスを開始できます。

```
# dscfadm -e
```

注 - Availability Suite ソフトウェアは、Availability Suite サービスが開始されるまで使用できません。

サービスの使用可能または使用不可への切り替え

あとでサービスを使用可能または使用不可にするには、-e および -d オプションを指定して dscfadm を実行してください。

```
# dscfadm -e
```

```
# dscfadm -d
```

デフォルトの動作はすべてのサービスに適用されますが、-r および -p オプションを使用すると、遠隔ミラーサービスまたはポイントインタイムコピーサービスを使用不可にするか使用可能にするかを別々に指定できます。

注 - サービスを使用不可にする前に、そのサービスが使用しているボリュームに対するすべてのアプリケーションからの書き込みを、確実に休止してください。

注 - 設定は、システムを起動しても維持されます。

ビットマップボリュームの使用

ポイントインタイムコピーソフトウェアおよび遠隔ミラーソフトウェアはどちらも、**raw** ボリュームを使用してビットマップを格納します。ビットマップファイルはサポートされません。

ビットマップボリュームの推奨場所

ビットマップの **raw** ボリュームは、関連するマスターおよびシャドウボリューム (ポイントインタイムコピーソフトウェアの場合)、または複製されたボリューム (遠隔ミラーソフトウェアの場合) が格納されたディスクとは別のディスクに格納してください。このようなビットマップボリュームには、ミラー化されたパーティションなどの RAID を構成して、ミラー化されたメンバーがマスターおよびシャドウボリュームや、複製されたボリュームと同じディスクに格納されないようにします。

クラスタ環境でポイントインタイムコピーソフトウェアを使用する場合、ビットマップボリュームが、関連するマスターまたはシャドウデータボリュームと同じディスクグループまたはクラスタリソースグループの一部になるように設定してください。

ビットマップボリュームのサイズの要件

ビットマップボリュームのサイズは、マスターボリュームのサイズおよび作成されるボリュームセットの種類 (独立、依存、または小型依存) に基づいて、次のように計算します。

- 独立または依存シャドウボリュームセットに必要なサイズ:

総 G バイト数から端数を切り上げたマスターボリュームのサイズの 1G バイトあたり 8K バイト、さらにオーバーヘッド用に 24K バイト

たとえば、3G バイトのマスターボリュームのシャドウを作成するには、ビットマップのサイズは、 $(3 \times 8K \text{ バイト}) + 24K \text{ バイト} = 48K \text{ バイト}$ にしてください。50G バイトのマスターボリュームには、424K バイトのビットマップボリュームが必要です。

- 小型依存シャドウボリュームセットに必要なサイズ:

総 G バイト数から端数を切り上げたマスターボリュームのサイズの 1G バイトあたり 264K バイト、さらにオーバーヘッド用に 24K バイト

たとえば、3G バイトのマスターボリュームのシャドウを作成するには、ビットマップのサイズは、 $(3 \times 264K \text{ バイト}) + 24K \text{ バイト} = 816K \text{ バイト}$ にしてください。小型依存シャドウボリュームセットの 50G バイトのマスターボリュームには、13224K バイトのビットマップボリュームが必要です。

サイズが大きすぎるビットマップを含むシャドウボリュームセットを使用可能になると、領域が無駄になる可能性があります。シャドウボリュームセットは作成されず。サイズが小さすぎるビットマップを含むシャドウボリュームセットを使用可能にすると、そのコマンドは失敗してエラーメッセージが表示されます。

Availability Suite ソフトウェアは、ポイントインタイムコピーのシャドウボリュームセットまたは遠隔ミラーのボリュームセットに必要なビットマップサイズを計算するための `dsbitmap` ユーティリティを提供しています。

1. ポイントインタイムコピーのビットマップサイズを計算するには、次のコマンドを使用します。

```
dsbitmap -p data-volume [bitmap-volume]
```

2. 遠隔ミラーのビットマップサイズを計算するには、次のコマンドを使用します。

```
dsbitmap -r data-volume [bitmap-volume]
```

`dsbitmap` ユーティリティの詳細は、`dsbitmap (1SCM)` のマニュアルページを参照してください。

ボリュームセットファイルの使用

遠隔ミラーソフトウェアを使用可能にすると、ボリューム、一次ホストおよび二次ホスト、ビットマップ、動作モードなどの、ボリュームセットに関する情報を含む「ボリュームセットファイル」を任意で指定できます。ボリュームセットファイルを使用する場合は、`sndradm` の `-f volset-file` オプションを使用します。

各ボリュームセットに関する情報は、コマンド行から入力することもできますが、複数のボリュームセットを使用する場合は、この情報をファイルに入力すると便利です。また、ほかのボリュームセットに影響することなく、特定のボリュームセットだけを操作の対象にできるという利点もあります。入出力グループにボリュームセットを追加する場合とは異なり、ボリュームセットファイルには複製モードを混在させることができます。ボリュームセットファイルに指定するフィールドは、次のとおりです。

```
phost pdev pbitmap shost sdev sbitmap ip {sync|async} [g io-groupname] [C tag] [q qdev]
```

表 2-2 に、フィールドの詳細を示します。ボリュームセットファイルの書式の詳細は、`rdc.cf` マニュアルページを参照してください。

次に、ファイルエントリの例を示します。

```
atm10 /dev/vx/rdisk/oracle816/oratest /dev/vx/rdisk/oracle816/oratest_bm \  
atm20 /dev/vx/rdisk/oracle816/oratest /dev/vx/rdisk/oracle816/oratest_bm \  
ip sync g oragroup
```

表 2-2 ボリュームセットファイルのフィールド

フィールド	意味	説明
<i>phost</i>	一次ホスト (Primary host)	一次ボリュームが格納されているサーバー。
<i>pdev</i>	一次デバイス (Primary device)	一次ボリュームのパーティション。フルパス名で指定します (例: /dev/rdisk/c0t1d0s4)。
<i>pbitmap</i>	一次ビットマップ (Primary bitmap)	一次パーティションのビットマップが格納されるボリュームパーティション。フルパス名で指定します。
<i>shost</i>	二次ホスト (Secondary host)	二次ボリュームが格納されているサーバー。
<i>sdev</i>	二次デバイス (Secondary device)	二次ボリュームのパーティション。フルパス名で指定します。
<i>sbitmap</i>	二次ビットマップ (Secondary bitmap)	二次パーティションのビットマップが格納されるボリュームパーティション。フルパス名で指定します。
<i>ip</i>	ネットワーク転送プロトコル (Network transfer protocol)	<i>ip</i> を指定します。
<i>sync async</i>	動作モード (Operating mode)	<ul style="list-style-type: none"> • <i>sync</i> は、遠隔ボリュームが更新されたあとに、入出力処理の完了を確認するモードです。 • <i>async</i> は、遠隔ボリュームが更新される前に、一次ホストの入出力処理の完了を確認するモードです。
<i>g io-groupname</i>	入出力グループ名 (I/O group name)	文字 <i>g</i> に続けて入出力グループの名前を指定します。前述の例では、 <i>oragroup</i> が入出力グループ名です。
<i>C tag</i>	クラスタタグ (Cluster tag)	処理の対象をクラスタリソースグループに属する遠隔ミラーセットのみに制限するタグです。
<i>q qdev</i>	ディスクキュー ボリューム (Disk queue volume)	非同期セットまたはグループ用の、ディスクベースの入出力キューとして使用されるボリュームです。フルパス名で指定してください。たとえば、/dev/rdisk/c1t2d0s6 のように指定します。

