

FUJITSU Server PRIMEQUEST 3000シリーズ

Linux設計ガイド

～Red Hat Enterprise Linux編～

はじめに

1. Linuxの概要
2. Linuxのシステム設計
3. 高可用設計
4. OS運用管理設計

■ 本書の読み方

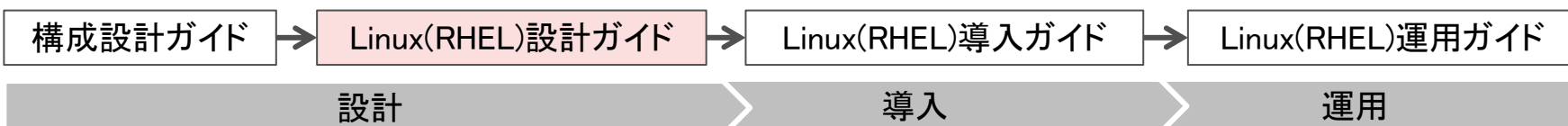
■ 本書の内容

PRIMEQUEST 3000シリーズにRed Hat Enterprise Linux (RHEL)を導入して使用される方を対象に、OS設計の概要、留意事項などについて記載しています。

- ・ 操作方法の詳細については、PRIMEQUEST 3000シリーズ本体およびRHELのマニュアルを参照してください。
- ・ ソフトウェアの動作環境は、各ソフトウェアのサイトを参照してください。

■ ガイド間の位置づけ

PRIMEQUEST 3000シリーズのガイドの位置づけです (RHEL使用時)。



■ 本文中の記号

本文中に記載されている記号には、次のような意味があります。

記号	意味
→	参照ページや参照ドキュメントを示しています。

■ 本文中の略称

名称	略称	
PRIMEQUEST 3400S Lite	PRIMEQUEST 3000シリーズ PRIMEQUEST	
PRIMEQUEST 3400S		
PRIMEQUEST 3400E		
PRIMEQUEST 3400L		
PRIMEQUEST 3800E		
PRIMEQUEST 3800L		
マネジメントボード	MMB	
システムボード	SB	
Red Hat® Enterprise Linux® 7 (for Intel64)	RHEL7 (for Intel64) RHEL7	RHEL Linux
ETERNUS SF AdvancedCopy Manager	ETERNUS SF ACM、ACM	

■ 本文中の略称

名称	略称
PRIMEQUEST 3000シリーズ システム構成図	システム構成図
PRIMEQUEST 3000シリーズ 製品概説	製品概説
PRIMEQUEST 3000シリーズ 導入マニュアル	導入マニュアル
Red Hat Enterprise Linux 7 Linuxユーザーズマニュアル (*1)	Linux(RHEL)ユーザーズマニュアル
ServerView Suite ServerView Installation Manager 取扱説明書	SVIMマニュアル
PRIMEQUEST 3000シリーズ 構成設計ガイド	構成設計ガイド
PRIMEQUEST 3000シリーズ Linux導入ガイド ～Red Hat Enterprise Linux編～	Linux(RHEL)導入ガイド
PRIMEQUEST 3000シリーズ Linux運用ガイド ～Red Hat Enterprise Linux編～	Linux(RHEL)運用ガイド
PRIMEQUEST 3000シリーズ クラスタ構成設計ガイド(Linux/PRIMECLUSTER編)	クラスタ構成設計ガイド(Linux/PRIMECLUSTER編)

(*1) 参照するにはSupportDesk契約が必要です。

■ 保守サービスについて

富士通では、お客様に安心してRed Hat Enterprise Linuxを使用していただくために、有償サポートサービス (SupportDesk Standard) を御用意しております。有償サポートサービスでは、Red Hat Enterprise Linuxに関する御質問、インストールや運用の際に発生する疑問やトラブルなどの問い合わせについて、富士通サポートセンター (OSC: One-stop Solution Center) で一括対応いたします。

また、この有償サポートサービスには、Red Hat Enterprise Linux を利用するために必要なサブスクリプション(利用権)が含まれています。

※サブスクリプションは、Red Hat Enterprise Linux を利用する権利であり、Red Hat 社からインストールイメージ (ISOファイル形式) やセキュリティアップデート、その他アップデートを入手するための権利を含みます。

Red Hat Enterprise Linuxの御利用に際しては、有償サポートサービスの御契約をお勧めします。有償サポートサービス (SupportDesk Standard) については、以下をご覧ください。

⇒ <http://www.fujitsu.com/jp/services/infrastructure/service-desk/menu/standard/supportdesk-standard/linux/>

- ・ Linux®は米国及びその他の国におけるLinus Torvaldsの登録商標です。
- ・ Red Hat、Red Hat Enterprise Linuxは米国およびその他の国において登録されたRed Hat,Inc.の商標です。
- ・ Intelは、アメリカ合衆国および/またはその他の国におけるIntel Corporationの商標です。
- ・ Microsoft、Windows、Windows Serverは、米国Microsoft Corporationの米国およびその他の国における商標または登録商標です。
- ・ OracleとJava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。
- ・ NetVaultは、米国、日本およびその他の国におけるDell,Inc.の商標もしくは登録商標です。
- ・ PRIMECLUSTERは、富士通株式会社の登録商標です。
- ・ その他、会社名と製品名はそれぞれ各社の商標、または登録商標です。

1. Linuxの概要

Linuxの概要について説明します。

1.1 利用可能なディストリビューション

製品名
Red Hat® Enterprise Linux® 7 (for Intel64)

■ 最新情報は、以下を参照

⇒ 『システム構成図』

⇒ 「カーネル・バージョン (PRIMEQUEST)」

<http://www.fujitsu.com/jp/products/computing/servers/primequest/products/3000/os/linux/support/kernel/>

■ 使用するディストリビューションにより、使用可能なソフトウェアが異なる場合がある

⇒ 詳細は各ソフトウェアの情報を参照

1.2 RHEL7の特長と機能

■ RHEL7で追加、強化された機能

- ファイルシステム機能の向上
 - XFSをデフォルトのファイルシステムとしてサポート
- 管理機能強化
 - systemdによるinitやログ機能を統合したシステム管理機能をサポート
- 仮想マシン機能の強化
 - KVMの性能改善
- pNFS (Parallel NFS) clientの強化
 - DBアクセスなどI/O負荷の高い処理のパフォーマンスが向上
- オンラインデフラグ
 - オンライン(マウント状態)でフラグメントを解消(対象: ext4、XFS)
- コマンド
 - ssmコマンド(system-storage-managerパッケージ)のサポート
 - 各論理ボリュームや各ファイルシステムの差異を吸収した管理ツール
 - snapperコマンド(snapperパッケージ)のサポート
 - 論理ボリュームのスナップショットを統合管理

1.3 サポートソフトウェア

PRIMEQUESTがサポートする、Linuxディストリビューションに含まれるソフトウェア

分類	用途	ソフトウェア名称
Linux基本システム	Linuxカーネル	RHEL7 : kernel 3.10系
	ブートローダ	RHEL7 : GRUB2
	標準Cライブラリ	glibc
	仮想マシン機能	KVM
サーバソフトウェア	スーパーサーバ	RHEL7 : systemd
	WWWサーバ	apache
	NFSサーバ	NFS
	NISサーバ	NIS
	PCファイルサーバ	samba
	リモートログイン	in.telnetd, telnet
	ファイル転送	vsftpd, ftp
	メール受信サーバ	dovecot, imap, pop
	メール配送エージェント	procmail
	メール送信サーバ	RHEL7 : postfix

分類	用途	ソフトウェア名称
サーバソフトウェア	DNSサーバ	bind
	DHCPサーバ	DHCP
	キャッシュサーバ	squid
	セキュアシェル	ssh
	標準時間サーバ	RHEL7 : chrony, ntp
	ファイアウォール	RHEL7 : firewalld, iptables
	ネットワーク監視	net-snmp
	論理ボリュームマネージャ	lvm2
	ディスク制限	quota
	ログ分析システム	logwatch, logrotate
	ディレクトリサーバ	OpenLDAP
	ソフトウェアRAID	multiple device(md)
	マルチパス機能	device mapper multipath(dm-mp)
デスクトップ	Xウィンドウ	X.Org

⇒ 「Linuxディストリビューション・サポートソフトウェア」
(<http://www.fujitsu.com/jp/products/computing/servers/primequest/products/3000/os/linux/support/software/>) を参照

1.4 富士通のLinuxサポート (1/3)

