

CLUSTERPRO MC RootDiskMonitor

CLUSTERPRO MC StorageSaver for BootDisk (for Linux)

導入ガイド

HP-UX版
Linux版

第 18 版

2024年 4月

日本電気株式会社

目次

・ 本製品のねらい	P.3
・ こんな障害が発生したら…	P.4
・ 導入効果	P.5
・ 適用例	P.6
・ RootDiskMonitor機能紹介	P.7
・ 運用管理製品との連携	P.12
・ システム要件	P.13
・ RootDiskMonitor MirrorOption 機能紹介	P.14
・ 保守/サービス関連情報	P.15
・ 構築編	P.16
・ OSディスク監視の概要	P.17
・ 導入の前に	P.18
・ 商標	P.32

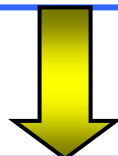
本製品のねらい

OSディスクを構成するI/Oパスの動作状況を定期監視し、I/Oパスに異常が発生した場合、syslogおよびコンソールにレポート通知を行うことで、OSディスク障害の早期発見・対応が可能になります。

また、クラスターウェアとの連携により、OSディスクが動作不能になった場合、ノードを切り替え、業務を継続することが可能になります。

こんな障害が発生したら…

OSディスクの故障やストールが発生!!

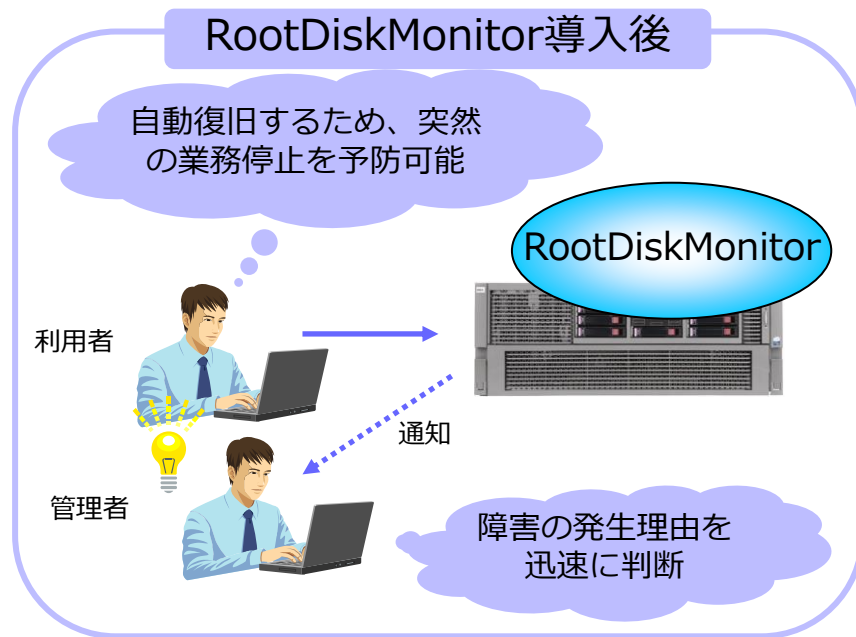
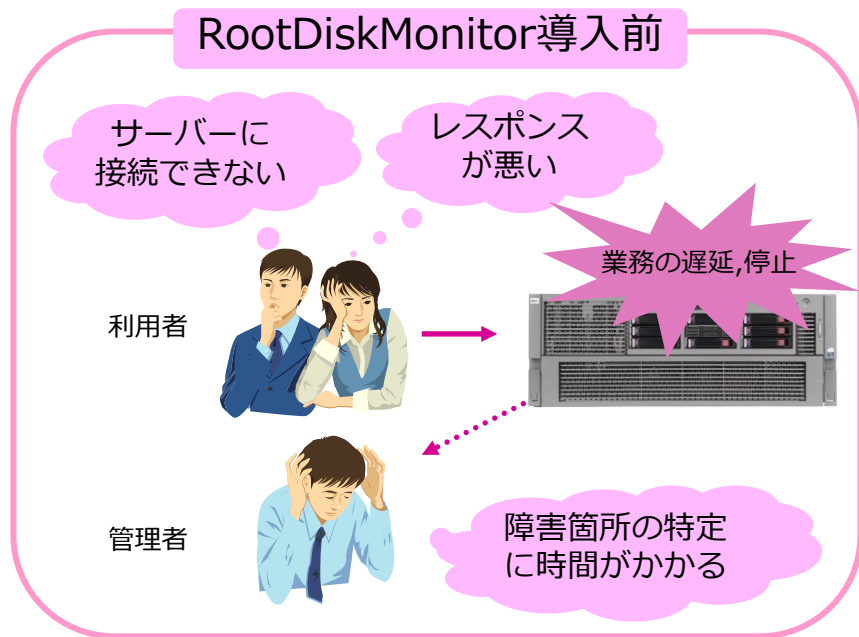


