

CLUSTERPRO

MC RootDiskMonitor 2.9 for Linux

CLUSTERPRO

MC StorageSaver for BootDisk 2.9 (for Linux)

ユーザーズガイド

© 2024(Apr) NEC Corporation

- 製品の概要
- OS ディスクの監視方式について
- 設定ファイル
- 操作・運用手順
- 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 RootDiskMonitor 2.9 for Linux (以後 RootDiskMonitor と記載します)、および CLUSTERPRO MC StorageSaver for BootDisk 2.9 (for Linux) の OS ディスク監視に関する設定について記載したものです。

(注) StorageSaver for BootDisk は、以後 RootDiskMonitor と表記します。

なお、StorageSaver for BootDisk では、以下の機能は利用できません。

- パトロールシーク機能
- HW-RAID 監視機能

基本機能として以下の運用が可能です。

- I/O パス監視機能
- 障害レポート機能
- クラスタウェア連携機能

拡張機能として以下の運用が可能です。

- ミラー構成監視機能 (本機能については次期バージョン以降でサポート予定)

(1) 商標および登録商標

- ✓ Red Hat、Red Hat Enterprise Linux は、米国およびその他の国における Red Hat, Inc. およびその子会社の商標または登録商標です。
- ✓ Oracle は、Oracle Corporation およびその子会社、関連会社の米国およびその他の国における登録商標です。文中の社名、商品名等は各社の商標または登録商標である場合があります。
- ✓ Linux は、米国およびその他の国における Linus Torvalds の登録商標です。
- ✓ Dell, EMC, および Dell, EMC が提供する製品およびサービスにかかる商標は、米国 Dell Inc. またはその関連会社の商標または登録商標です。
- ✓ その他記載の製品名および会社名は、すべて各社の商標または登録商標です。
- ✓ なお、本書では®、TM マークを明記しておりません。

(2) 本リリースの強化点について

RootDiskMonitor 2.9 (2024 年 4 月出荷版)では、下記の機能を強化しています。

- ・ サポート OS の記載を更新しました。
RootDiskMonitor がサポートする動作環境の詳細については、別紙『CLUSTERPRO MC RootDiskMonitor for Linux リリースメモ』の「動作環境」の章を参照してください。
- ・ OS によるディスクデバイスファイルの再割り当てに対応し、再割り当て後もディスクの監視を継続することが可能となりました。
機能概要については "2.6. 監視ディスクのデバイスファイル再解決機能について"、
運用手順については "4.8. 監視ディスクのデバイスファイル再解決手順"を
参照してください。

強化した内容は、次の 2 点です。

- ・ デバイス情報の変更をチェックする機能を追加しました。
デバイス情報の変更が発生したか1日に1度の定期監視または運用管理コマンドにてチェックすることができるようになりました。
- ・ デバイス情報の更新を行う機能を追加しました。
デバイス情報が変更されていた場合、構成復旧コマンドまたは運用管理コマンドにて RootDiskMonitor 内部のデバイス情報を更新し、監視を継続することができるようになりました。
- ・ 設定ファイル自動生成機能を強化しました。
Device Mapper を使用したソフトミラー構成をサポートしました。
ただし、Device Mapper を使用したソフトミラー構成でも、LVM を使用した構成はサポート対象外です。
- ・ 製品の 64bit 化を行いました。
- ・ 障害解析情報収集機能を追加しました。
障害発生時に手動で採取していた障害解析情報をツールを利用することで自動的に採取できるようになりました。

(3) これまでの強化点について

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

- ・ 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ーとして、設定ファイルに新規パラメーター (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ー TESTIO_DIRECT は本パラメーターに統合されました。

また、これに伴い TestUnitReady の発行抑止についても、本パラメーターで設定するよう変更しました。

パラメーターの詳細については "3.2 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、クラスターウェア連携用デーモン (rdmstat)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ー

(OVER_ACTION および VG_STALL_ACTION) に指定する値も変更しました

パラメーターの詳細については "3.2 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。

RootDiskMonitor 2.0 (2015 年 4 月出荷版)では、下記の機能を強化しています。

- ・ Red Hat Enterprise Linux 7.0 および Oracle Linux 7.0 をサポートしました。

Red Hat Enterprise Linux 7.0 および Oracle Linux 7.0 環境における監視機能をサポートしました。

Red Hat Enterprise Linux 7.0 および Oracle Linux 7.0 環境ではプロセスの起動、停止方法を変更しています。

詳細は "4.1 運用管理コマンドの操作手順" を参照してください。

- ・ Hyper-V のゲスト OS における監視機能をサポートしました。

Hyper-V のゲスト OS における OS ディスクの監視機能をサポートしました。

設定手順については「CLUSTERPRO MC RootDiskMonitor 仮想環境(ゲスト OS)での設定手順」を参照してください。

- ・ セクターサイズが 4096 バイトのディスク装置をサポートしました。
 セクターサイズが 4096 バイトのディスク装置に対する監視をサポートしました。
- ・ 設定ファイル(rdm.config)に定義するデバイスファイルの形式を変更しました。
 設定ファイル(rdm.config)に定義するデバイスファイルの形式を
 ブロックスペシャルファイル名から udev デバイスファイル名に変更しました。
 これにより構成変更等でブロックスペシャルファイル名が変更された場合でも
 設定ファイルの再作成が不要になります。
 なお、従来どおりブロックスペシャルファイル名も設定可能です。
- ・ デバイスファイルのチェック機能をサポートしました。
 設定ファイル(rdm.config)にスペシャルファイル名と udev デバイスファイル名の両方を
 指定した場合、デーモンプロセス起動時にスペシャルファイル名の妥当性チェックを
 行います。
 スペシャルファイル名が変更されていた場合には、本機能により当該デバイスを
 監視対象外とします。
 これにより、誤った設定のまま監視を開始して障害を誤検知することを防止します。
 なお、デバイスファイルのチェック機能は、設定ファイル自動生成(rdmconfig) の
 サポート対象外です。本機能を使用する場合は、手動で設定ファイルを編集してください。

RootDiskMonitor 2.1 (2016 年 4 月出荷版)では、下記の機能を強化しています。

- ・ KVM(Kernel-based Virtual Machine)のゲスト OS における監視機能をサポートしました。
 KVM のゲスト OS における OS ディスクの監視機能をサポートしました。
 設定手順については「CLUSTERPRO MC RootDiskMonitor 仮想環境(ゲスト OS)での
 設定手順」を参照してください。
- ・ Multiple Devices で構築したソフトウェア RAID 構成をサポートしました。
 Multiple Devices で構築したソフトウェア RAID 構成の OS ディスクの監視機能を
 サポートしました。設定手順については"8.4 Multiple Devices 構成における設定ファイル作
 成手順"を参照してください。

RootDiskMonitor 2.2 (2017 年 4 月出荷版)では、下記の機能を強化しています。

- ・ S.M.A.R.T.診断結果を表示する機能を追加しました。
本機能により、リソース監視の状態表示で監視対象ディスクが故障間近であるか確認できるようになります。
詳細は、"4.4 S.M.A.R.T.診断結果表示手順" を参照してください。

RootDiskMonitor 2.3 (2018 年 4 月出荷版)では、下記の機能を強化しています。

- ・ VMware vMotion に対応しました。(vSphere6.0 update3、vSphere6.5 update1)
以下 vMotion が実施されたときでも、継続して RootDiskMonitor での監視が行えるよう対応しました。なお、特に設定等おこなう必要はありません。
 1. コンピューティング vMotion
 2. ストレージ vMotion
 3. クロスホスト vMotion

RootDiskMonitor 2.4 (2019 年 4 月出荷版)では、下記の機能を強化しています。

- ・ コードワードが未登録でもインストールできるようになりました。
従来までは、インストール時にライセンスをチェックするため、事前にコードワードの登録が必要でしたが、コードワード未登録でもインストール可能となりました。

RootDiskMonitor 2.5 (2020 年 4 月出荷版)では、下記の機能を強化しています。

- ・ HW-RAID 監視機能において、次の構成をサポートしました。
 - ① HPE Smart Storage Administrator を使用した構成
 - ② HW-RAID を構成する1つの物理ディスク群から複数の論理ディスクを切り出した構成

RootDiskMonitor 2.6 (2021 年 4 月出荷版)では、下記の機能を強化しています。

- ・ Dell EMC 社製ストレージ装置をサポートしました。
Dell EMC 社製 Unity XT シリーズに対応しました。
- ・ ESXi 7.0 をサポートしました。
ESXi 7.0 上のゲスト OS における監視をサポートしました。
- ・ SAN-Boot 構成における I/O パスの閉塞・復旧機能をサポートしました。
SAN-Boot 構成において、障害を検知した I/O パスの自動閉塞が可能となりました。
また、手動にて I/O パスの閉塞・復旧が可能なオンライン保守機能をサポートしました。
- ・ Veritas InfoScale Storage 環境をサポートしました。
Veritas InfoScale Storage 環境での、Veritas Dynamic Multi-Pathing を使用した構成に対する監視をサポートしました。

RootDiskMonitor 2.7 (2022 年 4 月出荷版)では、下記の機能を強化しています。

- ・ 従来とくらべ障害の早期検知が可能になりました。
強化した内容は、次の 2 点です。
 - ・監視間隔をより短くすることで、障害の早期検知が可能になりました。
詳細については“4.7. 障害の検知、および、アクション実行の高速化の設定手順”を参照してください。
 - ・致命的なエラーを考慮した障害検出が可能になりました。
Test I/O (Inquiry)で致命的なエラーが発生した場合、障害として即時に通知することができるようになりました。
詳細については “2.5. 致命的なエラー発生時の障害検出の高速化について”ならびに “4.7. 障害の検知、および、アクション実行の高速化の設定手順”を参照してください。
- ・ Dell EMC 社製ストレージ装置をサポートしました。
Dell EMC 社製 PowerStore シリーズに対応しました。

RootDiskMonitor 2.8 (2023 年 4 月出荷版)では、下記の機能を強化しています。

- ・SELinux に対応しました

Linux システム用のセキュリティ・アーキテクチャである SELinux (Security-Enhanced Linux) を有効化した環境で本製品を利用可能になりました。

- ・NVMe SSD ブートディスク搭載 Express5800 シリーズをサポートしました。

NVMe SSD ブートディスク搭載 Express5800 シリーズに対応しました。

RootDiskMonitor がサポートする動作環境の詳細については、別紙『CLUSTERPRO MC RootDiskMonitor for Linux リリースメモ』の「動作環境」の章を参照してください。

- ・設定ファイル自動生成機能を強化しました。

マルチパス管理ソフトウェアのデバイス名を任意に変更した構成 (例.mpatha⇒root-disk01)をサポートしました。

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.....	2
2. OSディスクの監視方式について.....	4
2.1. OSディスクの監視.....	4
2.2. I/Oパスの監視手順について.....	9
2.3. I/Oパスの異常を検出すると.....	13
2.4. クラスタウェアとの連携について.....	15
2.5. 致命的なエラー発生時の障害検出の高速化について.....	17
2.6. 監視ディスクのデバイスファイル再解決機能について.....	21
2.6.1. デバイス情報のチェック機能.....	22
2.6.2. デバイス情報の更新機能.....	23
3. 設定ファイル.....	24
3.1. 本製品の導入.....	24
3.2. 設定ファイルの記述.....	28
4. 操作・運用手順.....	34
4.1. 運用管理コマンドの操作手順.....	34
4.2. CLUSTERPRO Xとの連携.....	41
4.2.1. カスタムモニタソースによるCLUSTERPRO Xとの連携.....	41
4.2.2. サーバー管理プロセス (clpnm) の強制終了によるCLUSTERPRO Xとの連携.....	44
4.2.3. システムメモリダンプ採取とOS強制停止によるCLUSTERPRO Xとの連携.....	46
4.3. HW-RAID状態表示手順.....	47
4.4. S.M.A.R.T.診断結果表示手順.....	51
4.5. 機能制限について.....	59
4.6. オンライン保守コマンドの操作手順.....	60
4.7. 障害の検知、および、アクション実行の高速化.....	63
4.7.1. 監視間隔のチューニングについて.....	63
4.7.2. 障害条件の変更について.....	64
4.7.3. 障害の検知、および、アクション実行の高速化の設定手順.....	65
4.8. 監視ディスクのデバイスファイル再解決手順.....	74
4.8.1. デバイスファイルのチェックについて.....	74
4.8.2. デバイスファイルの更新について.....	76
5. syslogメッセージ.....	77
5.1. syslog に出力するメッセージについて.....	77
5.2. 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.....	78
5.3. 運用管理製品との連携.....	81
6. 注意・制限事項について.....	82
6.1. 注意・制限事項.....	82
7. リファレンス.....	85

8. 付録.....	93
8.1. 本製品のテスト手順について.....	93
8.2. カスタムモニタリソースによる CLUSTERPRO X 4.0 以前との連携手順.....	112
8.2.1. CLUSTERPRO X連携設定.....	112
8.2.2. 動作確認.....	119
8.3. カスタムモニタリソースによる CLUSTERPRO X 4.1 以降との連携手順.....	120
8.3.1. CLUSTERPRO X連携設定.....	120
8.3.2. 動作確認.....	130
8.4. Multiple Devices構成における設定ファイル作成手順.....	131
8.5. S.M.A.R.T.診断結果の定期監視手順.....	135
8.5.1. 定期監視の利用手順.....	135
8.6. 本製品の障害解析情報.....	138

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 OS ディスクを構成する I/O パスの動作状態を定期監視します。I/O パスに異常が見られるとエラーレポートを通知し、さらに OS ディスクが動作不能になるとクラスターウェアと連携しノードを切り替えることでクラスターシステムでの可用性を向上させます。

- ・ I/O パス監視機能
OS ディスクを構成する I/O パスに対して死活監視、I/O リクエストのストール監視を行います。シングル構成、ミラー構成の OS ディスクを監視できます。
- ・ 障害レポート機能
I/O パスを定期監視し異常を検出すると syslog、コンソールに異常レポートを通知します。
- ・ I/O パスの自動閉塞機能 ※SAN-Boot 構成のみ
SAN-Boot 構成において、マルチパス管理ソフトウェアで冗長化された I/O パスを有する OS ディスクにおいて、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。
- ・ クラスターウェア連携機能
OS ディスクを構成する I/O パスがすべて障害となり、LUN(論理ディスク装置)へのアクセスが不可能になると、クラスターウェアと連携することによりノード切り替えを実現します。

クラスターウェアと連携しノード切り替えを実現するには以下の手法があります。

- (1) CLUSTERPRO X のサーバー管理プロセス (clpnm) を強制終了させる方式
- (2) OS ディスク監視専用のクラスターリソースを作成する方式
- (3) Linux のシステムメモリダンプを採取し、OS を強制的に停止させる方式

クラスターウェアを利用しない非クラスターシステムでは、ノード切り替え機能はご利用いただけません。

- ・ オンライン保守機能 ※SAN-Boot 構成のみ
SAN-Boot 構成において、FC 接続の OS ディスクに対して、マルチパス管理ソフトウェアと連携して I/O パスの閉塞・復旧を実現します。

1.2. 製品の構成について

(1) プロダクト構成

本製品は Linux の OS ディスク監視を行います。

(2) ソフトウェア構成

プロセス構成は以下のとおりです。

- ・ rdmdiagd OS ディスク監視エンジン
- ・ rdmping OS ディスク監視モニター
- ・ rdmsreduced 自動閉塞デーモン ※SAN-Boot 構成のみ
- ・ rdmsadmin 運用管理コマンド
- ・ rdmsconfig 設定ファイル自動生成コマンド
- ・ rdmsstat クラスターウェア連携用デーモン

ファイル、ディレクトリ構成は以下のとおりです。

- ・ 実行形式ディレクトリ /opt/HA/RDM/bin
- ・ 設定ファイル /opt/HA/RDM/conf/rdm.config
- ・ rc ファイル格納ディレクトリ

【Red Hat Enterprise Linux 6.x の場合】

【Oracle Linux 6.x の場合】

/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

- ・ Unit 定義ファイル格納ディレクトリ

【Red Hat Enterprise Linux 7.0 以降の場合】

【Oracle Linux 7.0 以降の場合】

【Amazon Linux 2 以降の場合】

/usr/lib/systemd/system

(3) サポート範囲

Linux OS ディスクが対象となります。

SCSI インタフェース接続の OS ディスク装置、増設ディスク装置

NVMe インタフェース接続の SSD 内蔵ディスク装置 ※1

※1 下記の構成を満たしている必要があります。

- ・NEC 社製 NVMe SSD ブートディスク搭載 Express5800 シリーズであること。
- ・Universal RAID Utility にて HW-RAID を構成していること。

FC 接続のディスクアレイ装置

- ・ NEC 社製 iStorage 全シリーズ(ただし、E1 シリーズ、HS シリーズは除きます)
- ・ Dell EMC 社製 CLARiX シリーズ
VNX シリーズ
Symmetrix DMX シリーズ
Symmetrix VMAX シリーズ、VMAX3 シリーズ、PowerMax ファミリー
XtremIO、XtremIO X2、VPLEX、Unity シリーズ、Unity XT シリーズ
PowerStore シリーズ
- ・ 日立 社製 SANRISE シリーズ
Hitachi USP シリーズ
Hitachi VSP シリーズ
- ・ HPE 社製 3PAR シリーズ

下記のボリューム管理製品を対象とします。

- LVM 上に構築された OS ディスク(/dev/VolGroup00,etc)
- 物理ディスク上に直接構築された OS ディスク(/dev/sdx)

注意:StorageSaver for BootDisk の場合は、サーバー搭載の内蔵ディスクの監視には
使用できません。

SAN-Boot 構成の場合は、下記のマルチパス管理ソフトウェアを対象とします。

- Device Mapper Multipath
- StoragePathSavior
- HA Dynamic Link Manager
- Dell EMC 社製 PowerPath
- 日立社製 Hitachi Dynamic Link Manager
- Veritas 社製 Veritas Dynamic Multi-Pathing

2. OS ディスクの監視方式について

2.1. OS ディスクの監視

(1) 監視のフレームワークについて

本製品では、OS ディスクの障害を検出するために、I/O パスに対して定期的に Test I/O を行います。Test I/O で監視対象となる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

Test I/O は SCSI パススレッドドライバー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

(2) 監視対象となる I/O パスについて

OS ディスクを構成する I/O パスが監視対象となります。
なお、コンフィグレーションファイルに OS ディスクを構成する I/O パスのスペシャルファイルと監視ルールの設定が必要です。

(3) I/O パスの異常を検出すると

Test I/O で異常を検出した I/O パスは、syslog、コンソールに障害レポートを通知します。
さらに、LUN(論理ディスク装置)へのアクセスが不可能になると、コンフィグレーションファイルで指定されたアクションを実行します。

また、SAN-Boot 構成の場合は、Test I/O で異常を検出した I/O パスを自動で閉塞させることも可能です。

(4) I/Oパスの状態について

LUNレベルのステータス(I/Oパスの監視状態)として、以下の状態をレポートします。

- UP
I/Oパスが正常に動作している状態です。
- DOWN
I/Oパスに異常があり、利用不可な状態です。

SAN-Boot 構成の場合は、追加で I/Oパスの組み込み状態をレポートします。

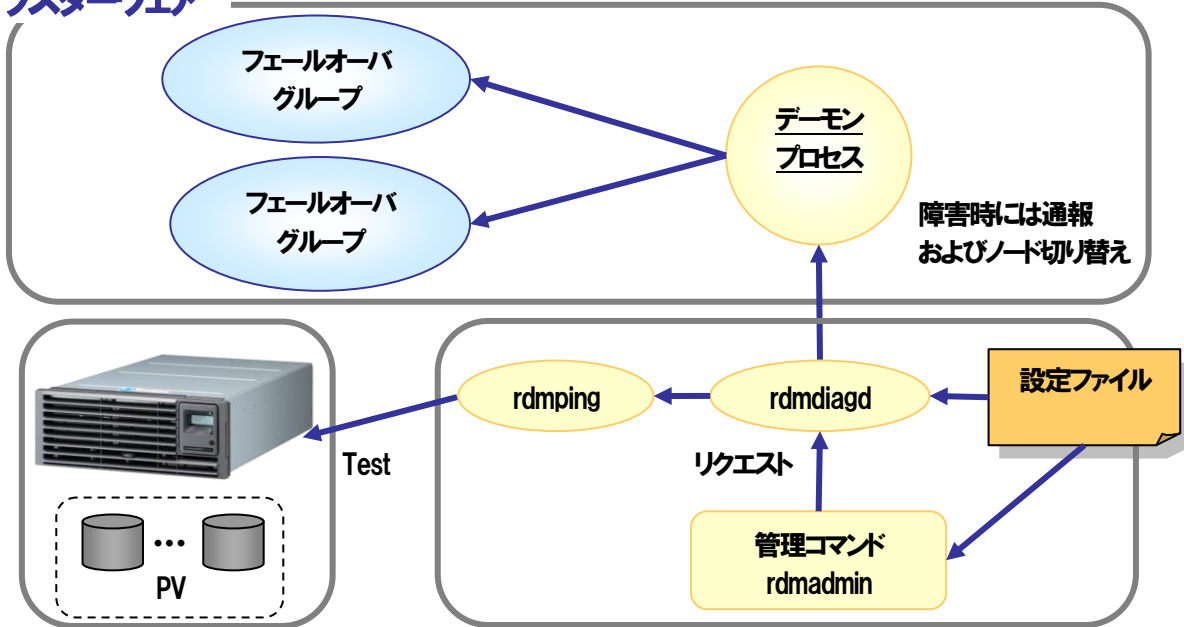
- extended
I/Oパスが正常に組み込まれている状態です。
- reduced
I/Oパスが閉塞されている状態です。
- unknown
I/Oパスの組み込み状態が不明な状態です。

- ※ I/Oパス障害が発生した際、障害の種類によってはOSが障害の発生したI/Oパスのsdデバイスを削除する場合があります。
物理環境のSANBoot構成で、TESTIO_FAULT_ACTIONにBLOCK_PATHを設定している場合、I/Oパスの障害を検出後、自動閉塞処理を実行する前に、OSによる対象のsdデバイスの削除が行われた場合は、I/Oパスの閉塞が行われず、Online statusがextendedのままとなる場合があります。
上記の場合でも、OS側で障害の発生したI/Oパスのsdデバイスの削除が行われているため、障害が発生したパスでI/Oが行われることはなく、運用上は問題ありません。

VGレベルのステータス(I/Oパスをミラーグループとして束ねた監視状態)として、以下の状態をレポートします。

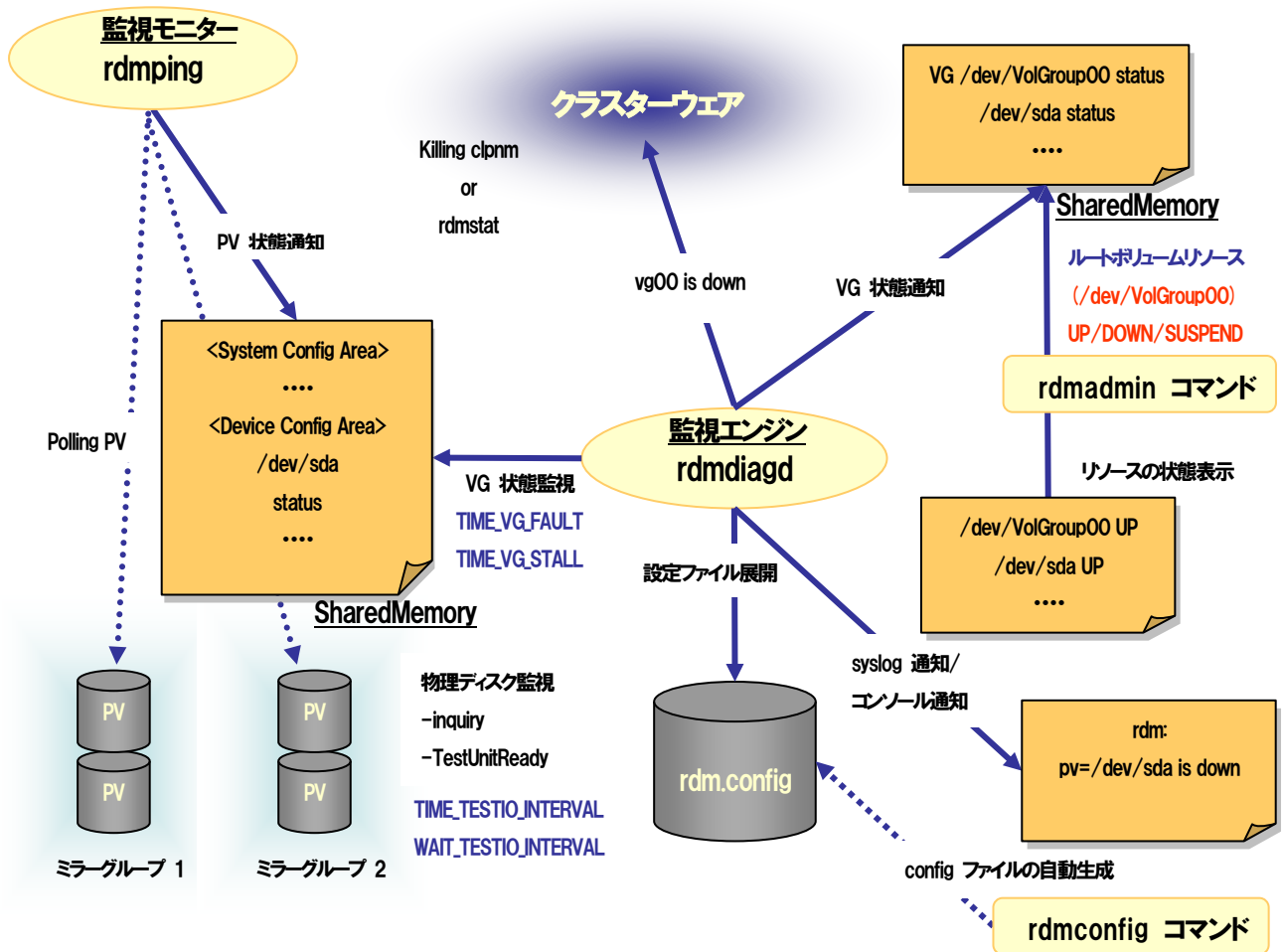
- UP
VGが正常に動作している状態です。
- SUSPEND
VGを構成するミラー片系のI/Oパスに異常がある状態です。
- DOWN
VGに異常があり、利用不可な状態です。

クラスターウェア



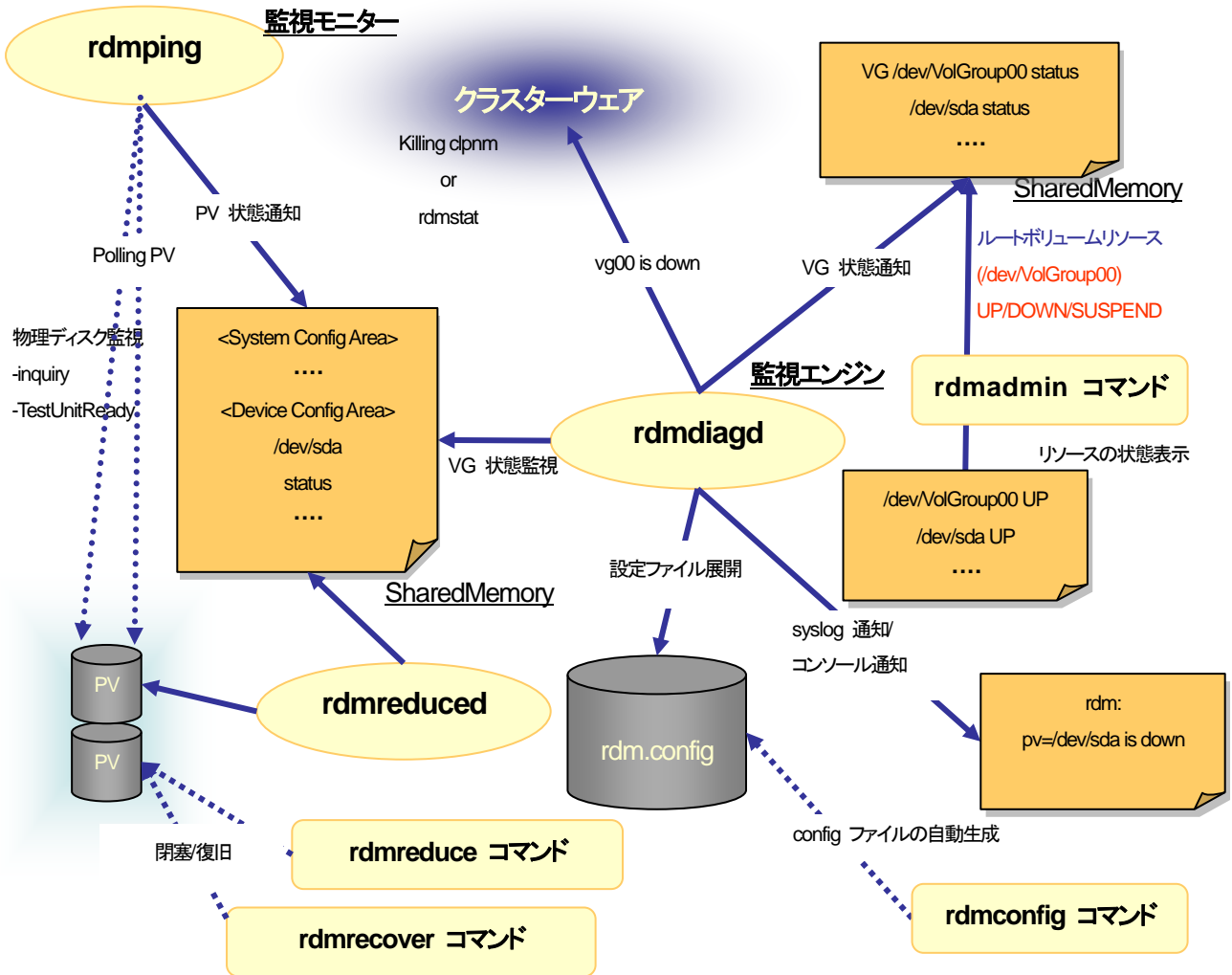
【RootDiskMonitor の構成】

内蔵ディスク構成での RootDiskMonitor のプロセス構成は次のとおりです。



【RootDiskMonitor のプロセスモデル (内蔵ディスク構成)】

SAN-Boot 構成での RootDiskMonitor のプロセス構成は次のとおりです。



【RootDiskMonitor のプロセスモデル (SAN-Boot構成)】

2.2. I/O パスの監視手順について

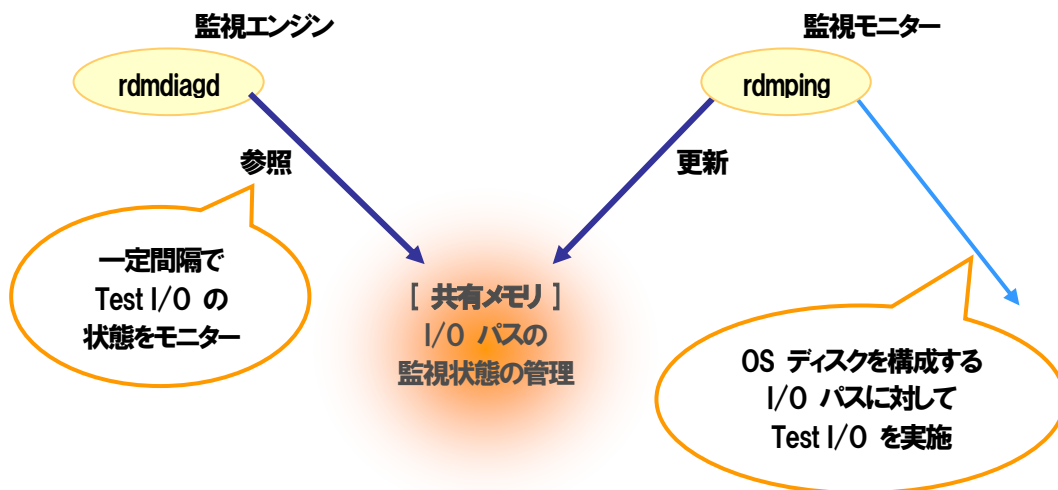
(1) I/O パスの死活監視

OS ディスクを構成する I/O パスに対し定期的に SCSI パススルー機能を利用して Test I/O を発行することで、I/O パスの動作状態を監視します。

