

Veritas Storage Foundation™ for Oracle® RAC ベストプラク ティスガイド



Veritas Storage Foundation for Oracle RAC ベストプラクティスガイド

このマニュアルで説明するソフトウェアは、使用許諾契約に基づいて提供され、その内容に同意する場合にのみ使用することができます。

Document version 1.0

法定通知

Copyright © 2009 Symantec Corporation. All rights reserved.

Symantec、Symantec ロゴ、Veritas、Veritas Storage Foundation、FlashSnap は、Symantec Corporation または同社の米国およびその他の国における関連会社の商標または登録商標です。その他の会社名、製品名は各社の登録商標または商標です。

本書に記載する製品は、使用、コピー、頒布、逆コンパイルおよびリバース・エンジニアリングを制限するライセンスに基づいて頒布されています。Symantec Corporation からの書面による許可なく本書を複製することはできません。

Symantec Corporation が提供する技術文書は Symantec Corporation の著作物であり、Symantec Corporation が保有するものです。保証の免責: 技術文書は現状有姿で提供され、Symantec Corporation はその正確性や使用について何ら保証いたしません。技術文書またはこれに記載される情報はお客様の責任にてご使用ください。本書には、技術的な誤りやその他不正確な点を含んでいる可能性があります。Symantec は事前の通知なく本書を変更する権利を留保します。

ライセンス対象ソフトウェアおよび資料は、FAR 12.212 の規定によって商用コンピュータソフトウェアとみなされ、場合に応じて、FAR 52.227-19 「Commercial Computer Licensed Software - Restricted Rights」、DFARS 227.7202 「Rights in Commercial Computer Licensed Software or Commercial Computer Licensed Software Documentation」、その後継規制の規定により制限された権利の対象となります。

Symantec Corporation
20330 Stevens Creek Blvd.
Cupertino, CA 95014

<http://www.symantec.com>

第 1 章	SF Oracle RAC のベストプラクティス	5
	概要	6
	対象者	6
	頭字語一覧	6
	Veritas Storage Foundation for Oracle RAC について	7
	SF Oracle RAC の利点	8
	標準的な SF Oracle RAC の設定	9
	システム設定の推奨事項	11
	ネットワーク設定について	12
	ネットワーク設定の推奨事項	13
	ストレージの推奨事項	15
	共有ストレージオプション	16
	ディスクアレイの必要条件	16
	Oracle 設定の推奨事項	16
	データベースストレージの推奨事項	17
	ボリューム設計の推奨事項	18
	ファイルシステムの推奨事項	18
	I/O フェンシングの推奨事項	19
	ディザスタリカバリの注意事項	19
	キャンパスクラスタ	20
	グローバルクラスタ	22
	データレプリケーションの注意事項	22
	データレプリケーションオプション	22
	リファレンス	23
付録 A	推奨される SF Oracle RAC 配備シナリオ	25
	VCS IPC および PrivNIC エージェントを含む SF Oracle RAC クラスタ	25
	UDP IPC および MultiPrivNIC エージェントを含む SF Oracle RAC クラスタ	28
	UDP IPC および MultiPrivNIC エージェントを含む複数のデータベース用の SF Oracle RAC	30
	分離された Oracle トラフィック および MultiPrivNIC エージェントを含む SF Oracle RAC クラスタ	32
	NIC 結合、VCS IPC、および PrivNIC エージェントを含む SF Oracle RAC クラスタ	34

NIC 結合、UDP IPC、および PrivNIC エージェントを含む SF Oracle RAC
クラスタ 36

SF Oracle RAC のベストプラクティス

この章では以下の項目について説明しています。

- 概要
- 対象者
- 頭字語一覧
- Veritas Storage Foundation for Oracle RAC について
- 標準的な SF Oracle RAC の設定
- システム設定の推奨事項
- ネットワーク設定について
- ストレージの推奨事項
- Oracle 設定の推奨事項
- データベースストレージの推奨事項
- ボリューム設計の推奨事項
- ファイルシステムの推奨事項
- I/O フェンシングの推奨事項
- ディザスタリカバリの注意事項
- リファレンス

概要

このマニュアルは、データセンターでの Veritas™ Storage Foundation for Oracle® RAC (SF Oracle RAC) の配備と管理に関して推奨されるベストプラクティスをまとめたものです。ここで説明されるさまざまな設定には、システム設定、ネットワーク設定、ストレージ設定、Oracle データベース設定、ディザスタリカバリ設定、インストール、およびアップグレードが含まれます。

対象者

このマニュアルは、稼働中の環境で SF Oracle RAC を配備および管理するシステム管理者、ストレージ管理者、データベース管理者を対象としています。

頭字語一覧

ここでは、この文書で使用される頭字語の一覧を示します。

表 1-1 頭字語の一覧

頭字語	定義
ASM	Automatic Storage Management
ASL	Array Support Library
CF	Cache Fusion
CFS	Cluster File System
CRS	Cluster Ready Services または Oracle Clusterware
CVM	Cluster Volume Manager
DLM	Distributed Lock Manager
DMP	Dynamic Multi-Pathing
EBN	Enclosure Based Naming
FAN	Fast Application Notification
FC	Fibre Channel
GAB	Group Membership Services/Atomic Broadcast
GCO	Global Cluster Option
GCS	Global Cache Service

頭字語	定義
GES	Global Enqueue Service
HCL	Hardware Compatibility List
IP	Internet Protocol
LLT	LLT (Low Latency Transport)
LMX	LLT Multiplexer
OCR	OCR (Oracle Cluster Registry)
OCSSD	Oracle Cluster Synchronization Services Daemon
ODM	Oracle Disk Manager
PGR	Persistent Group Reservation
RAC	Real Application Cluster
SCSI	Small Computer System Interface
SF	Storage Foundation
TCP	Transmission Control Protocol
VCS	Veritas Cluster Server
VCSIPC	VCS Interprocess Communication
VCSMM	Veritas Cluster Server Membership Module
VIP	Virtual Internet Protocol (address)
VXFEN	Veritas I/O Fencing
VVR	Veritas Volume Replicator
VxVM	Veritas Volume Manager

