

CLUSTERPRO

MC StorageSaver 1.1 for Linux

ユーザーズガイド

© 2013(Sep) NEC Corporation

- ☐ 製品の概要
- ☐ 製品の機能
- ☐ 設定ファイル
- ☐ 操作・運用手順
- ☐ CLUSTERPRO との連携
- ☐ RENS との連携
- ☐ syslog メッセージ
- ☐ 注意・制限事項について
- ☐ リファレンス
- ☐ 付録

はしがき

本書は、CLUSTERPRO MC StorageSaver 1.1 for Linux（以後 StorageSaver と記載します）のディスク装置監視に関する設定について記載したものです。

（１）商標および登録商標

- ✓ Red Hat は、米国およびその他の国における Red Hat, Inc.の商標または登録商標です。
- ✓ SUSE は、米国およびその他の国における Novell, Inc.の登録商標または商標です。
- ✓ Linux は、Linus Torvalds 氏の米国およびその他の国における、登録商標または商標です。
- ✓ EMC、Symmetrix DMX、Symmetrix VMAX、CLARiX、PowerPath は EMC Corporation の商標または登録商標です。
- ✓ VMware、VMware vSphere は、VMware, Inc.の米国および各国での商標または登録商標です。
- ✓ その他、本書に登場する会社名および商品名は各社の商標または登録商標です。
- ✓ なお、本書では®、TM マークを明記していません。

（２）参考ドキュメント

- ・ 『CLUSTERPRO MC StorageSaver 導入ガイド』
- ・ 『CLUSTERPRO MC StorageSaver 1.1 for Linux リリースメモ』

(3) 本リリースの強化点について

StorageSaver 1.1 (2013 年 10 月出荷版) では、下記の機能を強化しています。

- vSphere ESXi 上のゲスト OS における監視機能を強化しました。
vSphere ESXi 上のゲスト OS から、ESXi ホストに接続されている物理パスの監視を行えるよう監視機能を強化しました。
詳細については「CLUSTERPRO MC StorageSaver 1.1 for Linux ユーザーズガイド (vSphere 対応版)」を参照してください。
- TestI/O 発行方式の設定方法を変更しました。
TestI/O の発行方式を設定するパラメータとして、システム定義ファイルに新規パラメータ (TESTIO_MODE) を追加しました。
従来のバージョンにおけるパラメータ TESTIO_DIRECT は本パラメータに統合されました。
また、これに伴い TestUnitReady の発行抑止についても、本パラメータで設定するよう変更しました。
パラメータの詳細については “3.3 設定ファイルの記述” を参照してください。
なお、従来どおり TESTIO_DIRECT の指定、ならびに TIME_TUR_INTERVAL での TestUnitReady の発行抑止も可能ですので、1.0 以前のバージョンの設定ファイルもそのまま使用可能です。
- 障害検出時のアクション内容を強化しました。
I/O ストール検出時に実行するアクションとして、クラスタウェア連携用コマンド (srgstat) を使用した CLUSTERPRO 連携が指定可能となりました。
これに伴い、障害時のアクションを指定するパラメータ (VG_FAULT_ACTION および VG_STALL_ACTION) に指定する値も変更しました。
パラメータの詳細については “3.3 設定ファイルの記述” を参照してください。
なお、従来の設定値も本バージョンでは使用可能ですので、1.0 以前のバージョンの設定ファイルもそのまま使用可能です。

目次

1. 製品の概要	1
1.1. 製品概要について.....	1
1.2. 製品の構成について.....	5
1.3. 製品導入に関する注意	7
2. 製品の機能	9
2.1. ディスク装置のリソース監視手順.....	9
2.2. I/O パスの監視手順について.....	13
2.3. リソース監視で異常を検出すると.....	18
2.4. アクションの定義について	19
2.5. オンライン保守機能.....	24
2.6. RENS との連携機能について	25
2.7. クラスタウェアとの連携機能について	25
3. 設定ファイル	27
3.1. 本製品の導入	27
3.2. 設定ファイルの自動生成手順	29
3.3. 設定ファイルの記述.....	33
4. 操作・運用手順.....	47
4.1. 運用管理コマンドの操作手順.....	47
4.2. オンライン保守コマンドの操作手順.....	55
4.3. 障害復旧時の操作.....	65
5. CLUSTERPRO との連携.....	69
5.1. CLUSTERPRO との連携の概要.....	69
5.2. カスタムモニタリソースによる CLUSTERPRO との連携.....	70
5.3. CLUSTERPRO のサーバ管理プロセス (clpnm) の強制終了による CLUSTERPRO との 連携.....	73
5.4. システムメモリダンプ採取と OS 強制停止による CLUSTERPRO との連携.....	75
6. RENS との連携	77
6.1. RENS 連携の概要.....	77
6.2. RENS 連携の設定.....	78
6.3. RENS 連携用モニタプロセスの運用管理	80
6.4. RENS を利用した CLUSTERPRO 連携手順.....	81
7. syslog メッセージ	91
7.1. syslog に出力するメッセージについて	91
7.2. 警報対象として登録することを推奨するメッセージ一覧	92
7.3. その他のメッセージ.....	93
7.4. 運用管理製品との連携.....	93
8. 注意・制限事項について	95
8.1. 注意・制限事項.....	95
8.2. オンライン保守における注意事項.....	102

9. リファレンス.....	103
10. 付録.....	121
10.1. 運用管理コマンド.....	121
10.2. srgquery(1M) による設定ファイル自動生成手順.....	123
10.3. 本製品のテスト手順について.....	125
10.4. カスタムモニタリソースによる CLUSTERPRO との連携手順.....	132
10.4.1. CLUSTERPRO 連携設定.....	132
10.4.2. 動作確認.....	139

1. 製品の概要

1.1. 製品概要について

(1) 製品の提供する主な機能

本製品は、Linux のディスク装置の I/O パスの動作状態を定期監視します。I/O パスの異常を検出すると障害レポートを通知し、さらにディスク装置へのアクセスができなくなるとクラスタウェアと連携しノードを切り替えることでクラスタシステムでの可用性を向上させます。

- 共有ディスクのリソース監視機能

FC 接続や SCSI で接続されたディスク装置に対して、TestI/O によるリソース監視を実現します。

- ディスクアレイ装置コントローラの障害監視
- FC カード、SCSI カードの障害監視
- I/O パスの死活監視
- I/O リクエストのストール監視

- I/O パスの自動閉塞機能

StoragePathSavior(以降、SPS)または PowerPath で冗長化された I/O パスを有するディスク装置において、リソース監視で異常を検出すると障害の発生した I/O パスを閉塞し、すみやかに正常なパスへ切り替えます。

- 間欠的な FC リンクダウン障害に伴う頻繁なパス切り替えによる I/O パフォーマンスの低下や他のディスク装置への影響を未然に防止します。

注意: 閉塞機能・復旧機能は SPS または PowerPath を導入している環境でのみ有効です。

- クラスタウェア連携機能

ディスク装置へのすべてのインタフェース機構の異常により、ユーザーデータへのアクセスができなくなると、クラスタウェアと連携し、パッケージの移動やノード切り替えにより業務の継続的な運用を実現します。

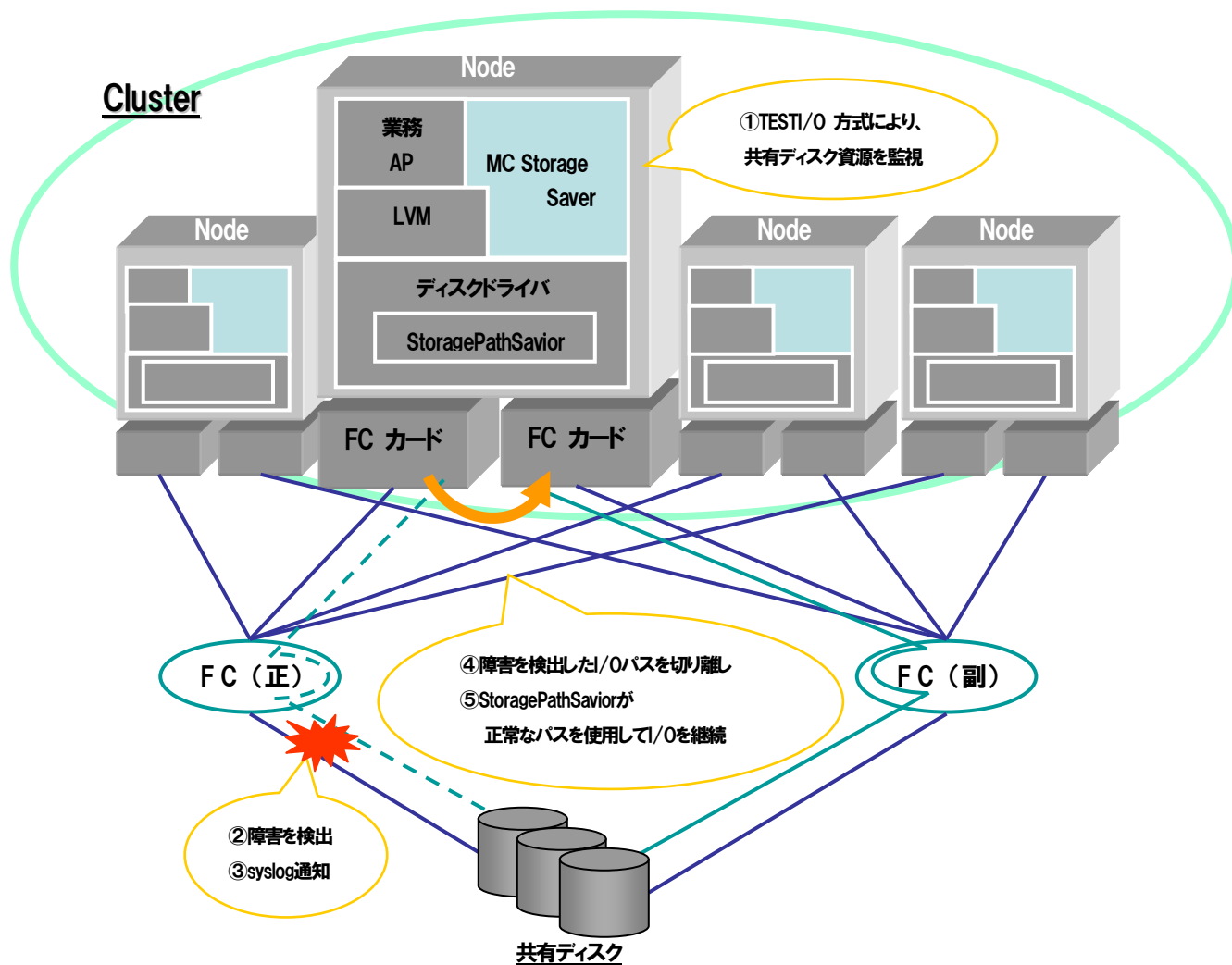
注意: 本機能は、クラスタウェアが導入されたクラスタシステムで使用可能です。

- ・ オンライン保守機能
FC 接続のディスク装置に対して、SPS または PowerPath における I/O パスの一括閉塞、一括復旧を実現します。
これにより、FC 上でリンクダウン時の障害が発生した際に、障害装置の特定化と保守員によるシステム無停止保守を可能とします。
- ・ 運用管理機能
I/O パスの監視状態の表示や手動による閉塞、復旧といった運用管理機能をコマンドインタフェースで提供します。

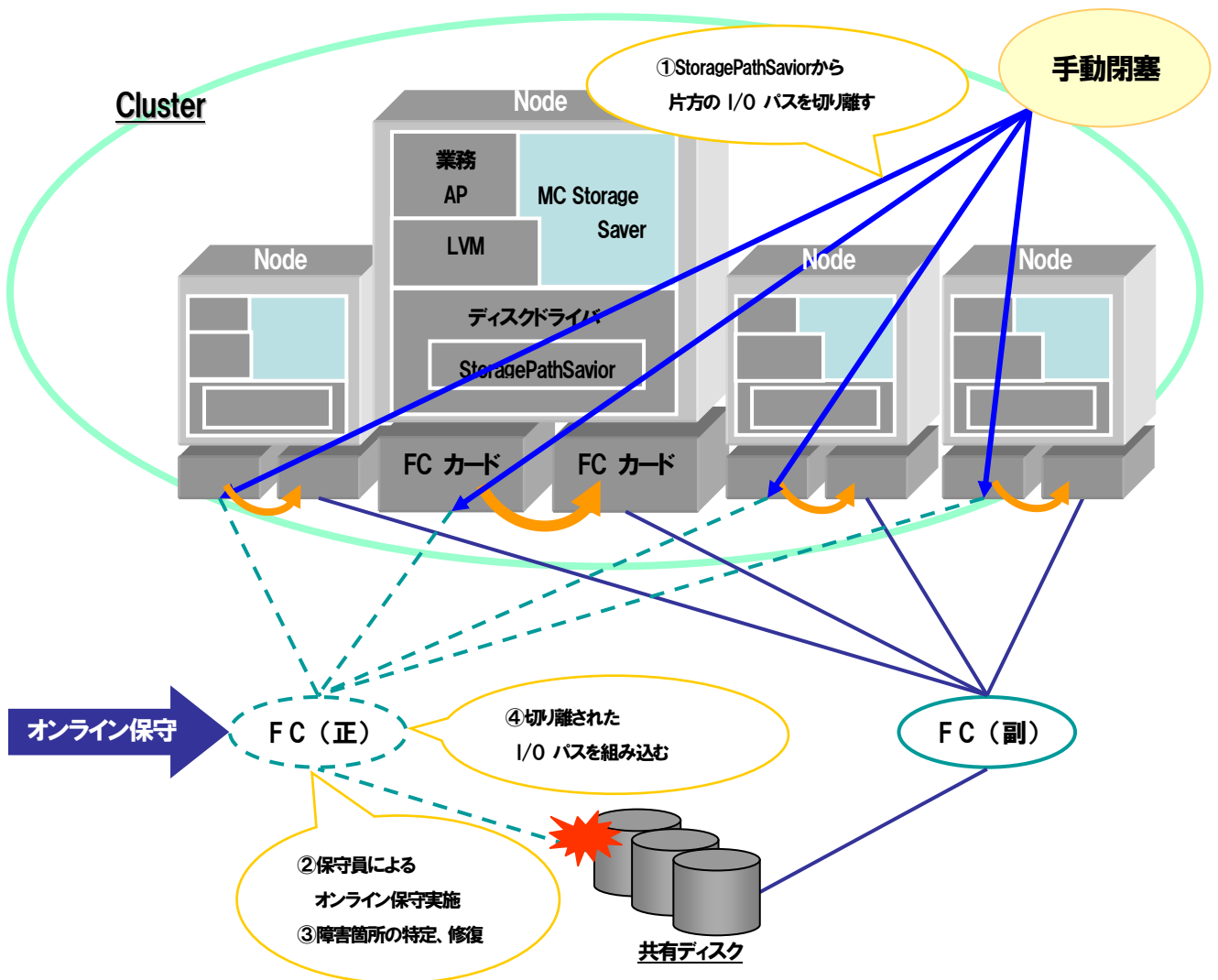
注意: 閉塞機能・復旧機能は SPS または PowerPath を導入している環境でのみ有効です。

- ・ オートコンフィグレーション機能
設定ファイルの自動生成機能です。
アクションに関しては、ユーザーカスタマイズ機能によりリソース単位での定義が可能です。
- ・ プロセス監視機能
本製品で提供するデーモンプロセスやリソース監視コマンドの動作状態を監視し、異常を検出すると自動的に再起動します。これにより、継続的なリソース監視を実現します。

(2) 自動閉塞の流れ



(3) オンライン保守の流れ



1.2. 製品の構成について

(1) StorageSaver は、下記のコンポーネントにより構成されます。

(a) StorageSaver

ディスク装置のリソース監視、I/O パスの運用管理を行う機能です。

下記のデーモンプロセスおよびコマンドにより構成されます。

・ srgd (1M)	リソース監視デーモン
・ srgping (1M)	リソース監視モニタ
・ srgadmin (1M)	運用管理コマンド
・ srgquery (1M)	設定ファイル生成コマンド
・ srgconfig (1M)	設定ファイル確認コマンド
・ srgstat (1M)	クラスタウェア連携用コマンド
・ srgwatch (1M)	プロセス監視デーモン
・ srgextend (1M)	手動復旧コマンド
・ srgreduce (1M)	手動閉塞コマンド
・ srgrecover (1M)	構成復旧コマンド

下記のディレクトリを使用します。

・ 実行形式ディレクトリ	/opt/HA/SrG/bin
・ 実行形式ディレクトリ	/opt/HA/SrG/local/bin
・ 設定ファイル管理ディレクトリ	/var/opt/HA/SrG/conf
・ ログ管理ディレクトリ	/var/opt/HA/SrG/log
・ 内部管理用ディレクトリ	/var/opt/HA/SrG/local/conf
・ rc ファイル格納ディレクトリ	【Red Hat Enterprise Linux ES/AS 4.x】 【Red Hat Enterprise Linux 5.x】 【Red Hat Enterprise Linux 6.x】 【Oracle Linux 5.x】 【Oracle Linux 6.x】 /etc/init.d /etc/rc.d/rc0.d /etc/rc.d/rc1.d /etc/rc.d/rc2.d /etc/rc.d/rc3.d /etc/rc.d/rc4.d /etc/rc.d/rc5.d /etc/rc.d/rc6.d 【SUSE Linux Enterprise Server 10 SP3】 【SUSE Linux Enterprise Server 11】 /etc/init.d /etc/rc.d/rc3.d /etc/rc.d/rc5.d

(b) StorageSaver RENS edition

StorageSaver の VG リソースを RENS(Resource Event Notification Service)フレームワークから
モニタし、クラスタウェアと連携するための機能です。

下記のデーモンプロセスおよびコマンドにより構成されます。

- | | |
|----------------|-----------------|
| ▪ ssdiagd (1M) | RENS 連携用モニタプロセス |
| ▪ ssreq (1M) | 運用管理コマンド |

下記のディレクトリを使用します。

- | | |
|-------------------|---|
| ▪ 実行形式ディレクトリ | /opt/HA/SrG/bin |
| ▪ 設定ファイル格納ディレクトリ | /var/opt/HA/SrG/conf/rens |
| ▪ rc ファイル格納ディレクトリ | 【Red Hat Enterprise Linux ES/AS 4.x】
【Red Hat Enterprise Linux 5.x】
【Red Hat Enterprise Linux 6.x】
【Oracle Linux 5.x】
【Oracle Linux 6.x】
/etc/init.d
/etc/rc.d/rc0.d
/etc/rc.d/rc1.d
/etc/rc.d/rc2.d
/etc/rc.d/rc3.d
/etc/rc.d/rc4.d
/etc/rc.d/rc5.d
/etc/rc.d/rc6.d

【SUSE Linux Enterprise Server 10 SP3】
【SUSE Linux Enterprise Server 11】
/etc/init.d
/etc/rc.d/rc3.d
/etc/rc.d/rc5.d |

1.3.製品導入に関する注意

本製品は、HW 構成、SW 構成、運用環境によってはご利用いただける機能が制約される場合があります。導入にあたっては、十分な検証を実施してください。

- (1) HW 構成的なサポート範囲は下記のとおりです。

インターフェース

- FC スイッチ接続
- FC 直結接続
- SCSI SE/FWD 接続

ディスク装置

- IA32 および x86_64 対応 CPU 搭載サーバ接続のディスクアレイ装置
- 増設ディスク装置

(注) NEC が正式販売しているディスク装置が対象となります。

2013 年 10 月時点でサポート済みのディスクアレイ装置は以下の通りです。

- ・NEC 社製 iStorage 全シリーズ (ただし、E1 シリーズは除きます)
- ・EMC 社製 CLARiX シリーズ
- ・EMC 社製 Symmetrix DMX シリーズ、Symmetrix VMAX シリーズ
- ・日立 社製 SANRISE シリーズ (個別対応となります)
Hitachi USP シリーズ (個別対応となります)
Hitachi VSP シリーズ (個別対応となります)

個別対応のディスク装置や上記以外のディスク装置を接続、監視する場合は、開発部門までお問い合わせください。

- (2) SW 構成的なサポート範囲は下記のとおりです。

ボリューム管理

- LVM

I/O パス管理製品

- StoragePathSavior
- EMC 社製 PowerPath
- 日立社製 HDLM (個別対応となります)

個別対応の I/O パス管理製品を使用する場合は、開発部門までお問い合わせください。

注意: LVM を構成しないディスク (/dev/sda 等) も監視可能です

2. 製品の機能

2.1. ディスク装置のリソース監視手順

ディスクアレイ装置を構成する I/O パスに対して定期的に TestI/O を発行することで、I/O パスの障害を早期に検出します。TestI/O で異常を検出した I/O パスについては障害状況をレポートし、障害の波及を防止するために I/O パスの自動閉塞やノード切り替え等のコンフィグレーションで規定されたアクションを実行します。

TestI/O の監視対象となる検査項目は下記のとおりです。

- I/O パスの死活監視
- I/O リクエストのストール監視

TestI/O は SCSI パススルードライバ(sg ドライバ)経由で行われますが、下記の SCSI コマンドを使用します。

- Inquiry command
- TestUnitReady command

注意: StorageSaver は内部で以下のパッケージを利用します。

sg3_utils Utils for Linux's SCSI generic driver devices + raw devices

本パッケージがインストールされていない場合、事前にインストールしてください。

- (1) 対象となるディスク装置
 - 増設 SCSI ディスク装置
 - SCSI 接続ディスクアレイ装置
 - FC 接続ディスクアレイ装置
- (2) ディスク装置のリソース監視手順
 - 1. ディスク装置コントローラに対して
TestI/O (Inquiry command)を発行します。
 - 2. LUN を構成する I/O パスに対して
TestI/O (TestUnitReady command) を発行します。
- (3) TestI/O による監視項目について
 - I/O パスの死活監視
 - I/O リクエストのストール監視

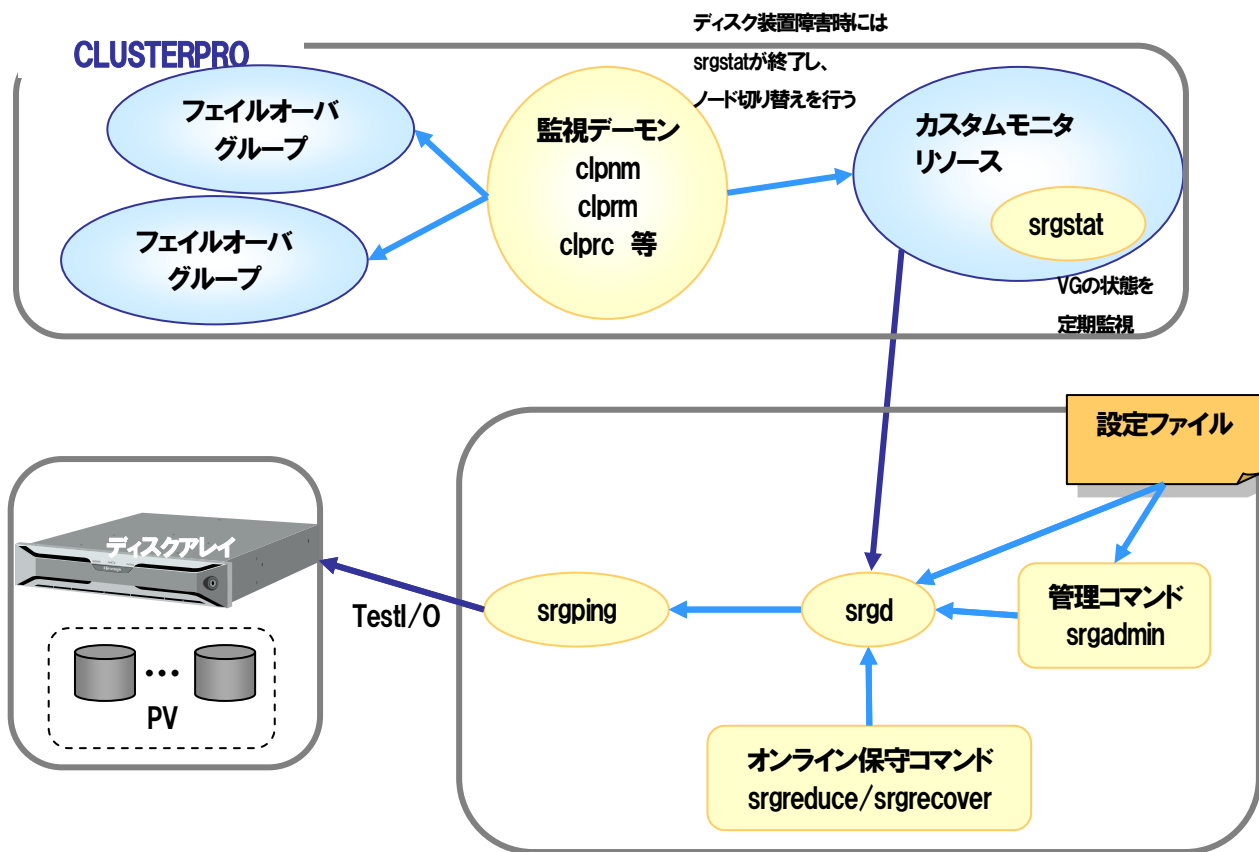
(注)ディスク装置のメディアエラーは検出できません。
- (4) TestI/O の対象となる I/O パスについて
設定ファイルに登録された I/O パスが TestI/O の対象となります。
下記のように監視対象から切り離されている I/O パスは TestI/O の対象とはなりません。
 - 閉塞状態、障害状態の I/O パス
 - オンライン保守実施中の I/O パス

- (5) コントローラに対する TestI/O 実行手順について
FC リンクダウンやコントローラ障害を検出するために、
ディスク装置配下のコントローラに対して TestI/O を発行します。
コントローラが正常応答すれば、デフォルト 20 秒間隔で TestI/O を繰り返します。
コントローラが正常応答しない場合は、デフォルト 180 秒の間 TestI/O を継続実行し
このリトライ時間以内に復旧しなければ、コントローラおよび配下の I/O パスを
障害状態として扱い TestI/O を終了します。
- (6) I/O パスに対する TestI/O 実行手順について
LUN の障害を検出するために I/O パス単位で TestI/O を実行します。
I/O パスが正常応答すれば、デフォルト 180 秒間隔で TestI/O を繰り返します。
LUN が正常応答しない場合は、デフォルト 180 秒の間 TestI/O を継続実行し
このリトライ時間以内に復旧しなければ、I/O パスを障害状態として扱い TestI/O を終了します。

注意:一部のディスク装置は、監視ソフトによる TestUnitReady command の発行を許可していないため LUN への TestI/O は実行できません。
このような装置を監視する場合には設定ファイルの TESTIO_MODE 値を INQ に設定する必要があります。
詳細は“3.2 設定ファイルの自動生成手順 (5)注意事項”を参照してください。
EMC 社製の CLARiX シリーズ、NEC 社製の iStorage M シリーズがこれに該当します。

注意:旧式のディスク装置は、LUN が故障した状態でも TestI/O に対して装置コントローラが正常な値を返す場合があるため、LUN の監視が有効に機能しない場合があります。

- (7) I/O パスの死活管理について
TestI/O の実行結果として、以下の状態をレポートします。
- UP
TestI/O が正常終了し I/O パスが正常に動作している状態です。
 - DOWN
TestI/O が異常終了し I/O パスが利用不可な状態です。
- VG レベル(I/O パスを SPS 等のマルチパス管理製品で冗長化した GROUP 単位)のステータスとして以下の状態をレポートします。
- UP
VG が正常に動作している状態です。
 - SUSPEND
VG を構成する片系の I/O パ스에異常を検出した状態です。
 - DOWN
VG に異常があり、利用不可な状態です。
- I/O パスの組み込み状態を示す Online status として、以下の状態をレポートします。
- extended
I/O パスが組み込まれた状態です。
 - reduced
I/O パスが閉塞された状態です。
 - unknown、alive
I/O パスの状態が不明です。



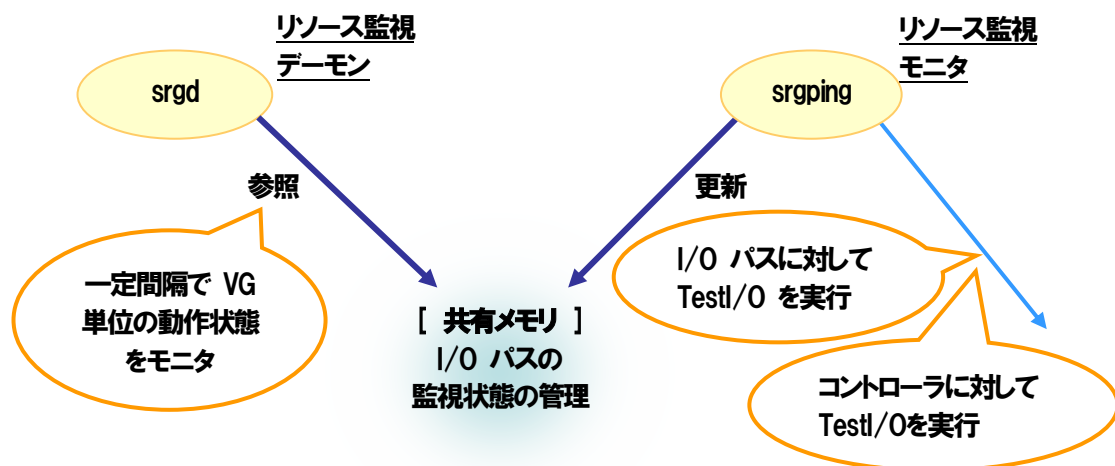
【StorageSaver の構成】

2.2.I/O パスの監視手順について

(1) I/O パスの死活監視

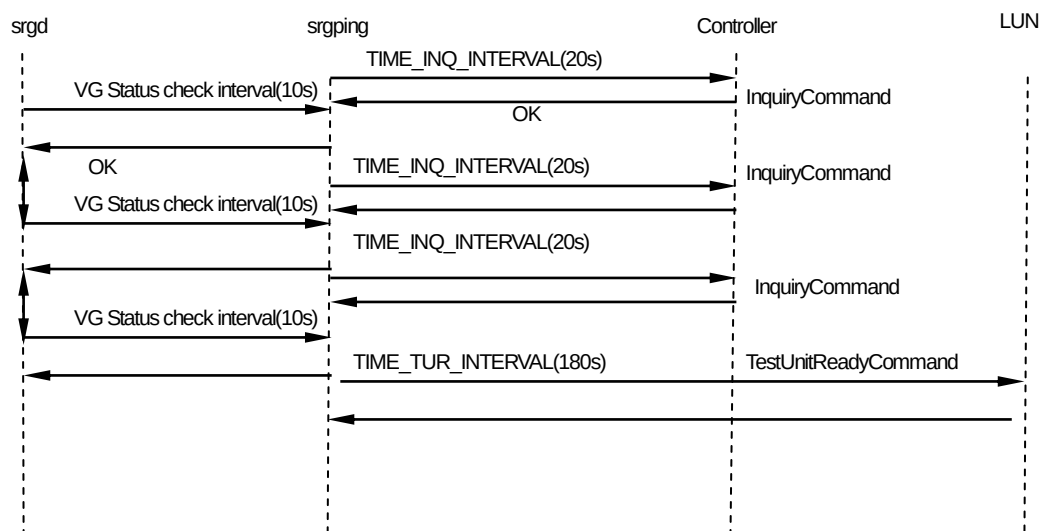
ディスク装置コントローラおよび LUN を構成する I/O パスに対して定期的に SCSI パススルーコマンドを利用して TestI/O を発行します。TestI/O が正常終了しない、または、タイムアウトした場合は I/O パスを異常と判定します。

【TestI/O のフレームワーク】

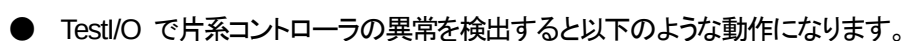


● TestI/O の正常なシーケンスは以下のような動作になります。

VG 障害検出時間	: TIME_VG_FAULT	: 60(秒)
I/O パスの障害検出時間	: TIME_LINKDOWN	: 180(秒)
コントローラ監視間隔	: TIME_INQ_INTERVAL	: 20(秒)
LUN 監視間隔	: TIME_TUR_INTERVAL	: 180(秒)