- 障害から復旧するまで、関連業務が停止する。
- 障害の深刻さによっては、バックアップされていないデータが消失する可能性がある。
- 原因の究明に人手と時間が必要になる。

導入効果

OSディスク障害が発生した場合

- 障害を検出すると、syslogおよびコンソールにレポートを通知します。
- ディスク障害が発生した場合、クラスターウェアと連携することにより、クラスターシステムでのノード切り替えを容易に実現します。
(クラスターシステムの可用性を向上)



RootDiskMonitor導入前

- 突然OSが停止し、原因の特定に人手・時間がかかった。
- 原因がOSディスク障害と判明したが、障害の特定に時間がかかったため、DBサーバーの停止時間が延びた。



RootDiskMonitor導入後

- 障害原因をsyslogおよびコンソールに通知するため、原因の特定が容易になり、障害復旧にかかる時間が短縮できる。
- クラスタウェアと連携することで、OSディスク障害時に自動的にノードを切り替えるため、DBサーバーの停止時間を最小に抑えることができる。

RootDiskMonitor機能紹介

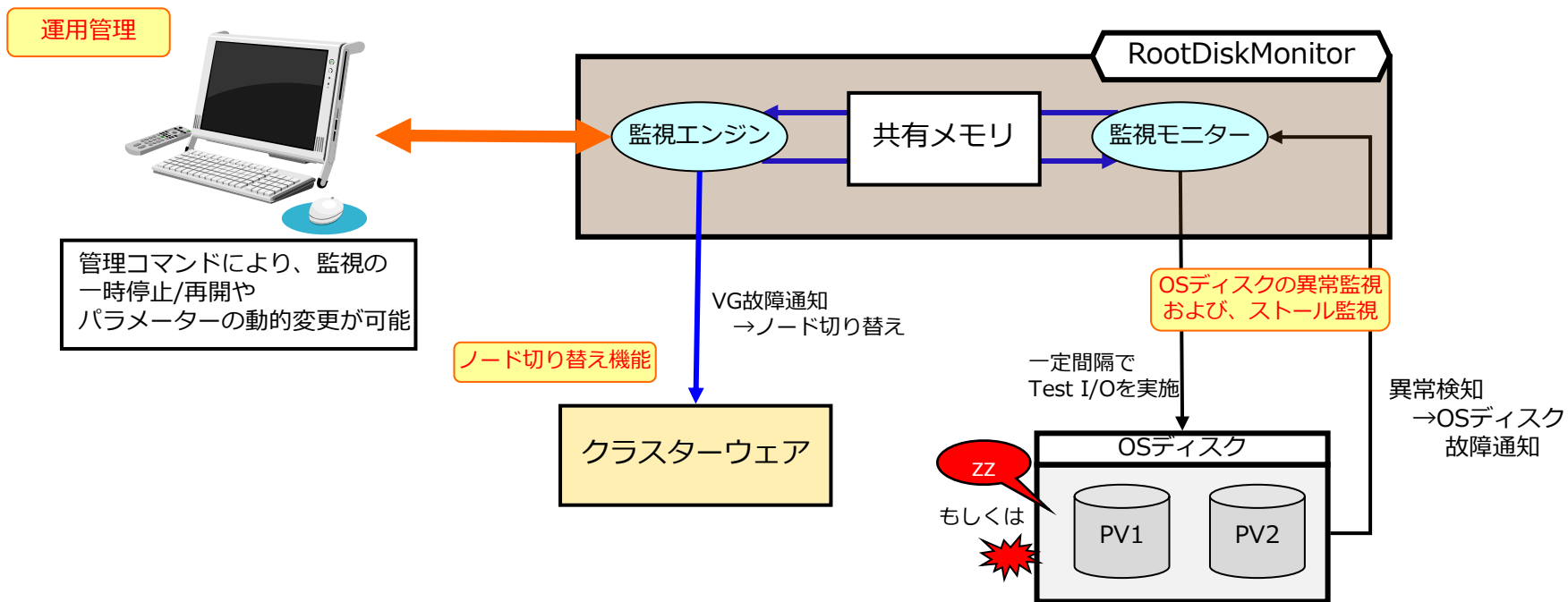
RootDiskMonitor は、OSディスクを構成する論理ディスク(LUN)の状態監視を行い、システムの高可用性を実現する製品です。以下の機能があります。