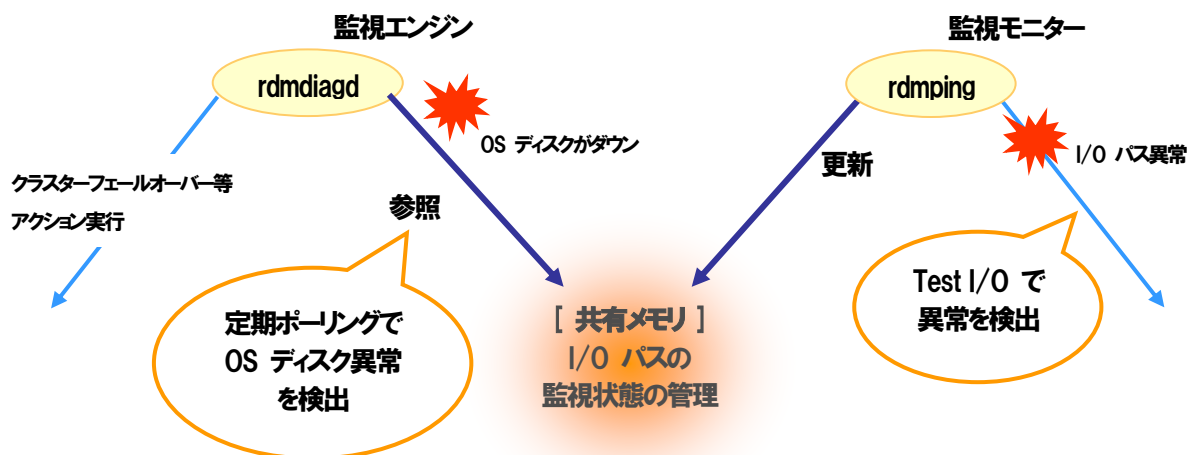
Test I/O が正常終了しない、またはタイムアウトした場合は I/O パスを異常と判定します。

注意 RootDiskMonitor を動作させる場合、Linux パススルードライバー(sg ドライバー)がインストールされ、事前にカーネルにロードされている必要があります。

【Test I/O のフレームワーク】

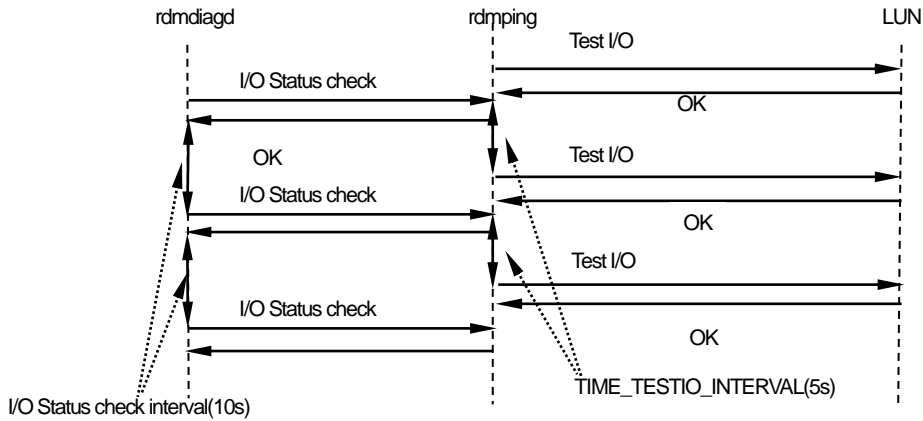


【Test I/O で OS ディスクの異常を検出】



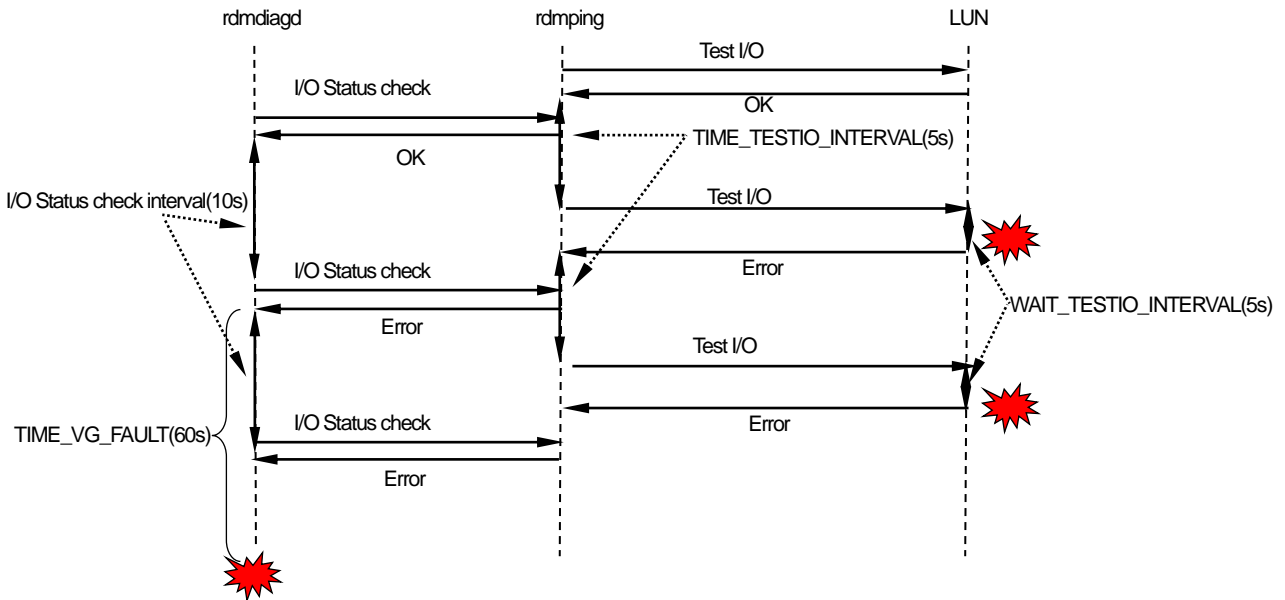
(2) Test I/O の正常なシーケンスは、以下のような動作になります。

基本タイマー(I/O Status Check Interval) :10(秒)
 TIME_TESTIO_INTERVAL :5(秒)



(3) Test I/O で異常を検出すると、以下のような動作になります。

基本タイマー(I/O Status Check Interval) :10(秒)
 TIME_VG_FAULT :60(秒)
 TIME_TESTIO_INTERVAL :5(秒)
 WAIT_TESTIO_INTERVAL :5(秒)

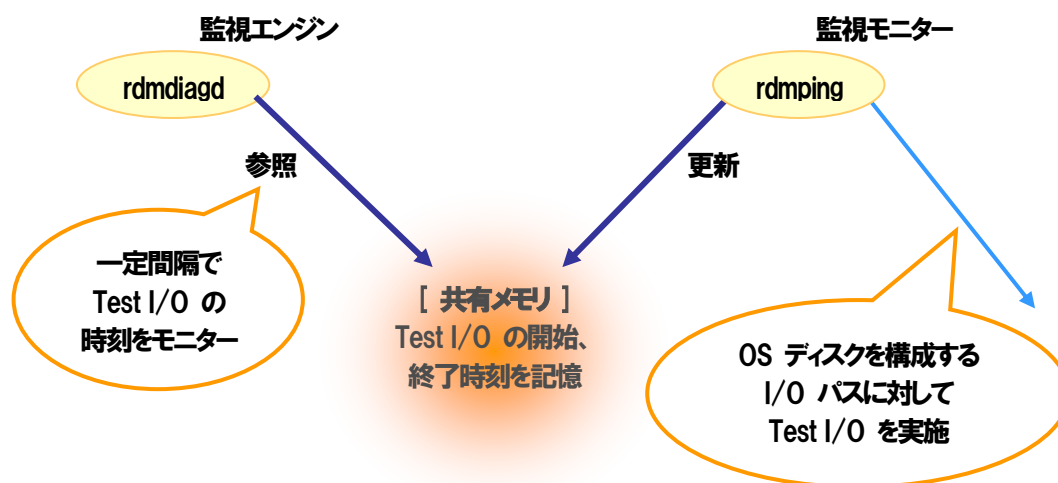


タイムオーバー
 syslog 通知 or クラスターウェアデーモン強制停止 or rdmsat 停止 or os system dump + panic

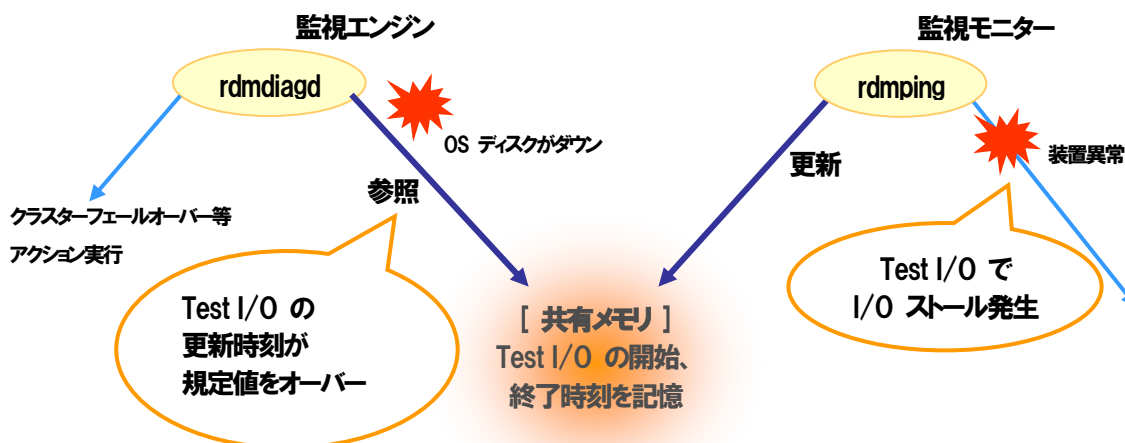
(4) I/Oパスのストール監視

OS ディスクを構成する I/O パスに対し定期的に SCSI パススルー機能を利用して Test I/O を発行することで、OS 全体のストール状態を監視します。
Test I/O が一定時間以内に正常完了しない場合は 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】

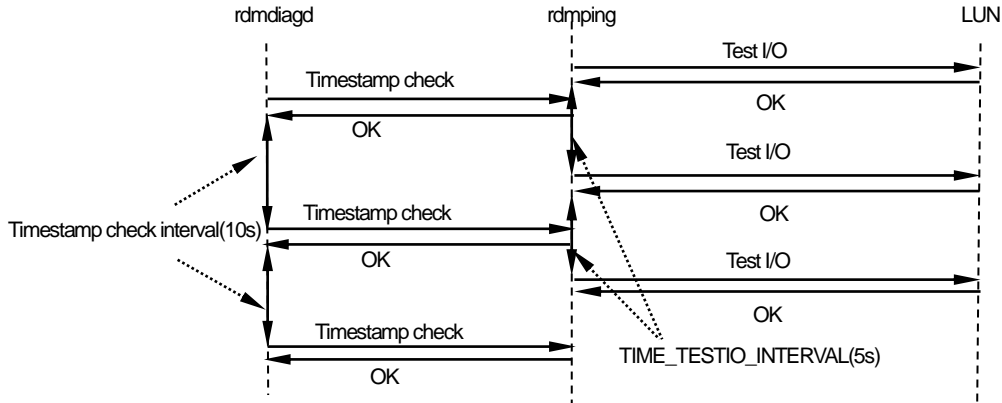


【I/O ストールを検出すると】



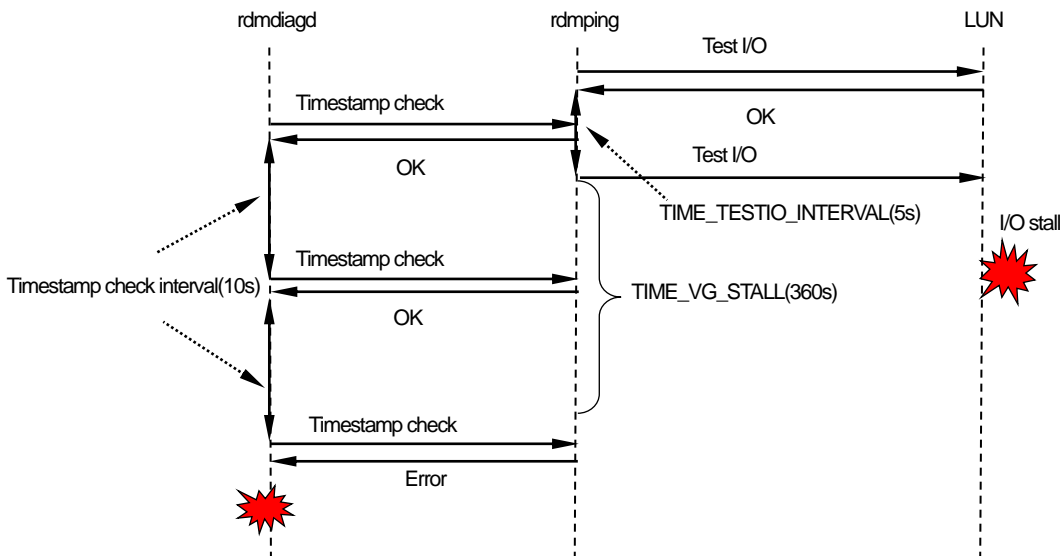
(5) Test I/O の正常なシーケンスは、以下のような動作になります。

基本タイマー(Timestamp Check Interval) :10(秒)
 TIME_VG_STALL :360(秒)
 TIME_TESTIO_INTERVAL :5(秒)
 WAIT_TESTIO_INTERVAL :5(秒)



(6) 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。

基本タイマー(Timestamp Check Interval) :10(秒)
 TIME_VG_STALL :360(秒)
 TIME_TESTIO_INTERVAL :5(秒)
 WAIT_TESTIO_INTERVAL :5(秒)



タイムオーバー
 syslog 通知 or クラスターウェアデーモン強制停止 or rdmstat 停止 or os system dump + panic

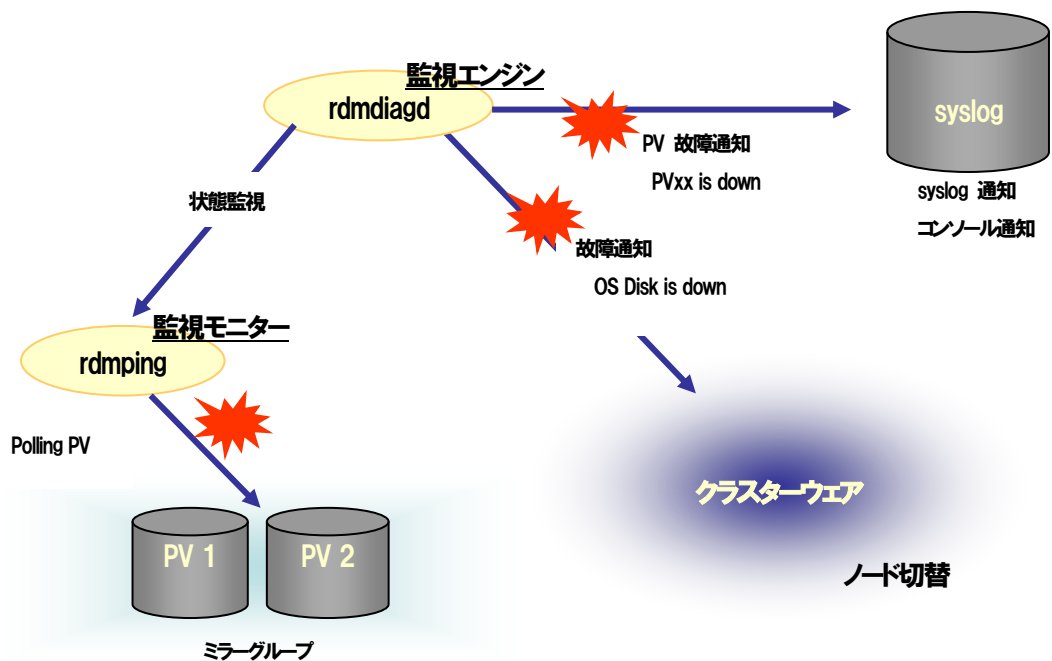
2.3. I/O パスの異常を検出すると

(1) I/O パスの異常を検出すると

I/O パスの異常を検出すると、syslog にエラーメッセージを出力します。
当該 I/O パスの監視は継続しますので I/O パスが復旧次第、
正常状態として監視を続けます。

SAN-Boot 構成の場合は、異常を検出した I/O パスを自動で閉塞させることが可能です。
自動閉塞を行った場合は、当該 I/O パスは監視対象から一時的に外れます。
I/O パスが復旧した場合は、構成復旧コマンド(rdmrecover)にて、復旧させるまで
監視が行われません。

【Test I/O (Polling) 方式によるディスク監視】



(2) メディアエラーで異常を検出すると ※本機能は次期バージョン以降でサポート予定です。

syslog の監視を行い、ログ内に監視キーワードを検出すると、
メディアエラーと判断し、障害ディスクの切り離しを行います。

(3) 両系障害レベルの異常を検出すると

OS ディスクを構成する I/O パスで異常が発生し、LUN へのアクセスが不可能になると、syslog、コンソールにエラーメッセージを出力します。

- シングルパス構成での I/O パス異常
- ソフトミラー構成での正副 LUN 異常

両系障害レベルで異常となると、設定ファイルで規定されたアクションを実行します。

- 警報通知のみ
- クラスターウェア連動によるノード切り替え
- システムメモリダンプを採取し、強制的に OS を停止しノード切り替え
- ユーザー定義コマンドの実行(本リリースでは未サポート)

2.4. クラスタウェアとの連携について

OS ディスクの障害で動作不能な状態に陥った場合にクラスタウェアと連携することで待機ノードへ切り替え業務を継続することができます。

本書ではクラスタウェア製品として、CLUSTERPRO を例にクラスタウェア連携について説明します。(以降の章でも特に断りがなければ、クラスタウェア連携については CLUSTERPRO を対象とします)

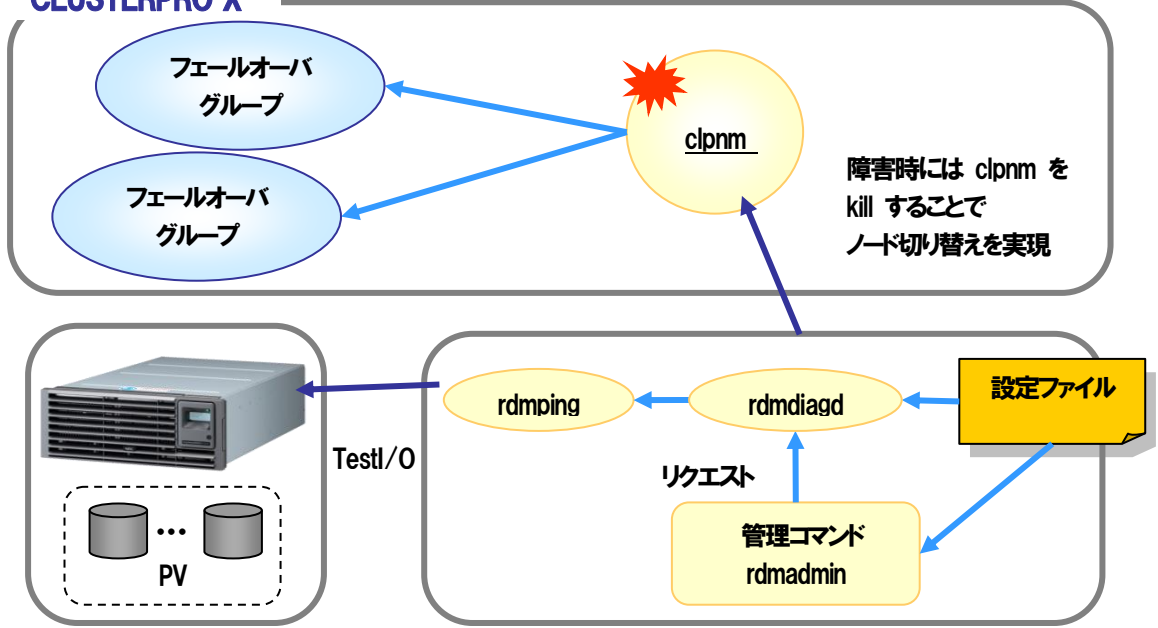
● CLUSTERPRO X との連携

RootDiskMonitor が CLUSTERPRO X と連携するには、以下の 3 つの方式があります。

1. CLUSTERPRO X 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(rdmstat)を登録する方式
2. CLUSTERPRO X のサーバー管理プロセス (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

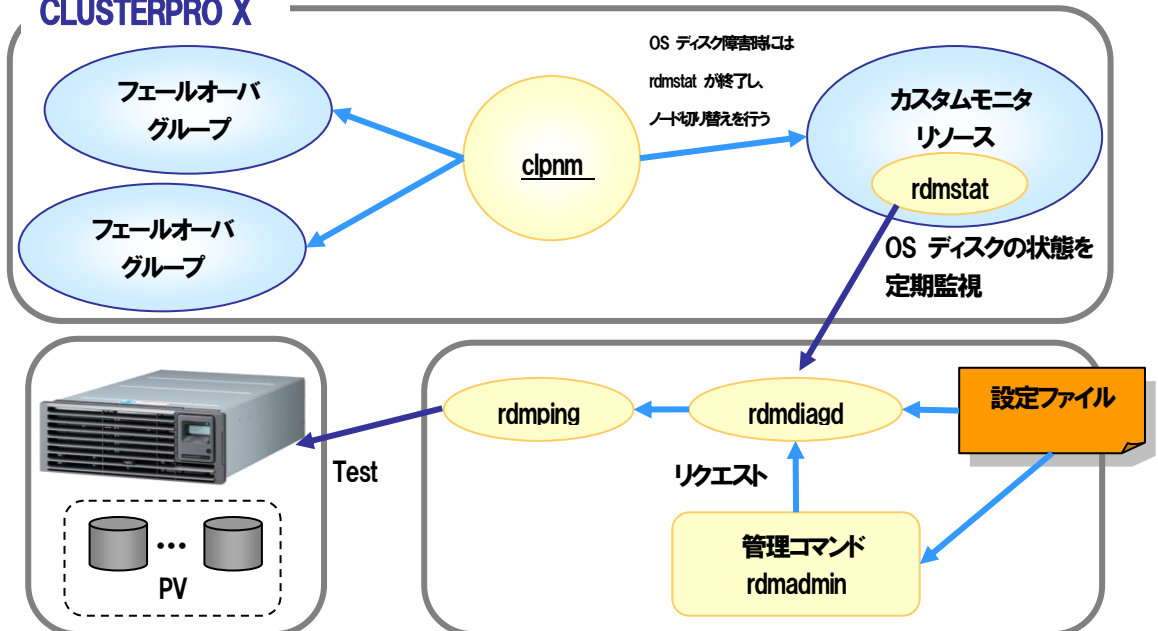
RootDiskMonitor としては 1 の方式を推奨します。

CLUSTERPRO X



【 clpnm を強制終了する手法】

CLUSTERPRO X



【カスタムモニタリソースによるフェールオーバーグループ連動】

2.5. 致命的なエラー発生時の障害検出の高速化について

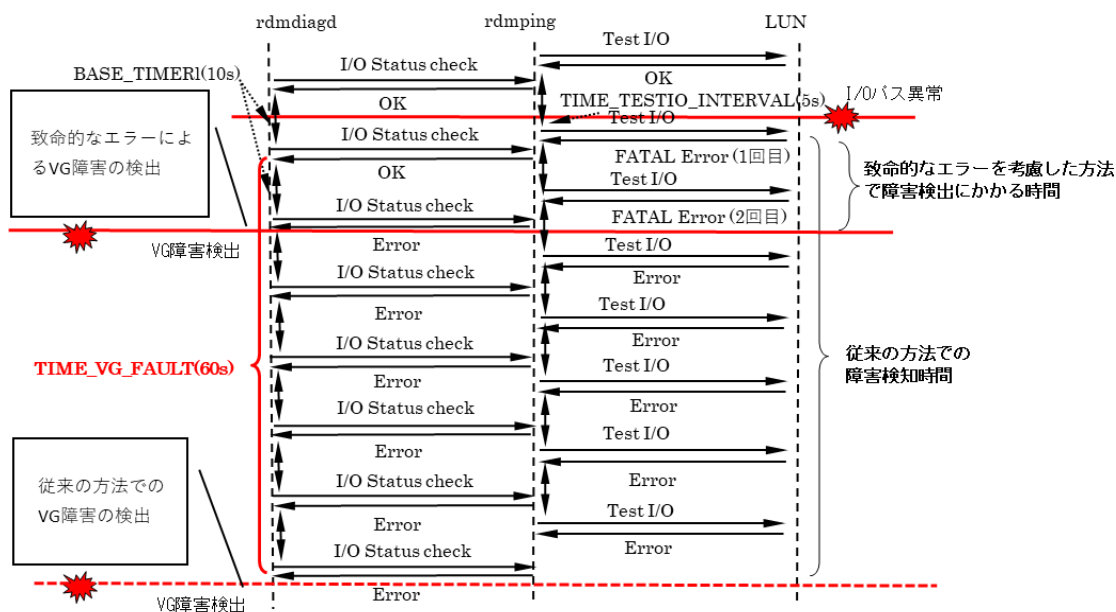
RootDiskMonitor では、物理的な障害が発生してから、障害判定時間(VG 障害:デフォルト 60 秒)内に発行された複数回の Test I/O 結果がすべて異常が返却された場合に、障害として検出されます。そのため、障害の検知には障害判定時間 + α の時間が必要でした。

CLUSTERPRO MC RootDiskMonitor 2.7 for Linux では、従来の検出ロジックに加えて、Test I/O (Inquiry) の結果において致命的なエラーが返却された場合に、障害判定時間に関係なく障害を検知する設定を追加しました。

これにより、致命的なエラーが発生した場合に、従来の障害検出方法では、VG 障害の場合は約 60 秒～75 秒程度かかっていた障害検知時間(デフォルト値の場合)が、約 5 秒～20 秒程度で障害を検出することができます。

VG 障害時の障害検知のシーケンスはつぎのとおりです。

致命的なエラーを考慮した障害検知の場合の方が、従来の障害検知に比べ、早期に障害を検出していることがわかります。



次に、従来の障害検出方法と、致命的なエラーを考慮した障害検出方法の違いを説明します。

(1) 従来の障害検出方法と致命的なエラーを考慮した障害検出方法との違い

従来の障害検出方法と致命的なエラーを考慮した障害検出方法との違いは、次のようになります。

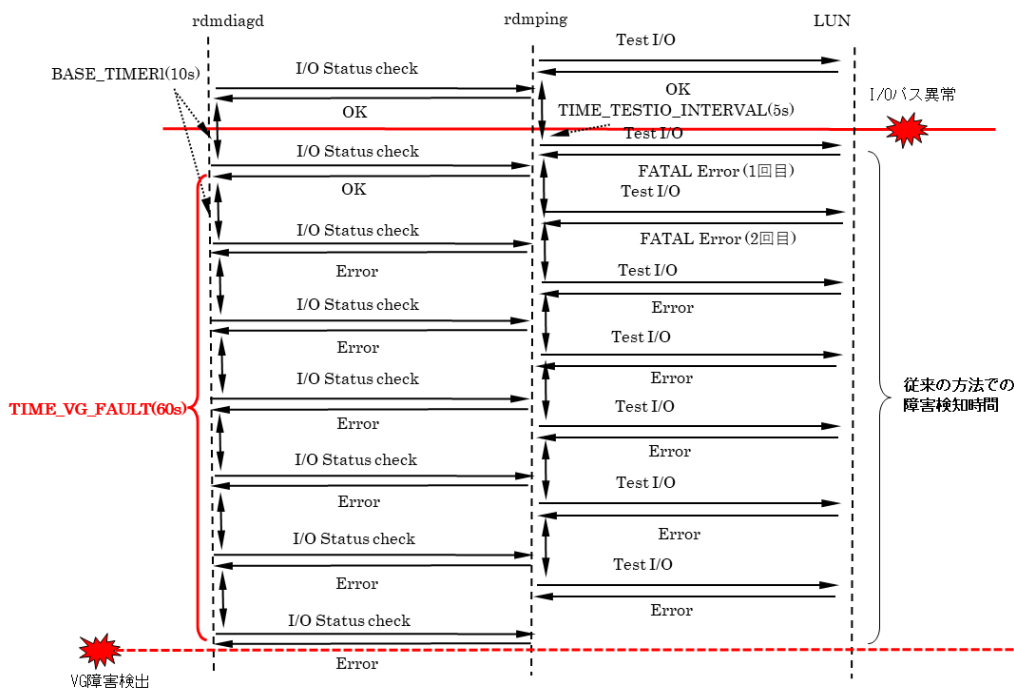
【従来の障害検出方法】

前述のとおり、RootDiskMonitor では、物理的な障害が発生してから、障害判定時間内に発行された複数回の Test I/O 結果がすべて異常が返却された場合に、障害として検出します。

これは、Test I/O の結果を正常・異常の 2 種類で取り扱っているため、回復可能な障害なのか、回復不可能な恒久障害なのかを、継続して障害状態が続いているかどうかで判断しているためです。

従来の障害検出方法の場合のシーケンスは次のとおり。

※VG 障害の検出



【致命的なエラーを考慮した障害検出方法】

致命的なエラーを考慮した障害検出方法の場合、Test I/O の結果として「正常」、「異常」以外に、新たに「致命的なエラー」を追加し、「致命的なエラー」となった場合は、別途指定したリトライ回数（デフォルト 1 回）内に、「致命的なエラー」から回復しなかった場合に障害と判断します。

・致命的なエラーに関する取り扱いについて

ディスク障害が発生した場合、RootDiskMonitor の Test I/O を発行する device の オープンに失敗し、エラーナンバーとして、以下の 2 つが返却されます。

```
ENOENT   ermo = 2
ENXIO    ermo = 6
```

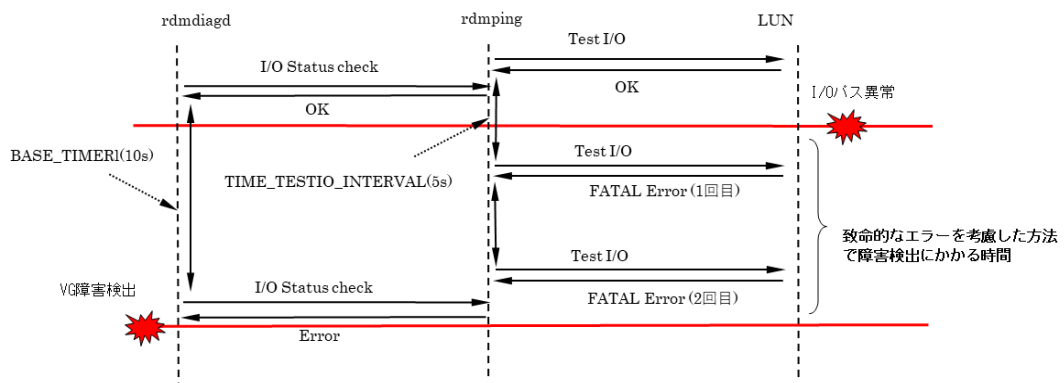
RootDiskMonitor では、Test I/O の結果として、上記のエラーナンバーが返却された場合に、致命的なエラーとして取り扱います。

※ 致命的なエラーとして取り扱うエラーナンバーに関しては、カスタマイズすることも可能です。
カスタマイズをご希望の場合は、開発部門へお問い合わせください。

なお、致命的なエラーを考慮した障害検出方法の設定を行った場合でも、上記の「致命的なエラー」による障害判定に加え、従来どおりの、継続して「異常」となった場合による障害判定も実施されます。

致命的なエラーを考慮した障害検出方法の場合のシーケンスは次のとおり。

※VG 障害の検出



つまり、障害検出にかかる時間が障害判定時間に依存しなくなり、従来の障害検出方法に比べ、より迅速に障害と検出することが可能となりました。

(2) 致命的なエラーを考慮した障害検出方法の設定手順について

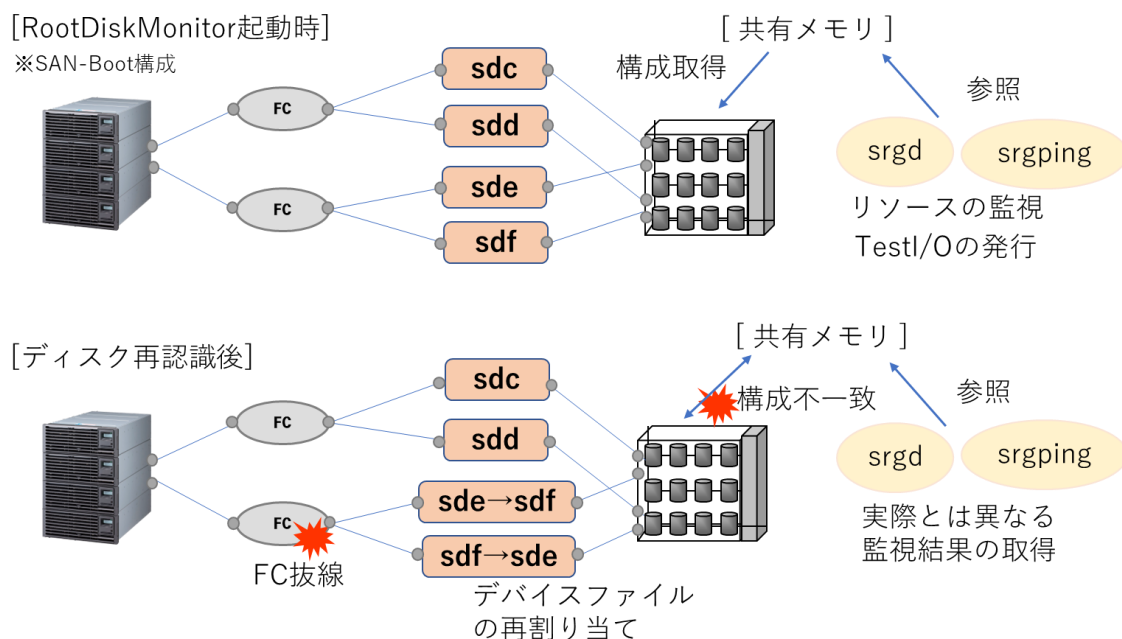
致命的なエラーを考慮した障害検出方法を使用するための設定手順については、後述の "4.7 章. 障害の検知、および、アクション実行の高速化" を参照してください。