- ## 【TestI/O で片系コントローラ異常を検出】



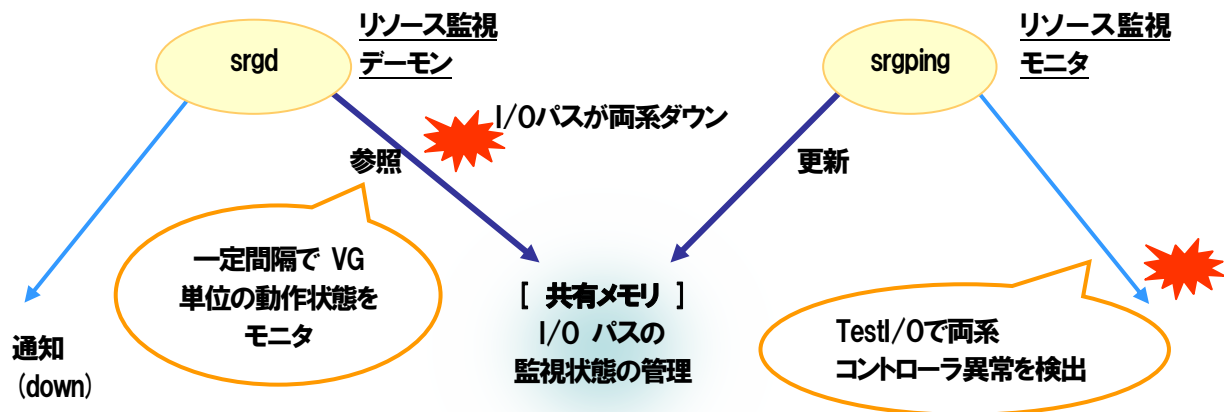
```

sequenceDiagram
    participant srgd
    participant srgping
    participant Controller

    srgd->>srgping: VG Status check interval(10s)
    srgping->>Controller: TIME_INQ_INTERVAL(20s)
    Controller-->>srgping: OK
    srgping-->>srgd: OK
    srgd->>srgping: RETRY
    srgd->>srgping: VG Status check interval(10s)
    srgping->>Controller: TIME_INQ_INTERVAL(20s)
    Controller-->>srgping: Error
    srgping-->>srgd: Error
    srgd->>srgping: RETRY
    srgd->>srgping: VG Status check interval(10s)
    srgping->>Controller: InquiryCommand
    Controller-->>srgping: Error
    srgping-->>srgd: Error
    srgd->>srgping: RETRY
    srgd->>srgping: VG Status check interval(10s)
    srgping->>Controller: Error
    Controller-->>srgping: Error
    srgping-->>srgd: Error
    srgd->>srgping: TIME_LINKDOWN(180s)
    
```

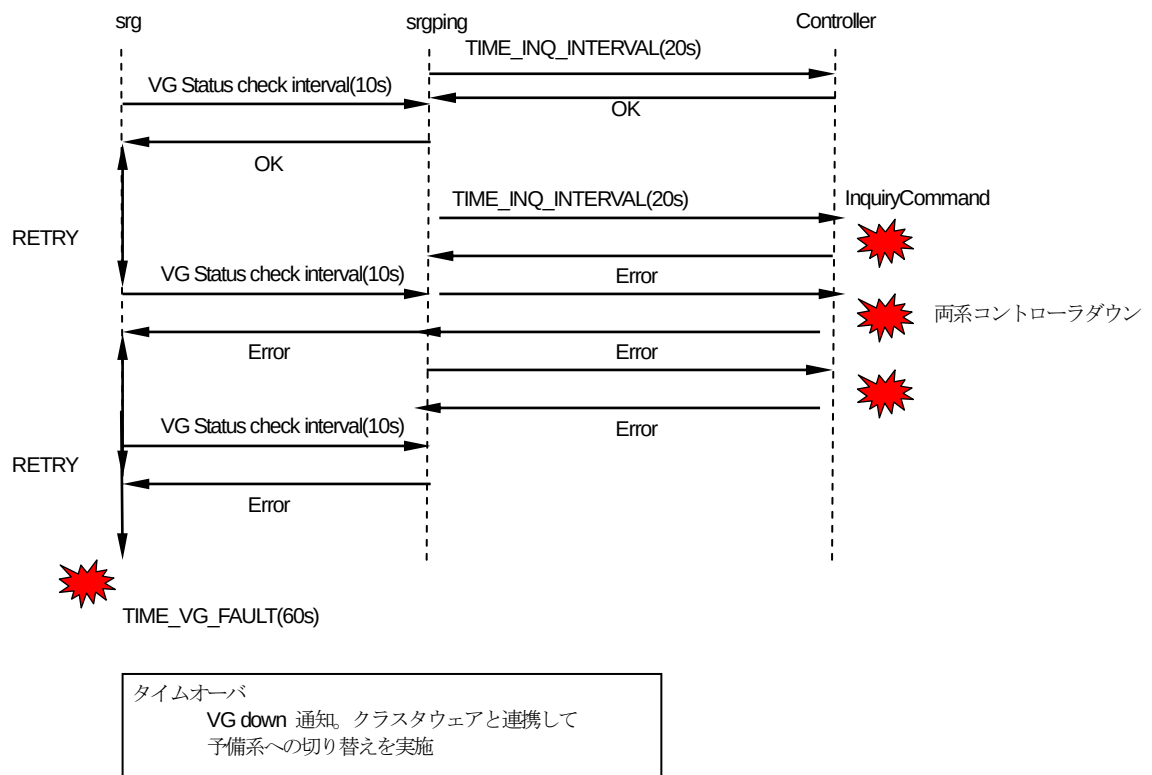
タイムオーバ
障害の発生した FC 配下 VG suspend 通知

- (3) I/O パスの死活監視で両系障害を検出
冗長化された I/O パスを構成するディスク装置コントローラ、
FC スwitch、FC カードのいずれかの部品の両系が故障した場合、予備ノードへ切り替えます。



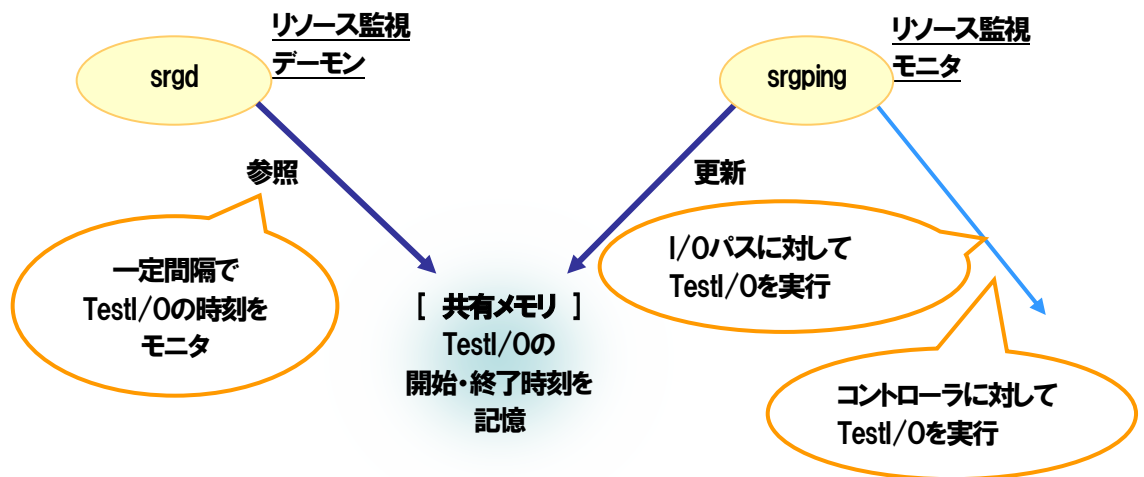
- TestI/O で両系コントローラの異常を検出すると以下のような動作になります。

VG 障害検出時間	: TIME_VG_FAULT	: 60(秒)
I/O パスの障害検出時間	: TIME_LINKDOWN	: 180(秒)
コントローラ監視間隔	: TIME_INQ_INTERVAL	: 20(秒)



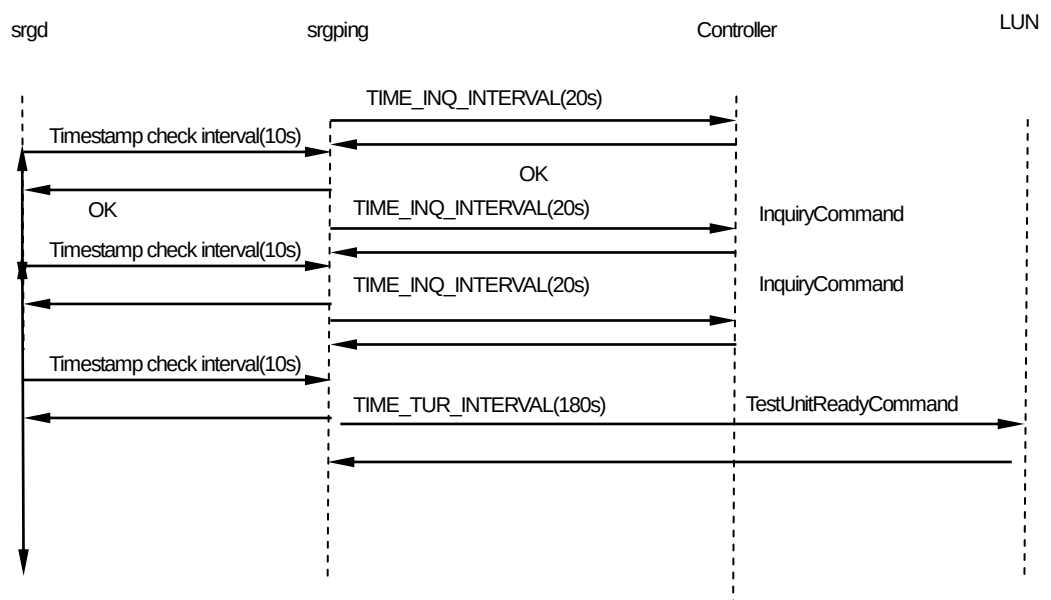
- (4) I/O リクエストのストール監視
 I/O パスに対して定期的に発行する TestI/O の実行時刻を検査することで、
 OS 全体のストール状態を監視します。
 TestI/O が一定時間以内に正常完了しなければ、I/O パスを異常と判定します。

【I/O ストール監視のフレームワーク】

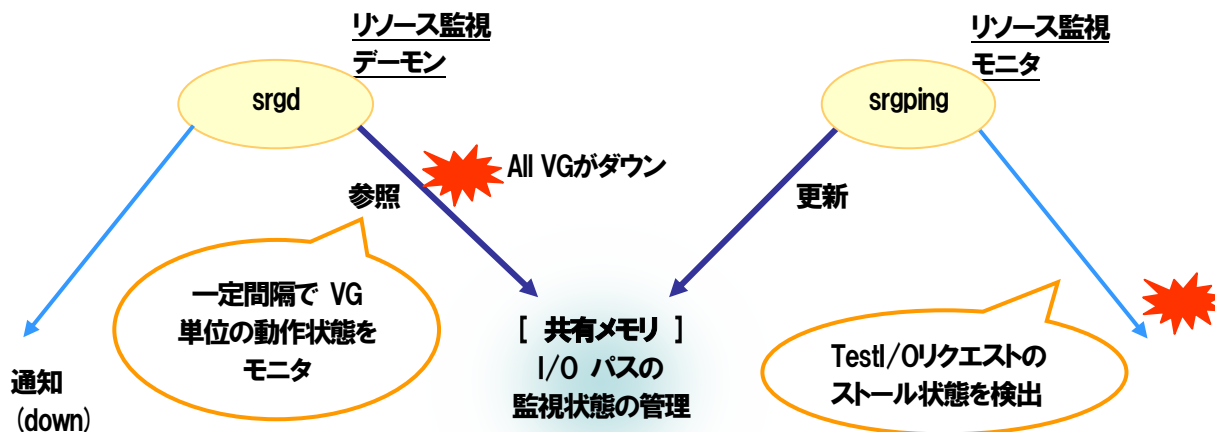


- TestI/O の正常なシーケンスは、以下のような動作になります。

I/O パスのストール監視時間	:	TIME_VG_STALL	:	360(秒)
コントローラ監視間隔	:	TIME_INQ_INTERVAL	:	20(秒)
LUN 監視間隔	:	TIME_TUR_INTERVAL	:	180(秒)



【I/O ストールを検出】

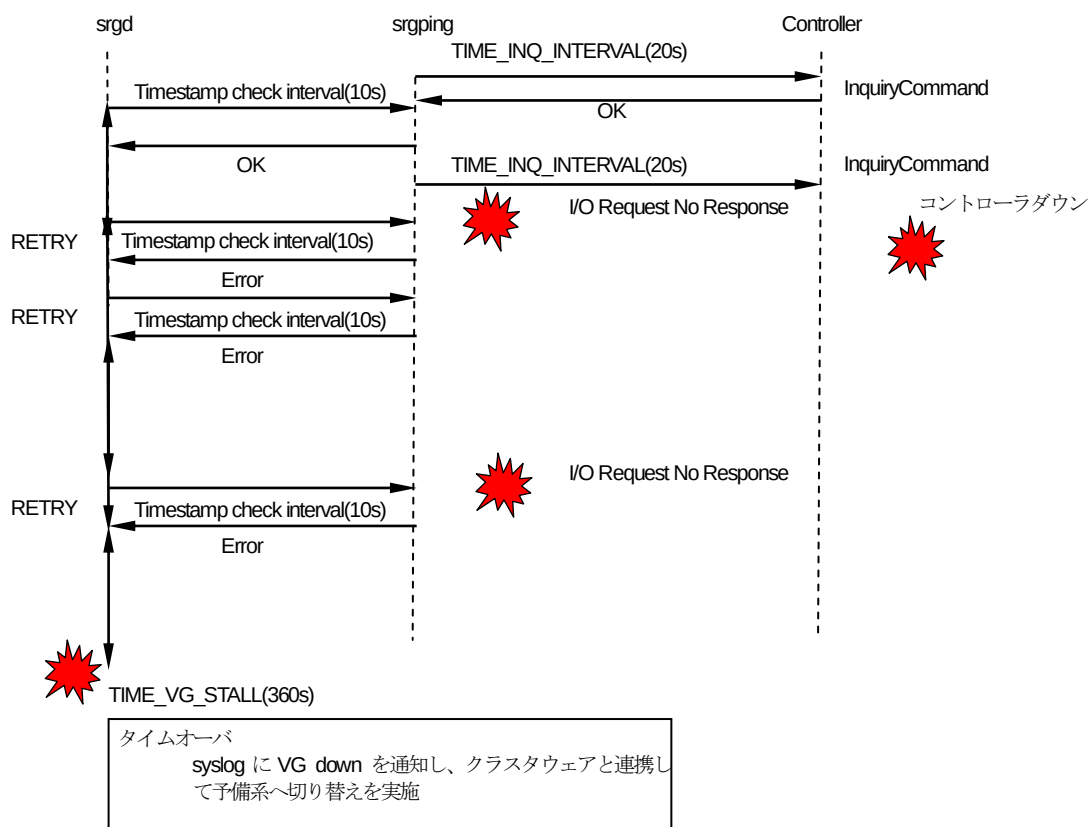


- Test/I/O で I/O ストールを検出すると、以下のような動作になります。

```

I/Oパスのストール監視時間 : TIME_VG_STALL : 360(秒)
コントローラ監視間隔 : TIME_INQ_INTERVAL : 20(秒)
LUN 監視間隔 : TIME_TUR_INTERVAL : 180(秒)

```



2.3.リソース監視で異常を検出すると

TestI/O で I/O パスの異常を検出すると、コンフィグレーションで指定されたオペレーションを実行します。

TestI/O で検査できる監視項目は下記のとおりです。

- I/O パスの死活監視で異常を検出
- I/O リクエストのストール状態を検出
- LUN へのアクセス不可を検出

(1) TestI/O で I/O パス死活監視の異常を検出

TestI/O に対して異常応答、タイムアウトを検出した場合、以下のオペレーションを選択できます。

- I/O パスを閉塞する
- I/O パスを閉塞しない

注意:I/O パスの自動閉塞は、SPS または PowerPath でパスを冗長化している場合にのみ可能です。SPS または PowerPath を導入していない場合には、自動閉塞機能は利用できません。

(2) TestI/O で I/O リクエストのストールを検出

TestI/O に対して、OS レベルで無応答を検出すると I/O ストール状態と判定します。
I/O リクエストが I/O ストール監視時間(デフォルト 360 秒)以内に終了しない場合
以下のオペレーションを選択できます。

- ノードを切り替える
- ノードを切り替えない

(3) TestI/O で LUN へのアクセス不可を検出

LUN への TestI/O に対して異常応答を検出した場合、以下のオペレーションを選択できます。

- ノードを切り替える
- ノードを切り替えない

発生要因として以下の障害が考えられます。

- すべての I/O パス(全経路)で障害を検出
 - 全 FC スイッチ障害
 - 全 FC カード障害
 - 全 SCSI カード障害
 - ディスク装置本体の故障
- ソフトミラー構成で両系ディスク障害を検出
 - ディスク装置本体の故障

2.4. アクションの定義について

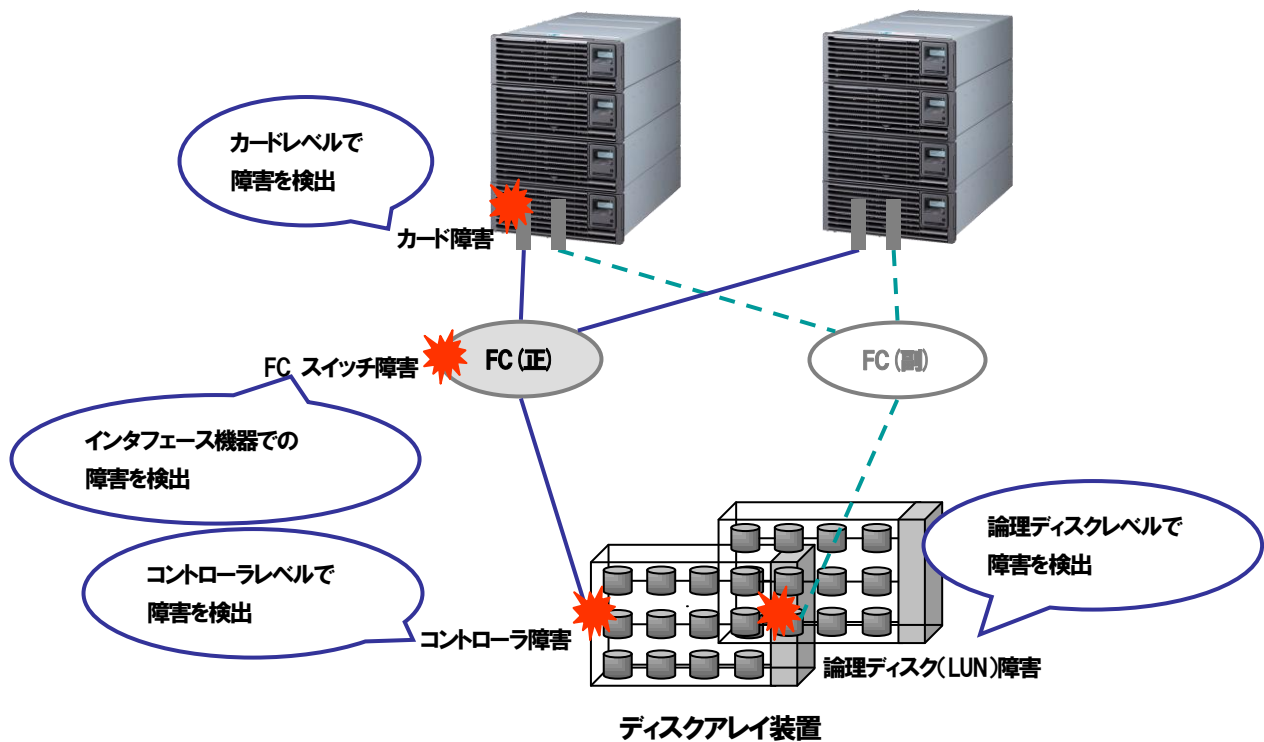
TestI/O で異常を検出した場合、下記のアクションを指定できます。

- I/O パスを自動閉塞する
- ノードを切り替える(クラスタウェア連動)

(1) アクションを選択しない場合

アクションを選択しない場合でも、syslog ファイルに障害メッセージを出力します。

ディスクアレイ装置コントローラ、論理ディスクに対して定期的に TestI/O を発行することにより、ディスクアレイ装置、インタフェース機器の故障、間欠障害を検出し、障害情報をレポートします。



(2) I/O パスを自動閉塞する

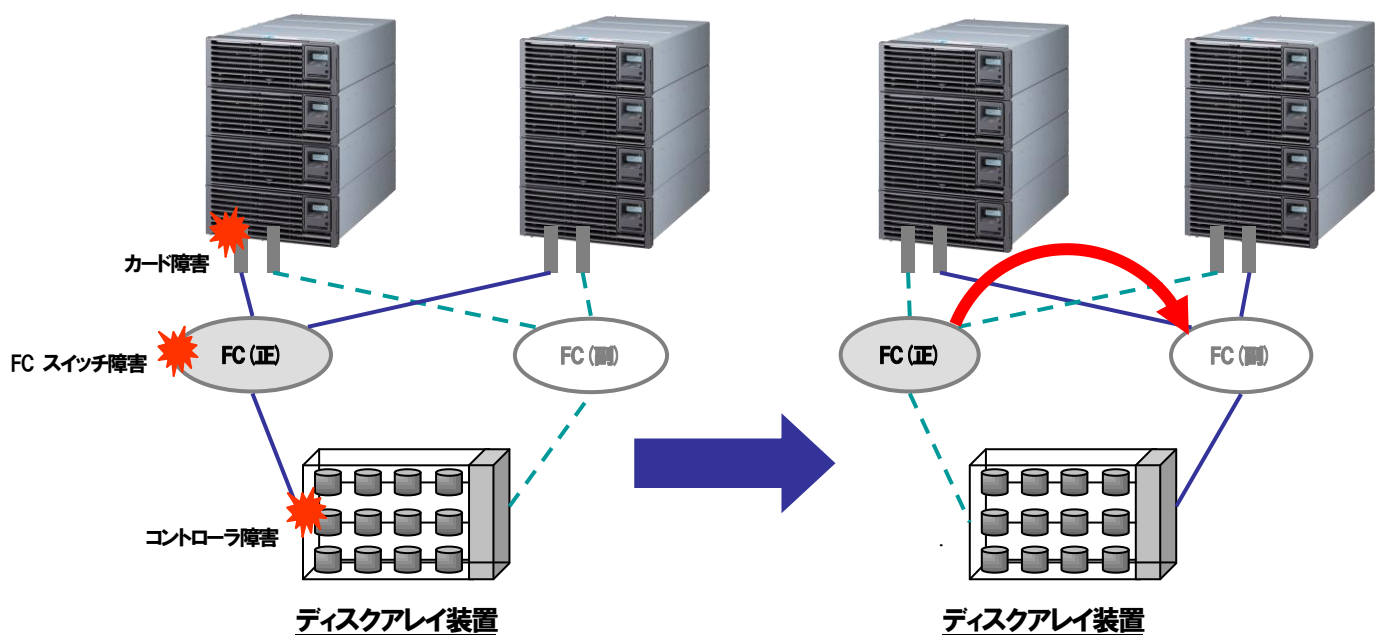
FC インタフェース上でリンクダウンが断続的に発生すると I/O パスの切り替えが多発し、ユーザー I/O のリトライにより I/O 遅延が発生します。

この機能はリンクダウン等の障害を検出した I/O パスを速やかに FC レイヤから切り離すことで、正常な I/O パスでの運用に切り替えます。

ディスクアレイ装置の信頼性を向上させるために、

StoragePathSavior を使って I/O パスを冗長化する手法があります。

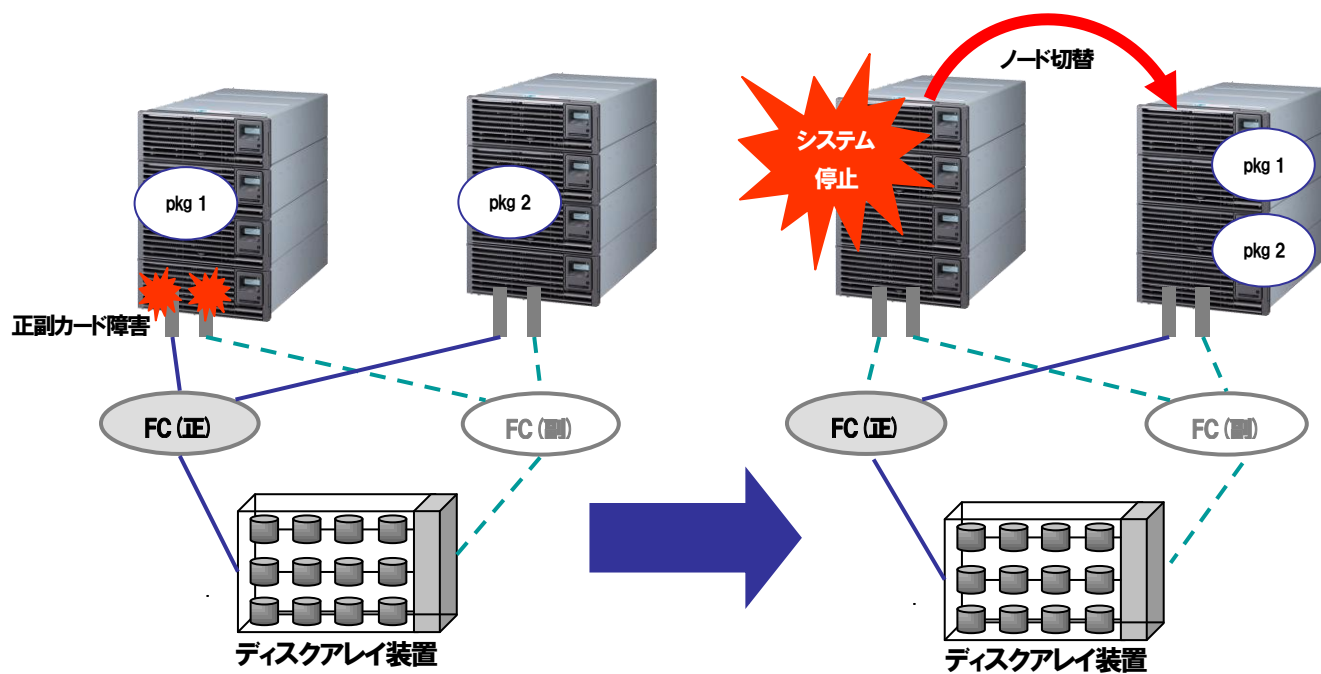
代替パス構成では、リンクダウンや機器故障により間欠障害が発生すると、I/O が遅延する問題を含んでいますが、障害箇所を早期に特定し故障箇所を切り離すことで、業務プロセスの I/O 遅延を防止します。



(3) ノードを切り替える

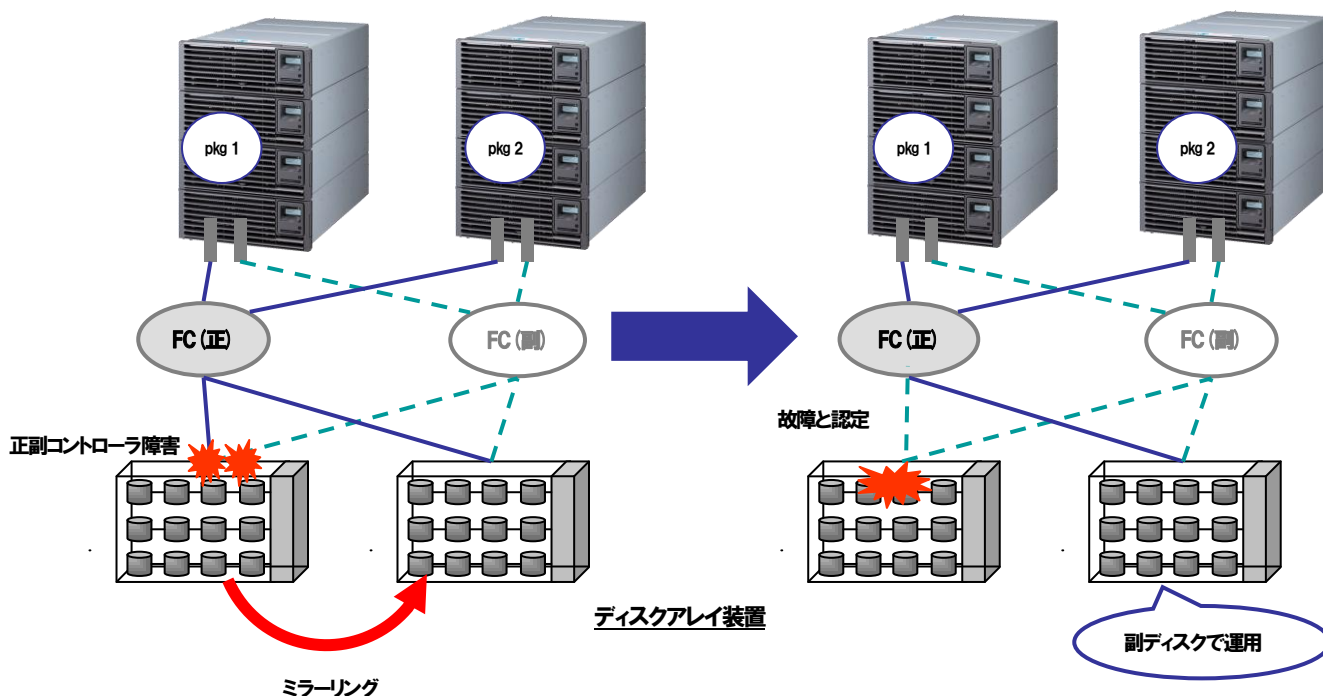
この機能は FC カードやスイッチ等の二重故障でディスク装置が利用できなくなり、業務の続行が不可能な状況に陥った場合に実行中の業務を待機ノードに切り替えます。クラスタウェアにより構築されたクラスタシステムで利用可能です。

共有ディスクへの I/O 処理が不可能になると（両系カード故障、両系 FC スイッチ故障）、待機ノードへ切り替えることで、業務を継続します。

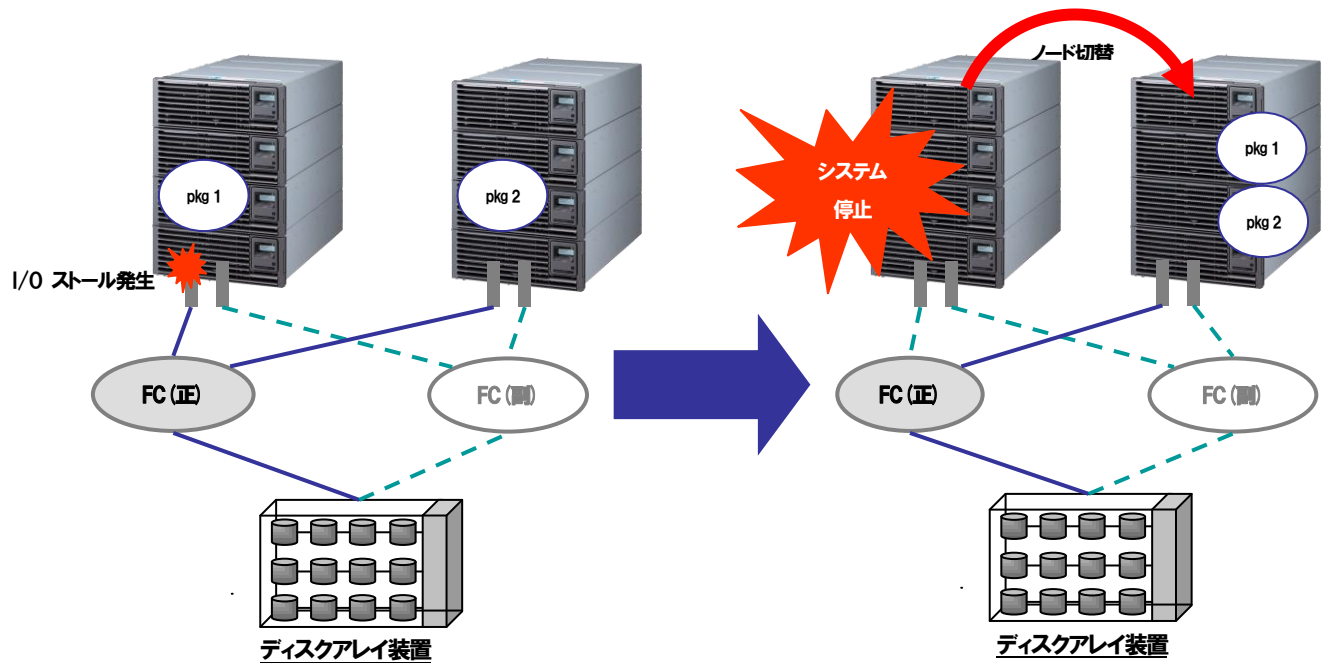


共有ディスクが Linux mdデバイス管理ツールでソフトウェアミラーリングされている場合に、片系のディスクが使用不可になると、故障箇所を切り離して、他系のディスクで I/O を継続します。なお、ミラーを構成するすべてのディスクが使用不可になると、待機ノードへ切り替えることで業務を継続します