構成情報のバックアップ

Sun StorageTek、VERITAS Volume Manager、および Solaris ボリュームマネージャーの構成情報は、定期的にバックアップしてください。ボリュームセットに関する変更を行うときは、`/usr/sbin/iiadm` コマンドを使用します。このコマンドの詳細は、『Sun StorageTek Availability Suite 4.0 ポイントインタイムコピーソフトウェア管理マニュアル』に記載されています。また、次の事項を検討してください。

- バックアップコマンドをシェルスクリプトに記述し、毎日の `cron(1M)` ジョブの一部としてこのスクリプトを実行する
- コマンドの出力を特定の場所に格納し、定期的にテープにバックアップする。

定期的に構成データベースを自動的にバックアップするために、`cron(1M)` ジョブの設定が必要になる場合があります。また、ボリュームの追加や削除などで構成を変更したときにも、構成情報をバックアップしてください。

▼ 構成情報をバックアップする

- 構成データベース (`/etc/dscfg_local`) を安全な場所にコピーします。

```
# cp /etc/dscfg_local /var/backups/dscfg_db
```

遠隔ミラーソフトウェアの構成

Sun StorageTek Availability Suite 遠隔ミラーソフトウェアは、Solaris 10 (Update 1 以降) のオペレーティングシステムにボリュームレベルの「複製」機能を提供します。遠隔ミラーソフトウェアを使用すると、物理的に離れた場所にある一次サイトと二次サイトとの間で、リアルタイムにディスクボリュームへの書き込み処理を複製できます。遠隔ミラーソフトウェアは、Sun のすべてのネットワークアダプタおよび TCP/IP をサポートするネットワークリンクで使用できます。

遠隔ミラーソフトウェアはボリュームベースであるため、ストレージに依存せず、Sun の製品か Sun 以外の製品かを問わず、raw ボリュームおよびすべてのボリュームマネージャーをサポートします。また、この製品は、Solaris OS が動作する単一のホストでデータを書き込むすべてのアプリケーションまたはデータベースをサポートします。Solaris OS が動作する複数のホストが共有ボリュームにデータを書き込むことを許可するように構成された、Oracle® 9iRAC, Oracle® Parallel Server などのデータベース、アプリケーション、またはファイルシステムはサポートされていません。

災害回復計画および業務継続計画の一環として遠隔ミラーソフトウェアを使用すると、重要なデータのコピーを遠隔サイトで最新の状態に保つことができます。また、遠隔ミラーソフトウェアを使用して、業務継続計画のリハーサルおよびテストを行うことができます。高可用性ソリューションでは、Sun StorageTek Availability Suite ソフトウェアを、Sun™ Cluster 3.x 環境内でフェイルオーバーを実行するように構成できます。

遠隔ミラーソフトウェアは、アプリケーションがデータボリュームにアクセスするとき動作して、データを継続的に遠隔サイトに複製するか、または変更をスコアボードに記録してあとで高速再同期を行えるようにします。

遠隔ミラーソフトウェアを使用すると、一次サイトから二次サイトへの再同期 (通常「順方向の同期」と呼ぶ) または二次サイトから一次サイトへの再同期 (通常「逆方向の同期」と呼ぶ) を手動で起動できます。

遠隔ミラーソフトウェアの複製および構成は、ボリュームセット単位で行われます。遠隔ミラーセットは、一次ボリューム、二次ボリューム、一次サイトおよび二次サイトの両方のビットマップボリューム (高速再同期用に変更を追跡してスコアボードに記録するためのボリューム) と、「非同期複製」モード用の任意の「非同期キュー」ボリューム (オプション) で構成されます。一次ボリュームおよび二次ボリューム

は、同じサイズにすることをお勧めします。dsbitmap ツールを使用すると、必要なビットマップボリュームのサイズを確認できます。遠隔ミラーセットの構成または dsbitmap ツールの詳細は、『Sun StorageTek Availability Suite 4.0 遠隔ミラーソフトウェア管理マニュアル』を参照してください。

複製

複製は、同期または非同期のいずれのモードで実行することもできます。同期モードでは、アプリケーションの書き込み処理に対する応答は、一次サイトおよび二次サイトの両方で書き込みが確定するまで通知されません。非同期モードでは、アプリケーションの書き込み処理に対する応答は、書き込みがローカルのストレージで確定し、その処理が非同期キューに入ったときに通知されます。この非同期キューによって、二次サイトへの書き込み処理が非同期で実行されます。

同期複製

同期処理のデータフローは、次のとおりです。

1. ビットマップボリュームでスコアボードのビットがセットされます。
2. ローカルの書き込み処理とネットワークの書き込み処理が並行して開始されます
3. 両方の書き込み処理が完了すると、スコアボードのビットがクリアされます (遅延クリア)。
4. 書き込み処理の完了がアプリケーションに通知されます。

同期複製の利点は、一次サイトおよび二次サイトが常に同期をとっていることです。この複製は、接続の待ち時間が短く、接続がアプリケーションの帯域幅の要件を満たしている場合にのみ使用できます。通常、この制約があるため、同期ソリューションの使用は構内または都市部に制限されます。

この場合、書き込み処理の平均サービス時間は、次のようになります。

ビットマップの書き込み時間 + MAX (ローカルデータの書き込み, ネットワーク上の往復 + 遠隔データの書き込み)

構内および都市部ではネットワーク上の往復にかかる時間はごくわずかであるため、サービス時間は遠隔ミラーソフトウェアがインストールされていない場合の約 2 倍になります。

書き込みに 5 ミリ秒かかるとすると、次のようになります。

5 ミリ秒 + MAX (5 ミリ秒, 1 ミリ秒 + 5 ミリ秒) = 11 ミリ秒

注 - 5 ミリ秒という値は、負荷の小さいシステムでの妥当な想定値です。より現実的な負荷のかかるシステムでは、未処理の操作がキューに蓄積されるため、値が大きくなります。

しかし、ネットワーク上の往復時間が約 50 ミリ秒 (長距離複製の標準値) になると、ネットワークの待ち時間は、次の例のように同期ソリューションにとって非現実的な値になります。

$$5 \text{ ミリ秒} + \text{MAX} (5 \text{ ミリ秒}, 50 \text{ ミリ秒} + 5 \text{ ミリ秒}) = 60 \text{ ミリ秒}$$

非同期複製

非同期複製は、遠隔の書き込み処理とアプリケーションの書き込み処理とを区別しません。このモードでは、ネットワークの書き込み処理が非同期キューに追加されるときに完了が通知されます。つまり、すべての書き込み処理が二次サイトに配信されるまで、二次サイトは一次サイトと同期をとりません。このモードのデータフローは、次のとおりです。