Veritas Storage Foundation for Oracle RAC について

Veritas Storage Foundation™ for Oracle® RAC (SF Oracle RAC) は、独自のストレージ管理および高可用性テクノロジーを活用して、UNIX プラットフォームへの Oracle RAC の堅牢で管理性と拡張性に優れた配備を可能にします。ソリューションで使用するクラスタファイルシステムテクノロジーには、容易なファイルシステム管理と使い慣れたオペレーティングシステムツールおよびユーティリティを使用したデータベース管理という2つの利点があります。

ソリューションスタックは、Veritas Cluster Server (VCS)、Veritas Cluster Volume Manager (CVM)、Veritas Cluster File System (CFS)、および Veritas Storage Foundation (基本の Veritas Volume Manager (VxVM) と Veritas File System (VxFS) を含む) で構成されます。

SF Oracle RAC の利点

SF Oracle RAC には次の利点があります。

- ファイルシステムベースの管理のサポート。SF Oracle RAC は、Oracle データファイルおよびその他のアプリケーションデータを保存して管理するための汎用のクラスタ化されたファイルシステムテクノロジーを提供します。
- クラスタ相互接続の高可用性のサポート。LMX/LLT プロトコルと PrivNIC/MultiPrivNIC エージェントの組み合わせにより、最大帯域幅およびスイッチの冗長性を含むクラスタ相互接続の高可用性が提供されます。
- Oracle Cluster Registry および投票ディスクの配置に対するクラスタ化されたファイルシステムの使用。クラスタ化されたファイルシステムおよびボリューム管理テクノロジーは、Oracle Cluster Registry および投票ディスクを配置するための堅牢な共有ブロックと RAW インターフェースを提供します。SF Oracle RAC がない場合、OCR および投票ディスク用として個別に LUN を設定する必要があります。
- アプリケーションおよびデータベース管理に対する標準のアプローチのサポート。すべての SF Oracle RAC ソフトウェアスタックの単一ベンダによるソリューションは、アプリケーションおよびデータベース管理に対する標準のアプローチの作成を可能にします。さらに、管理者は SF Oracle RAC で Veritas テクノロジーの既存の専門知識を応用できます。
- DMP (Dynamic Multipathing) を使った可用性とパフォーマンスの向上。DMP は幅広いストレージレイサポートを提供し、HBA スイッチや SAN (Storage Area Network) スイッチでの故障やパフォーマンスのボトルネックから保護します。
- 単一の Web コンソールからの SF Oracle RAC クラスタの容易な管理および監視。
- さまざまな種類のアプリケーションおよびデータベースのサポート。
- Oracle Disk Manager (ODM) を使用したファイルシステムアクセス時間の向上。
- ASM ディスクグループを CVM ボリュームに対して設定して DMP (Dynamic Multi-Pathing) の利点を活用する機能。
- クラスタ内のデータベースごとに複数の Oracle RAC インスタンスへアクセスできることによる拡張性と可用性の向上。
- ボリュームレベルとファイルシステムレベルのスナップショットテクノロジーを使用するバックアップおよび回復ソリューションのサポート。SF Oracle RAC は、オフホスト処理でのフルボリュームレベルのスナップショットと、効率的なバックアップ処理とロールバック処理のためのファイルシステムレベルのスナップショットに対応します。

- クラスタ化されたファイルシステムテクノロジーを使用してダウンタイムを発生させずにアプリケーションをフェールオーバーする機能。
- 堅牢な SCSI-3 PGR (Persistent Reservation) ベースの I/O フェンシングによるスプリットブレインのシナリオでのデータ破損の防止。
- Oracle データベースファイルだけでなくあらゆる種類のファイルのノード間での共有をサポート。
- ユーザーに対するダウンタイムと中断を最小限に抑えた迅速なディザスタリカバリ。ユーザーは、ローカルでの高可用性サイトから、プライマリサイトとセカンダリサイトを備えた広域ディザスタリカバリ環境へ移行できます。1 つのノードに障害が発生した場合、そのノードに接続していたクライアントは障害が発生していないノードに接続し直して、アクセスを再開できます。SF Oracle RAC 環境では、障害後のリカバリが、フェールオーバーされたデータベースのリカバリよりもはるかに高速です。
- 稼働中のシステムに影響を与えない、ファイアドリルテクノロジーを使用したディザスタリカバリ設定の確認。
- 広範なハードウェアレプリケーションテクノロジーおよび VVR を使用したブロックレベルレプリケーションのサポート。