注意:現時点で本機能は非サポートです。次期バージョンでサポート予定です。



共有ディスクへの I/O リクエストが一定時間経過しても完了しない場合は、
I/O ストール状態と判断し、待機ノードへ切り替えることで、業務を継続します。



2.5.オンライン保守機能

(1) オンライン保守機能の目的

本製品は、Linux に接続されたディスク装置のオンライン保守を円滑に行うために専用のコマンドを提供します。

マルチパス管理製品で冗長化された I/O パスに対して、Linux OS、業務ソフトウェアを停止することなくディスク装置の故障個所の特定、交換、FW update といったオンライン保守の作業環境を提供します。

(2) オンライン保守ユーティリティの機能

本製品の提供する機能は以下の通りです。

- ・ FC 単位の一括閉塞機能

マルチパス構成の I/O パスを FC 単位で一括閉塞することでディスク装置の特定のコントローラへの I/O をブロックします。

この間、冗長化された残りの I/O パスでの read、write が保証されるため業務ソフトウェアを停止することなく、ディスク装置のオンライン保守が可能となります。

また、FC カードやコントローラの HW 障害が発生した場合に、オペレータ介入により I/O パスを一括閉塞することで、速やかに正常系の I/O パスへの切り替えを実現します。

- ・ FC 単位の一括復旧機能

閉塞した I/O パスを、FC 単位で一括復旧します。

- ・ PV 単位の一括閉塞機能

マルチパス構成の I/O パスを PV 単位で一括閉塞することで特定のディスクへの I/O をブロックします。

ディスク障害が発生した場合に、オペレータ介入により I/O パスを一括閉塞することで、速やかなディスク交換作業を実現します。

- ・ PV 単位の一括復旧、構成復旧

閉塞した I/O パスを、PV 単位で一括復旧します。

- ・ I/O パスの構成復旧機能

すべての I/O パスに対して構成復旧を実行します。

- ・ 状態表示機能、TestI/O 機能

I/O パスの運用状態を FC 単位、PV 単位で表示します。

コマンドベースで TestI/O を発行することで、I/O パスの稼動状態を知ることができます。

注意:オンライン保守機能は、SPS または PowerPath でパスを冗長化している場合にのみ可能です。SPS または PowerPath を導入していない場合には、オンライン保守機能は利用できません。

2.6.RENS との連携機能について

RENS が導入されたシステムの場合、RENS と連携して StorageSaver の監視リソースの状態を他の製品へ通知することが可能です。

例えば、RENS のリソース通知をサポートしているクラスタウェアなどはこの RENS 連携機能を利用して、StorageSaver の監視しているリソースの状態に応じてフェイルオーバーの実行などが可能となります。

RENS 連携を行う場合、RENS 連携用モニタプロセスを別途インストールする必要があります。

RENS 連携の詳細については "6 章 RENS との連携" を参照してください。

2.7.クラスタウェアとの連携機能について

本製品の提供するクラスタ連携機能を導入することで、クラスタウェアで構築したクラスタシステムで高速なノード切り替えが実現できます。

以降、本製品ではクラスタウェア製品として、CLUSTERPRO を例に説明します。

本体系障害、すべてのインタフェースカード障害、すべての FC スイッチ障害において有効ですが、ノードを切り替える手段として以下の4通りの手法があります。

- ① CLUSTERPRO のカスタムモニタリソースにクラスタウェア連携デーモン(srgstat)を登録する方式
- ② CLUSTERPRO のサーバ管理プロセス (clpnm) を強制終了し、ノードを切り替える方式
- ③ システムメモリダンプを採取し、強制的に OS を停止し、ノードを切り替える方式
- ④ RENS 経由で CLUSTERPRO へリソース状態を通知し、ノードを切り替える方式

具体的な連携の設定手順については、後述の "CLUSTERPRO との連携" を参照してください。

3. 設定ファイル

3.1. 本製品の導入

(1) インストール

ストレージシステムの監視を行うには、StorageSaver のインストールが必要です。

- RPM パッケージ名
clusterpro-mc-ss-w.x.y-z.i386.rpm

- インストール

```
# rpm -ivh /mnt/cdrom/Linux/rpm/clusterpro-mc-ss-w.x.y-z.i386.rpm
```

インストールが完了した場合以下のコマンドでインストールの確認を行ってください。

```
# rpm -qa | grep clusterpro-mc-ss  
clusterpro-mc-ss-w.x.y-z
```

注意: Red Hat Enterprise Linux 6.x (64bit) または Oracle Linux 6.x (64bit) のシステムに
本製品をインストールする場合、事前に互換ライブラリ(glibc-x-y.el6.i686.rpm)が
インストールされている必要があります。

- アンインストール

```
# rpm -e clusterpro-mc-ss-w.x.y-z
```

注意: w, x, y, z にはバージョン番号が入ります。
機能強化があるとバージョン番号が更新されます。

※インストール手順についての詳細は、『CLUSTERPRO MC StorageSaver 1.1 for Linux リリースメモ』
をご覧ください。

(2) セットアップ

ディスク装置を監視するには、設定ファイルの作成が必要です。

設定ファイルは `/var/opt/HA/SrG/conf` 配下に作成します。

ファイル名は以下のとおりで、サンプルファイルが `/var/opt/HA/SrG/conf/sample` 配下に提供されています。

- システム定義ファイル (`srg.config`)
ノード一意で使用する設定を定義したファイルです。
- 構成定義ファイル(`srg.map`)
I/O パス情報の論理構成を定義したファイルです。
- リソース定義ファイル(`srg.rsc`)
HBA カード(FC/SCSI カード)および I/O パスのリソース情報を定義したファイルです。

設定ファイル自動生成コマンド `/opt/HA/SrG/bin/srgquery(1M)` を利用するとデバイス情報を検索し設定ファイルのテンプレートを自動生成できます。

3.2. 設定ファイルの自動生成手順

srgquery(1M) による設定ファイル自動生成の手順を説明します。

- (1) はじめに
クラスタウェア のクラスタ環境構築(LVM の VG や LV、ソフトミラーの構築)が完了している場合には、srgquery(1M) により設定ファイルを自動生成することができます。
特に、ソフトミラー構成を採用する場合は、LV の設定が完了していなければミラーを構成する PV の組み合わせを取得できません。

また、クラスタウェアを利用しないシングルノードの場合でも LVM の VG や LV、ソフトミラーの構築を完了してから srgquery(1M) により設定ファイルを自動生成してください。

なお、LVM を利用しないディスク構成の場合も srgquery(1M)により設定ファイルを自動生成可能です。この場合もディスク構成等の設定が完了している状態で実行してください。

注意: 本バージョンではソフトミラー構成の設定ファイル自動生成は非サポートです。
次期バージョンでサポート予定です。

- (2) srgquery(1M) による設定ファイル自動生成
設定ファイルは、srgquery(1M) により自動的に作成されます。
一般的に使用する引数は下記のいずれかです。

- FC 接続の場合

```
# /opt/HA/SrG/bin/srgquery -s <格納ディレクトリ>
```

- SCSI 接続の場合

FC 接続に加え SCSI 接続のディスク装置を対象にした場合

```
# /opt/HA/SrG/bin/srgquery -a -s <格納ディレクトリ>
```

注意:

- 仮想環境 (ゲスト OS) では、srgquery コマンドに -a オプションを指定する必要があります。また、自動生成された設定ファイル(srg.config) の TESTIO_MODE に READ を設定する必要があります。
- srgquery コマンドにて設定ファイルの自動生成を行った際、マルチパス管理製品 (SPS, PowerPath) による I/O パス冗長設定が完了していない可能性がある場合、以下の警告メッセージを出力します。

```
srgquery: <デバイスファイル> is not found in StoragePathSavior.  
srgquery: <デバイスファイル> is not found in PowerPath.
```

マルチパス管理製品 (SPS, PowerPath) の I/O パス冗長設定が完了していない場合は設定完了後、再度 srgquery コマンドにて設定ファイルの自動生成を行ってください。

- (3) 設定ファイルの確認、適用手順
設定ファイルを新規に作成、または変更した場合、srgconfig(1M) コマンドによりその妥当性および相関関係を確認した後にシステムに適用してください。

設定ファイルを実行環境に適用した場合は、デーモンプロセスの再起動が必要です。

1. 設定ファイルの妥当性の確認手順

```
# /opt/HA/SrG/bin/srgconfig -c -s <確認対象設定ファイルの格納ディレクトリ>
```

注意: 本バージョンでは srg.config ファイルの妥当性チェック機能は非サポートです。
次期バージョンでサポート予定です。

2. 設定ファイルの実行環境への適用手順

```
# /opt/HA/SrG/bin/srgconfig -a -s <確認対象設定ファイルの格納ディレクトリ>
```

3. デーモンプロセスの再起動

```
# /opt/HA/SrG/bin/srgconfig -r
```

- (4) H/W 構成を変更する際の手順
FC 接続構成や LUN 構成等、H/W 構成を変更する場合は、設定ファイルの再作成および適用操作を行う必要があります。
以下の手順を実行してください。

1. デーモンプロセスの自動起動を抑制する

```
# touch /var/opt/HA/SrG/conf/srg.ignore
```

2. OS 停止

3. H/W 構成変更

4. OS 起動

5. 設定ファイルを再作成する

```
# /opt/HA/SrG/bin/srgquery -s /tmp
```

6. 設定ファイルの整合性をチェックする

```
# /opt/HA/SrG/bin/srgconfig -c -s /tmp
```

7. 設定ファイルを実行環境に適用する

```
# /opt/HA/SrG/bin/srgconfig -a -s /tmp
```

8. デーモンプロセスの自動起動抑制を解除する

```
# rm /var/opt/HA/SrG/conf/srg.ignore
```

(5) 注意事項

- ・ 設定ファイルのバックアップについて
設定ファイル(/var/opt/HA/SrG/conf 配下)は、バックアップすることをお奨めします。再インストールする場合の復旧手順が容易となります。
- ・ 設定ファイルの更新時の注意
接続されているディスクの構成や LVM の VG や PV 構成を変更した場合、また、監視ポリシーを変更した場合は、設定ファイルの再作成、デーモンプロセス再起動が必要です。
- ・ 設定ファイル自動生成のサポート構成について
設定ファイル自動生成機能は設定作業軽減のため、設定ファイルのテンプレートを作成する機能であり、すべてのディスク構成をサポートしているわけではありません。
したがって、構成によっては作成できない場合もあります。必ず作成された設定ファイルを確認し、実際の環境と差異がある場合には手動で修正してください。

また、本バージョンではソフトミラー構成の設定ファイル自動生成は非サポートです。

- ・ EMC 社製 CLARiX シリーズ が接続されている場合の注意
CLARiX シリーズは、LUN への TestI/O がサポートされていないので、TestUnitReady を発行しないようにシステム定義ファイル(/var/opt/HA/SrG/conf/srg.config)の TESTIO_MODE 値を INQ に設定してください。

注意:srgquery(1M) で設定ファイル自動生成を行った際、監視対象に CLARiX シリーズが含まれている場合は、自動で設定ファイルの TESTIO_MODE 値を INQ に設定します。

```
#####  
# StorageSaver #  
# system configuration file for StorageSaver #  
#####  
:  
# TestI/O mode  
# select INQ, INQ_TUR_READ, READ, DIRECT, INQ_TUR(default)  
TESTIO_MODE INQ  
:
```

- ・ NEC 社製ストレージ iStorage M シリーズを使用する場合の注意
iStorage M シリーズに対して TestUnitReady を発行することができません。
TestUnitReady を発行しないようにシステム定義ファイル(/var/opt/HA/SrG/conf/srg.config)の TESTIO_MODE 値を INQ に設定してください。

※ 詳細については『CLUSTERPRO MC StorageSaver iStorage M シリーズ使用時の設定手順』をご覧ください。

注意:StorageSaver 1.0.0-2 以降では、srgquery(1M) で設定ファイル自動生成を行った際に、自動でシステム定義ファイルの TIME_TUR_INTERVAL 値を 0 に設定します。
StorageSaver 1.1 以降ではパラメータ変更に伴い TESTIO_MODE 値を INQ に設定します。TIME_TUR_INTERVAL 値はデフォルト値(180)のままとなります。

- SANboot 環境において設定ファイル自動生成コマンド (srgquery) にて設定ファイルを生成した場合、OS 情報が格納されたディスク装置も FC 接続構成であるため、他のデータディスク装置同様、監視対象として定義します。設定ファイルから OS 情報が格納されたディスク装置の記述を手動で削除してください。
- IDE 接続のデバイスが接続されている環境において設定ファイル自動生成コマンド (srgquery) にて設定定義ファイルを生成した際、IDE 接続のデバイスを監視対象として定義する場合があります。
IDE 接続のデバイスが設定ファイルに含まれている場合は手動で削除してください。
- 仮想環境 (ゲスト OS) において設定ファイルを生成する場合、srgquery コマンドに `-a` オプションを指定する必要があります。
また、自動生成されたシステム定義ファイル (/var/opt/HA/SrG/conf/srg.config) の `TESTIO_MODE` に `READ` を設定する必要があります。

3.3. 設定ファイルの記述

- (1) システム定義ファイル の設定について
設定ファイル名は以下のとおりです。

/var/opt/HA/SrG/conf/srg.config

以下に使用するキーワードを記述します。

項目	説明
TIME_VG_FAULT	ボリュームグループの障害検出時間を指定します。 各ボリュームグループ配下の I/O パスに対する TestI/O が失敗し始めてから VG リソースを異常と判定する時間を指定します。 このパラメータはデフォルト値を使用することを推奨します。 最小値は 30 秒、デフォルト 60 秒
TIME_LINKDOWN	I/O パスの障害検出時間を指定します。 TestI/O が失敗し始めてから I/O パス障害と判定する時間を指定します。 最小値は 30 秒、デフォルト 180 秒
TIME_INQ_INTERVAL	コントローラ監視間隔を指定します。 ディスク装置コントローラへの TestI/O インターバルを指定します。 FC カード、インタフェース機器、ディスク装置コントローラの障害検出時間を短縮したい場合は、本パラメータを調整してください。 最小値は 10 秒、デフォルト 20 秒
TIME_TUR_INTERVAL	LUN 監視間隔を指定します。 ディスク装置論理ディスクへの TestI/O インターバルを指定します。 LUN の障害検出時間を短縮したい場合は、本パラメータを調整してください。 なお、指定値は TIME_INQ_INTERVAL で指定した値よりも大きく、かつ、整数倍である必要があります。 また、LUN の監視が不要な場合は、0 秒を指定すると論理ディスクへの TestI/O は行いません。 最小値は 0 秒、デフォルト 180 秒
TIME_READ_INTERVAL	LUN データ読込間隔を指定します。 ディスク装置論理ディスクへの Read TestI/O のインターバルを指定します。 本パラメータは TESTIO_MODE に READ が指定された場合に有効となり、本設定値の間隔で設定ファイルに定義されたデバイスに対して Read TestI/O が発行されます。 LUN の障害検出時間を短縮したい場合は、本パラメータを調整してください。 最小値は 10 秒、デフォルト 180 秒

TESTIO_FAULT_ACTION	TestI/O 異常検出時のアクションを指定します。
ACTION_NONE	アクション指定なし。デフォルトです。 syslog ファイルに障害情報を通知します。
BLOCK_PATH	I/O パスを自動閉塞します。 障害が発生した I/O パスの切り離しを行います。
VG_FAULT_ACTION	ボリュームグループ異常検出時のアクションを指定します。
SERVICE_CMD_DISABLE	アクション指定なし。デフォルトです。 syslog ファイルに障害情報を通知します。
SERVICE_CMD_ENABLE	クラスタウェア連携用コマンド(srgstat)を使用してクラスタウェアと連携することでノードを切り替えます。
CLPNM_KILL_ENABLE	clpnm(1M) を 強制終了させます。 CLUSTERPRO のサーバ管理プロセス (clpnm) を強制終了することでノードを切り替えます。
TOC_EXEC	システムメモリダンプを採取し、OS を強制停止することでノードを切り替えます。
POWER_OFF	ソフトウェア watchdog を利用し、OS を停止します。
RENS_REPORT_ENABLE	RENS に障害イベントを通知します。RENS を利用したクラスタ連携を行う場合にはこの値を設定します。 RENS 連携を行う場合、別途 RENS 連携用モニタプロセスをインストールする必要があります。詳細は後述の「RENS との連携」の章を参照してください。
AUTO_RECOVERY	監視ステータス自動復旧の実行可否を指定します。 障害発生時、故障箇所が障害状態から復旧したことをオペレータが確認した上で監視ステータスを復旧して頂くため、本パラメータは DISABLE に設定することを推奨しています。 運用上オペレータによる確認が困難である場合は、本パラメータを ENABLE に設定することで自動復旧機能を利用することができます。
ENABLE	自動復旧を行う。 障害状態から復旧した I/O パスを自動的に組み込み、監視を再開します。
DISABLE	自動復旧を行わない(デフォルト)。 障害状態から復旧した場合にはオペレータによる確認、手動での復旧が必要です。

TESTIO_MODE	TestI/O の発行方法を指定します。
INQ	パススルードライバ経由で Inquiry を発行します。
INQ_TUR	パススルードライバ経由で Inquiry と TestUnitReady を発行します。 デフォルトです。
INQ_TUR_READ	パススルードライバ経由で Inquiry と TestUnitReady と Read を発行します。
DIRECT	システム定義ファイルに定義されたデバイス H/W パスに対応するスペシャルファイルに対して直接 Inquiry と TestUnitReady を発行します。
READ	システム定義ファイルに定義されたデバイス H/W パスに対応するスペシャルファイルに対して直接 read を発行します。 仮想環境 (ゲスト OS) で使用する場合は、この値を指定します。
TIO_MONITOR	間欠障害監視機能の使用可否を指定します。 デフォルトではコメントアウトされています。 間欠障害監視機能を使用する場合にのみ本パラメータを有効にしてください。
DISABLE	間欠障害監視機能を使用しません。 デフォルトです。
ENABLE	間欠障害監視機能を使用します。
TIO_FILE_PATH	間欠障害監視機能のログ出力先を指定します。 デフォルトではコメントアウトされています。 間欠障害監視機能を使用する場合にのみ本パラメータを有効にしてください。
TIO_FILE_SIZE	間欠障害監視機能のログファイルのサイズを Mbyte 単位で指定します。 デフォルトではコメントアウトされています。 間欠障害監視機能を使用する場合にのみ本パラメータを有効にしてください。 最小値は 1Mbyte、デフォルトは 1Mbyte

注意:

- ・ CLUSTERPRO を導入し、クラスタを起動している場合は CLUSTERPRO によるソフトウェア watchdog を利用した OS 停止機能を優先するため、POWER_OFF 機能を利用することができません。
同等の機能が CLUSTERPRO 側に実装されているので、そちらの機能を利用してください。
また、CLUSTERPRO と他の方式にて連携を行う場合は “5 CLUSTERPRO との連携” を参照してください。
- ・ TOC_EXEC 機能、POWER_OFF 機能を利用する場合は事前に kdump の設定を行い、SysRq キーを発行することによりカーネルパニックが発生することを確認してください。
- ・ POWER_OFF 機能を利用する場合は softdog モジュールが必要となります。
事前に softdog モジュールがインストールされていることを確認してください。
- ・ POWER_OFF 機能を利用する場合、ソフトウェア watchdog を利用した OS 停止に失敗した際には自動的に TOC_EXEC 機能にて OS を停止させます。
- ・ 仮想環境 (ゲスト OS) で本製品を使用する場合は、TESTIO_MODE に READ を設定してください。
- ・ 上記タイマ値の上限値は MAXINT まで指定可能ですが、常識的な運用での適用を推奨します。

これ以降のパラメータは変更できません。変更する場合は開発部門までお問い合わせください。

項目	説明
BASE_TIMER	基本タイマを指定します。 最小値は 10 秒、デフォルト 10 秒
TIME_VG_STALL	監視リソースの I/O ストールを判定する時間を指定します。 このパラメータはデフォルト値を使用することを推奨します。 最小値は 60 秒、デフォルト 360 秒 なお、0 秒を指定すると I/O ストール監視を行いません。
VG_STALL_ACTION	I/O ストール検出時のアクションを指定します。 SERVICE_CMD_DISABLE アクション指定なし。 デフォルトです。syslog ファイルに障害情報を通知します。 SERVICE_CMD_ENABLE クラスタウェア連携用コマンド (srgstat) を使用してクラスタウェアと連携することでノードを切り替えます。 CLPNM_KILL_ENABLE clpnm(1M) を強制終了させます。 CLUSTERPRO のサーバ管理プロセス (clpnm) を強制終了することでノードを切り替えます。 TOC_EXEC システムメモリダンプを採取し、OS を強制停止することでノードを切り替えます。 POWER_OFF ソフトウェア watchdog を利用し、OS を停止します。 RENS_REPORT_ENABLE RENS に障害イベントを通知します。 RENS を利用したクラスタ連携を行う場合にはこの値を設定します。 RENS 連携を行う場合、別途 RENS 連携用モニタプロセスをインストールする必要があります。詳細は後述の「RENS との連携」の章を参照してください。
WAIT_TESTIO_INTERVAL	TestI/O でパススルードライバに指定する I/O 待ち合わせ時間を指定します。 このパラメータはデフォルト値を使用することを推奨します。 最小値は 1 秒、デフォルト 5 秒
DAILY_CHECK_TIME	障害の発生した I/O パスを定期通知する時刻を指定します。 指定値は 0~23、デフォルト 10(10:00)です。 自動復旧機能を使用する場合、ここで指定した時刻に自動復旧を行います。

TESTIO_USE	TestI/O 実行可否を指定します。 ENABLE TestI/O を行う(デフォルト) DISABLE TestI/O を行わない。
EXEC_SYNC_ENABLE	I/O パスの状態について定期的に同期を取るかを指定します。 ENABLE 定期同期を行う(デフォルト) DISABLE 定期同期を行わない。
SHM_BUFF_SIZE	リソーステーブルで使用する共有メモリサイズ です。 Mbyte 単位で指定します。指定値は 1～、デフォルト 2(Mbyte)です。
LOG_SIZE	リソース監視デーモンのログファイルのサイズ です。 Mbyte 単位で指定します。指定値は 1～40、デフォルト 20(Mbyte)です。

注意:

- ・ CLUSTERPRO を導入し、クラスタを起動している場合は CLUSTERPRO によるソフトウェア watchdog を利用した OS 停止機能を優先するため、POWER_OFF 機能を利用することができません。同等の機能が CLUSTERPRO 側に実装されているので、そちらの機能を利用してください。
また、CLUSTERPRO と他の方式にて連携を行う場合は“5 CLUSTERPRO との連携”を参照してください。
- ・ TOC_EXEC 機能、POWER_OFF 機能を利用する場合は事前に kdump の設定を行い、SysRq キーを発行することによりカーネルパニックが発生することを確認してください。
- ・ POWER_OFF 機能を利用する場合は softdog モジュールが必要となります。事前に softdog モジュールがインストールされていることを確認してください。
- ・ POWER_OFF 機能を利用する場合、ソフトウェア watchdog を利用した OS 停止に失敗した際には自動的に TOC_EXEC 機能にて OS を停止させます。
- ・ 上記タイマ値の上限値は MAXINT まで指定可能ですが、常識的な運用での適用を推奨します。

- (2) リソース定義ファイルの設定について
設定ファイル名は以下のとおりです。

/var/opt/HA/SrG/conf/srg.rsc

以下に使用するキーワードを記述します。

項目	説明										
FC	HBA 情報を定義します。 管理対象となる HBA カード情報の定義です。 SCSI 接続のカードの場合は、FC ではなく SCSI を指定します。 alias 名 FC を特定する任意の名称を指定します。 デフォルトは、fcx (x は1からの通番) FC HW パス FC の HW パスを指定します。 udev デバイスファイル名に記載されている HBA デバイスを識別するパスを指定します。 (例)udev デバイスファイル名の以下の斜体の部分 pci-<u>0000:0b:00.0</u>-fc-0x2001000013840322:0x0000000000000000										
PV	I/O パス情報を定義します。 FC 配下の I/O パス情報を定義します。 複数の I/O パスが存在する場合は、本パラメータを列記します。 ディスクタイプ ディスク種別を指定します。 <table> <tr> <td>iStorage_Series</td><td>NEC 社製 iStorage シリーズ</td></tr> <tr> <td>CLARiX_Series</td><td>EMC 社製 CLARiX シリーズ</td></tr> <tr> <td>Symmetrix_Series</td><td>EMC 社製 Symmetrix DMX シリーズ、 EMC 社製 Symmetrix VMAX シリーズ</td></tr> <tr> <td>SANRISE_Series</td><td>日立社製 SANRISE シリーズ 日立社製 Hitachi USP シリーズ 日立社製 Hitachi VSP シリーズ</td></tr> <tr> <td>Other</td><td>その他ディスク装置</td></tr> </table> デバイス HW path I/O パスの経路を示すデバイス HW パスを指定します。 デバイス HW パスは udev デバイスファイル名です。	iStorage_Series	NEC 社製 iStorage シリーズ	CLARiX_Series	EMC 社製 CLARiX シリーズ	Symmetrix_Series	EMC 社製 Symmetrix DMX シリーズ、 EMC 社製 Symmetrix VMAX シリーズ	SANRISE_Series	日立社製 SANRISE シリーズ 日立社製 Hitachi USP シリーズ 日立社製 Hitachi VSP シリーズ	Other	その他ディスク装置
iStorage_Series	NEC 社製 iStorage シリーズ										
CLARiX_Series	EMC 社製 CLARiX シリーズ										
Symmetrix_Series	EMC 社製 Symmetrix DMX シリーズ、 EMC 社製 Symmetrix VMAX シリーズ										
SANRISE_Series	日立社製 SANRISE シリーズ 日立社製 Hitachi USP シリーズ 日立社製 Hitachi VSP シリーズ										
Other	その他ディスク装置										

注意:FC パラメータと複数の PV パラメータの組み合わせをひとつのセットで指定してください。

(3) 構成定義ファイルの設定について

設定ファイル名は以下のとおりです。

/var/opt/HA/SrG/conf/srg.map

以下に使用するキーワードを記述します。

項目	説明
< 構成定義 >	
[リソース情報]	I/O パスの論理的な構成を指定します。
PKG	クラスタウェア で規定されたパッケージ名を指定します。 このパラメータには 64 文字以内のクラスター意の英数字を指定してください。 クラスタウェア の指定ではパッケージ名の長さの制約はありませんが、 64 文字を超える場合はパッケージ名の変更が必要です。 なお、パッケージ名が付与されていない VG を指定する場合は、パッケージ名なしである PKG_NONE を指定してください。
VG	LVM で規定された VG 名を指定します。64 文字以内を指定してください。 LVM を構成しないディスクについては管理上"VG_NONE"という仮想 VG 名で管理します。
RSC_ACTION	デフォルトでは使用しません。省略してください。 VG 単位で VG 障害発生時のアクションを変えたい場合に指定します。
SERVICE_CMD_DISABLE	アクション指定なし。 syslog ファイルに障害情報を通知します。
SERVICE_CMD_ENABLE	クラスタウェア連携用コマンド(srgstat)を使用してクラスタウェアと連携することでノードを切り替えます。
CLPNM_KILL_ENABLE	clpnm(1M) を強制終了させます。 CLUSTERPRO のサーバ管理プロセス (clpnm) を強制終了することでノードを切り替えます。
TOC_EXEC	システムメモリダンプを採取し、OS を強制停止することでノードを切り替えます。
POWER_OFF	ソフトウェア watchdog を利用し、OS を停止します。
RENS_REPORT_ENABLE	RENS に障害イベントを通知します。RENS を利用したクラスタ連携を行う場合にはこの値を設定します。 RENS 連携を行う場合、別途 RENS 連携用モニタプロセスをインストールする必要があります。詳細は後述の「RENS との連携」の章を参照してください。
(*) 指定可能な値は srg.config の VG_FAULT_ACTION に指定する値と同一です。 省略された場合は VG_FAULT_ACTION に指定されているアクションを実行します。	

FS_TYPE

I/O パスの管理方式を指定します。

SpsDevice	SPS でマルチパス管理している場合に指定します。
EmcDevice	PowerPath でマルチパス管理している場合に指定します。
HdlmDevice	HDLM でマルチパス管理している場合に指定します。
RawDevice	マルチパス管理製品で管理していない I/O パス、または SPS、PowerPath 以外のマルチパス管理製品を使用されている場合に指定します。

GROUP

任意の文字列を GROUP 名として指定します。

GROUP 名は groupxxxx (xxxx は 0001 ~ 9999) となるノード一意の数字です。

GROUP 定義には PV 定義が必須となります。

LVM でソフトミラー構成を定義している場合は、ミラー番号を指定します。

GROUP 名とミラー番号の間にはスペースが必要です。

ミラー番号は mirrorxxxx (xxxx は 0001 ~ 9999) となるノード一意の数字です。

LVM でミラー構成を定義していない、またはミラー構成であるがノード切り替えのアクションを使用しない場合は、ミラー番号を指定する必要はありません。

設定しても無視されます。

PV

経路を表す I/O パス情報をすべて指定します。

udev デバイスファイル名をデバイス HW パスとして指定します。

複数の I/O パスが存在する場合は、本パラメータを列記します。

注意:

- ・ CLUSTERPRO を導入し、クラスタを起動している場合は CLUSTERPRO によるソフトウェア watchdog を利用した OS 停止機能を優先するため、POWER_OFF 機能を利用することができません。同等の機能が CLUSTERPRO 側に実装されているので、そちらの機能を利用してください。
また、CLUSTERPRO と他の方式にて連携を行う場合は“5 CLUSTERPRO との連携”を参照してください。
- ・ TOC_EXEC 機能、POWER_OFF 機能を利用する場合は事前に kdump の設定を行い、SysRq キーを発行することによりカーネルパニックが発生することを確認してください。
- ・ POWER_OFF 機能を利用する場合は softdog モジュールが必要となります。事前に softdog モジュールがインストールされていることを確認してください。
- ・ POWER_OFF 機能を利用する場合、ソフトウェア watchdog を利用した OS 停止に失敗した際には自動的に TOC_EXEC 機能にて OS を停止させます。
- ・ 上記タイム値の上限値は MAXINT まで指定可能ですが、常識的な運用での適用を推奨します。