■ 検証とサポート範囲

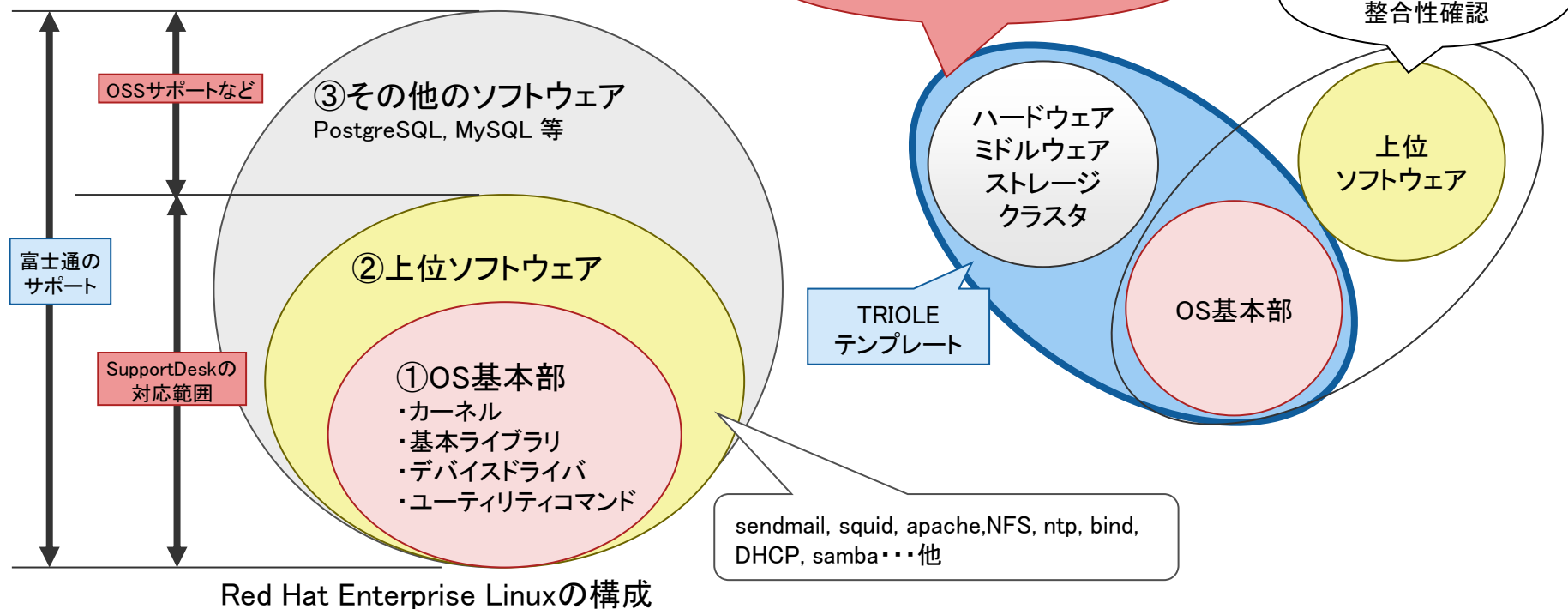
- 基幹システムに必要な十分な範囲について徹底的な整合性検証を実施
- サポート範囲において、お客様のトラブル解決を支援

・ OS基本部

- ・ ソフトウェア検査部門による、富士通品質基準に基づく整合性検証の実施
- ・ ハードウェア、OS、ミドルウェア、ストレージ、およびクラスタを組み合わせた整合性検証を実施し、TRIOLEテンプレート化

・ 上位ソフトウェア

- ・ OS基本部との整合性確認を実施



Red Hat Enterprise Linuxの構成

1.4 富士通のLinuxサポート (2/3)

■ お客様の要望に応えるサポート内容を用意

■ SupportDesk Standard (標準サポート)

標準でEUS (Extended Update Support) を提供

- ・ 特定のマイナーリリースに対するerrata提供期間を拡張
 - ・ RHEL6 (全マイナーリリース) : リリース後、24か月間
 - ・ RHEL7 (7.0を除く全マイナーリリース) : リリース後、24か月間
- ・ マイナーリリースのアップデートに対し、余裕を持った準備期間の確保が可能

■ SupportDesk Standard (拡張プラスサポート)

EUSに加え、AUS (Advanced Mission Critical Update Support) を提供

- ・ 特定のマイナーリリースに対するerrata提供期間をEUSよりさらに拡張
 - ・ RHEL6 (6.4,6.5,6.6) : リリース後、6年間
 - ・ RHEL7 (7.2) : リリース後、6年間
- ・ マイナーリリースのアップデート回数を最小限に抑えることが可能
- ・ 原則として無停止での稼働が求められるシステム向け

■ Linuxシステムの円滑な運用に向けたツールの提供

■ 富士通 Linux サポートパッケージ (FJ-LSP)

システム情報採取ツール、メモリダンプ支援ツール、性能情報一括採取ツールなどのツールを提供

⇒ 詳細は、『Linux(RHEL)ユーザーズマニュアル』を参照

1.5 ハードウェア添付ソフト

ハードウェア添付ソフトは必ず導入する

※SVIM (ServerView Installation Manager) を使用する場合、
ハードウェア添付ソフトが自動的に導入される

⇒ 詳細は、『製品概説』の「添付ソフトウェア」を参照

2. Linuxのシステム設計

Linuxの各設定項目の事前設計について説明します。

詳細は、『Linux(RHEL)ユーザーズマニュアル』を参照してください。

2.1 Linuxシステム設計の流れ

- SVIMを使用したインストール項目の設計
 - ⇒「[2.2 SVIMを使用したインストール項目の設計](#)」を参照
 - パッケージ選択の設計項目
 - 認証の設計項目
- セキュリティの設計
 - ⇒「[2.3 セキュリティの設計項目](#)」を参照
- ハードディスクの運用設計
 - ⇒「[2.4 ハードディスクの運用設計](#)」を参照
 - ディスクパーティション設計項目
 - 各パーティションの容量見積り
- ネットワークの運用設計
 - ⇒「[2.5 ネットワークの運用設計](#)」を参照
- デバイス名ずれ対策設計
 - ⇒「[2.6 デバイス名ずれ対策設計](#)」を参照
 - デバイス名ずれの原因と対策
- ログファイルの運用設計
 - ⇒「[2.7 ログファイルの運用設計](#)」を参照
 - ログローテーションの必要性
- ダンプ環境設計
 - ⇒「[2.8 ダンプ環境設計](#)」を参照
 - メモリダンプ機能

■ 設計のポイント

- スワップ領域、ダンプデバイス、ダンプ退避域はメモリ容量によって見積る
領域の設定はガイドモードで行う

- rootのパスワードを必ず設定する

- ブートローダのパスワードを必ず設定する

SVIMを使用してクイックモードでインストールした場合、パスワードは設定されないガイドモードでインストールする場合にはパスワードの設定が可能

セキュリティ確保のため、インストール完了後にパスワードを変更すること

- DNSサーバを必要に応じて設定する

■ SVIMでインストールできるツール

ServerView Suite (監視ソフトウェア)、FJ-LSP(*)、そのほかツール類

(* SVIMからインストールの誘導を行う)

⇒ 詳細は、『SVIMマニュアル』を参照

■ パッケージ選択の設計項目

以下のパッケージがデフォルトで選択済み

- X Window System
- GNOME Desktop Environment

■ 認証の設計項目

■ MD5を使う

ユーザーパスワードに対して、MD5を使用。通常はMD5を有効にする

MD5を使用しない場合は、DESが使用され、パスワードは8文字までの英数字に制限される

※ MD5: 認証やデジタル署名に使用されるハッシュ関数。ファイルの改ざん検知が可能

DES: 56ビットの秘密鍵暗号化アルゴリズム

■ シャドウパスワードを使う

ユーザーパスワードに対して、シャドウパスワードを使用。通常はシャドウパスワードを有効にする

※ シャドウパスワード: /etc/shadowに暗号化したユーザーパスワードを格納し、root権限だけにファイルの参照権限を許可。/etc/passwdには、「x」などの文字でユーザーパスワードを格納

■ ファイアーウォール

別のコンピュータからのアクセスを許可するプロトコルやポート番号を設定する

⇒ 詳細は、『SVIMマニュアル』を参照

2.3 セキュリティの設計項目

以下の対策を行う

■ OSのセキュリティ対策

⇒ 詳細は、Red Hat カスタマーポータル (<https://access.redhat.com/>)の「製品およびサービス」→「ドキュメント」→「Red Hat Enterprise Linux」から『セキュリティガイド』を参照

■ 修正適用

⇒ 詳細は、「[4.1 修正適用の概要](#)」を参照

※セキュリティソフト (ISV製品) の情報は、以下のURLを参考にしてください。

<http://www.fujitsu.com/jp/products/computing/servers/partner/product-line/security/>

2.4 ハードディスクの運用設計 (1/2)