- (1) I/Oパス監視機能
- (2) クラスタウェア連携機能
- (3) 障害レポート機能

(1) I/Oパス監視機能

I/Oパス監視機能

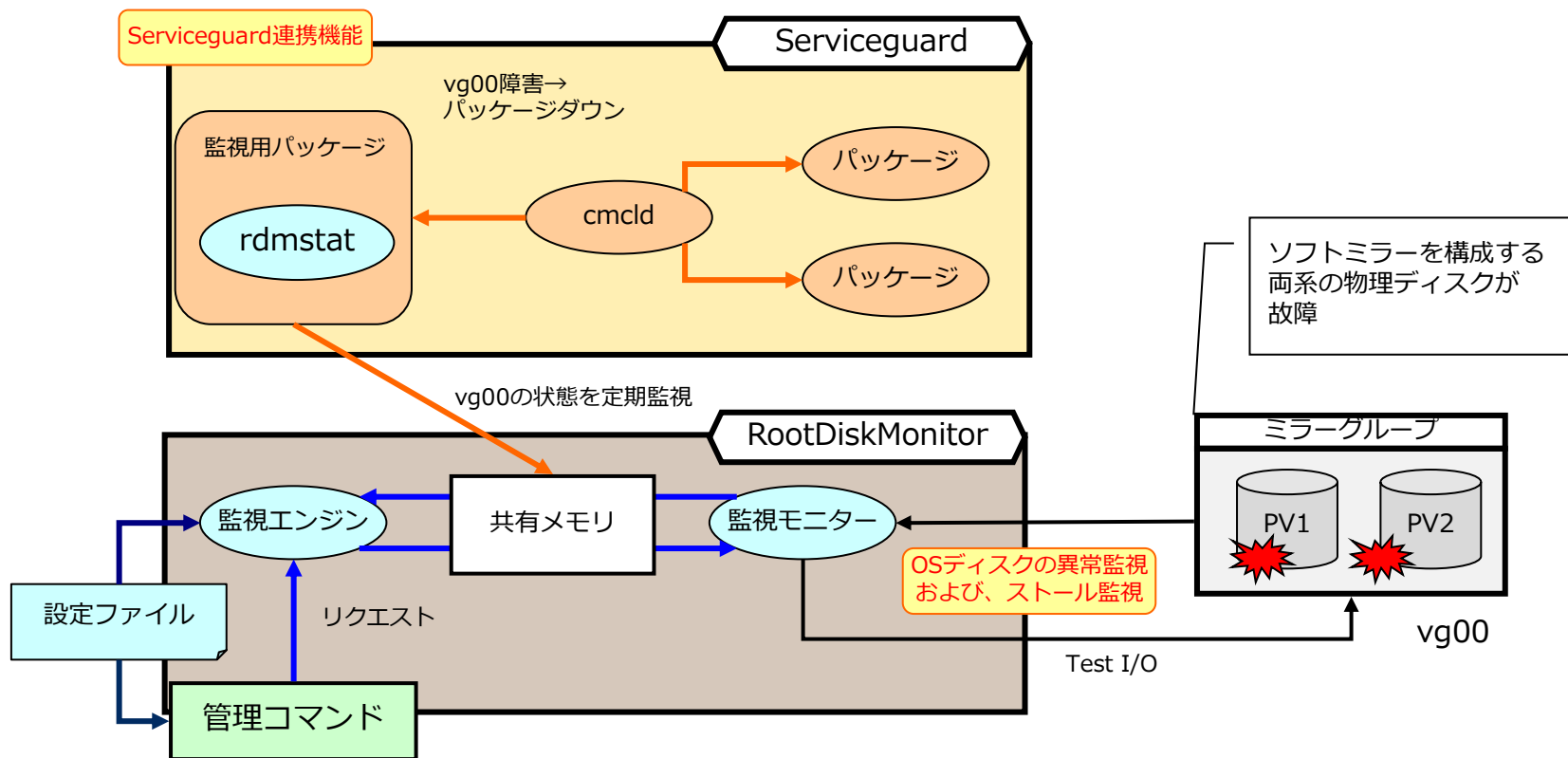
- OSディスクを構成する論理ディスク(LUN)に対して死活監視を行います。
- I/Oリクエストのストール監視を行います。
- シングル構成、ミラー構成のOSディスクの監視が可能です。



(2) クラスターウェア連携機能

Serviceguard連携機能(HP-UX版)