(4) 設定ファイルの設定例

[システム定義ファイル]

```
#####
# StorageSaver                                     #
# system configuration file for StorageSaver #
#####

#####
# User Config Area
#####

# TestI/O interval timer for vg is failed (seconds)
#     vg status changes fail between this timer
# minimum = 30, default = 60
TIME_VG_FAULT          60

# FC linkdown detected timer value (seconds)
# minimum = 10, default = 180
TIME_LINKDOWN          180

# TestI/O(Inquiry) interval timer value (seconds)
#     exec normal TestI/O for PV between this timer
# minimum = 10, default = 20.
TIME_INQ_INTERVAL      20

# TestI/O(TestUnitReady) interval timer value (seconds)
#     exec normal TestI/O for PV between this timer
# minimum = 0, default = 180. 0 mean TestUnitReady not run
TIME_TUR_INTERVAL      180

# TestI/O(Read) interval timer value (seconds)
#     exec normal TestI/O for PV between this timer
# minimum = 10, default = 180.
TIME_READ_INTERVAL     180

# TestI/O fault action
# select ACTION_NONE(default), BLOCK_PATH
TESTIO_FAULT_ACTION    ACTION_NONE

# vg fault action
# select SERVICE_CMD_DISABLE(default), SERVICE_CMD_ENABLE, CLPNM_KILL_ENABLE, TOC_EXEC, POWER_OFF
VG_FAULT_ACTION        SERVICE_CMD_DISABLE

# Auto recovery flag
#     used = ENABLE : unused = DISABLE(default)
AUTO_RECOVERY          DISABLE
```

```

# TestI/O mode
# select INQ, INQ_TUR_READ, READ, DIRECT, INQ_TUR(default)
TESTIO_MODE          INQ_TUR

#####
# For STSINFO
#####

# Monitor intermittent fault.
# monitor = ENABLE : not monitor = DISABLE (default)
#TIO_MONITOR    DISABLE

# Set the file path for log directory (for output).
# The path is full set.
#TIO_FILE_PATH  /var/opt/HA/SrG/log

# Size (M byte) of backup file.
# minmum = 1, default = 1
#TIO_FILE_SIZE  1

#####
# Development Config Area
# do not touch this field
#####

# srgping status check timer (seconds)
# default = 10
BASE_TIMER          10

# I/O stall interval timer for Volume Group is failed (seconds)
# Volume Group status changes fail between this timer
# minimum = 60, default = 360. 0 mean I/O stall nocheck.
TIME_VG_STALL       360

# Volume Group stall find action
# select SERVICE_CMD_DISABLE(default), SERVICE_CMD_ENABLE, CLPNM_KILL_ENABLE, TOC_EXEC, POWER_OFF
VG_STALL_ACTION      SERVICE_CMD_DISABLE

# Wait I/O for sg driver timer value (seconds)
# wait TestI/O between this timer
# minimum = 1, default = 5
WAIT_TESTIO_INTERVAL 5

# Daily check time for check pv status (o'clock)
# default = 10
DAILY_CHECK_TIME    10

# TestI/O use flag
# used = ENABLE(default) : unused = DISABLE
TESTIO_USE           ENABLE

```



```
# PV status sync flag
#      used = ENABLE(default) : unused = DISABLE
EXEC_SYNC_ENABLE      ENABLE

# Shared memory size (M byte)
# default = 2
SHM_BUFF_SIZE         2

# Trace log file size (M byte)
# default = 20
LOG_SIZE               20
```

[構成定義ファイル]

```
# srg.map
#####
# StorageSaver #
# LVM configuration file #
#####

# [FORMAT]
# PKG          pkg_name
# VG           vg_name
# RSC_ACTION    rsc_action
# FS_TYPE       fs_type
# GROUP         group0001
# PV           H/W Path
# PV           H/W Path
#
PKG      PKG_NONE
VG       VolGroup01
FS_TYPE  SpsDevice
GROUP    group0001
## PV Name: /dev/sdd
PV       pci-0000:13:00.0-fc-0x2100001697120ca7:0x0000000000000000
## PV Name: /dev/sdf
PV       pci-0000:13:00.0-fc-0x2900001697120ca7:0x0000000000000000
## PV Name: /dev/sdj
PV       pci-0000:13:00.1-fc-0x2100001697120ca7:0x0000000000000000
## PV Name: /dev/sdh
PV       pci-0000:13:00.1-fc-0x2900001697120ca7:0x0000000000000000

VG       VolGroup02
FS_TYPE  SpsDevice
GROUP    group0002
## PV Name: /dev/sde
PV       pci-0000:13:00.0-fc-0x2100001697120ca7:0x0001000000000000
## PV Name: /dev/sdg
PV       pci-0000:13:00.0-fc-0x2900001697120ca7:0x0001000000000000
## PV Name: /dev/sdk
PV       pci-0000:13:00.1-fc-0x2100001697120ca7:0x0001000000000000
## PV Name: /dev/sdi
PV       pci-0000:13:00.1-fc-0x2900001697120ca7:0x0001000000000000
```

[リソース定義ファイル]

```
# srg.rsc
#####
# StorageSaver #
# resource configuration file #
#####

# [FORMAT]
# [FC or SCSI] AliasName [FC Path or SCSI Path]
# PV      DiskType      H/W Path
# PV      DiskType      H/W Path
#
## FC Name: fc1
FC      fc1      0000:13:00.0
## PV Name: /dev/sdd
PV      iStorage_Series pci-0000:13:00.0-fc-0x2100001697120ca7:0x0000000000000000
## PV Name: /dev/sde
PV      iStorage_Series pci-0000:13:00.0-fc-0x2100001697120ca7:0x0001000000000000
## PV Name: /dev/sdf
PV      iStorage_Series pci-0000:13:00.0-fc-0x2900001697120ca7:0x0000000000000000
## PV Name: /dev/sdg
PV      iStorage_Series pci-0000:13:00.0-fc-0x2900001697120ca7:0x0001000000000000

## FC Name: fc2
FC      fc2      0000:13:00.1
## PV Name: /dev/sdj
PV      iStorage_Series pci-0000:13:00.1-fc-0x2100001697120ca7:0x0000000000000000
## PV Name: /dev/sdk
PV      iStorage_Series pci-0000:13:00.1-fc-0x2100001697120ca7:0x0001000000000000
## PV Name: /dev/sdh
PV      iStorage_Series pci-0000:13:00.1-fc-0x2900001697120ca7:0x0000000000000000
## PV Name: /dev/sdi
PV      iStorage_Series pci-0000:13:00.1-fc-0x2900001697120ca7:0x0001000000000000
```


4. 操作・運用手順

4.1. 運用管理コマンドの操作手順

(1) リソース監視の状態を表示します。

```
# /opt/HA/SrG/bin/srgadmin
(monitor status = TRUE)
```

type	device	HostBusAdapter	L status	P status	Online status
VG	VolGroup01	—	up		
PV	/dev/sdj	pci-0000:13:00.0	up	up	extended
PV	/dev/sdh	pci-0000:13:00.0	up	up	extended
PV	/dev/sdf	pci-0000:13:00.1	up	up	extended
PV	/dev/sdd	pci-0000:13:00.1	up	up	extended
VG	VolGroup02	—	up		
PV	/dev/sdk	pci-0000:13:00.0	up	up	extended
PV	/dev/sdi	pci-0000:13:00.0	up	up	extended
PV	/dev/sgd	pci-0000:13:00.1	up	up	extended
PV	/dev/sde	pci-0000:13:00.1	up	up	extended

# /opt/HA/SrG/bin/srgadmin (monitor status = TRUE)					
リソース監視の有効／無効を表示します					
type : device : HostBusAdapter : L status : P status : Online status					
VG : VolGroup01 : — : up : extended					
PV : /dev/sdj : pci-0000:13:00.0 : up : up : extended					
PV : /dev/sdh : pci-0000:13:00.0 : up : up : extended					
PV : /dev/sdf : pci-0000:13:00.1 : up : up : extended					
PV : /dev/sdd : pci-0000:13:00.1 : up : up : extended					
VG : VolGroup02 : — : up : extended					
PV : /dev/sdk : pci-0000:13:00.0 : up : up : extended					
PV : /dev/sdi : pci-0000:13:00.0 : up : up : extended					
PV : /dev/sdg : pci-0000:13:00.1 : up : up : extended					
PV : /dev/sde : pci-0000:13:00.1 : up : up : extended					

① VG の監視状態を表示します。

up	VG または PKG を構成するすべての I/O パスが正常に動作している状態です。
suspend	VG または PKG を構成する I/O パスの一部に異常があり、片パスで運用されている状態、または、保守員によるオンライン保守実施中のため、監視を停止している状態です。
down	VG または PKG を構成する I/O パスに異常があり、利用不可の LUN がある状態です。

② I/O パスの論理ステータス(管理状態)を L status として表示します。

表示	意味
up	正常動作中
down	障害状態
-	監視停止中

③ I/O パスの物理ステータス(TestI/O の実行状態)を P status として表示します。

表示	意味
up	正常動作中
down	障害状態
-	監視停止中

④ I/O パスの組み込み状態(Online status)を表示します。

表示	意味
extended	組み込み済み
reduced	閉塞状態
alive	状態不明
unknown	状態不明

(2) I/O パスが異常になるとステータスがダウン状態になります。

# /opt/HA/SrG/bin/srgadmin (monitor status = TRUE)					
type	device	HostBusAdapter	L status	P-stat	
VG	VolGroup01	—	<i>suspend</i>		
PV	/dev/sdj	pci-0000:13:00.0	<i>down</i>	<i>down</i>	extended
PV	/dev/sdh	pci-0000:13:00.0	up	up	extended
PV	/dev/sdf	pci-0000:13:00.1	up	up	extended
PV	/dev/sdd	pci-0000:13:00.1	up	up	extended
VG	VolGroup02	—	up		
PV	/dev/sdk	pci-0000:13:00.0	up	up	extended
PV	/dev/sdi	pci-0000:13:00.0	up	up	extended
PV	/dev/sdg	pci-0000:13:00.1	up	up	extended
PV	/dev/sde	pci-0000:13:00.1	up	up	extended

- ①、② TestI/O で片系の I/O 障害を検出すると、障害発生から TIME_LINKDOWN 秒後に、論理、物理ステータスが down になります。
- ③ さらに、VG の監視状態が suspend になります。

設定ファイルの TESTIO_FAULT_ACTION 値を BLOCK_PATH に設定している場合、I/O パスの異常検出時に自動で閉塞します。

# /opt/HA/SrG/bin/srgadmin (monitor status = TRUE)					
type	device	HostBusAdapter	L status	P-stat	
VG	VolGroup01	—	<i>suspend</i>		
PV	/dev/sdj	pci-0000:13:00.0	<i>down</i>	<i>down</i>	<i>reduced</i>
PV	/dev/sdh	pci-0000:13:00.0	up	up	extended
PV	/dev/sdf	pci-0000:13:00.1	up	up	extended
PV	/dev/sdd	pci-0000:13:00.1	up	up	extended
VG	VolGroup02	—			
PV	/dev/sdk	pci-0000:13:00.0	up	up	extended
PV	/dev/sdi	pci-0000:13:00.0	up	up	extended
PV	/dev/sdg	pci-0000:13:00.1	up	up	extended
PV	/dev/sde	pci-0000:13:00.1	up	up	extended

- ①、② TestI/O で片系の I/O 障害を検出すると、障害発生から TIME_LINKDOWN 秒後に、論理、物理ステータスが down になります。
- ③ さらに、VG の監視状態が suspend になります。
- ④ I/O パスの状態を示す Online status が閉塞状態である reduced になります。

注意:閉塞機能は SPS または PowerPath を導入している環境でのみ有効です。

(3) 両系の I/O パスが異常になると VG レベルのステータスもダウン状態になります。

```
# /opt/HA/SrG/bin/srgadmin
(monitor status = TRUE)
```

type	device	HostBusAdapter	L status	P status	
VG	VolGroup01	—	<u>down</u>	<u>down</u>	extended
PV	/dev/sdj	pci-0000:13:00.0	<u>down</u>	<u>down</u>	extended
PV	/dev/sdh	pci-0000:13:00.0	<u>down</u>	<u>down</u>	extended
PV	/dev/sdf	pci-0000:13:00.1	<u>down</u>	<u>down</u>	extended
PV	/dev/sdd	pci-0000:13:00.1	<u>down</u>	<u>down</u>	extended
VG	VolGroup02	—	up	up	extended
PV	/dev/sdk	pci-0000:13:00.0	up	up	extended
PV	/dev/sdi	pci-0000:13:00.0	up	up	extended
PV	/dev/sgd	pci-0000:13:00.1	up	up	extended
PV	/dev/sde	pci-0000:13:00.1	up	up	extended

③VG のステータス
①I/O パスの論理ステータス
②I/O パスの物理ステータス

①、②、③ TestI/O で両系の I/O 障害を検出すると、障害発生から TIME_VG_FAULT 秒後に、VG の監視状態が down になります。

(4) -i オプションを付与すると詳細情報を表示します。

```
# /opt/HA/SrG/bin/srgadmin -i
(monitor status = TRUE)
```

type	device	L status	P status	Online status
VG	VolGroup01	up	—	
PV	/dev/sdj	up	up	extended
	<u>pci-0000:13:00.0-fc-0x2100001697120ca7:0x0000000000000000</u>			
PV	/dev/sdh	up	up	extended
	pci-0000:13:00.0-fc-0x2900001697120ca7:0x0000000000000000			
PV	/dev/sdf	up	up	extended
	pci-0000:13:00.1-fc-0x2100001697120ca7:0x0000000000000000			
PV	/dev/sdd	up	up	extended
	pci-0000:13:00.1-fc-0x2900001697120ca7:0x0000000000000000			
VG	VolGroup02	up	—	
PV	/dev/sdk	up	up	extended
	pci-0000:13:00.0-fc-0x2100001697120ca7:0x0001000000000000			
PV	/dev/sdi	up	up	extended
	pci-0000:13:00.0-fc-0x2900001697120ca7:0x0001000000000000			
PV	/dev/sgd	up	up	extended
	pci-0000:13:00.1-fc-0x2100001697120ca7:0x0001000000000000			
PV	/dev/sde	up	up	extended
	pci-0000:13:00.1-fc-0x2900001697120ca7:0x0001000000000000			

①LUN への経路を示す udev デバイスファイル名

① LUN への経路を示す udev デバイスファイル名を表示します。

(5) リソース監視の停止と再開について

TestI/O の一時的な停止、再開をノード一意で指定します。

リソース監視停止中は、モニタステータスが FALSE になります。

```
# /opt/HA/SrG/bin/srgadmin -c stop
Change TESTIO.
START -> STOP
```

```
# /opt/HA/SrG/bin/srgadmin
(monitor status = FALSE)
```

type	device	HostBusAdapter	L status	P status	Online status
VG	VolGroup01	—	up		
PV	/dev/sdj	pci-0000:13:00.0	up	up	extended
PV	/dev/sdh	pci-0000:13:00.0	up	up	extended
PV	/dev/sdf	pci-0000:13:00.1	up	up	extended
PV	/dev/sdd	pci-0000:13:00.1	up	up	extended
VG	VolGroup02	—	up		
PV	/dev/sdk	pci-0000:13:00.0	up	up	extended
PV	/dev/sdi	pci-0000:13:00.0	up	up	extended
PV	/dev/sgd	pci-0000:13:00.1	up	up	extended
PV	/dev/sde	pci-0000:13:00.1	up	up	extended

再開する場合は start を指定してください。

```
# /opt/HA/SrG/bin/srgadmin -c start
Change TESTIO.
STOP -> START
```

(6) 一定間隔でリソースの状態を定期表示します。

<表示間隔として3秒を指定した場合>

```
# /opt/HA/SrG/bin/srgadmin - c status - t 3
(monitor status = TRUE)
```

type	device	HostBusAdapter	L status	P status	Online status
VG	VolGroup01	—	up		
PV	/dev/sdj	pci-0000:13:00.0	up	up	extended
PV	/dev/sdh	pci-0000:13:00.0	up	up	extended
PV	/dev/sdf	pci-0000:13:00.1	up	up	extended
PV	/dev/sdd	pci-0000:13:00.1	up	up	extended
VG	VolGroup02	—	up		
PV	/dev/sdk	pci-0000:13:00.0	up	up	extended
PV	/dev/sdi	pci-0000:13:00.0	up	up	extended
PV	/dev/sdg	pci-0000:13:00.1	up	up	extended
PV	/dev/sde	pci-0000:13:00.1	up	up	extended

3 秒経過

```
(monitor status = TRUE)
```

type	device	HostBusAdapter	L status	P status	Online status
VG	VolGroup01	—	up		
PV	/dev/sdj	pci-0000:13:00.0	up	up	extended
PV	/dev/sdh	pci-0000:13:00.0	up	up	extended
PV	/dev/sdf	pci-0000:13:00.1	up	up	extended
PV	/dev/sdd	pci-0000:13:00.1	up	up	extended

注意: コマンドを終了させたい場合、ctrl+c で終了できます。

(7) I/O パスの組み込み状態(Online status)と同期を取ります。

現在の状態と I/O パスのステータスの同期を取ります。

コマンド投入等、手動で I/O パスの切り離しを行った場合には、実際の I/O パスの状態と Online status が異なっている可能性がありますので、Online status を更新してください。

```
# /opt/HA/SrG/bin/srgadmin -c sync
```

(8) デーモンプロセスの起動、終了について

本製品は、インストール時に rc ファイルがインストールされますので、OS 起動時に自動的に監視を開始します。

- rc ファイルからの起動、終了
OS 起動(boot)を契機に自動起動、OS 終了を契機に自動終了されます。
- 手動起動、終了
以下のコマンドを root 権限で投入することで起動できます。

```
# cd /  
#/etc/init.d/srgctl start
```

以下のコマンドを root 権限で投入することで終了できます。

```
# cd /  
#/etc/init.d/srgctl stop
```

上記コマンドで終了しない場合は、ps -ef | grep srg で srg から始まるプロセスの pid を検索して、kill -9 <pid> で終了させてください。

また、RENS と連携している場合、RENS 連携用モニタプロセスは以下のコマンドで起動、停止可能です。

起動コマンド

```
# cd /  
#/etc/init.d/ssdiagctl start
```

停止コマンド

```
# cd /  
#/etc/init.d/ssdiagctl stop
```

(9) デーモンプロセスの自動起動を一時停止したい場合

メンテナンス等で OS 起動時にディスク監視を行いたくない場合には、以下の手順で監視の一時停止が可能です。

- OS の再起動前に、一時ファイルを作成します。

```
# touch /var/opt/HA/SrG/conf/srg.ignore
```

- OS を再起動すると、syslog に監視停止中のメッセージが表示されます。

```
srgignore[<pid>]: Start waiting for /var/opt/HA/SrG/conf/srg.ignore
```

- 一時ファイルを削除すると、監視を開始します。

```
# rm /var/opt/HA/SrG/conf/srg.ignore
```

- StorageSaver が監視を開始したメッセージが syslog に表示されます。

```
srgignore[<pid>]: /var/opt/HA/SrG/conf/srg.ignore is deleted. Start srgd.
```

4.2.オンライン保守コマンドの操作手順

(1) srgreduce コマンドは、指定された I/O パスの閉塞を実行します。

注意: I/O パスが SPS または PowerPath で冗長化されている場合のみ有効です。

- ・ -H を付与すると、HBA カード単位で I/O パスを切り離します。
 - 状態確認

```
# /opt/HA/SrG/bin/srgadmin - c status
(monитор status = TRUE)
```

type	device	HBA	L status	P status	Online status
VG	VolGroup01	—	up		
PV	/dev/sdj	pci-0000:13:00.0	up	up	extended
PV	/dev/sdh	pci-0000:13:00.0	up	up	extended
PV	/dev/sdf	pci-0000:13:00.1	up	up	extended
PV	/dev/sdd	pci-0000:13:00.1	up	up	extended
VG	VolGroup02	—	up		
PV	/dev/sdk	pci-0000:13:00.0	up	up	extended
PV	/dev/sdi	pci-0000:13:00.0	up	up	extended
PV	/dev/sdg	pci-0000:13:00.1	up	up	extended
PV	/dev/sde	pci-0000:13:00.1	up	up	extended

- ① 2 つ以上の FC(上記の場合は 0000:13:00.0, 0000:13:00.1)配下の I/O パスが共に組み込み済(extended)であることが前提です。
- ②、③ 両 FC 配下の I/O パスが up であることを確認してください。

- 閉塞実行

```
# /opt/HA/SrG/bin/srgreduce - H 0000:13:00.1
```

- ① 0000:13:00.1 配下の I/O パスを一括で閉塞します。

- 状態確認

```
# /opt/HA/SrG/bin/srgadmin - c status
(monitor status = TRUE)
```

① I/O パスの組み込み状態				
type	device	HostBusAdapter	L status	P status : Online status
② VG のステータス				
VG	VolGroup01	—	<i>suspend</i>	
PV	/dev/sdj	pci-0000:13:00.0	up	up : extended
PV	/dev/sdh	pci-0000:13:00.0	up	up : extended
PV	/dev/sdf	pci-0000:13:00.1	up	up : <i>reduced</i>
PV	/dev/sdd	pci-0000:13:00.1	up	up : <i>reduced</i>
VG	VolGroup02	—	<i>suspend</i>	
PV	/dev/sdk	pci-0000:13:00.0	up	up : extended
PV	/dev/sdi	pci-0000:13:00.0	up	up : extended
PV	/dev/sdg	pci-0000:13:00.1	up	up : <i>reduced</i>
PV	/dev/sde	pci-0000:13:00.1	up	up : <i>reduced</i>

- ① 0000:13:00.1 配下の I/O パスを閉塞すると Online status が reduced になります。
- ② VG status が suspend になります。

注意: I/O パスの一括閉塞は FC 単位で実行しますが、両系の FC カード配下の I/O パス (冗長化されたすべての I/O パス) を同時に閉塞することは出来ません。
事前に、他系の FC カード配下の I/O パスが組み込まれていることを確認してください。

- ・ -F を付与すると、PV の udev デバイスファイル単位で I/O パスを切り離します。
 - 状態確認

```
# /opt/HA/SrG/bin/srgadmin -i
(monitor status = TRUE)

=====
type : device status : P status : Online status
=====
VG : vordrupper : up : —
PV : /dev/sdi : up : up : extended
PV : /dev/sdh : up : up : extended
    : pci-0000:13:00.0-fc-0x2900001697120ca7:0x0000000000000000
PV : /dev/sdf : up : up : extended
    : pci-0000:13:00.1-fc-0x2100001697120ca7:0x0000000000000000
PV : /dev/sdd : up : up : extended
    : pci-0000:13:00.1-fc-0x2900001697120ca7:0x0000000000000000
VG : VolGroup02 : up : —
PV : /dev/sdk : up : up : extended
    : pci-0000:13:00.0-fc-0x2100001697120ca7:0x0001000000000000
PV : /dev/sdi : up : up : extended
    : pci-0000:13:00.0-fc-0x2900001697120ca7:0x0001000000000000
PV : /dev/sdg : up : up : extended
    : pci-0000:13:00.1-fc-0x2100001697120ca7:0x0001000000000000
PV : /dev/sde : up : up : extended
    : pci-0000:13:00.1-fc-0x2900001697120ca7:0x0001000000000000
```

- ① pci-0000:13:00.1-fc-0x2900001697120ca7:0x0000000000000000 の示す I/O パスが組み込み済(extended)であることが前提です。
- ②、③ I/O パスが up であることを確認してください。

- 閉塞実行

①閉塞対象 udev デバイスファイル

```
# /opt/HA/SrG/bin/srgreduce - F pci-0000:13:00.1-fc-0x2900001697120ca7:0x0000000000000000
```

- ① I/O パスを閉塞します。

- 状態確認

```
# /opt/HA/SrG/bin/srgadmin -i
(monitor status = TRUE)
```

① I/O パスの組み込み状態				
type	device	L status	P status	Online status
② VG のステータス				
VG	VolGroup01	<i>suspend</i>	—	
PV	/dev/sdj	up	up	extended
	pci-0000:13:00.0-fc-0x2100001697120ca7:0x0000000000000000			
PV	/dev/sdh	up	up	extended
	pci-0000:13:00.0-fc-0x2900001697120ca7:0x0000000000000000			
PV	/dev/sdf	up	up	extended
	pci-0000:13:00.1-fc-0x2100001697120ca7:0x0000000000000000			
PV	/dev/sdd	up	up	<i>reduced</i>
	pci-0000:13:00.1-fc-0x2900001697120ca7:0x0000000000000000			
VG	VolGroup02	up	—	
PV	/dev/sdk	up	up	extended
	pci-0000:13:00.0-fc-0x2100001697120ca7:0x0001000000000000			
PV	/dev/sdi	up	up	extended
	pci-0000:13:00.0-fc-0x2900001697120ca7:0x0001000000000000			
PV	/dev/sdg	up	up	extended
	pci-0000:13:00.1-fc-0x2100001697120ca7:0x0001000000000000			
PV	/dev/sde	up	up	extended
	pci-0000:13:00.1-fc-0x2900001697120ca7:0x0001000000000000			

- ① Online status が reduced になります。
- ② VG status が suspend になります。

- (2) srgextend コマンドは、指定された I/O パスの復旧を実行します。
- ・ -H を付与すると、HBA カード単位で I/O パスを組み込みます。
 - 状態確認

```
# /opt/HA/SrG/bin/srgadmin - c status
(monitor status = TRUE)
```

① I/O パスの組み込み状態					
type	device	HostBusAdapter	L status	P status	Online status
② VG のステータス					
VG	VolGroup01	—	<i>suspend</i>		
PV	/dev/sdj	pci-0000:13:00.0	up	up	extended
PV	/dev/sdh	pci-0000:13:00.0	up	up	extended
PV	/dev/sdf	pci-0000:13:00.1	up	up	<i>reduced</i>
PV	/dev/sdd	pci-0000:13:00.1	up	up	<i>reduced</i>
VG	VolGroup02	—	<i>suspend</i>		
PV	/dev/sdk	pci-0000:13:00.0	up	up	extended
PV	/dev/sdi	pci-0000:13:00.0	up	up	extended
PV	/dev/sgd	pci-0000:13:00.1	up	up	<i>reduced</i>
PV	/dev/sde	pci-0000:13:00.1	up	up	<i>reduced</i>

- ① I/O パスが閉塞(reduced)されていることが前提です。
- ② 0000:13:00.1 配下の I/O パスが閉塞されているため、VG status が *suspend* になっています。

- 復旧実行

① 復旧対象 FC の HW パス

```
# /opt/HA/SrG/bin/srgextend - H 0000:13:00.1
```

- ① 0000:13:00.1 配下の I/O パスを一括で復旧します。

- 状態確認

```
# /opt/HA/SrG/bin/srgadmin - c status
(monитор status = TRUE)
```

① I/O パスの組み込み状態					
type	device	HostBusAdapter	L status	P status	Online status
② VG のステータス					
VG	VolGroup01	—	up		
PV	/dev/sdj	pci-0000:13:00.0	up	up	extended
PV	/dev/sdh	pci-0000:13:00.0	up	up	extended
PV	/dev/sdf	pci-0000:13:00.1	up	up	<i>extended</i>
PV	/dev/sdd	pci-0000:13:00.1	up	up	<i>extended</i>
VG	VolGroup02	—	up		
PV	/dev/sdk	pci-0000:13:00.0	up	up	extended
PV	/dev/sdi	pci-0000:13:00.0	up	up	extended
PV	/dev/sgd	pci-0000:13:00.1	up	up	<i>extended</i>
PV	/dev/sde	pci-0000:13:00.1	up	up	<i>extended</i>

- ① 0000:13:00.1 配下の I/O パスを復旧すると Online status が extended になります。
- ② VG status が up になります。

- ・ -F を付与すると、PV の udev デバイスファイル単位で I/O パスを組み込みます。
 - 状態確認

```
# /opt/HA/SrG/bin/srgadmin -i
(monitor status = TRUE)
```

① I/O パスの組み込み状態			
type	device	L status	P status
VG	VolGroup01	<i>suspend</i>	—
PV	/dev/sdj	up	extended
	pci-0000:13:00.0-fc-0x2100001697120ca7:0x0000000000000000		
PV	/dev/sdh	up	extended
	pci-0000:13:00.0-fc-0x2900001697120ca7:0x0000000000000000		
PV	/dev/sdf	up	extended
	pci-0000:13:00.1-fc-0x2100001697120ca7:0x0000000000000000		
PV	/dev/sdd	up	<i>reduced</i>
	pci-0000:13:00.1-fc-0x2900001697120ca7:0x0000000000000000		
VG	VolGroup02	up	—
PV	/dev/sdk	up	extended
	pci-0000:13:00.0-fc-0x2100001697120ca7:0x0001000000000000		
PV	/dev/sdi	up	extended
	pci-0000:13:00.0-fc-0x2900001697120ca7:0x0001000000000000		
PV	/dev/sdg	up	extended
	pci-0000:13:00.1-fc-0x2100001697120ca7:0x0001000000000000		
PV	/dev/sde	up	extended
	pci-0000:13:00.1-fc-0x2900001697120ca7:0x0001000000000000		

- ① pci-0000:13:00.1-fc-0x2900001697120ca7:0x0000000000000000 の示す I/O パスが閉塞(reduced)されていることが前提です。

- 復旧実行

① 復旧対象 udev デバイスファイル

```
# /opt/HA/SrG/bin/srgextend - F pci-0000:13:00.1-fc-0x2900001697120ca7:0x0000000000000000
```

- ① I/O パスを復旧します。

- 状態確認

```
# /opt/HA/SrG/bin/srgadmin -i
(monitor status = TRUE)
```

① I/O パスの組み込み状態				
type	device	L status	P status	Online status
② VG のステータス				
VG	VolGroup01	up	—	
PV	/dev/sdj	up	up	extended
	pci-0000:13:00.0-fc-0x2100001697120ca7:0x0000000000000000			
PV	/dev/sdh	up	up	extended
	pci-0000:13:00.0-fc-0x2900001697120ca7:0x0000000000000000			
PV	/dev/sdf	up	up	extended
	pci-0000:13:00.1-fc-0x2100001697120ca7:0x0000000000000000			
PV	/dev/sdd	up	up	extended
	pci-0000:13:00.1-fc-0x2900001697120ca7:0x0000000000000000			
VG	VolGroup02	up	—	
PV	/dev/sdk	up	up	extended
	pci-0000:13:00.0-fc-0x2100001697120ca7:0x0001000000000000			
PV	/dev/sdi	up	up	extended
	pci-0000:13:00.0-fc-0x2900001697120ca7:0x0001000000000000			
PV	/dev/sdg	up	up	extended
	pci-0000:13:00.1-fc-0x2100001697120ca7:0x0001000000000000			
PV	/dev/sde	up	up	extended
	pci-0000:13:00.1-fc-0x2900001697120ca7:0x0001000000000000			

- ① Online status が extended になります。
- ② VG status が up になります。

(3) srgrecover コマンドは、すべての I/O パスの復旧を実行します。

- 状態確認

```
# /opt/HA/SrG/bin/srgadmin - c status
(monitor status = TRUE)
```

① I/O パスの組み込み状態					
type	device	HostBusAdapter	L status	P status	Online status
② VG のステータス					
VG	VolGroup01	—	<i>suspend</i>		
PV	/dev/sdj	pci-0000:13:00.0	up	up	extended
PV	/dev/sdh	pci-0000:13:00.0	up	up	extended
PV	/dev/sdf	pci-0000:13:00.1	up	up	<i>reduced</i>
PV	/dev/sdd	pci-0000:13:00.1	up	up	<i>reduced</i>
VG	VolGroup02	—	<i>suspend</i>		
PV	/dev/sdk	pci-0000:13:00.0	up	up	extended
PV	/dev/sdi	pci-0000:13:00.0	up	up	extended
PV	/dev/sgd	pci-0000:13:00.1	up	up	<i>reduced</i>
PV	/dev/sde	pci-0000:13:00.1	up	up	<i>reduced</i>

- ① I/O パスが閉塞されています。
- ② 0000:13:00.1 配下の I/O パスが閉塞されているため、VG status が *suspend* になっています。

- 復旧実行

```
# /opt/HA/SrG/bin/srgrecover - v
HW-path: pci-0000:13:00.0-fc-0x2100001697120ca7:0x0000000000000000 result: OK
HW-path: pci-0000:13:00.0-fc-0x2900001697120ca7:0x0000000000000000 result: OK
HW-path: pci-0000:13:00.1-fc-0x2100001697120ca7:0x0000000000000000 result: OK
HW-path: pci-0000:13:00.1-fc-0x2900001697120ca7:0x0000000000000000 result: OK
HW-path: pci-0000:13:00.0-fc-0x2100001697120ca7:0x0001000000000000 result: OK
HW-path: pci-0000:13:00.0-fc-0x2900001697120ca7:0x0001000000000000 result: OK
HW-path: pci-0000:13:00.1-fc-0x2100001697120ca7:0x0001000000000000 result: OK
HW-path: pci-0000:13:00.1-fc-0x2900001697120ca7:0x0001000000000000 result: OK
srgrecover complete.
```