■ ディスクパーティション設計項目

設計項目	説明
マウントポイント	各ディスクパーティションのパスを設定する ・ / :root ディレクトリ ・ /boot :ブートストラップのプロセス中に使用されるファイルと、OS のカーネル用の領域 ・ swap :システムメモリの容量が足りないとき、仮想記憶域として使用される領域
ファイルシステムタイプ	ファイルシステムタイプを設定する。以下に主なファイルシステムを示す 【RHEL6で利用できるファイルシステム】 ext2: ext3が登場するまで、多くのディストリビューターでデフォルトのファイルシステムとして採用されていたファイルシステム ext3: ext2にジャーナリング機能を加えたファイルシステム 最大ファイルシステムサイズが16TB、最大ファイルサイズが2TB ext4: ext3と同様のジャーナリング機能を備えたファイルシステム 最大ファイルシステムサイズが16TB、最大ファイルサイズが16TB 【RHEL7で利用できるファイルシステム】 XFS : RHEL7からデフォルトになったファイルシステム 最大ファイルシステムサイズが500TBの大容量のファイルシステム ext4: 最大ファイルシステムサイズが50TB、最大ファイルサイズが16TB ext2、ext3: RHEL7でも利用可能。
ドライブ	ディスクパーティションを作成するドライブを設定する 例) sda: 1台目の内蔵ハードディスクドライブ sdb: 2台目の内蔵ハードディスクドライブ

⇒ 詳細は、『Linux(RHEL)ユーザーズマニュアル』を参照

2.4 ハードディスクの運用設計 (2/2)

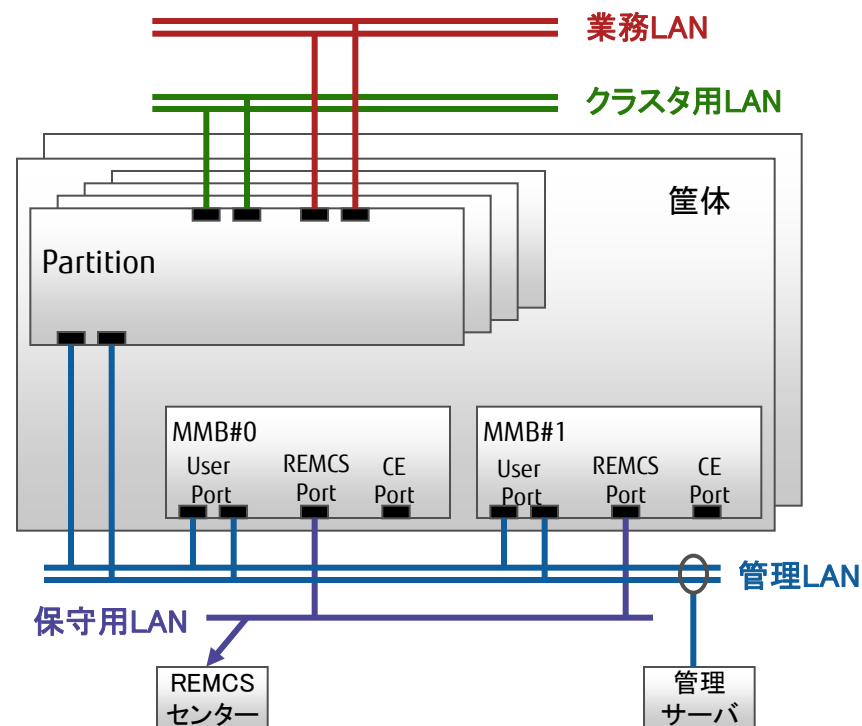
■ 各パーティションの容量見積り

パーティション	推奨容量	ポイント
	RHEL7	
/ (rootディレクトリ)	10GB以上	オプションソフトウェアなどによる機能追加や修正パッチの適用、ログファイルの保存に対応できる容量を確保する
/ boot	500MB以上	—
swap	<ul style="list-style-type: none">・メモリが2GB以下:メモリ容量の2倍・メモリが2GB超から8GB以下:メモリ容量と同量・メモリが8GB超から64GB以下:メモリ容量の0.5倍・メモリが64GB超:4GB	システムのメモリ負荷に応じて設定

⇒ 詳細は、『Linux(RHEL)ユーザーズマニュアル』を参照

■ 設計のポイント

- 管理LAN用のNIC (Network Interface Card) と業務LAN用のNICのOS認識順番に注意する
- 管理LAN(MMBとの通信に使用)に接続するスイッチのSTP機能をオフにする
- 各LANのIPアドレスは、以下の点を考慮して割り当てる
 - 管理LAN、業務LANを分離する場合は、サブネットを分ける
 - コンソールリダイレクションのIPアドレスは、管理LANと同一サブネットを設定する



2.6 デバイス名ずれ対策設計

■ デバイス名ずれの原因と対策

種類	要因	原因(例)	対策
ディスク系 デバイス名ずれ	ソフトウェア変更	<ul style="list-style-type: none">カーネルアップデートをした時のドライバ適用状況	<ul style="list-style-type: none">ハードウェアによる防止策 (RAID構成、ハード故障時のシステム起動の自動停止)OS機能による防止策 (ラベル名の利用、udev機能の利用、UUIDの利用)ミドルウェア機能による防止策 (PRIMECLUSTER GD などの製品の導入)
	ハードウェア構成変更	<ul style="list-style-type: none">ディスク(リムーバブルディスク)、ストレージ装置の増設または減設PCI スロットへのカード増設または減設	
	ハードウェア故障	<ul style="list-style-type: none">ディスク故障カード故障ストレージ装置故障、電源断、ケーブル故障	
ネットワーク系 デバイス名ずれ	NICの増設または故障	<ul style="list-style-type: none">NICの増設または故障時にネットワークデバイス名が変更される	<ul style="list-style-type: none">MACアドレス指定による、ネットワークデバイス名の固定

⇒ 詳細は、『Linux(RHEL)ユーザーズマニュアル』を参照

■ ログローテーションの必要性

ログのロストを避けるために、複数のログファイルをローテーションして運用

■ システムログファイルの運用

設計項目	内容
実施方法	ローテーション運用を行うタイミングでlogrotate コマンドを起動 以下の起動タイミングから選択 <ul style="list-style-type: none">• コマンドライン入力やスクリプトから任意のタイミングで起動• スケジュール設定による定期的な起動(cron、Systemwalkerなどのスケジュール管理プログラムを使用)
スケジュール	システムログファイル出力の負荷が高い時間帯を避け、業務に影響しない時間帯に実施する
切替え周期と保有世代数	システムログファイルの総量がシステム運用に適切になるよう設計する (デフォルトは週に1回の切替え (Weekly) と4世代保持(ログ切替え後のファイル))
バックアップ	切替え済みのシステムログファイルをバックアップすることを推奨

⇒ 詳細は、『Linux(RHEL)ユーザーズマニュアル』を参照

2.7 ログファイルの運用設計 (2/2)

■ ユーザー固有ログファイルの運用

設計項目	内容		
対象ログファイル	ローテーションを実施する対象のログファイルを決定する		
実行ユーザー	ログファイルのパーミッションに基づき、どのユーザーIDで実施すべきかを決定する		
ローテーション実施コマンド	ログローテーションを実施するコマンドを決定する(システム標準はlogrotate)		
運用方法 (logrotateコマンドの場合)	ログを出力するアプリケーションのリロード機能の有無を確認し、運用方法を決定する		
	リロード機能	運用方法	指定方法
	あり	ユーザー固有ログファイル自体を切り替えて保存	logrotate設定ファイルに「create」ディレクティブを記述
なし	ユーザー固有ログファイルのコピーを保存	logrotate設定ファイルに「copytruncate」ディレクティブを記述	

⇒詳細は、『Linux(RHEL)ユーザーズマニュアル』を参照

■ メモリダンプ機能

⇒ 詳細は、『Linux(RHEL)ユーザーズマニュアル』を参照

■ kdump機能

RHELの標準ダンプ機能

運用中に異常が発生した場合にメモリの内容を採取

■ sadump機能

PRIMEQUEST固有のダンプ機能

kdump機能でダンプが採取できない状態やダンプ採取に失敗した場合に、メモリの内容を採取

■ 設計のポイント

■ メモリ容量

メモリダンプ機能には、システム運用とは別に以下のメモリ領域の確保が必要

Intel 64: 256MB (固定)

システム運用で使用するメモリ容量の設計は、上記を考慮して設計する

■ ハードディスク容量

- ・ メモリダンプ機能は、大量のディスク容量が必要。メモリダンプ機能で占有されるディスク容量を考慮して設計する
- ・ iSCSI接続およびFCoE接続のハードディスクには、kdump機能でメモリダンプを採取できない別途、メモリダンプ採取消用の内蔵ハードディスクを用意する