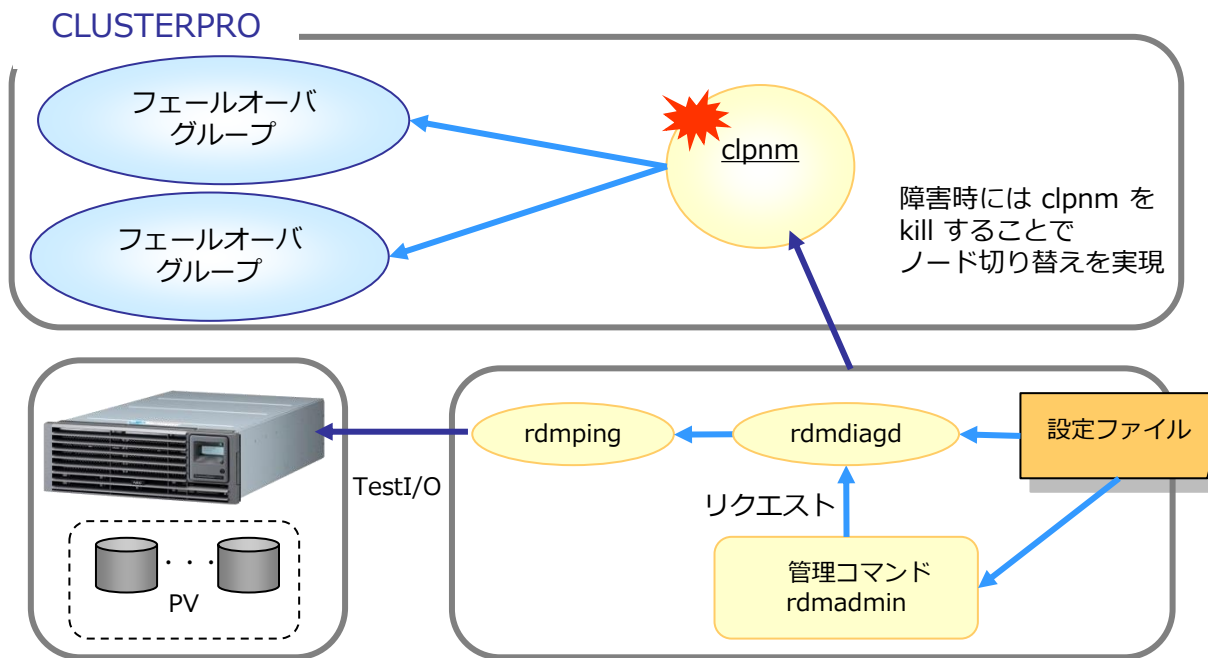
- OSディスクの動作状態をモニターするコマンド rdmstatをServiceguardのサービスコマンドとして登録することで、OSディスク障害時のノード切り替えを実現します。



(2)つづき

CLUSTERPRO連携機能(Linux版)

- OSディスクの状態を監視し、ディスク障害を検出するとCLUSTERPRO管理デーモン(clpnm)を強制停止することで、OSディスク障害時のノード切り替えを実現します。