- ① すべての I/O パスを一括で復旧します。

- 状態確認

```
# /opt/HA/SrG/bin/srgadmin - c status
(monитор status = TRUE)
```

① I/O パスの組み込み状態				
type	device	hostbusadapter	L status	P status : Online status
② VG のステータス				
VG	VolGroup01	—	up	
PV	/dev/sdj	pci-0000:13:00.0	up	extended
PV	/dev/sdh	pci-0000:13:00.0	up	extended
PV	/dev/sdf	pci-0000:13:00.1	up	<i>extended</i>
PV	/dev/sdd	pci-0000:13:00.1	up	<i>extended</i>
VG	VolGroup02	—	up	
PV	/dev/sdk	pci-0000:13:00.0	up	extended
PV	/dev/sdi	pci-0000:13:00.0	up	extended
PV	/dev/sgd	pci-0000:13:00.1	up	<i>extended</i>
PV	/dev/sde	pci-0000:13:00.1	up	<i>extended</i>

- ① すべての I/O パスの Online status が extended になります。
- ② VG status が up になります。

4.3.障害復旧時の操作

障害を検出し、該当障害箇所を交換等して復旧が完了すると、StorageSaver の監視ステータスを復旧する必要があります。

本復旧操作を実施しない状態で運用を継続すると、別の装置故障を契機に両系障害を誤検出する可能性がありますので必ず実施してください。

監視ステータスの復旧は srgrecover コマンドを実行します。

※本手順の詳細については別冊の『CLUSTERPRO MC StorageSaver ハードウェア障害復旧後の運用手順』に詳述しておりますのでそちらもご参照ください。

(1) srgrecover コマンドで、すべての I/O パスの復旧を実行します。

① 状態確認

```
# /opt/HA/SrG/bin/srgadmin -c status
(monitor status = TRUE)
```

type	device	HostBusAdapter	L status	P status	
VG	VolGroup01	—	down	down	extended
PV	/dev/sdj	pci-0000:13:00.0	down	down	extended
PV	/dev/sdh	pci-0000:13:00.0	down	down	extended
PV	/dev/sdf	pci-0000:13:00.1	down	down	extended
PV	/dev/sdd	pci-0000:13:00.1	down	down	extended
VG	VolGroup02	—	up	up	extended
PV	/dev/sdk	pci-0000:13:00.0	up	up	extended
PV	/dev/sdi	pci-0000:13:00.0	up	up	extended
PV	/dev/sgd	pci-0000:13:00.1	up	up	extended
PV	/dev/sde	pci-0000:13:00.1	up	up	extended

② 復旧実行

```
# /opt/HA/SrG/bin/srgrecover -v
HW-path: pci-0000:13:00.0-fc-0x2100001697120ca7:0x0000000000000000 result: OK
HW-path: pci-0000:13:00.0-fc-0x2100001697120ca7:0x0001000000000000 result: OK
HW-path: pci-0000:13:00.0-fc-0x2900001697120ca7:0x0000000000000000 result: OK
HW-path: pci-0000:13:00.0-fc-0x2900001697120ca7:0x0001000000000000 result: OK
HW-path: pci-0000:13:00.1-fc-0x2100001697120ca7:0x0000000000000000 result: OK
HW-path: pci-0000:13:00.1-fc-0x2100001697120ca7:0x0001000000000000 result: OK
HW-path: pci-0000:13:00.1-fc-0x2900001697120ca7:0x0000000000000000 result: OK
HW-path: pci-0000:13:00.1-fc-0x2900001697120ca7:0x0001000000000000 result: OK
srgrecover complete.
```

③ 状態確認

```
# /opt/HA/SrG/bin/srgadmin - c status
(monitor status = TRUE)
```

type	device	HostBusAdapter	L status	P status	Online status
VG	VolGroup01	—	up		
PV	/dev/sdj	pci-0000:13:00.0	up	up	extended
PV	/dev/sdh	pci-0000:13:00.0	up	up	extended
PV	/dev/sdf	pci-0000:13:00.1	up	up	extended
PV	/dev/sdd	pci-0000:13:00.1	up	up	extended
VG	VolGroup02	—	up		
PV	/dev/sdk	pci-0000:13:00.0	up	up	extended
PV	/dev/sdi	pci-0000:13:00.0	up	up	extended
PV	/dev/sdg	pci-0000:13:00.1	up	up	extended
PV	/dev/sde	pci-0000:13:00.1	up	up	extended

(2) 自動復旧機能について

ディスクアレイ装置の故障等により障害が発生し、装置交換を行った場合、ハードウェアとして I/O パスを復旧させた後に、手動で `srgrecover` コマンドを実行して StorageSaver の管理ステータスを復旧する必要があります。

自動復旧機能を使用すると、I/O パスの状態を定期的にチェックし、障害状態から復旧した I/O パスを検出すると、`srgrecover` コマンドを実行して復旧した I/O パスを自動的に組み込みます。

ただし、一時的に I/O が通るような間欠故障が発生した場合に、自動復旧機能を使用して I/O パスの自動組み込みを行うと、I/O の切り替えが頻発することによる I/O の遅延が発生する可能性があります。

そのため、障害が発生した場合には障害箇所の点検をおこない、確実に I/O パスの状態が復旧したことを確認した後に、手動で `srgrecover` コマンドを実行する運用を推奨します。

運用上オペレータの介入が困難であるなど、復旧作業を自動化せざるをえない場合はシステム定義ファイルの以下のパラメータを `ENABLE` に変更し、デーモンプロセスを再起動することで、自動復旧機能を利用することができます。

- システム定義ファイル (`/var/opt/HA/SrG/conf/srg.config`) の変更

```
# Auto recovery flag
#      used = ENABLE : unused = DISABLE(default)
AUTO_RECOVERY      ENABLE      ← ENABLE に変更します。
```

- デーモンプロセスの再起動
システム定義ファイルの変更後、デーモンプロセスを再起動します。

```
# /opt/HA/SrG/bin/srgconfig -r
```

(3) オンライン保守中の自動復旧機能について

オンライン保守中は自動復旧機能を一時停止し、メンテナンス中の機器が自動で組み込まれることを防止します。

自動復旧機能を使用されている環境で、srgreduce コマンド実行時に以下のメッセージが出力されると、自動復旧機能により、メンテナンス中の機器を組み込んでしまいます。以下のメッセージが出力された場合は、再度 srgreduce コマンドを実行するか、メッセージで指定されているコマンドを実行してください。

Can't create target file(/var/opt/HA/SrG/conf/.online_maintaining)(xx).
Please execute the command "srgreduce" again.
Or please execute the following command "touch /var/opt/HA/SrG/conf/.online_maintaining".

自動復旧機能を使用されている環境で、srgrecover, srgextend コマンド実行時に以下のメッセージが出力されると、自動復旧機能が動作しません。以下のメッセージが出力された場合は、再度 srgrecover, srgextend コマンドを実行するかメッセージで指定されているコマンドを実行してください。

Can't delete target file(/var/opt/HA/SrG/conf/.online_maintaining)(xx).
Please execute the command "srgextend/srgrecover" again.
Or please execute the following command "rm /var/opt/HA/SrG/conf/.online_maintaining".

5. CLUSTERPRO との連携

5.1. CLUSTERPRO との連携の概要

CLUSTERPRO と連携して、I/O パスの両系障害発生時に待機ノードへフェイルオーバーグループを切り替えて業務を継続することができます。

StorageSaver が CLUSTERPRO と連携する場合、次の4通りの方式があります。

- ① CLUSTERPRO のカスタムモニタリソースにクラスタウェア連携デーモン (srgstat) を登録する方式
 - ② CLUSTERPRO のサーバ管理プロセス (clpnm) を強制終了し、ノードを切り替える方式
 - ③ システムメモリダンプを採取し、強制的に OS を停止し、ノードを切り替える方式
 - ④ RENS 経由で CLUSTERPRO へリソース状態を通知し、ノードを切り替える方式
- RENS を利用した CLUSTERPRO 連携の詳細については後述の「RENS との連携」の章をご覧ください。

StorageSaver としては①の方式を推奨します。

また、①の方式で連携する場合、障害時に確実にフェイルオーバーできるよう、フェイルオーバー時の CLUSTERPRO の動作設定は「クラスタサービス停止と OS シャットダウン」を選択してください。

「クラスタサービス停止と OS シャットダウン」を選択していない場合、I/O パス障害の影響でフェイルオーバー処理が正しく完了せず、フェイルオーバーに失敗したり、フェイルオーバー完了が遅延したりする場合があります。

各連携方式の説明は後述の章を参照してください。

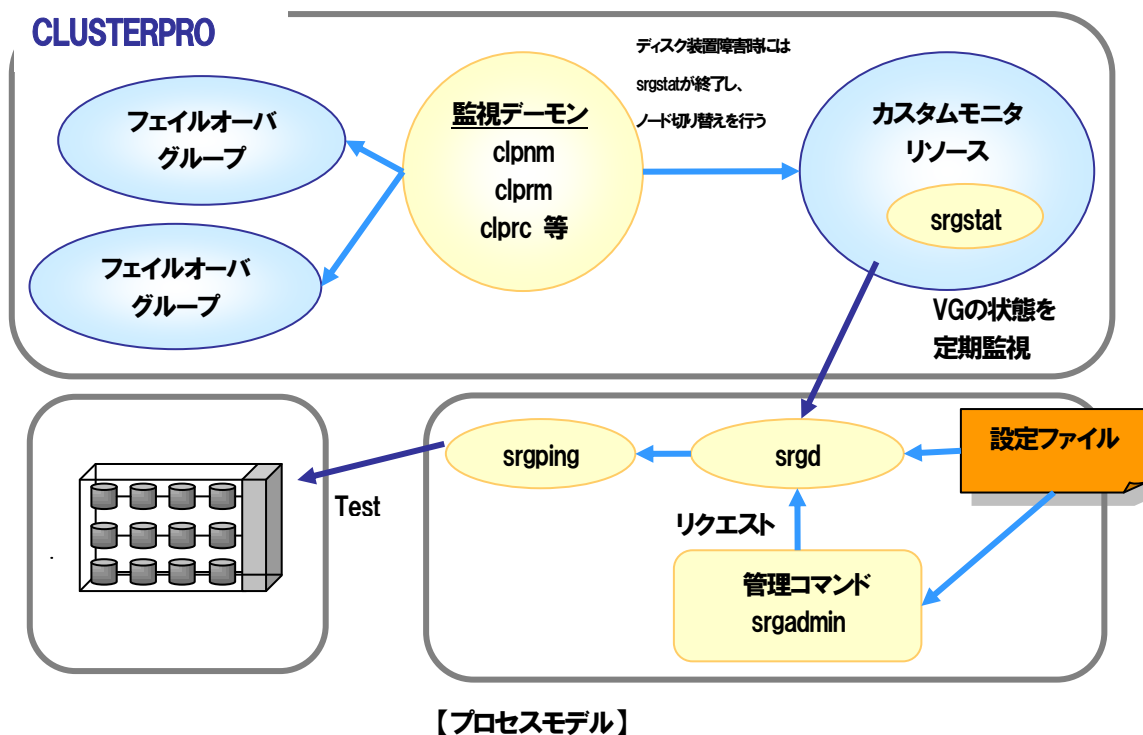
5.2.カスタムモニタリソースによる CLUSTERPRO との連携

(1) CLUSTERPRO との連携について

ディスク装置の動作状態をモニタするコマンド `srgstat(1M)` を CLUSTERPRO のカスタムモニタリソースとして登録することで、ディスク装置の障害時のノードダウン、ノード切り替えを実現します。

本機能を利用する場合は、不必要に CLUSTERPRO のサーバ管理プロセス (`clpnm`) を kill しないために StorageSaver のコンフィグレーション(`srg.config`)の **`VG_FAULT_ACTION`**、**`VG_STALL_ACTION`**には **`SERVICE_CMD_ENABLE`**を指定してください。

この方式であれば、複数ノードクラスタシステムでのノード切り替えだけでなく縮退した状態でのノードダウンや1ノードのクラスタシステムでのノードダウンを実現できますので、非常に有用な手法です。



(2) srgstat の運用について

ディスク装置に障害が発生すると、srgd(1M)が I/O パスおよび VG レベルの管理ステータスを down 状態に変更し、syslog、コンソールにエラーメッセージを出力します。

srgstat(1M)は共有メモリを経由して VG レベルの管理ステータスをモニタします。
VG が down 状態に遷移した時点で、srgstat(1M)は異常終了し、CLUSTERPRO がカスタムモニタリソースのダウンを検出しノード切り替え、ノードダウンが発生します。

srgstat(1M)は、srgd(1M)および srgping(1M)のプロセスが起動され、ディスク装置の監視を行っている場合に有効に機能します。

以下のようなリソース監視を停止している場合は、VG 障害を検出できません。

- srgd(1M)および srgping(1M)のプロセスが起動されていない。
- srgadmin(1M)のオペレーション操作でリソース監視停止を指示されている。

<syslog メッセージの出力例>

LVM かつシングルディスク構成での事例です。

下記の順序で syslog にメッセージが出力されます。

ディスク装置へのファイル I/O が停止すると、syslog に記録されない場合もあります。

● 最初に、I/O パスの片系障害(正系 LUN 障害)を検出

Jan 10 18:03:28 node1 srgd[xxxxx]: PV status change fail .

[hwpath = pci-0000:1a:00.1-fc-0x2001000013840322:0x0000000000000000: s.f = /dev/sdb]

● 次に、I/O パスの両系障害(副系 LUN 障害)を検出

Jan 10 18:05:29 node1 srgd[xxxxx]: PV status change fail .

[hwpath = pci-0000:1a:00.0-fc-0x2002000013840322:0x0000000000000000: s.f = /dev/sdc]

Jan 10 18:05:29 node1 srgd[xxxxx]: VG status change down .(vg=VolGroup01)

● VG へのアクセス不可を検出し、srgstat が異常終了しノードダウン、ノード切り替えを実行

Jan 10 18:05:29 node1 srgstat[xxxxx]: found VG status is down

(3) カスタムモニタリソースの設定手順について

srgstat を CLUSTERPRO のカスタムモニタリソースに登録する手順については、後述の

“10.4 カスタムモニタリソースによる CLUSTERPRO との連携手順”

を参照してください。

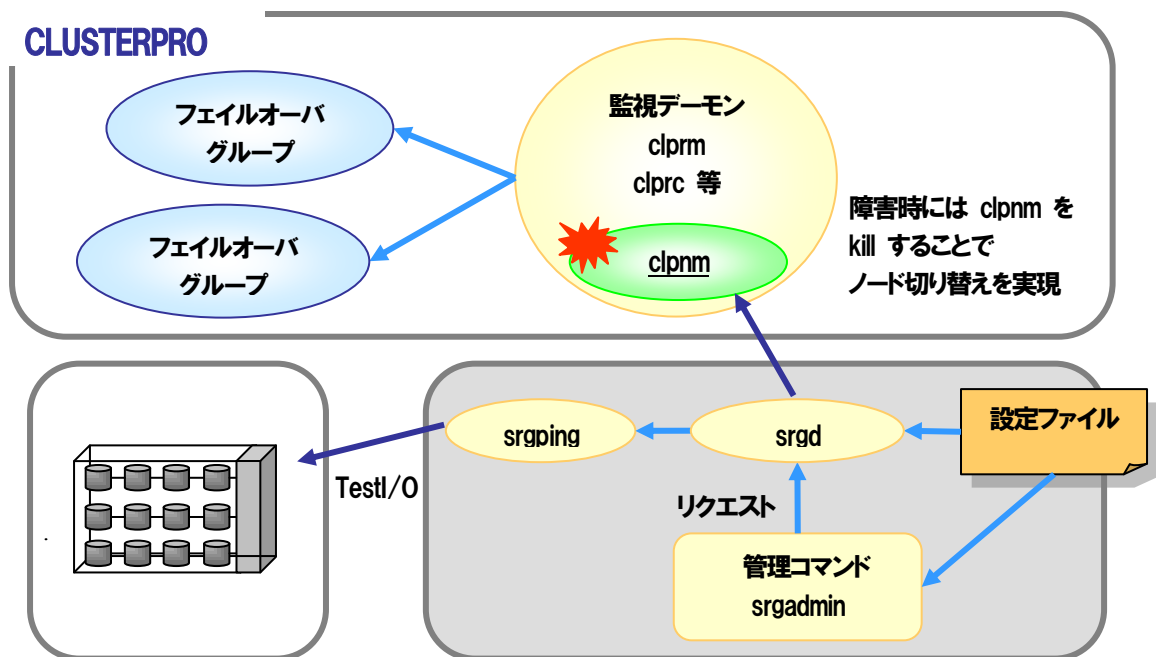
5.3.CLUSTERPRO のサーバ管理プロセス (clpnm) の強制終了による CLUSTERPRO との 連携

(1) CLUSTERPRO との連携について

ディスク装置の障害時に、CLUSTERPRO のサーバ管理プロセス (clpnm) を強制終了させることで、ノード切り替えを実現する方式です。

StorageSaver のコンフィグレーションである **VG_FAULT_ACTION**、**VG_STALL_ACTION** に **CLPNM_KILL_ENABLE** を指定すると、ディスク装置の障害時に clpnm を kill することができます。

このノード切り替え機能は、2ノード以上のクラスタシステムで有効です。



【 プロセスモデル 】

(2) I/O パスの障害を検出すると

ディスク装置に障害が発生すると、srgd(1M)が I/O パスおよび VG レベルの管理ステータスを down 状態に変更し、syslog、コンソールにエラーメッセージを出力します。

srgd(1M)はコンフィグレーションに CLPNM_KILL_ENABLE を指定していると CLUSTERPRO のサーバ管理プロセス (clpnm) を強制終了させ、CLUSTERPRO によるノード切り替えを行います。
また、srgd(1M)自身も abort します。

/(root) 配下に core ファイルを出力しますので、ディスク装置故障時にこれらのファイルを確認してください。
なお、故障パターンによっては core ファイル等が残っていないケースもあります。

<syslog メッセージの出力例>

- 最初に、I/O パスの片系障害(正系 LUN 障害)を検出
Jan 10 18:03:28 node1 srgd[xxxxx]: PV status change fail .
[hwpath = pci-0000:1a:00.1-fc-0x2001000013840322:0x0000000000000000: s.f = /dev/sdb]
- 次に、I/O パスの両系障害(副系 LUN 障害)を検出
Jan 10 18:05:29 node1 srgd[xxxxx]: PV status change fail .
[hwpath = pci-0000:1a:00.0-fc-0x2002000013840322:0x0000000000000000: s.f = /dev/sdc]
Jan 10 18:05:29 node1 srgd[xxxxx]: VG status change down .(vg=VolGroup01)
- VG へのアクセス不可を検出し、予備ノードへ切り替え
Jan 10 18:05:29 node1 srgd[xxxxx]: start KILL clpnm.
Jan 10 18:05:29 node1 srgd[xxxxx]: send signal clpnm.
Jan 10 18:05:29 node1 srgd[xxxxx]: abort srgd.

5.4. システムメモリダンプ採取と OS 強制停止による CLUSTERPRO との連携

(1) CLUSTERPRO との連携について

ディスク装置故障時にシステムメモリダンプの採取と OS 強制停止 (panic) により CLUSTERPRO と連携してノード切り替えを実現します。

ディスク装置故障時には OS やその他監視製品なども正常に動作できない場合がありますので、この方式による OS 強制停止でノード切り替えを行うことは有効です。

また、システムメモリダンプが採取されますので、障害状態の解析なども可能です。

注意: ディスク装置の故障パターンによっては、正しくシステムメモリダンプが採取できない場合があります。
システムメモリダンプが採取できない場合も、CLUSTERPRO が待機ノードから現用ノードの異常を検出しますので系切り替えは可能です。

本機能を使ってシステムメモリダンプを採取する場合には、あらかじめ kdump の設定が完了している必要があります。

また、システムメモリダンプは /var/crash 配下に作成されます。

万が一システムメモリダンプ採取に失敗した場合、続いて CLUSTERPRO のサーバ管理プロセス (clpnm) を強制停止することで待機ノードへの切り替えを試みます。

6. RENS との連携

6.1. RENS 連携の概要

RENS が導入されたシステムの場合、RENS と連携して StorageSaver の監視リソースの状態を他の製品へ通知することが可能です。

例えば、RENS のリソース通知をサポートしているクラスタウェアなどはこの RENS 連携機能を利用して、StorageSaver の監視しているリソースの状態に応じてフェイルオーバーの実行などが可能となります。

StorageSaver と RENS を連携するためには RENS 連携用の rpm パッケージを別途インストールする必要があります。

RENS 連携には StorageSaver 本体のリソース監視デーモンプロセスとは別に、RENS 連携用モニタプロセス (ssdiagd) を利用します。

6.2.RENS 連携の設定

(1) RENS 連携 rpm パッケージのインストール

- RPM パッケージ名
clusterpro-mc-ss-renc-w.x.y-z.x86_64.rpm

- インストール

```
# rpm -ivh /mnt/cdrom/Linux/rpm/clusterpro-mc-ss-renc-w.x.y-z.x86_64.rpm
```

注意:RENS がインストールされていない環境に RENS 連携 rpm パッケージをインストールすることは出来ません。
RENS 連携を行う場合は、事前に RENS をインストールしてください。

インストールが完了した場合以下のコマンドでインストールの確認を行ってください。

```
# rpm -qa | grep clusterpro-mc-ss-renc  
clusterpro-mc-ss-renc-w.x.y-z
```

- アンインストール

```
# rpm -e clusterpro-mc-ss-renc-w.x.y-z
```

注意:w, x, y, z にはバージョン番号が入ります。
機能強化があるとバージョン番号が更新されます。

※インストール手順についての詳細は、『CLUSTERPRO MC StorageSaver 1.1 for Linux リリースメモ』をご覧ください。

(2) RENS 連携設定

RENS 連携を行う場合、システム定義ファイル (*srg.config*)中の以下のパラメータを RENS 連携の指定に変更します。

VG_FAULT_ACTION → RENS_REPORT_ENABLE

これにより、RENS 連携用モニタプロセス(ssdiagd)の起動と、監視対象リソース障害時の RENS へのメッセージ通知が行われるようになります。

また、srgwatchの設定ファイル(/var/opt/HA/SrG/local/conf/srgwatch.config)を修正し、srgwatchの監視対象に RENS 連携用モニタプロセス(ssdiagd)を追加します。以下の記述を追加してください。

```
#
# ssdiagd
#
ssdiagd {
    PROCNAME = /opt/HA/SrG/bin/ssdiagd -c 60
    EXEC_CMD = /etc/init.d/ssdiagctl start
    EXECLOCK = /var/opt/HA/SrG/conf/.ssdiagd_lock
    INTERVAL = 1
    WATCH    = WATCHON
    HUP       = ACTION_NONE
    RETRY     = 10
}
```

設定変更後は必ず以下の手順で StorageSaver デーモンプロセスの再起動および RENS 連携用モニタプロセス(ssdiagd)の起動を実行してください。

```
# /etc/init.d/srgctl restart
# /etc/init.d/ssdiagctl start
```

6.3.RENS 連携用モニタプロセスの運用管理

RENS 連携用モニタプロセスは、通常 OS 起動時に rc スクリプト経由で起動されます。

メンテナンス等で手動による起動停止をする場合、rc スクリプトをコマンドラインから実行することで実現できます。

- (1) RENS 連携用モニタプロセスの起動

```
# cd /  
# /etc/init.d/ssdiagctl start
```

- (2) RENS 連携用モニタプロセスの停止

```
# cd /  
# /etc/init.d/ssdiagctl stop
```

注意:RENS 連携中に RENS 連携用モニタプロセス(ssdiagd)を手動停止した場合、syslog に以下のメッセージが出力されます。動作に問題はありませんので、無視してください。

```
XXX XX XX:XX:XX XX hamon(ssdiagd)[xxxxx]: --RENS-- WARNING: RENS monitor is  
already stopped: ssdiagd  
XXX XX XX:XX:XX XX hamon(ssdiagd)[xxxxx]: --RENS-- FATAL: Failed to  
rens_get_next_lowuser_HAevent. errcode = 8, Monitor = ssdiagd  
XXX XX XX:XX:XX XX hamon(ssdiagd)[xxxxx]: --RENS-- INFO: Monitor process finished.  
Monitor = ssdiagd  
XXX XX XX:XX:XX XX hamon(ssdiagd)[xxxxx]: --RENS-- ERROR: RENS monitor is not  
found: ssdiagd  
XXX XX XX:XX:XX XX hamon(ssdiagd)[xxxxx]: --RENS-- INFO: Monitor unregistered.  
(monitor-id=<モニタ ID>, monitor-name=ssdiagd)
```

6.4.RENS を利用した CLUSTERPRO 連携手順

(1) StorageSaver の RENS 連携設定

StorageSaver の設定手順については前述の 6.2 章を参照してください。

(2) RENS の設定

詳細手順については、RENS ユーザーズガイド『サーバ管理基盤 利用の手引き(リリース 1.1)』の“3.3.2 SW イベント監視コンポーネントとの連携手順”を参照してください。

- ssdiagd の辞書ファイルを登録します。

```
# cp /var/opt/HA/SrG/conf/rens/* /opt/mcl/rens/dict/
```

- RENS SW イベント監視コンポーネント設定ファイルを作成します。

```
# cp /opt/mcl/rens/conf/lower/buffer/monitor_buf.conf.template  
/opt/mcl/rens/conf/lower/buffer/ssdiagd_buf.conf
```

- モニタプロセス設定ファイルを作成します。

```
# cp /opt/mcl/rens/conf/monitor.conf.template  
/opt/mcl/rens/conf/ssdiagd.conf
```

上記コマンド実行後、/opt/mcl/rens/conf/ssdiagd.conf の shm_key 値を他のモニタプロセス設定ファイルと競合しない値に変更してください。

注意: 他のモニタプロセス設定ファイルの shm_key 値は以下のコマンドで取得できます。
grep "shm_key" /opt/mcl/rens/conf/*.conf

- モニタ起動スクリプトを編集します。

/opt/mcl/rens/script/monitor_run.sh に以下の 1 行を追加します。

```
${RENS_BIN}/hamon -n ssdiagd &
```

- (3) 監視対象 VG が大量に存在する際の手順
- RENS 連携機能を使用する際、監視対象 VG が大量に(101 以上)存在する場合には、RENS のモニタプロセス(hamon)に-t オプションを付加し、登録リソース数の上限を拡張する必要があります。
- 以下の 2 ファイルを編集してください。

注意:<登録リソース数上限> には VG 数より大きい値(非負の整数値)を指定して下さい。

・/opt/mcl/rens/script/monitor_run.sh

【 変更前 】

`${RENS_BIN}/hamon -n ssdiagd &`

【 変更後 】

`${RENS_BIN}/hamon -t <登録リソース数上限> -n ssdiagd &`

・/etc/init.d/ssdiagctl

【 変更前 】

`/opt/mcl/rens/bin/hamon -n ssdiagd &`

【 変更後 】

`/opt/mcl/rens/bin/hamon -t <登録リソース数上限> -n ssdiagd &`

- (4) CLUSTERPRO の設定
- 詳細手順については、RENS ユーザーズガイド『サーバ管理基盤 利用の手引き(リリース 1.1)』の“3.3.3 CLUSTERPRO X for MC Linux との連携手順”を参照してください。

- メッセージ受信モニタリソースを登録します。

【設定パラメータ(デフォルト値から変更する必要があるパラメータのみ記載)】

タイプ	“message receive monitor”
監視タイプ	“HA/SS”
監視対象	監視するデバイス名 (下記【監視対象リソース名の取得方法】を参照)
回復対象	クラスタ全体
最終動作	“クラスタデーモン停止と OS シャットダウン”

【監視対象リソース名の取得方法】

RENS を起動します。

```
# /opt/mcl/rens/script/rens_start.sh
```

RENS が起動していることを確認します。

```
# ps -ef | grep rensd
```

StorageSaver を起動します。

```
# /etc/init.d/srgctl start
# /etc/init.d/ssdiagctl start
```

RENS 管理コマンドを実行し、監視対象リソースの情報を確認します。

Monitor Name が “ssdiagd” である行の Alias 部が監視対象リソース名となります。

```
# /opt/mcl/rens/bin/rensadmin show -r
```

ID	ResourceName	Alias	Status	LastUpdateTime	MonitorName
0	0000:07:00.0	eth0	up	2009/08/27 16:22:31	e1000
1	0000:13:00.0	host7	up	2009/08/27 16:22:31	ssdiagd
2	0000:07:00.1	eth1	up	2009/08/27 16:22:31	e1000
3	0000:13:00.1	host8	up	2009/08/27 16:22:31	lpfc
4	VolGroup02_status	VolGroup02_status	up	2009/08/27 16:22:13	ssdiagd
5	VolGroup01_status	VolGroup01_status	up	2009/08/27 16:22:13	ssdiagd

- (5) 動作確認手順
設定内容が正しく反映されていることを確認します。

- 片系障害時の動作確認
障害発生前のリソース状態を確認します。

```
# /opt/HA/SrG/bin/srgadmin - c status
(monitor status = TRUE)
```

type	device	HostBusAdapter	L status	P status	Online status
VG	VolGroup01	—	up		
PV	/dev/sdj	pci-0000:13:00.0	up	up	extended
PV	/dev/sdh	pci-0000:13:00.0	up	up	extended
PV	/dev/sdf	pci-0000:13:00.1	up	up	extended
PV	/dev/sdd	pci-0000:13:00.1	up	up	extended
VG	VolGroup02	—	up		
PV	/dev/sdk	pci-0000:13:00.0	up	up	extended
PV	/dev/sdi	pci-0000:13:00.0	up	up	extended
PV	/dev/sdg	pci-0000:13:00.1	up	up	extended
PV	/dev/sde	pci-0000:13:00.1	up	up	extended

```
# /opt/mcl/rens/bin/rensadmin show
```

ID	ResourceName	Alias	Status	LastUpdateTime	MonitorName
0	0000:07:00.0	eth0	up	2009/08/27 16:22:31	e1000
1	0000:13:00.0	host7	up	2009/08/27 16:22:31	lpfc
2	0000:07:00.1	eth1	up	2009/08/27 16:22:31	e1000
3	0000:13:00.1	host8	up	2009/08/27 16:22:31	lpfc
4	VolGroup02_status	VolGroup02_status	up	2009/08/27 17:16:14	ssdiagd
5	VolGroup01_status	VolGroup01_status	up	2009/08/27 17:16:14	ssdiagd

ID	TargetName	Type	Priority
0	syslog	syslog	middle
1	textlog	textlog	middle
2	clpx	clusterpro	middle low

ID	MonitorName	Pid	Status	Commandline
0	e1000	8519	run	/opt/mcl/rens/bin/nicmon -n e1000
1	ssdiagd	8522	run	/opt/mcl/rens/bin/hamon -n ssdiagd
2	lpfc	8529	run	/opt/mcl/rens/bin/fcomon -n lpfc
3	targetregclpd	8526	run	/opt/mcl/rens/bin/targetregclpd -f

/opt/mcl/rens/conf/targetclp.conf

片系障害を擬似的に発生させます。

```
# /opt/HA/SrG/bin/srgadmin - c debug - v on - F 0000:13:00.1
Change debug value.
FC devfile = 0000:13:00.1
0 -> 1
FC devfile = 0000:13:00.1
0 -> 1
FC devfile = 0000:13:00.1
0 -> 1
FC devfile = 0000:13:00.1
0 -> 1
```

約 180 秒後

```
# /opt/HA/SrG/bin/srgadmin - c status
(monitor status = TRUE)
```

type	device	HostBusAdapter	L status	P status	Online status
VG	VolGroup01	—	<i>suspend</i>		
PV	/dev/sdj	pci-0000:13:00.0	up	up	extended
PV	/dev/sdh	pci-0000:13:00.0	up	up	extended
PV	/dev/sdf	pci-0000:13:00.1	<i>down</i>	<i>down</i>	extended
PV	/dev/sdd	pci-0000:13:00.1	<i>down</i>	<i>down</i>	extended
VG	VolGroup02	—	<i>suspend</i>		
PV	/dev/sdk	pci-0000:13:00.0	up	up	extended
PV	/dev/sdi	pci-0000:13:00.0	up	up	extended
PV	/dev/sdg	pci-0000:13:00.1	<i>down</i>	<i>down</i>	extended
PV	/dev/sde	pci-0000:13:00.1	<i>down</i>	<i>down</i>	extended

```
# /opt/mcl/rens/bin/rensadmin show
```