なお、致命的なエラーを考慮した障害検出方法は、Test I/O (Inquiry) に対する設定となります。TESTIO_MODE が READ の場合は、Test I/O (Inquiry)を使用しないため、致命的なエラーを考慮した障害検出方法は利用できません。TESTIO_MODE が READ 以外の場合は、使用可能です。

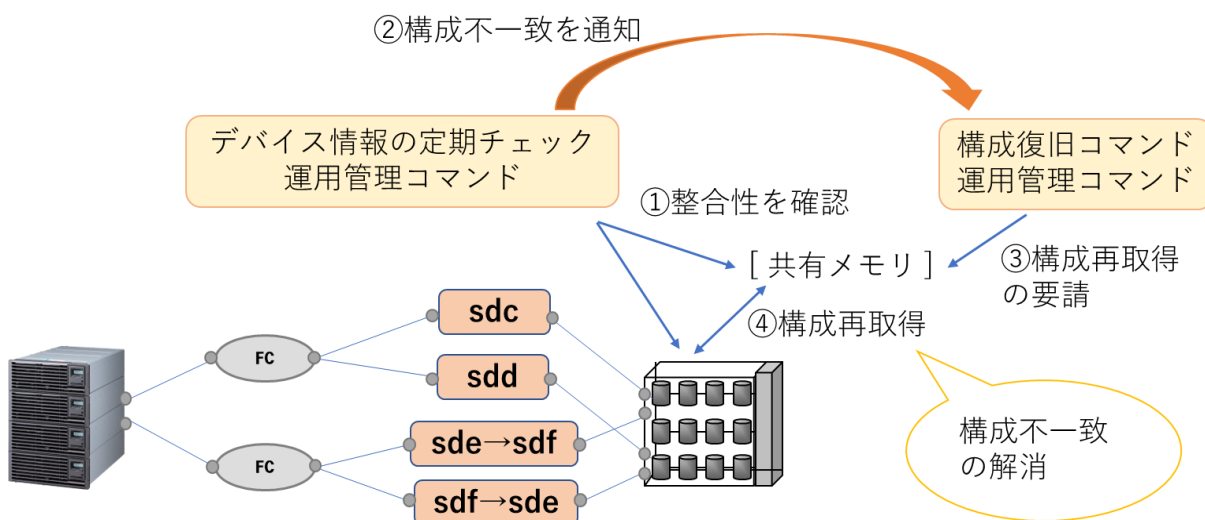
2.6. 監視ディスクのデバイスファイル再解決機能について

FC ケーブルの抜線などにより、OS からディスクの認識が一時的に外れた後、ディスクが再認識された場合、デバイスファイルの再割り当てが行われる可能性があります。

RootDiskMonitor では、起動時に監視対象のデバイス情報を共有メモリに格納し、その情報を使用して監視を行っています。監視対象のデバイスファイルが更新された場合、監視が出来なくなる場合があります。



本機能を使用することで、デバイスファイルの変更が行われた場合でもデバイスファイルを更新し、監視を継続することができます。



本機能は、デバイス情報のチェック機能とデバイス情報の更新機能で構成されており、それぞれの機能について説明します。

2.6.1. デバイス情報のチェック機能

何らかの契機でデバイスファイルの再割り当てが行われると、監視に使用している RootDiskMonitor の共有メモリ内のデバイスファイルと実際のデバイスファイルにずれが生じ、監視が出来なくなる場合があります。本機能では、現在のデバイスファイルを取得し直し、RootDiskMonitor の共有メモリ内のデバイスファイルと比較を行い、デバイスファイルが変更されたことを syslog メッセージで通知します。メッセージは以下のように変更された監視対象のデバイスファイル名が出力されます。

```
Special file changed [hwpath=pci-0000:13:00.1-fc0x2100001697120ca7:0x0001000000000000  
s.f =sdh (from sdg)].
```

上記のメッセージが出力された場合、後述するデバイス情報の更新機能を使用することにより、RootDiskMonitor 内部のデバイスファイル名を現在のデバイスファイル名に更新し、監視を継続することができます。

また、RootDiskMonitor 起動時に監視対象が存在していない場合は、下記メッセージが出力されます。本メッセージが出力された場合、デーモンを再起動することにより、該当パスの監視を再開することができます。

```
unknown path may be recoverable HW-path=pci-0000:13:00.1-fc-0x2100001697120ca7:0x0001000000000000
```

デバイス情報のチェック機能は次の契機で実行されます。

(1) 定期的なデバイス情報のチェック

cron にデバイス情報のチェックコマンドを登録することで、定期的にデバイスファイルの変更が行われているか自動的にチェックし、変更が行われていた場合、syslog ファイルに変更が行われた監視対象のデバイスファイル名が出力されます。

cron の設定手順は" 4.8. 監視ディスクのデバイスファイル再解決手順" を参照してください。

(2) 運用管理コマンドによるデバイス情報のチェック

運用管理コマンドにてデバイス情報のチェックを行うことが可能です。

```
# /opt/HA/RDM/bin/rdmadmin -c devcheck
```

チェック結果と変更が行われた監視対象のデバイスファイル名が出力されます。コマンドの詳細は" 7. リファレンス" を参照してください。

2.6.2. デバイス情報の更新機能

前述のデバイス情報のチェック機能でデバイスファイルの変更が通知された場合に本機能を使用することで、監視を継続することができます。

本機能は RootDiskMonitor の共有メモリ内のデバイスファイル名と現在のデバイスファイルで比較を行い、デバイスファイルの変更が行われていた場合、共有メモリ内のデバイスファイル名を現在のデバイスファイルに更新し、更新結果をメッセージ出力で通知することができます。

メッセージは以下のように更新された監視対象のデバイスファイル名が出力されます。

```
Update: HW-path=pci-0000:13:00.1-fc-0x2100001697120ca7:0x0001000000000000 s.f=sdh
```

また、デバイス情報のチェック機能と同様に、RootDiskMonitor 起動時に監視対象が存在しておらず、本機能のデバイス情報のチェック時にデバイスファイルが取得できた場合に、RootDiskMonitor を再起動することで該当パスの復旧が可能であることをメッセージ出力で通知します。

メッセージはデバイス情報のチェック機能と同様です。

デバイス情報の更新機能は次の契機で実行されます。

(1) 構成復旧コマンドによるデバイス情報の更新

構成復旧コマンドにてデバイス情報の更新を行います。

```
# /opt/HA/RDM/bin/rdmrecover -v
```

更新が行われていた場合、syslog ファイルに更新が行われた監視対象のデバイス情報が出力されます。

(2) 運用管理コマンドによるデバイス情報の更新

運用管理コマンドにてデバイス情報の更新を行うことが可能です。

```
# /opt/HA/RDM/bin/rdmadmin -c devsync
```

このコマンド契機でデバイスファイルが更新された場合、更新結果と更新された監視対象のデバイスファイル名が出力されます。

コマンドの詳細は"7. リファレンス" を参照してください。

3. 設定ファイル

3.1. 本製品の導入

本製品の導入手順について説明します。

本製品のインストール手順については、『CLUSTERPRO MC RootDiskMonitor 2.9 for Linux リリースメモ』をご覧ください。

HW-RAID 構成の場合、別冊

『CLUSTERPRO MC RootDiskMonitor 2.9 for Linux HW-RAID 監視機能 ユーザーズガイド』および『CLUSTERPRO MC RootDiskMonitor 2.9 for Linux HW-RAID 監視機能 リリースメモ』もご覧ください。

(1) セットアップ

OS ディスクを監視するには、設定ファイルの作成が必要です。

設定ファイル名は、`/opt/HA/RDM/conf/rdm.config` です。
サンプルファイルが `/opt/HA/RDM/conf/rdm.config.default` として提供されていますので、このファイルをコピーした後に、OS ディスクを構成するデバイス情報を登録してください。

設定ファイル自動生成コマンド `/opt/HA/RDM/bin/rdmconfig` 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。
既に、設定ファイルが存在する場合は、上書き要否を問い合わせますので「Y/N」から Y を選択してください。N を選択すると、設定ファイル自動生成を中止します。
SAN-Boot 構成の場合は、`-S` オプションを指定して、実行してください。

【内蔵ディスク構成】

```
# /opt/HA/RDM/bin/rdmconfig
```

【SAN-Boot 構成 (Device Mapper Multipath 環境の場合)】

```
# /opt/HA/RDM/bin/rdmconfig -S dmmultipath
```

設定ファイルの自動生成コマンドについては、「7. リファレンス」にて、詳細な説明が記載されております。そちらを参照ください。

自動生成したテンプレートファイルについては、監視ルール、OS ディスクを構成するデバイス情報の妥当性を確認してください。

なお、OS ディスク以外については自動生成対象とはなりませんので、手動で設定ファイルを編集してください。

注意:

- 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。その場合はつぎの手順で、設定ファイルを作成してください。

【設定ファイル手動作成手順】

- ① OS ディスクのデバイスファイル名を確認

```
# df
Filesystem                1K-blocks    Used Available Use% Mounted on
/dev/mapper/mpatha2        8181760     874112   7307648  11% /
/dev/mapper/mpatha1        201388      84216   117172   42% /boot
```

- ② マルチパスデバイスを使用している場合は、対応する sd デバイスファイル名、udev デバイスファイル名を確認

(例) Device Mapper Multipath のデバイスを使用している場合

```
# multipath -ll /dev/mapper/mpatha
size=20G features='0' hwhandler='0' wp=rw
|+- policy='service-time 0' prio=1 status=active
| '- 7:0:0:0 sdb 8:16 active ready running
'+- policy='service-time 0' prio=1 status=enabled
' - 8:0:0:0 sdl 8:176 active ready running
```

```
# ls -l /dev/disk/by-path
lrwxrwxrwx 1 root root 9 Mar 9 18:03 pci-0000:15:00.0-fc-0x220000255c3a0266-lun-0 -> ../../sdb
lrwxrwxrwx 1 root root 9 Mar 9 18:03 pci-0000:15:00.1-fc-0x2a0000255c3a0266-lun-0 -> ../../sdl
```

- ③ 設定ファイルのサンプルファイルをコピー

```
# cp -p /opt/HA/RDM/conf/rdm.config.default /opt/HA/RDM/conf/rdm.config
```

- ④ 設定ファイルの編集

【CLUSTERPRO MC StorageSaver for BootDisk 2.0 以降の場合】

```
# vi /opt/HA/RDM/conf/rdm.config
<中略>
VG          VG_NONE
MIRROR      group01
## PV Name: /dev/sdb
PV          pci-0000:15:00.0-fc-0x220000255c3a0266-lun-0
## PV Name: /dev/sdl
PV          pci-0000:15:00.1-fc-0x2a0000255c3a0266-lun-0
```

【CLUSTERPRO MC StorageSaver for BootDisk 1.2 以前の場合】

```
# vi /opt/HA/RDM/conf/rdm.config
<中略>
VG          VG_NONE
MIRROR      group01
PV          /dev/sdb
PV          /dev/sdl
```

(2) 設定ファイルの変更

- CLUSTERPRO との連携方式によって、設定ファイルの変更が必要です。
カスタムモニタリソースによる CLUSTERPRO X との連動を使ったノード切り替えを行う場合は、下記のパラメーターを変更してください。

パラメーター名 :OVER_ACTION
設定値 :SERVICE_CMD_DISABLE を SERVICE_CMD_ENABLE へ変更

CLUSTERPRO X のサーバー管理プロセス (clpnm) を強制終了する手法によるノード切り替えを行う場合は、下記のパラメーターを変更してください。

パラメーター名 :OVER_ACTION
設定値 :SERVICE_CMD_DISABLE を CLPNM_KILL へ変更

OS ディスク故障時に、システムメモリダンプを採取し OS を強制終了する手法によるノード切り替えを行う場合は、下記のパラメーターを変更してください。

パラメーター名 :OVER_ACTION
設定値 :SERVICE_CMD_DISABLE を TOC_EXEC へ変更

CLUSTERPRO と連携したノード切り替えを行わない場合は OVER_ACTION の変更は不要です。

- 仮想環境 (ゲスト OS) で本製品を使用する場合は、設定ファイルの変更が必要です。
仮想環境 (ゲスト OS) で本製品を使用する場合は、下記のパラメーターを変更してください。

パラメーター名 :TESTIO_MODE
設定値 :INQ_TUR を READ へ変更

- Hyper-V または KVM 上のゲスト OS で本製品を使用する場合は、設定ファイルに udev デバイスファイル名を定義することができません。
設定ファイルを自動生成する際、以下のコマンドを実行してください。

```
# /opt/HARDM/bin/rdmconfig -sd
```

- SAN-Boot 構成で、障害が発生した I/O パスを自動閉塞させる場合は、設定ファイルの変更が必要です。
障害が発生した I/O パスの自動閉塞を行う場合は、下記のパラメーターを変更してください。

パラメーター名 :TESTIO_FAULT_ACTION
設定値 :ACTION_NONE を BLOCK_PATH へ変更

(3) プロセスの再起動

① モニタープロセスの再起動

設定ファイルを作成した後にモニタープロセスを再起動してください。

【Red Hat Enterprise Linux 7.0 以降、Oracle Linux 7.0 以降、Amazon Linux 2 以降の場合】

```
# systemctl stop rdmd  
# systemctl start rdmd
```

【Red Hat Enterprise Linux 6.x、Oracle Linux 6.x の場合】

```
# /etc/init.d/rdmd stop  
# /etc/init.d/rdmd start
```

② モニタープロセスの確認

モニタープロセスが起動されていることを確認してください。

```
# ps -ef|grep rdm  
root 2169 1 0 10:43:40 ? 0:00 /opt/HA/RDM/bin/rdmdiagd  
root 2179 2169 0 10:43:40 ? 0:00 rdmping  
root 2181 2169 0 10:43:40 ? 0:00 rdmreduced
```

※rdmreduced プロセスは、SAN-Boot 環境でかつ自動閉塞が有効になっている場合にのみ存在します。

③ ステータスの確認

監視リソースの状態を確認してください。

```
# /opt/HA/RDM/bin/rdmadmin
```

3.2. 設定ファイルの記述

- (1) 設定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。

`/opt/HA/RDM/conf/rdm.config`

以下に使用するキーワードを記述します。

監視ルール	
項目	説明
TIME_VG_FAULT	監視リソースを異常と判定する時間を指定します。 このパラメーターはデフォルト値を使用することを推奨します。 指定値は 6 秒～、デフォルト 60 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。 このパラメーターはデフォルト値を使用することを推奨します。 指定値は 6 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。
TIME_TESTIO_INTERVAL	Test I/O の発行間隔を指定します。 このパラメーターはデフォルト値を使用することを推奨します。 指定値は 1 秒～、デフォルト 5 秒
WAIT_TESTIO_INTERVAL	Test I/O でパススルードライバーに指定する I/O 待ち合わせ時間を指定します。 このパラメーターはデフォルト値を使用することを推奨します。 指定値は 1 秒～、デフォルト 5 秒
BASE_TIMER	基本タイマーを指定します。 最小値は 1 秒、デフォルト 10 秒
OVER_ACTION	OS ディスク異常検出時のアクションを指定します。 VG 単位に異常検出時のアクションを制御する場合は後述の VOLTYPE パラメーターを指定してください。
SERVICE_CMD_DISABLE	アクション指定なし、デフォルトです。 syslog にメッセージを出力します。
SERVICE_CMD_ENABLE	クラスターウェア連携用デーモン (rdmstat) を使用してクラスターウェア連携を行うことで、ノードを切り替えます。
CLPNM_KILL	CLUSTERPRO X のサーバー管理プロセス (clpnm) を強制終了することでノードを切り替えます。 クラスターウェア連携時にはこの設定を推奨します。
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。

VG_STALL_ACTION	I/O ストール検出時のアクションを指定します。 ※本パラメーターは変更しないことを推奨します。	
	SERVICE_CMD_DISABLE	アクション指定なし、デフォルトです。 I/O ストールを検出すると、syslog にメッセージを出力します。
	SERVICE_CMD_ENABLE	クラスターウェア連携用デーモン(rdmstat)を使用してクラスターウェア連携を行うことで、ノードを切り替えます。
	CLPNM_KILL	CLUSTERPRO X のサーバー管理プロセス (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。
TESTIO_MODE	TestI/O の発行方法を指定します。 ※本パラメーターは特別な要件がないかぎり変更しないことを推奨します。	
	INQ	パススルードライバー経由で Inquiry を発行します。
	INQ_TUR	パススルードライバー経由で 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ー経由で Inquiry と TestUnitReady と Read を発行します。
	DIRECT	設定ファイルに定義されたデバイスファイルに対して直接 Inquiry と TestUnitReady を発行します。
	READ	設定ファイルに定義されたデバイス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。
TESTIO_PERFORMANCE_TYPE	TestI/O の結果から致命的なエラーの検出要否を指定します。 デフォルトでは NORMAL に設定されています。 障害検知、および、アクションの実行を高速化したい場合に本パラメーターを HIGH に設定してください。	
	NORMAL	TEST I/O の結果は、正常・異常の2種類で扱われます。 デフォルトです。
	HIGH	TEST I/O の結果は、正常・異常・致命的なエラーの3種類で扱われます。 致命的なエラーが検出された場合、FATAL_ERROR_RETRY_COUNT で指定された回数連続で返却された場合に、障害を検出します。
FATAL_ERROR_RETRY_COUNT	TESTI/O の結果が致命的なエラーの場合に障害判定までのリトライ回数を指定します。 最小値は 1 回、デフォルト値は 2 回	

SAN_BOOT 用定義	
項目	説明
BOOT_TYPE	BootDisk の構成。
	LOCAL_DISK OS ディスクが内蔵されたディスクに格納された構成。デフォルトです。
	SAN_BOOT SAN-Boot 構成。
TESTIO_FAULT_ACTION	Test/I/O 異常検出時のアクションを指定します。 ※本パラメータは、BOOT_TYPE が SAN_BOOT の場合に有効です。
	ACTION_NONE アクション指定なし、デフォルトです。
	BLOCK_PATH I/O パスを自動閉塞します。
MULTIPATH_TYPE	マルチパス管理ソフトウェアの種類。 ※本パラメータは、BOOT_TYPE が SAN_BOOT の場合に有効です。
	DM_MULTIPATH DeviceMapper Multipath を使用した構成。
	SPS StoragePathSavior を使用した構成。
	POWER_PATH EMC PowerPath を使用した構成。
	HDLM HitachiDynamicLinkManager または、HADynamicLinkManager を使用した構成。
	DMP Veritas DynamicMultiPassing を使用した構成。
	NO_USE マルチパス管理ソフトウェアを使用しない構成。デフォルトです。

デバイス定義					
項目	説明				
VG	OS ディスクのボリューム名を指定します。 LVM 構成の場合は VG 名(/dev/VolGroup00 など)を指定します。 物理ディスク構成の場合はダミーの VG 名として VG_NONE を指定します。 複数の VG を監視する場合は VG-MIRROR-PV の組み合わせを指定してください。				
VOLTYPE ※通常は指定しない。省略可	<p>ディスクの種別を指定します。 ※本パラメーターは通常指定する必要はありません。 OVER_ACTION の動作を変更する必要がない場合は指定しないでください。</p> <table border="1"> <tr> <td>ROOT_VOLUME</td> <td>通常の OS ディスクの場合に指定します。また、指定されていない場合のデフォルトです。 VG ダウンを検出した場合に通常どおり OVER_ACTION の動作を実行します。</td> </tr> <tr> <td>OTHER</td> <td>OS ディスク以外のデータディスクの場合に指定します。OTHER が指定された VG は、VG ダウンを検出した場合でも OVER_ACTION の動作を実行せず、syslog への通報のみとなります。</td> </tr> </table>	ROOT_VOLUME	通常の OS ディスクの場合に指定します。また、指定されていない場合のデフォルトです。 VG ダウンを検出した場合に通常どおり OVER_ACTION の動作を実行します。	OTHER	OS ディスク以外のデータディスクの場合に指定します。OTHER が指定された VG は、VG ダウンを検出した場合でも OVER_ACTION の動作を実行せず、syslog への通報のみとなります。
ROOT_VOLUME	通常の OS ディスクの場合に指定します。また、指定されていない場合のデフォルトです。 VG ダウンを検出した場合に通常どおり OVER_ACTION の動作を実行します。				
OTHER	OS ディスク以外のデータディスクの場合に指定します。OTHER が指定された VG は、VG ダウンを検出した場合でも OVER_ACTION の動作を実行せず、syslog への通報のみとなります。				
MIRROR	OS ディスクのミラーグループを構成する組み合わせを指定します。 指定値は 16 文字以内のノード一意である任意の英数字です。 設定ファイル自動生成コマンドを使用すると、groupxy(xy は 01 からの連番)を登録します。 OS ディスクの中で同一データを構成する I/O パスをグルーピングします。				
PV	<p>I/O パスへの udev デバイスファイルまたはスペシャルファイルを指定します。 スペシャルファイルはブロック型のファイル名を指定します。 (たとえば/dev/sda など) なお、スペシャルファイルは OS 再起動時にファイル名が変更される可能性があります。そのため、PV 行は udev デバイスファイルによる指定を推奨します。 物理ディスクをパーティション分割して利用している場合は、パーティション名の指定ではなく物理ディスク名を指定してください。</p> <ul style="list-style-type: none"> ■ シングルパス構成の場合は、LUN の udev デバイスファイルまたはスペシャルファイルをひとつ指定してください。 ■ ソフトミラー構成の場合は、正副 LUN 両方の udev デバイスファイルまたはスペシャルファイルを指定してください。 ■ FC 接続の代替パス構成の場合は、LUN への正副 I/O パス両方の udev デバイスファイルまたはスペシャルファイルを指定してください。 <p>設定ファイルにスペシャルファイル名と udev デバイスファイル名の両方を指定した場合、デーモンプロセス起動時にスペシャルファイル名の妥当性チェックを行います。スペシャルファイル名が変更されていた場合には、本機能により当該デバイスを監視対象外とします。これにより、誤った設定のまま監視を開始して障害を誤検知することを防止します。スペシャルファイル名と udev デバイスファイル名の間にはスペースが必要です。</p>				

注意:

- ・ 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。
- ・ 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、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。
- ・ CLUSTERPRO を導入し、クラスター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 停止機能を優先させるため、RootDiskMonitor の POWER_OFF 機能を利用することができません。どちらも同等の機能のため、CLUSTERPRO の softdog 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、前述の "2.4 クラスターウェアとの連携について"を参照してください。
- ・ 仮想環境 (ゲスト OS) で本製品を使用する場合は、TESTIO_MODE 値に READ を設定してください。
- ・ 上記タイマー値の上限値は 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。
- ・ デバイスファイルのチェック機能を使用する場合は、"/opt/HARDM/local" ディレクトリが存在するか確認してください。存在しない場合は、該当のディレクトリを作成してください。

(2) 設定ファイルの設定例について

```
# rdm.config (RootDiskMonitor (Linux) Configuration)
```

```
#####  
# System Config Area  
#####
```

```
# Test/I/O interval timer for Root Volume is failed (seconds)  
# Root Volume status changes fail between this timer  
# minimum = 6, default = 60  
TIME_VG_FAULT 60
```

```
# I/O stall interval timer for Root Volume is failed (seconds)  
# Root Volume status changes fail between this timer  
# minimum = 6, default = 360. 0 mean I/O stall no check.  
TIME_VG_STALL 360
```

```
# Test/I/O interval timer value (seconds)  
# exec normal Test/I/O for PV between this timer  
# minimum = 1, default = 5  
TIME_TESTIO_INTERVAL 5
```

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

```
# status check timer (seconds)  
# minimum = 1, default = 10  
BASE_TIMER 10
```

```
# Root Volume fault action  
# select SERVICE_CMD_DISABLE(default),SERVICE_CMD_ENABLE,TOC_EXEC,CLPNM_KILL,POWER_OFF  
OVER_ACTION CLPNM_KILL
```

CLUSTERPRO のサーバー管理プロセス (clpnm) 強制停止によるノード切り替えを行う事例

```
# Root Volume stall find action
# select SERVICE_CMD_DISABLE(default),SERVICE_CMD_ENABLE,TOC_EXEC,CLPNM_KILL,POWER_OFF
VG_STALL_ACTION          SERVICE_CMD_DISABLE
```

変更しないことを推奨

```
# Test/I/O mode
# select INQ,INQ_TUR_READ,READ,DIRECT,INQ_TUR(default)
TESTIO_MODE              INQ_TUR
```

```
# Test I/O performance type
# select NORMAL(default),HIGH
TESTIO_PERFORMANCE_TYPE  NORMAL
```

```
# Fatal error retry count
# minimum = 1, default = 2
FATAL_ERROR_RETRY_COUNT  2
```

```
#####
```

```
# For SANBOOT
#####
```

```
# boot type
# select LOCAL_DISK(default),SAN_BOOT
BOOT_TYPE                 LOCAL_DISK
```

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

```
# multipath type
# select NO_USE(default),DM_MULTIPATH,POWER_PATH,SPS,HDLM,DMP
MULTIPATH_TYPE            NO_USE
```

```
#####
```

```
# Device Config Area
#####
```

```
# VG          volume group for LVM(VolGroup00 , etc ....)
# MIRROR      PV set for Mirror(set any number)
# PV          PV Block Device File(/dev/sda , etc ....)
```

自動生成にて定義されるフォーマット

```
VG          /dev/VolGroup00
MIRROR      group01
PV          pci-0000:0c:00.0-scsi-0:2:0:0
```

ブロックデバイスファイル名を指定する場合の定義例

```
MIRROR      group02
PV          /dev/sda
MIRROR      group03
PV          /dev/sdb pci-0000:0c:00.0-scsi-0:2:0:0
```

ブロックデバイスファイルのチェック機能を利用する場合の定義例
ブロックデバイスファイル名と udev デバイスファイル名を定義します

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/RDM/bin/rdmadmin
(monitor status = TRUE)
```

type		Logical	I/O
		status	status
VG	:/dev/VolGroup00	: up	
MIRROR	:	:	
PV	:/dev/sda	: up	: up
MIRROR	:	:	
PV	:/dev/sdb	: up	: up

リソース監視の on/off を表示します

VG レベルの監視状態を表示します

I/O パスの論理・物理状態を表示します

SAN-Boot 構成の場合は、追加で I/O パスの組み込み状態も表示します。

```
# /opt/HA/RDM/bin/rdmadmin
(monitor status = TRUE)
```

type		Logical	I/O	Online
		status	status	status
VG	:/dev/rhel00	: up		
MIRROR	:	:		
PV	:/dev/sdc	: up	: up	: extended
	pci-0000:15:00.0-fc-0x22008cdf9dcaf0ea-lun-0			
PV	:/dev/sdd	: up	: up	: extended
	pci-0000:15:00.0-fc-0x24008cdf9dcaf0ea-lun-0			
PV	:/dev/sde	: up	: up	: extended
	pci-0000:15:00.1-fc-0x21008cdf9dcaf0ea-lun-0			
PV	:/dev/sdf	: up	: up	: extended
	pci-0000:15:00.1-fc-0x23008cdf9dcaf0ea-lun-0			

リソース監視の on/off を表示します

VG レベルの監視状態を表示します

I/O パスの論理・物理状態を表示します

I/O パスの組み込み状態を表示します

- (2) ディスクに対するすべての 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。

以下の例ではシングル構成のため、/dev/sda が故障した時点で VG ダウンとなります。

```
# /opt/HA/RDM/bin/rdmadmin
(monitor status = TRUE)
```

		: Logical	: I/O
type	: H/W Path	: status	: status
VG	: /dev/VolGroup00	: down	
MIRROR	:	:	
PV	: /dev/sda	: down	: down
MIRROR	:	:	
PV	: /dev/sdb	: up	: up

すべての経路が障害となるため VG レベルで down となります

障害を検出

- (3) 内蔵ディスクがソフトミラー構成で冗長化されている場合は、片系の I/O パスのみが異常になると PV レベルのステータスがダウン状態になります。

このとき VG レベルでは縮退状態で運用されているため、VG のステータスは suspend になります。

```
# /opt/HA/RDM/bin/rdmadmin
(monitor status = TRUE)
```

		: Logical	: I/O
type	: H/W Path	: status	: status
VG	: /dev/VolGroup00	: suspend	
MIRROR	:	:	
PV	: /dev/sda	: down	: down
PV	: /dev/sdb	: up	: up
MIRROR	:	:	
PV	: /dev/sdc	: up	: up
PV	: /dev/sdd	: up	: up

ソフトミラーの片系が障害となった状態です

(4) 機能制限中のリソース監視の状態表示について

コードワードの登録なしに本製品をインストールした場合、インストールから 30 日を経過した後に RootDiskMonitor の機能が制限され、障害を検知なくなります。
機能制限については、「4.5. 機能制限について」を参照してください。

機能制限中にリソース監視の状態を表示した場合、最終行に機能制限中を示すメッセージ「Monitoring stop until activation succeeded.」を表示します。

注意: 機能制限中は障害を検出しないため、以下のステータスは実際の状態と異なる場合があります。

- ・ Logical status
- ・ I/O status

```
# /opt/HA/RDM/bin/rdmadmin
(monitor status = TRUE)
=====
type      :                               : Logical   : I/O
          : H/W Path                          : status    : status
=====
VG        : /dev/VolGroup00                    : up
MIRROR    :                               :
PV        : /dev/sda                          : up        : up
MIRROR    :                               :
PV        : /dev/sdb                          : up        : up
Monitoring stop until activation succeeded.
```

機能制限中を示すメッセージを表示します

- (5) リソース監視の停止と再開についてリソース監視を一時的に停止および再開する場合は以下のコマンドで行います。

```
# /opt/HA/RDM/bin/rdmadmin -c stop
Change TESTIO.
START -> STOP
```

```
# /opt/HA/RDM/bin/rdmadmin -c start
Change TESTIO.
STOP -> START
```

なお、リソース監視停止中は、モニターステータスが FALSE になります。

```
# /opt/HA/RDM/bin/rdmadmin
(monitor status = FALSE) FALSE になります
```

type	:	H/W Path	:	Logical status	:	I/O status
VG	:	/dev/VolGroup00	:	up	:	
MIRROR	:		:		:	
PV	:	/dev/sda	:	up	:	up
MIRROR	:		:		:	
PV	:	/dev/sdb	:	up	:	up

- (6) 3 秒間隔でリソースの状態を定期表示します。

```
# /opt/HA/RDM/bin/rdmadmin -c status -t 3
(monитор status = TRUE)
=====
type      :      : Logical   : I/O
          : H/W Path : status   : status
=====
VG        : /dev/VolGroup00 : up
MIRROR    :
PV        : /dev/sda       : up      : up
MIRROR    :
PV        : /dev/sdb       : up      : up

<... 3 秒経過 ...>

(monитор status = TRUE)
=====
type      :      : Logical   : I/O
          : H/W Path : status   : status
=====
VG        : /dev/VolGroup00 : up
MIRROR    :
PV        : /dev/sda       : up      : up
MIRROR    :
PV        : /dev/sdb       : up      : up
```

(注)コマンドを終了させたい場合、ctrl+c で終了できます。

(7) コンフィグレーション情報を表示します。

```
# /opt/HA/RDM/bin/rdmadmin -c param
SG parameters.
-----
TIME_VG_FAULT           60
TIME_VG_STALL           360
TIME_TESTIO_INTERVAL    5
WAIT_TESTIO_INTERVAL    5
OVER_ACTION             SERVICE_CMD_DISABLE
VG_STALL_ACTION          SERVICE_CMD_DISABLE
TUR_TESTIO_USE           TRUE
TESTIO_DIRECT            INQ_TUR
testioYN                 TRUE
TESTIO_MODE              INQ_TUR
TESTIO_PERFORMANCE_TYPE NOMAL
FATAL_ERROR_RETRY_COUNT 2
BOOT_TYPE                LOCAL_DISK
TESTIO_FAULT_ACTION      ACTION_NONE
MULTIPATH_TYPE           NO_USE
AUTO_REDUCE_INTERVAL    5