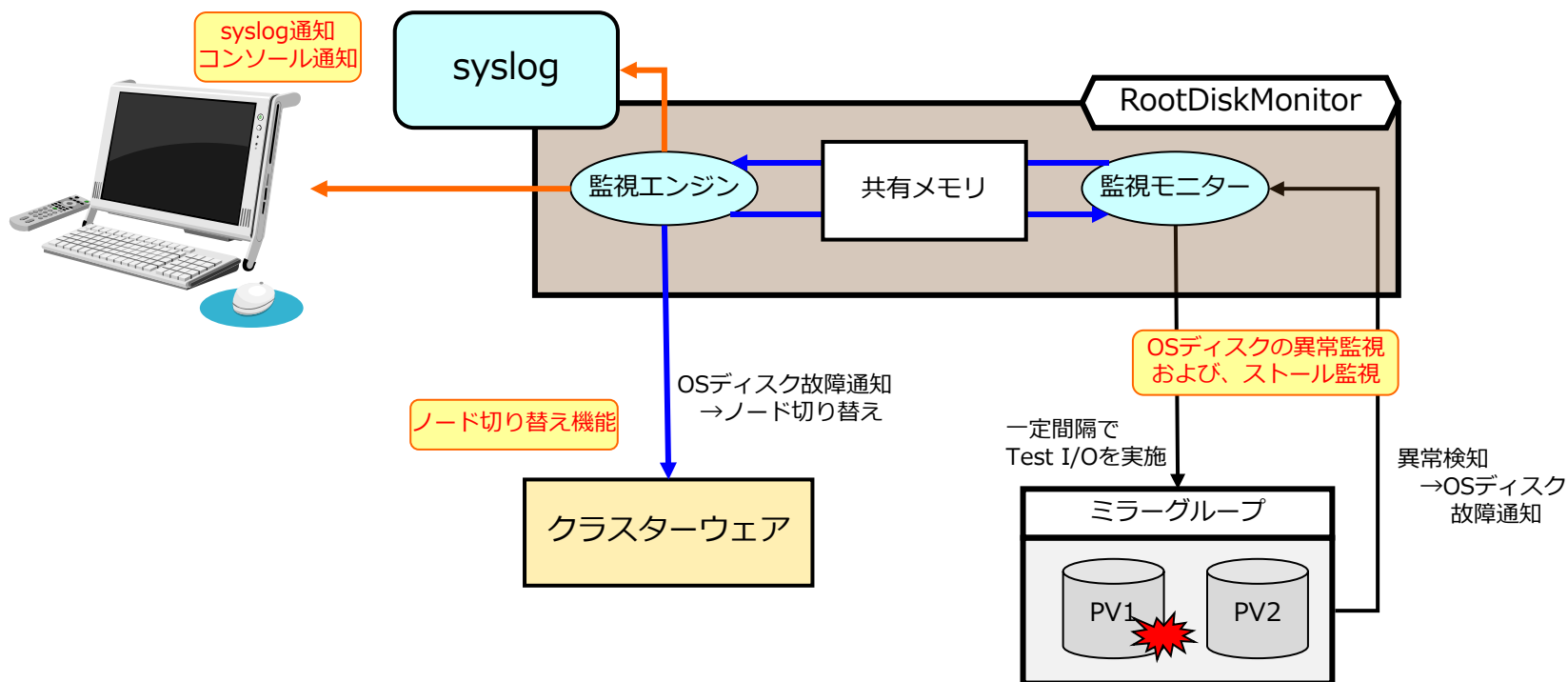


【 clpnm を強制終了する手法】

(3)障害レポート機能

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。



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 [HP-UX版]

- 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。

システム要件

■ HP-UX版

対応機種	NX7000 シリーズ、NX7700i シリーズ
対応OS	RootDiskMonitor R4.x … HP-UX 11i v3 RootDiskMonitor R3.1x … HP-UX 11i v3 RootDiskMonitor R2.1x … HP-UX 11i v1、 HP-UX 11i v2
使用ディスク容量	6Mバイト以上
使用メモリ容量	61Mバイト以上
対応ソフトウェア ※	Serviceguard A11.09 ~ A11.20 Serviceguard Extention for RAC A11.15 ~ A11.20

※クラスター構成の場合のみ必要になります。

■ Linux版

対応機種	以下のアーキテクチャのサーバー • x86_64
対応OS	Red Hat Enterprise Linux 6.0~6.10 Red Hat Enterprise Linux 7.0~7.9 Red Hat Enterprise Linux 8.0~8.9 Red Hat Enterprise Linux 9.0~9.3 Oracle Linux 6.2~6.10、 Oracle Linux 7.0~7.9、 Oracle Linux 8.0~8.9、 Oracle Linux 9.0~9.3、 Amazon Linux 2、 Amazon Linux 2023
使用ディスク容量	3Mバイト以上
使用メモリ容量	13Mバイト以上
対応ソフトウェア ※	CLUSTERPRO X 3.0 ~ X 3.3, X 4.0 ~ X 4.3, X 5.0 ~ X 5.2

※クラスター構成の場合のみ必要になります。

RootDiskMonitor MirrorOption 機能紹介(HP-UX版のみ)

RootDiskMonitor のオプション製品、RootDiskMonitor MirrorOption について説明します。

RootDiskMonitor MirrorOptionは、MirrorDisk/UX(LVM Mirroring) を使ったソフトミラー構成のOSディスクの監視を行い、システムの高可用性を実現する製品です。以下の機能があります。

- (1) メディアエラー検出機能 ※
- (2) ソフトミラー構成監視機能
- (3) パトロールシーク機能

※ HP-UX 11i v3 (NX7700i/5000 シリーズ) におけるメディアエラー検出機能は RootDiskMonitor MirrorOption R3.1b からサポートしました。

ただし、本機能は HP-UX 11i v3 0903 版以降で EMS Hardware Monitors をご使用の環境でのみ使用可能です。HP-UX 11i v3 0809 版以前のバージョンの OS を使用されている場合はメディアエラー検出機能は使用できません。使用するためには OS のバージョンを上げていただく必要があります。

また、NX7700i/7000 シリーズにおけるメディアエラー検出機能は RootDiskMonitor MirrorOption R4.1からサポートしました。

以下の環境に該当する場合は、NX7700i/7000 シリーズ用の RootDiskMonitor MirrorOption を使用してください。NX7700i/7000 シリーズ用の RootDiskMonitor MirrorOption は、他の NX7700i シリーズの機種でも動作可能です。

- NX7700i/5000 シリーズにおいて、HP-UX 11i v3 1104 版以降の DiagProdCollections に含まれる DASProvider をインストールしている環境
- NX7700i/7000, 8000 シリーズの環境