1. スコアボードのビットがセットされます。
2. ローカルの書き込み処理と非同期キューへの書き込み処理が並行して行われます。
3. 書き込み処理の完了がアプリケーションに通知されます。
4. フラッシュスレッドが非同期キューのエントリを読み取り、ネットワークの書き込みを実行します。
5. スコアボードのビットがクリアされます (遅延クリア)。

サービスに必要な時間は、次のとおりです。

$$\text{ビットマップの書き込み時間} + \text{MAX} (\text{ローカルの書き込み}, \text{非同期キューへのデータのエン트리})$$

書き込み処理に要するサービス時間を 5 ミリ秒とすると、非同期の書き込み処理の推定サービス時間は次のようになります。

$$5 \text{ ミリ秒} + \text{MAX} (5 \text{ ミリ秒}, 5 \text{ ミリ秒}) = 10 \text{ ミリ秒}$$

ボリュームまたは整合グループで、書き込み速度がネットワークへの排出速度を上回る状態が長く続くと、非同期キューはいっぱいになります。キューを適切なサイズに設定することが重要です。適切なボリュームサイズを算出する方法の詳細は後述します。

非同期のディスクキューがいっぱいになった場合の遠隔ミラーソフトウェアの動作は、次の2つのモードで管理されます。

■ ブロックモード

ブロックモードはデフォルトの設定です。このモードでは、遠隔ミラーソフトウェアは一時停止し、非同期のディスクキューが特定のポイントまで排出されるのを待ってから非同期キューに書き込みを追加します。これによって、アプリケーションの書き込み処理は影響を受けますが、書き込み順序は接続先でも保持されます。

■ 非ブロックモード

非ブロックモードは、メモリーベースのキューでは使用できません。このモードでは、遠隔ミラーソフトウェアは非同期のディスクキューがいっぱいになっても停止せず、記録モードになって書き込みをスコアボードに記録します。次の更新同期では、これらはビット0から順に読み込まれるため、書き込み順序が保持されません。このモードが設定されているときに、非同期のディスクキューがいっぱいになって書き込み順序が失われると、関連するボリュームまたは整合グループの一貫性が失われます。

注 - autosync デモンなどを使用して更新同期を開始する前には、二次サイトでポイントインタイムコピーをとっておくことを強くお勧めします。

整合グループ

同期モードでは、多数のボリュームにわたるアプリケーションの書き込み順序が保証されます。これは、書き込み順序の保持が必要であるときに、アプリケーションは書き込みが完了するまではほかの入出力処理を発行せずに待機し、遠隔ミラーソフトウェアは書き込み処理が一次サイトおよび二次サイトの両方で終了するまで完了を通知しないためです。

デフォルトの非同期モードでは、各ボリュームのキューは1つ以上の独立したスレッドによって排出されます。この処理はアプリケーションとは区別されているため、複数のボリュームへの書き込み処理間では書き込み順序は保持されません。

アプリケーションが書き込み順序の保持を必要とする場合のために、遠隔ミラーソフトウェアは整合グループ機能を提供します。各整合グループは1つのネットワークキューを使用します。並行して複数の書き込み処理を行うことができますが、シーケンス番号を使用して書き込み順序が保持されます。

遠隔複製の計画

遠隔複製を計画している場合は、業務上の要件、アプリケーションの書き込み負荷、および使用するネットワークの特性を考慮してください。

業務上の要件

業務データの複製を決定したら、最大の遅延を考慮してください。これは、二次サイトのデータが古い状態になっている時間を、どのくらい許容するかを意味します。これによって、複製のモードと、スナップショットのスケジューリングが決まります。また、複製の対象になるアプリケーションで、二次ボリュームへの書き込み処理が正しい順序で複製される必要があるかどうかを知ることが重要です。

アプリケーションの書き込み負荷

一次サイトと二次サイトとの間に必要なネットワーク接続のタイプを決定するには、平均的な書き込み負荷と最大の書き込み負荷を知っておく必要があります。構成を決定するために、次の情報を収集します。

- データ書き込み処理の平均速度およびサイズ

平均速度とは、アプリケーションが標準的な負荷で動作しているときのデータ書き込み処理の量です。アプリケーションの読み取り処理は、遠隔複製の準備および計画には重要ではありません。

- データ書き込み処理の最大速度およびサイズ

最大速度とは、測定期間中にアプリケーションによって書き込まれたデータの最大量です。

- 最大書き込み速度の持続時間および頻度

持続時間とは、最大書き込み速度が続く時間です。頻度とは、最大書き込み速度が発生する回数です。

アプリケーションの特性が不明な場合は、`iostat`、`sar`などのツールを使用して、アプリケーションが動作している間の書き込みトラフィックを計測してください。

ネットワークの特性

アプリケーションの書き込み負荷を把握したら、ネットワークリンクの要件を決定します。もっとも重要なネットワーク属性は、一次サイトと二次サイトとの間のネットワーク帯域幅およびネットワーク待ち時間です。Sun StorageTek Availability Suite ソフトウェアをインストールする前にネットワークリンクがすでに存在する場合は、ping などのツールを使用してサイト間の接続の特性を確認できます。

同期複製を使用するには、各書き込み処理がネットワーク上を往復する時間がアプリケーションの応答時間にあまり影響を及ぼさないように、ネットワークの待ち時間を短くする必要があります。また、アプリケーションの最大書き込み速度の時間中に生成された書き込みトラフィックを処理できるだけのネットワーク帯域幅が必要です。ネットワークがどの時点でも書き込みトラフィックを処理できるようになっていないと、アプリケーションの応答時間に影響を及ぼします。

非同期複製を使用するには、アプリケーションの平均書き込み速度の時間中に生成された書き込みトラフィックを処理できるだけのネットワークリンクの帯域幅が必要です。アプリケーションの最大書き込み速度の時間中に処理できなかった書き込み処理は、ローカルの非同期キューに書き込まれ、あとでネットワークトラフィックが空いたときに二次サイトに書き込まれます。非同期キューのサイズを適切に設定すると、ネットワークの制限を超えた書き込みトラフィックをバースト転送する間のアプリケーションの応答時間を最短に抑えられます。

このマニュアルの 28 ページの「非同期キューの設定」を参照してください。遠隔ミラーの非同期のオプションモード (ブロックまたは非ブロック) の選択によって、キューがいっぱいになったときの遠隔ミラーソフトウェアの動作が決まります。

非同期キューの設定

非同期複製を使用する場合は、この節で説明する構成設定を計画する必要があります。これらの設定は、遠隔ミラーのセットまたは整合グループ単位で行います。

ディスクまたはメモリーのキュー

version 3.2 では、遠隔ミラーソフトウェアは新たにディスクベースの非同期キューをサポートします。以前のバージョンからのアップグレードを容易にするため、メモリーベースのキューもまだサポートしています。ただし、新機能であるディスクベースのキューを使用すると、非常に大きくより効率的なキューを作成できます。大きいキューを使用すると、アプリケーションの応答時間に影響を与えることなく、より大きいバースト書き込みが可能になります。また、ディスクベースのキューは、メモリーベースのキューほどシステム資源に影響を与えません。