■ sadumpのデバイス

- ・ ダンプデバイスに使用しているディスクを活性交換する運用の場合、sadump初期設定後にディスクをバックアップする必要がある。ダンプデバイスに用意するディスクと同数のディスクを用意する
- ・ ローカルテープ装置がなく、ダンプをローカルテープに吸い上げる運用ができない場合、ダンプデバイスに採取したダンプを一度ファイルに変換する必要がある。ファイルに変換する場合、ダンプデバイスと同じサイズの空き容量があるファイルシステムをもう1つ用意する

3. 高可用設計

システム冗長化のポイントについて説明します。
詳細は、以下のマニュアルを参照してください。

- 『構成設計ガイド』
- 『Linux(RHEL)ユーザーズマニュアル』

3.1 ハードディスク冗長化のポイント (1/2)

以下の機能を利用してフレキシブルにストレージを管理し、データの可用性と信頼性を向上

■ LVM (Logical Volume Manager)

ユーザーが扱うパーティションとして論理ボリュームを作成。物理的なディスクの存在を隠ぺいし、ディスクデバイスの管理の柔軟性を向上

■ GD (PRIMECLUSTER Global Disk Services)

⇒ [「3.2 ソフトウェアRAIDの種類と特長」](#)を参照

複数台のハードディスクを組み合わせ、ディスク故障時の運用継続とデータ損失防止、およびディスクI/Oの負荷分散を実現

■ MD (Multi Disk)

⇒ [「3.2 ソフトウェアRAIDの種類と特長」](#)を参照

複数台のハードディスクを組み合わせ、ディスク故障時にデータの損失防止やディスクアクセスの処理速度の向上を実現

■ Symantec Storage Foundation

⇒ [「3.2 ソフトウェアRAIDの種類と特長」](#)を参照

ディスク故障時の運用継続とデータ損失防止に加え、省電力化 (MAID) や初期導入コストの削減 (シンプロビジョニング) を実現

3.1 ハードディスク冗長化のポイント (2/2)

■ DM-MP (Device-Mapper Multipath)

冗長性を持たせたディスクアレイ装置へのアクセスパスを統合して、アクセスパス障害の発生に備えたシステム運用の継続性を実現

■ DMP (Dynamic Multi-Pathing)

冗長性を持たせたディスクアレイ装置へのアクセスパスを統合して、アクセスパス障害の発生に備えたシステム運用の継続性を実現
(Symantec Storage Foundation 製品に同梱)

3.2 ソフトウェアRAIDの種類と特長

■ 富士通のミドルウェア

■ PRIMECLUSTER GD (Global Disk)

- ・ RAID-1に特化した冗長化機能を提供
- ・ ディスク単位の単純設計
- ・ GUIによる簡単で確実な運用管理
- ・ すぐれた管理機能により、ディスク障害だけでなく、操作ミスによるトラブルも軽減

■ OS標準

■ MD (Multi Disk)

- ・ RHELに同梱されているソフトウェア
- ・ ドライバ層でRAID(RAID 0/1/10)を構成可能

■ ISV製品

■ Symantec Storage Foundation

- ・ 多数のソフトウェアRAID機能(RAID 0/1/5/01/10 など)を提供
- ・ オンラインストレージ管理機能(ディスク交換、サイズ変更、RAID変更など)を提供
- ・ ストレージの筐体間でRAIDを構成可能
- ・ GUIによる簡単で確実な運用管理
- ・ ユーザーデータを他のOS (Solaris)と共有可能

各製品のRHEL7の対応状況については、弊社担当営業までお問い合わせください。

⇒ 詳細は、『Linux(RHEL)ユーザーズマニュアル』を参照
ハードウェアRAIDについては、『構成設計ガイド』を参照

3.3 ソフトウェアRAIDの比較

機能	PRIMECLUSTER GD	MD	Symantec Storage Foundation
RAID 0(ストライピング)	△	△	△
RAID 1(ミラーリング)	○	△	△
RAID 01(ストライプのミラーリング)	△	—	△
RAID 10(ミラーのストライピング)	—	△	△
RAID 4	—	—	—
RAID 5	—	—	△
RAID 6	—	—	—
コンカチネーション	△	△	△
ホットスペア	△	△	△
ホットスワップ	○	△	△
システムダウン後のミラー回復の高速化	○	—	△
リードエラー時のライトバックによるミラー継続	○	—	—
誤操作によるディスクアクセスからのデータ保護	○	—	—
オンラインボリューム拡張	○	—	△
ボリューム縮小	—	—	△
スナップショット	○ (*1)	—	△
GUI	○	—	△

○: システムディスクとユーザーデータ用ディスクに対応

△: ユーザーデータ用ディスクだけ

—: なし

(*1) オプション製品
「PRIMECLUSTER GD Snapshot」が必要

3.4 ネットワーク冗長化のポイント (1/2)

■ LAN冗長化(二重化)の検討対象

管理LAN、業務LAN

■ ソフトウェアの選択指針

二重化ソフトの使用により、ネットワークの高信頼化が可能

■ PRIMECLUSTER GLS (Global Link Services) (推奨オプション)

bondingドライバより高信頼性を実現(業務の即時再開と継続、故障箇所の特定と復旧が可能)

次ページの、「■高信頼化基準に基づくPRIMECLUSTER GLSとbondingの選択指針」を参照

⇒ PRIMECLUSTER GLSの詳細は、『PRIMECLUSTER Global Link Services 説明書(伝送路二重化機能編)(Linux版)』を参照

■ bondingドライバ(OS標準)

PRIMECLUSTER GLSと比べると、機能に制限あり

⇒ 詳細は、『Linux(RHEL)ユーザーズマニュアル』を参照

機能	PRIMECLUSTER GLS (NIC切替え方式)	bondingドライバ (mode1)
使用するスイッチの制限	○(*1)	○
LANカードの故障検出	○	○
ネットワークの故障検出	○	△(*2)
マルチプラットフォーム対応	○	—(*3)
サポート/提供形態	富士通/有償製品(*4)	RHEL/OS標準添付ソフト
PRIMECLUSTERとの連携	○	×

○:可能

△:一部可能

×:不可

—:該当せず

(*1) GLSのpingによるネットワーク監視先として、どれかのスイッチにIPアドレスを割り当てる必要がある

(*2) 隣接するスイッチだけ検出

(*3) SolarisとWindowsはOS標準機能でチーミング可能

(*4) SupportDesk契約によるサポート

3.4 ネットワーク冗長化のポイント (2/2)

■ 高信頼化基準に基づくPRIMECLUSTER GLSとbondingの選択指針

高信頼化要件	業務の即時再開と継続		故障箇所の特定と復旧	
	PRIMECLUSTER GLS	bonding	PRIMECLUSTER GLS	bonding
物理サーバでアプリケーションの通信を高信頼化する	○ ルータの故障を検出して、必要に応じて切替えが可能	× mode1の場合、ルータの故障を検出できず、ネットワークの構成によっては切り替わらない可能性がある	○ ネットワーク機器ごとの監視が可能で、故障箇所の特定と復旧が容易である	× mode1の場合、物理アダプターの状態(リンク状態やパケットの送受信状態)だけ監視するため、ネットワークの故障箇所の特定が不可能
仮想化環境(ゲストOS)のアプリケーションの通信を高信頼化する	○ 上記に加えてネットワーク切替え後、即時にゲストOSの通信を再開させることが可能	× 数分間、通信ができないことがある		

※Oracle DBを導入する場合など、bondingだけ使用可能な場合があります。

○:適、×:不適

■ PRIMECLUSTERを利用したクラスタ構成

Linuxでクラスタ構成する場合は、PRIMECLUSTER(クラスタ構成ソフト)の使用を推奨
PRIMECLUSTERには、以下の特長がある

■ ウィザード機能

クラスタシステムを簡単に構築できる

■ 高可用性

異常発生時に、業務が停止しないよう、速やかにノードを引き継ぐ
伝送路を多重化することにより、ネットワーク異常やネットワーク機器の故障時にも
通信の継続が可能

■ 拡張性

1つ以上のアプリケーションを複数のノードで同時に処理できる
並列データベースやロードシェアおよびロードバランスを利用した並列業務を行うことができる

⇒ 詳細は、『構成設計ガイド』の「3.4 クラスタ構成のポイント」、ServerViewのマニュアルを参照

■ Symantec Cluster Serverを利用したクラスタ構成

ワールドワイドの実績を重視される場合、プラットフォームにかかわらずすでにSymantec Cluster Server (*) を使用している場合に推奨

Symantec Cluster Server の特長は以下のとおり

■ 高可用性

- リソースのリアルタイム監視機能で、故障検知(=フェイルオーバ開始)までの時間を短縮することが可能
- SFCFSHA (*)を使用することで、高速フェイルオーバ(ダウンタイム短縮)が可能