標準的な SF Oracle RAC の設定

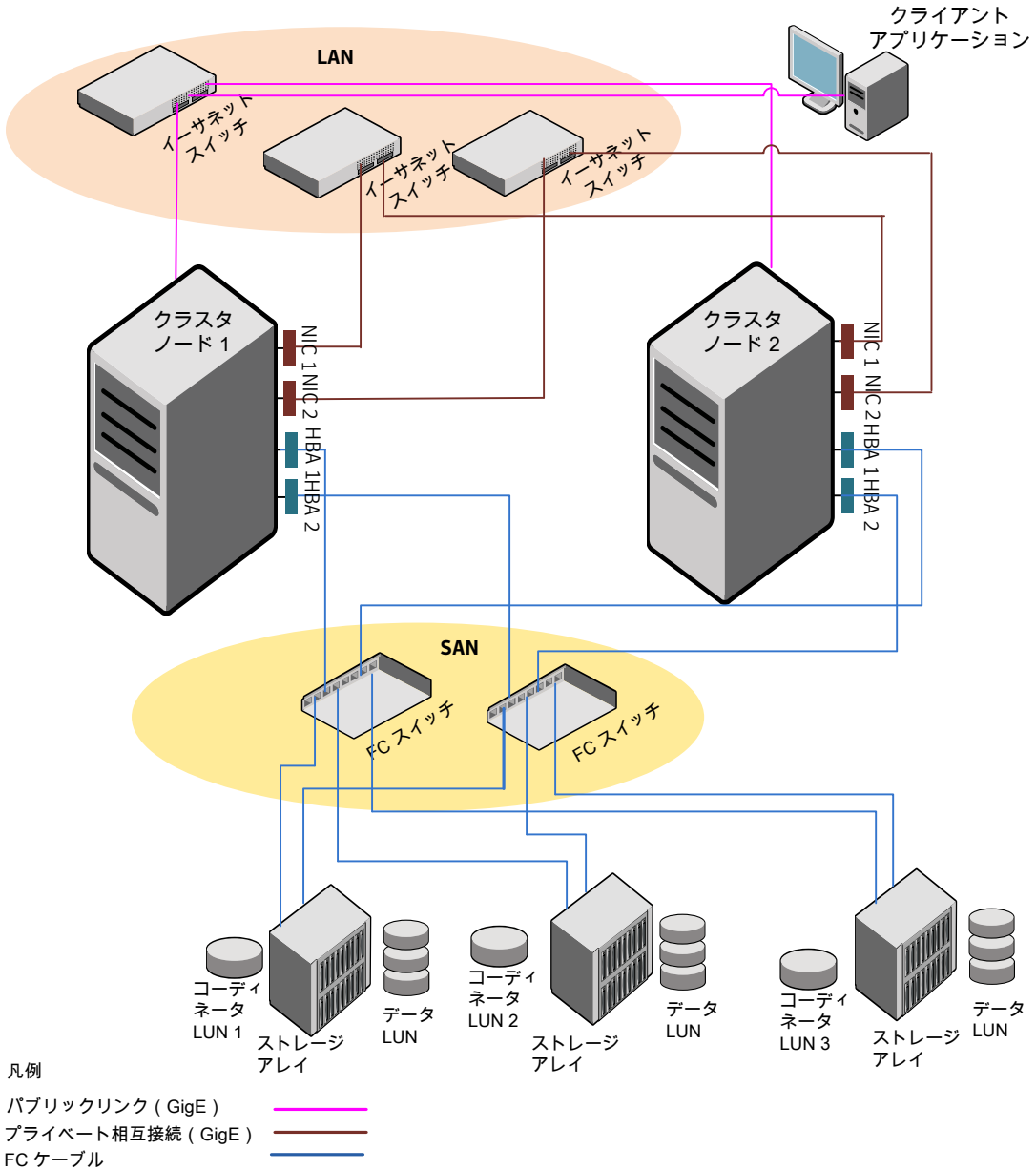
この項では、基本的な 2 ノードの SF Oracle RAC クラスタの物理的な設定を図で示します。

その他の推奨される SF Oracle RAC 配備設定については、

p.25 の [付録A](#) を参照してください。

[図 1-1](#) に、基本的な 2 ノードの SF Oracle RAC 設定を示します。

図 1-1 基本的な 2 ノードの SF Oracle RAC 設定



システム設定の推奨事項

次の推奨事項は、SF Oracle RAC クラスタでノードを一貫して設定するためのガイドラインです。

アーキテクチャ	Oracle RAC は、クラスタ内のすべてのノードが同じアーキテクチャであることを必要とします。x86、SPARC、IA/PA などのアーキテクチャが混在するクラスタはサポートされません。
CPU およびメモリの注意事項	SF Oracle RAC クラスタ内のサーバーの CPU 速度およびメモリは同一にする必要があります。
スワップ領域の必要条件	Oracle RAC 10g のスワップ領域の最小必要条件は 4 GB です。 Oracle RAC 11g のスワップ領域の最小必要条件は 8 GB です。 最小スワップに対するオペレーティングシステムの必要条件は、スワップ領域用の RAM サイズの 2 倍です。 Oracle RAC の最小必要条件とオペレーティングシステムの最小必要条件のうち、高いほうの必要条件を満たすようにします。たとえば、Oracle RAC 11g システムで最小スワップ領域に対するオペレーティングシステムの必要条件が 5 GB と計算される場合は、Oracle RAC の最小スワップ領域必要条件である 8 GB を満たすようにしてください。
OS バージョンおよびパッチ	サーバーのオペレーティングシステムのバージョンおよびパッチは同一にする必要があります。
SF Oracle RAC バージョン	クラスタ内のすべてのノードで、SF Oracle RAC バージョンが同じである必要があります。
Oracle バージョン	最新の Oracle パッチセットを使用します。サポートされる Oracle データベースバージョンの最新情報については、次の TechNote を参照してください。 http://entsupport.symantec.com/docs/280186 SF Oracle RAC は、物理クラスタ内で複数のバージョンの Oracle データベースをサポートします。たとえば、Oracle 11g R1 Clusterware を使って、同じクラスタに Oracle 10g R2 データベースと Oracle 11g R1 データベースをインストールできます。
クラスタノードのシステム時間	クラスタ内のすべてのノードで、システム時間が同じである必要があります。各ノードの時間設定を同期するための NTP (Network Time Protocol) デーモンをお勧めします。

メモ: SF Oracle RAC をインストールまたはアップグレードする前に、Veritas Installation Assessment Service (VIAS) ユーティリティを実行して、環境内のシステムが必要条件を満たしているかどうかを確認します。VIAS のダウンロードおよび実行に関する情報については、<https://vias.symantec.com/> を参照してください。