非同期キューには、アプリケーションの最大書き込み時の書き込みトラフィックのバーストを処理できるだけのサイズが必要です。キューを大きくすると、書き込み処理のバーストを長時間処理できますが、二次サイトと一次サイトの同期のずれがさらに大きくなる可能性もあります。最大書き込み速度および最大書き込み速度の持続時間、書き込みのサイズ、ネットワークリンクの特性によって、キューのサイズを決定する必要があります。33 ページの「ディスクベースの非同期キューの適切なサイズの設定」を参照してください。

選択したキューオプション (ブロックまたは非ブロック) によって、ディスクキューがいっぱいになったときの遠隔ミラーソフトウェアの動作が決まります。dsstat ツールを使用して、これまでに使用されたキューの最大量を示すハイウォーターマーク (hwm) などの非同期キューの統計情報を確認してください。非同期キューを遠隔ミラーセットまたは整合グループに追加するには、sndradm コマンドに -q オプションを指定して実行します (sndradm -q a)。

キューのサイズ

「ハイウォーターマーク」(hwm)を確認するには、dsstat(1SCM) コマンドを使用して非同期キューを監視します。アプリケーションの書き込みがキューのデータ処理を上回って、hwm がキューの合計サイズの 80 ~ 85% に頻繁に達した場合は、キューのサイズを大きくします。この原則は、ディスクベースおよびメモリーベースの両方のキューに適用できます。ただし、各タイプのキューのサイズを変更する手順は異なります。

メモリーベースのキュー

- キューの書き込み処理のデフォルトの最大数 (調整可能) は 4096 です。この値を変更するには、sndradm -W コマンドを使用します。
- 512 バイトのデータブロックのデフォルトの最大数 (デフォルトのキューのサイズ) (調整可能) は 16384 です。これは、約 8M バイトのデータに相当します。この値を変更するには、sndradm -F コマンドを使用します。

ディスクベースのキュー

ディスクキューの実際のサイズは、ディスクキューのボリュームのサイズになります。ディスクキューは、別のサイズのボリュームに置き換えることによってのみ、サイズを変更できます。たとえば、キューのサイズが 16384 ブロックであれば、hwm が 13000 ~ 14000 ブロックを超えていないことを確認します。この量を超えた場合は、次の手順でキューのサイズを変更します。

▼ キューのサイズを変更する

1. `sndradm -l` コマンドを使用して、ボリュームを記録モードに設定します。
2. キューのサイズを変更します。
 - メモリーベースの場合: `sndradm -F` コマンドを使用します。
 - ディスクベースの場合: `sndradm -q` コマンドを使用して、既存のディスクキューのボリュームを、よりサイズの大きいボリュームに置き換えます。
3. `sndradm -u` コマンドを使用して、更新同期を実行します。

▼ 現在のキューのサイズ、長さ、および hwm を表示する

1. キューのサイズを表示するには、次のように入力します。
 - メモリーベースの場合:

```
# sndradm -P
/dev/vx/rdsk/data_t3_dg/vol0 -> priv-2-
230:/dev/vx/rdsk/data_t3_dg/vol0
autosync: off, max q writes: 4096, max q fbas: 16384, async
threads: 8, mode: async, state: replicating
```

キューのサイズは、`max q fbas` にブロック単位で表示されます (この例では 16384 ブロック)。キューに格納できる項目の最大数は、`max q writes` に表示されます (この例では 4096)。この例は、キュー内の項目の平均サイズが 2K バイトであることを示しています。

- ディスクベースの場合:

```
# sndradm -P
/dev/vx/rdsk/data_t3_dg/vol0 -> priv-
230:/dev/vx/rdsk/data_t3_dg/vol0
autosync: off, max q writes: 4096, max q fbas: 16384, async
threads: 1, mode: async, blocking diskqueue:
/dev/vx/rdsk/data_t3_dg/dq_single, state: replicating
```

ディスクキューのボリューム (`/dev/vx/rdsk/data_t3_dg/dq_single`) が表示されます。ボリュームのサイズを確認すると、キューのサイズを特定できます。

2. 現在のキューの大きさおよび hwm を表示するには、次のように入力します。

```
# dsstat -m sndr -d q
name          q role   qi     qk    qhwi   qhwk
data_a5k_dg/vol10 D net    4      13     5     118
```

各情報には、次の意味があります。

- qi は、現在キューに入っている項目の数を示します
 - qk は、現在キューに入っているデータの合計サイズを K バイト単位で示します
 - qhwi は、これまでにキューに入った項目の最大数を示します
 - qhwk は、これまでにキューに入った最大のデータサイズを K バイト単位で示します
3. ストリーミングの概要およびディスクキューの情報を表示するには、次のように入力します。

```
# dsstat -m sndr -r bn -d sq 2
```

4. より詳細な情報を表示するには、ほかの表示オプションを指定して dsstat(1SCM) を実行します。

キューのサイズが適切な場合の dsstat の出力例

注 - この例では、コマンド出力の関連する部分だけを示します。dsstat コマンドは、実際にはより詳細な情報を表示します。

カーネル統計情報を表示する dsstat(1SCM) の次の出力例は、非同期キューに関する情報を示しています。この例では、キューのサイズは適切で、現在いっぱいにはなっていません。この例が示す設定および統計情報の意味は、次のとおりです。

```
# dsstat -m sndr -r n -d sq -s priv-2-230:/dev/vx/rdisk/data_t3_dg/vol167
name          q role   qi     qk    qhwi   qhwk   kps   tps   svt
data_t3_dg/vol167 D net    48    384   240   1944   10    1    54
```

各情報には、次の意味があります。

- qi エントリは、キューに入った書き込みトランザクションの合計数が 48 であることを示しています
- qk エントリは、キューに入ったデータが 384K バイトであることを示しています

- qhwi エントリは、キューに入った項目数の hwm が 240 で、現在はその値に達していないことを示しています
- qhwk エントリは、キューに入ったデータの hwm (K バイト単位) が 1944 で、現在はその値に達していないことを示しています

ディスクキューのボリュームサイズが 1G バイト (2097152 ディスクブロック) である場合、1944 ブロックの hwm は全体の 80% よりかなり小さいものです。このディスクキューのサイズは、書き込み負荷に対して正しく設定されています。

キューのサイズが不適切な場合の dsstat の出力例

カーネル統計情報を表示する dsstat(1SCM) の次の出力例は、サイズが不適切な非同期キューに関する情報を表示しています。

```
# sndradm -P
/dev/vx/rdsk/data_a5k_dg/vol0 -> priv-230:/dev/vx/rdsk/data_a5k_dg/vol0
autosync: off, max q writes: 4096, max q fbas: 16384, async threads: 2, mode:
async, state: replicating

# dsstat -m sndr -d sq
name          q role   qi    qk  qhwi  qhwk   kps   tps   svt
data_a5k_dg/vol0 M net   3609  8060 3613  8184    87    34    57
k/bitmap_dg/vol0  bmp    -     -    -     -     0     0     0
```