SG device list.
-----
SYSTEM
VG          /dev/VolGroup00
VOLTYPE     ROOT_VOLUME
MIRROR
PV          /dev/sda
MIRROR
PV          /dev/sdb
```

(8) モニタープロセスの起動、終了

- 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。

デフォルトで自動起動が設定されるのはランレベル 3 および 5 です。

- 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。

【Red Hat Enterprise Linux 7.0 以降、Oracle Linux 7.0 以降、Amazon Linux 2 以降の場合】

```
# systemctl start rdmd
```

【Red Hat Enterprise Linux 6.x、Oracle Linux 6.x の場合】

```
# /etc/init.d/rdmd 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。

【Red Hat Enterprise Linux 7.0 以降、Oracle Linux 7.0 以降、Amazon Linux 2 以降の場合】

```
# systemctl stop rdmd
```

【Red Hat Enterprise Linux 6.x、Oracle Linux 6.x の場合】

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

上記コマンドで終了しない場合は、`ps -ef|grep rdmd` で rdmd 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、以下のコマンドを root 権限で投入することで、デーモンプロセスを再起動できます。

【Red Hat Enterprise Linux 7.0 以降、Oracle Linux 7.0 以降、Amazon Linux 2 以降の場合】

```
# systemctl restart rdmd
```

【Red Hat Enterprise Linux 6.x、Oracle Linux 6.x の場合】

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

Red Hat Enterprise Linux 7.0 以降、Oracle Linux 7.0 以降、Amazon Linux 2 以降において、RootDiskMonitor のプロセスが一部残存していた場合には、start コマンドではサービスが起動できません。そのため、RootDiskMonitor を再起動する場合には、start コマンドではなく、restart コマンドを使用してください。

4.2. CLUSTERPRO X との連携

4.2.1. カスタムモニタリソースによる CLUSTERPRO X との連携

(1) CLUSTERPRO X との連携について

OS ディスクの動作状態をモニターするコマンド `rdmstat` を CLUSTERPRO X のカスタムモニタリソースとして登録することで、OS ディスクの障害時のノードダウン、ノード切り替えを実現します。

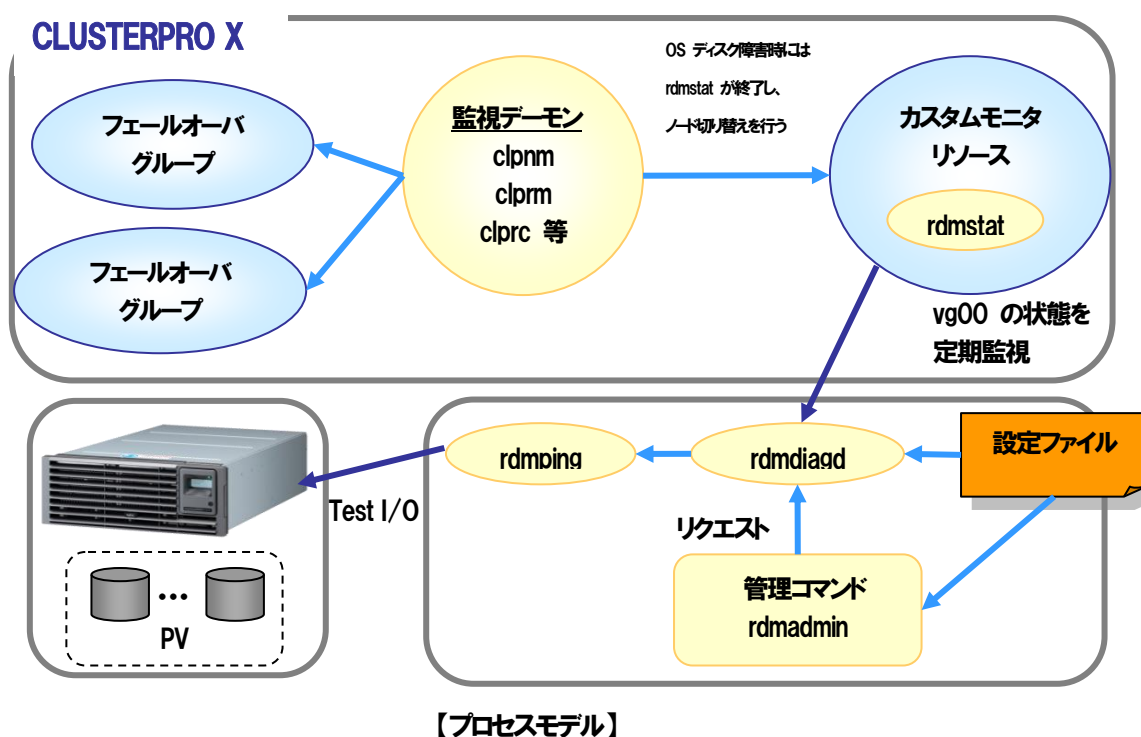
本機能を利用する場合は、不必要に CLUSTERPRO X のサーバー管理プロセス (`clpnm`) を kill しないために RootDiskMonitor のコンフィグレーションの **OVER_ACTION** には、**SERVICE_CMD_ENABLE** を指定してください。

この方式であれば、複数ノードクラスター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 ノードのクラスター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。

RootDiskMonitor としては本方式を推奨します。

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

「クラスターサービス停止と 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ールオーバー処理が正しく完了せず、フェールオーバーに失敗したり、フェールオーバー完了が遅延したりする場合があります。



(2) rdmstat の運用について

OS ディスクに障害が発生すると、rdmdiagd 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。

rdmstat 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ーします。VG が down 状態に遷移した時点で、rdmstat は異常終了し、CLUSTERPRO X 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。

rdmstat は、rdmdiagd および rdmping のプロセスが起動され、OS ディスク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。

- rdmdiagd および rdmping のプロセスが起動していない。
- rdmadmin のオペレーション操作でリソース監視停止を指示している。

<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。

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

- OS ディスクの障害を検出し VG ステータスをダウンに変更します。
Jan 10 18:05:29 node1 rdm[24031]: PV down find .(sf=/dev/sda)
Jan 10 18:05:29 node1 rdm[24031]: VG status change down .(vg=/dev/VolGroup00)

また、LVM かつソフトウェアミラー構成の例は以下となります。

- 最初に、ソフトミラーの片系障害(正系 LUN 障害)を検出
Jan 10 18:03:28 node1 rdm[24031]: PV down find .(sf=/dev/sda)
Jan 10 18:03:28 node1 rdm[24031]: VG status change suspend .(vg=/dev/VolGroup00)
- 次に、ソフトミラーの両系障害(副系 LUN 障害)を検出
Jan 10 18:05:29 node1 rdm[24031]: PV down find .(sf=/dev/sdb)
Jan 10 18:05:29 node1 rdm[24031]: VG status change down .(vg=/dev/VolGroup00)

(3) カスタムモニタリソースの設定手順について

rdmstat を CLUSTERPRO X 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

「8.2. カスタムモニタリソースによる CLUSTERPRO X 4.0 以前との連携手順」
または

「8.3. カスタムモニタリソースによる CLUSTERPRO X 4.1 以降との連携手順」

を参照してください。

4.2.2. サーバー管理プロセス (clpnm) の強制終了による CLUSTERPRO X との連携

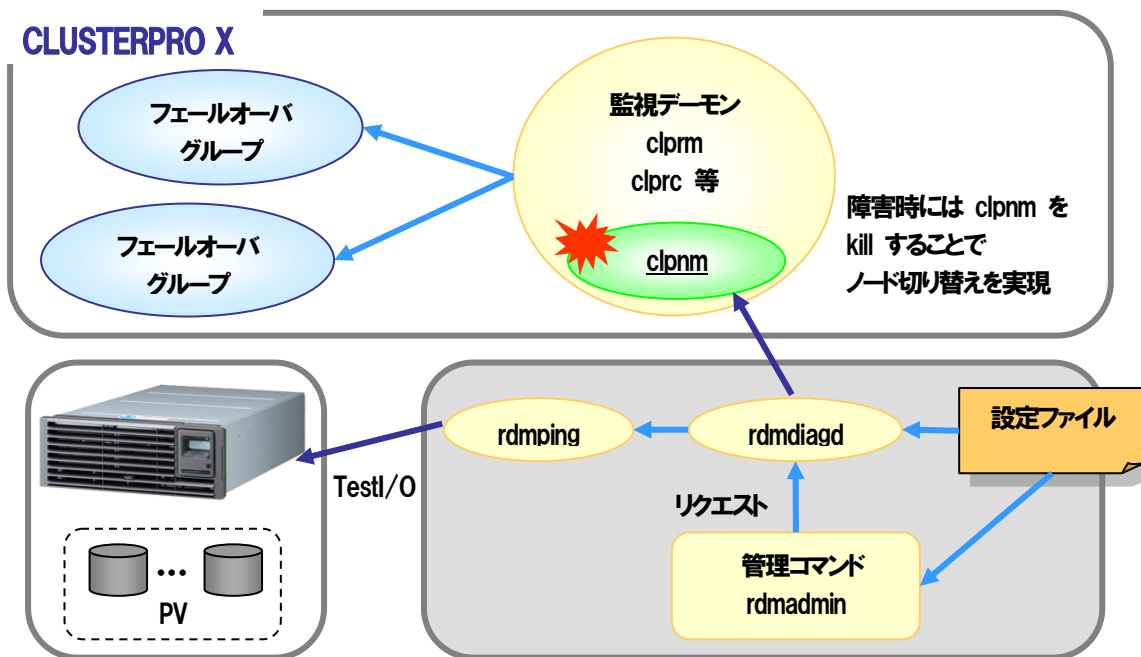
(1) CLUSTERPRO X との連携について

OS ディスクの障害時に、CLUSTERPRO X のサーバー管理プロセス (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。

RootDiskMonitor のコンフィグレーションである **OVER_ACTION** に **CLPNM_KILL** を指定すると、OS ディスクの障害時に CLUSTERPRO X のサーバー管理プロセス (clpnm) を kill することができます。

この方式であれば、内蔵ディスク障害が発生した場合でも、確実にノード切り替えが実現可能です。

このノード切り替え機能は、2 ノード以上のクラスターシステムで有効です。



【プロセスモデル】

(2) OS ディスクの障害を検出すると

OS ディスクに障害が発生すると、rdmdiagd 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。

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

/opt/HA/RDM/bin 配下には、core ファイルやトレースファイルを出力しますので、OS ディスク故障時にこれらのファイルを確認してください。
なお、OS ディスク故障時には、ファイル I/O が失敗する可能性がありますので 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>

LVM かつシングルディスクの構成の例です。

下記の順序で syslog にメッセージが出力されます。

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

- OS ディスクの障害を検出し VG ステータスをダウンに変更
Jan 10 18:05:29 node1 rdm[24031]: PV down find .(sf=/dev/sda)
Jan 10 18:05:29 node1 rdm[24031]: VG status change down .(vg=/dev/VolGroup00)
- 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 rdm[24031]: start KILL clpnm.
Jan 10 18:05:29 node1 rdm[24031]: send signal clpnm.
Jan 10 18:05:29 node1 rdm[24031]: abort rdmdiagd.

また、LVM かつソフトウェアミラー構成の例は以下となります。

- 最初に、ソフトミラーの片系障害(正系 LUN 障害)を検出
Jan 10 18:03:28 node1 rdm[24031]: PV down find .(sf=/dev/sda)
Jan 10 18:03:28 node1 rdm[24031]: VG status change suspend .(vg=/dev/VolGroup00)
- 次に、ソフトミラーの両系障害(副系 LUN 障害)を検出
Jan 10 18:05:29 node1 rdm[24031]: PV down find .(sf=/dev/sdb)
Jan 10 18:05:29 node1 rdm[24031]: VG status change down .(vg=/dev/VolGroup00)
- 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 rdm[24031]: start KILL clpnm.
Jan 10 18:05:29 node1 rdm[24031]: send signal clpnm.
Jan 10 18:05:29 node1 rdm[24031]: abort rdmdiagd.

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

(1) CLUSTERPRO X との連携について

OS ディスク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 X と連携してノード切り替えを実現します。

OS ディスク故障時には 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)内蔵ディスク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 X が予備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 X のサーバー管理プロセス (clpnm) を強制停止することで待機系への切り替えを試みます。

4.3. HW-RAID 状態表示手順

RootDiskMonitor HW-RAID 監視機能をインストールすることで、運用管理コマンドを用いて HW-RAID 構成の論理ディスクおよび物理ディスクの状態表示を行うことが可能となります。

ただし、監視対象ディスクを構成している論理ディスクおよび物理ディスクのみの表示となります。

RootDiskMonitor HW-RAID 監視機能 については別冊『CLUSTERPRO MC RootDiskMonitor 2.9 for Linux HW-RAID 監視機能 ユーザーズガイド』を参照してください。

(1) 運用管理コマンドによる HW-RAID 状態表示

以下のコマンドにて HW-RAID 状態表示を行います。

```
# /opt/HA/RDM/bin/rdmadmin -c pdstatus
```

表示結果は、Universal RAID Utility の場合と HPE Smart Storage Administrator の場合で異なります。

それぞれの表示例は、次のページをご覧ください。

【 Universal RAID Utility の場合 】

```
# /opt/HA/RDM/bin/rmdadmin -c pdstatus
(monitor status = TRUE)
```

=====			
type	:	Logical	I/O
:	:	status	status
=====			
VG	: VG_NONE	: up	
MIRROR	:	:	
PV	:/dev/sda	: up	: up
LD	: 1	: Online	
PD	: 1	: Online	
PD	: 2	: Online	
MIRROR	:	:	
PV	:/dev/sdb	: up	: up
LD	: 2	: Degraded	
PD	: 3	: Failed	
PD	: 4	: Online	

① 論理ディスクのステータス

② 物理ディスクのステータス

① 論理ディスク(LD) の状態を表示します。

- Online 論理ディスクの冗長性が保たれている状態です。
- Degraded 論理ディスクの冗長性が失われているか、冗長性が低下した状態です。
論理ディスクへのアクセスは可能です、
- Offline 論理ディスクが停止し、論理ディスクへのアクセスも不可能な状態です。
- Unknown 論理ディスクの状態が不明な場合に表示されます。

② 物理ディスク(PD) の状態を表示します。

- Online 物理ディスクが論理ディスクに組み込まれており、正常に動作している状態です。
- Failed 物理ディスクが故障している状態です。
- Rebuilding 物理ディスクがリビルド中の状態です。
- Hot Spare 物理ディスクがホットスペアに設定されている状態です。
- Ready 物理ディスクが論理ディスクに組み込まれていない状態です。
- Unknown 物理ディスクの状態が不明な場合に表示されます。

【 HPE Smart Storage Administrator の場合 】

```
# /opt/HA/RDM/bin/rdmadmin -c pdstatus
(monitor status = TRUE)
```

=====			
type	:	Logical	I/O
=====			
	:	status	status
=====			
VG	:	VG_NONE	up
MIRROR	:		
PV	:	/dev/sda	up
SLOT	:	0	
LD	:	1	OK
PD	:	11:3:1	OK
PD	:	11:3:2	OK
MIRROR	:		
PV	:	/dev/sdb	up
SLOT	:	0	up
LD	:	2	Interim Recovery Mode
PD	:	11:3:3	Failed
PD	:	21:2:1	OK

①論理ディスクのステータス

②物理ディスクのステータス

① 論理ディスク(LD) の状態を表示します。

- OK 論理ディスクの冗長性が保たれている状態です。
- Interim Recovery Mode 論理ディスクの冗長性が失われているか、冗長性が低下した状態です。論理ディスクへのアクセスは可能です、
- Failed 論理ディスクが停止し、論理ディスクへのアクセスも不可能な状態です。

② 物理ディスク(PD) の状態を表示します。

- OK 物理ディスクが論理ディスクに組み込まれており、正常に動作している状態です。
- Failed 物理ディスクが故障している状態です。

(2) 注意・制限事項

- HW-RAID の冗長性が低下しても OS からは正常に見えるため PV のステータスは up になります。
- 論理ディスクが Offline の場合は論理ディスクおよび物理ディスクの状態は表示されなくなります。
- 本コマンドで障害を検出した場合は HW 部門に確認を取ってください。
- RAID コントローラーが複数ある環境はサポート対象外です。
- LD に付加されるデバイス名(/dev/sdX)は補助的な情報のため故障箇所の特定は HW 構成を確認してください。

4.4. S.M.A.R.T.診断結果表示手順

(1) S.M.A.R.T.について

S.M.A.R.T.(Self-Monitoring, Analysis and Reporting Technology)とは、障害の早期発見、および故障予測を目的としたディスクに内蔵されている自己診断機能です。

本機能では、運用管理コマンドのリソース監視の状態表示にS.M.A.R.T.診断結果を表示します。

本機能により、リソース監視の状態表示で監視対象ディスクが故障間近であるか確認することができます。

なお、HW-RAID 構成の場合とHW-RAID 構成以外の場合で本機能を利用する手順が異なります。使用する環境にあわせてご確認ください。

(2) 前提条件と環境確認手順

本機能は以下の条件を満たす環境で使用できます。

- NEC 社製 Express5800 シリーズ
- 監視対象ディスクでS.M.A.R.T.診断結果を使用できる
- Red Hat Enterprise Linux 6.8 以上 または 7.2 以上
- 依存パッケージをインストールしている

本機能を使用するための環境確認手順を以下に記述します。

● HW-RAID 構成の場合

① 依存パッケージの確認

HW-RAID 構成でS.M.A.R.T.診断結果を表示する場合、RootDiskMonitor HW-RAID 監視機能が必要です。

パッケージ名

clusterpro-mc-rdmhwraid-w.x.y-z

※w, x, y, z は任意のバージョン番号が入ります

RootDiskMonitor HW-RAID 監視機能 については別冊

『CLUSTERPRO MC RootDiskMonitor 2.9 for Linux HW-RAID 監視機能 ユーザーズガイド』を参照してください。

② S.M.A.R.T.診断結果の取得確認

"Physical Device" 内に "S.M.A.R.T." 行が表示されることを確認します。

"Physical Device" は、HW-RAID を構成する物理ディスクの数だけ表示します。

各物理ディスクで "S.M.A.R.T." 行が表示されることを確認します。

※HPE Smart Storage Administrator を使用した構成の場合は、本作業は不要です。

```
# raidcmd property -tg=all
RAID Controller #1
ID : 0
Vendor : Avago
Model : LSI MegaRAID SAS 9272-8i
Firmware Version : 1.40.32-0580
Cache Size : 128MB
Battery Status : Normal
Rebuild Priority : Low
Consistency Check Priority : Low
Patrol Read : Enable
Patrol Read Priority : Low
Buzzer Setting : Disable

RAID Controller #1 Disk Array #1
ID : 0
Sector Format : 512
Capacity : 148GB
Unused Capacity : 0GB
Type : Disk Array
Physical Device : e252s0

RAID Controller #1 Logical Drive #1
ID : 0
Disk Array Information : 1 (order 1/1)
RAID Level : RAID 0
Sector Format : 512
Capacity : 148GB
Stripe Size : 64KB
Cache Mode (Setting) : Auto Switch
Cache Mode (Current) : Write Through
Type : Logical Drive
Status : Online

RAID Controller #1 Physical Device e252s0
Enclosure : 252
Enclosure Position : Internal
Slot : 0
ID : 0
Device Type : HDD
Interface : SATA
Vendor/Model : ATA ST3160815AS
Firmware Version : A
Serial Number : 9RXJL8KF
Sector Format : 512
Capacity : 148GB
Status : Online
S.M.A.R.T. : Normal
```

HW-RAID を構成する
物理ディスク

S.M.A.R.T.診断結果

● HW-RAID 構成以外の場合

① 依存パッケージの確認

HW-RAID 以外の構成で S.M.A.R.T.診断結果を表示する場合、以下のパッケージが必要です。

smartmontools **Tools for monitoring SMART capable hard disks**

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 smartmontools
smartmontools-w.x.y.z
```

- ・ 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:w, x, y, z には smartmontools 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。

② デバイスファイル名の確認

以下のコマンドで "/"、および "/boot" のデバイスファイル名(/dev/sdX)を確認します。
ご使用の環境によりパーティション番号が表示されますが、パーティション番号を除外した
名称(/dev/sdX)を確認します。

```
# df
ファイルシス 1K-ブロック 使用 使用可 使用% マウント位置
/dev/sda2      20511312 3040512 16405840 16% /
/dev/sda1      508588 119136 389452 24% /boot
```

デバイスファイル名

③ S.M.A.R.T.診断結果の利用確認

smartctl コマンドで、前述のデバイスファイル名(/dev/sdX) を指定し、S.M.A.R.T.診断結果の利用を確認します。

以下のどちらかのメッセージが出力されることを確認してください。

どちらのメッセージも表示されない場合、S.M.A.R.T.診断結果を表示できません。

・メッセージ 1

SMART support is: Available - device has SMART capability.

SMART support is: Enabled

・メッセージ 2

Device supports SMART and is Enabled

```
# smartctl -i /dev/sda
smartctl 5.43 2012-06-30 r3573 [x86_64-linux-2.6.32-642.el6.x86_64]
Copyright (C) 2002-12 by Bruce Allen, http://smartmontools.sourceforge.net

=== START OF INFORMATION SECTION ===
Model Family:      Seagate Barracuda 7200.10
Device Model:      ST3160815AS
Serial Number:     9RXJL8KF
Firmware Version: 4.AAA
User Capacity:     160,000,081,920 bytes [160 GB]
Sector Size:       512 bytes logical/physical
Device is:         In smartctl database [for details use: -P show]
ATA Version is:    7
ATA Standard is:   Exact ATA specification draft version not indicated
Local Time is:     Thu Dec 22 11:16:34 2016 JST
SMART support is: Available - device has SMART capability.
SMART support is: Enabled
```

デバイスファイル名

確認メッセージ

または

```
# smartctl -i /dev/sda
smartctl 5.43 2012-06-30 r3573 [x86_64-linux-2.6.32-642.el6.x86_64]
Copyright (C) 2002-12 by Bruce Allen, http://smartmontools.sourceforge.net

Vendor:            LSI
Product:           MegaRAID 8708EM2
Revision:          1.40
User Capacity:     159,450,660,864 bytes [159 GB]
Logical block size: 512 bytes
Logical Unit id:   0x600605b0017bdd6012a04c4212fb665a
Serial number:     005a66fb12424ca01260dd7b01b00506
Device type:       disk
Local Time is:     Tue Dec 27 15:13:41 2016 JST
Device supports SMART and is Enabled
Temperature Warning Enabled
```

デバイスファイル名

確認メッセージ

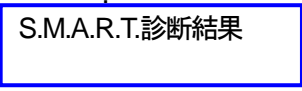
(3) S.M.A.R.T.診断結果の表示手順

① S.M.A.R.T.診断結果の表示手順

S.M.A.R.T.診断結果は、本製品の運用管理コマンドで確認することができます。
本製品の導入については、"3.1 本製品の導入" を参照してください。


● HW-RAID 構成の場合

```
# /opt/HA/RDM/bin/rdmadmin -c pdstatus -s
(monitor status = TRUE)
=====
type      :                               : Logical   : I/O      :
          : HW Path                          : status    : status    : S.M.A.R.T.
=====
VG        : VG_NONE                          : up        :           :
MIRROR    :                               :           :           :
PV        : /dev/sda                            : up        : up        :
LD        : 1                                    : Online    :           :
PD        : 1                                    : Online    :           : Normal
PD        : 2                                    : Online    :           : Normal
=====
```



● HW-RAID 構成以外の場合

```
# /opt/HA/RDM/bin/rdmadmin -s
(monitor status = TRUE)
=====
type      :                               : Logical   : I/O      :
          : HW Path                          : status    : status    : S.M.A.R.T.
=====
VG        : /dev/VolGroup00                    : up        :           :
MIRROR    :                               :           :           :
PV        : /dev/sda                            : up        : up        : Normal
MIRROR    :                               :           :           :
PV        : /dev/sdb                            : up        : up        : Normal
=====
```



S.M.A.R.T.診断結果として、以下の状態を表示します。

- Normal
正常な状態です。
- Detected
障害が発生している、または障害発生間近の状態です。
- Unknown
S.M.A.R.T.診断結果を取得できない状態です。

② S.M.A.R.T.診断結果異常時の対応

S.M.A.R.T.診断結果で Detected を検出した場合、該当ディスクの障害発生が予想されます。

早急に該当ディスクの点検を行ってください。

HW-RAID 構成以外の場合、Detected を検出したディスクの障害情報が出力されます。

詳細は HW 部門に確認してください。

実行例を以下に記載します。

● HW-RAID 構成の場合

```
# /opt/HA/RDM/bin/rdmadmin -c pdstatus -s
(monитор status = TRUE)
=====
type      :      : Logical      : I/O      :
          : HW Path  : status     : status    : S.M.A.R.T.
=====
VG        : VG_NONE   : up
MIRROR    :
PV        : /dev/sda   : up        : up
LD        : 1          : Online
PD        : 1          : Online    :           : Detected
PD        : 2          : Online    :           : Normal
```

障害発生が予想されるディスク

● HW-RAID 構成以外の場合

```
# /opt/HA/RDM/bin/rdmadmin -s
WARN: S M A R T error is detected (sf=/dev/sda)
===== START OF READ SMART DATA SECTION =====
SMART overall-health self-assessment test result: FAILED!
Drive failure expected in less than 24 hours. SAVE ALL DATA.
Failed Attributes:
ID# ATTRIBUTE_NAME          FLAG      VALUE WORST THRESH TYPE
UPDATED WHEN_FAILED RAW_VALUE
  5 Reallocated_Sector_Ct   0x0033   134   134   140   Pre-fail Always
Failing_NOW 1961
```

障害情報

```
(monitor status = TRUE)
=====
type      :      : Logical      : I/O      :
          : HW Path  : status     : status    : S.M.A.R.T.
=====
VG        : /dev/VolGroup00 : up
MIRROR    :
PV        : /dev/sda   : up        : up        : Detected
MIRROR    :
PV        : /dev/sdb   : up        : up        : Normal
```

障害発生が予想されるディスク

本機能では、以下のメッセージをコンソールに出力します。

● HW-RAID 構成の場合

- **rdm_hwraidmon is not installed.**
説明:HW-RAID 監視機能がインストールされていません。
処置:HW-RAID 監視機能をインストールしてください。

- **cannot get hwraid info [x].**
説明:Universal RAID Utility がインストールされていません。
処置:Universal RAID Utility をインストールしてください。

- **cannot get S.M.A.R.T. status.(pd=x)**
説明:該当ディスクの S.M.A.R.T.診断結果の取得に失敗しました。
処置:本章の "(2) 前提条件と環境確認手順" を参照し、ご使用の環境を確認してください。

- **get invalid S.M.A.R.T. status.(xxx)(pd=x)**
説明:該当ディスクの S.M.A.R.T.診断結果で不正な値を取得しました。
処置:本章の "(2) 前提条件と環境確認手順" を参照し、ご使用の環境を確認してください。

- **cannot get LD/PD info.(pv=/dev/sdX)**
説明:該当ディスクの情報を取得できませんでした。
処置:本章の "(2) 前提条件と環境確認手順" を参照し、ご使用の環境を確認してください。

- **cannot get PD information(xxx)**
説明:該当ディスクの情報で不正な値を取得しました。
処置:本章の "(2) 前提条件と環境確認手順" を参照し、ご使用の環境を確認してください。

● HW-RAID 構成以外の場合

- **S.M.A.R.T. error is detected.(sf=/dev/sdX)**
説明:障害が発生している、または障害発生間近の状態です。
処置:早急に該当ディスクの点検を行ってください。

- **smartmontools is not installed.**
説明:smartmontools がインストールされていません。
処置:smartmontools をインストールしてください。

- **cannot get S.M.A.R.T. support.(sf=/dev/sdX)**
説明:該当ディスクの S.M.A.R.T.の使用可否を取得できませんでした。
処置:本章の "(2) 前提条件と環境確認手順" を参照し、ご使用の環境を確認してください。

- **S.M.A.R.T. support is Unavailable.(sf=/dev/sdX)**
説明:該当ディスクでは S.M.A.R.T.を使用できません。
処置:該当ディスクでは S.M.A.R.T.診断結果を表示できません。

- **cannot get S.M.A.R.T. Enabled/Disabled.(sf=/dev/sdX)**
説明:該当ディスクの S.M.A.R.T.の使用可否を取得できませんでした。
処置:本章の "(2) 前提条件と環境確認手順" を参照し、ご使用の環境を確認してください。

- **S.M.A.R.T. support is Disabled.(sf=/dev/sdX)**
説明:該当ディスクでは S.M.A.R.T.を使用できません。
処置:該当ディスクでは S.M.A.R.T.診断結果を表示できません。

- **cannot get S.M.A.R.T. status.(sf=/dev/sdX)**
説明:該当ディスクの S.M.A.R.T.診断結果を取得できませんでした。
処置:本章の "(2) 前提条件と環境確認手順" を参照し、ご使用の環境を確認してください。

- **get invalid S.M.A.R.T. status.(sf=/dev/sdX)**
説明:該当ディスクの S.M.A.R.T.診断結果で不正な値を取得しました。
処置:本章の "(2) 前提条件と環境確認手順" を参照し、ご使用の環境を確認してください。

4.5. 機能制限について

CLUSTERPRO MC RootDiskMonitor 2.3 for Linux までは、インストール時にライセンスをチェックするため、インストール前にコードワードの登録が必要でした。

CLUSTERPRO MC RootDiskMonitor 2.4 for Linux より、コードワードが未登録でもインストールが可能となりました。なお、インストールから 30 日を経過してもコードワードが未登録の場合には RootDiskMonitor の機能に制限がかかり、障害発生時でも障害を検知なくなります。

コードワードが未登録の場合、以下の契機でメッセージが出力されます。

- ・ インストールから 30 日以内
 - After YYYYMMDD, monitoring function is stopped.
説明:ライセンス認証に失敗しました。
YYYYMMDD までは通常どおり使用できますが、経過後は機能制限を行います。
機能制限中は障害が検知されません。
- ・ インストールから 30 日経過後
 - Monitoring stop until activation succeeded.
説明:ライセンス認証に失敗しました。
正しいコードワードの登録が確認できるまで RootDiskMonitor の機能が制限されます。
機能制限中は障害が検知されません。

これらのメッセージが表示される場合、コードワードを登録してください。

手順は『CLUSTERPRO MC RootDiskMonitor 2.9 for Linux リリースメモ』の「1.2.2. ソフトウェアパッケージのインストール後にコードワードを登録する方法」を参照してください。

4.6. オンライン保守コマンドの操作手順

SAN-Boot 構成の場合は、オンライン保守コマンドにより、I/O パスの閉塞・復旧が可能です。

(1) rdmreduce 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: 設定ファイル(rdm.config)の「BOOT_TYPE」が「SAN_BOOT」の場合のみ有効です。
また、RootDiskMonitor が起動している状態でのみ使用可能です。

- ・ -p を付与し、udev デバイス名を指定して I/O パスを切り離します。
 - 状態確認

```
# /opt/HA/RDM/bin/rdmadmin
(monitor status = TRUE)
=====
type      :                               : Logical  : I/O      : Online
          : H/W Path                            : status   : status   : status
=====
VG        : /dev/rhel00                        : up
MIRROR    :                               :
PV        : /dev/sdc                          : up       : up       : extended
          : pci-0000:15:00.0-fc-0x22008cdf9dcaf0ea-lun-0
PV        : /dev/sdd                          : up       : up       : extended
          : pci-0000:15:00.0-fc-0x24008cdf9dcaf0ea-lun-0
PV        : /dev/sde                          : up       : up       : extended
          : pci-0000:15:00.1-fc-0x21008cdf9dcaf0ea-lun-0
PV        : /dev/sdf                          : up       : up       : extended
          : pci-0000:15:00.1-fc-0x23008cdf9dcaf0ea-lun-0
=====
```

② I/O パスの物理ステータス

③ I/O パスの物理ステータス

① I/O パスの組み込み状態

- ① 2 つ以上の I/O パスがともに組み込み済(extended)であることが前提です。
- ② 、③ 組み込み済(extended)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

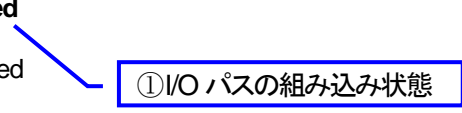
① 閉塞対象の I/O パスの udev

```
# /opt/HA/RDM/bin/rdmreduce -p pci-0000:15:00.0-fc-0x22008cdf9dcaf0ea-lun-0
```

- ① 指定した I/O パスを閉塞します。

- 状態確認