ID	ResourceName	Alias	Status	LastUpdateTime	MonitorName
0	0000:07:00.0	eth0	up	2009/08/27 16:22:31	e1000
1	0000:13:00.0	host7	up	2009/08/27 16:22:31	lpfc
2	0000:07:00.1	eth1	up	2009/08/27 16:22:31	e1000
3	0000:13:00.1	host8	up	2009/08/27 16:22:31	lpfc
4	VolGroup02_status	VolGroup02_status	<i>suspend</i>	2009/08/27 17:24:21	ssdiagd
5	VolGroup01_status	VolGroup01_status	<i>suspend</i>	2009/08/27 17:24:21	ssdiagd

ID	TargetName	Type	Priority
0	syslog	syslog	middle
1	textlog	textlog	middle
2	clpx	clusterpro	middle low

ID	MonitorName	Pid	Status	Commandline
0	e1000	8519	run	/opt/mcl/rens/bin/nicmon -n e1000
1	ssdiagd	8522	run	/opt/mcl/rens/bin/hamon -n ssdiagd
2	lpfc	8529	run	/opt/mcl/rens/bin/fcomon -n lpfc
3	targetregclpd	8526	run	/opt/mcl/rens/bin/targetregclpd -f

```
/opt/mcl/rens/conf/targetclp.conf
```

syslog に以下のメッセージが出力されます。

```
srgd[xxxxx]: PV status change fail . [hwpath =
pci-0000:13:00.1-fc-0x2100001697120ca7:0x0001000000000000: s.f = /dev/sdg].
srgd[xxxxx]: PV status change fail . [hwpath =
pci-0000:13:00.1-fc-0x2900001697120ca7:0x0001000000000000: s.f = /dev/sde].
srgd[xxxxx]: PV status change fail . [hwpath =
pci-0000:13:00.1-fc-0x2100001697120ca7:0x0000000000000000: s.f = /dev/sdf].
srgd[xxxxx]: PV status change fail . [hwpath =
pci-0000:13:00.1-fc-0x2900001697120ca7:0x0000000000000000: s.f = /dev/sdd].
hamon(ssdiagd) [xxxxx]: RENS detected the CRITICAL event. <resource=VolGroup02_status
(id=4)><event#=3><severity=CRITICAL><summary= VolGroup02 : SUSPEND><event seq#=41836>
hamon(ssdiagd) [xxxxx]: RENS detected the CRITICAL event. <resource=VolGroup01_status
(id=5)><event#=3><severity=CRITICAL><summary= VolGroup01 : SUSPEND><event seq#=41837>
```

- 両系障害時の動作確認
障害発生前のリソース状態を確認します。
既に片系障害が発生していることを確認します。

```
# /opt/HA/SrG/bin/srgadmin - c status
(monitor status = TRUE)
```

type	device	HostBusAdapter	L status	P status	Online status
VG	VolGroup01	—	<i>suspend</i>		
PV	/dev/sdj	pci-0000:13:00.0	up	up	extended
PV	/dev/sdh	pci-0000:13:00.0	up	up	extended
PV	/dev/sdf	pci-0000:13:00.1	<i>down</i>	<i>down</i>	extended
PV	/dev/sdd	pci-0000:13:00.1	<i>down</i>	<i>down</i>	extended
VG	VolGroup02	—	<i>suspend</i>		
PV	/dev/sdk	pci-0000:13:00.0	up	up	extended
PV	/dev/sdi	pci-0000:13:00.0	up	up	extended
PV	/dev/sdg	pci-0000:13:00.1	<i>down</i>	<i>down</i>	extended
PV	/dev/sde	pci-0000:13:00.1	<i>down</i>	<i>down</i>	extended

```
# /opt/mcl/rens/bin/rensadmin show
```

ID	ResourceName	Alias	Status	LastUpdateTime	MonitorName
0	0000:07:00.0	eth0	up	2009/08/27 16:22:31	e1000
1	0000:13:00.0	host7	up	2009/08/27 16:22:31	lpfc
2	0000:07:00.1	eth1	up	2009/08/27 16:22:31	e1000
3	0000:13:00.1	host8	up	2009/08/27 16:22:31	lpfc
4	VolGroup02_status	VolGroup02_status	<i>suspend</i>	2009/08/27 17:43:20	ssdiagd
5	VolGroup01_status	VolGroup01_status	<i>suspend</i>	2009/08/27 17:43:20	ssdiagd

ID	TargetName	Type	Priority
0	syslog	syslog	middle
1	textlog	textlog	middle
2	clpx	clusterpro	middle low

ID	MonitorName	Pid	Status	Commandline
0	e1000	8519	run	/opt/mcl/rens/bin/nicmon -n e1000
1	ssdiagd	8522	run	/opt/mcl/rens/bin/hamon -n ssdiagd
2	lpfc	8529	run	/opt/mcl/rens/bin/fcomon -n lpfc
3	targetregclpd	8526	run	/opt/mcl/rens/bin/targetregclpd -f /opt/mcl/rens/conf/targetclp.conf

両系障害を擬似的に発生させます。

```
# /opt/HA/SrG/bin/srgadmin -c debug -v on -F 0000:13:00.0
Change debug value.
FC devfile = 0000:13:00.0
0 -> 1
FC devfile = 0000:13:00.0
0 -> 1
FC devfile = 0000:13:00.0
0 -> 1
FC devfile = 0000:13:00.0
0 -> 1
```

約 60 秒後

```
# /opt/HA/SrG/bin/srgadmin -c status
(monitor status = TRUE)
```

type	device	HostBusAdapter	L status	P status	Online status
VG	VolGroup01	—	<i>down</i>		
PV	/dev/sdj	pci-0000:13:00.0	<i>down</i>	<i>down</i>	extended
PV	/dev/sdh	pci-0000:13:00.0	<i>down</i>	<i>down</i>	extended
PV	/dev/sdf	pci-0000:13:00.1	down	down	extended
PV	/dev/sdd	pci-0000:13:00.1	down	down	extended
VG	VolGroup02	—	<i>down</i>		
PV	/dev/sdk	pci-0000:13:00.0	<i>down</i>	<i>down</i>	extended
PV	/dev/sdi	pci-0000:13:00.0	<i>down</i>	<i>down</i>	extended
PV	/dev/sdg	pci-0000:13:00.1	down	down	extended
PV	/dev/sde	pci-0000:13:00.1	down	down	extended

```
# /opt/mcl/rens/bin/rensadmin show
```

ID	ResourceName	Alias	Status	LastUpdateTime	MonitorName
0	0000:07:00.0	eth0	up	2009/08/27 16:22:31	e1000
1	0000:13:00.0	host7	up	2009/08/27 16:22:31	lpfc
2	0000:07:00.1	eth1	up	2009/08/27 16:22:31	e1000
3	0000:13:00.1	host8	up	2009/08/27 16:22:31	lpfc
4	VolGroup02_status	VolGroup02_status	down	2009/08/27 17:51:37	ssdiagd
5	VolGroup01_status	VolGroup01_status	down	2009/08/27 17:51:37	ssdiagd

ID	TargetName	Type	Priority
0	syslog	syslog	middle
1	textlog	textlog	middle
2	clpx	clusterpro	middle low

ID	MonitorName	Pid	Status	Commandline
0	e1000	8519	run	/opt/mcl/rens/bin/nicmon -n e1000
1	ssdiagd	8522	run	/opt/mcl/rens/bin/hamon -n ssdiagd
2	lpfc	8529	run	/opt/mcl/rens/bin/fcmon -n lpfc
3	targetregclpd	8526	run	/opt/mcl/rens/bin/targetregclpd -f /opt/mcl/rens/conf/targetclp.conf

syslog に以下のメッセージが出力されます。

```
srgd[xxxx]: VG status change down . (vg=VolGroup01)
srgd[xxxx]: VG status change down . (vg=VolGroup02)
hamon(ssdiagd)[xxxx]: RENS detected the CRITICAL event. <resource=VolGroup02_status
(id=4)><event#=2><severity=CRITICAL><summary= VolGroup02 : DOWN><event seq#=43182>
hamon(ssdiagd)[xxxx]: RENS detected the CRITICAL event. <resource=VolGroup01_status
(id=5)><event#=2><severity=CRITICAL><summary= VolGroup01 : DOWN><event seq#=43183>
hamon(ssdiagd)[xxxx]: —RENS— INFO: Succeeded to send notification. Target = clpx.
clusterpro: <type: mm><event: 903> An error of HA/SS type and VolGroup01_status device has been
detected
. (VolGroup01 : DOWN)
clusterpro: <type: mm><event: 905> An error has been detected in monitoring VolGroup01_status. (-1)
clusterpro: <type: apisv><event: 12> There was a request to shutdown server from the
mm(IP=XXX.XXX.XXX.XXX).
clusterpro: <type: rc><event: 91> The server was shut down.
```

注意: サーバが shutdown するため、ステータスは確認することができない場合があります。

7. syslog メッセージ

7.1. syslog に出力するメッセージについて

本製品では、リソース監視で致命的な異常を検出すると syslog にメッセージを出力します。
syslog ファイルおよび syslog の facility と level は以下のとおりです。

syslog ファイル名	<i>/var/log/messages</i>
facility	LOG_DAEMON
level	LOG_ALERT、LOG_ERR、LOG_WARNING

7.2. 警報対象として登録することを推奨するメッセージ一覧

特に重要度の高い syslog メッセージを記述します。

これらのメッセージが出力された場合は、HW 保守担当者に HW 検査を依頼してください。

(1) TestI/O のリソース監視で異常を検出した場合

LOG_ERROR の出力契機は以下のとおりです。

PV status change fail .[hwpath = xxx: s.f = /dev/xxx].

説明 : TestI/O で I/O パスの異常を検出
処置 : I/O パス異常を検出したので、早急に該当ディスクの点検を行ってください。

I/O request uncomplete in time .[hwpath = xxx: s.f = /dev/xxx].

説明 : TestI/O で I/O ストールを検出
処置 : I/O パス異常を検出したので、早急に該当ディスクの点検を行ってください。

VG status change down .(vg=xx)

説明 : TestI/O で VG レベルの異常(down)を検出
処置 : すべての I/O パスが障害となっています。
早急に該当ディスクの点検を行ってください。

(2) リソース状態の定期通知で異常を検出した場合

LOG_ERROR の出力契機は以下のとおりです。

Monitor Status is reported, P-stat is down.

説明 : I/O パスの異常を検出
処置 : TestI/O で異常を検出後、復旧していない可能性があります。
早急に該当ディスクの点検を行ってください。
障害状態からの復旧後は、srgrecover を実行して I/O パスの管理状態を復旧させてください。

Monitor Status is reported, L-stat is reduced.

説明 : I/O パスの閉塞状態を検出
処置 : TestI/O で異常を検出後、復旧していない可能性があります。
早急に該当ディスクの点検を行ってください。
さらに、I/O パスの管理状態を復旧させてください。
障害状態からの復旧後は、srgrecover を実行して I/O パスの管理状態を復旧させてください。

7.3. その他のメッセージ

(1) ライセンス不正を検出した場合

LOG_ERROR の出力契機は以下のとおりです。

Activation failed. Product key name is not been entry.

説明: ライセンス認証に失敗しました。有償ロックキーが登録されていません。

処置: ライセンスファイルに有償ロックキーを登録してください。

Activation failed. Cord word is generated by different product key name.

説明: ライセンス認証に失敗しました。有償ロックキーが一致していません。

処置: 発行されたコードワードが正しく登録できていることを確認してください。

Activation failed. Cord word is generated by different host ID.

説明: ライセンス認証に失敗しました。ホスト情報が一致していません。

処置: 発行されたコードワードが正しく登録できていることを確認してください。

Activation failed. Trial term is expired.

説明: ライセンス認証に失敗しました。試用期限を過ぎています。

処置: 正式版のライセンスを登録してください。

7.4. 運用管理製品との連携

本製品は、運用管理製品と連携し syslog メッセージを監視することができます。

これにより、syslog に出力される重要なログをアラートとしてリアルタイムで通知でき、障害発生時も早急な発見、迅速な対応が可能になります。

本製品で連携可能な運用管理製品は、以下となります。

◆ WebSAM SystemManager

StorageSaver が異常を検知し syslog にその内容が出力されると、

WebSAM SystemManager のログ監視機能にて通知が行われます。

※ 連携手順については、「CLUSTERPRO MC (HA シリーズ) WebSAM SystemManager メッセージ監視連携手順書」を参照してください。

◆ ESMPRO/ServerAgent

StorageSaver が異常を検知し syslog にその内容が出力されると、

ESMPRO/ServerAgent のアラート通報機能にて通知が行われます。

※ 連携手順については、「ESMPRO/ServerAgent アラート通報機能との連携手順」を参照してください。

8. 注意・制限事項について

8.1. 注意・制限事項

(1) 下記の注意事項があります。

- 本製品を運用中には共有メモリをデフォルトで約 2MB 程度使用します。
- ログ用のディレクトリ(`/var/opt/HA/SrG/log`)配下に、ログファイルを保存するために、約 60MB 程度使用します。
トレースファイルは、サイクリックとなっていますので、60MB を超えることはありません
ただし、オンライン保守コマンドである `srgreduce`, `srgextend`, `srgrecover` のログファイルは
ログ用のディレクトリ(`/var/opt/HA/SrG/log/util`)に日単位に追記型で保存されるため、
これらのコマンドを数分単位で連続実行し続けると、約100パスで1日100MB 程度のディスクを
消費することがあります。
オンライン保守コマンドを定期的に行う場合は、ログファイルを定期的に削除するような
運用をお願いします。
- ディスクの間欠障害、部分的なメディアエラー等で異常を検出できない場合があります。
- 本製品における管理リソースの上限値は以下のとおりです。
HBA カード ノードあたり最大 32 枚まで
I/O パス HBA カードあたり最大 512 パスまで
- StorageSaver で障害を検出すると、障害を検出した I/O パスの監視を停止します。
自動閉塞機能を有効にしている場合は、障害を検出した I/O パスの閉塞、監視の停止を
行います。
障害復旧後は必ず `srgrecover` コマンドで閉塞した I/O パスの復旧と監視の再開を
行う必要があります。
本復旧操作を実施しない状態で運用を継続すると、別の装置故障を契機に両系障害を誤検出する
可能性がありますので必ず実施してください。
詳細については『CLUSTERPRO MC StorageSaver ハードウェア障害復旧後の運用手順』をご覧ください。
- 一時的に I/O が通るような間欠故障が発生した場合に、自動復旧機能を使用して I/O パスの
自動組み込みを行うと、I/O の切り替えが頻発することによる I/O の遅延が発生する可能性が
あります。
障害が発生した場合には障害箇所の点検をおこない、確実に I/O パスの状態が復旧したことを
確認した後に、手動で `srgrecover` コマンドを実行する運用を推奨します。
- 本製品を利用する場合、Linux SCSI パススルードライバ(`sg`ドライバ)がインストールされ、
事前にカーネルモジュールとしてロードされている必要があります。

モジュールがロードされているかどうかは `lsmod(8)`コマンドの結果から確認できます。
以下の `sg`ドライバのエントリ行が出力されることを確認してください。

```
# lsmod | grep sg  
sg
```

```
38369 2
```

```
sgドライバのエントリ行
```

- 内部で以下のパッケージを利用します。

sg3_utilsUtils for Linux's SCSI generic driver devices + raw devices

本パッケージがインストールされていない場合、事前にインストールしてください。
以下のコマンドでインストールの有無を確認できます。

```
# rpm -qa sg3_utils  
sg3_utils-x.y-z.el5
```

※インストールされていない場合、何も出力されません

注意: x, y, z には sg3_utils パッケージのバージョン番号が入ります。

本パッケージは標準で OS インストール媒体中に含まれます。

- Red Hat Enterprise Linux 6.x (64bit) または Oracle Linux 6.x (64bit) のシステムに本製品をインストールする場合、事前に互換ライブラリ(glibc-x.y.el6.i686.rpm)がインストールされている必要があります。

互換性ライブラリがインストールされていない場合、事前にインストールしてください。
以下のコマンドでインストールの有無を確認できます。

```
# rpm -qa glibc  
:  
glibc-x.y.el6.i686
```

※インストールされていない場合、" glibc-x.y.el6.i686" の行が出力されません

注意: x, y には互換性ライブラリのバージョン番号が入ります。

本パッケージは標準で OS インストール媒体中に含まれます。

- 障害等で監視対象のディスクアレイ装置等が OS 起動時に認識されていない状態で StorageSaver が起動した場合、監視対象に組み込むには故障したディスクアレイ装置等の障害を復旧し、OS が正常に認識出来ていることを確認した後、デーモンプロセスの再起動が必要です。

- FC 接続構成の変更等、H/W 構成を変更する場合は、設定ファイルの再作成および適用操作を行う必要があります。以下の手順を実行してください。
 1. デーモンプロセスの自動起動を抑制する


```
# touch /var/opt/HA/SrG/conf/srg.ignore
```
 2. OS 停止
 3. H/W 構成変更
 4. OS 起動
 5. 設定ファイルを再作成する


```
# /opt/HA/SrG/bin/srgquery -s /tmp
```
 6. 設定ファイルの整合性をチェックする


```
# /opt/HA/SrG/bin/srgconfig -c -s /tmp
```
 7. 設定ファイルを実行環境に適用する


```
# /opt/HA/SrG/bin/srgconfig -a -s /tmp
```
 8. デーモンプロセスの自動起動抑制を解除する


```
# rm /var/opt/HA/SrG/conf/srg.ignore
```
- StorageSaver はマルチパス管理製品(SPS, PowerPath)のステータス・LVM 構成等を取得し、取得した情報にしたがって設定ファイルを生成します。
設定ファイル自動生成コマンド(srgquery)を実行する前に
サーバに接続されているディスクアレイ装置の設定(I/O パス冗長構成・LVM 構成等)が正しいことを確認してください。
- EMC 社製 symmetrix DMX シリーズ、Symmetrix VMAX シリーズのディスク装置を接続した環境において、設定ファイル自動生成コマンド (srgquery) にて設定ファイルを生成した場合、ディスク容量が一定値 (128MB) 以下のものは管理用ディスクと判断し、監視対象に含めません。
ディスク容量が 128MB 以下であるデータディスク装置を接続している場合は、設定ファイルを手動で編集して監視定義を追加してください。
- NEC 社製ストレージ iStorage M シリーズを接続した環境の場合、設定ファイルの自動生成後に必要な手順があります。
詳細については『CLUSTERPRO MC StorageSaver iStorage M シリーズ使用時の設定手順』をご覧ください。
- SANboot 環境において設定ファイル自動生成コマンド (srgquery) にて設定ファイルを生成した場合、OS 情報が格納されたディスク装置も FC 接続構成であるため他のデータディスク装置同様、監視対象として定義します。
設定ファイルから OS 情報が格納されたディスク装置の記述を手動で削除してください。

- IDE 接続のデバイスが接続されている環境において設定ファイル自動生成コマンド (srgquery) にて設定ファイルを生成した際、IDE 接続のデバイスを監視対象として定義する場合があります。
IDE 接続のデバイスが設定ファイルに含まれている場合は手動で削除してください。
- 仮想環境 (ゲスト OS) にて本製品を使用する場合は、設定ファイル自動生成コマンド (srgquery) 実行時に -a オプションを指定する必要があります。
また、自動生成されたシステム定義ファイル (srg.config) の TESTIO_MODE に READ を設定する必要があります。
- 設定ファイル自動生成コマンド (srgquery) にて設定ファイルの自動生成を行った際、マルチパス管理製品 (SPS, PowerPath) による I/O パス冗長設定が完了していない可能性がある場合、以下の警告メッセージを出力します。
srgquery: <デバイスファイル> is not found in StoragePathSavior.
srgquery: <デバイスファイル> is not found in PowerPath.
マルチパス管理製品 (SPS, PowerPath) の I/O パス冗長設定が完了していない場合は設定完了後、再度 srgquery コマンドにて設定ファイルの自動生成を行ってください。
- CLUSTERPRO を導入している場合、CLUSTERPRO による softdog を利用した OS 停止機能を優先するため、StorageSaver の POWER_OFF 機能は利用することができません。
同等の機能が CLUSTERPRO 側に実装されているので、そちらの機能を利用してください。
- 特定の旧式の SCSI 接続のディスク装置を監視した場合、syslog に以下のメッセージが定期的に出力されます。
これらのディスク装置はサポート対象外のため、監視することができません。
設定ファイルにこれらのディスク装置の記述を含めないようにしてください。

XXX XX XX:XX:XX XX kernel: targetX:X:X: FAST-XX SCSI XX MB/s ST (XXX ns, offset XX)

- 監視対象ディスク装置へのアクセスができなくなり、クラスタウェアと連携しノードを切り替える際、障害発生時に情報を残すため、ルートディレクトリ("/")配下に StorageSaver デーモンプロセスの core ファイルが作成されることがあります。
- 対象 I/O パスが LVM の VG に含まれていない場合、管理上"VG_NONE"という仮想的な VG 名で登録しますので、LVM の VG 名に"VG_NONE"を使用しないようにしてください。
- 閉塞機能・復旧機能は SPS または PowerPath が導入された環境でのみ有効です。
それ以外の構成の場合はパスの閉塞・復旧を行いません。
- SIGMABLADE 等の一般 Linux サーバを用いる場合は、RENS 連携を行うことが出来ません。
その場合は、クラスタウェア連携デーモン(srgstat)を用いることでクラスタウェア連携を行います。
詳細は "5.2 カスタムモニタリソースによる CLUSTERPRO との連携" を参照してください。

- RENS 連携機能を使用する際、監視対象 VG が大量に(101 以上)存在する場合には、RENS のモニタプロセス(hamon)に-t オプションを付加し、登録リソース数の上限を拡張する必要があります。以下の 2 ファイルを編集してください。

注意:<登録リソース数上限> には VG 数より大きい値(非負の整数値)を指定して下さい。

•/opt/mcl/rens/script/monitor_run.sh

【 変更前 】

`${RENS_BIN}/hamon -n ssdiagd &`

【 変更後 】

`${RENS_BIN}/hamon -t <登録リソース数上限> -n ssdiagd &`

•/etc/init.d/ssdiagctl

【 変更前 】

`/opt/mcl/rens/bin/hamon -n ssdiagd &`

【 変更後 】

`/opt/mcl/rens/bin/hamon -t <登録リソース数上限> -n ssdiagd &`

- RENS がインストールされていない環境に RENS 連携 rpm パッケージ (clusterpro-mc-ss-rens-w.x.y-z.x86_64.rpm)をインストールすることは出来ません。RENS 連携を行う場合は、事前に RENS をインストールしてください。
- RENS 連携中に RENS 連携用モニタプロセス(ssdiagd)を手動停止した場合、syslog に以下のメッセージが出力されます。動作に問題はありませんので、無視してください。

```
XXX XX XX:XX:XX XX hamon(ssdiagd)[xxxxx]: --RENS-- WARNING: RENS monitor is
already stopped: ssdiagd
XXX XX XX:XX:XX XX hamon(ssdiagd)[xxxxx]: --RENS-- FATAL: Failed to
rens_get_next_lowuser_HAevent. errcode = 8, Monitor = ssdiagd
XXX XX XX:XX:XX XX hamon(ssdiagd)[xxxxx]: --RENS-- INFO: Monitor process finished.
Monitor = ssdiagd
XXX XX XX:XX:XX XX hamon(ssdiagd)[xxxxx]: --RENS-- ERROR: RENS monitor is not
found: ssdiagd
XXX XX XX:XX:XX XX hamon(ssdiagd)[xxxxx]: --RENS-- INFO: Monitor unregistered.
(monitor-id=<モニタ ID>, monitor-name=ssdiagd)
```

(2) 下記の制限事項があります。

- iStorage RV のデバイスの監視はできません。
- Linux md デバイス管理ツールをつかったソフトミラー構成では設定ファイル自動生成できません。
- 設定ファイル自動生成機能はすべてのデバイス構成をサポートするものではありません。
構成によっては自動生成できない場合がありますので、その場合はエディタ等で直接ファイルを編集してください。
- srgconfig -c コマンド(設定ファイル整合性チェック機能)による srg.config のチェック機能は本バージョンでは非サポートです。

(3) 障害発生時の対応について

本製品では、リソース監視で異常を検出すると、syslog に警告メッセージを出力します。
このときの動作履歴やオンライン保守による閉塞／復旧の運用履歴をトレースファイルに取得
していますので、障害解析資料として、以下の情報を採取してください。
なお、トレースファイルは、サイクリックとなっているためディスク容量を圧迫することはありません。

監視構成ファイル	<i>/var/opt/HA/SrG/conf</i> 配下の全ファイル <i>/var/opt/HA/SrG/local/conf</i> 配下の全ファイル
トレースファイル	<i>/var/opt/HA/SrG/log</i> 配下の全ファイル
syslog ファイル	<i>/var/log/messages</i> ファイル
コマンド実行一時ファイル	<i>/var/opt/HA/SrG/local</i> 配下の全ファイル
コマンド出力結果	<i>rpm -qa</i> <i>sg_scan -i /dev/sd*</i> <i>ls -l /dev/disk/by-path/*</i> <i>cat /proc/scsi/sps/dd* (※1)</i> <i>powermt display dev=all (※2)</i> <i>vgdisplay -v (※3)</i> <i>lvdisplay -v (※3)</i> <i>pvdisplay -v (※3)</i> (※1) SPS を利用したマルチパス構成の場合 (※2) PowerPath を利用したマルチパス構成の場合 (※3) LVM 構成の場合
クラスタ関連ファイル	(※)クラスタ関連ファイルについては各クラスタウェア製品により異なりますので、製品ごとにマニュアルを参照してください。

8.2.オンライン保守における注意事項

- (1) 本製品のオンライン保守機能で I/O パスを閉塞しても、一時的に I/O が発行されることがありますのでご注意ください。
- オンライン保守中にフェイルオーバーグループが起動されると一時的に I/O が発生することがあります。
 - I/O パスを閉塞しても FC アダプタを制御する FC ドライバから TestI/O のフレームが送出される場合があります。
 - FC 上に DLT 装置が接続されている場合は、DLT 装置に対する I/O の停止、再開はできません。
- (2) オンライン保守開始前に
- 特定の FC 配下 I/O パスへの TestI/O を停止する場合は、クラスタ配下の全ノードに対して FC 単位で閉塞コマンドを実行してください。
- (3) オンライン保守終了後の取り扱いについて
- オンライン保守終了後は、srgrecover(1M) で構成復旧を実施してください。
 - FC ケーブルを抜いたままでノードを立ち上げた場合では、構成復旧コマンドを操作しても正常に動作しません。
OS が監視対象のディスク装置を正常に認識出来ていることを確認した上で、srgrecover(1M) で構成復旧を実施してください。
- (4) 自動復旧機能を ON(ENABLE)にした状態でオンライン保守を行った場合、以下の制限事項があります。
- オンライン保守中は自動復旧機能は動作しません。オンライン保守終了後、srgrecover(1M) もしくは srgextend(1M) コマンドの実行によりオンライン保守終了と判断し、自動復旧機能が動作するようになります。
 - 複数の PV が閉塞された状態で 1PV のみ指定して srgextend を実行した場合、オンライン保守中を示すファイルが削除されるため自動復旧機能が再開されます。その場合、保守作業が完了していない状態で PV が復旧される可能性があります。PV 単位で閉塞、復旧を行う場合には PV 毎に閉塞、復旧を行うか、すべての保守作業が完了した時点で srgrecover コマンドにより全 PV を一括で復旧させてください。
 - srgreduce(1M) コマンドの実行を中断した場合、オンライン保守中を示すファイルが作成されたまま残ってしまい、自動復旧機能が動作できなくなる可能性があります。srgreduce(1M) コマンドを中断した場合は、コマンドを再度実行してパスの閉塞を行うか、srgrecover(1M) もしくは srgextend(1M) コマンドを実行してパスを復旧させてください。

9. リファレンス

srgd(1M)

名称

srgd - StorageSaver リソース監視デーモン

構文

srgd

機能説明

srgd は、StorageSaver のリソース監視デーモンです。

srgd は、クラスタウェア で構築されたシステムで使用される FC 接続および SCSI 接続のディスク装置の監視を行い、シングルノードまたはクラスタシステムの保全性を最大限に確保する機能を提供します。

srgd は、rc(/etc/init.d/srgctl)から起動されます。

srgd を停止、再開させる場合、以下のコマンドを実行してください。

停止時

```
# /etc/init.d/srgctl stop
```

再開時

```
# /etc/init.d/srgctl start
```

ファイル

/opt/HA/SrG/bin/*

/var/opt/HA/SrG/conf/*

/var/opt/HA/SrG/log/*

関連項目

srgadmin(1M), srgconfig(1M), srgquery(1M),
srgextend(1M), srgreduce(1M), srgrecover(1M)

srgadmin(1M)

名称

srgadmin - 運用管理コマンド

構文

```
srgadmin [-c status [-f devfile] [-t time]]  
[-c start] [-c stop] [-c sync]  
[-i]
```

機能説明

srgadmin は、StorageSaver を制御する運用管理コマンドです。

オプション

srgadmin のオプションは以下のとおりです。

- | | |
|----------------------------------|---|
| -c status [-f devfile] [-t time] | I/O パスの監視状態を表示します。
-f オプションに PV のデバイスファイルを指定すると、指定した I/O パスの状態のみ表示します。
-t オプションに時間 (単位は秒) を指定すると、指定した時間ごとに status を表示します。
-t オプションを省略すると、srgadmin は status を一度だけ表示して終了します。 |
| -c start | すべての I/O パスの監視を開始します。 |
| -c stop | すべての I/O パスの監視を停止します。 |
| -c sync | すべての I/O パスの Online status (組み込み状態)を現在の状態と同期します。 |
| -i | I/O パス上の情報を詳細表示します。 |

実行例

- すべての I/O パスの状態を表示します。

```
# /opt/HA/SrG/bin/srgadmin -c status
```

- すべての I/O パスの状態を30秒間隔で表示します。

```
# /opt/HA/SrG/bin/srgadmin -c status -t 30
```

関連項目

srgd(1M)

srgwatch(1M)

名称

srgwatch - プロセス監視デーモン

構文

srgwatch

機能説明

srgwatch は StorageSaver の常駐プロセスを監視するデーモンです。

srgwatch は監視対象プロセスの動作状態を、一定時間ごとに監視します。

監視対象プロセスが存在しないことを検出すると、ただちに指定されたコマンドを実行し、監視対象プロセスの再起動を行います。

srgwatch で監視するプロセスの情報は、設定ファイルに記述します。

設定ファイルの記述方法については、次項で説明します。

設定ファイル

以下に、設定ファイルの形式について説明します。

- 設定ファイルは、/var/opt/HA/SrG/local/conf/srgwatch.config です。変更はできません。
- コメントを記述する場合、"#" または ";" を使用します。コメント文字から改行までをコメントとして解釈します。
- 一つの監視対象プロセスごとに、設定ファイルに以下の形式で記述します。
- 設定値設定の記述を省略した場合、default 値が使用されます。

```
title {
    PROCNAME    =   監視対象プロセス名
    EXECCMD      =   再起動コマンド
    EXECLOCK     =   起動ロックファイル
    INTERVAL     =   監視間隔 (単位: 分)
    WATCH       =   監視要否
    HUP          =   送信シグナル
    RETRY        =   リトライ回数
}
```

以下に、各設定値について説明します。

title {..}

"{" から "}" に囲まれた区間を、一つの監視対象プロセスに関する設定情報とします。

"{" の前にある title には、対象プロセスを表す任意の文字列を記述してください。

最大 255 文字まで指定可能です。256 文字以降は認識されません。

PROCNAME	監視対象プロセス名を記述します。起動パスも含めたプロセス名を指定してください。 最大 255 文字まで指定可能です。256 文字以降は認識されません。 PROCNAME の指定を省略することはできません。
EXECCMD	PROCNAME に指定されたプロセス名が存在しない場合に、実行するコマンドを記述します。 EXECCMD に記述した文字列は、そのまま標準シェル (/bin/sh) 上で実行されます。 最大 511 文字まで指定可能です。512 文字以降は認識されません。 EXECCMD を省略した場合、ログファイルにプロセスが存在しないというメッセージだけが採取され、プロセスの再起動は行いません。
EXECLOCK	起動ロックファイルのファイル名を指定します。ファイルは絶対パスで指定してください。 srgwatch は、PROCNAME に指定されたプロセスが存在しないことを検出した場合、起動ロックファイルがあれば、指定コマンドの実行を行いません。 これは、ユーザーが監視対象プロセスを意図的に停止させ、srgwatch から自動的に再起動されるのを防ぐ場合に使用します。 srgwatch 起動時に、各監視対象プロセスに記述された起動ロックファイルは消去されます。 EXECLOCK を省略した場合、常に起動ロックファイルはないと解釈されます。
INTERVAL	監視間隔を指定します。単位は分です。 INTERVAL を省略した場合、default 値である 1 分が設定されます。
WATCH	設定ファイルに指定したプロセスを、srgwatch の監視対象に含めるかどうかを指定します。 以下の値が指定可能です。 WATCHON 設定ファイルに記述した内容にしたがい、srgwatch は対象プロセスの監視を行います。 IGNORE 設定ファイルに記述した情報を無視します。srgwatch は、対象プロセスの監視を行いません。 WATCH を省略した場合、default 値である IGNORE が設定されます。