この例はデフォルトのキュー設定を示していますが、アプリケーションはキューの処理能力を超えるデータの書き込みを行なっています。qhwk の値である 8184K バイトを max q fbas の値の 16384 ブロック (8192K バイト) と比較すると、アプリケーションの書き込みが 512 バイトブロックの最大許容限界値に近づいていることがわかります。このあと数回で、入出力命令がキューに入らなくなる可能性があります。

この場合は、キューのサイズを大きくして問題を解決します。ただし、長期的に改善するには、より大きい帯域幅インタフェースを使用するなどの方法で、ネットワークリンクを向上させる必要があります。または、ボリュームのポイントインタイムコピーをとって、シャドウボリュームを複製することを検討します。『Sun StorageTek Availability Suite 4.0 ポイントインタイムコピーソフトウェア管理マニュアル』を参照してください。

要約すると次のようになります。

- 流入量が排出量以下である場合は、キューのサイズはデフォルトのサイズで十分です。
- 排出量が流入量より小さい場合は、キューのサイズを大きくすることによって問題を一時的に解決できる場合があります。ただし、書き込み処理が長時間継続すると、キューは最終的にいっぱいになります。

データベースの非同期キューの適切なサイズの設定

次の例について考えます。この例では、複製される入出力の負荷の特徴を確認するために、iostat を 1 時間ごとに実行しました。この例では、DS3 (45M バイト/秒) 接続を使用していることを想定しています。また、このアプリケーションは 1 つの整合グループを使用すること、つまり 1 つのキューが含まれることを想定しています。

24 時間にわたって統計データを収集し、それが対象のアプリケーションの日常的な値であると想定すると、平均的な書き込み速度および非同期キューの適切なサイズ、1 日の間で遠隔サイトが古い状態になっている時間、選択したネットワーク帯域幅がこのアプリケーションに対して適切であるかどうかを確認できます。

表 3-1 データベースのキューの適切なサイズの設定例

時間	書き込み K バイト/秒	書き込み数/ 秒	ネットワーク スループット	キューの増加	キューの サイズ
	A	B	C	$(A/1000 - C) \times 3600$	
午前 6 時	0	0	4MBps*		
午前 7 時	1000	400	4MBps		
午前 8 時	2000	1000	4MBps		
午前 9 時	2000	1000	4MBps		
午前 10 時	4000	1800	4MBps		
午前 11 時	5000	2400	4MBps	3.6G バイト	3.6G バイト
午前 0 時	1000	400	4MBps	-10G バイト	
午後 1 時	1200	600	4MBps		
午後 2 時	1000	500	4MBps		
午後 3 時	1200	400	4MBps		
午後 4 時	2000	600	4MBps		
午後 5 時	1000		4MBps		
午後 6 時	800		4MBps		
午後 7 時	800		4MBps		
午後 8 時	3200	1000	4MBps		
午後 9 時	8000	2500	4MBps	14G バイト	14G バイト
午後 10 時	8000	2500	4MBps	14G バイト	28G バイト
午後 11 時	1000	400	4MBps	-10	18

表 3-1 ディスクベースのキューの適切なサイズの設定例 (続き)

時間	書き込み K バイト/秒	書き込み数/ 秒	ネットワーク スループット	キューの増加	キューの サイズ
午前 0 時	0		4MBps	-14	4
午前 1 時	0		4MBps	-14	
午前 2 時	0		4MBps		
午前 3 時	0		4MBps		
午前 4 時	0		4MBps		
午前 5 時	0		4MBps		
平均 帯域幅	1.8MBps				

* M バイト/秒

表に記入してキューの増加およびサイズを計算すると、30G バイトのキューで十分対処できることがわかります。キューが大きくなると二次サイトの同期のずれも大きくなりますが、夕方にバッチ処理を行うことによって、通常の業務時間までにキューを空にして 2 つのサイトの同期をとることができます。

この動作テストによって、ネットワークの帯域幅がアプリケーションの書き込み負荷に対して適切であるかどうかを検証できます。

非同期キューのフラッシュスレッドの構成

Sun StorageTek Availability Suite ソフトウェアでは、非同期キューをフラッシュするスレッドの数を設定できます。この数を変更することによって、ネットワーク上のボリュームごとまたは整合グループごとに、同時に複数の入出力が可能になります。二次ノードの遠隔ミラーソフトウェアは、シーケンス番号を使用して入出力の書き込み順序を処理します。

複製の構成に対してもっとも効率のよいキューのフラッシュスレッド数を決定するには、多くの変数を考慮する必要があります。変数とは、セットまたは整合グループの数、使用できるシステム資源、ネットワークの特性、ファイルシステムの有無などです。セットまたは整合グループの数が少ない場合は、フラッシュスレッドの数を大きくするとより効率的です。この変数の値を少しずつ変えながら基本的なテストまたは試用を行い、使用する構成にもっとも効率的な設定を特定してください。

構成、ネットワークの特質、および遠隔ミラーソフトウェアの動作に関する知識は、ネットワークスレッドの数を適切に選択するためのガイドラインになります。遠隔ミラーソフトウェアは、転送機構として Solaris RPC を使用します。この RPC は同期式です。各ネットワークスレッドで、個々のスレッドが実現できる最大スループット

は、入出力サイズ/往復時間で求められます。入出力が主に 2K バイトで、往復時間が 60 ミリ秒の作業負荷について考えます。各ネットワークスレッドが処理できる量は、次のようになります。

$$2\text{K バイト}/0.060 \text{ 秒} = 33\text{K バイト}/\text{秒}$$

1 つのボリューム、または多数のボリュームを含む 1 つの整合グループでは、ネットワークスレッドがデフォルトの 2 に設定されていると、ネットワーク複製は 66K バイト/秒に制限されます。この数を増やして調整することをお勧めします。複製用のネットワークが 4M バイト/秒に対応している場合、2K バイトの作業負荷に最適なネットワークスレッドの数は、理論上は次のようになります。

$$(4096\text{K バイト}/\text{秒}) / (2\text{K バイト}/0.060 \text{ 入出力}/\text{秒}) = 123 \text{ スレッド}$$

この数字は、直線的に拡張できる場合を想定しています。実際には、ネットワークスレッドの数を 65 以上に増やしても効果はありません。整合グループが存在しない場合に、30 のボリュームが 4M バイト/秒の接続で複製されていて、8K バイトの入出力がある場合について考えてみます。デフォルトではボリュームごとに 2 つのネットワークスレッドが設定されているため、60 のネットワークスレッドが作成されます。作業負荷が各ボリューム間で均等に分散される場合、理論上の帯域幅は次のようになります。

$$60 \times (8\text{K バイト}/0.060 \text{ 入出力}/\text{秒}) = 8\text{M バイト}/\text{秒}$$