```
# /opt/HA/RDM/bin/rmdadmin
(monitor status = TRUE)
=====
type      : Logical      : I/O      : Online
           : H/W Path      : status   : status   : status
=====
VG        : /dev/rhel00      : up
MIRROR    :
PV        : /dev/sdc         : up       : up       : reduced
           : pci-0000:15:00.0-fc-0x22008cdf9dcaf0ea-lun-0
PV        : /dev/sdd         : up       : up       : extended
           : pci-0000:15:00.0-fc-0x24008cdf9dcaf0ea-lun-0
PV        : /dev/sde         : up       : up       : extended
           : pci-0000:15:00.1-fc-0x21008cdf9dcaf0ea-lun-0
PV        : /dev/sdf         : up       : up       : extended
           : pci-0000:15:00.1-fc-0x23008cdf9dcaf0ea-lun-0
```



① I/O パスを閉塞すると Online status が reduced になります。

注意: 冗長化された I/O パスに関して、すべての I/O パスを閉塞することはできません。
閉塞を行う場合は、事前に、2 パス以上 I/O パスが組み込まれていることを確認してください。

(2) rdmrecover 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/RDM/bin/rdmadmin
(monitor status = TRUE)
=====
① I/O パスの組み込み状態
=====
Logical   : I/O           : Online
type      : H/W Path      : status   : status   : status
=====
VG        : /dev/rhel00    : up
MIRROR    :                :
PV        : /dev/sdc       : up       : up       : reduced
          : pci-0000:15:00.0-fc-0x22008cdf9dcaf0ea-lun-0
PV        : /dev/sdd       : up       : up       : reduced
          : pci-0000:15:00.0-fc-0x24008cdf9dcaf0ea-lun-0
PV        : /dev/sde       : up       : up       : extended
          : pci-0000:15:00.1-fc-0x21008cdf9dcaf0ea-lun-0
PV        : /dev/sdf       : up       : up       : extended
          : pci-0000:15:00.1-fc-0x23008cdf9dcaf0ea-lun-0
```

① I/O パスが閉塞されています。

- 復旧実行

```
# /opt/HA/RDM/bin/rdmrecover -v
HW-path: pci-0000:15:00.0-fc-0x22008cdf9dcaf0ea-lun-0 result: OK
HW-path: pci-0000:15:00.0-fc-0x24008cdf9dcaf0ea-lun-0 result: OK
HW-path: pci-0000:15:00.1-fc-0x21008cdf9dcaf0ea-lun-0 result: OK
HW-path: pci-0000:15:00.1-fc-0x23008cdf9dcaf0ea-lun-0 result: OK
rdmrecover 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/RDM/bin/rdmadmin
(monitor status = TRUE)
=====
① I/O パスの組み込み状態
=====
Logical   : I/O           : Online
type      : H/W Path      : status   : status   : status
=====
VG        : /dev/rhel00    : up
MIRROR    :                :
PV        : /dev/sdc       : up       : up       : extended
          : pci-0000:15:00.0-fc-0x22008cdf9dcaf0ea-lun-0
PV        : /dev/sdd       : up       : up       : extended
          : pci-0000:15:00.0-fc-0x24008cdf9dcaf0ea-lun-0
PV        : /dev/sde       : up       : up       : extended
          : pci-0000:15:00.1-fc-0x21008cdf9dcaf0ea-lun-0
PV        : /dev/sdf       : up       : up       : extended
          : pci-0000:15:00.1-fc-0x23008cdf9dcaf0ea-lun-0
```

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

4.7. 障害の検知、および、アクション実行の高速化

CLUSTERPRO MC RootDiskMonitor 2.7 for Linux より、従来に比べ早期に障害を検知し、アクションを実行することが可能となりました。高速化の設定を行う場合、次の2通りの方法があります。

- ①監視間隔を短く設定することにより、障害を素早く検出する方式
- ②障害条件を変更することにより、致命的なエラーによる障害が発生した場合、障害を素早く検出する方式

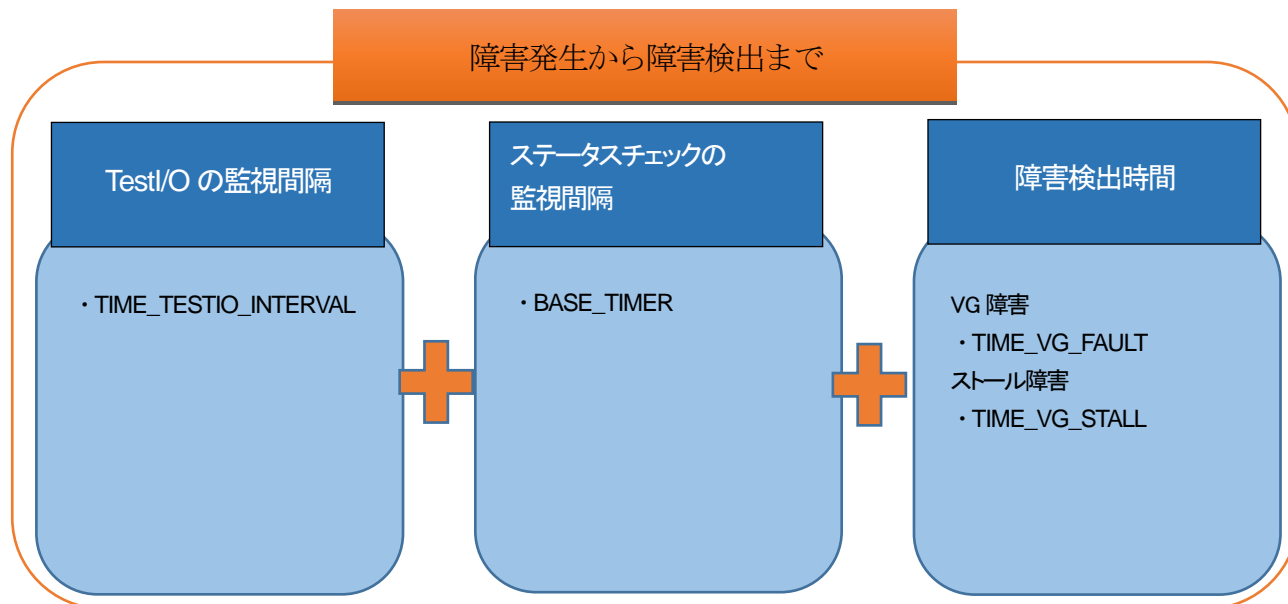
各方式の説明と設定手順は後述の章を参照してください。

4.7.1. 監視間隔のチューニングについて

障害検出に関する監視間隔の下限値を引き下げ、監視間隔をより短くすることで従来に比べ、早期に障害を検出することが可能となりました。

障害の監視間隔と障害検出時間は、システム定義ファイル(rdm.config)で設定されているパラメーターに応じて変化します。

障害発生から障害検出までの時間は次のような計算式で算出可能となります。パラメーターの詳細については3.2章を参照してください。



注意:

- ・ 障害発生から障害検出までの時間はアクションの実行を含みません。アクションの実行時間は別途加算されます。

4.7.2. 障害条件の変更について

従来は TestI/O の結果を正常・異常の 2 種類で取り扱っており、回復可能な障害か、回復不可能な恒久障害なのか継続して障害状態が続いているかで障害の判定を行っていました。

CLUSTERPRO MC RootDiskMonitor 2.7 for Linux より、TESTI/O の結果を正常・異常・致命的なエラーの 3 種類で取り扱い、致命的なエラーが発生した場合は、別途指定したリトライ回数内に致命的なエラーが継続した場合に、即座に障害と判定、アクションを実行する機能を追加しました。致命的なエラーが発生した場合、従来より早く障害を検出することが可能となります。機能の詳細については 2.5 章を参照してください。また、本機能は 4.7.1 章の監視間隔のチューニングを併用可能です。

本機能を利用するには、以下のパラメーターを変更し、RootDiskMonitor を再起動してください。設定手順の詳細は 4.7.3 章を参照してください。

パラメーター	説明
TESTIO_PERFORMANCE_TYPE	TestI/O の結果から致命的なエラーの検出可否を指定します。 本パラメーターを HIGH に設定することにより、致命的なエラーを検知した場合、即座に障害を検知することができます。 指定可能な値 HIGH NORMAL デフォルト値は NORMAL です。
FATAL_ERROR_RETRY_COUNT	TESTI/O の結果が致命的なエラーの場合に障害判定までのリトライ回数を指定します。 最小値は 1 回 デフォルト値は 2 回です。

注意:

- 致命的なエラーを検出できる障害は VG 障害のみとなります。
TESTI/O の結果が返却されないストール障害につきましては対象外となります。
- TESTIO_MODE が READ に設定されている場合は、致命的なエラーを検出できません。

4.7.3. 障害の検知、および、アクション実行の高速化の設定手順

設定方法は3パターン存在します。監視要件に合わせて選択してください。

A:監視間隔のチューニング

すべての障害に対して検出時間を短縮します。

B:障害条件の変更

従来の障害検知ロジックに加え、より詳細に Test I/O のエラー内容が判別されます。

致命的なエラーによる障害が発生した場合の障害検知時間を短くしたい場合に設定してください。

C:監視間隔のチューニングと障害条件の変更

AとBを併用する設定です。

すべての障害に対して検出時間を短縮し、致命的なエラーが発生した場合も検出時間を短縮します。

パターンごとの障害検出時間は以下のとおりです。

なお、ここで示す障害検出時間は想定される最長の時間となります。

	デフォルト	A	B	C
VG 障害	75 秒	8 秒	75 秒	8 秒
VG 障害 (致命的なエラー)	-	-	20 秒	3 秒
ストール障害	375 秒	8 秒	375 秒	8 秒

障害の検知、および、アクション実行の高速化の設定手順は、次のようになります。

(1) デーモンプロセスの停止

設定ファイルを編集する前にデーモンプロセスを停止してください。

【Red Hat Enterprise Linux 7.0 以降、Oracle Linux 7.0 以降、Amazon Linux 2 以降の場合】

```
# systemctl stop rdmd
```

【Red Hat Enterprise Linux 6.x、Oracle Linux 6.x の場合】

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

(2) デーモンプロセスの確認

デーモンプロセスが停止されていることを確認してください。

以下のコマンド実行後、rdmdiagd、rdmping が停止していることを確認してください。

```
# ps -ef|grep rdm
```

(3) 設定ファイル(rdm.config)のバックアップを取得

設定ファイルのバックアップを取得してください。
設定ファイル名は以下のとおりです。

/opt/HA/RDM/conf/rdm.config

(4) 設定ファイルの修正

A) 監視間隔のチューニング

次のパラメーターを設定ファイル(rdm.config)で書き換えてください。

TIME_VG_FAULT	VG 障害判定時間
TIME_TESTIO_INTERVAL	Test I/O 発行間隔
BASE_TIMER	基本タイマー
TIME_VG_STALL	I/O ストール障害判定時間

また、SAN-Boot 構成の場合で、TESTIO_FAULT_ACTION に BLOCK_PATH を設定している場合は、以下のパラメーターを設定ファイルに追記し、パラメーター値を1としてください。

AUTO_REDUCE_INTERVAL 自動閉塞実行間隔

下記の設定例では VG 障害の検知時間が最大でも 8 秒以内になります。

	パラメーター名	デフォルト値	設定値
①	TIME_VG_FAULT	60	6
②	TIME_TESTIO_INTERVAL	5	1
③	BASE_TIMER	10	1
④	TIME_VG_STALL	360	6

また、障害の障害検知時間は、次の計算式で計算可能です。
計算式をもとに、監視要件に合わせて、パラメーターを調整してください。
※パラメーターを調整して障害検知時間の短縮を行った場合は、調整後の設定値によっては障害の誤検知が発生する可能性があります。
パラメーター調整に関しては、十分な検証を行ったうえで実施してください。

【VG 障害】

BASE_TIMER + TIME_TESTIO_INTERVAL + TIME_VG_FAULT

【I/O ストール障害】

BASE_TIMER + TIME_TESTIO_INTERVAL + TIME_VG_STALL

以下はシステム定義ファイルの設定例となります。

内蔵ディスクの場合

```
# Test/I/O interval timer for Root Volume is failed (seconds)
#       Root Volume status changes fail between this timer
# minimum = 6, default = 60
TIME_VG_FAULT      6
```

① VG 障害検出時間

```
# I/O stall interval timer for Root Volume is failed (seconds)
#       Root Volume status changes fail between this timer
# minimum = 6, default = 360. 0 mean I/O stall nocheck.
TIME_VG_STALL     6
```

④ I/O ストール障害検出時間

```
# Test/I/O interval timer value (seconds)
#       exec normal Test/I/O for PV between this timer
# minimum = 1, default = 5
TIME_TESTIO_INTERVAL  1
```

② Test I/O 発行間隔

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

③ 基本タイマー

```
# status check timer (seconds)
# minimum = 1, default = 10
BASE_TIMER         1
```

SAN-BOOT 構成の場合

```

# Test/I/O interval timer for Root Volume is failed (seconds)
#       Root Volume status changes fail between this timer
# minimum = 6, default = 60
TIME_VG_FAULT           6

# I/O stall interval timer for Root Volume is failed (seconds)
#       Root Volume status changes fail between this timer
# minimum = 6, default = 360. 0 mean I/O stall nocheck
TIME_VG_STALL           6

# Test/I/O interval timer value (seconds)
#       exec normal Test/I/O for PV between this timer
# minimum = 1, default = 5
TIME_TESTIO_INTERVAL    1

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

# status check timer (seconds)
# minimum = 1, default = 10
BASE_TIMER              1

~~~~ 中略 ~~~~

#####
# For SANBOOT
#####
# boot type
# select LOCAL_DISK(default),SAN_BOOT
BOOT_TYPE                SAN_BOOT

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

# multipath type
# select NO_USE(default),DM_MULTIPATH,POWER_PATH,SPS,HDLM,DMP
MULTIPATH_TYPE           SPS

AUTO_REDUCE_INTERVAL     1

```

① VG 障害検出時間

④ I/O ストール障害検出時間

② Test I/O 発行間隔

③ 基本タイマー

※TESTIO_FAULT_ACTION に、BLOCK_PATH を
設定している場合に、追記してください。

B) 障害条件の変更

次のパラメータを設定ファイル(srg.config)で書き換えてください。
 下記の設定例では VG 障害の検知時間が、最大でも 20 秒以内になります。

	パラメーター名	デフォルト値	設定値
⑤	TESTIO_PERFORMANCE_TYPE	NORMAL	HIGH
⑥	FATAL_ERROR_RETRY_COUNT	2	2(※1)

※1 FATAL_ERROR_RETRY_COUNT に関しては監視要件に応じて変更してください。

また、致命的なエラーによる障害の障害検知時間は、次の計算式で計算可能です。
 計算式をもとに、監視要件に合わせて、パラメータを調整してください。
 ※パラメータを調整して障害検知時間の短縮を行った場合は、
 調整後の設定値によっては障害の誤検知が発生する可能性があります。
 パラメータ調整に関しては、十分な検証を行ったうえで実施してください。

【VG 障害】

$BASE_TIMER + (TIME_TESTIO_INTERVAL \times FATAL_ERROR_RETRY_COUNT)$

以下はシステム定義ファイルの設定例となります。

# Test I/O performance type # select NORMAL(default),HIGH TESTIO_PERFORMANCE_TYPE	HIGH	⑤ Test I/O の結果の分類種別
# Fatal error retry count # minimum = 1, default = 2 FATAL_ERROR_RETRY_COUNT	2	⑥ 致命的なエラーの際のリトライ回数

C) 監視間隔のチューニングと障害条件の変更

次のパラメーターを設定ファイル(rdm.config)で書き換えてください。

下記の設定例では VG 障害の検知時間が最大でも 8 秒以内、
致命的なエラーによる障害が発生した場合は最大でも 3 秒以内になります。

	パラメーター名	デフォルト値	設定値
①	TIME_VG_FAULT	60	6
②	TIME_TESTIO_INTERVAL	5	1
③	BASE_TIMER	10	1
④	TIME_VG_STALL	360	6
⑤	TESTIO_PERFORMANCE_TYPE	NORMAL	HIGH
⑥	FATAL_ERROR_RETRY_COUNT	2	2(※1)

※障害検出時間の算出方法は A,B の設定手順を参照ください。

以下はシステム定義ファイルの設定例となります。

Test I/O interval timer for Root Volume is failed (seconds)

Root Volume status changes fail between this timer

minimum = 6, default = 60

TIME_VG_FAULT

6

① VG 障害検出時間

I/O stall interval timer for Root Volume is failed (seconds)

Root Volume status changes fail between this timer

minimum = 6, default = 360. 0 mean I/O stall nocheck.

TIME_VG_STALL

6

④ I/O ストール障害検出時間

Test I/O interval timer value (seconds)

exec normal Test I/O for PV between this timer

minimum = 1, default = 5

TIME_TESTIO_INTERVAL

1

② Test I/O 発行間隔

Wait I/O for sg driver timer value (seconds)

wait Test I/O between this timer

minimum = 1, default = 5

WAIT_TESTIO_INTERVAL

5

status check timer (seconds)

minimum = 1, default = 10

BASE_TIMER

1

③ 基本タイマー

... 中略 ...

Test I/O performance type

select NORMAL(default),HIGH

TESTIO_PERFORMANCE_TYPE

HIGH

⑤ Test I/O の結果の分類種別

Fatal error retry count

minimum = 1, default = 2

FATAL_ERROR_RETRY_COUNT

2

⑥ 致命的なエラーの際のリトライ回数

(6) デーモンプロセスの再開

デーモンプロセスを起動してください。

【Red Hat Enterprise Linux 7.0 以降、Oracle Linux 7.0 以降、Amazon Linux 2 以降の場合】

```
# systemctl start rdmd
```

【Red Hat Enterprise Linux 6.x、Oracle Linux 6.x の場合】

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

(7) デーモンプロセスの確認

デーモンプロセスが開始されていることを確認してください。

以下のコマンド実行後、rdmdiagd、rdmping が起動していることを確認してください。

```
# ps -ef | grep rdm
root      2211      1  0 Jan04 ?        00:00:25 /opt/HA/RDM/bin/rdmdiagd
root      2367     2211  0 Jan04 ?        00:01:43 rdmping
root      20366   20345  0 13:40 pts/7    00:00:00 grep rdm
```

(8) コンフィグレーション情報を確認します。

(5)で設定したパラメーターが反映されているか確認してください。

以下は設定例「(C) 監視間隔のチューニングと障害条件の変更」で設定されたパラメーターが正しく反映された例です。

内蔵ディスクの場合

```
# /opt/HA/RDM/bin/rdmadmin -c param
SG parameters.
-----
TIME_VG_FAULT 6
TIME_VG_STALL 6
TIME_TESTIO_INTERVAL 1
WAIT_TESTIO_INTERVAL 5
BASE_TIMER 1
OVER_ACTION SERVICE_CMD_DISABLE
VG_STALL_ACTION SERVICE_CMD_DISABLE
TUR_TESTIO_USE TRUE
TESTIO_DIRECT READ
testioYN TRUE
TESTIO_MODE READ
TESTIO_PERFORMANCE_TYPE HIGH
FATAL_ERROR_RETRY_COUNT 2
BOOT_TYPE LOCAL_DISK
TESTIO_FAULT_ACTION ACTION_NONE
MULTIPATH_TYPE NO_USE
AUTO_REDUCE_INTERVAL 5
```

SAN-BOOT 構成の場合

```
# /opt/HA/RDM/bin/rdmadmin -c param
SG parameters.
-----
TIME_VG_FAULT 6
TIME_VG_STALL 6
TIME_TESTIO_INTERVAL 1
WAIT_TESTIO_INTERVAL 5
BASE_TIMER 1
OVER_ACTION SERVICE_CMD_DISABLE
VG_STALL_ACTION SERVICE_CMD_DISABLE
TUR_TESTIO_USE TRUE
TESTIO_DIRECT READ
testioYN TRUE
TESTIO_MODE READ
TESTIO_PERFORMANCE_TYPE HIGH
FATAL_ERROR_RETRY_COUNT 2
BOOT_TYPE SAN_BOOT
TESTIO_FAULT_ACTION BLOCK_PATH
MULTIPATH_TYPE SPS
AUTO_REDUCE_INTERVAL 1
```

4.8. 監視ディスクのデバイスファイル再解決手順

FC ケーブルの抜線などにより、OS からディスクの認識が一時的に外れた後、ディスクが再認識された際にデバイスファイルが変更される可能性があります。

デバイスファイルの変更が行われた場合、RootDiskMonitor 内のデバイス情報を更新する必要があります。

4.8.1. デバイスファイルのチェックについて

何らかの契機でデバイスファイルの再割り当てが行われると、監視に使用している RootDiskMonitor の共有メモリ内のデバイスファイルと実際のデバイスファイルにずれが生じ、監視が出来なくなる場合があります。このデバイスファイルのずれをチェックします。

デバイスファイルのチェックは `rdmadmin -c devcheck` コマンドで行います。なお、`rdmadmin -c devcheck` コマンドを OS の日時指定ジョブ実行デーモン(cron)に登録することで、自動でチェックを行うこともできます。

(1) 自動チェックについて

デバイス情報の自動チェックは OS の日時指定ジョブ実行デーモン(cron)を利用して実現します。

設定手順については cron の設定方式に準拠します。cron についての詳細は cron のオンラインマニュアルを参照してください。

デバイス情報のチェックコマンドと定期実行する時刻を cron の設定ファイル(crontab ファイル)に登録します。

【cron の設定手順】

①設定したいスケジュールから cron の設定ファイル(crontab ファイル)に登録する文字列を準備します。crontab へのスケジュール登録方式は以下のとおりです。

```
minute hour monthday month weekday command
```

各フィールドに設定したい時刻や日付を整数値で指定します。
また "*" (アスタリスク) を指定すると、すべての意味となります。
複数指定したい場合は "," (コンマ) で区切って列記します。

以下に設定例を示します。

例 毎日 10:00 にデバイス情報の自動チェックを行う場合

```
0 10 * * * /opt/HA/RDM/bin/rdmadmin -c devcheck -S
```

②crontab コマンドを使って、①で作成した登録文字列を cron の設定ファイル(crontab ファイル)へ登録します。

以下のコマンドを実行し、最後の行に登録文字列を入力してください。

```
#crontab -e
```

③正しく登録されていることを確認します。

確認は crontab コマンドに -l オプションを指定して実行します。登録した文字列が出力されれば登録は完了です。

```
#crontab -l  
0 10 * * * /opt/HA/RDM/bin/rdmadmin -c devcheck -S
```

前述の手順で cron にデバイス情報のチェックコマンドを登録後、指定した時刻でデバイス情報のチェックが自動的に行われます。

デバイスファイルの変更が行われていた場合、出力されるメッセージは、次の通りです。

```
Special file changed [hwpath=pci-0000:13:00.1-fc0x2100001697120ca7:0x0001000000000000  
s.f =sdh (from sdg)].
```

このメッセージが表示された場合、デバイスファイルの更新を行ってください。
更新手順は"4.8.2. デバイスファイルの更新について"を参照してください。

また、RootDiskMonitor 起動時に監視対象が存在していない場合は、下記メッセージが出力されます。
本メッセージが出力された場合、デーモンを再起動することにより、該当パスの監視を再開することができます。

```
unknown path may be recoverable HW-path=pci-0000:13:00.1-fc-0x2100001697120ca7:0x0001000000000000
```

(2) 手動チェックについて

デバイスファイルのチェックを手動で行うことができます。

手動でチェックを行う場合は rdmadmin -c devcheck コマンドを実行します。

```
# /opt/HA/RDM/bin/rdmadmin -c devcheck
```

チェックに成功した場合、「Device information check complete.」と出力されます。なお、チェックに失敗した場合は「Device information check fail.」と出力されます。サポートセンターに連絡して下さい。

デバイスファイルの変更が行われていた場合、出力されるメッセージは、次の通りです。

```
Special file changed [hwpath=pci-0000:13:00.1-fc0x2100001697120ca7:0x0001000000000000  
s.f =sdh (from sdg)].
```

このメッセージが表示された場合、デバイスファイルの更新を行ってください。
更新手順は"4.8.2. デバイスファイルの更新について"を参照してください。

また、RootDiskMonitor 起動時に監視対象が存在していない場合は、下記メッセージが出力されます。
本メッセージが出力された場合、デーモンを再起動することにより、該当パスの監視を再開することができます。

```
unknown path may be recoverable HW-path=pci-0000:13:00.1-fc-0x2100001697120ca7:0x0001000000000000
```

4.8.2. デバイスファイルの更新について

デバイスファイルの更新は構成復旧コマンド、もしくは、運用管理コマンドで行います。構成復旧コマンドを実行し、デバイス情報を更新してください。

```
# /opt/HA/RDM/bin/rdmrecover -v
```

構成復旧を行いたくない(パスの復旧を行いたくない)場合は構成復旧コマンドの代わりに運用管理コマンドを実行します。

```
# /opt/HA/RDM/bin/rdmadmin -c devsync
```

更新に成功した場合、「rdmrecover complete.」、もしくは、「Device information sync complete.」と出力されます。なお、チェックに失敗した場合は「rdmrecover fail.」、もしくは、「Device information sync fail.」と出力されます。サポートセンターに連絡して下さい。

デバイスファイルを更新した場合、出力されるメッセージは、次の通りです。

```
Update: HW-path=pci-0000:13:00.0-fc-0x2100001697120ca7:0x0000000000000000 s.f.=sdj
```

5. syslog メッセージ

5.1. 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 の facility と level は以下のとおりです。

facility: LOG_DAEMON

level : LOG_ALERT、LOG_ERROR 、 LOG_NOTICE

LOG_ALERT、LOG_ERROR は、ハードエラーなどの致命的な障害の場合に使用されます。

5.2. 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。(下線部はメッセージの固定部分を示します)
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) Test I/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 down find .(sf=/dev/sdx)

説明:Test I/O で PV レベルの異常を検出

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

シングル構成の場合は、その後 VG レベルも down になります。

ミラー構成の場合は、VG レベルが suspend になります。

I/O path reduced.(sf='/dev/sdx')

説明: I/O パスが閉塞されました。

処置: ディスク装置の点検を行ってください。

閉塞した I/O パスは RootDiskMonitor の構成復旧コマンドで復旧させてください。

VG status change suspend .(vg=/dev/VolGroupXX)

説明:Test I/O で VG レベルの異常(suspend)を検出

処置:ミラー構成の片系のディスク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。

VG status change down . (vg=/dev/VolGroupXX)

説明:Test I/O で VG レベルの異常(down)を検出

処置:ミラー構成の両系のディスク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。

VG status change up .(vg=/dev/VolGroupXX)

説明:Test I/O で VG レベルの復旧(up)を検出

処置:ディスクの障害から復旧しました。

ディスクの監視を開始していますので、特に必要ありません。

I/O stall find , timeover occurred (sf=/dev/sdx)

説明:Test I/O で I/O ストールタイムオーバーを検出

処置:ディスクが故障している可能性があります。

早急に該当ディスクの点検を行ってください。

Test I/O fatal error found. (errno = x, retry count = x, sf = /dev/sdx)

説明:Test I/O で致命的なエラーを検出

処置:ディスク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。

Special file changed [hwpath='ハードウェアパス名':

s.f = '変更後のスペシャルファイル名' (from '変更前のスペシャルファイル名')].

説明:監視対象のデバイスファイル名の変更を検知

処置:監視対象のデバイスファイルが変更されたため、変更された監視対象に対する監視結果が正しいものと保証できません。構成復旧コマンドまたは運用管理コマンドによるデバイス情報の更新を実行してください。

- (3) コンフィグレーション関係

LOG_ERROR の出力契機は以下のとおりです。

ConfigFile error(xxx).

説明:設定ファイルの設定誤りを検出

処置:設定ファイル(/opt/HA/RDM/conf/rdm.config)を確認してください。

記述例は『3.2 設定ファイルの記述』を参照してください。

- (4) システムメモリダンプ採取と OS 強制停止による CLUSTERPRO 連携

LOG_ALERT の出力契機は以下のとおりです。

start kernel system dump.

説明:システムメモリダンプの採取と OS 強制停止を開始

処置:内蔵ディスクの障害により、クラスター系切り替えが発生しています。

早急に該当ディスクの点検を行ってください。

- (5) CLUSTERPRO のサーバー管理プロセス (clpnm) 強制終了による CLUSTERPRO 連携

LOG_ALERT の出力契機は以下のとおりです。

start KILL clpnm.

説明:CLUSTERPRO のサーバー管理プロセス (clpnm) 強制終了を開始

処置:内蔵ディスクの障害により、クラスター系切り替えが発生しています。

早急に該当ディスクの点検を行ってください。

send signal clpnm.

説明:CLUSTERPRO のサーバー管理プロセス (clpnm) 強制終了を完了

処置:内蔵ディスクの障害により、クラスター系切り替えが発生しています。

早急に該当ディスクの点検を行ってください。

abort rdmdiagd.

説明:rdmdiagd が abort しました

処置:内蔵ディスクの障害により、クラスター系切り替えが発生しています。

早急に該当ディスクの点検を行ってください。

5.3. 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
RootDiskMonitor 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
RootDiskMonitor 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。

6. 注意・制限事項について

6.1. 注意・制限事項

(1) 下記の注意事項があります。

- 本製品を利用する場合、Linux SCSI パススルードライバー(sg ドライバー)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 ドライバーのエントリー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 ドライバーのエントリー行
```

- RootDiskMonitor 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。
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-w.x.y.z
```

- ・ 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:w, 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。

- 本製品を運用中には共有メモリを約 1MB 程度使用します。
- ログ用のディレクトリ(/opt/HA/RDM/log)配下に、トレースファイルや core ファイルを保存するために、約 3MB 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、3MB を超えることはありません。
※SAN-Boot 構成で、自動閉塞機能を使用する場合は、自動閉塞用のトレースログが最大 4MB 出力されます。その場合は、使用する領域は最大で約 7MB となります。
- 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。
- アンインストール時に /opt/HA ディレクトリは削除されません。不要な場合、手動で削除してください。
- CLUSTERPRO を導入し、クラスター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、RootDiskMonitor の POWER_OFF 機能を利用することができません。
どちらも同等の機能のため、CLUSTERPRO の softdog 機能を利用してください。

- ミラー元の Lvol が複数の PV にまたがっている構成の場合は、設定ファイルの自動生成の対象外となります。このような構成の場合には手動で設定ファイルを作成してください。
- Device Mapper を使用したソフトミラー構成の場合は、物理ディスク構成のみサポートし、LVM を使用した構成は設定ファイル自動生成 (rdmconfig) のサポート対象外です。手動で設定ファイルを編集してください。
- 仮想環境 (ゲスト OS) において本製品を使用する場合は、設定ファイルの TESTIO_MODE に READ を設定する必要があります。
- StorageSaver for BootDisk の場合は、サーバー搭載の内蔵ディスクの監視には使用できません。
- Hyper-V または KVM のゲスト OS において本製品を使用する場合は、設定ファイルに udev デバイスファイル名を定義することができません。設定ファイルを自動生成する際、以下のコマンドを実行してください。

```
# /opt/HARDM/bin/rdmconfig -sd
```

- コードワードが未登録の状態インストールした場合でも、インストールから 30 日までは本製品のすべての機能を通常どおり使用できます。なお、インストールから 30 日を経過してもコードワードが未登録の場合には RootDiskMonitor の機能に制限がかかり、障害発生時でも障害を検知なくなります。詳細は「4.5. 機能制限について」を参照してください。
- Red Hat Enterprise Linux 7.0 以降、Oracle Linux 7.0 以降、Amazon Linux 2 以降において、RootDiskMonitor のプロセスが一部残存していた場合には、start コマンドではサービスは起動できません。そのため、RootDiskMonitor を再起動する場合には、start コマンドではなく、restart コマンドを使用してください。
- 障害発生から障害検出までの時間はアクションの実行を含みません。アクションの実行時間は別途加算されます。
- 致命的なエラーを検出できる障害は VG 障害のみとなります。TESTI/O の結果が返却されないストール障害につきましては対象外となります。
- TESTIO_MODE が READ に設定されている場合は、致命的なエラーを検出できません。
- I/O パス障害が発生した際、障害の種類によっては OS が障害の発生した I/O パスの sd デバイスを削除する場合があります。物理環境の SANBoot 構成で、TESTIO_FAULT_ACTION に BLOCK_PATH を設定している場合、I/O パスの障害を検出後、自動閉塞処理を実行する前に、OS による対象の sd デバイスの削除が行われた場合は、I/O パスの閉塞が行われず、Online status が extended のままとなる場合があります。上記の場合でも、OS 側で障害の発生した I/O パスの sd デバイスの削除が行われているため、障害が発生したパスで I/O が行われることはなく、運用上は問題ありません。

(2) 下記の制限事項があります。

- LVM で構築した OS ディスクについては、OS ディスクが設定ファイル自動生成の対象になります。
それ以外の VG については、手動で設定ファイル情報を設定してください。

また、設定ファイル自動生成が可能な構成は、LVM および物理ディスク構成で、ディスク 1 つによるシングル構成、およびディスク 2 つによるソフトミラー構成の場合のみとなります。

ディスク 4 つでのソフトミラー構成や OS ディスク以外に同一 VG 内にデータディスクを有する構成の場合などでは、正しく設定ファイル自動生成できない場合があります。このような構成の場合には手動で設定ファイルを作成してください。

- FC 接続のディスクアレイ装置を OS ディスクとして監視する場合、VG が活性化されていない、またはファイルシステムとしてマウントされていないと、正常に動作しない場合があります。必ず、ファイルシステムとしてマウントした状態で監視してください。

7. リファレンス

rdmadmin

名称

rdmadmin - OSディスク監視モニターの制御

構文