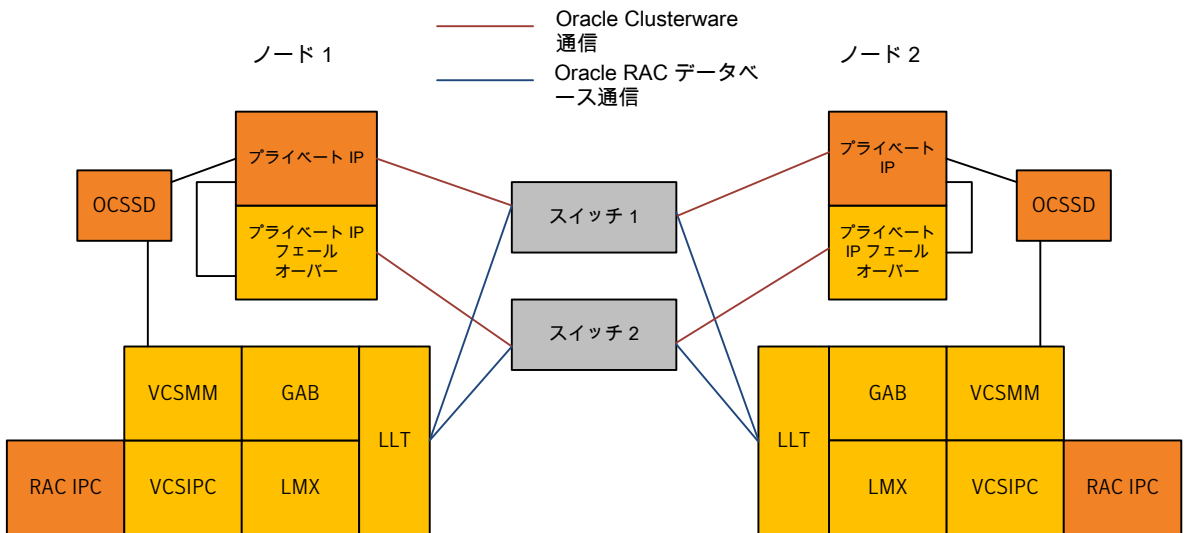
ネットワーク設定について

SF Oracle RAC および Oracle RAC のネットワーク接続の必要条件は次のように分類されます。

- プライベートネットワーク(クラスタ内の通信用)
- パブリックネットワーク(外部との通信用)

図 1-2 は、ノード間の通信層を示しています。

図 1-2 通信層



SF Oracle RAC および Oracle RAC コンポーネントは、クラスタ内でプライベートネットワークを介して通信します。

通信は次のように分類できます。

クラスタ通信	<p>クラスタ通信は、SF Oracle RAC のさまざまなコンポーネント (Veritas Cluster Server、Cluster File System、および Cluster Volume Manager) 間の通信を構成します。</p> <p>この通信は、LLT 設定ファイルで指定された複数の LLT リンクで多重化されます。</p>
Oracle Clusterware 通信	<p>Oracle Clusterware 通信は、Cluster Synchronization Service (ハートビート) を含む Oracle Clusterware トラフィックを構成します。</p> <p>Oracle Clusterware 通信は、単一のサブネット上の単一の論理接続を使用します。スプリットプレインシナリオを防止するために、このネットワークをリンク障害に対して保護する必要があります。</p>
Oracle RAC データベース通信 (データベース単位)	<p>Oracle RAC データベース通信は、GCS (Global Cache Service) および GES (Global Enqueue Service) (以前の名称は CF (Cache Fusion) および DLM (Distributed Lock Manager)) 用の RAC インスタンスピアツーピア通信を構成します。</p> <p>この通信の帯域幅の必要条件は、設定されているデータベースの数によって異なります。データベース サービスが中断されないように、ネットワークをリンク障害に対して保護する必要があります。</p>

ネットワーク設定の推奨事項

耐障害性に優れたネットワークセットアップを行うために、次の方法をお勧めします。

- 複数の専用ギガビットイーサネットリンク上でプライベートクラスタ相互接続を設定します。ネットワークインターフェースカード (NIC)、スイッチ、相互接続などの単一障害点をすべて取り除く必要があります。
- NIC とスイッチポートが速度を自動的にネゴシエートできるようにします。プライベートクラスタ相互接続に使用される NIC は、すべてのノードで速度、MTU、二重などの特性を備えている必要があります。
- プライベートクラスタ相互接続に対してルーティング不能な IP アドレスを設定します。
- ネットワーク通信に対してジャンボフレームを有効/無効にします。インスタンス間通信に LLT (Low Latency Transport) を使用する Veritas 提供の Cache Fusion ライブラリを使用する場合は、ネットワーク通信用のジャンボフレームはお勧めしません。UDP (User Datagram Protocol) を使用する場合は、ジャンボフレームをお勧めします。UDP およびジャンボフレームの利点を活用するには、インスタンス間での極度の Cache Fusion 負荷が必要です。

- 結合された NIC インターフェースまたは PrivNIC/MultiPrivNIC エージェントを使用して、リンク障害に対して Oracle Clusterware クラスタ相互接続を保護する必要があります。

リンク集約/NIC 結合 SF Oracle RAC 5.0 MP3 以降、LLT は結合された NIC を認識し、結合内部でフェールオーバー/フェールバック操作を処理します。LLT は結合された NIC を単一リンクとして扱います。このため、NIC 結合に基づくリンクが個別の LLT リンクとして設定されていないことを確認してください。

PrivNIC/MultiPrivNIC エージェント PrivNIC エージェントは、Oracle RAC 10g 用に提供されました。MultiPrivNIC エージェントは、SF Oracle RAC 5.0 MP3 以降で使用できます。

リンクが PrivNIC/MultiPrivNIC エージェントで設定されている場合、リンク障害が発生すると、エージェントは障害が発生したリンクから接続済み/使用可能な LLT リンクに Oracle Clusterware をフェールオーバーします。PrivNIC/MultiPrivNIC を使用するには、Oracle Clusterware 相互接続と Oracle RAC データベース通信リンクの両方を LLT リンク上で設定する必要があります。

- Oracle データベースクライアントは、データベースサービスにパブリックネットワークを使用します。ノードまたはネットワークに障害が発生した場合、クライアントは、既存接続と新規接続の両方を、接続可能なクラスタ内の障害が発生していないノードにフェールオーバーします。クライアントフェールオーバーは、Oracle Fast Application Notification、VIP フェールオーバー、およびクライアント接続 TCP タイムアウトの結果として発生します。Oracle Cache Fusion トラフィックをパブリックネットワーク経由で送信しないことを強くお勧めします。

- パブリックリンクに障害が発生した場合に Oracle が仮想 IP アドレスをフェールオーバーできるよう、複数のパブリックネットワークを設定して冗長性を確保します。

- LLT ピアの非アクティビティタイムアウトおよび CSS miss-count に推奨される値を設定します。

LLT ピアの非アクティビティタイムアウト値は、あるノード上の SF Oracle RAC が、クラスタ内のその他のノードを、そのノードからのネットワーク通信 (ハートビート) がない場合にダウンしたと宣言するまでの間隔を示します。