■ 拡張性

最大64ノードまで拡張可能

SFCFSHAを使用することで、スケールアウト型のNFSサーバを構築することが可能

■ 柔軟性

フェイルオーバポリシー、クラスタとの連携機能、事象通知機能などが充実しており、柔軟なクラスタシステムを構築することが可能

(*) Symantec Cluster Server は、以下の製品に同梱

- Symantec Storage Foundation Standard HA (SFHA)
- Symantec Storage Foundation Enterprise Cluster File System HA (SFCFSHA)

※ SFHA6.2、SFCFSHA6.2からRHEL7に対応

■ クラスタ構成のポイント (PRIMECLUSTER)

■ アプリケーション監視

PRIMECLUSTER では、アプリケーションの状態(正常状態、sleep 状態、ハング状態など)を監視できないが、PRIMECLUSTER が提供する各種ウィザード製品を導入することにより可能となる

■ OSのハング状態

Linuxでは、OS がハング状態であってもノードの異常を検出することができない場合がある
検出するには、ServerView のウォッチドッグタイマー値を設定する

■ クラスタ構成のポイント (Symantec Cluster Server)

■ アプリケーション監視

Symantec Cluster Server標準の Application Agent(汎用的なAgent)または、お客様やアプリケーション開発ベンダーが作成した Agent を使用する(クラスタに登録する)ことで監視可能

■ 異機種間クラスタ

異機種間クラスタもサポート対象だが、クラスタのサービスが性能の低いサーバ機種に偏っても耐えられる設定が必要

⇒ 詳細は、『構成設計ガイド』の「3.4 クラスタ構成のポイント」、ServerViewのマニュアルを参照

4. OS運用管理設計

Linuxの運用管理設計について説明します。

4.1 修正適用の概要 (1/3)

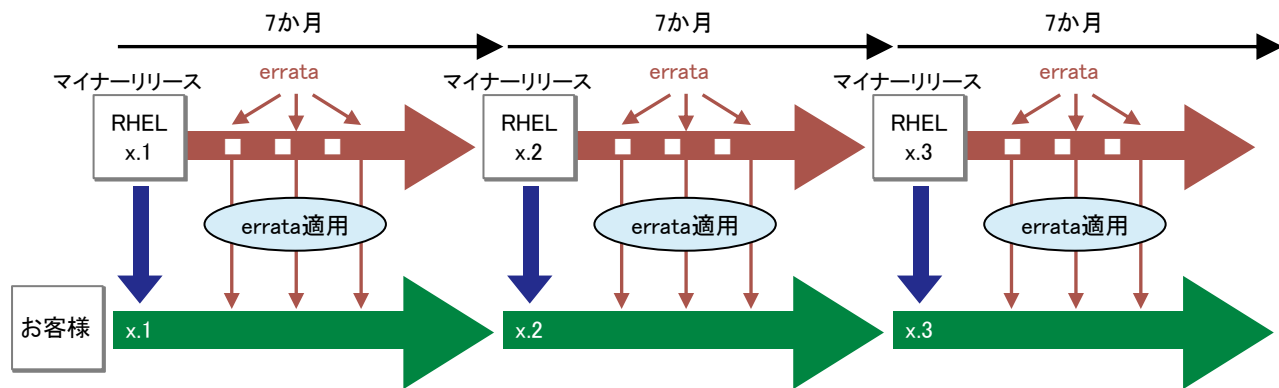
■ 修正の種類

修正の種類	説明	入手先
errata	OS の重大障害修正(セキュリティ修正を含む)	Red Hat社 カスタマーポータル (https://access.redhat.com/)
マイナーリリース	新ハード対応、機能追加、障害修正を含んだリリースで、定期的に提供される	

■ 修正適用の運用方式

■ 標準運用方式【最新化運用向け】

最新のマイナーリリースでだけerrataが提供される

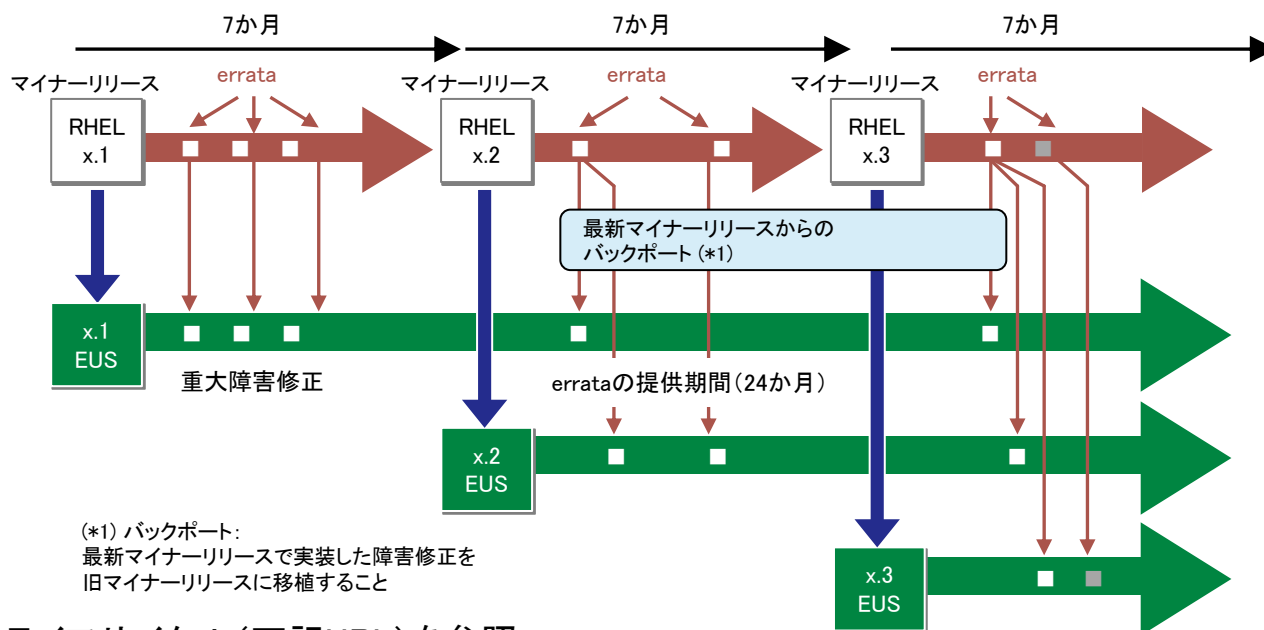


⇒ 詳細は、Red Hat社のライフサイクル(下記URL)を参照
<https://access.redhat.com/ja/support/policy/updates/errata>

4.1 修正適用の概要 (2/3)

■ EUS (Extended Update Support) 運用方式【計画保守向け】

- ・ 特定のマイナーリリースに対するerrataを標準運用方式よりも長期間継続して提供
 - ・ 特定のマイナーリリース: RHEL7.1以降のすべてのマイナーリリース
 - ・ errataの提供期間: 当該マイナーリリースのリリース後、24か月間
- ・ 定期保守に対し余裕を持った準備期間が確保できるため、無理のない保守計画が可能



⇒ 詳細は、Red Hat社のライフサイクル(下記URL)を参照
<https://access.redhat.com/ja/support/policy/updates/errata>

4.1 修正適用の概要 (3/3)

■ EUSによる修正運用を標準サポートで利用可能

- EUSを「SupportDesk Standard」で提供
- 富士通がサブスクリプション登録に使用するプロダクトID(1年更新)、およびEUS-ID(1年更新)を送付
- EUSに対応したerrataは、Red Hat社 カスタマーポータルでのEUS専用チャンネル(EUSチャンネル)から入手

■ 修正適用に関する情報やツールを提供

SupportDesk契約者様向けサイト (SupportDesk Web) で、修正運用に役立つ様々な情報やツールを提供

- ・ 修正適用方法
- ・ 検証情報
- ・ 一括修正 (マイナーリリース適用ツール)

⇒ 詳細は、『Linux(RHEL)ユーザーズマニュアル』を参照

4.2 バックアップとリストアの設計概要 (1/2)

自然災害などのリスクから資産、データを守るため、定期的なバックアップを行い、緊急時にリストアできるようにしてください

■ システムボリュームを含めたバックアップ

■ オンラインバックアップ

オンラインバックアップ対応製品か、業務への影響がないかの事前確認が必要

■ オフラインバックアップ

- ・ 業務の停止が必要(業務停止時間はバックアップ容量に比例)
- ・ オンラインバックアップより運用が簡単
- ・ 方法は、フルバックアップ、差分バックアップ、増分バックアップの3種類

方法、ソフトウェア	種類	オンライン	オフライン
dump/restoreコマンド	OS標準ユーティリティ	×	○
SystemcastWizard Professional	富士通自社開発のイメージバックアップソフト	×	○
PRIMECLUSTER GD Snapshot	PRIMECLUSTER GDのオプションソフトウェア	○	○
ETERNUS SF AdvancedCopy Manager	富士通のSANブート対応ソフトウェア	×	○
NetVault Backup 11	デル・ソフトウェア社	×	○