この値は、ネットワークの帯域幅を超えています。調整は必要ありません。

非同期キューのフラッシュスレッドの数は、デフォルトで 2 に設定されています。この設定を変更するには、`sndradm` コマンド行インタフェースを `-A` オプションを指定して実行します。`sndradm` の `-A` オプションには、セットが非同期モードで複製を行なっている場合に非同期キューを処理するために作成できるスレッドの最大数 (デフォルトは 2) を指定します。

非同期キューのサービスに対して現在設定されているフラッシュスレッドの数を確認するには、`sndradm -P` コマンドを使用します。たとえば、次のセットには 2 つの非同期フラッシュスレッドが設定されています。

```
# sndradm -P
/dev/md/rdsk/d52 -> lh1:/dev/md/sdsdg/rdsk/d102
autosync: off, max q writes: 4096, max q fbas: 16384, async threads: 2, mode:
async, group: butch, blocking diskqueue: /dev/md/rdsk/d100, state: replicating
```

`sndradm -A` オプションを使用して非同期キューのフラッシュスレッドの数を 3 に変更するには、次のように入力します。

```
# sndradm -A 3 lh1:/dev/md/sdsdg/rdsk/d102
```

ネットワークの調整

遠隔ミラーソフトウェアはシステムの入出力パスに直接関与し、すべてのトラフィックを監視して、遠隔ミラーボリュームへのトラフィックであるかどうかを確認します。遠隔ミラーボリュームに対する入出力コマンドが追跡され、その書き込み処理の複製が管理されます。遠隔ミラーソフトウェアがシステムの入出力パスに直接関与するため、システムのパフォーマンスに影響を与えることが予想されます。ネットワークの複製に必要な TCP/IP 処理が追加されるため、ホストの CPU 資源も消費されます。この節で説明する手順は、一次および二次の遠隔ミラーホストで実行してください。

TCP バッファのサイズ

TCP バッファのサイズとは、完了応答の待機に入るまでに TCP (Transfer Control Protocol) によって転送できるバイト数です。最大のスループットを得るには、使用する接続に最適な TCP 送受信ソケットのバッファサイズを使用することが重要です。バッファが小さすぎると、TCP の輻輳ウィンドウが完全には開きません。受信バッファが大きすぎると、TCP のフロー制御が失われて、送信側の処理が受信側を上回ることによって TCP ウィンドウが閉じる場合があります。これは、送信ホストの速度が受信ホストよりも早い場合によく発生します。送信側のウィンドウが大きすぎても、大容量のメモリーを使用しているかぎり問題はありません。

注 – 共有ネットワークでバッファのサイズを極端に大きくすると、ネットワークのパフォーマンスに影響を与えることがあります。サイズの調整については、Solaris System Administrator Collection を参照してください。

表 3-2 に、100BASE-T ネットワークの最大スループットの推定値を示します。

表 3-2 ネットワークのスループットおよびバッファサイズ

待ち時間	バッファサイズ=24K バイト	バッファサイズ=256K バイト
10 ミリ秒	18.75 MBps*	100 MBps
20 ミリ秒	9.38 MBps	100 MBps
50 ミリ秒	3.75 MBps	40 MBps
100 ミリ秒	1.88 MBps	20 MBps
200 ミリ秒	0.94 MBps	10 MBps

* M バイト/秒

TCP バッファースizeの表示および調整

/usr/bin/netstat(1M) および /usr/sbin/ndd(1M) コマンドを使用すると、TCP バッファースizeを表示および調整できます。調整を検討する必要のある TCP パラメータは、次のとおりです。

- tcp_max_buf
- tcp_cwnd_max
- tcp_xmit_hiwat
- tcp_recv_hiwat

これらのパラメータの1つを変更したら、shutdown コマンドで遠隔ミラーソフトウェアを再起動し、ソフトウェアが新しいバッファースizeを使用できるようにします。ただし、サーバーを停止して再起動すると、TCP バッファースizeはデフォルトのサイズに戻ります。変更を保持するには、この節の後半の説明に従って、起動スクリプトの値を設定します。

TCP バッファースizeおよび値を表示するためのネットワークの調整

次に、TCP バッファースizeおよび値を表示する手順を示します。

▼ すべての TCP バッファースizeを表示する

- すべての TCP バッファースizeを表示するには、次のように入力します。

```
# /usr/sbin/ndd /dev/tcp ? | more
```

▼ バッファースizeごとに設定を表示する

- バッファースizeごとに設定を表示するには、次のように入力します。

```
# /usr/sbin/ndd /dev/tcp tcp_max_buf
1073741824
```

このコマンドは、値が 1073741824 であることを示しています。

▼ ソケットのバッファサイズを表示する

- 特定のネットワークソケットのバッファサイズを表示するには、`/usr/bin/netstat(1M)` コマンドを実行します。

たとえば、遠隔ミラーのデフォルトのポート 121 のサイズを表示するには、次のように入力します。

```
# netstat -na |grep "121 "  
*.121 *.* 0 0 262144 0 LISTEN  
192.168.112.2.1009 192.168.111.2.121 263536 0 263536 0 ESTABLISHED  
192.168.112.2.121 192.168.111.2.1008 263536 0 263536 0 ESTABLISHED  
  
# netstat -na |grep rdc  
*.rdc *.* 0 0 262144 0 LISTEN  
ip229.1009 ip230.rdc 263536 0 263536 0 ESTABLISHED  
ip229.rdc ip230.ufsd 263536 0 263536 0 ESTABLISHED
```

この例の値 263536 は、256K バイトのバッファサイズであることを示しています。一次ホストおよび二次ホストでは、同じバッファサイズを設定する必要があります。

▼ 起動スクリプトのバッファサイズを設定して確認する

注 – このスクリプトは、一次ホストおよび二次ホストで作成してください。

1. 次の値を使用して、テキストエディタでスクリプトファイルを作成します。

```
#!/bin/sh  
nnd -set /dev/tcp tcp_max_buf 16777216  
nnd -set /dev/tcp tcp_cwnd_max 16777216  
  
# increase DEFAULT tcp window size  
nnd -set /dev/tcp tcp_xmit_hiwat 262144  
nnd -set /dev/tcp tcp_recv_hiwat 262144
```

2. ファイルを `/etc/rc2.d/S68nnd` という名前で保存して、終了します。
3. `/etc/rc2.d/S68nnd` ファイルに、アクセス権および所有権を設定します。

```
# /usr/bin/chmod 744 /etc/rc2.d/S68nnd  
# /usr/bin/chown root /etc/rc2.d/S68nnd
```

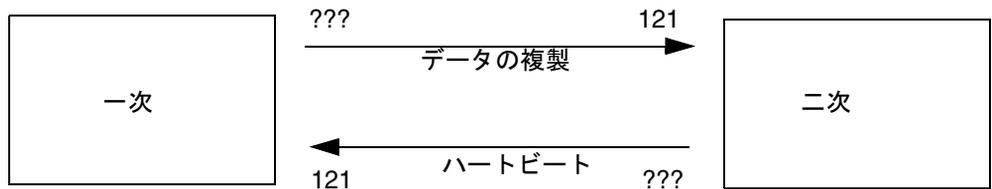
4. サーバーを停止して再起動します。

```
# /usr/sbin/shutdown -y g0 -i6
```