(注) 機能の詳細については、RootDiskMonitor MirrorOption の各種ドキュメントを参照してください。

(注) RootDiskMonitor MirrorOption をご利用いただくためには、RootDiskMonitor のライセンスとメディアが必要です。

PP・サポートサービス

- CLUSTERPRO MC シリーズのサポートは、保守アンバンドル製品をご購入の場合、PP・サポートサービスをご契約いただくか、PPSupportPack製品をご購入ください。詳細は以下をご参照ください。

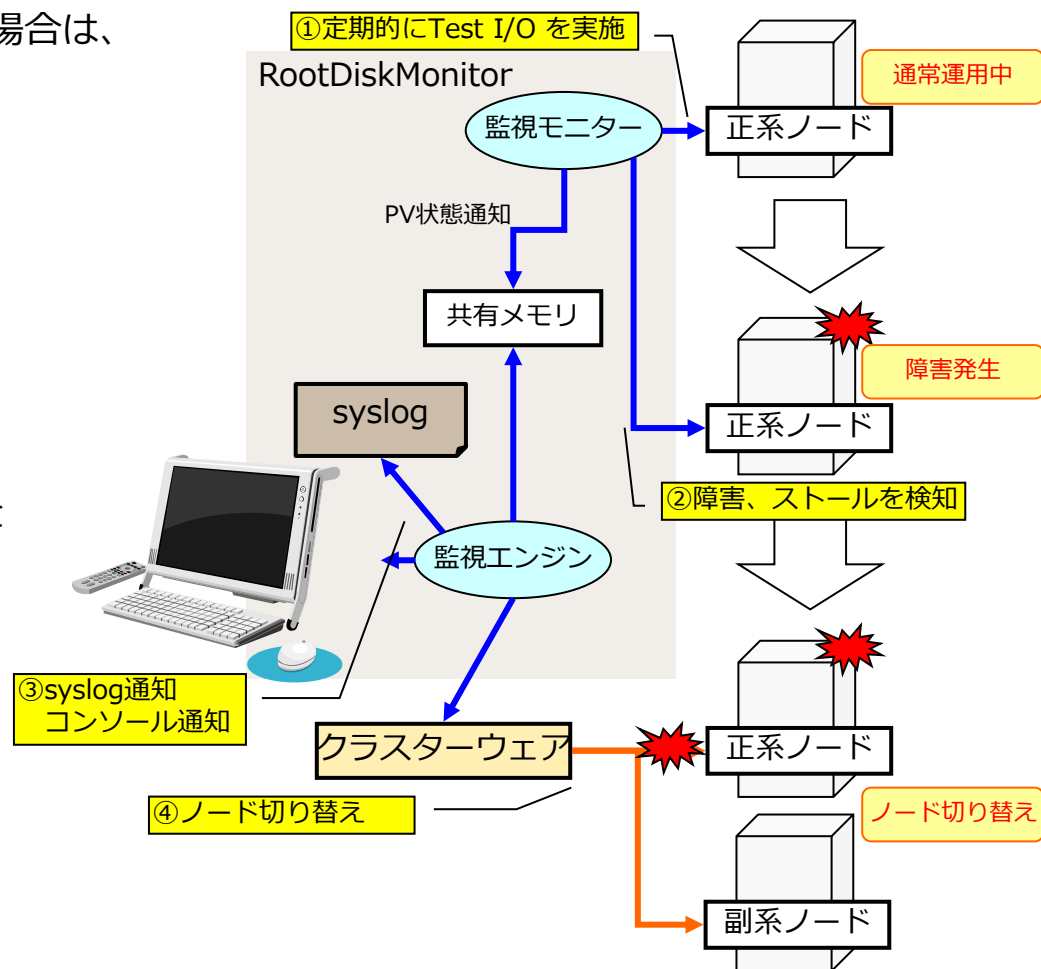
https://jpn.nec.com/clusterpro/mc_ha/support.html

構築編

OSディスク監視の概要

OSディスクに対し、定期的にTest I/O を実行することにより、死活監視、ノード自動切り替えを実現します。
また、クラスターウェアと連携している場合は、ノード自動切り替えを実現します。

- ① 定期的にTest I/Oを実施することで、運用中のノードの状態を監視します。
- ② 対象ノードの障害、ストール状態を検知します。
- ③ 障害を検知すると、syslogおよびコンソールにレポート通知を行い、障害情報を通知します。
設定により、障害が発生したディスクを自動的に切り離すこともできます。
- ④ クラスターウェアと連携している場合、自動的にノードを切り替えます。