HUP	srgwatch が SIGHUP シグナルを受信した場合、その延長で、監視対象プロセスにもシグナルを送信するかどうかを指定します。 HUP には以下の値が指定可能です。 ACTION_NONE 何も行いません。 SIG_HUP 監視対象プロセスに対して、SIGHUP シグナルを送信します。 SIG_QUIT 監視対象プロセスに対して、SIGQUIT シグナルを送信します。 SIG_KILL 監視対象プロセスに対して、SIGKILL シグナルを送信します。 HUP を省略した場合、default 値である ACTION_NONE が設定されます。
RETRY	指定コマンドの実行に連続して失敗した場合、再起動を試みる回数の上限値を指定します。 0 を指定すると、対象プロセスの起動が確認されるまで、無限に繰り返します。 RETRY を省略した場合、default 値である 10 が設定されます。

- 複数のプロセスを監視する場合、監視対象プロセス毎の設定値を設定ファイルに記述します。
- 設定値を記述する順番は、特に制約はありません。
- 区間内に同一の設定情報名が存在する場合、最後に記述されたものが有効になります。

注意事項

- srgwatch は、スーパーユーザーのみ実行可能です。
- srgwatch は監視対象プロセスが存在しないことを検出した場合、ただちに指定されたコマンドを実行しますが、次の監視対象プロセスの検索は次回の監視間隔時間経過時に行います。したがって、例えば毎回 3 分で終了するアプリケーションの監視について "INTERVAL=5" と指定した場合、srgwatch はプロセスの再起動に失敗したと解釈します。

関連ファイル

/var/opt/HA/SrG/local/conf/srgwatch.config srgwatch の設定ファイル

実行例

- srgwatch を起動します。

```
# /etc/init.d/srgctl start
```

- srgwatch を停止します。

```
# /etc/init.d/srgctl stop
```

設定ファイル記述例

<srgd というプロセスを監視する例>

```
srgd{
;; プロセス名を絶対パスから記述。
PROCNAME = /opt/HA/SrG/bin/srgd
;; 再起動コマンドを指定。標準エラー出力をファイルに採取する例。
EXECCMD  = /opt/HA/SrG/bin/srgd >> /tmp/srgd.log 2>&1
;; .srgd_lock ファイルが存在する場合、再起動は行わない。
EXECLOCK = /var/opt/HA/SrG/conf/.srgd_lock
;; 監視間隔は 1 分に 1 回。
INTERVAL = 1
;; この設定ファイル情報を有効にする。
WATCH    = WATCHON
;; srgwatch が SIGHUP を受けたら、srgd には SIGKILL を送る。
HUP      = SIG_KILL
;; 再起動に失敗しても、連続 10 回まで再起動を試みる。
RETRY    = 10
}
```

srgextend(1M)

名称

srgextend - StorageSaver の監視リソースの復旧機能

構文

```
srgextend -H FCname|HWpath
srgextend -F udev_devicefile
```

機能説明

srgextend は、指定された HWpath、FC 名、PV 名に対応する I/O パスを復旧します。

オプション

srgextend のオプションは以下のとおりです。

- | | |
|--------------------|--|
| -H HWpath | HBA カードの HWpath を指定します。
入力で指定された HWpath 配下の全 I/O パスが対象となります。
HBA カードの Hwpath は udev デバイスファイル名の
以下の斜体部分です。
pci- <u>0000:0b:00.0</u> -fc-0x2001000013840322:0x0000000000000000 |
| -H FCname | HWpath の代わりに FC 名を指定することができます。 |
| -F udev_devicefile | PV の udev デバイスファイル名を指定します。
入力で指定された PV に繋がる全 I/O パスが対象となります。 |

終了ステータス

成功すると 0 を返し、失敗するとそれ以外を返します。

注意事項

srgextend を実行するとオンライン保守中を示す以下のファイルを削除します。

/var/opt/HA/SrG/conf/.online_maintaining

自動復旧機能を使用していた場合には自動復旧機能を再開します。

複数の PV が閉塞された状態で 1PV のみ指定して srgextend を実行した場合、
上記ファイルが削除されるため自動復旧機能が再開されます。
その場合、保守作業が完了していない状態で PV が復旧される可能性があります。
PV 単位で閉塞、復旧を行う場合には PV 毎に閉塞、復旧を行うか、
すべての保守作業が完了した時点で srgrecover コマンドにより
全 PV を一括で復旧させてください。

実行例

- HBA カード 0000:0a:00.0 配下の I/O パスを復旧します。

```
# /opt/HA/SrG/bin/srgextend -H 0000:0a:00.0
```

- FC 名 fc1 配下の I/O パスを復旧します。

```
# /opt/HA/SrG/bin/srgextend -H fc1
```

- PV 名 pci-0000:0a:00.0-fc-0x2100001697120ca7:0x0000000000000000 配下の I/O パスを復旧します。

```
# /opt/HA/SrG/bin/srgextend -F pci-0000:0a:00.0-fc-0x2100001697120ca7:0x0000000000000000
```

関連項目

srgreduce(1M), srgrecover(1M)

srgreduce (1M)

名称

srgreduce - StorageSaver の監視リソースの閉塞機能

構文

```
srgreduce -H FCname|HWpath
srgreduce -F udev_devicefile
```

機能説明

srgreduce は、指定された HWpath、FC 名、PV 名に対応する I/O パスを閉塞します。

オプション

srgreduce のオプションは以下のとおりです。

- | | |
|--------------------|--|
| -H HWpath | HBA カードの HWpath を指定します。
入力で指定された HWpath 配下の全 I/O パスが対象となります。
HBA カードの Hwpath は udev デバイスファイル名の
以下の斜体部分です。
pci- <u>0000:0b:00.0</u> -fc-0x2001000013840322:0x0000000000000000 |
| -H FCname | HWpath の代わりに FC 名を指定することができます。 |
| -F udev_devicefile | PV の udev デバイスファイル名を指定します。
入力で指定された PV に繋がる全 I/O パスが対象となります。 |

終了ステータス

成功すると 0 を返し、失敗するとそれ以外を返します。

注意事項

srgreduce を実行するとオンライン保守中を示す以下のファイルを作成します。

/var/opt/HA/SrG/conf/.online_maintaining

自動復旧機能を使用していた場合には自動復旧機能を一時停止します。

srgreduce を中断するとオンライン保守中を示すファイルが作成されたまま残ってしまい、自動復旧機能を使用していた場合には自動復旧機能が動作できなくなる可能性があります。srgreduce を中断した場合は、コマンドを再度実行してパスの閉塞を行うか、srgextend もしくは srgrecover を実行してパスを復旧させてください。

実行例

- HBA カード 0000:0a:00.0 配下の I/O パスを閉塞します。

```
# /opt/HA/SrG/bin/srgreduce -H 0000:0a:00.0
```

- FC 名 fc1 配下の I/O パスを閉塞します。

```
# /opt/HA/SrG/bin/srgreduce -H fc1
```

- PV 名 pci-0000:0a:00.0-fc-0x2100001697120ca7:0x0000000000000000 配下の I/O パスを閉塞します。

```
# /opt/HA/SrG/bin/srgreduce -F pci-0000:0a:00.0-fc-0x2100001697120ca7:0x0000000000000000
```

関連項目

srgextend(1M), srgrecover(1M)

srgrecover(1M)

名称

srgrecover - StorageSaver の監視リソースの構成復旧コマンド

構文

```
srgrecover [-v]
srgrecover [-v] [-H FCname|FC_HWpath]
srgrecover [-v] [-F udev_devicefile]
srgrecover [-n]
```

機能説明

srgrecover は、指定された HWpath、FC 名、PV 名に対応する I/O パスを復旧します。

復旧されたことは srgadmin コマンドの状態表示オプションで確認します。

オプション

srgrecover のオプションは以下のとおりです。

- | | |
|--------------------|--|
| -v | I/O パスの復旧結果を表示します。 |
| -n | syslog へのコマンド実行結果の出力を抑止します。 |
| -H FC_HWpath | HBA カードの HWpath を指定します。
入力で指定された HWpath 配下の全 I/O パスが対象となります。
HBA カードの HWpath は udev デバイスファイル名の
以下の斜体部分です。
pci- <u>0000:0b:00.0</u> -fc-0x2001000013840322:0x0000000000000000 |
| -H FCname | HWpath の代わりに FC 名を指定することができます。 |
| -F udev_devicefile | PV の udev デバイスファイル名を指定します。
入力で指定された PV に繋がる全 I/O パスが対象となります。 |

終了ステータス

成功すると 0 を返し、失敗するとそれ以外を返します。

注意事項

srgrecover を実行するとオンライン保守中を示す以下のファイルを削除します。

/var/opt/HA/SrG/conf/.online_maintaining

自動復旧機能を使用していた場合には自動復旧機能を再開します。

複数の PV が閉塞された状態で 1PV のみ指定して srgrecover を実行した場合、上記ファイルが削除されるため自動復旧機能が再開されます。

その場合、保守作業が完了していない状態で PV が復旧される可能性があります。

PV 単位で閉塞、復旧を行う場合には PV 毎に閉塞、復旧を行うか、すべての保守作業が完了した時点で srgrecover コマンドにより全 PV を一括で復旧させてください。

実行例

- すべての I/O パスに対して構成復旧を実行します。

```
# /opt/HA/SrG/bin/srgrecover -v
```

自動復旧機能について

ディスクアレイ装置の故障等により障害が発生し、装置交換を行った場合、ハードウェアとして I/O パスを復旧させた後に、手動で srgrecover コマンドを実行して StorageSaver の管理ステータスを復旧する必要があります。

自動復旧機能を使用すると、I/O パスの状態を定期的にチェックし、障害状態から復旧した I/O パスを検出すると、srgrecover コマンドを実行して復旧した I/O パスを自動的に組み込みます。

ただし、一時的に I/O が通るような間欠故障が発生した場合に、自動復旧機能を使用して I/O パスの自動組み込みを行うと、I/O の切り替えが頻発することによる I/O の遅延が発生する可能性があります。

そのため、障害が発生した場合には障害箇所の点検をおこない、確実に I/O パスの状態が復旧したことを確認した後に、手動で srgrecover コマンドを実行する運用を推奨します。

運用上オペレータの介入が困難であるなど、復旧作業を自動化せざるをえない場合はシステム定義ファイルの以下のパラメータを ENABLE に変更し、デーモンプロセスを再起動することで、自動復旧機能を利用することができます。

- システム定義ファイル (/var/opt/HA/SrG/conf/srg.config) の変更

```
# Auto recovery flag
#      used = ENABLE : unused = DISABLE(default)
AUTO_RECOVERY      ENABLE      ← ENABLE に変更します。
```

- デーモンプロセスの再起動

システム定義ファイルの変更後、デーモンプロセスを再起動します。

```
# /opt/HA/SrG/bin/srgconfig -r
```

関連項目

srgextend(1M), srgreduce(1M)

srgconfig(1M)

名称

srgconfig - StorageSaver の設定ファイルの確認、適用コマンド

構文

```
srgconfig -c [-s 適用対象ディレクトリ]
srgconfig -a [-s 適用対象ディレクトリ]
srgconfig -r
```

機能説明

設定ファイルの妥当性を確認し、実行環境に適用します。

オプション

srgconfig のオプションは以下のとおりです。

- c 指定されたディレクトリにある設定ファイルの妥当性、整合性をチェックします。
- a 指定されたディレクトリにある設定ファイルを srg 実行環境に適用します。
この場合 srg デーモンの再起動契機で設定ファイルが有効となります。
また、新しく作成した設定ファイルを適用する前に現在の設定ファイルの
バックアップを /var/opt/HA/SrG/conf/backup 配下に作成します。
- r デーモンプロセスを再起動します。
- s 対象となる設定ファイルのディレクトリを指定します。

一番目の書式および二番目の書式では、-s オプションを省略した場合
カレントディレクトリ配下の設定ファイルを対象とします。

終了ステータス

成功すると 0 を返し、失敗するとそれ以外を返します。

注意事項

特になし。

実行例

- ・/tmp 配下に作成した設定ファイルの妥当性を確認する。

```
# /opt/HA/SrG/bin/srgconfig -c -s /tmp
```

- ・/tmp 配下に作成した設定ファイルを実行環境に適用する。

```
# /opt/HA/SrG/bin/srgconfig -a -s /tmp
```

- ・デーモンプロセスを再起動する。

```
# /opt/HA/SrG/bin/srgconfig -r
```

関連項目

srgd(1M), srgquery(1M)

srgquery(1M)

名称

srgquery - StorageSaver の設定ファイルテンプレート自動生成コマンド

構文

srgquery [-s sgfile directory]

srgquery [-a] [-s sgfile directory]

機能説明

StorageSaver の設定ファイルのテンプレートを自動生成します。

オプション

srgquery のオプションは以下のとおりです。

- s 設定ファイルのテンプレートを作成するディレクトリを指定します。
-s 指定のない場合は、カレントディレクトリにテンプレートを作成します。
- a FC の配下だけでなく、SCSI 配下の I/O パスについてもテンプレートに組み込みます。
-a 指定のない場合は、FC 配下の I/O パス のみテンプレートに組み込みます。

終了ステータス

成功すると 0 を返し、失敗するとそれ以外を返します。

注意事項

- 本コマンドは、sg_scan(1M)、vgdisplay(1M)、powermt(1)コマンドの実行結果および/proc の情報を参照して、テンプレートを作成します。
コマンドの実行結果が不完全であれば、自動生成したテンプレートも不完全となりますので手作業で更新してください。
- 本コマンドは対象 I/O パスが LVM の VG に含まれていない場合、管理上"VG_NONE"という仮想的な VG 名で登録しますので、LVM の VG 名に"VG_NONE"を使用しないようにしてください。
- SANboot 環境において設定ファイル自動生成コマンド (srgquery) にて設定ファイルを生成した場合、OS 情報が格納されたディスク装置も FC 接続構成であるため、他のデータディスク装置同様、監視対象として定義します。
設定ファイルから OS 情報が格納されたディスク装置の記述を手動で削除してください。
- IDE 接続のデバイスが接続されている環境において設定ファイル自動生成コマンド (srgquery) にて設定ファイルを生成した際、IDE 接続のデバイスを監視対象として定義する場合があります。
IDE 接続のデバイスが設定ファイルに含まれている場合は手動で削除してください。
- srgquery コマンドにて設定ファイルの自動生成を行った際、マルチパス管理製品 (SPS, PowerPath) による I/O パス冗長設定が完了していない可能性がある場合、以下の警告メッセージを出力します。
srgquery: <デバイスファイル> is not found in StoragePathSavior.
srgquery: <デバイスファイル> is not found in PowerPath.
マルチパス管理製品 (SPS, PowerPath) の I/O パス冗長設定が完了していない場合は設定完了後、再度 srgquery コマンドにて設定ファイルの自動生成を行ってください。

実行例

/tmp 配下に、設定ファイルのテンプレートを作成する。

```
# /opt/HA/SrG/bin/srgquery -s /tmp
```

ファイル

srg.config	システム定義ファイル
srg.map	構成定義ファイル
srg.rsc	リソース定義ファイル

関連項目

srgconfig(1M),srgd(1M)

srgstat(1M)

名称

srgstat - クラスタウェア連携用コマンド

構文

srgstat [-h] [-v] [-V vgname] [-w wait-time]

機能説明

StorageSaver がレポートするリソースステータスを監視します。
srgstat のプロセス状態、あるいは終了ステータスを参照することで、
StorageSaver の監視対象リソースの状態を知ることができます。

-w オプションを指定すると srgstat は常駐して StorageSaver がレポートするステータスを監視し続け、指定された VG が異常状態(down)になると異常終了します。
CLUSTERPRO と連携する場合のカスタムモニタリソースとして有効です。
StorageSaver のプロセスが動作していない場合、また srgadmin(1M) でリソース監視の停止を指示された場合は、VG の異常を検出できないため、正常状態として報告します。

-w オプションを指定しない場合は、一回だけ StorageSaver がレポートするステータスを調べ終了します。

オプション

srgstat のオプションは以下のとおりです。

- | | |
|--------------|---|
| -h | コマンドの説明を表示します。 |
| -v | 動作トレースを標準出力に表示します。 |
| -V vgname | 監視したい VG 名を指定します。
対象 VG 名を一つだけ指定できます。
-V を指定しない場合は、StorageSaver が監視する全ての VG が監視対象になります。
複数の VG が監視対象になっている場合、少なくとも一つの VG で
DOWN が検出された時点で、srgstat は「StorageSaver の監視対象ディスクが異常」と判断します。 |
| -w wait-time | ディスク監視を常駐させたい場合に指定します。
wait-time には、監視間隔の時間(単位は秒)を指定します。
1 以上の値を指定してください。
もし 0 を指定した場合は、強制的に 1 に補正されます。 |

終了ステータス

-w オプションと StorageSaver がレポートするステータスによって、以下のように動作します。

-w オプション有りの場合:

StorageSaver が監視するディスクの状態	srgstat コマンド	終了ステータス
ディスク正常時	終了せずに常駐	—
ディスク異常時	終了	1 を返す
StorageSaver が動作していない	終了せずに常駐	—

-w オプション無しの場合:

StorageSaver が監視するディスクの状態	srgstat コマンド	終了ステータス
ディスク正常時	終了	0 を返す
ディスク異常時	終了	1 を返す
StorageSaver が動作していない	終了	2 を返す

注意事項

- 本コマンドはメモリ上に常駐します。swap 領域に退避されません。
- StorageSaver のプロセスを起動していない状態、srgadmin(1M) でリソース監視の停止を指示された状態であれば VG リソースの異常を検出できません。

実行例

ヘルプを表示する。

```
# /opt/HA/SrG/bin/srgstat -h
```

動作の詳細なトレースを表示します。監視対象ディスクは、StorageSaver が監視する全ての VG です。

```
# /opt/HA/SrG/bin/srgstat -v
```

VolGroup01 のみを監視対象とします。

```
# /opt/HA/SrG/bin/srgstat -V VolGroup01
```

srgstat を常駐させ、StorageSaver のステータスを 5 秒間隔で調べます。

```
# /opt/HA/SrG/bin/srgstat -w 5
```

関連項目

srgadmin(1M),srgd(1M)

10.付録

10.1. 運用管理コマンド

StorageSaver の運用管理コマンドの操作手順は下記のとおりです。

(1) デーモンプロセス起動

コマンドラインからデーモンプロセスを起動する場合は以下を実行します。

```
# /etc/init.d/srgctl start
```

注意: 通常は rc ファイルから自動起動されます

起動状態を確認

```
# ps -ef|grep srg
root  8720      8098   0 10 月  9  ?        0:40 srgping
root  8098         1   0 10 月  9  ?        0:19 /opt/HA/SrG/bin/srgd
root  8111         1   0 10 月  9  ?        0:54 /opt/HA/SrG/local/bin/srgwatch
```

注意: srgping は srgd を起動後、約1分後に起動されます。

(2) デーモンプロセス終了

コマンドラインからデーモンプロセスを終了する場合は以下を実行します。

```
# /etc/init.d/srgctl stop
```

注意: 通常は rc ファイルから自動終了します

終了状態を確認

```
# ps -ef|grep srg
```

前述の“srgxxx”プロセスが表示されていないことを確認してください。

(3) 設定値の参照

システム定義ファイル /var/opt/HA/SrG/conf/srg.config

構成定義ファイル /var/opt/HA/SrG/conf/srg.map

リソース定義ファイル /var/opt/HA/SrG/conf/srg.rsc

(4) 監視リソースの状態確認

- リソース一覧の表示

```
# /opt/HA/SrG/bin/srgadmin -c status
```

- リソース詳細表示

```
# /opt/HA/SrG/bin/srgadmin -i
```

(5) 監視リソースの一括復旧

```
# /opt/HA/SrG/bin/srgrecover -v
```

注意: 自動閉塞等により障害となった場合の復旧に使用します

(6) オンライン保守

- 障害箇所の閉塞

```
# /opt/HA/SrG/bin/srgreduce -H <FCパス>
```

- 状態表示

```
# /opt/HA/SrG/bin/srgadmin -c status
```

注意: reduced になっていることを確認します。

- オンライン保守実施

この間、障害部位への Test/I/O, LVM からの I/O が停止

- 障害箇所の復旧

```
# /opt/HA/SrG/bin/srgrecover -v
```

注意: -v オプションは復旧結果を表示します。

- 状態表示

```
# /opt/HA/SrG/bin/srgadmin -c status
```

注意: extended になっていることを確認します。

10.2. srgquery(1M) による設定ファイル自動生成手順

クラスタウェアで共有ディスク上に VG を構築している場合は、
下記の手順で設定ファイルを自動生成します。

以降はクラスタウェアとして CLUSTERPRO を例に説明します。

なお、下記の手順はクラスタ内のすべてのノードで実行してください。

(1) 設定ファイルの自動生成

```
# cd /tmp
```

```
# /opt/HA/SrG/bin/srgquery -s ./
```

CLUSTERPRO と連携を行う場合は、自動生成後に、srg.config の VG_FAULT_ACTION の値を
CLPNM_KILL_ENABLE に変更してください。

自動生成後、srg.config、srg.map、srg.rsc の内容を確認します。

(2) 設定ファイルの妥当性確認

```
# /opt/HA/SrG/bin/srgconfig -c -s ./
```

(3) 設定ファイルの適用

```
# /opt/HA/SrG/bin/srgconfig -a -s ./
```

(4) デーモンプロセス再起動

```
# /opt/HA/SrG/bin/srgconfig -r
# ps -ef | grep srg
root 8720      8098   0 10月 9   ?        0:40 srgping
root 8098          1   0 10月 9   ?        0:19 /opt/HA/SrG/bin/srgd
root 8111          1   0 10月 9   ?        0:54 /opt/HA/SrG/local/bin/srgwatch
```

注意:srgping は srgd を起動後、約1分後に起動されます。

(5) 監視リソースの状態が正常であることの確認

```
# /opt/HA/SrG/bin/srgadmin - c status  
(monitor status = TRUE)
```

type	device	HostBusAdapter	L status	P status	Online status
VG	VolGroup01	—	up		
PV	/dev/sdj	pci-0000:13:00.0	up	up	extended
PV	/dev/sdh	pci-0000:13:00.0	up	up	extended
PV	/dev/sdf	pci-0000:13:00.1	up	up	extended
PV	/dev/sdd	pci-0000:13:00.1	up	up	extended
VG	VolGroup02	—	up		
PV	/dev/sdk	pci-0000:13:00.0	up	up	extended
PV	/dev/sdi	pci-0000:13:00.0	up	up	extended
PV	/dev/sgd	pci-0000:13:00.1	up	up	extended
PV	/dev/sde	pci-0000:13:00.1	up	up	extended

10.3. 本製品のテスト手順について

(1) はじめに

StorageSaver を導入するシステムにおいて、設定ファイルの検証および性能チューニングの検証を擬似的に行う手順を説明します。

コマンドオペレーションでディスク障害を擬似的に発生することにより、物理ディスクの抜き差し等の操作を行う必要がなくなり、システムへ影響を与えず評価が実現できます。

注意: 本作業は必ずスーパーユーザーアカウント (root 権限) で実施してください。

本作業は、ディスクアレイ装置が正しく構築されていることを確認した後に実施してください。

クラスタウェア 連携を行う場合、クラスタウェア が正しく設定されていることを確認した後に実施してください。

● StorageSaver の擬似評価のパターン

擬似障害試験のパターンは以下のとおりです。

1. I/O パスの片系故障
2. I/O パスの両系故障
3. I/O ストール障害

運用管理コマンドをデバッグモードで実行することで、TestI/O の実行結果を擬似的に異常に遷移させ、監視ステータスを up から down に変更します。

さらに両系の I/O パスを異常に変更することで VG ステータスを down 状態に擬似することができます。

● コマンド書式

擬似障害を発生させるためのコマンドラインは以下のとおりです。

```
srgadmin -c debug -v [ on | off ] [ -f devicefile | -H HW path | -F fc_hwpath ]
```

on I/O status modify down I/O ステータスを down にします。

off I/O status modify up I/O ステータスを up にします。

また、以下のオプションがあります。

-H HW パス(udev デバイスファイル)ごとに切り替えることができます。

-F FC 単位 で切り替えることができます。指定できるのは FC の HW パスです。

(2) I/O パスの片系故障

マルチパス構成の I/O パスの片系障害を擬似的に発生させる手順を説明します。

1. 障害前に現在の状態をモニタリングします。

```
# /opt/HA/SrG/bin/srgadmin - c status  
(monitor status = TRUE)
```

type	device	HostBusAdapter	L status	P status	Online status
VG	VolGroup01	—	up		
PV	/dev/sdj	pci-0000:13:00.0	up	up	extended
PV	/dev/sdh	pci-0000:13:00.0	up	up	extended
PV	/dev/sdf	pci-0000:13:00.1	up	up	extended
PV	/dev/sdd	pci-0000:13:00.1	up	up	extended
VG	VolGroup02	—	up		
PV	/dev/sdk	pci-0000:13:00.0	up	up	extended
PV	/dev/sdi	pci-0000:13:00.0	up	up	extended
PV	/dev/sdg	pci-0000:13:00.1	up	up	extended
PV	/dev/sde	pci-0000:13:00.1	up	up	extended

2. `srgadmin -c debug` にて、片系のディスクを指定します。
指定されたディスクは強制的に I/O ステータスが up から down に変更され擬似的に片系障害を起こすことができます。

```
# /opt/HA/SrG/bin/srgadmin -c debug -v on -F 0000:13:00.1
Change debug value.
FC devfile = 0000:13:00.1
0 -> 1
FC devfile = 0000:13:00.1
0 -> 1
FC devfile = 0000:13:00.1
0 -> 1
FC devfile = 0000:13:00.1
0 -> 1
```

約 180 秒後

```
# /opt/HA/SrG/bin/srgadmin -c status
(monitor status = TRUE)
```

type	device	HostBusAdapter	L status	P status	Online status
VG	VolGroup01	—	<i>suspend</i>		
PV	/dev/sdj	pci-0000:13:00.0	up	up	extended
PV	/dev/sdh	pci-0000:13:00.0	up	up	extended
PV	/dev/sdf	pci-0000:13:00.1	<i>down</i>	<i>down</i>	extended
PV	/dev/sdd	pci-0000:13:00.1	<i>down</i>	<i>down</i>	extended
VG	VolGroup02	—	<i>suspend</i>		
PV	/dev/sdk	pci-0000:13:00.0	up	up	extended
PV	/dev/sdi	pci-0000:13:00.0	up	up	extended
PV	/dev/sdg	pci-0000:13:00.1	<i>down</i>	<i>down</i>	extended
PV	/dev/sde	pci-0000:13:00.1	<i>down</i>	<i>down</i>	extended

3. 片系障害を発生させた場合、`syslog` にメッセージが出力されます。
以下のメッセージが出力されることを確認してください。

```
srgd[xxxxx]: PV status change fail . [hwpath = pci-0000:13:00.1-fc-0x2100001697120ca7:0x0001000000000000: s. f = /dev/sdg].
srgd[xxxxx]: PV status change fail . [hwpath = pci-0000:13:00.1-fc-0x2900001697120ca7:0x0001000000000000: s. f = /dev/sde].
srgd[xxxxx]: PV status change fail . [hwpath = pci-0000:13:00.1-fc-0x2100001697120ca7:0x0000000000000000: s. f = /dev/sdf].
srgd[xxxxx]: PV status change fail . [hwpath = pci-0000:13:00.1-fc-0x2900001697120ca7:0x0000000000000000: s. f = /dev/sdd].
```

(3) I/O パスの両系故障

マルチパス構成の I/O パスの両系障害を擬似的に発生させる手順を説明します。
これによりクラスタウェア とのパッケージ連動の評価が可能になります。

1. 障害前に現在の状態をモニタリングします。既に片系の I/O パスが異常であることを確認します。

```
# /opt/HA/SrG/bin/srgadmin -c status  
(monitor status = TRUE)
```

type	device	HostBusAdapter	L status	P status	Online status
VG	VolGroup01	—	<i>suspend</i>		
PV	/dev/sdj	pci-0000:13:00.0	up	up	extended
PV	/dev/sdh	pci-0000:13:00.0	up	up	extended
PV	/dev/sdf	pci-0000:13:00.1	<i>down</i>	<i>down</i>	extended
PV	/dev/sdd	pci-0000:13:00.1	<i>down</i>	<i>down</i>	extended
VG	VolGroup02	—	<i>suspend</i>		
PV	/dev/sdk	pci-0000:13:00.0	up	up	extended
PV	/dev/sdi	pci-0000:13:00.0	up	up	extended
PV	/dev/sdg	pci-0000:13:00.1	<i>down</i>	<i>down</i>	extended
PV	/dev/sde	pci-0000:13:00.1	<i>down</i>	<i>down</i>	extended

2. 両系の I/O パスを異常状態にします。

```
# /opt/HA/SrG/bin/srgadmin -c debug -v on -F 0000:13:00.0
Change debug value.
FC devfile = 0000:13:00.0
0 -> 1
FC devfile = 0000:13:00.0
0 -> 1
FC devfile = 0000:13:00.0
0 -> 1
FC devfile = 0000:13:00.0
0 -> 1
```

約 60 秒後

```
# /opt/HA/SrG/bin/srgadmin -c status
(monitor status = TRUE)
```

type	device	HostBusAdapter	L status	P status	Online status
VG	VolGroup01	—	<i>down</i>		
PV	/dev/sdj	pci-0000:13:00.0	<i>down</i>	<i>down</i>	extended
PV	/dev/sdh	pci-0000:13:00.0	<i>down</i>	<i>down</i>	extended
PV	/dev/sdf	pci-0000:13:00.1	down	down	extended
PV	/dev/sdd	pci-0000:13:00.1	down	down	extended
VG	VolGroup02	—	<i>down</i>		
PV	/dev/sdk	pci-0000:13:00.0	<i>down</i>	<i>down</i>	extended
PV	/dev/sdi	pci-0000:13:00.0	<i>down</i>	<i>down</i>	extended
PV	/dev/sdg	pci-0000:13:00.1	down	down	extended
PV	/dev/sde	pci-0000:13:00.1	down	down	extended

3. 両系障害を発生させた場合、syslog にメッセージが出力されます。

以下のメッセージが出力されることを確認してください。

```
srgd[xxxxx]: VG status change down . (vg=VolGroup01)
srgd[xxxxx]: VG status change down . (vg=VolGroup02)
srgd[xxxxx]: PV status change fail . [hwpath = pci-0000:13:00.0-fc-0x2100001697120ca7:0x0001000000000000: s. f = /dev/sdk].
srgd[xxxxx]: PV status change fail . [hwpath = pci-0000:13:00.0-fc-0x2900001697120ca7:0x0001000000000000: s. f = /dev/sdi].
srgd[xxxxx]: PV status change fail . [hwpath = pci-0000:13:00.0-fc-0x2100001697120ca7:0x0000000000000000: s. f = /dev/sdj].
srgd[xxxxx]: PV status change fail . [hwpath = pci-0000:13:00.0-fc-0x2900001697120ca7:0x0000000000000000: s. f = /dev/sdh].
```

両系のディスクのステータスが down になったため、VG 状態も down となります。

注意: CLUSTERPRO の設定によってはサーバが shutdown または再起動するため、ステータスは確認することができない場合があります。その場合は設定に従った動作が実行されることを確認してください。