5. 38 ページの「ソケットのバッファサイズを表示する」の説明に従って、サイズを確認します。

遠隔ミラーによる TCP/IP ポートの使用

一次および二次の両方のノードの遠隔ミラーソフトウェアは、`/etc/services` に指定された既知のポート (ポート 121) で待機します。遠隔ミラーの書き込みトラフィックは、一次サイトの任意に割り当てられたアドレスと二次サイトの既知のアドレスが指定されたソケットを使用して、一次サイトから二次サイトへ流れます。健全性の監視用のハートビートは、二次ホストの任意に割り当てられたアドレスと一次ホストの既知のアドレスが指定された別の接続を使用して流れます。遠隔ミラープロトコルは、これらの接続に SUN RPC を使用します。



ポート 121 は、デフォルトの既知のアドレス

図 3-1 遠隔ミラーによる TCP ポートアドレスの使用

デフォルトの TCP 待機ポート

ポート 121 は、遠隔ミラーの `sndrd` デーモンが使用するデフォルトの TCP ポートです。このポート番号を変更するには、テキストエディタを使用して `/etc/services` ファイルを編集します。

ポート番号を変更する場合は、この構成セット内のすべての遠隔ミラーホスト (つまり、一次ホストと二次ホスト、1 対多、多対 1、およびマルチホップ構成のすべてのホスト) で、ポート番号を変更してください。また、影響を受けるすべてのホストを停止および再起動して、ポート番号の変更を有効にする必要があります。

遠隔ミラーとファイアウォールの使用

RPC 要求には応答が必要であるため、ファイアウォールは、パケットの発信元または宛先フィールドのいずれかが既知のポートアドレスであっても許可するように開いている必要があります。オプションを指定できる場合は、RPC トラフィックも許可するようにファイアウォールを構成してください。

書き込み複製トラフィックの場合、二次ホストへのパケットでは、宛先フィールドに既知のポート番号が設定されている必要があります。RPC の応答では、発信元フィールドに既知のアドレスが設定されます。

健全性の監視の場合、二次ホストから流れるハートビートでは、宛先フィールドに既知のアドレスが設定されます。応答では、発信元フィールドにこのアドレスが設定されます。

遠隔ミラーソフトウェアとポイントインタイムコピーソフトウェアの併用

通常の運用を行なっているときに、どちらのサイトでもデータの完全性とシステムのパフォーマンスをもっとも高いレベルで維持するには、Sun StorageTek Availability Suite 4.0 ポイントインタイムコピーソフトウェアと遠隔ミラーソフトウェアを併用することをお勧めします。

ポイントインタイムコピーを物理的な遠隔地に複製できるため、全体的な災害回復計画の一環としてボリュームの一貫性のあるコピーを作成できます。これは通常、バッチ複製と呼ばれます。この方法の処理および利点については、最良事例ガイド『Sun StorageTek Availability Suite Software - Improving Data Replication over a Highly Latent Link』に記載されています。

一次サイト (一次ボリュームが存在するサイト) から二次ボリュームへの同期処理は、遠隔ミラーの二次ボリュームのポイントインタイムコピーが作成されたあとに開始されます。再同期が開始される前に、ポイントインタイムコピーソフトウェアを使用可能にして二次サイトの複製データのポイントインタイムコピーを作成することによって、二重の障害に対する保護を実現できます。再同期中に二次障害が発生した場合には、ポイントインタイムコピーを代替として使用して、二次障害が解決したときに再同期を復元再開できます。二次サイトと一次サイトが完全に同期化されたあとは、ポイントインタイムコピーソフトウェアのボリュームセットを使用不可にするか、二次サイトで遠隔バックアップ、遠隔データ解析などのその他の機能に使用することができます。

使用可能への切り替え、コピー、更新処理中に内部で実行されるポイントインタイムコピーソフトウェアの入出力処理は、入出力スタックに対する新規の入出力を行うことなく、シャドウボリュームの内容を変更できます。このとき、入出力は sv 層で横

取りされません。シャドウボリュームが遠隔ミラーボリュームにもなっている場合、遠隔ミラーソフトウェアはこれらの入出力処理を認識しません。この状況では、入出力によって変更されたデータは、対象となる遠隔ミラーボリュームに複製されません。

この複製を行うには、ポイントインタイムコピーソフトウェアが遠隔ミラーソフトウェアに変更されたビットマップを提供するように構成します。遠隔ミラーソフトウェアは、記録モードになっているときにはビットマップを受け取り、ポイントインタイムコピーソフトウェアのビットマップと、遠隔ミラー自身のそのボリュームのビットマップとの論理和をとります。こうしてポイントインタイムコピーソフトウェアの変更を自身の変更リストに追加して、データを遠隔ノードに複製します。遠隔ミラーソフトウェアがボリュームの複製モードになっているときには、ポイントインタイムコピーソフトウェアからのビットマップは拒否されます。その結果、使用可能への切り替え、コピー、更新処理が失敗します。遠隔ミラーの記録がふたたび使用可能になったとき、ポイントインタイムコピーソフトウェアの処理を再実行できます。

注 - 遠隔ミラーボリューム上でポイントインタイムコピーソフトウェアの使用可能への切り替え、コピー、更新、またはリセット処理を正常に実行するには、遠隔ミラーボリュームセットを記録モードにする必要があります。記録モードになっていないと、ポイントインタイムコピー処理は失敗し、処理が拒否されたことが遠隔ミラーソフトによって報告されます。

遠隔複製の構成

遠隔ミラーソフトウェアでは、1対多、多対1、およびマルチホップのボリュームセットを作成できます。

- 1対多の複製では、1つの一次ボリュームから、1つまたは複数のホスト上にある多数の二次ボリュームにデータを複製できます。1つの一次サイトボリュームと各二次サイトボリュームが、1つのボリュームセットになります。たとえば、1つの一次ホストボリュームと3つの二次ホストボリュームがある場合は、一次 A と二次 B1、一次 A と二次 B2、一次 A と二次 B3 の、3つのボリュームセットを構成する必要があります。
- 多対1の複製では、3つ以上のホストで2つ以上のネットワーク接続を介してボリュームを複製できます。遠隔ミラーソフトウェアは、多数の異なるホスト上のボリュームを1つのホスト上のボリュームに複製することができます。1対多構成の「1」および「多」がボリュームを表すのとは異なります。
- マルチホップの複製では、1つのボリュームセットの二次ホストボリュームが、ほかのボリュームセットの一次ホストボリュームとしても機能します。1つの一次ホストボリューム A と1つの二次ホストボリューム B がある場合、二次ホストボリューム B は、二次ホストボリューム B1 からは一次ホストボリューム A1 として見えます。