```
/opt/HA/RDM/bin/rdmadmin [-h]
                             [-s]
                             [-c param]
                             [-c status [-f file] [-t time] [-s]]
                             [-c pdstatus [-s]]
                             [-c start]
                             [-c stop]
                             [-c trace [-l diag|ping]]
                             [-c devcheck]
                             [-c devsync]
                             [-L]
```

説明

rdmadmin コマンドは、OSディスク監視モニター (RootDiskMonitor)を制御するコマンドです。パラメーターの表示やOSディスク監視の開始/停止、OSディスク監視の設定ファイルの生成等を行います。オプションを省略した場合、-c status を指定した場合と同様の情報を表示します。

オプション

-h

コマンドの説明を表示します。

-s

OSディスクの状態とS.M.A.R.T.診断結果を表示します。

-c param

OSディスク監視モニターのパラメーター、監視リソースの一覧を表示します。

-c status [-f file] [-t time] [-s]

OSディスクの状態を表示します。

-f オプションにスペシャル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。

-f オプションを省略すると、すべてのOSディスクの状態を表示します。

-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 を実行します。

-t オプションを省略すると、rdmadmin は status を一度だけ表示して終了します。

-s オプションを指定すると、OSディスクの状態とS.M.A.R.T.診断結果を表示します。

-c start

OSディスクの監視を開始します。

-c stop

OSディスクの監視を停止します。

-c trace [-l diag|ping]

OSディスク監視コマンドの内部トレースを、標準出力に表示します。

-l diag は、rdmdiagd コマンドのトレースを出力します。

-l ping は、rdmping コマンドのトレースを出力します。

-l オプションを省略すると、rdmdiagd と rdmping の両方のコマンドのトレースを出力します。

-c pdstatus

監視対象ディスクがHW-RAID構成の場合に、論理ディスクおよび物理ディスクの状態を表示します。

-s オプションを指定すると、物理ディスクの状態とS.M.A.R.T.診断結果を表示します。

-c devcheck

I/Oパスに紐づくデバイスファイルが最新の状態かチェックし、画面上にチェック結果を出力します。

-c pdstatus

I/Oパスに紐づくデバイスファイルを最新の状態に更新します。

-L

現在登録されているパスワードを確認して反映します。

使用例

・すべてのOSディスクの状態を表示します。

```
# /opt/HA/RDM/bin/rdmadmin
```

または

```
# /opt/HA/RDM/bin/rdmadmin -c status
```

・OSディスク監視モニターのパラメーターを表示します。

```
# /opt/HA/RDM/bin/rdmadmin -c param
```

・すべてのOSディスクの状態を30秒間隔で表示します。

```
# /opt/HA/RDM/bin/rdmadmin -c status -t 30
```

・HW-RAID構成の各ディスクの状態を表示します。

```
# /opt/HA/RDM/bin/rdmadmin -c pdstatus
```

関連項目

rdmconfig

関連ファイル

/opt/HA/RDM/bin/rdmadmin

rdmadmin コマンドの標準のパス。

rdmstat

名称

rdmstat - OSディスク カスタムモニタリソース

構文

/opt/HA/RDM/bin/rdmstat [-h] [-v] [-V vgname] [-w wait-time]

機能説明

rdmstat は、OSディスク監視モニター(Root Disk Monitor)がレポートするリソースステータスを監視するコマンドです。rdmstat 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、RootDiskMonitor 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、rdmstat は常駐して RootDiskMonitor 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なパラメータです。

RootDiskMonitor のプロセスが動作していない場合、また、rdmadminでリソース監視の停止を指示した場合は、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 RootDiskMonitor がレポートするステータスを調べ、終了します。

rdmstat は、-w オプションと RootDiskMonitor 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 オプションありの場合:

RootDiskMonitor が監視するディスクの状態	rdmstat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
RootDiskMonitor が動作していない	終了せずに常駐	---

-w オプションなしの場合:

RootDiskMonitor が監視するディスクの状態	rdmstat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
RootDiskMonitor が動作していない	終了	2 を返す

オプション

- 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 vgroupname
監視したい VG 名を指定します。
-V を指定した場合は、対象 VG 名を1つだけ指定できます。
-V を指定しない場合は、RootDiskMonitor が監視するすべての 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1つのVG で DOWN が検出された時点で、rdmstat は「RootDiskMonitor 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 に補整されます。

補足事項

本コマンドはメモリ上に常駐します。swap 領域に待避されません。
RootDiskMonitor のプロセスを起動していない状態、rdmadm 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。

使用例

- ヘルプを表示します。
/opt/HA/RDM/bin/rdmstat -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、RootDiskMonitor が監視するすべての VG です。
/opt/HA/RDM/bin/rdmstat -v

- /dev/VolGroup00 のみを監視対象とします。
/opt/HA/RDM/bin/rdmstat -V /dev/VolGroup00

- /dev/VolGroup01 で異常が発生するまで、rdmstat を常駐させます。
このとき、RootDiskMonitor のステータスを 5 秒間隔で調べます。
/opt/HA/RDM/bin/rdmstat -V /dev/VolGroup01 -w 5

rdmconfig

名称

rdmconfig - OSディスク監視モニターの設定ファイルテンプレートの自動生成

構文

```
/opt/HA/RDM/bin/rdmconfig [-sd] [-S dmmultipath | sps | powerpath | hdlm | dmp | nouse]
```

説明

rdmconfig は、OS ディスク監視モニター (RootDiskMonitor) の設定ファイルのテンプレートを自動生成します。本コマンドで設定ファイルを作成した後に、監視ポリシーの設定、監視リソースの妥当性を確認してください。

rdmconfig を実行すると、/opt/HA/RDM/conf/rdm.config.default を元に /opt/HA/RDM/conf/rdm.config を生成します。

コマンド実行前に既に /opt/HA/RDM/conf/rdm.config ファイルが存在する場合は、上書き要否の問い合わせを行います。上書きした場合、元の設定ファイルは/opt/HA/RDM/conf/rdm.config.old として保存されます。

オプション

-sd

設定ファイルに定義する PV 情報をブロックデバイスファイル名 (/dev/sda等) で定義します。

-S dmmultipath | sps | powerpath | hdlm | dmp | nouse

SAN-Boot環境で設定ファイルを生成する場合に指定します。

SAN-Boot構成でない場合は、指定する必要はありません。

-Sの後には、使用しているマルチパス管理ソフトウェアを指定します。

dmmultipath Device Mapper Multipath を使用している場合に指定します。

sps StoragePathSavior を使用している場合に指定します。

powerpath EMC powerpath を使用している場合に指定します。

hdlm HitachiDynamicLinkManager または、HADynamicLinkManagerを使用している場合に指定します。

dmp Veritas DynamicMultiPassing を使用している場合に指定します。

nouse マルチパス管理ソフトウェアを使用していない場合に指定します。

注意事項

- ・ 本コマンドで作成した設定ファイルは OS ディスクの構築状況によっては監視対象リソースの修正が必要です。特に、LVM で規定したミラーを構成する LUN(PV) の組み合わせについては、妥当性を確認してください。
- ・ OS ディスク以外については自動生成対象とはなりませんので、手動で設定ファイルを編集してください。
- ・ ミラー元のvolが複数のPVIにまたがっている構成は設定ファイル自動生成 (rdmconfig) のサポート対象外です。手動で設定ファイルを編集してください。
- ・ Device Mapperを使用したソフトミラー構成の場合は、物理ディスク構成のみサポートし、LVMを使用した構成は設定ファイル自動生成 (rdmconfig) のサポート対象外です。手動で設定ファイルを編集してください。

- ・ デバイスファイルのチェック機能は、設定ファイル自動生成 (rdmconfig) のサポート対象外です。本機能を使用する場合は、" 3.2 設定ファイルの記述 (2) 設定ファイルの記述例について"を参照し、手動で設定ファイルを編集してください。
- ・ iStorage Vシリーズでは、使用するマルチパス管理ソフトウェアは、従来のStoragePathSaviorではなく HA Dynamic Link Managerとなります。そのため、SAN-Boot環境でiStorage Vシリーズをご利用の場合に設定ファイルの自動生成を行う際は、-Sオプションの引数として "hdlm" を指定してください。

関連ファイル

/opt/HA/RDM/bin/rdmadmin

rdmadmin コマンドの標準のパス。

/opt/HA/RDM/conf/rdm.config

rdmconfigで生成される、RootDiskMonitor の設定ファイル。

/opt/HA/RDM/conf/rdm.config.default

rdm.config を生成する際に、元となるファイル。

rdmreduce

名称

rdmreduce - RootDiskMonitorの監視リソースの閉塞コマンド

構文

```
/opt/HA/RDM/bin/rdmreduce -p udev_devicefile
```

説明

rdmreduce は、指定されたudevデバイスファイル名に対応するI/Oパスを閉塞します。
SAN-Boot構成でのみ使用可能です。

オプション

-p udev_devicefile
udevデバイスファイル名を指定します。
指定された I/O パスが閉塞の対象となります。

注意事項

- ・ 本コマンドは、RootDiskMonitorが起動している場合のみ、使用することができます。
- ・ 設定ファイル(rdm.config)の「BOOT_TYPE」が「SAN_BOOT」となっている場合のみ使用可能です。

関連ファイル

/opt/HA/RDM/bin/rdmadmin
RootDiskMonitor の運用管理コマンド
/opt/HA/RDM/bin/rdmrecover
RootDiskMonitor の構成復旧コマンド。
/opt/HA/RDM/conf/rdm.config
RootDiskMonitor の設定ファイル。

rdmrecover

名称

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

構文

```
/opt/HA/RDM/bin/rdmrecover [-v] [-p udev_devicefile]
```

説明

rdmrecover は、指定したudevデバイス名に対応したI/Oパスを復旧します。
udevデバイス名を指定しない場合、すべてのI/Oパスを復旧します。
SAN-Boot構成でのみ使用可能です。

復旧されたことは rdmadmin 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

-v

I/Oパスの復旧結果をコンソールに表示します。

-p udev_devicefile

udev デバイスファイル名を指定します。

指定された I/O パスが復旧の対象となります。

注意事項

- ・ 本コマンドは、RootDiskMonitorが起動している場合のみ、使用することができます。
- ・ 設定ファイル(rdm.config)の「BOOT_TYPE」が「SAN_BOOT」となっている場合のみ使用可能です。

関連ファイル

/opt/HA/RDM/bin/rdmadmadmin

RootDiskMonitor の運用管理コマンド。

/opt/HA/RDM/bin/rdmreduce

RootDiskMonitor の手動閉塞コマンド。

/opt/HA/RDM/conf/rdm.config

RootDiskMonitor の設定ファイル。

8. 付録

8.1. 本製品のテスト手順について

■ はじめに

RootDiskMonitor 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。

○ RootDiskMonitor の評価

- ・ 物理ディスク故障(OS ディスク障害、クラスタウェア連携)
- ・ 致命的なエラーを伴う物理ディスク故障
- ・ SAN-Boot 構成の片系 I/O パス障害、両系 I/O パス障害
- ・ I/O ストール障害

※SAN-Boot 構成の場合でも、致命的なエラーを検出することは可能です。

疑似障害の評価手順については、前述の致命的なエラーを伴う物理ディスク故障の手順を参照ください。

■ 物理ディスク擬似障害

物理ディスクの障害には以下のパターンがあります。

- ・ 物理ディスクの障害
ソフトミラー構成の場合には片系障害を経て両系障害となります
- ・ 致命的なエラーを伴う物理ディスクの障害
ソフトミラー構成の場合には片系障害を経て両系障害となります
- ・ CLUSTERPRO 連携

物理ディスクの擬似的な障害発生の手順について説明します。

【通常の物理ディスク障害】

以下のコマンドにより、Test I/O の実行結果を擬似的に異常にすることで、監視ステータスを up から down に切り替えることや、VG ステータスを down 状態にすることができますので、評価を容易に行うことができます。

【コマンド書式】

```
/opt/HARDM/bin/rdmadmin -c debug -v [ on | off ] [ -f SpecialFile ]
```

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

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

【致命的なエラーを伴う物理ディスク障害】

以下のコマンドにより、Test I/O の実行結果を擬似的に致命的なエラーにすることで、監視ステータスを up から down に切り替えることや、VG ステータスを down 状態にすることができますので、評価を容易に行うことができます。

【コマンド書式】

```
/opt/HARDM/bin/rdmadmin -c debug3 -v [ on | off ] [-e errno] [ -f SpecialFile ]
```

off -> Test I/O の結果として正常が返却されるようになります。

on -> Test I/O の結果として致命的なエラーが返却されるようになります。

※-e オプションで errno を指定した場合は、該当の errno が返却されます。

■ シングルディスク構成の擬似障害試験手順

シングルディスク構成で物理ディスクの障害を擬似す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/RDM/bin/rdmadmin
(monitor status = TRUE)
=====
type      :                : Logical      : I/O
          : H/W Path         : status       : status
=====
VG        : /dev/VolGroup00 : up           :
MIRROR    :                :
PV        : /dev/sda        : up           : up
=====
```

VG ステータスが up となっていることを確認

PV ステータスが up となっていることを確認

- (2) -f SpecialFile オプションで障害を擬似する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障害を起こすことができます。

```
# /opt/HA/RDM/bin/rdmadmin -c debug -v on -f /dev/sda
(monitor status = TRUE)
Change debug value.
special file = /dev/sda
0 -> 1
```

約 60 秒後

```
# /opt/HA/RDM/bin/rdmadmin
(monitor status = TRUE)
=====
type      :                : Logical      : I/O
          : H/W Path         : status       : status
=====
VG        : /dev/VolGroup00 : down       :
MIRROR    :                :
PV        : /dev/sda        : down       : down
=====
```

シングルディスク構成では down となります

メモリ上のステータスを強制的に書き換えます

- (3) 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。

```
Apr 19 10:24:24 host1 rdm[16517]: PV down find .(sf=/dev/sda)
```

```
Apr 19 10:24:24 host1 rdm[16517]: VG status change down .(vg=/dev/VolGroup00) → ①
```

```
Apr 19 10:24:24 host1 rdm[16517]: start KILL clpnm.
```

```
Apr 19 10:24:24 host1 rdm[16517]: send signal clpnm. → ②
```

```
Apr 19 10:24:24 host1 rdm[16517]: abort rdmdiagd. → ③
```

- ① OS ディスク(全パス)のステータスが down になったため、VG の状態も down になります。
- ② CLUSTERPRO X のサーバー管理プロセス (clpnm) の強制終了。
- ③ 監視エンジン(rdmdiagd) の強制終了。
- (4) CLUSTERPRO X のサーバー管理プロセス (clpnm) が強制終了されていることを確認します。

```
# ps -ef |grep clpnm
```

何も表示されません。(強制終了したことを示します。)

- (5) 監視エンジン(rdmdiagd)、監視モニター(rdmping)が強制終了されていることを確認します。

```
# ps -ef |grep rdm
```

何も表示されません。(強制終了したことを示します。)

CLUSTERPRO X のサーバー管理プロセス (clpnm) の強制終了後、クラスターを再度起動する際は以下どちらかの方法で復旧を行ってください。

方法 1: マシンの再起動を行う。

方法 2: CLUSTERPRO X のサーバー管理プロセス (clpnm) のコマンド起動

```
# clpcl -s
```

■ ソフトミラー構成の擬似障害試験手順

ソフトミラー構成のディスクで片系の物理ディスクの障害を擬似す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/RDM/bin/rdmadmin
(monitor status = TRUE)
=====
type      :      : Logical   : I/O
          : H/W Path : status   : status
-----
VG        : /dev/VolGroup00 : up
MIRROR    :
PV        : /dev/sda      : up      : up
PV        : /dev/sdb      : up      : up
```

VG ステータスが up であることを確認

PV ステータスが up であることを確認

- (2) `-f SpecialFile` にて、ミラーを構成する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/RDM/bin/rdmadmin -c debug -v on -f /dev/sdb
(monitor status = TRUE)
Change debug value.
special file = /dev/sdb
0 -> 1
```

約 60 秒後

```
# /opt/HA/RDM/bin/rdmadmin
(monitor status = TRUE)
=====
type      :      : Logical   : I/O
          : H/W Path : status   : status
-----
VG        : /dev/VolGroup00 : suspend
MIRROR    :
PV        : /dev/sda      : up      : up
PV        : /dev/sdb      : down    : down
```

ソフトミラー構成では suspend になります

メモリ上のステータスを強制的に書き換えます

- (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。

```
Apr 19 10:19:44 host1 rdm[16517]: PV down find .(sf=/dev/sdb)
Apr 19 10:19:44 host1 rdm[16517]: VG status change suspend .(vg=/dev/VolGroup00)
```

次に、ソフトミラー構成のディスクで両系の物理ディスクの障害を擬似する手順を説明します。

本手順により、CLUSTERPRO X のサーバー管理プロセス (clpnm) を強制終了し、CLUSTERPRO X との連携を実現することが可能になります。

CLUSTERPRO X との連携方式は 2 通りあり、ここでは『 CLUSTERPRO X のサーバー管理プロセス (clpnm) の強制終了による CLUSTERPRO X との連携 』の動作確認方法を記載します。

- (1) 設定ファイル内の OVER_ACTION に CLPNM_KILL を指定します。

```
# vi /opt/HA/RDM/conf/rdm.config
...
# Root Volume fault action
# select SERVICE_CMD_DISABLE(default),SERVICE_CMD_ENABLE,TOC_EXEC,
# CLPNM_KILL,POWER_OFF
#OVER_ACTION          SERVICE_CMD_DISABLE
OVER_ACTION           CLPNM_KILL
...
...
...
```

設定ファイルを変更した後、デーモンプロセスを再起動してください

【Red Hat Enterprise Linux 7.0 以降、Oracle Linux 7.0 以降、Amazon Linux 2 以降の場合】

```
# systemctl restart rdmd
```

【Red Hat Enterprise Linux 6.x、Oracle Linux 6.x の場合】

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

- (2) CLUSTERPRO X のサーバー管理プロセス (clpnm) が起動していることを確認します。

```
# ps -ef |grep clpnm
root 15691 15690 0 15:36:45 ?          0:00 clpnm
```

- (3) RootDiskMonitor のプロセス rdmdiagd と rdmping が起動していることを確認します。

```
# ps -ef |grep rdm
root 9351      1 0 14:53:12 ?          0:03 /opt/HA/RDM/bin/rdmdiagd
root 9354 9351 0 14:53:12 ?          0:00 rdmping
```

- (4) 障害前に現在の状態をモニタリングします。(既に片系のディスクが異常であること。)

```
# /opt/HA/RDM/bin/rdmadmin
(monitor status = TRUE)
=====
type      :      : Logical      : I/O
          : H/W Path : status      : status
=====
VG        : /dev/VolGroup00 : suspend
MIRROR    :
PV        : /dev/sda      : up          : up
PV        : /dev/sdb      : down        : down
```

片系障害でsuspend状態になっていることを確認

upになっていることを確認

downになっていることを確認

- (5) ソフトミラー構成のディスクの両系の物理ディスクを異常状態にします。

すでに片系が異常な状態で、残りの正常なディスクを異常にすることで両系障害を擬似します。

```
# /opt/HA/RDM/bin/rdmadmin -c debug -v on -f /dev/sda
(monitor status = TRUE)
Change debug value.
special file = /dev/sda
0 -> 1
```

約 60 秒後

```
# /opt/HA/RDM/bin/rdmadmin
(monitor status = TRUE)
=====
type      :      : Logical      : I/O
          : H/W Path : status      : status
=====
VG        : /dev/VolGroup00 : down
MIRROR    :
PV        : /dev/sda      : down        : down
PV        : /dev/sdb      : down        : down
```

両系障害でdown状態になっていることを確認

メモリ上のステータスを強制的に書き換えます

- (6) I/O パスの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。

```
Apr 19 10:24:24 host1 rdm[16517]: PV down find.(sf=/dev/sda)
```

```
Apr 19 10:24:24 host1 rdm[16517]: VG status change down.(vg=/dev/VolGroup00) → ①
```

```
Apr 19 10:24:24 host1 rdm[16517]: start KILL clpnm.
```

```
Apr 19 10:24:24 host1 rdm[16517]: send signal clpnm. → ②
```

```
Apr 19 10:24:24 host1 rdm[16517]: abort rdmdiagd. → ③
```

- ①ソフトミラーを構成する両系のディスクのステータスが down になったため、VG の状態も down になります。
- ②CLUSTERPRO X のサーバー管理プロセス (clpnm) の強制終了。
- ③監視エンジン(rdmdiagd) の強制終了。

- (7) CLUSTERPRO X のサーバー管理プロセス (clpnm) が強制終了されていることを確認します。

```
# ps -ef |grep clpnm
```

何も表示されません。(強制終了したことを示します。)

- (8) 監視エンジン(rdmdiagd)、監視モニター(rdmping)が強制終了されていることを確認します。

```
# ps -ef |grep rdm
```

何も表示されません。(強制終了したことを示します。)

CLUSTERPRO X のサーバー管理プロセス (clpnm) の強制終了後、クラスターを再度起動する際は以下どちらかの方法で復旧を行ってください。

方法 1: マシンの再起動を行う。

方法 2: CLUSTERPRO X のサーバー管理プロセス (clpnm) のコマンド起動

```
# clpcl -s
```

■ シングルディスク構成の致命的なエラーを伴うディスク障害の擬似障害試験手順

シングルディスク構成で致命的なエラーを伴うディスク障害を擬似す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/RDM/bin/rdmadmin
(monitor status = TRUE)
=====
type      :                : Logical  : I/O
          : H/W Path       : status  : status
=====
VG        : /dev/rhel00    : up
MIRROR   :                :
PV       : /dev/sda       : up      : up
          : pci-0000:09:00.0-scsi-0:2:0:0
```

VG ステータスが up となっていることを確認

PV ステータスが up となっていることを確認

(2) -f SpecialFile にて、対象の sd デバイスを指定します。
 指定された sd デバイスに対する Test I/O の結果が致命的なエラーとなり、
 I/O ステータスが up から down に変更され、致命的なエラーを伴う物理ディスク障害を
 起こすことができます。

```
# /opt/HA/RDM/bin/rdmadmin -c debug3 -v on -f /dev/sda
Change debug_fatal_error value.
special file = /dev/sda
0 -> 1  ermo = 2
```

※-e オプションで ermo を指定しない場合は、ermo は 2 が設定されて動作します。

約 20 秒後

```
# /opt/HA/RDM/bin/rdmadmin
(monitor status = TRUE)
=====
type      :                : Logical  : I/O
          : H/W Path       : status  : status
=====
VG        : /dev/rhel00    : down
MIRROR   :                :
PV       : /dev/sda       : down    : down
          : pci-0000:09:00.0-scsi-0:2:0:0
```

VG のステータスは down となります

PV のステータスも down となります。。

- (3) 致命的なエラーにより物理ディスクの VG 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。

Jan 20 11:32:40 host1 rdm[17023]: Test I/O fatal error found. (errno = 2, retry count = 2, sf = /dev/sda) → ①

Jan 20 11:32:40 host1 rdm[17023]: VG status change down .(vg=/dev/rhel00) → ②

Jan 20 11:32:40 host1 rdm[17023]: Action is not define. → ③

- ① 致命的なエラーにより障害を検出した際に出力されるメッセージです。
- ② VG 障害を検出した際に出力されるメッセージです。
- ③ OVER_ACTION に SERVICE_CMD_DISABLE が設定されている場合に出力されるメッセージです。

TOC_EXEC などが設定されている場合は、OS の停止等を優先し、syslog のメッセージ出力が行われない場合があります。

その場合は、OS 停止などアクションが実行されたことで、障害検知が正常に行われたとお考えください。

■ SAN-Boot 構成の擬似障害試験手順

SAN-Boot 構成で片系および両系の I/O パス障害を擬似する手順を説明します。

ここでは、自動閉塞ありの場合の動作確認方法を記載します。

自動閉塞の有無は、設定ファイル(rdm.config)の TESTIO_FAULT_ACTION の項目で指定します。自動閉塞なしの場合は、ACTION_NONE を、自動閉塞ありの場合は、BLOCK_PATH を指定します。

(4) 障害前に現在の状態をモニタリングします。

```
# /opt/HA/RDM/bin/rdmadmin
(monitor status = TRUE)
```

type	:	H/W Path	:	Logical : status	:	I/O : status	:	Online : status
VG	:	/dev/rhel01	:	up	:		:	
MIRROR	:		:		:		:	
PV	:	/dev/sda	:	up	:	up	:	extended
	:		:		:	pci-0000:15:00.0-fc-0x22008cdf9dcaf0ea-lun-0	:	
PV	:	/dev/sdb	:	up	:	up	:	extended
	:		:		:	pci-0000:15:00.0-fc-0x24008cdf9dcaf0ea-lun-0	:	
PV	:	/dev/sdc	:	up	:	up	:	extended
	:		:		:	pci-0000:15:00.1-fc-0x21008cdf9dcaf0ea-lun-0	:	
PV	:	/dev/sdd	:	up	:	up	:	extended
	:		:		:	pci-0000:15:00.1-fc-0x23008cdf9dcaf0ea-lun-0	:	

VG ステータスが up となっていることを確認

PV ステータスが up となっていることを確認

(5) -f SpecialFile にて、ミラーを構成する片系の I/O パスを指定します。
指定された I/O パスは強制的に 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/RDM/bin/rdmadmin -c debug -v on -f /dev/sda
(monitor status = TRUE)
Change debug value.
special file = /dev/sda
0 -> 1
```

約 60 秒後

```
# /opt/HA/RDM/bin/rdmadmin
(monitor status = TRUE)
```

type	:	H/W Path	:	Logical status	:	I/O status	:	Online status
VG	:	/dev/rhel01	:	suspend	:		:	
MIRROR	:		:		:		:	
PV	:	/dev/sda	:	down	:	down	:	reduced
	:		:		:	pci-0000:15:00.0-fc-0x22008cdf9dcaf0ea-lun-0	:	
PV	:	/dev/sdb	:	up	:	up	:	extended
	:		:		:	pci-0000:15:00.0-fc-0x24008cdf9dcaf0ea-lun-0	:	
PV	:	/dev/sdc	:	up	:	up	:	extended
	:		:		:	pci-0000:15:00.1-fc-0x21008cdf9dcaf0ea-lun-0	:	
PV	:	/dev/sdd	:	up	:	up	:	extended
	:		:		:	pci-0000:15:00.1-fc-0x23008cdf9dcaf0ea-lun-0	:	

VG のステータスは suspend となります

メモリ上のステータスを強制的に書き換えます。

自動閉塞されると reduced となります。

- (6) 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。

```
Apr 19 10:24:24 host1 rdm[16517]: PV down find.(sf=/dev/sda)
```

```
Apr 19 10:24:24 host1 rdm[16517]: VG status change suspend.(vg= /dev/rhel01)
```

```
Apr 19 10:24:24 host1 rdmreduced[16519]: I/O path reduced.(sf=/dev/sda)
```

- (7) -f SpecialFile にて、残りの I/O パスを指定します。
指定された I/O パスは強制的に I/O ステータスが up から down に変更され擬似的に両系障害を起こすことができます。

```
# /opt/HA/RDM/bin/rdmadmin -c debug -v on -f /dev/sdb
(monitor status = TRUE)
Change debug value.
special file = /dev/sdb
0 -> 1
# /opt/HA/RDM/bin/rdmadmin -c debug -v on -f /dev/sdc
(monitor status = TRUE)
Change debug value.
special file = /dev/sdc
0 -> 1
# /opt/HA/RDM/bin/rdmadmin -c debug -v on -f /dev/sdd
(monitor status = TRUE)
Change debug value.
special file = /dev/sdd
0 -> 1
```

約 60 秒後

```
# /opt/HA/RDM/bin/rdmadmin
(monitor status = TRUE)
=====
type      :                : Logical : I/O   : Online
          : H/W Path             : status : status : status
=====
VG        : /dev/rhel01       : down
MIRROR    :                   :
PV        : /dev/sda           : down   : down   : reduced
          : pci-0000:15:00.0-fc-0x22008cdf9dcaf0ea-lun-0
PV        : /dev/sdb             : down   : down   : reduced
          : pci-0000:15:00.0-fc-0x24008cdf9dcaf0ea-lun-0
PV        : /dev/sdc             : down   : down   : reduced
          : pci-0000:15:00.1-fc-0x21008cdf9dcaf0ea-lun-0
PV        : /dev/sdd             : down   : down   : extended
          : pci-0000:15:00.1-fc-0x23008cdf9dcaf0ea-lun-0
```

両系障害で down 状態になっていることを確認

メモリ上のステータスを強制的に書き換えます

※自動閉塞の場合でも、最後の 1 パスは reduced にはなりません。

- (8) 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。

Apr 19 10:26:32 host1 rdm[16517]: PV down find .(sf=/dev/sdb)

Apr 19 10:26:32 host1 rdm[16517]: PV down find .(sf=/dev/sdc)

Apr 19 10:26:32 host1 rdm[16517]: PV down find .(sf=/dev/sdd)

Apr 19 10:26:32 host1 rdm[16517]: VG status change down .(vg=/dev/rhel01) →①

- ① VG を構成する両系の I/O パスのステータスが down になったため、
VG の状態も down になります。

■ CLUSTERPRO X 連携

CLUSTERPRO X 連携による デバッグ 手順を説明します。

本手順により、カスタムモニタリソースによる CLUSTERPRO X との連携による OS ディスクの障害時の、ノードを切り替えをデバッグすることが可能になります。クラスタ環境の構築がされていることが前提となります。

- (1) クラスタ環境構築
クラスタ環境構築については、
『 8 章(8.2) カスタムモニタリソースによる CLUSTERPRO X 4.0 以前との連携手順 』または
『 8 章(8.3) カスタムモニタリソースによる CLUSTERPRO X 4.1 以降との連携手順 』を
ご覧ください。
- (2) クラスタの起動
クラスタ環境構築後、クラスタの起動を行います。

クラスタの起動によりサービスとして登録した rdmstat が状態監視を開始します。

※ rdmstat は/opt/HA/RDM/bin 配下から任意のディレクトリ配下へコピーしてご利用ください。

以下は rdmstat を/usr/local/bin 配下にコピーした場合の例です。

```
root 26949 26929 0 10:40:04 ? 0:00 /usr/local/bin/rdmstat -w 10
```

(3) 物理ディスク障害による VG ステータスのダウン

- ① 前述の手順により、物理ディスクの障害を発生させ VG ステータスを down 状態にします。
- ② rdmstat が VG の異常を検出します。
- ③ rdmstat 自身が異常終了します。
- ④ 障害が発生したノードをダウンさせ、ノード切り替えを行います。

syslog を確認してください。

```
Apr 22 10:42:53 host1 rdm[19483]: PV down find .(sf=/dev/sdb)
```

```
Apr 22 10:42:53 host1 rdm[19483]: VG status change down .(vg=/dev/VolGroup00) → ①
```

```
Apr 22 10:42:53 host1 rdm[19483]: Action is not define.
```

動作確認後は、マシンの再起動を行ってください。

■ 物理ディスク疑似障害の復旧

通常の内蔵ディスク構成の場合は、疑似障害状態を解除すれば復旧いたします。

※致命的なエラーを伴う物理ディスク疑似障害の場合も同様です。

SAN-Boot 構成で、I/O パスの自動閉塞を行っている場合は、一時的に監視対象から外されているため、疑似障害状態を解除後に、構成復旧コマンド(rdmrecover)により、復旧する必要があります。

【物理ディスク疑似障害の復旧手順】

- (1) 物理ディスク(/dev/sda)の疑似障害の復旧を行います。

通常の物理ディスク疑似障害の場合

```
# /opt/HA/RDM/bin/rdmadmin -c debug -v off -f /dev/sda
Change debug value.
special file = /dev/sda
1 -> 0
```

致命的なエラーを伴う物理ディスク疑似障害の場合

```
# /opt/HA/RDM/bin/rdmadmin -c debug3 -v off -f /dev/sda
Change debug_fatal_error value.
special file = /dev/sda
1 -> 0
```

```
# /opt/HA/RDM/bin/rdmadmin
(monitor status = TRUE)
```

		Logical	I/O
type	H/W Path	status	status
VG	:/dev/VolGroup00	up	
MIRROR	:	:	
PV	:/dev/sda	up	up

up になることを確認します

- (2) syslog メッセージを確認してください。

```
Apr 19 11:57:40 host1 rdm[16517]: VG status change up.(vg=/dev/VolGroup00)
```