※ソフトウェアのサポート状況は、2016年6月現在のものです。

ソフトウェア情報は、以下のURLを参考にしてください。

<http://www.fujitsu.com/jp/products/software/resources/condition/syskou/>

<http://www.fujitsu.com/jp/products/computing/servers/partner/product-line/system/>

○:サポート ×:サポートしない

⇒詳細は、『Linux(RHEL)ユーザーズマニュアル』、各ツールのマニュアルを参照

■ データボリュームのバックアップだけ

■ NetBackup

Symantec社のソフトウェア

■ NetWorker

EMC社のソフトウェア

■ Symantec Storage Foundation の Flash Snap (*)

Symantec社のソフトウェア

(*) Flash Snap は以下の製品に同梱

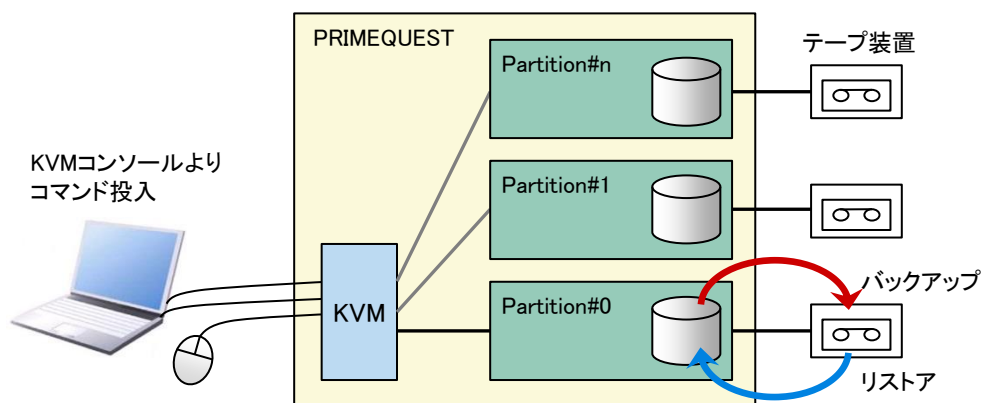
- Symantec Storage Foundation Enterprise
- Symantec Storage Foundation Enterprise Cluster File System HA

⇒ 詳細は、各ソフトウェアのマニュアルを参照

バックアップまたはリストアしたいシステムパーティションのOSコンソールからdump/restoreコマンドを実行

長所	短所
<ul style="list-style-type: none">OS標準ユーティリティであり、特別にソフトウェアを購入する必要がないバックアップサーバ不要システムボリュームのバックアップとリストアが可能(※レスキューモードの利用)	<ul style="list-style-type: none">パーティションごとに別々に操作が必要退避先として対応しているのはシングルユニットだけDBのオンラインバックアップができない

【標準ユーティリティによるローカルバックアップ】

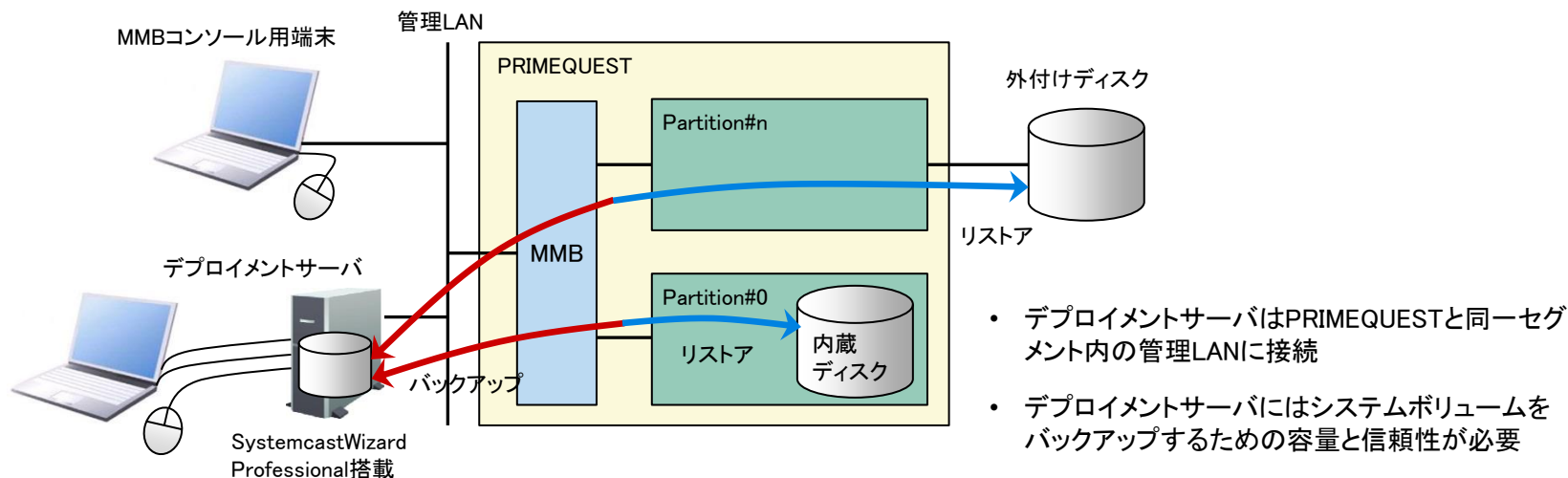


KVM: Keyboard, Video, Mouse

バックアップ対象システムのハードディスクイメージを、ネットワーク経由でデプロイメントサーバに退避し、ターゲットシステムを復旧し展開

長所	短所
<ul style="list-style-type: none"> • 複数パーティションを対象としたバックアップまたはリストアを自動的に行うことができる • リモートからの操作だけでよい • GUI操作で実行できる 	<ul style="list-style-type: none"> • デプロイメントサーバが別途必要 • 差分バックアップおよび増分バックアップができない • バックアップデータの自動的な世代管理機能がない • バックアップ前にパーティションの停止が必要 • ハードディスクへのバックアップが基本で、二次記憶媒体へのバックアップにはWindows上で動作するバックアップソフトが別途必要

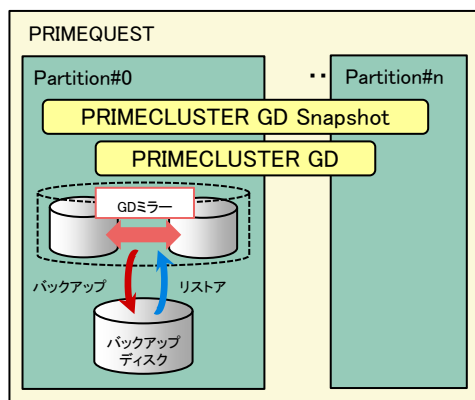
【SystemcastWizard Professionalによるバックアップの構成】



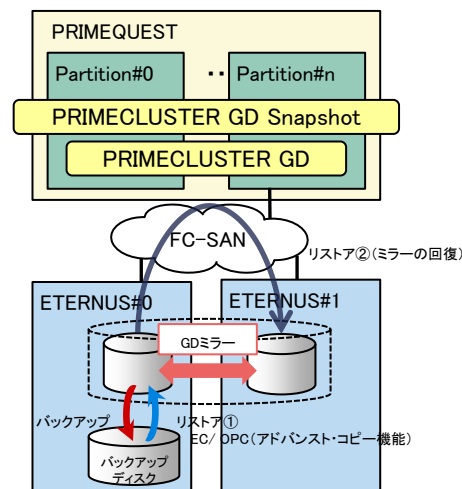
高速バックアップと高速リストア

長所	短所
<ul style="list-style-type: none"> PRIMECLUSTER GDで管理するボリューム単位、およびディスク単位でのスナップショット作成が可能 システムディスクのスナップショット機能により、運用中にシステムのバックアップやパッチの事前適用ができる ボリュームの構成やストレージ装置の持つ機能から、最適なスナップショットの方式を自動的に選択 高速コピー機能で作成されたスナップショットデータをリストアする場合、高速コピー機能が利用できない構成では、コピー方式を自動的にソフトウェアコピー機能に切り替えて、高速リストアを行う 	<ul style="list-style-type: none"> PRIMECLUSTER GD Snapshotを利用するには、PRIMECLUSTER GDが必要 共用ディスク装置を共用するすべてのサーバにPRIMECLUSTER GD Snapshotが必要