LLT ピアの非アクティビティタイムアウトのデフォルト値は 16 秒です。値は、サービス可用性の必要条件、およびキャンパスクラスタセットアップの場合はクラスタノード間の伝播遅延に基づいて設定する必要があります。

CSS miss-count 値は、Oracle Clusterware が、相互接続で応答していない別のノードをクラスタから退去させる前に待機する時間を示します。

SF Oracle RAC の場合、CSS miss-count のデフォルト値は 600 秒です。2 つのクラスタウェア (VCS と Oracle Clusterware) がネットワークでスプリットブレインが発生し

た場合にクラスタ内でどのノードを残すかについての互いの決定に干渉しないよう、このパラメータの値は LTT ピアの非アクティビティタイムアウトを大きく上回る値に保たれます。Veritas I/O フェンシングは、残るノードについて最初に決定することを許可されています。その後 Oracle Clusterware が続きます。

ストレージの推奨事項

SF Oracle RAC スタックを効率的に機能させるために、次の方法をお勧めします。

- パフォーマンスを最適化するには、次の方法でローカルストレージと共有ストレージでファイルを分散します。

ローカルストレージ

次のバイナリをローカルディスクに配置します。

- SF Oracle RAC バイナリ
 - Oracle Clusterware バイナリ
 - Oracle バイナリ
- Oracle バイナリをローカルディスクに配置すると、クラスタを段階的にアップグレードできます。

共有ストレージ

次のバイナリを共有ディスクに配置します。

- SF Oracle RAC フェンシングコーディネータディスク
- SF Oracle RAC データベースストレージ管理リポジトリ
- Oracle RAC データベース
- Oracle Clusterware レジストリ(OCR)および投票ディスク

データベースデータファイル

管理しやすくするために、Raw デバイスまたは CVM Raw デバイスではなく CFS に Oracle データベースファイルを格納します。Oracle データベースごとに別個のクラスタファイルシステムを作成します。Oracle データベースのホームを別個のマウントポイント上に維持すると、メンテナンスの際に他のデータベースに影響を与えずにデータベースをマウント解除できます。

Oracle データベースを ASM に格納する予定がある場合は、ASM ディスクグループを CVM ボリューム上で設定して、動的マルチパスの利点を活用します。

- データベースリカバリデータ(アーカイブ、フラッシュリカバリ) REDO ログおよびアーカイブログは、ローカルファイルシステムではなく、CFS に配置します。

- OCR および投票ディスク

OCR および投票ディスクを CVM ボリュームまたは CFS ファイルシステムに配置して、動的マルチパスでディスクを投票する高可用性を実現します。

冗長性のため、OCR デバイスと投票デバイスを含む CVM ボリュームをミラー化します。

- CVM は、ネイティブの命名 (cXtXdX) とエンクロージャに基づく命名 (EBN) を提供します。
ストレージを管理しやすくするために、エンクロージャに基づく命名を使用します。エンクロージャに基づく命名を使用すると、LUN のオペレーティングシステム名に関係なく、すべてのノードで共有 LUN に同じ名前が付けられます。

共有ストレージオプション

SF Oracle RAC には、次の共有ストレージオプションがあります。

- CVM
- CFS
- CVM 上の Oracle ASM

ASM はデータファイル、コントロールファイル、オンラインおよびアーカイブ REDO ログファイル、バックアップファイル専用のストレージであることに注意してください。Oracle バイナリ、トレースファイル、警告ログ、エクスポートファイル、tar ファイル、コアファイル、Oracle クラスタレジストリデバイス (OCR) とクォーラムデバイス (投票ディスク)、およびアプリケーションバイナリとデータはサポートされません。

ディスクアレイの必要条件

ディスクアレイの次の必要条件を満たしていることを確認します。

- SF Oracle RAC は、I/O フェンシング用のコーディネータディスクとデータディスクが SCSI-3 PR 互換であることを必要とします。
- 最新の互換性リストを参照して、ストレージアレイと SF Oracle RAC の互換性を確認します。
http://www.symantec.com/business/products/otherresources.jsp?pcid=2245&pvid=203_1
- 正しい ASL (Array Support Library) がすべてのクラスタノードにインストールされていることを確認します。

Oracle 設定の推奨事項

Oracle ストレージの推奨事項については、

p.15 の「[ストレージの推奨事項](#)」を参照してください。

ストレージに関する推奨事項に加えて、次の方法で Oracle を設定することをお勧めします。

Oracle データベース管理	Oracle データベースの管理には VCS Oracle エージェントをお勧めします。VCS は、この設定において Oracle データベースを制御します。VCS Oracle エージェントを使用しない設定は、単一のデータベースの設定でのみ使用できます。
複数の RAC データベース	<p>デフォルトでは、Oracle Clusterware 通信および Oracle RAC データベース通信は同じプライベートネットワークを共有します。複数の RAC データベースが同じ SF Oracle RAC クラスタに配備される場合、この単一の物理ネットワークが提供する帯域幅ではすべてのデータベースの必要条件を満たすのに十分でない可能性があります。このようなシナリオでは、複数のプライベートネットワークを使用し、Oracle RAC データベース通信の負荷をこれらのリンク間で分散するように RAC データベースを設定することができます。</p> <p>これを行うには、次の例に示されているように、Oracle データベース初期化パラメータ CLUSTER_INTERCONNECT を定義し、複数の値を設定します。</p> <pre>CLUSTER_INTERCONNECTS = ip1:ip2:...:ipn</pre> <p>Oracle データベースは、指定されている相互接続がすべて機能している限り、これらの相互接続すべてを使用して、物理リンク間でトラフィックを均等に分散します。</p>
ASM	<ul style="list-style-type: none">■ 外部冗長性を持つ動的マルチパスで VxVM ミラーボリュームを使い、高可用性を実現します。■ ASM が使う VxVM ボリュームを、ファイルシステムの作成など他の目的で使わないでください。■ データベースが ASM にインストールされているときは、ODM を有効にしないでください。