【SAN-Boot 構成での物理ディスク擬似障害の復旧手順】

- (1) 擬似障害を発生させた I/O パスの擬似障害の復旧を行います。

```
# /opt/HA/RDM/bin/rdmadmin -c debug -v off -f /dev/sda
Change debug value.
special file = /dev/sda
1 -> 0
# /opt/HA/RDM/bin/rdmadmin -c debug -v off -f /dev/sdb
Change debug value.
special file = /dev/sdb
1 -> 0
# /opt/HA/RDM/bin/rdmadmin -c debug -v off -f /dev/sdc
Change debug value.
special file = /dev/sdc
1 -> 0
# /opt/HA/RDM/bin/rdmadmin -c debug -v off -f /dev/sdd
Change debug value.
special file = /dev/sdd
1 -> 0
```

- (2) rdmrecover コマンドを実行し、VG と PV のステータスの復旧を行います。

```
# /opt/HA/RDM/bin/rdmrecover
rdmrecover complete.
```

- (3) rdmadmin コマンドで、VG と PV の状態を確認してください。

```
# /opt/HA/RDM/bin/rdmadmin
(monitor status = TRUE)
=====
type      : Logical   : I/O      : Online
           : H/W Path   : status   : status   : status
=====
VG        : /dev/rhel01      : up
MIRROR   :                  :
PV       : /dev/sda         : up       : up       : extended
           : pci-0000:15:00.0-fc-0x22008cdf9dcaf0ea-lun-0
PV       : /dev/sdb         : up       : up       : extended
           : pci-0000:15:00.0-fc-0x24008cdf9dcaf0ea-lun-0
PV       : /dev/sdc         : up       : up       : extended
           : pci-0000:15:00.1-fc-0x21008cdf9dcaf0ea-lun-0
PV       : /dev/sdd         : up       : up       : extended
           : pci-0000:15:00.1-fc-0x23008cdf9dcaf0ea-lun-0
```

up になっていることを確認します

up になっていることを確認します。

extended になっていることを確認します。

■ I/O ストール擬似障害

本手順により、I/O ストール状態を擬似的に発生させることができます。
設定ファイル内の VG_STALL_ACTION に SERVICE_CMD_DISABLE を指定し、
syslog に I/O ストールのエラーメッセージが出力されることを確認してください。

【コマンド書式】

```
/opt/HA/RDM/bin/rmdadmin -c debug2 -v [ on | off ]
```

off -> I/O stall stop I/O ストール擬似障害を復旧します。

on -> I/O stall start I/O ストール擬似障害を開始します。

- (1) 設定ファイル内の VG_STALL_ACTION に SERVICE_CMD_DISABLE を指定します。

```
# vi /opt/HA/RDM/conf/rdm.config
.....
# RootVolume stall find action
# select SERVICE_CMD_DISABLE(default),SERVICE_CMD_ENABLE,TOC_EXEC,
# CLPNM_KILL,POWER_OFF
VG_STALL_ACTION          SERVICE_CMD_DISABLE
.....
.....
```

設定ファイルを変更した後、デーモンプロセスを再起動してください

【Red Hat Enterprise Linux 7.0 以降、Oracle Linux 7.0 以降、Amazon Linux 2 以降の場合】

```
# systemctl restart rdm
```

【Red Hat Enterprise Linux 6.x、Oracle Linux 6.x の場合】

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

- (2) RootDiskMonitor のプロセス rmddiagd、rdmping が起動していることを確認します。

```
# ps -ef |grep rdm
root 9351    1  0 15:37:35 ?        0:03 /opt/HA/RDM/bin/rmddiagd
root 9354  9351  0 15:37:35 ?        0:00 rdmping
```

- (3) rmdadmin コマンドにより擬似的に I/O ストール擬似障害を開始します。

実行後、以下のメッセージが出力されます。

```
# /opt/HA/RDM/bin/rmdadmin -c debug2 -v on
DEBUG: I/O STALL start.
```


- (4) I/O ストールのエラーメッセージを確認します。
※ (3)の rdmadmin コマンド実行後、設定ファイルに記載の
TIME_VG_STALL(デフォルト 360 秒)経過した場合に syslog に出力されます。

```
# vi /var/log/messages
Jul 21 17:25:06 host1 rdm[11930]: I/O stall find , timeover occurred.(sf=/dev/sda) ★
Jul 21 17:25:06 host1 rdm[11930]: Action is not define.(VG STALL)
```

評価完了後、I/O ストール擬似障害を復旧します。

- (5) I/O ストール擬似障害の復旧を行います。

実行後、以下のメッセージが出力されます。

```
# /opt/HA/RDM/bin/rdmadmin -c debug2 -v off
DEBUG: I/O STALL stop.
```

また syslog に以下のメッセージが出力されます。

```
# vi /var/log/messages
Jul 21 17:26:06 host1 rdm[11930]: DEBUG:I/O stall DEBUG mode end. I/O restore.
```

8.2. カスタムモニタリソースによる CLUSTERPRO X 4.0 以前との連携手順

8.2.1. CLUSTERPRO X 連携設定

本製品は、カスタムモニタリソースによる CLUSTERPRO X 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ーの実 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/

CLUSTERPRO X 2.x および 3.x の場合、上記手順にて WebManager が表示されます。

CLUSTERPRO X 4.0 の場合、上記手順にて Cluster WebUI が表示されます。

Cluster WebUI のメニューバーから WebManager を選択してください。

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)	rdmstat_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	クラスターサービス停止と 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(rdmstat_mon) が rdmstat の消滅を検知し、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 を例とした設定手順となります。
CLUSTERPRO X のバージョンにより画面が異なる場合がありますが、同様の手順で設定可能です。

1. 設定ファイルの確認

RootDiskMonitor の設定を確認します。

- (1) RootDiskMonitor の設定ファイル (rdm.config) にて OVER_ACTION が SERVICE_CMD_ENABLE になっていることを確認します。

```
# /bin/cat /opt/HA/RDM/conf/rdm.config | /bin/grep "OVER_ACTION"  
OVER_ACTION          SERVICE_CMD_ENABLE
```

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

※OVER_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(rdmd)の再起動を行ってください。

- (2) 設定ファイル (rdm.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 /opt/HA/RDM/conf/rdm.config
```

- (3) 設定ファイル (rdm.config) が正しく変更されていることを確認します。

```
# /bin/cat /opt/HA/RDM/conf/rdm.config | /bin/grep "OVER_ACTION"  
OVER_ACTION          SERVICE_CMD_ENABLE
```

- (4) デーモンプロセス (rdmd) を再起動します。

【Red Hat Enterprise Linux 7.0 以降、Oracle Linux 7.0 以降、Amazon Linux 2 以降の場合】

```
# systemctl stop rdmd
```

```
# systemctl start rdmd
```

【Red Hat Enterprise Linux 6.x、Oracle Linux 6.x の場合】

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

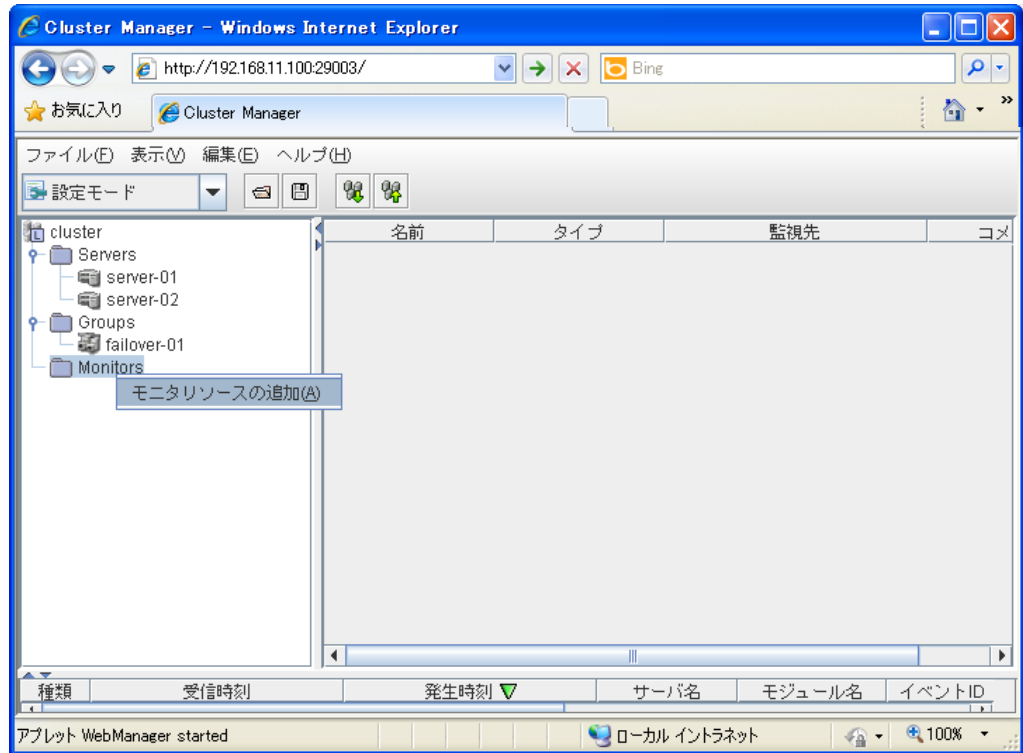
```
# /etc/init.d/rdmd start
```

2. カスタムモニタリソースの作成

rdmstat を監視するカスタムモニタリソースを作成します。

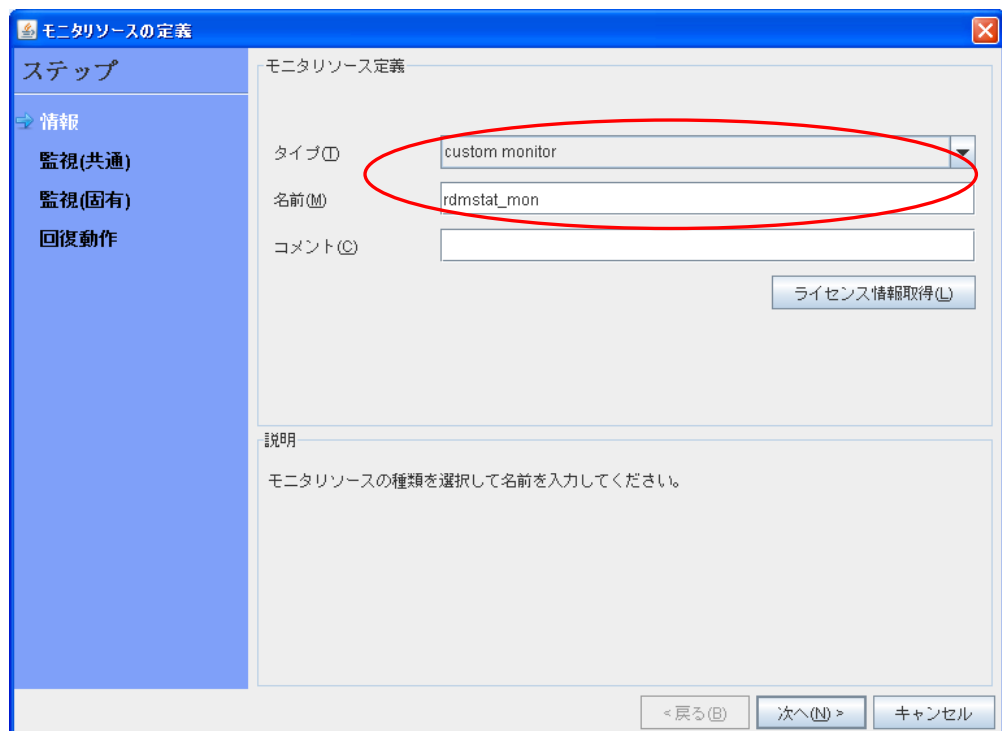
CLUSTERPRO WebManager の「表示」メニューより「設定モード」を選択し、設定します。

(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。



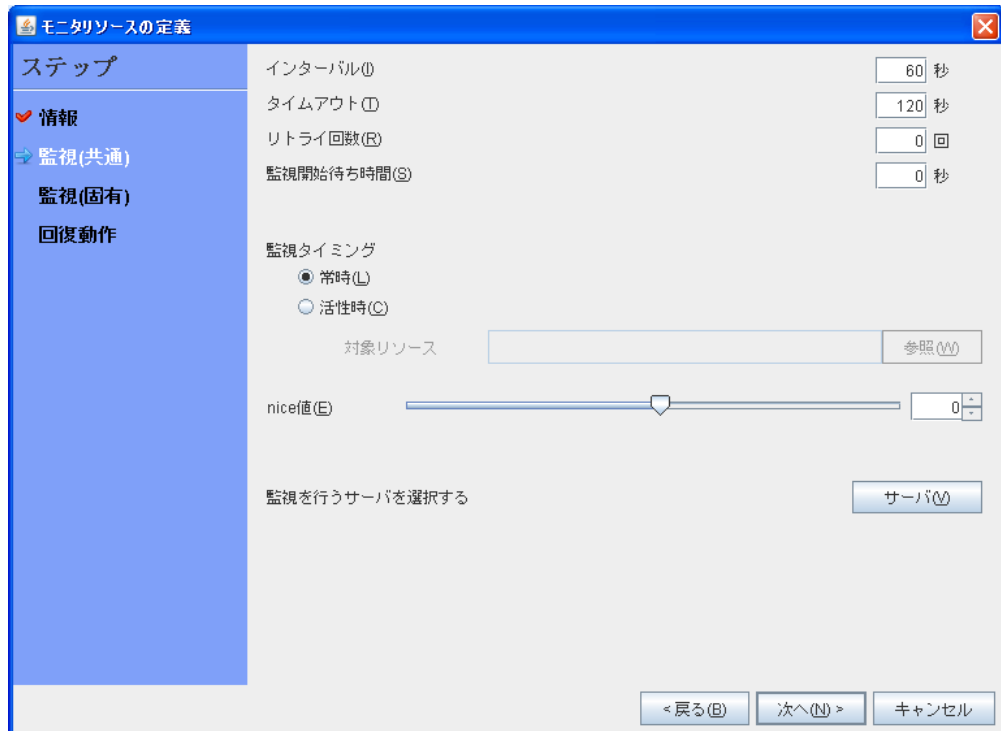
(2) 「モニタリソースの定義」ダイアログボックスが開きます。

「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"rdmstat_mon"を入力し、「次へ」をクリックします。



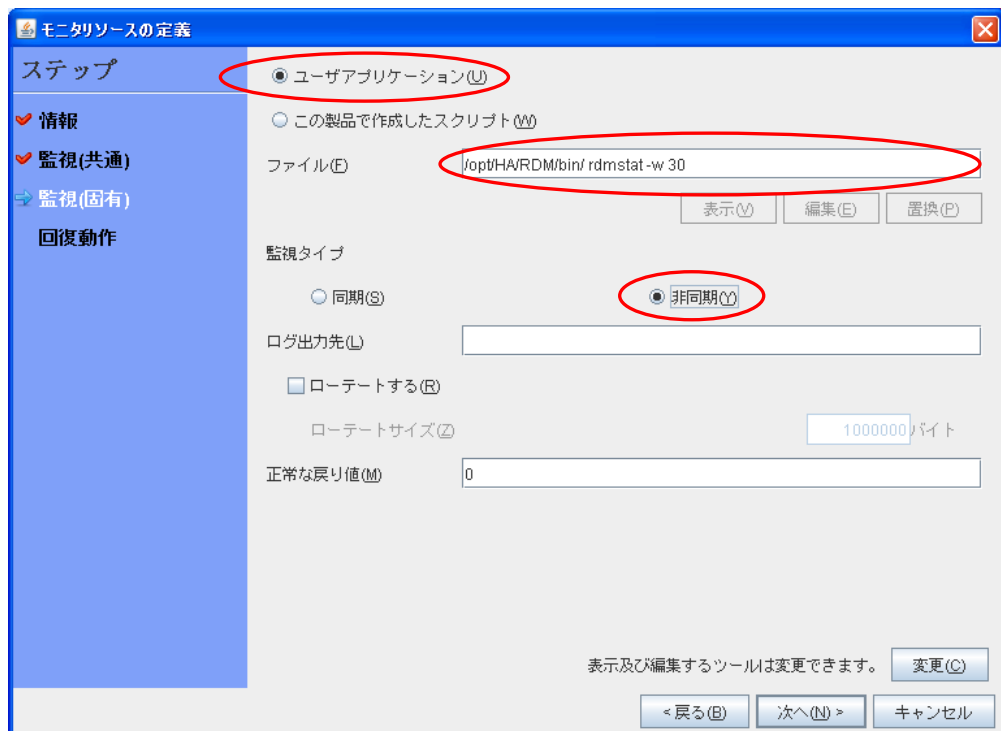
(3) 監視条件を設定します。

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 rdmstat 起動処理 (/opt/HA/RDM/bin/rdmstat -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。



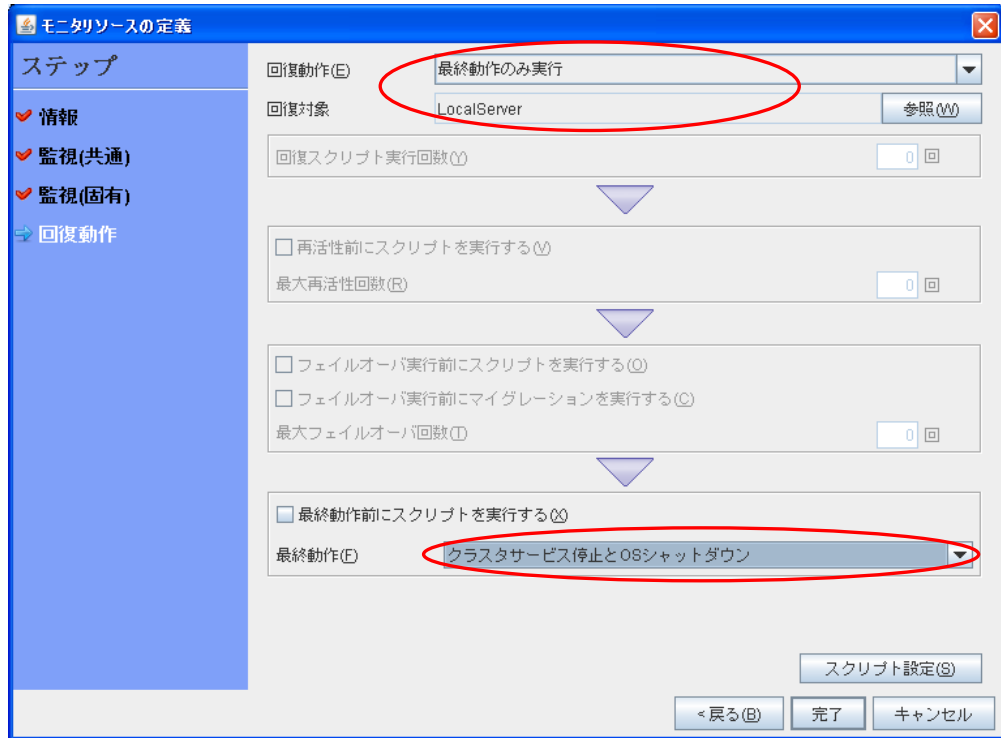
(5) 回復動作を設定します。

「回復動作」で「最終動作のみ実行」を選択します。

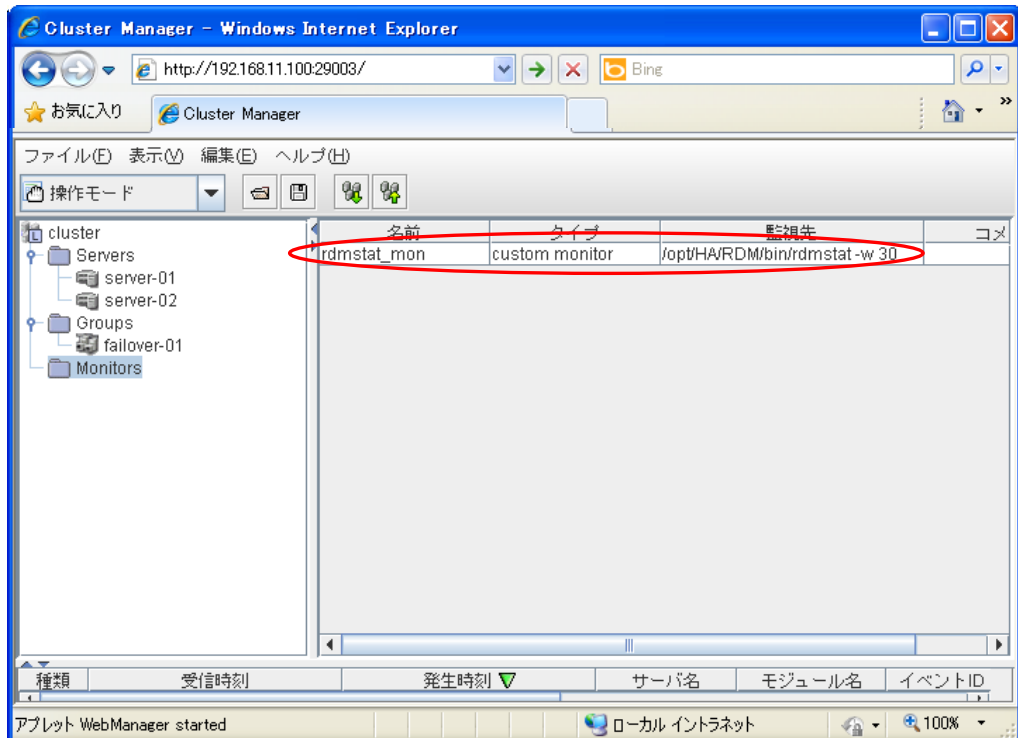
「回復対象」の「参照」をクリックし、表示されるツリービューで"LocalServer"を選択して

「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「クラスターサービス停止とOS シャットダウン」を選択し、「完了」をクリックします。

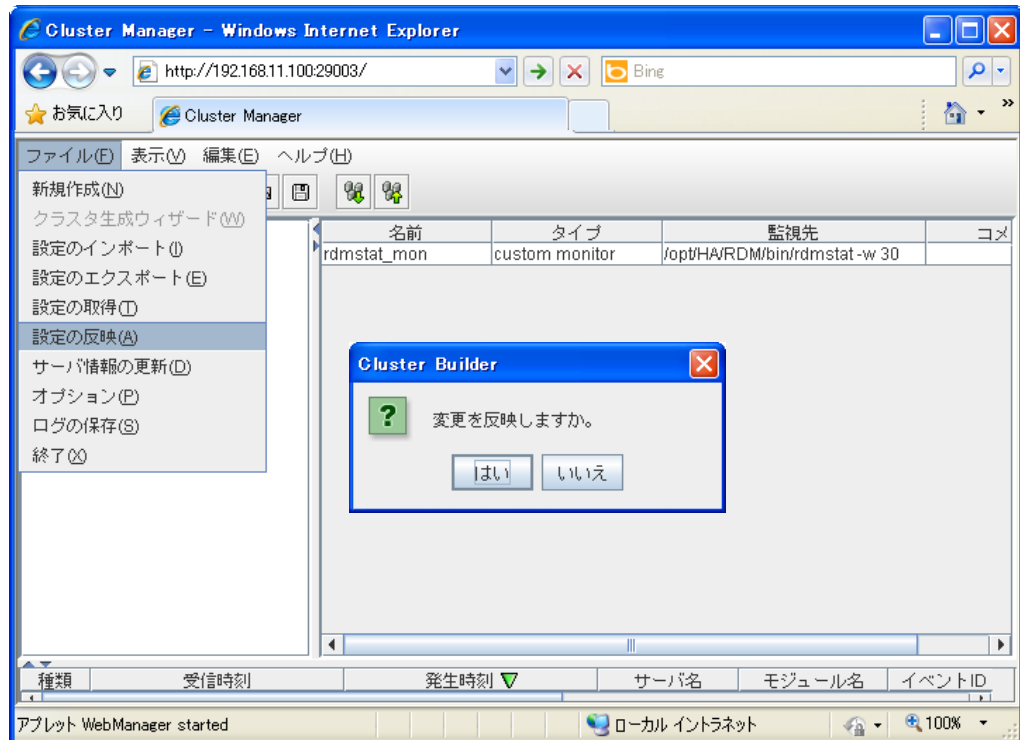


(6) 「Monitors」をクリックし、カスタムモニタリソースが登録されていることを確認します。



3. クラスター構成情報のアップロード

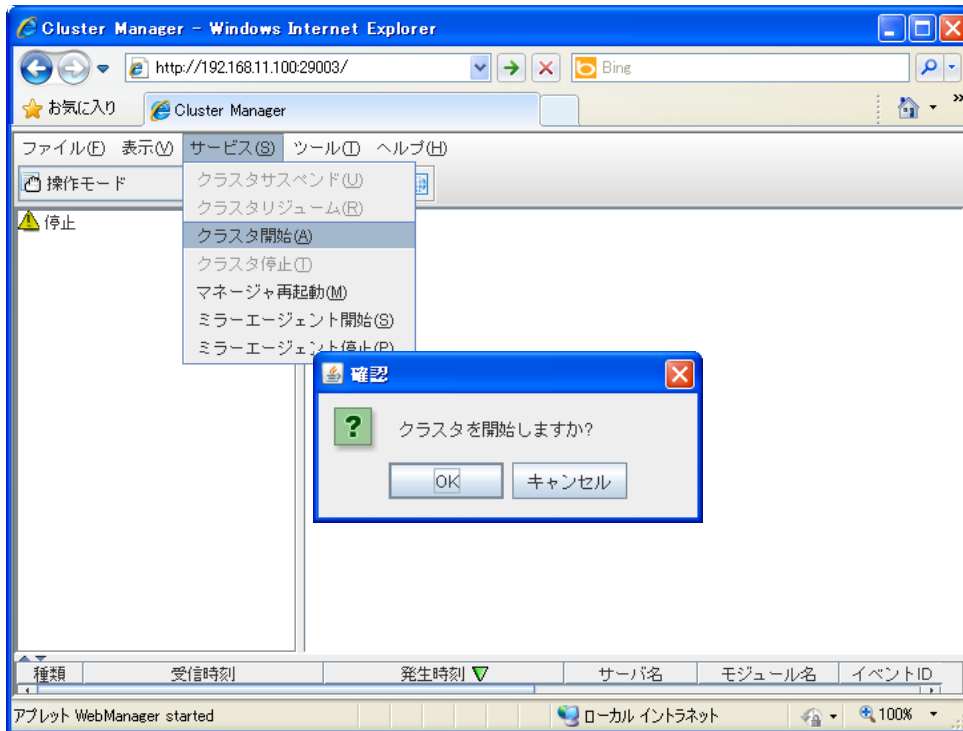
- (1) クラスター構成情報の内容を、CLUSTERPRO X 本体の環境に反映します。
「ファイル」メニューから「設定の反映」を選択します。
確認ダイアログが表示されます。「OK」をクリックします。



- ※ 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 の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。

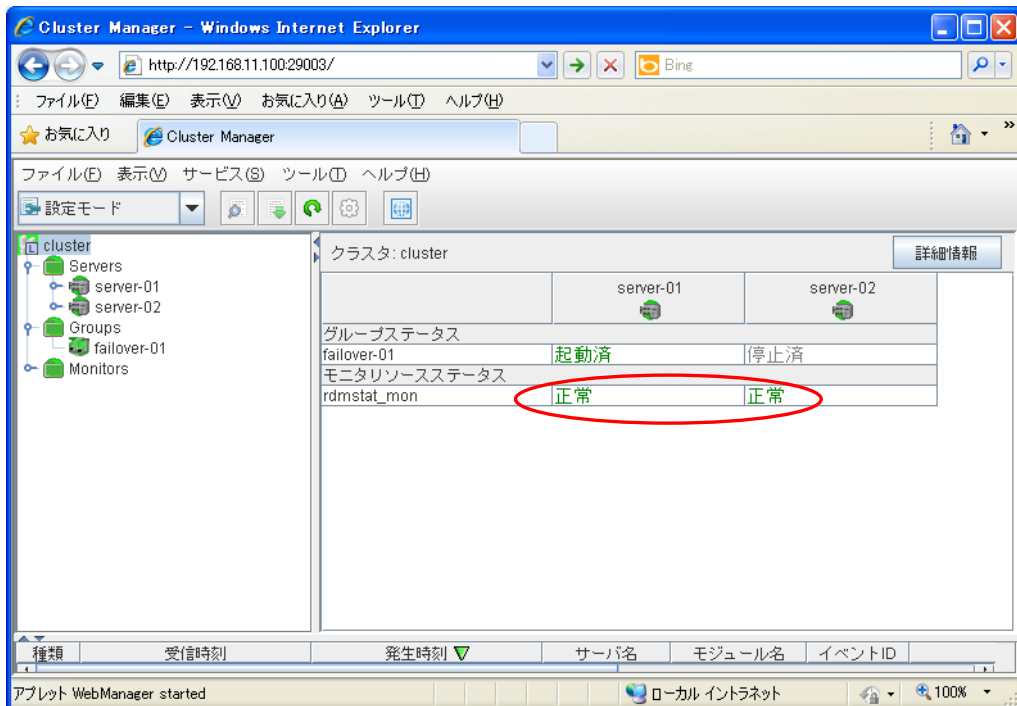
クラスターが停止状態の場合は、クラスターを開始します。
 WebManager の「表示」メニューより「操作モード」を選択し、「サービス」メニューから「クラスター開始」を選択し、クリックしてください。



(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。

- ・ 現用系サーバー、待機系サーバーにて rdmstat 監視用のカスタムモニタリソース「rdmstat_mon」のステータスが「正常」であることを確認してください。



以上で、CLUSTERPRO の設定は終了です。

8.2.2. 動作確認

以降の手順で RootDiskMonitor および CLUSTERPRO X の設定の動作確認を行います。
コマンドオペレーションで擬似的にディスク障害を発生させ、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ールオーバーが発生することを確認します。

1. 現用系サーバーでの RootDiskMonitor 動作確認

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ます。
擬似障害テスト手順については、「8.1 本製品のテスト手順について」の「物理ディスク擬似障害」の章を参照してください。

2. 待機系サーバーへのフェールオーバー確認

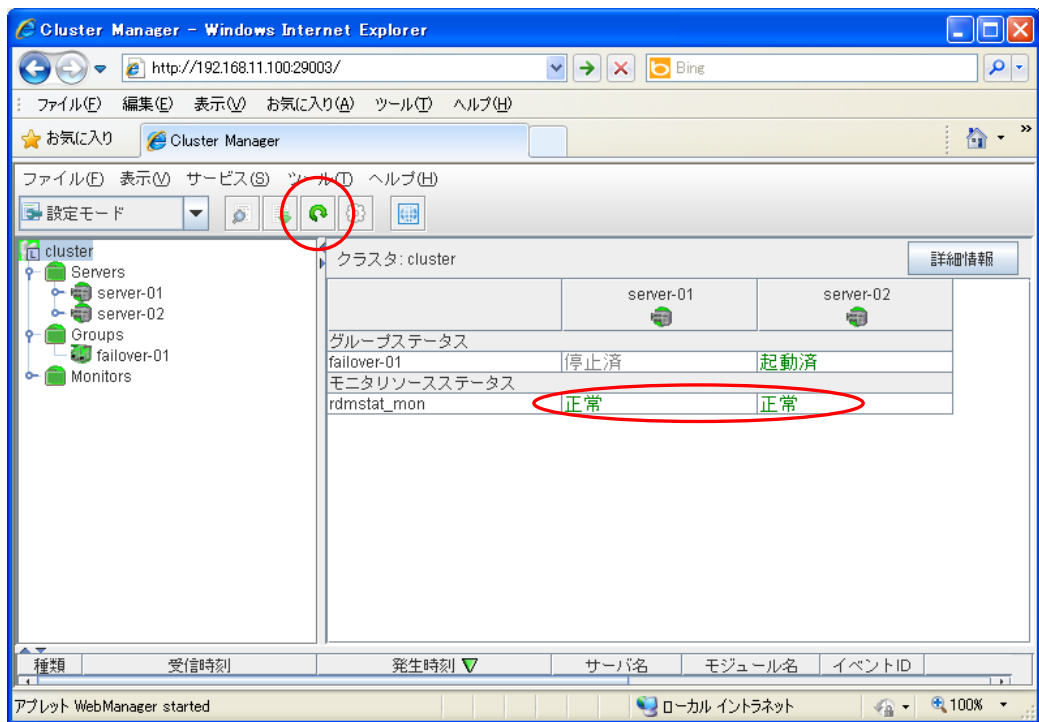
(1) 手順 1. で shutdown したサーバーを起動させてください。

(2) syslog に 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: [!] <type: rm><event: 16> Stopping the system has been required because an error was detected in monitoring rdmstat_mon.  
clusterpro: [!] <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。

- rdmstat 監視用のカスタムモニタリソース「rdmstat_mon」のステータスが現用系、待機系にて「正常」であることを確認してください。



以上で、動作確認は終了となります。

8.3. カスタムモニタリソースによる CLUSTERPRO X 4.1 以降との連携手順

8.3.1. CLUSTERPRO X 連携設定

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

以下の作業は、CLUSTERPRO Cluster WebUI にて実施します。

本書は、CLUSTERPRO Server をインストールしたサーバーの実 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/

Cluster WebUI のメニューバーから [設定モード] を選択してください。

また、本書では『CLUSTERPRO Cluster WebUI』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)	カスタムモニタ
名前 (カスタムモニタリソース名)	rdmstat_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	クラスターサービス停止と 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(rdmstat_mon) が rdmstat の消滅を検知し、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 4.1 を例とした設定手順となります。
CLUSTERPRO X のバージョンにより画面が異なる場合がありますが、同様の手順で設定可能です。

1. 設定ファイルの確認

RootDiskMonitor の設定を確認します。

- (1) RootDiskMonitor の設定ファイル (rdm.config) にて OVER_ACTION が SERVICE_CMD_ENABLE になっていることを確認します。