RootDiskMonitorの導入について

RootDiskMonitorを使用するには、設定ファイルの作成が必要です。

(1)設定ファイルの構成

- ・ 監視ルールについて
- ・ デバイス定義について

(2)設定ファイルの自動生成

(1)設定ファイルの構成

設定ファイル作成時には、以下のキーワードを参照してください。
設定ファイル名は /opt/HA/RDM/conf/rdm.config となります。

■ HP-UX版、Linux版共通

[監視ルール]	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。 指定値:30 秒～、デフォルト値:60 秒 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。
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。 指定値:60 秒～、デフォルト値:360 秒 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。 例外として、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	TestI/O の発行間隔を指定します。 指定値:1 秒～、デフォルト値:5 秒 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。
WAIT_TESTIO_INTERVAL	TestI/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。 指定値:1 秒～、デフォルト値:5 秒 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。

(注)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。

■ HP-UX版、Linux版共通

[監視ルール]	オプション	説明
TESTIO_MODE	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)
	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。 ※Linux版のみです。
	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。

(注) 本パラメーターは、HA/RootDiskMonitor R4.4 (HP-UX版)以降または CLUSTERPRO MC RootDiskMonitor 1.1 for Linux 以降で使用可能です。

■ HP-UX版

[監視ルール]	オプション	説明
VG_STALL_ACTION	ACTION_NONE	I/O ストール検出時のアクション指定を行いません。(デフォルト) 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。
	TOC_EXEC	I/O ストール検出時に Serviceguard の管理デーモンプロセスである cmclد を強制終了することでノードを切り替えます。 本オプションは Serviceguard を利用したクラスターシステムでのみ有効です。

(注) 本パラメーターは、変更しないことを推奨します。

[監視ルール]	オプション	説明
OVER_ACTION	ACTION_NONE	OS ディスク異常検出時のアクション指定を行いません。(デフォルト) Serviceguard と連携したノード切り替えを行わない場合や、サービスコマンドによる Serviceguard パッケージ連動でノード切り替えを行う場合はこの値を指定します。
	TOC_EXEC	OS ディスク異常検出時に、Serviceguard の管理デーモンプロセスである cmclد を強制終了させることによりノード切り替えを行います。 本オプション は Serviceguard を利用したクラスターシステムでのみ有効です。

(続き)

(注)

Serviceguard との連携方式を変更すると、設定ファイルの変更が必要になる場合があります。デフォルトの監視ルールから、cmclد を強制終了する手法によるノード切り替えを行うルールに変更した場合、下記のパラメーターを変更してください。

OVER_ACTION

ACTION_NONE -> TOC_EXEC に変更

また、cmclد を強制終了する手法によるノード切り替えを行うルールから、デフォルトの監視ルールに変更した場合、下記のパラメーターを変更してください。

OVER_ACTION

TOC_EXEC -> ACTION_NONE に変更

■ HP-UX版

[監視ルール]	オプション	説明
TESTIO_FAULT_ACTION	ACTION_NONE	TestI/O 異常検出時のアクション指定を行いません。(デフォルト) TestI/O 失敗によるvg の異常を検出すると、syslog にメッセージを出力します。
	REDUCE_EXEC	TestI/O 異常検出時に異常を検知したディスクの切り離しを行います。 TestI/O 失敗によるvg の異常を検出すると、障害発生ディスクの切り離し を行います。 本オプションは、SAN boot 環境の場合、内蔵ディスクがPV-Link 構成の場合は使用できません。

(注) 本パラメーターは、変更しないことを推奨します。

(注) 本パラメーターは、RootDiskMonitor R3.1 (HP-UX版)以降で使用可能です。

■ HP-UX版 RootDiskMonitor R3.1x , R4.x (HP-UX版)の場合

[デバイス定義]	説明
VG	OS ディスクのボリューム名を指定します。 通常 /dev/vg00 (vgxx) または、/dev/vx/rootdg が有効となります。 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。
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	I/O パスへの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 esdisk (LegacyView の場合、sdisk)で定義されるキャラクター型のファイル名を指定し、スペース区切りで lunpath H/W Path を指定します。 (例) PersistentDSF (Agile View) 構成 /dev/rdisk/disk1_p2 0/1/1/0.0x0.0x0 LegacyDSF (Legacy View) 構成 /dev/rdisk/c3t0d0s2 0/1/1/0.0x0.0x0 ※ lunpath H/W Path は以下のコマンドで確認することができます ① # ioscan -Nfnk <disk名> で H/W Path を確認 (例) # ioscan -Nfnk /dev/rdisk/disk1_p2 (LegacyView 構成の場合、N は不要) disk 1 <u>64000/0xfa00/0x2</u> esdisk CLAIMED DEVICE HP 73.4GMAS3735NC ② # ioscan -m hwpath -H <①で確認した H/W Path> (例) # ioscan -m hwpath -H 64000/0xfa00/0x2 64000/0xfa00/0x2 <u>0/1/1/0.0x0.0x0</u> 0/1/1/0.0.0 ・シングルパス構成の場合は、LUN の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 のスペシャルファイルを両方指定してください。 ・FC接続の代替パス構成の場合は、LUNへの正副I/Oパス(PV-link)のスペシャルファイルを両方指定してください。