【内蔵ディスクを使用したバックアップの構成例】



【ETERNUSを使用したバックアップの構成例】

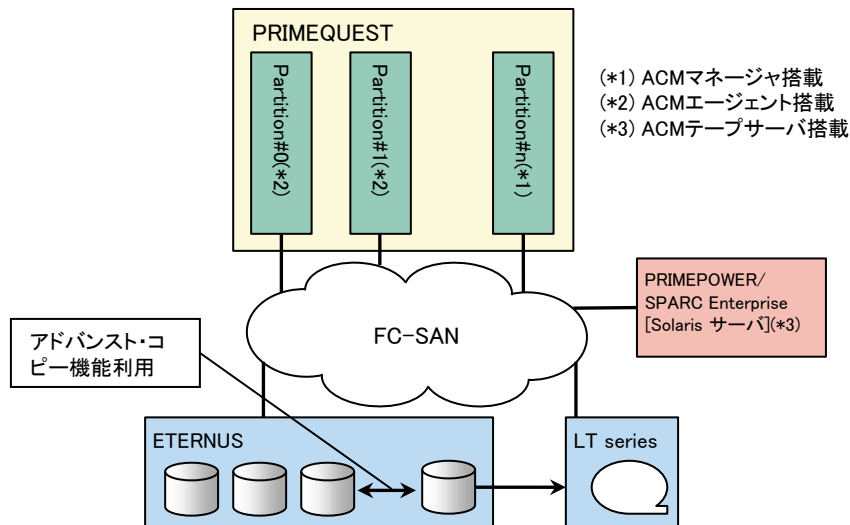


- バックアップディスクには、バックアップ元のディスクと同等以上のサイズのディスクが必要

4.3 バックアップとリストア – ETERNUS SF ACM

- ディスクアレイ装置 (ETERNUS) と組み合わせて使用
- ディスクアレイのデータを業務サーバ、LANを経由せずバックアップまたはリストア

長所	短所
<ul style="list-style-type: none"> ・ 業務サーバやLANに負荷をかけずにバックアップまたはリストア可能 ・ 業務停止時間の短い高速なバックアップまたはリストアが可能 (※ETERNUSのアドバンスド・コピー機能利用) ・ 遠隔地への高速レプリケーションが可能 (※ETERNUSのリモート・アドバンスド・コピー機能利用) ・ ETERNUSの消費電力を削減する省電力バックアップが可能 ・ DB (Oracle, Symfoware) のオンラインバックアップ可能 ・ SANブート時のシステムボリュームバックアップまたはリストア可能 	<ul style="list-style-type: none"> ・ 内蔵ディスクのバックアップに対応していない



【ACMによるETERNUS上のデータバックアップの構成】

- ・ ACMエージェント
ETERNUSのアドバンスド・コピー機能を制御
バックアップまたはリストア対象の業務サーバにインストール
- ・ ACMマネージャ
全ACMエージェントのデバイス情報やポリシーを一元管理

⇒ ETERNUS SF ACM のサポートOSについては、以下を参照
<http://www.fujitsu.com/jp/products/computing/storage/software/sf-acm/>

4.3 バックアップとリストア – NetVault Backup

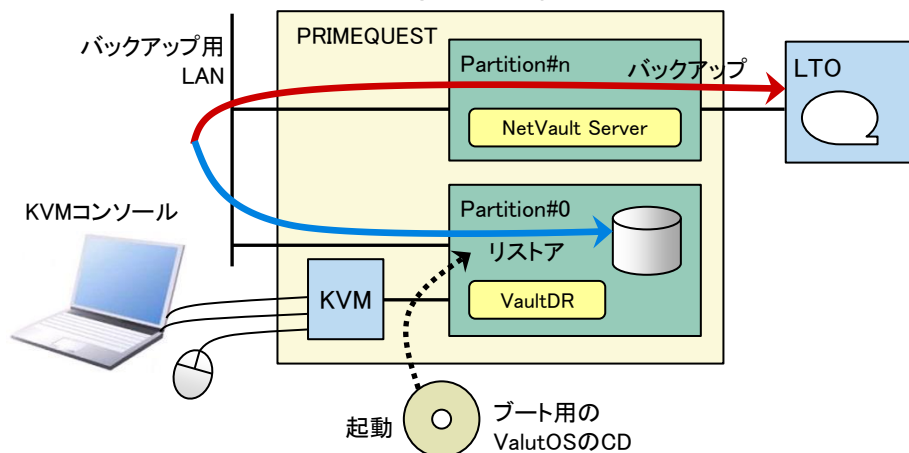
日本国内のLinuxサーババックアップ市場でシェアNo.1のソフトウェア

長所	短所
<ul style="list-style-type: none">バックアップのスケジュール定義が可能で、処理を自動化できる差分データの合成バックアップ可能 (*1)異なるOSでも統一されたGUIによる操作が可能PRIMECLUSTERとの親和性が高いDB (Oracle) のオンラインバックアップが可能 (オプション製品が必要)システムボリュームのバックアップとリストアが可能 (オプション製品が必要)OS混在 (Linux、Windows) のバックアップも可能長期サポート商品 (サポート契約継続で最大7年) を用意 (*2)	<ul style="list-style-type: none">業務サーバのシステムバックアップを取得する場合には、オプション製品が必要

(*1) フルバックアップと差分バックアップを統合し、最新のフルバックアップ相当を合成

(*2) 「基本サポート」に加え、「長期サポート」を前提とした商品を用意。RHELの拡張プラスサポートと組み合わせたサポートが可能。

【KVMコンソールを接続した構成例】



NetVault Backupによるシステムボリュームのバックアップの仕組み

- ・ VaultOSで起動する時に、リストアするシステムパーティションにKVMコンソールを接続して操作
- ・ MMBを利用したリモートコンソールによる操作

4.3 バックアップとリストア – NetBackup

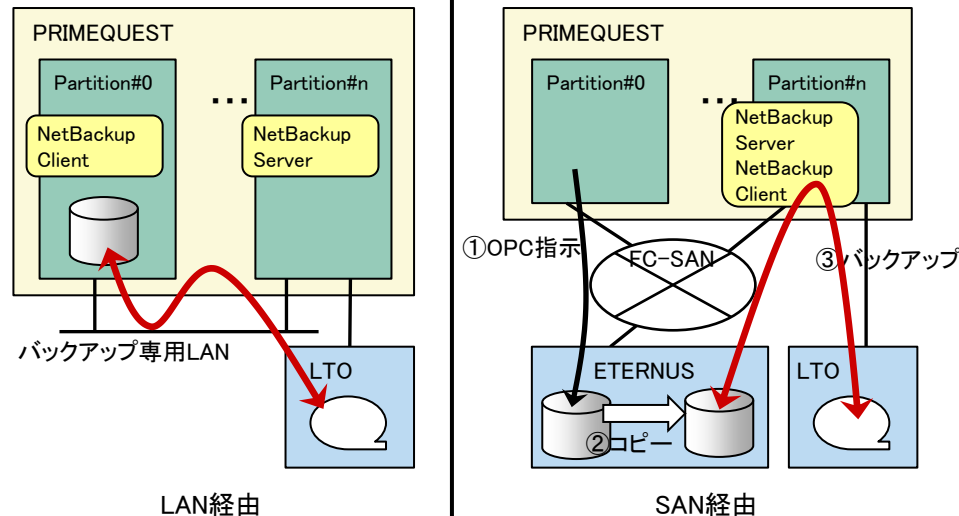
大規模システムにおけるバックアップの実績があるソフトウェア

長所	短所
<ul style="list-style-type: none"> 重複排除バックアップで、ストレージ使用量を削減(*1) アクセラレータで、完全バックアップを高速化(*1) Oracle DBのインスタンスを自動検出してオンラインバックアップ(*1) Auto Image Replication(AIR)で、遠隔地に自動で複製(*1) 複数のプラットフォームが混在している環境でも1台のサーバで管理 Snapshot Clientを使用したファイル数が多い環境でも高速でバックアップ 差分データの合成バックアップ (*2) 	<ul style="list-style-type: none"> 日本語に対応していない (Linux版だけ) システムボリュームのバックアップとリストア不可