(4) テスト完了後の復旧

擬似障害テスト完了後の復旧手順を説明します。

1. マシンが shutdown した場合は、マシンを再起動して復旧させます。

注意: / (root) 配下に情報収集のため srgd の core が作成される場合がありますが
不要なため削除してください。

srgadmin コマンドにてデバッグモードを off にした後、srgrecover -v を実行して構成を復旧させます。

```
# /opt/HA/SrG/bin/srgadmin -c debug -v off -F 0000:13:00.1
Change debug value.
FC devfile = 0000:13:00.1
1 -> 0
FC devfile = 0000:13:00.1
1 -> 0
FC devfile = 0000:13:00.1
1 -> 0
FC devfile = 0000:13:00.1
1 -> 0
```

```
# /opt/HA/SrG/bin/srgadmin -c debug -v off -F 0000:13:00.0
Change debug value.
FC devfile = 0000:13:00.0
1 -> 0
FC devfile = 0000:13:00.0
1 -> 0
FC devfile = 0000:13:00.0
1 -> 0
FC devfile = 0000:13:00.0
1 -> 0
```

```
# /opt/HA/SrG/bin/srgrecover -v
HW-path: pci-0000:13:00.0-fc-0x2100001697120ca7:0x0000000000000000 result: OK
HW-path: pci-0000:13:00.0-fc-0x2100001697120ca7:0x0001000000000000 result: OK
HW-path: pci-0000:13:00.0-fc-0x2900001697120ca7:0x0000000000000000 result: OK
HW-path: pci-0000:13:00.0-fc-0x2900001697120ca7:0x0001000000000000 result: OK
HW-path: pci-0000:13:00.1-fc-0x2100001697120ca7:0x0000000000000000 result: OK
HW-path: pci-0000:13:00.1-fc-0x2100001697120ca7:0x0001000000000000 result: OK
HW-path: pci-0000:13:00.1-fc-0x2900001697120ca7:0x0000000000000000 result: OK
HW-path: pci-0000:13:00.1-fc-0x2900001697120ca7:0x0001000000000000 result: OK
srgrecover complete.
```


2. 復旧後、srgadmin コマンドにて正常にディスク監視が行われていることを確認します。

```
# /opt/HA/SrG/bin/srgadmin -c status
(monitor status = TRUE)
```

type	device	HostBusAdapter	L status	P status	Online status
VG	VolGroup01	—	up		
PV	/dev/sdj	pci-0000:13:00.0	up	up	extended
PV	/dev/sdh	pci-0000:13:00.0	up	up	extended
PV	/dev/sdf	pci-0000:13:00.1	up	up	extended
PV	/dev/sdd	pci-0000:13:00.1	up	up	extended
VG	VolGroup02	—	up		
PV	/dev/sdk	pci-0000:13:00.0	up	up	extended
PV	/dev/sdi	pci-0000:13:00.0	up	up	extended
PV	/dev/sdg	pci-0000:13:00.1	up	up	extended
PV	/dev/sde	pci-0000:13:00.1	up	up	extended

10.4. カスタムモニタリソースによる CLUSTERPRO との連携手順

10.4.1. CLUSTERPRO 連携設定

本製品は、カスタムモニタリソースによる CLUSTERPRO との連携を行うことが可能です。

以下の作業は、CLUSTERPRO WebManager にて実施します。

本書は、CLUSTERPRO Server をインストールしたサーバの実 IP アドレスを「192.168.11.100」、ポート番号を「29003(デフォルト値)」とした場合の例です。

接続例) <http://192.168.11.100:29003/>

また、本書では CLUSTERPRO WebManager のモニタリソースの設定を以下としています。

プロパティ	設定値
タイプ (モニタリソースのタイプ)	custom monitor
名前 (カスタムモニタリソース名)	srgstat_mon
監視タイプ	非同期
回復対象	LocalServer
回復動作	最終動作のみ実行
最終動作	クラスタサービス停止と OS シャットダウン

上記設定を行うことによって、VG 障害発生時にカスタムモニタリソース(srgstat_mon) が srgstat の消滅を検知して、現用系ノードを shutdown させた後、待機系へノード切り替えを行います。

※ 本書で設定している各種プロパティの値は一例です。構築時にはそれぞれの環境に応じた値を設定してください。

※ 以下は、CLUSTERPRO X 3.1 の場合の設定手順となります。

1. システム定義ファイルの確認

StorageSaver の設定を確認します。

- (1) StorageSaver のシステム定義ファイル (srg.config) にて VG_FAULT_ACTION が SERVICE_CMD_ENABLE になっていることを確認します。

```
# /bin/cat /var/opt/HA/SrG/conf/srg.config | /bin/grep "VG_FAULT_ACTION"
VG_FAULT_ACTION          SERVICE_CMD_ENABLE
```

※ 旧バージョンとの互換を保証するため、VG_FAULT_ACTION が ACTION_NONE の場合にも CLUSTERPRO 連携は可能です。

※ VG_FAULT_ACTION が SERVICE_CMD_ENABLE または ACTION_NONE になっていない場合は、以下 (2) ~ (4) の手順で設定変更およびデーモンプロセス(srgd)の再起動を行ってください。

- (2) システム定義ファイル (srg.config) を編集します。

※ 以下は vi コマンドを使用して編集を行う場合の例です。

```
# /bin/vi /var/opt/HA/SrG/conf/srg.config
```

- (3) システム定義ファイル (srg.config) が正しく変更されていることを確認します。

```
# /bin/cat /var/opt/HA/SrG/conf/srg.config | /bin/grep "VG_FAULT_ACTION"
VG_FAULT_ACTION          SERVICE_CMD_ENABLE
```

- (4) デーモンプロセス (srgd) を再起動します。

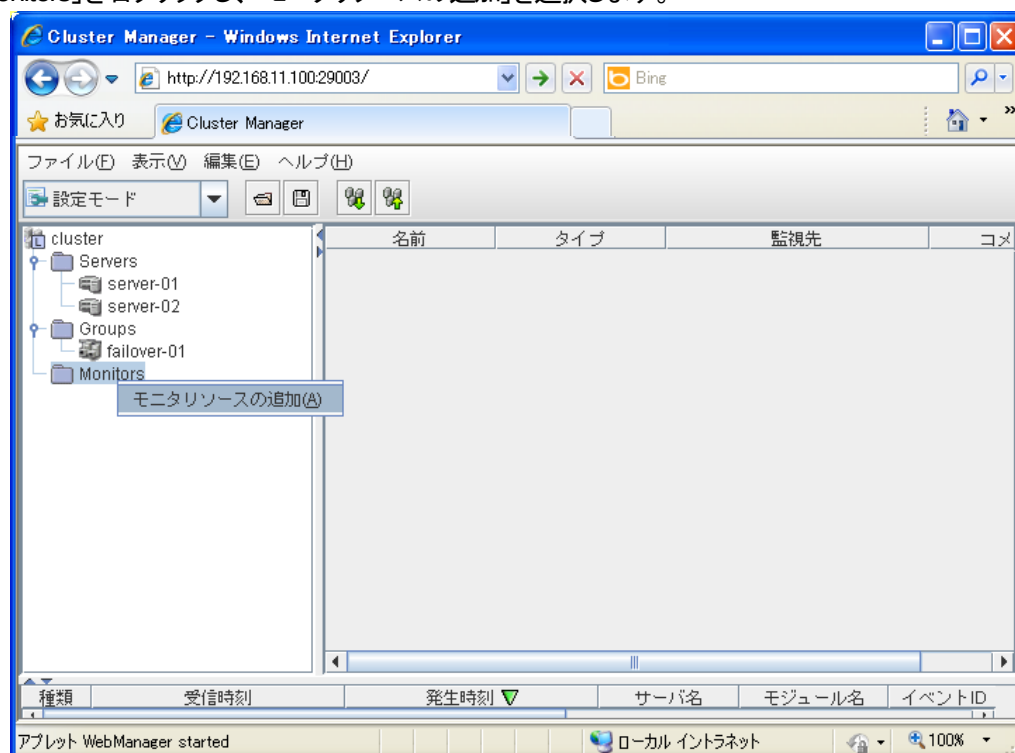
```
# /etc/init.d/srgctl restart
```

2. カスタムモニタリソースの作成

srgstat を監視するカスタムモニタリソースを作成します。

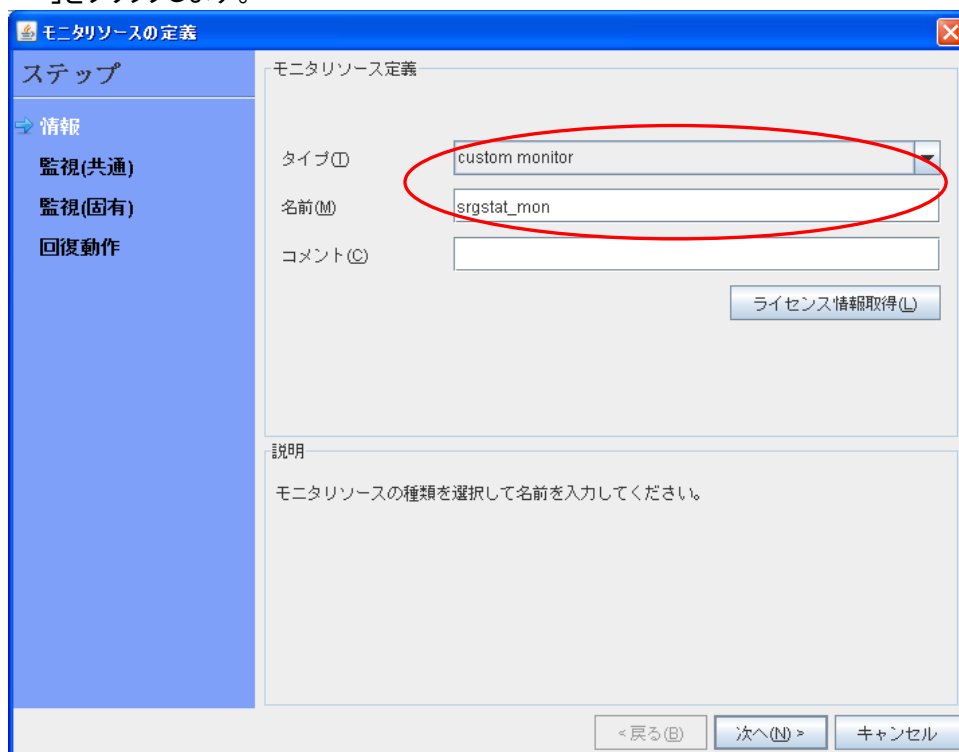
CLUSTERPRO WebManager の「表示」メニューより「設定モード」を選択し、設定します。

(1) 「Monitors」を右クリックし、「モニタリソースの追加」を選択します。



(2) 「モニタリソースの定義」ダイアログボックスが開きます。

「タイプ」で「custom monitor」を選択し、「名前」にカスタムモニタリソース名“srgstat_mon”を入力し、「次へ」をクリックします。



(3) 監視条件を設定します。

本書ではデフォルトのまま変更しません。「次へ」をクリックします。

(4) 監視条件を設定します。

「ユーザアプリケーション」を選択し、「ファイル」に `srgstat 起動処理 (/opt/HA/SrG/bin/srgstat -w 30)` を入力します。

監視タイプで「非同期」を選択し、「次へ」をクリックします。

(5) 回復動作を設定します。

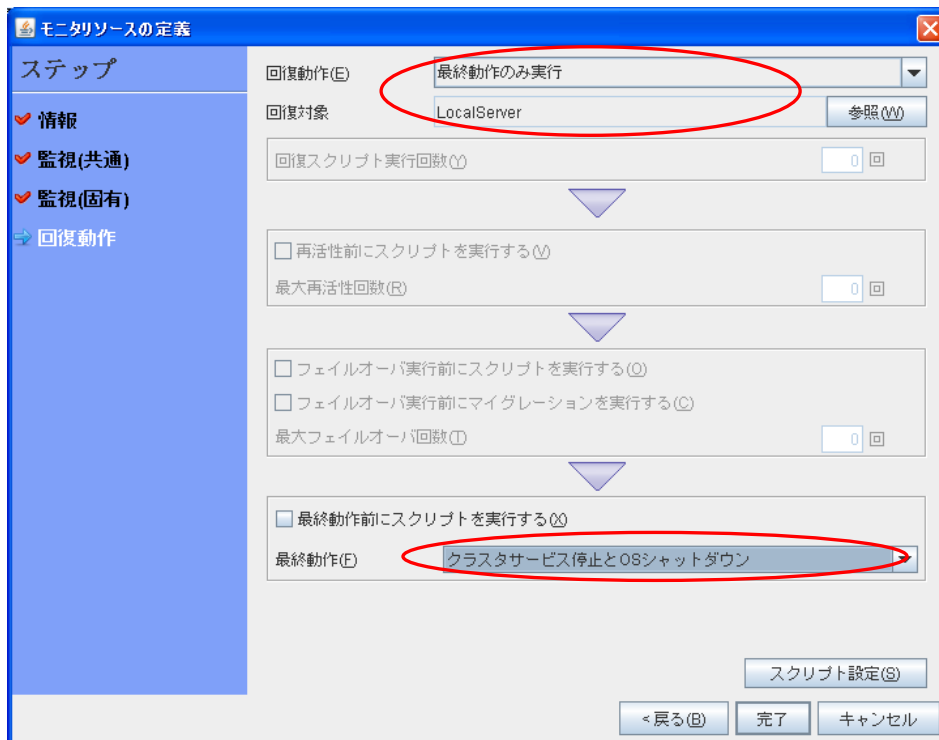
「回復動作」で「最終動作のみ実行」を選択します。

「回復対象」の「参照」をクリックし、表示されるツリービューで“LocalServer”を選択して

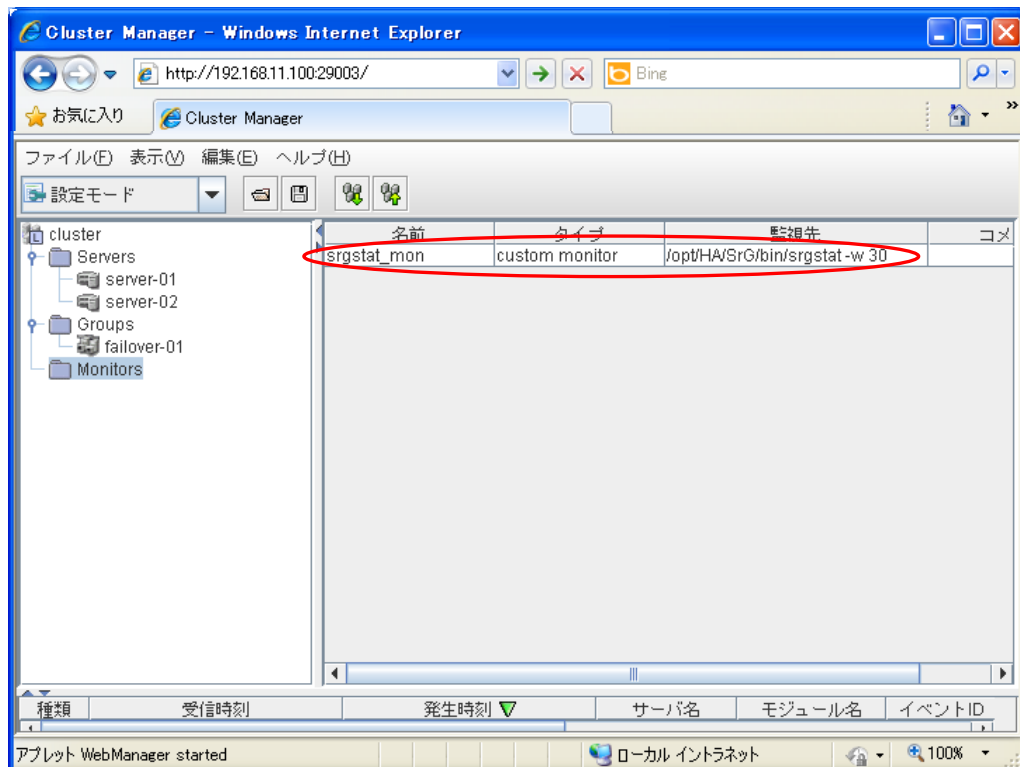
「OK」をクリックします。「回復対象」に“LocalServer”が追加されたことを確認します。

「最終動作」で「クラスタサービス停止と OS シャットダウン」を選択し、

「完了」をクリックします。



(6) 「Monitors」をクリックし、カスタムモニタリソースが登録されていることを確認します。

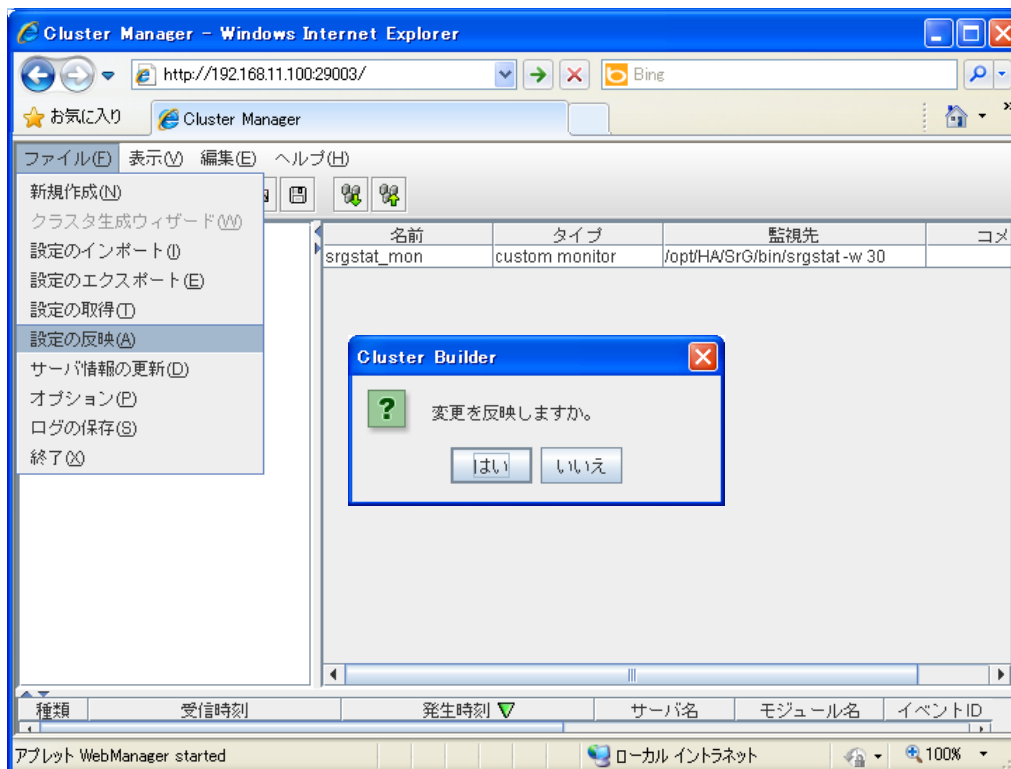


3. クラスタ構成情報のアップロード

(1) クラスタ構成情報の内容を、CLUSTERPRO 本体の環境に反映します。

「ファイル」メニューから「設定の反映」を選択します。

確認ダイアログが表示されます。「OK」をクリックします。



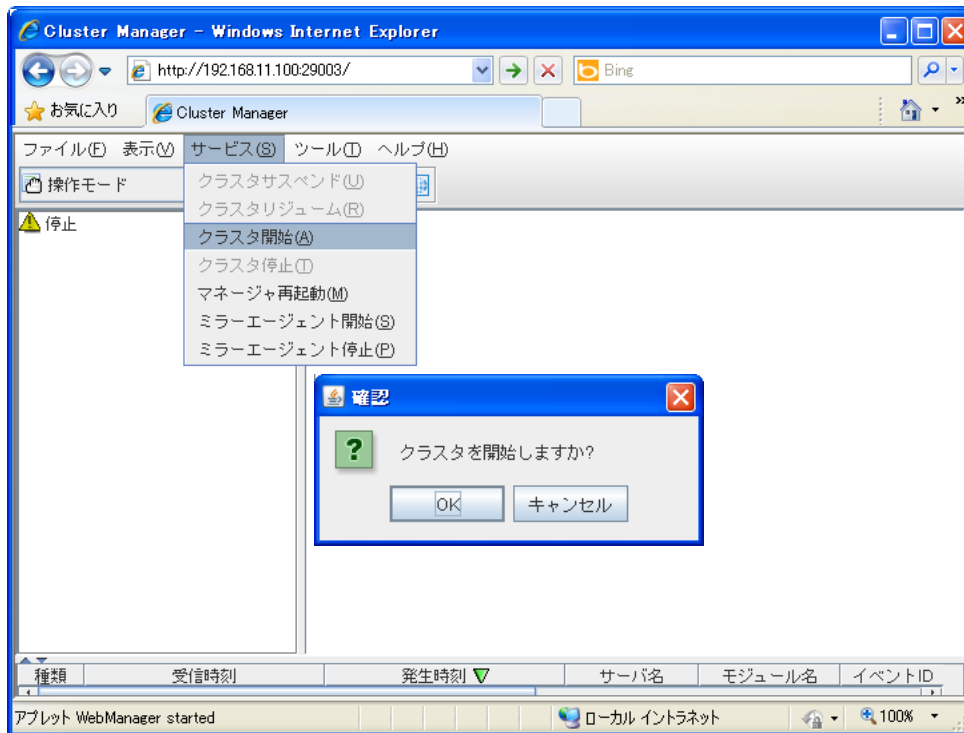
※ 変更した情報によって表示されるメッセージが異なりますので、表示されたメッセージにしたがって操作を行ってください。

詳細は、CLUSTERPRO X 3.1 for Linux マニュアルを参照してください。

反映に成功すると確認ダイアログが表示されます。「了解」をクリックしてダイアログを閉じます。

クラスタが停止状態の場合は、クラスタを開始します。

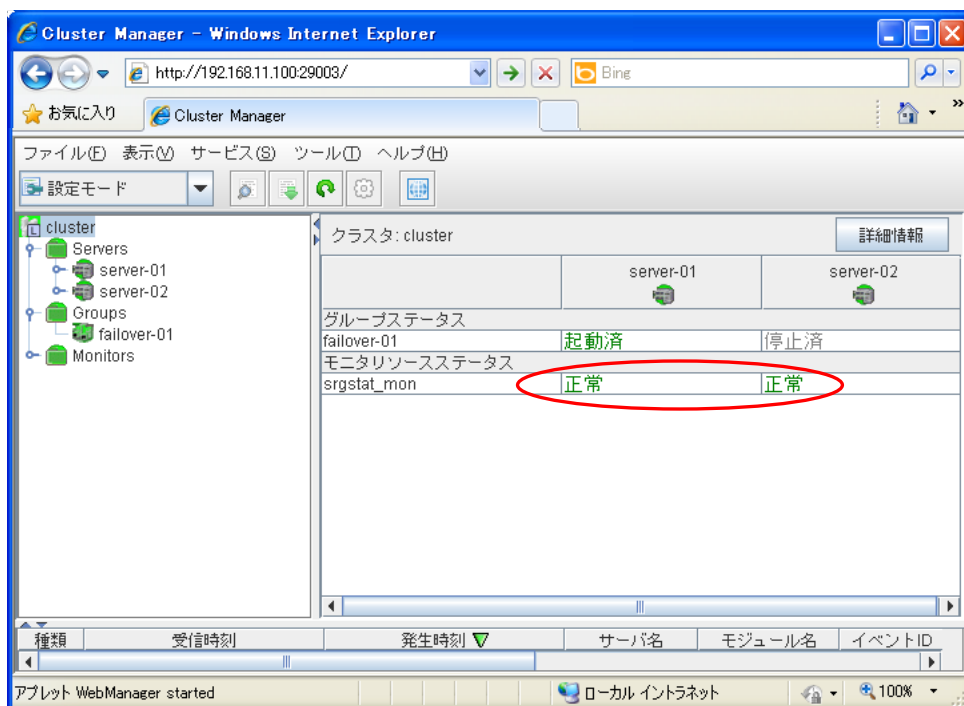
WebManager の「表示」メニューより「操作モード」を選択し、「サービス」メニューから「クラスタ開始」を選択し、クリックしてください。



(2) 設定が反映されていることを確認します。

WebManager の「表示」メニューより「操作モード」を選択し、以下の項目を確認してください。

- ・ 現用系サーバ、待機系サーバにて StorageSaver 監視用のカスタムモニタリソース「srgstat_mon」のステータスが「正常」である事を確認してください。



以上で、CLUSTERPRO の設定は終了です。

10.4.2. 動作確認

以降の手順で StorageSaver および CLUSTERPRO の設定の動作確認を行います。

FC 抜線により障害を発生させ、ディスクにアクセスすることができなくなった際にフェイルオーバーが発生することを確認します。

また、コマンドオペレーションでディスク障害を擬似的に発生させることも可能です。

擬似障害テスト手順については、「10.3 本製品のテスト手順について」の章を参照してください。

1. 現用系サーバでの StorageSaver 動作確認

FC ケーブルが 2 本接続されている環境における動作確認手順を記載します。

- (1) 片方の FC ケーブルを抜線し、片系障害を発生させます。
- (2) 約 180 秒後に片系障害を検出することを確認してください。

```
# /opt/HA/SrG/bin/srgadmin -c status
(monitor status = TRUE)
```

type	device	HostBusAdapter	L status	P status	Online status
VG	VolGroup01	—	<i>suspend</i>		
PV	/dev/sdj	pci-0000:13:00.0	up	up	extended
PV	/dev/sdh	pci-0000:13:00.0	up	up	extended
PV	/dev/sdf	pci-0000:13:00.1	<i>down</i>	<i>down</i>	extended
PV	/dev/sdd	pci-0000:13:00.1	<i>down</i>	<i>down</i>	extended
VG	VolGroup02	—	<i>suspend</i>		
PV	/dev/sdk	pci-0000:13:00.0	up	up	extended
PV	/dev/sdi	pci-0000:13:00.0	up	up	extended
PV	/dev/sgd	pci-0000:13:00.1	<i>down</i>	<i>down</i>	extended
PV	/dev/sde	pci-0000:13:00.1	<i>down</i>	<i>down</i>	extended

- (3) 片系障害を発生させた場合、syslog にメッセージが出力されます。
以下のメッセージが出力されることを確認してください。

```
srgd[xxxx]: PV status change fail . [hwp path = pci-0000:13:00.1-fc-0x2100001697120ca7:0x0001000000000000: s. f = /dev/sgd].
srgd[xxxx]: PV status change fail . [hwp path = pci-0000:13:00.1-fc-0x2900001697120ca7:0x0001000000000000: s. f = /dev/sde].
srgd[xxxx]: PV status change fail . [hwp path = pci-0000:13:00.1-fc-0x2100001697120ca7:0x0000000000000000: s. f = /dev/sdf].
srgd[xxxx]: PV status change fail . [hwp path = pci-0000:13:00.1-fc-0x2900001697120ca7:0x0000000000000000: s. f = /dev/sdd].
```

- (4) 続いて、もう片方の FC ケーブルを抜線して両系障害を発生させます。

(5) 約 60 秒後に障害を検出し、StorageSaver のステータスが down になることを確認してください。

```
# /opt/HA/SrG/bin/srgadmin - c status
(monitor status = TRUE)
```

type	device	HostBusAdapter	L status	P status	Online status
VG	VolGroup01	—	down		
PV	/dev/sdj	pci-0000:13:00.0	down	down	extended
PV	/dev/sdh	pci-0000:13:00.0	down	down	extended
PV	/dev/sdf	pci-0000:13:00.1	down	down	extended
PV	/dev/sdd	pci-0000:13:00.1	down	down	extended
VG	VolGroup02	—	down		
PV	/dev/sdk	pci-0000:13:00.0	down	down	extended
PV	/dev/sdi	pci-0000:13:00.0	down	down	extended
PV	/dev/sgd	pci-0000:13:00.1	down	down	extended
PV	/dev/sde	pci-0000:13:00.1	down	down	extended

(6) 両系障害を発生させた場合、syslog にメッセージが出力されます。

```
srgd[xxxxx]: VG status change down . (vg=VolGroup01)
srgd[xxxxx]: VG status change down . (vg=VolGroup02)
srgstat[xxxxx]: found VG status is down
srgd[xxxxx]: PV status change fail . [hwpath = pci-0000:13:00.0-fc-0x2100001697120ca7:0x0001000000000000: s. f =
/dev/sdk].
srgd[xxxxx]: PV status change fail . [hwpath = pci-0000:13:00.0-fc-0x2900001697120ca7:0x0001000000000000: s. f =
/dev/sdi].
srgd[xxxxx]: PV status change fail . [hwpath = pci-0000:13:00.0-fc-0x2100001697120ca7:0x0000000000000000: s. f =
/dev/sdj].
srgd[xxxxx]: PV status change fail . [hwpath = pci-0000:13:00.0-fc-0x2900001697120ca7:0x0000000000000000: s. f =
/dev/sdh].
```

両系のディスクのステータスが down になったため、VG 状態も down となります。

(7) FC ケーブルの抜線を行ったサーバが shutdown することを確認してください。

※ 本書の CLUSTERPRO の設定ではサーバが shutdown するため、(5) (6) は確認できない場合があります。

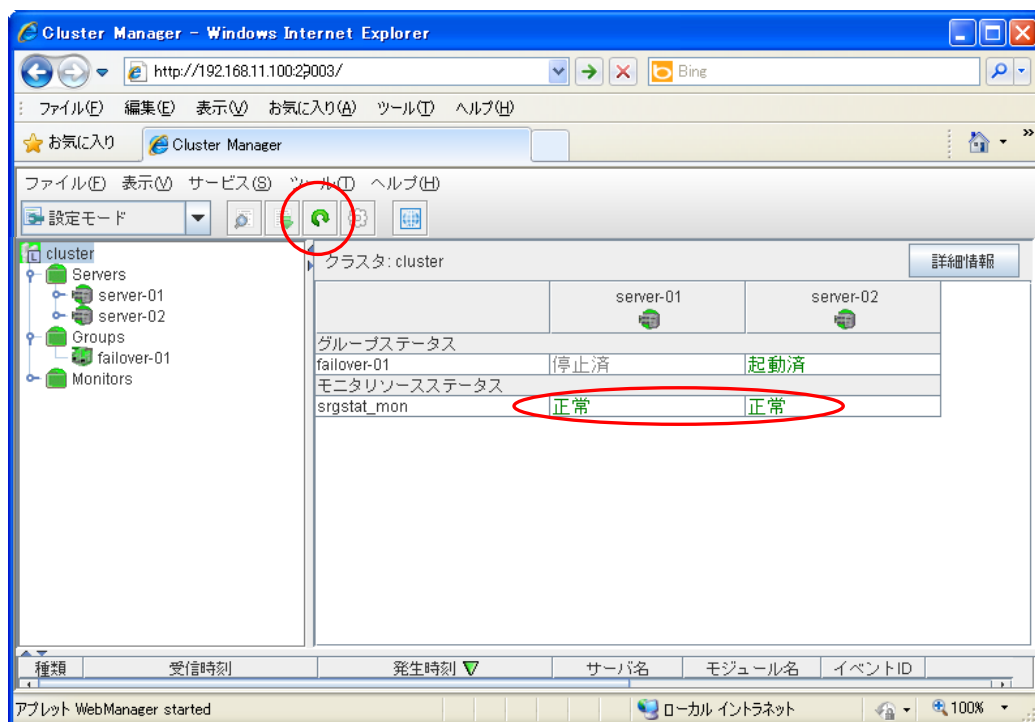
2. 待機系サーバへのフェイルオーバー確認

- (1) 手順 1.で拔線した FC ケーブルを接続後、shutdown したサーバを起動させてください。
- (2) syslog 以下のようなメッセージが出力されていることを確認してください。

```
clusterpro: [I] <type: rm><event: 16> Stopping the system has been required because an error  
was detected in monitoring srgstat_mon.  
clusterpro: [I] <type: pm><event: 30> Received a request to stop the system from internal (rc).
```

- (3) CLUSTERPRO WebManager の「リロード」をクリックし、以下の項目を確認してください。

・srgstat 監視用のカスタムモニタリソース「srgstat_mon」のステータスが現用系、待機系にて「正常」である事を確認してください。



以上で、動作確認は終了となります。

CLUSTERPRO
MC StorageSaver 1.1 for Linux
ユーザーズガイド

2013 年 9 月 第 5 版
日本電気株式会社
東京都港区芝五丁目7番地1号
TEL (03) 3454-1111 (代表)



© NEC Corporation 2013

日本電気株式会社の許可なく複製、改変などを行うことはできません。
本書の内容に関しては将来予告なしに変更することがあります。

保護用紙