データベースストレージの推奨事項

次の推奨事項に従うことで、ストレージのパフォーマンスと可用性を向上させることができます。

- アレイの障害に対する保護を確実にするためには、可能な場合、複数のストレージアレイを使います。推奨される最小の設定は、2 つの HBA (ホストごと)、2 つのスイッチ、および 2 つのストレージアレイです。
- 処理効率および高可用性の必要条件を念頭において、ストレージレイアウトを設計してください。ストライプおよびミラーといったテクノロジーを活用します。
- I/O の処理効率を最適化するには、適切なストライプの幅と深度を使います。
- SCSI-3 PGR 準拠のストレージを使います。

- HBA/スイッチの組み合わせによってディスクへのアクセスパスを複数提供して、ストレージリンク障害に対する高可用性を確保し、負荷分散を実現します。
- データベース DST (Dynamic Storage Tiering) (DBDST) を使用してストレージコストを最適化します。DBDST を使用すると、使用頻度の低いデータを低速の安価なディスクに移動することができます。また、これにより、アクセス頻度の高いデータがより高速なディスクに保存されるため、データの取得が高速になります。

ボリューム設計の推奨事項

次の推奨事項に従うことで、VxVM/CVM ボリュームのレイアウトを最適化できます。

- VxVMミラーを使用する場合は、ボリュームを2つ以上のストレージアレイにまたがってミラー化します。
- 配置ポリシーを設計するときは、Oracleリカバリ構造をデータベースファイルから分離して、高可用性を確実にします。
- より高い処理効率のためには、REDO ログを分離し、最速のストレージ(たとえば、RAID 1+0)に配置します。
- Oracle ログボリュームのクローン作成には「サードミラーブレイクオフ」スナップショットを使用します。領域最適化(SO)スナップショット上に Oracle ログボリュームを作成しないでください。
- Oracle データボリュームに領域最適化(SO)スナップショットを使用する場合は、可能な限り多くのキャッシュオブジェクト(CO)を作成します。
- 複数のキャッシュオブジェクトを作成する場合は、すべてのキャッシュオブジェクトで均等に I/O 負荷を分散します。
- VxVMミラーを使用する場合は、Fast Mirror Resync リージョンサイズをデータベースブロックサイズと等しくして、コピーオンライト(COW)のオーバーヘッドを削減します。リージョンサイズを小さくすると、キャッシュオブジェクトの割り当て量が増え、パフォーマンスのオーバーヘッドが発生します。
- 共有ストレージへのアクセスを制御するには、ゾーンを実装します。物理ディスクは複数のサーバーまたはアプリケーションで共有されることがあるため、偶発的なアクセスから保護する必要がある点に注意してください。
- ストレージネットワークポロジおよびアプリケーション I/O パターンに基づいて DMP I/O ポリシーを選択します。
- シンプロビジョニングを活用して、投資利益率(ROI)を高めます。

ファイルシステムの推奨事項

次の推奨事項に従うことで、データベースのファイルシステム設計を最適化できます。

- VxVMミラー化を使用する場合は、パフォーマンスを高めるためにODMとCFSを使用します。SmartSyncを使ったODMでは、Oracle Resilveringを使ったミラーボリュームの高速リカバリが有効になります。
- Oracle バイナリ、データ、REDO ログ、およびアーカイブログのために別個のファイルシステムを作成します。これにより、データベースのデータファイルストレージで問題が発生した場合、リカバリデータを使用できます。
- REDO ログとアーカイブログは、ローカルファイルシステムではなく、必ず CFS ファイルシステムに配置します。

I/O フェンシングの推奨事項

SF Oracle RAC クラスタでは、Oracle Clusterware が OPROCD による I/O フェンシングを提供するにもかかわらず、Veritas フェンシングがスプリットブレインシナリオを処理します。このため、SF Oracle RAC でフェンシングを有効にすることが必要です。

次の方法をお勧めします。

- 3 つの LUN をコーディネータディスクとして設定します。vxfsentsthdw (1M) ユーティリティを使用して、LUN が SCSI-3 互換であることを確認します。
- DMP および複数の物理アクセスパスを使用して、コーディネータディスクへの高可用性を確保します。
- ストレッチクラスタ内の SF Oracle RAC の場合は、第 3 の場所に 3 つ目のコーディネータを配置して、サイト障害を伴うスプリットブレインシナリオでのクラスタの可用性を確保します。

SF Oracle RAC 5.0 MP3 以降、iSCSI デバイスを I/O フェンシング用のコーディネータディスクとして使用できます。

メモ: SF Oracle RAC は、ディスクポリシーが DMP に設定されている場合のみ、I/O フェンシング用の iSCSI デバイスをサポートします。iSCSI デバイスを使う場合は、`/etc/vxfsentsthdw` ファイルのディスクポリシーが DMP に設定されていることを確認します。

- 次のシマンテック社のテクニカルノートを参照して、スプリットブレインのリスクを軽減するために使用できるさまざまなオプションを確認します。

<http://support.veritas.com/docs/252635>

ディザスタリカバリの注意事項

SF Oracle RAC は、マルチサイトクラスタに対して、キャンパスクラスタやグローバルクラスタなど、さまざまなディザスタリカバリ設定を提供します。マルチサイトクラスタでは、建物

のさまざまな場所、別々の建物、または別々の都市にノードを配置できます。ノード間の距離は、保護を必要とする災害の種類およびデータをレプリケートするために使用されるテクノロジーによって異なります。**SF Oracle RAC** は、データレプリケーション用のさまざまなレプリケーションテクノロジーをサポートします。