```
# /bin/cat /opt/HA/RDM/conf/rdm.config | /bin/grep "OVER_ACTION"  
OVER_ACTION          SERVICE_CMD_ENABLE
```

※旧バージョンとの互換を保証するため、OVER_ACTION が ACTION_NONE の場合にも CLUSTERPRO X 連携は可能です。
※OVER_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(rdmd)の再起動を行ってください。

- (2) 設定ファイル (rdm.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 /opt/HA/RDM/conf/rdm.config
```

- (3) 設定ファイル (rdm.config) が正しく変更されていることを確認します。

```
# /bin/cat /opt/HA/RDM/conf/rdm.config | /bin/grep "OVER_ACTION"  
OVER_ACTION          SERVICE_CMD_ENABLE
```

- (4) デーモンプロセス (rdmd) を再起動します。

【Red Hat Enterprise Linux 7.0 以降、Oracle Linux 7.0 以降、Amazon Linux 2 以降の場合】

```
# systemctl stop rdmd
```

```
# systemctl start rdmd
```

【Red Hat Enterprise Linux 6.x、Oracle Linux 6.x の場合】

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

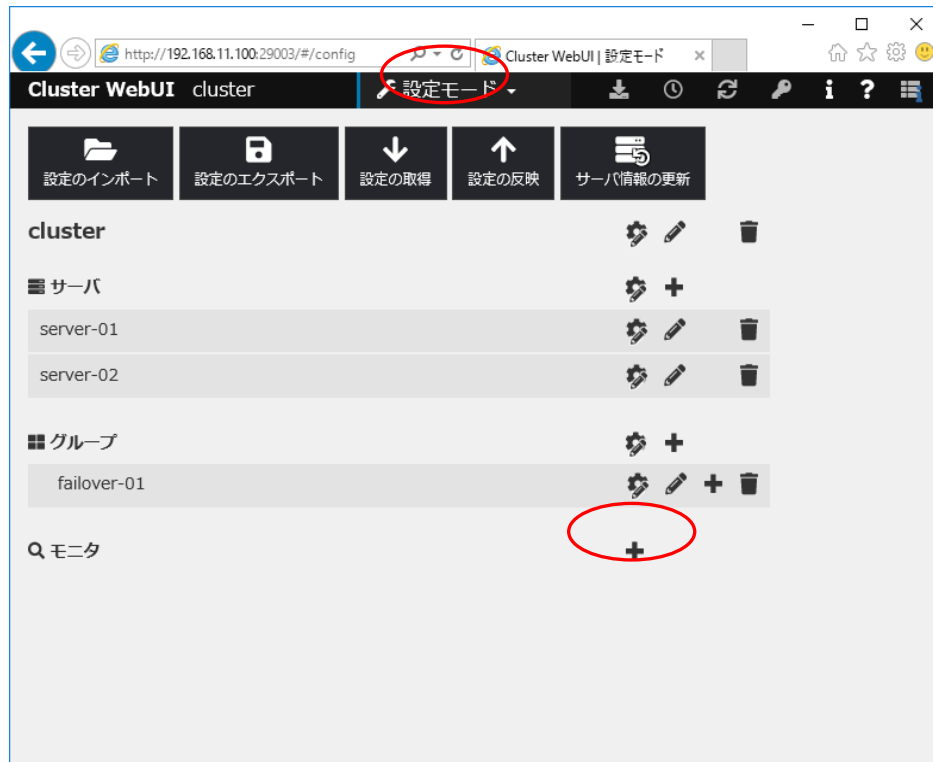
```
# /etc/init.d/rdmd start
```

2. カスタムモニタリソースの作成

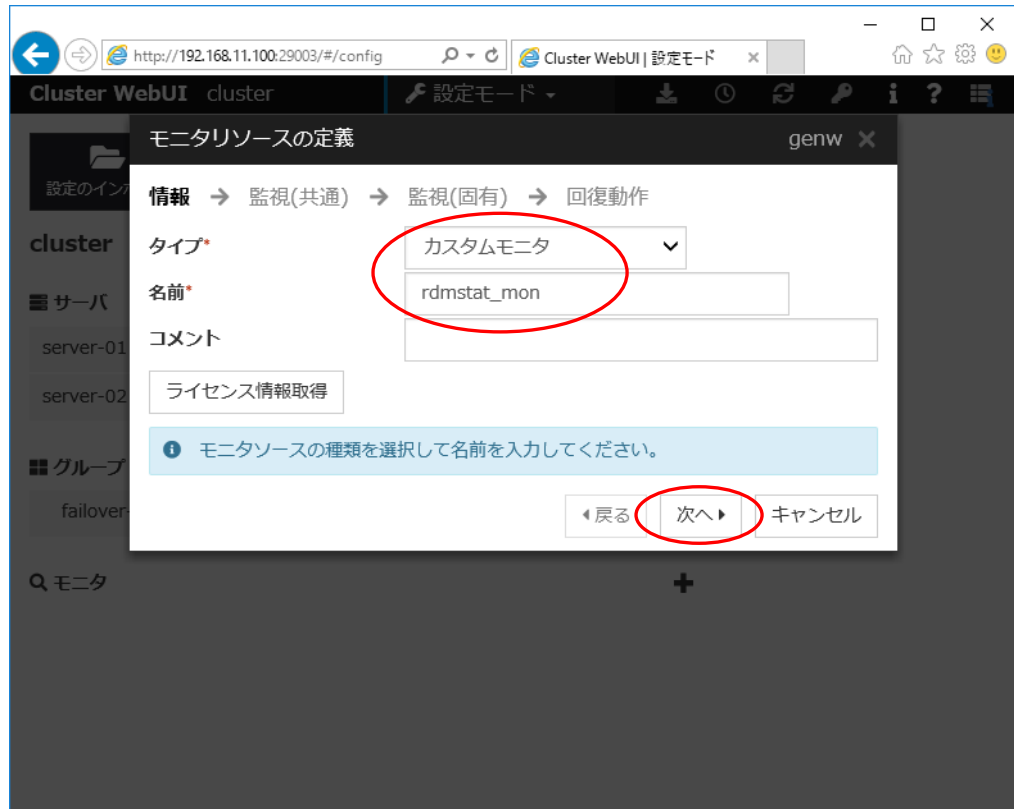
rdmstat を監視するカスタムモニタリソースを作成します。

Cluster WebUI ツールバーのドロップダウンメニューで [設定モード] を選択します。

(1) モニタの追加 [+] をクリックします。

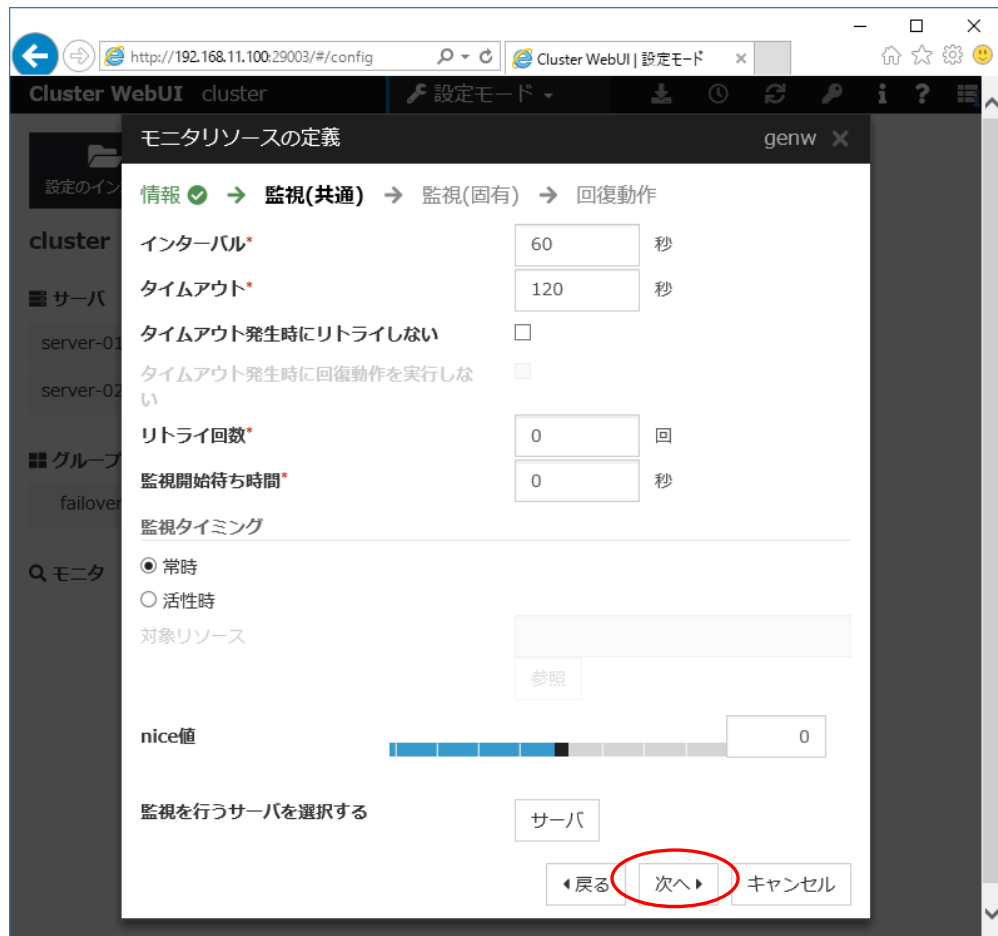


- (2) 「モニタリソースの定義」ダイアログボックスが開きます。
「タイプ」で「カスタムモニタ」を選択し、「名前」にカスタムモニタリソース名 "rdmstat_mon" を入力し、「次へ」をクリックします。



(3) 監視条件を設定します。

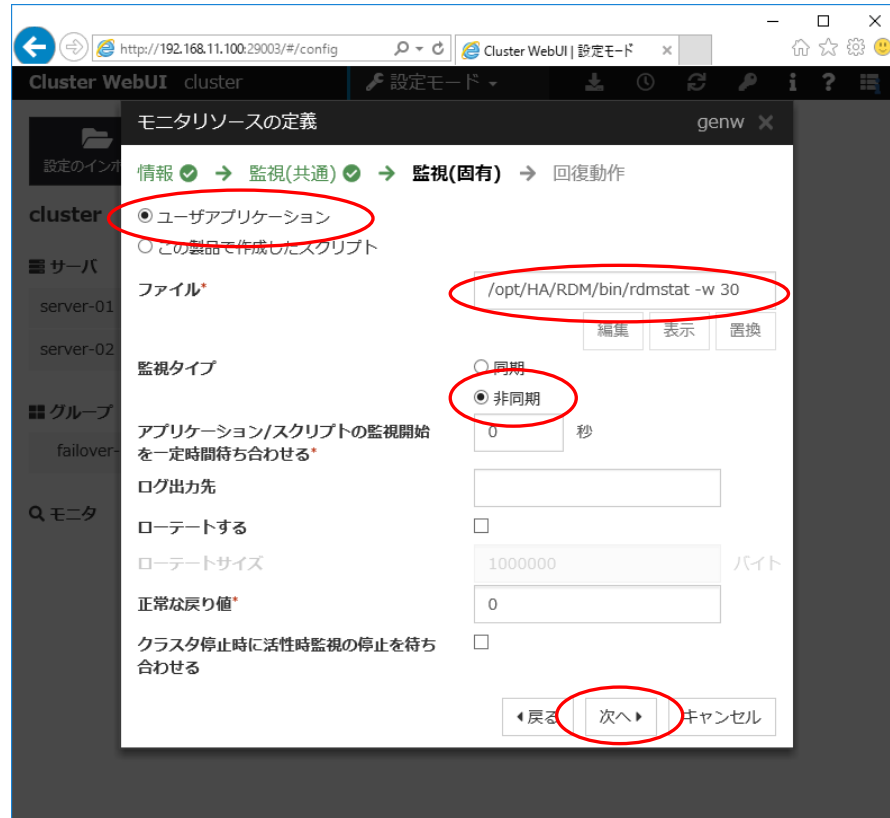
本書ではデフォルトのまま変更しません。「次へ」をクリックします。



(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 rdmstat 起動処理 (/opt/HA/RDM/bin/rdmstat -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。

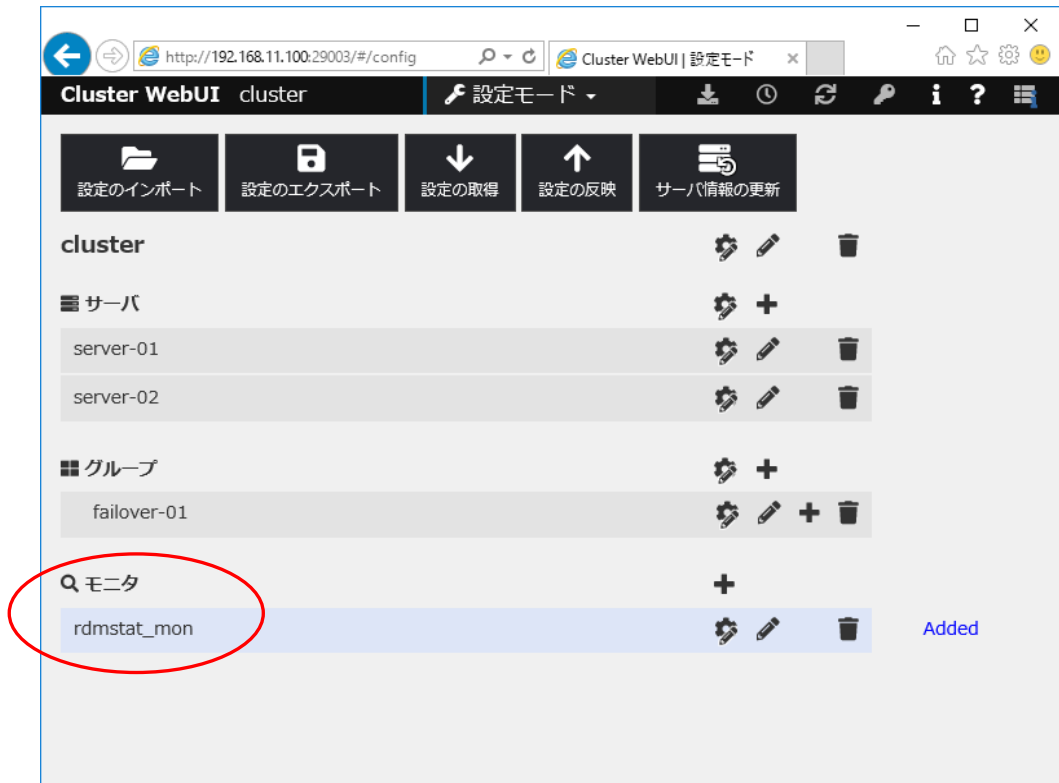
「回復対象」の「参照」をクリックし、表示されるツリービューで"LocalServer"を選択して

「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。

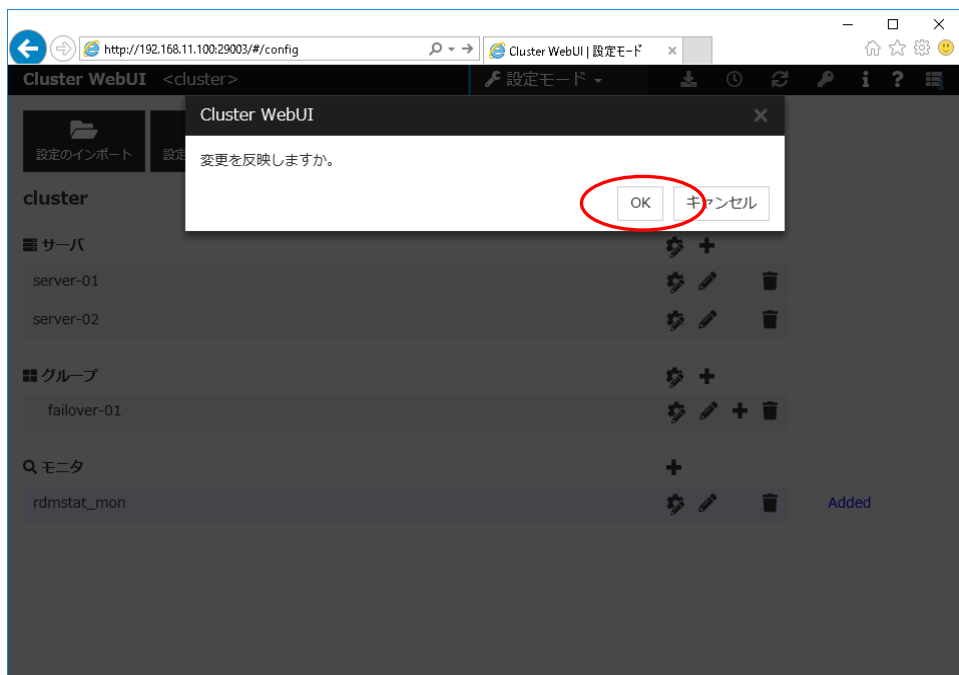
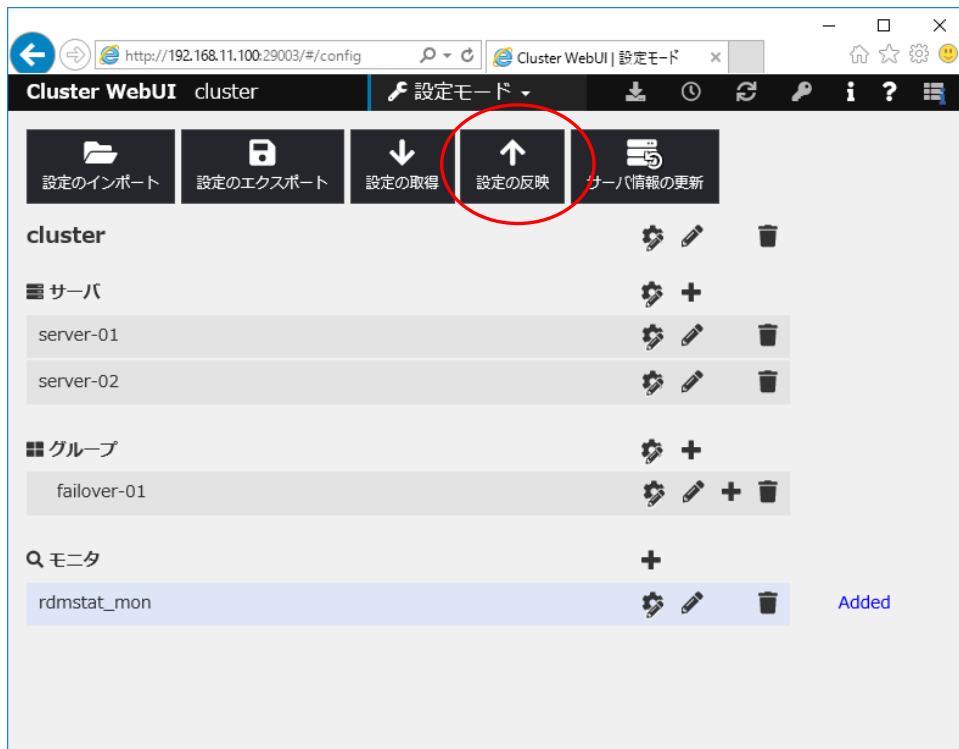


(6) モニタに、カスタムモニタリソースが登録されていることを確認します。



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 X 本体の環境に反映します。
「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。



クラスタサスペンド状態、またはクラスタを停止している場合は、
クラスタリジュームもしくは、クラスタを開始してください。

(2) 設定が反映されていることを確認します。

Cluster WebUI ツールバーのドロップダウンメニューより「操作モード」を選択し、「ステータス」タブで、以下の項目を確認してください。

- ・ 現用系サーバー、待機系サーバーにて rdmstat 監視用のカスタムモニタリソース「rdmstat_mon」のステータスが「正常」であることを確認してください。



以上で、CLUSTERPRO の設定は終了です。

8.3.2. 動作確認

以降の手順で RootDiskMonitor および CLUSTERPRO X の設定の動作確認を行います。
コマンドオペレーションで擬似的にディスク障害を発生させ、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ールオーバーが発生することを確認します。

1. 現用系サーバーでの RootDiskMonitor 動作確認

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ます。
擬似障害テスト手順については、「8.1 本製品のテスト手順について」の「物理ディスク擬似障害」の章を参照してください。

2. 待機系サーバーへのフェールオーバー確認

(1) 手順 1. で shutdown したサーバーを起動させてください。

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

(3) Cluster WebUi の「最新情報を取得」をクリックし、以下の項目を確認してください。

- rdmstat 監視用のカスタムモニタリソース「rdmstat_mon」のステータスが現用系、待機系にて「正常」であることを確認してください。



以上で、動作確認は終了となります。

8.4. Multiple Devices 構成における設定ファイル作成手順

Multiple Devices で構築したソフトウェア RAID 構成(以後、Multiple Device 構成)の OS ディスクを監視することができます。

Multiple Device 構成の設定ファイルは手動で作成します。下記手順に従って作成してください。

設定ファイルの作成手順は次のとおりです。

- (1) 設定ファイルテンプレートの作成
- (2) Multiple Device 構成の確認
- (3) マルチパス管理製品構成の確認
- (4) sd デバイスに対応する udev の確認
- (5) 設定ファイルの編集(Multiple Device 構成の反映)

手順について説明します。

- (1) 設定ファイルテンプレートの作成
設定ファイル自動生成コマンド(rdmconfig)を用いて設定ファイル(rdm.config)のテンプレートを作成します。
設定ファイル自動生成コマンド(rdmconfig)の実行方法については「3.1 本製品の導入」の「(4)セットアップ」を参照してください。
- (2) Multiple Device 構成の確認
Multiple Device で管理しているデバイスのデバイス名を確認します。
/proc/mdstat を参照し、Multiple Device 名(mdx) と
管理対象デバイス(mdx : 行の 4 フィールド目以降)を確認します。

```
# cat /proc/mdstat
Personalities : [raid6] [raid5] [raid4] [raid0]
md0 : active raid1 dda[0] ddb[1]
                1036288 blocks super 1.2 level 1, 512k chunk, algorithm 1 [2/2] [UU]

unused devices: <none>
```

Multiple Device 名

管理対象デバイス

- (3) マルチパス管理製品構成の確認
マルチパス管理製品を導入している場合、前項(2)で確認した管理対象デバイスはマルチパス管理製品のデバイスとなります。当該デバイスで束ねている sd デバイスを確認します。
確認手順はマルチパス管理製品により異なります。該当する製品の手順で確認してください。
なお、マルチパス管理製品を導入していない場合は次項(4)に進みます。

- StoragePathSavior

マルチパスデバイスの一覧から sd デバイス名を確認します。

spsadmin --lun コマンドを実行します。表示されたマルチパスデバイスの一覧から、SPS デバイス(ddX)に束ねられている sd デバイス(sdX)を確認します。

```
# spsadmin --lun

+++ LogicalUnit 7:0:0:0 /dev/dda [Normal] +++
  SerialNumber=0000000939747987, LUN=0x000C1
  LoadBalance=LeastSectors
  0: ScsiAddress=1:0:0:0, ScsiDevice=/dev/sdb, Priority=1, Status=Active
  10: ScsiAddress=2:0:0:0, ScsiDevice=/dev/sdc, Priority=2, Status=Active
```

SPS デバイス

sd デバイスファイル

- PowerPath

マルチパスデバイスの一覧から sd デバイス名を確認します。

powermt display dev=all コマンドを実行します。表示されたマルチパスデバイスの一覧から、emcpower デバイス(emcpowerX)に束ねられている sd デバイス(sdX)を確認します。

```
# powermt display dev=all

Pseudo name=emcpowera
VNX ID=CKM00134600693 [SG_Lin]
Logical device ID=6006016016903600ED105E92C0FDE411 [LUN 24]
state=alive; policy=CLAROpt; queued-IOs=0
Owner: default=SP B, current=SP B Array failover mode: 4

----- Host ----- - Stor - -- I/O Path -- -- Stats ---
### HW Path          I/O Paths  Interf.  Mode   State  Q-IOs Errors
-----
  2 lpfc              sdb      SPA0    active alive   0      0
  3 lpfc              sdc      SP B0   active alive   0      0
```

emcpower デバイス

sd デバイスファイル

- Device-Mapper Multipath

マルチパスデバイスの一覧から sd デバイス名を確認します。

multipath -ll コマンドを実行します。表示されたマルチパスデバイスの一覧から、mpath デバイス (mpathX)に束ねられている sd デバイス(sdX)を確認します。

```
# multipath -ll
mpath01(200255c3a02660091) dm-18 NEC,iStorage 1000
[size=1.0G][features=0][hw_handler=0][rw]
¥_ round-robin 0 [prio=1][enabled]
¥_ 7:0:0:16426 sdb 66:192 [active][ready]
¥_ round-robin 0 [prio=1][enabled]
¥_ 8:0:0:16426 sdc 129:32 [active][ready]
```

mpath デバイス

sd デバイス

(4) sd デバイスに対応する udev の確認

sd デバイスに対応する udev を確認します。

/dev/disk/by-path を参照し、sd デバイスに対応する udev を確認します。

```
# ls -l /dev/disk/by-path
lrwxrwxrwx 1 root root          9 Jul 9 13:20
pci-0000:13:00.0-fc-0x210000255c3a0266:0x0000000000000000 -> ../sdb
lrwxrwxrwx 1 root root          9 Jul 9 13:20
pci-0000:13:00.0-fc-0x210000255c3a0266:0x0001000000000000 -> ../sdc
```

sd デバイス

udev

(5) 設定ファイルの編集(Multiple Device 構成の反映)

前項(1)で作成した設定ファイルテンプレート(rdm.config)に Multiple Device 構成を反映します。

rdm.config の Device Config Area の PV を前項(2)~(4)で確認した Multiple Device 構成にあわせて同一 MIRROR エントリー配下に配置します。

下記に OS ディスク(sdb、sdc)をソフトウェアミラーした Multiple Device 構成に対する rdm.config の修正を例示します。

- 修正前(設定ファイルテンプレート)

```
#####
# Device Config Area
#####
# VG          volume group for LVM(/dev/VolGroup00 , etc ....)
# MIRROR      PV set for Mirror(set any number)
# PV          PV Block Device File(/dev/sda , etc ....) or
# PV          udev Device(pci-0000:0c:00.0-scsi-0:2:0:0 , etc ....)

VG          VG_NONE
MIRROR      group01
## PV Name: /dev/sdb
PV          pci-0000:13:00.0-fc-0x210000255c3a0266:0x0000000000000000
MIRROR      group02
## PV Name: /dev/sdc
PV          pci-0000:13:00.0-fc-0x210000255c3a0266:0x0001000000000000
```

ミラー構成ではない

- 修正後(Multiple Device 構成を反映)

```
#####
# Device Config Area
#####
# VG          volume group for LVM(/dev/VolGroup00 , etc ....)
# MIRROR      PV set for Mirror(set any number)
# PV          PV Block Device File(/dev/sda , etc ....) or
# PV          udev Device(pci-0000:0c:00.0-scsi-0:2:0:0 , etc ....)

VG          VG_NONE
MIRROR      group01
## PV Name: /dev/sdb
PV          pci-0000:13:00.0-fc-0x210000255c3a0266:0x0000000000000000
## PV Name: /dev/sdc
PV          pci-0000:13:00.0-fc-0x210000255c3a0266:0x0001000000000000
```

ミラー構成とする

8.5. S.M.A.R.T.診断結果の定期監視手順

S.M.A.R.T.診断結果を定期監視するためには、smartmontools の設定が必要です。
ただし、HW-RAID 構成で定期監視を行う場合、HW-RAID 監視機能の依存パッケージである Universal RAID Utility が定期監視を行うため、本設定は不要です。

8.5.1. 定期監視の利用手順

定期監視を行うことで、障害検知時に syslog にメッセージを出力します。
定期監視の利用手順は以下のとおりです。

- (1) smartmontools 設定ファイルの修正
- (2) 定期監視サービスの自動起動手順
- (3) 定期監視サービスの起動手順

手順について説明します。

(1) smartmontools 設定ファイルの修正
smartmontools の設定ファイルを修正します。
ご使用の環境に合わせて以下の設定ファイルを修正します。

【Red Hat Enterprise Linux 7.2 以上】

`/etc/smartmontools/smartd.conf`

【Red Hat Enterprise Linux 6.8 以上】

`/etc/smartd.conf`

修正内容を以下に記載します。

- ① smartd.conf ファイルの"DEVICSCAN" で始まる行の先頭に "#" を追加し、コメントアウトします。

```
# The word DEVICSCAN will cause any remaining lines in this
# configuration file to be ignored: it tells smartd to scan for all
# ATA and SCSI devices.  DEVICSCAN may be followed by any of the
# Directives listed below, which will be applied to all devices that
# are found.  Most users should comment out DEVICSCAN and explicitly
# list the devices that they wish to monitor.
# DEVICSCAN xxxxx
:
```

コメントアウト

- ② smartd.conf ファイルに定期監視するデバイスファイル名の追加
定期監視を行うデバイスファイルを任意の場所に記載します。
"-H" オプションを指定することで指定したデバイスファイルの障害検知時に syslog に
メッセージを出力します。

```
/dev/sda -H
```

(2) 定期監視サービスの自動起動手順
OS 起動時の自動起動設定を確認します。

【Red Hat Enterprise Linux 7.2 以上】

```
# systemctl is-enabled smartd.service
disabled
```

自動起動が有効の場合は「enabled」、無効の場合は「disabled」が表示されます。

【Red Hat Enterprise Linux 6.8 以上】

```
# chkconfig --list smartd
smartd          0:off  1:off  2:off  3:off  4:off  5:off  6:off
```

自動起動が有効の場合は「0:off 1:off 2:on 3:on 4:on 5:on 6:off」、
自動起動が無効の場合は「0:off 1:off 2:off 3:off 4:off 5:off 6:off」
が表示されます。

自動起動が無効になっている場合は、以下のコマンドを実行して自動起動を有効にします。

【Red Hat Enterprise Linux 7.2 以上】

```
# systemctl enable smartd.service
```

【Red Hat Enterprise Linux 6.8 以上】

```
# chkconfig smartd on
```

自動起動を無効にしたい場合は、以下のコマンドを実行します。

【Red Hat Enterprise Linux 7.2 以上】

```
# systemctl disable smartd.service
```

【Red Hat Enterprise Linux 6.8 以上】

```
# chkconfig smartd off
```

設定変更後に、再度、定期監視サービスの自動起動設定を確認します。

(3) 定期監視サービスの起動手順
定期監視サービスの起動を確認します。

【Red Hat Enterprise Linux 7.2 以上】

```
# ps -ef | grep smartd  
root    26290    1  0 11:08 ?        00:00:00 /usr/sbin/smartd -n -q never
```

【Red Hat Enterprise Linux 6.8 以上】

```
# ps -ef | grep smartd  
root    25732    1  0 13:59 ?        00:00:00 /usr/sbin/smartd -q never
```

定期監視サービスが停止している場合は以下のコマンドを実行して起動します。
定期監視サービスが起動済みの場合は必要ありません。

【Red Hat Enterprise Linux 7.2 以上】

```
# systemctl start smartd
```

【Red Hat Enterprise Linux 6.8 以上】

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

定期監視サービス起動後、再度、起動を確認します。

以上で定期監視の設定は完了です。

8.6. 本製品の障害解析情報

本製品運用中に何らかの障害が発生した場合は、情報採取を行ってください。

RootDiskMonitor 本体、StorageSaver for BootDisk 本体に情報採取ツールを用意しております。
ツールを使用して情報を採取される場合は、下記のマニュアルをご参照ください。

『CLUSTERPRO MC RootDiskMonitor 2.9 for Linux

CLUSTERPRO MC StorageSaver for BootDisk 2.9 (for Linux) syslog メッセージ一覧』

5.1. 障害解析情報収集ツール

CLUSTERPRO
MC RootDiskMonitor 2.9 for Linux

CLUSTERPRO
MC StorageSaver for BootDisk 2.9 (for Linux)

ユーザーズガイド

2024 年 4 月 第 13 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 2024

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