■ HP-UX版 RootDiskMonitor R2.1x(HP-UX版)の場合

[デバイス定義]	説明
VG	OS ディスクのボリューム名を指定します。 通常 /dev/vg00 (vgxx) または、/dev/vx/rootdg が有効となります。 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。
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	I/O パスへの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 sdisk で定義される キャラクター型のファイル名を指定します。(例 /dev/rdisk/c3t0d0) ・シングルパス構成の場合は、LUN の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 のスペシャルファイルを両方指定してください。 ・FC 接続の代替パス構成の場合は、LUN への正副 I/O パス(PV-link)のスペシャルファイルを 両方指定してください。

■ Linux版

[監視ルール]	オプション	説明
VG_STALL_ACTION	SERVICE_CMD_DISABLE	I/O ストール検出時のアクション指定を行いません。(デフォルト) 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。 CLUSTERPRO と連携したノード切り替えを行わない場合はこの値を指定します。
	SERVICE_CMD_ENABLE	I/O ストール検出時に CLUSTERPRO のカスタムモニタリソースと連携してノードを切り替えます。
	CLPNM_KILL	I/Oストール検出時にCLUSTERPROの管理デーモンプロセスである、clpnmプロセスを強制終了することでノードを切り替えます。 本オプションはCLUSTERPROを利用したクラスターシステムでのみ有効です。
	TOC_EXEC	I/O ストール検出時に OSシステムダンプの採取とOS強制停止、リブートを実施します。 CLUSTERPROを利用している場合ノードを切り替えることができます。
	POWER_OFF	ソフトウェアwatchdogを利用し、OSを停止します。

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。
- 本パラメーターは変更しないことを推奨します。

■ Linux版

[監視ルール]	オプション	説明
OVER_ACTION	SERVICE_CMD_DISABLE	OS ディスク異常検出時のアクション指定を行いません。(デフォルト) OS ディスク異常を検出すると syslog にメッセージを出力します。 CLUSTERPRO と連携したノード切り替えを行わない場合はこの値を指定します。
	SERVICE_CMD_ENABLE	OS ディスク異常検出時に、CLUSTERPRO のカスタムモニタリソースと連携してノードを切り替えを行います。
	CLPNM_KILL	OS ディスク異常検出時に、CLUSTERPROの管理デーモンプロセスである clpnm を強制終了させることによりノード切り替えを行います。 本オプション はCLUSTERPROを利用したクラスターシステムでのみ有効です。
	TOC_EXEC	OSディスク異常検出時に OSシステムダンプの採取とOS強制停止、リブートを実施します。 CLUSTERPROを利用している場合ノードを切り替えることができます。
	POWER_OFF	ソフトウェアwatchdogを利用し、OSを停止します。

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。

CLUSTERPRO MC RootDiskMonitor 2.9 for Linuxの場合

[デバイス定義]	説明	
VG	OS ディスクのボリューム名を指定します。 通常ボリュームグループ名 /dev/VolGroupXX が有効となります。(Xはシステムにより任意) 複数のボリュームを監視する場合は VG-MIRROR-PV の組み合わせを指定してください。 OSディスクボリュームがLVM構成ではない場合は、VG_NONEを指定してください	
VOLTYPE	ディスクの種別を指定します。 ※本パラメータ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。	
	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	I/O パスへの udev デバイスファイル名またはブロックスペシャルファイル名を指定します。 <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ブロック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 の udev デバイスファイルまたはブロックスペシャルファイルを両方指定してください。	

(2)設定ファイルの自動生成

設定ファイル自動生成コマンド `rdmconfig`

上記コマンドにより、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します。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。

“Y” を選択した場合、設定ファイルは上書きされます。

“N” を選択した場合、設定ファイルの自動生成を中止します。

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必ず確認してください。

注意(Linux版のみ):

- SANBoot 構成は設定ファイル自動生成 (`rdmconfig`) のサポート対象外です。サンプルファイル (`/opt/HA/RDM/conf/rdm.config.default`) が提供されていますので `/opt/HA/RDM/conf/rdm.config` というファイル名でコピーし、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。

- 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。
- CLUSTERPRO は、日本電気株式会社の登録商標です
- 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。

 **Orchestrating** a brighter world

NEC