災害によって引き起こされる停止からクラスタを保護するには、クラスタコンポーネントを地理的に分離する必要があります。

キャンパスクラスタ

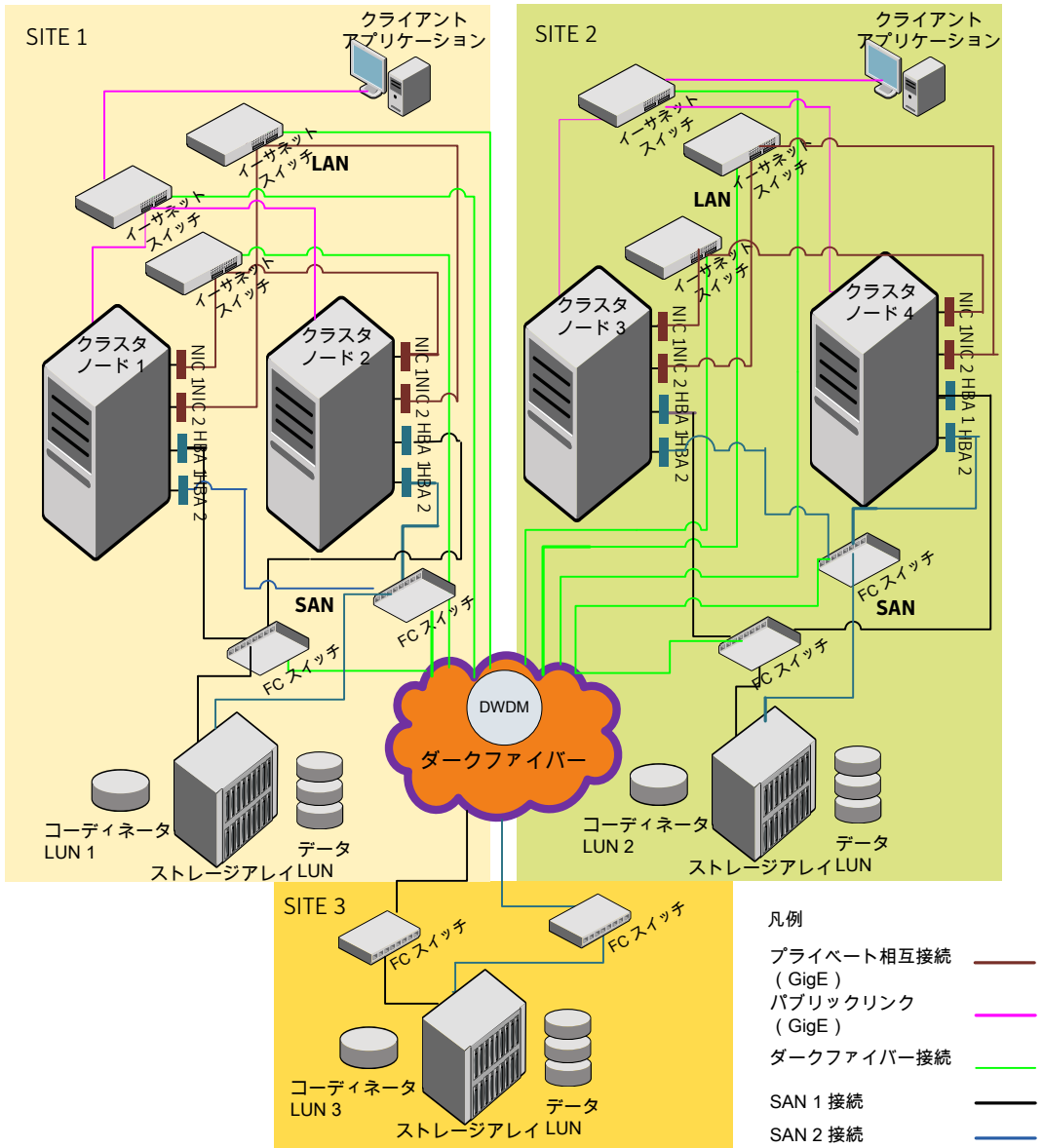
キャンパスクラスタは、ストレッチクラスタまたはリモートミラー設定とも呼ばれます。キャンパスクラスタでは、クラスタのホストおよびストレージは、数マイル間隔の複数のサイトにまたがります。これらのサイトは通常、ストレージへのアクセスとクラスタノード間のプライベートリンク通信を提供する、冗長性を備えた大容量ネットワークを使用して接続されます。

SF Oracle RAC キャンパスクラスタを設定するときは、次のベストプラクティスを考慮してください。

- 2 ノードのキャンパスクラスタの場合は、第 3 のサイトに 3 つ目のコーディネータディスクを配置します。第 3 のサイトでは、ダークファイバーで接続された **FC-SAN** ディスクの代わりに **iSCSI** ディスクを使用できます。
- サイト対応ブックス切断のためのサイト一貫性機能を有効にします。サイト一貫性のタグ付けはエンクローチャレベルで使用します。
- ボリュームがサイト間で自動的にミラー化されるよう、ボリュームで「**allsites**」属性を設定します。
- ディスクグループでグローバル切断ポリシーを設定します。ローカル切断ポリシーは **SF Oracle RAC** キャンパスクラスタではサポートされません。

図 1-3 に、4 ノードの **SF Oracle RAC** キャンパスクラスタの物理的な設定を示します。

図 1-3 4 ノードの SF Oracle RAC キャンパスクラスタ



グローバルクラスタ

グローバルクラスタは、災害が発生した場合に、地理的に分散したクラスタ間でアプリケーションをフェールオーバーする機能を提供します。

グローバルなクラスタ化には 2 つのステップがあります。

1. サイト間でのデータのレプリケーション
2. 災害発生時のアプリケーションの移行

ディザスタリカバリソリューションを設計するときは、次の要素を考慮する必要があります。

- 災害発生時に失われるデータの量
- 災害発生後の許容リカバリ時間

データレプリケーションの注意事項

レプリケーションソリューションを選択するとき、考慮する必要がある重要な要素の 1 つはデータスループットの必要レベルです。データスループットとは、アプリケーションに対して期待されるデータの書き込み速度です。レプリケーションに対する書き込み操作の影響は読み取り操作よりも重要です。

前に説明したビジネスニーズに加えて、レプリケーションオプションを選択するときは、次の要素を考慮する必要があります。

- レプリケーションのモード
- ネットワーク帯域幅
- 2 つのサイト間のネットワーク遅延
- リモートサイトが最初のサイトでのデータ変更に遅れずに対応できること