(*1) オプションライセンスが必須

(*2) フルバックアップと差分バックアップを統合し、最新のフルバックアップ相当を合成

【NetBackupによるデータバックアップ】



4.3 バックアップとリストア – NetWorker

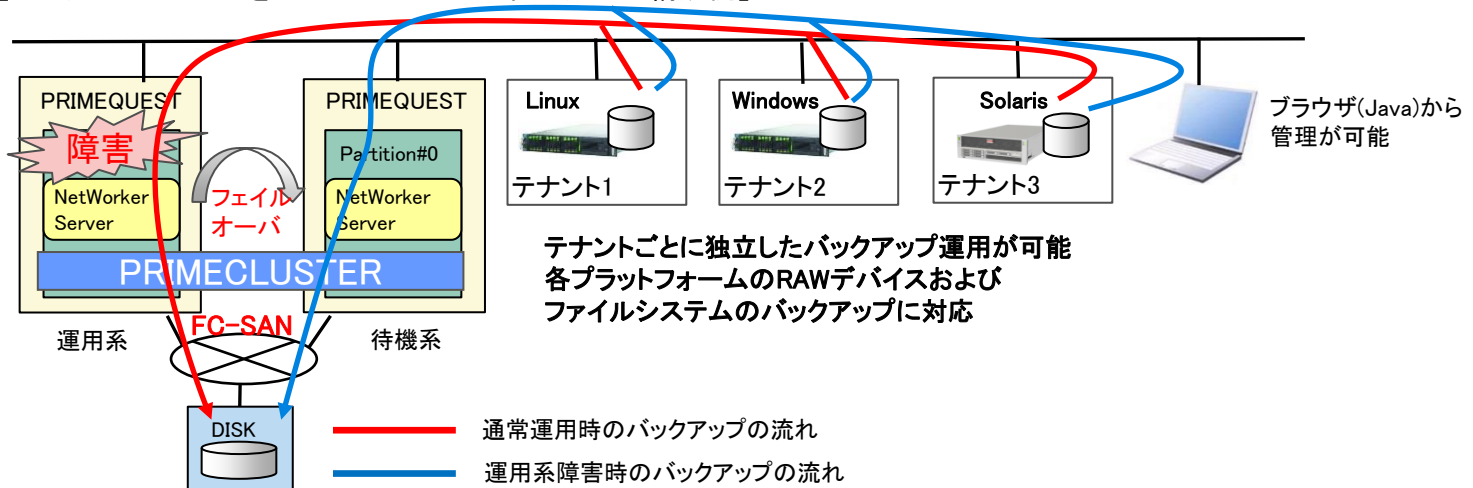
PRIMECLUSTERとの親和性がNo.1のバックアップソフトウェア

長所	短所
<ul style="list-style-type: none"> PRIMECLUSTERとの親和性が高い (UNIX Solarisで実績が多く、Solarisから移行する場合に最適) マルチテナントに対応 (テナントごとに独立したバックアップ運用が可能) VMware上のLinuxゲストOSのオンラインバックアップ (VADP)が可能 (オプション製品が必要) OS混在 (Linux、Windows、Solaris)のバックアップが可能 (rawデバイス含む) 異なるOSでも統一されたGUIによる操作が可能 バックアップのスケジュール定義が可能で、処理を自動化できる 多数ファイル (10万ファイルなど) 環境のバックアップが速い (*1) DB (Oracle) のオンラインバックアップが可能 (オプション製品が必要) 差分データの合成バックアップ可能 (*2) 	<ul style="list-style-type: none"> システムボリュームのバックアップとリストア不可 (OS標準のバックアップユーティリティ使用を推奨)

(*1) バックアップを自動的に分割させ、OSリソースを有効活用 (並列保存ストリーム機能)

(*2) フルバックアップと差分バックアップを統合し、最新のフルバックアップ相当を合成

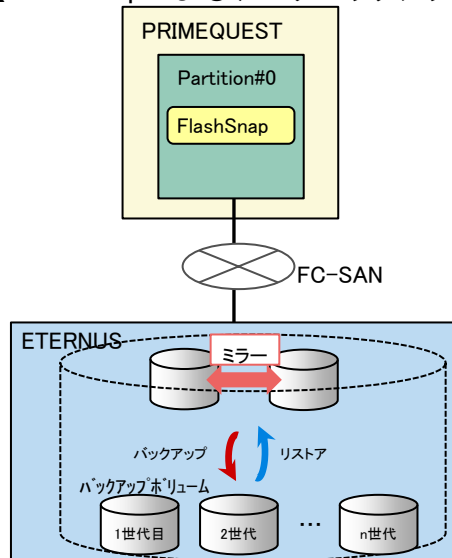
【バックアップサーバをPRIMECLUSTERに組み込んだ構成例】



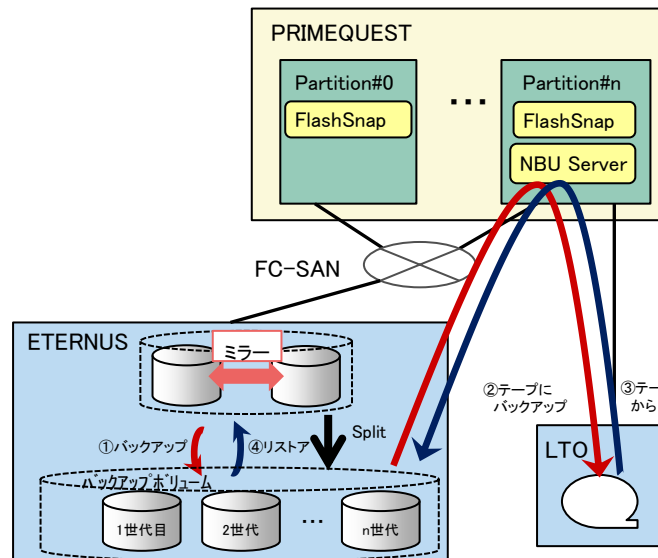
柔軟性を備えた高速バックアップ

長所	短所
<ul style="list-style-type: none"> • 通常はボリューム単位でスナップショットを作成するため、ディスク単位とは異なり、未使用領域に対しての同期処理が省かれる • バックアップ元ボリュームを含め、最大32世代まで管理可能 • 高速再同期、ディスクグループのSplit/Join 機能で、オフホストバックアップなどのオフホスト処理が可能 • ミラー切離し方式(ETERNUSのECに相当)と、インスタントスナップショット方式(ETERNUSのOPCに相当)を選択可能 • サポート対象のストレージの範囲であれば、バックアップ元およびバックアップ先のストレージの組み合わせは自在 • ボリューム単位ではなく、ファイル単位やディレクトリ単位といった局所的なスナップショットも可能 	<ul style="list-style-type: none"> • ストレージ側のコピー機能と比較して、切離し時や再接続時に、業務サーバやI/Oパスに対して負荷がかかる • 内蔵ディスクに対応していない

【FlashSnapによるデータバックアップ】



【FlashSnapとNetBackupによるオフホストバックアップ】



は、ディスクグループ

- バックアップボリュームは、別のストレージ筐体でも可能
- SFHA6.2、SFCFSHA6.2からRHEL7に対応

■ NTPの設定

- Reserved SB使用時や、複数SB(2SB以上)でパーティションを構成する場合は必ずNTP運用を行う
- NTP運用で時刻を設定する対象
 - ⇒詳細は、『構成設計ガイド』および『導入マニュアル』を参照
 - ・ MMB
NTP クライアントとなって、外部の NTP サーバと時刻の同期を取る
 - ・ 各パーティション
OSのNTP クライアント設定によって補正
- 設定のポイント
 - ・ NTPサーバを各パーティションに3台以上指定する
 - ・ MMBがNTP運用を行っている場合、各パーティションもMMBと同じサーバをNTPサーバに指定する

4.4 時刻補正のポイント (2/2)

■ Linuxにおける時刻の運用

■ システム起動時

ハードウェア時計から取得した時刻をシステム時計に設定

■ システム運用中

ハードウェア時計とシステム時計は別々に進行
(NTP運用中は、システム時計だけ補正される)

■ システム停止時

システム時計の時刻をハードウェア時計に反映

■ 時刻補正の設定

⇒詳細は、『製品概説』、『Linux(RHEL)ユーザーズマニュアル』を参照

ハードウェア時計とシステム時計の定期的な同期実施を推奨
(hwclockコマンドの定期実施)

システムの長期運用中は、システム時計、
ハードウェア時計とも独自に計時され、同期しません

システム時計:システム稼働中にタイマー割り込みを使用して時刻を管理しているソフトウェア時計
ハードウェア時計:システム停止時もバッテリーにより駆動され時刻を保持する、ハードウェア搭載の時計

■ 著作権・商標権・その他の知的財産権について

コンテンツ(文書・画像・音声など)は、著作権・商標権・その他の知的財産権で保護されています。本コンテンツは、個人的に使用する範囲でプリントアウトまたはダウンロードできます。ただし、これ以外の利用(御自分のページへの再利用やほかのサーバへのアップロードなど)については、当社または権利者の許諾が必要となります。


■ 保証の制限

本コンテンツについて、当社は、その正確性、商品性、御利用目的への適合性などに関して保証するものではなく、その御利用により生じた損害について、当社は法律上のいかなる責任も負いかねます。本コンテンツは、予告なく変更・廃止されることがあります。

不明な点は、「本製品のお問い合わせ」

(<http://www.fujitsu.com/jp/products/computing/servers/primequest/contact/>) よりお尋ねください。

無断転載を禁じます。



FUJITSU

shaping tomorrow with you