遠隔ミラーソフトウェアでは、これらを組み合わせた構成もすべてサポートされま
す。

用語集

dsstat	遠隔ミラーおよびポイントインタイムスナップショット製品のカーネル統計情報を表示する Sun StorageTek Availability Suite のツールセットに含まれるツール。
hwm	「ハイウォーターマーク」を参照。
TCP バッファ	TCP (Transfer Control Protocol) バッファのサイズとは、応答の待機に入る前に TCP によって転送できるバイト数である。
一次またはローカル: ホストまたはボリューム	ホストアプリケーションが主に依存するシステムまたはボリューム。たとえば、製品データベースによってアクセスされるシステムまたはボリュームなどである。このデータは、遠隔ミラーソフトウェアによって二次ボリュームに複製される。
完全同期	完全なボリューム間コピー。同期操作の中でもっとも時間を要する操作である。ほとんどの場合、一次ボリュームを基準に対応する二次ボリュームの同期がとられる。ただし、一次ディスクで障害が発生した場合は、復元のため残った遠隔ミラーを基準に逆方向の同期が必要な場合もある。
逆方向の同期	回復リハーサルで使用される操作。リハーサル中に二次システムに適用されたテスト更新は記録によって追跡される。一次システムが復元されたら、テスト更新に一次イメージのブロックを上書きして、一致する遠隔セットを復元する。
記録	各入出力イベントの実行記録ではなく、ビットマップによってディスクへの書き込みを追跡するモード。遠隔サービスで中断や障害が発生している間に、遠隔にコピーされなかったディスクの更新内容を追跡する。各ソースボリュームについて、遠隔セットと一致しないブロックが識別される。ソフトウェアは、完全なボリューム間コピーではなく、このログを使用して最適化された更新同期を行うことで、遠隔ミラーを再確立する。
更新同期	記録によって特定されたディスクブロックだけをコピーすることで、遠隔ミラーボリュームセットの復元時間を削減する同期。

構成情報の場所	Sun StorageTek Availability Suite ソフトウェアが使用する、使用可能なボリュームのすべてに関する構成情報を格納する場所。
自動同期	一次ホストで自動同期オプションを使用可能にすると、システムの再起動または接続障害が発生したときに、同期デーモン (autosyncd) によってボリュームセットの再同期が行われる。
順方向の再同期	「更新同期」を参照。
整合グループ	書き込み順序を保持するために 1 つの非同期キューを共有する遠隔ボリュームのグループ。
遅延クリア	コア内のビットをクリアする動作。ただし、別のビットが設定されるか、コア内のコピーが再利用されるまでは、ディスクへのビットマップブロックの書き込みを行わない。これは、システム障害後に変更が単に再転送される動作であるため、安全である。
同期	遠隔ミラーソフトウェアによるミラー化を開始する前の状態として、ソースディスクと同一のコピーをターゲットディスクに作成すること。
同期複製	短い距離 (数十キロメートル) に制限される複製。距離が長いと、伝播遅延が入出力応答時間に悪影響を及ぼす。
二次または遠隔: ホスト またはボリューム	一次ホストまたは一次ボリュームの遠隔の複製。データのコピーが書き込まれて読み取られる。遠隔コピーは、ピアサーバー間でホストの介入なしに転送される。1 つのサーバーが、あるボリュームの一次記憶域として機能するとともに、ほかのボリュームの二次 (遠隔) 記憶域として使用される場合がある。
ハイウォーターマーク	これまでに使用された非同期キューの最大数。
非同期キュー	遠隔サイトに複製する書き込みを格納するために使用する、ローカルにあるディスクまたはメモリー上の領域。書き込みがキューに入ると、アプリケーションに完了が通知される。そのあとネットワーク能力に空きができたときに、書き込みが遠隔サイトに転送される。
非同期複製	遠隔イメージを更新する前に、一次ホストの入出力トランザクションが完了したことを元のホストに通知する複製の形式。つまり、ローカルの書き込み操作が終了して遠隔の書き込み命令がキューに入ると、ホストに入出力トランザクションが完了したことが通知される。二次コピーをあとで行うことで、長距離による伝播遅延が入出力応答時間に含まれないようにする。
非ブロック	(非同期キュー) 非ブロックモードでは、非同期キューがいっぱいになると、遠隔ミラーソフトウェアはスコアボーディングモードになり、キューの内容は破棄される。非ブロックモードでは、二次サイトへのパケットの書き込み順序は保証されない。ただし、非同期キューがいっぱいになった場合でも、アプリケーションへの応答時間には影響を及ぼさない。
ファイアウォール	2 つのネットワーク間のインタフェースとして動作し、ネットワーク間のトラフィックを調整するコンピュータ。外部ネットワークからの電子攻撃から内部ネットワークを保護することを目的とする。

複製 ボリュームセットが最初に同期化されたあと、遠隔ミラーソフトウェアは、一次ボリュームと二次ボリュームに継続して同じデータが含まれることを保証する。複製は、ユーザー層のアプリケーション書き込み操作によって起動される継続的なプロセスである。

ブロック (非同期キュー) ブロックモードでは、非同期キューがいっぱいになると、そのあとのすべての書き込みはキューの内容が排出されて書き込める状態になるまで遅延される。ブロックモードは、デフォルトの非同期動作モードで、二次サイトへのパケットの書き込み順序を保証する。ブロックモードが設定されているときに非同期キューがいっぱいになると、アプリケーションへの応答時間に影響を及ぼす場合がある。

**ボリュームセット
ファイル**

特定のボリュームセットに関する情報を含むテキストファイル。このテキストファイルは、遠隔ミラーソフトウェアおよびポイントインタイムコピーソフトウェアが使用する構成されたボリュームセットのすべてに関する情報を含む構成情報の場所とは異なる。

索引

記号

/etc/hosts, 7
/etc/nsswitch.conf ファイル
編集, 12
/etc/services ファイル
編集, 11
/usr/kernel/drv/rdc.conf, 13

A

Availability Suite ソフトウェア
アップグレード, 2
アンインストール, 4
インストール, 3

D

dscfgadm ユーティリティー, 15

S

Sun StorageTek
インストール後, 5
構成, 5

T

TCP / IP ポート, 39

あ

アップグレード、Availability Suite ソフトウェア
, 2
アンインストール、Availability Suite ソフトウェア
, 4

い

インストール、Availability Suite ソフトウェア, 3
インストール後
構成, 5
インターネットプロトコルバージョン 6 (IPv6), 8

こ

構成
IPv6 アドレス, 8
インストール後, 5
ファイル, 7
ファイル (省略可能), 20

構成情報
バックアップ, 22

構成手順, 6

せ

整合グループ, 26

そ

ソフトウェアの設定
変更, 13

と

同期複製, 24

ね

ネットワークの調整, 36
TCP バッファのサイズ, 36

ひ

ビットマップ
要件, 18
ビットマップボリューム
サイズの要件, 18
推奨場所, 18
非同期キュー
構成, 28
設定サイズ, 33
フラッシュスレッドの構成, 34
非同期複製, 25

ふ

ファイル
/etc/hosts, 7
/usr/kernel/drv/rdc.conf, 13
複製
遠隔, 27
遠隔構成, 41
同期, 24
非同期, 25

ほ

ボリュームセットファイル
使用, 20