データレプリケーションオプション

SF Oracle RAC は、さまざまなハードウェアおよびソフトウェアレプリケーションテクノロジーをサポートします。

- ハードウェアレプリケーションオプション
 - Hitachi True Copy
 - IBM Metro Mirror
 - IBM SVC
 - EMC Mirror View

- ソフトウェアレプリケーションオプション
 - Veritas Volume Replicator (VVR)
 - Oracle Data Guard

リファレンス

SF Oracle RAC について詳しくは、次のマニュアルを参照してください。

Veritas Storage Foundation for Oracle RAC 5.0 sfrac_install.pdf

MP3 インストールガイド

SF Oracle RAC マニュアル CD に含まれて
います

Veritas Storage Foundation for Oracle RAC 5.0 sfrac_admin.pdf

MP3 インストールガイド

SF Oracle RAC マニュアル CD に含まれて
います

推奨される SF Oracle RAC 配備シナリオ

この付録では以下の項目について説明しています。

- VCS IPC および PrivNIC エージェントを含む SF Oracle RAC クラスタ
- UDP IPC および MultiPrivNIC エージェントを含む SF Oracle RAC クラスタ
- UDP IPC および MultiPrivNIC エージェントを含む複数のデータベース用の SF Oracle RAC
- 分離された Oracle トラフィック および MultiPrivNIC エージェントを含む SF Oracle RAC クラスタ
- NIC 結合、VCS IPC、および PrivNIC エージェントを含む SF Oracle RAC クラスタ
- NIC 結合、UDP IPC、および PrivNIC エージェントを含む SF Oracle RAC クラスタ

VCS IPC および PrivNIC エージェントを含む SF Oracle RAC クラスタ

この項では、次のシナリオに対して推奨される設定について説明します。

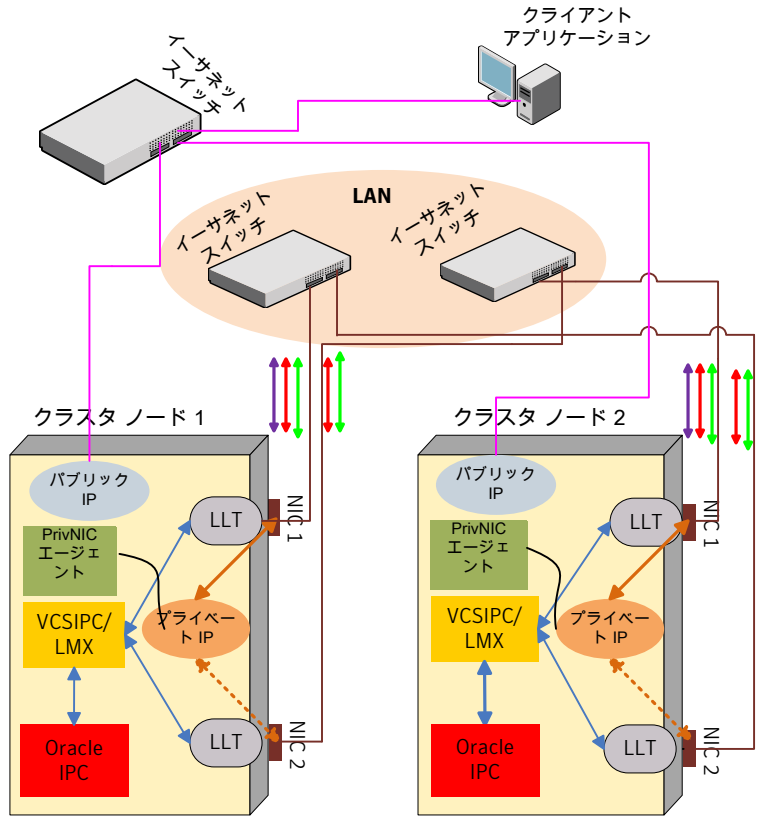
配備シナリオ	Oracle RAC 10g データベース Cache Fusion トラフィックが、LMX/LLT 上の VCS IPC を使用して複数の LLT リンクに分散されます。
推奨事項	PrivNIC エージェントを使用します。

この項で説明する設定の概要は次のとおりです。

- Oracle は、Cache Fusion のために LMX/LLT 上で VCS IPC を使用します。

- 1つの共通プライベートネットワーク IP アドレスが、Oracle データベース Cache Fusion と Oracle Clusterware 通信に使用されます。
 - Oracle Clusterware 通信は、LLT プライベート相互接続リンクのいずれかで行われます。
 - CFS/CVM/VCS メタデータおよび Oracle データベース Cache Fusion トラフィックは、LLT リンクとして設定されている 2 つの物理リンク上で LLT を介して行き来します。
 - リンクに障害が発生した場合、PrivNIC は障害が発生したリンクから使用可能な LLT リンクへとプライベートネットワーク IP アドレスをフェールオーバーします。
- 図 A-1 に、VCS IPC および PrivNIC エージェントを含む 2 ノードの SF Oracle RAC クラスタの論理設定を示します (Oracle RAC 10g)。

図 A-1 VCS IPC および PrivNIC エージェントを含む SF Oracle RAC クラスタ (Oracle RAC 10g)



凡例

- | | | | |
|---------------------------|------------------|------------------------------------|--------------------------------|
| パブリックリンク (GigE) | — (Magenta line) | Oracle inter-process communication | ↔ (Blue double-headed arrow) |
| プライベート相互接続 (GigE) | — (Brown line) | アクティブな接続 | ↔ (Orange double-headed arrow) |
| Oracle Clusterware ハートビート | — (Purple line) | PrivNIC エージェント用のフェールオーバー接続 | — (Dashed orange arrow) |
| Oracle DB Cache Fusion | — (Red line) | | |
| CFS/CVM/VCS メタデータ | — (Green line) | | |

UDP IPC および MultiPrivNIC エージェントを含む SF Oracle RAC クラスタ

この項では、次のシナリオに対して推奨される設定について説明します。

配備シナリオ	データベース Cache Fusion 用の UDP IPC を使って設定されている Oracle RAC 10g または Oracle RAC 11g。
推奨事項	MultiPrivNIC エージェントを使用します。

この項で説明する設定の概要は次のとおりです。

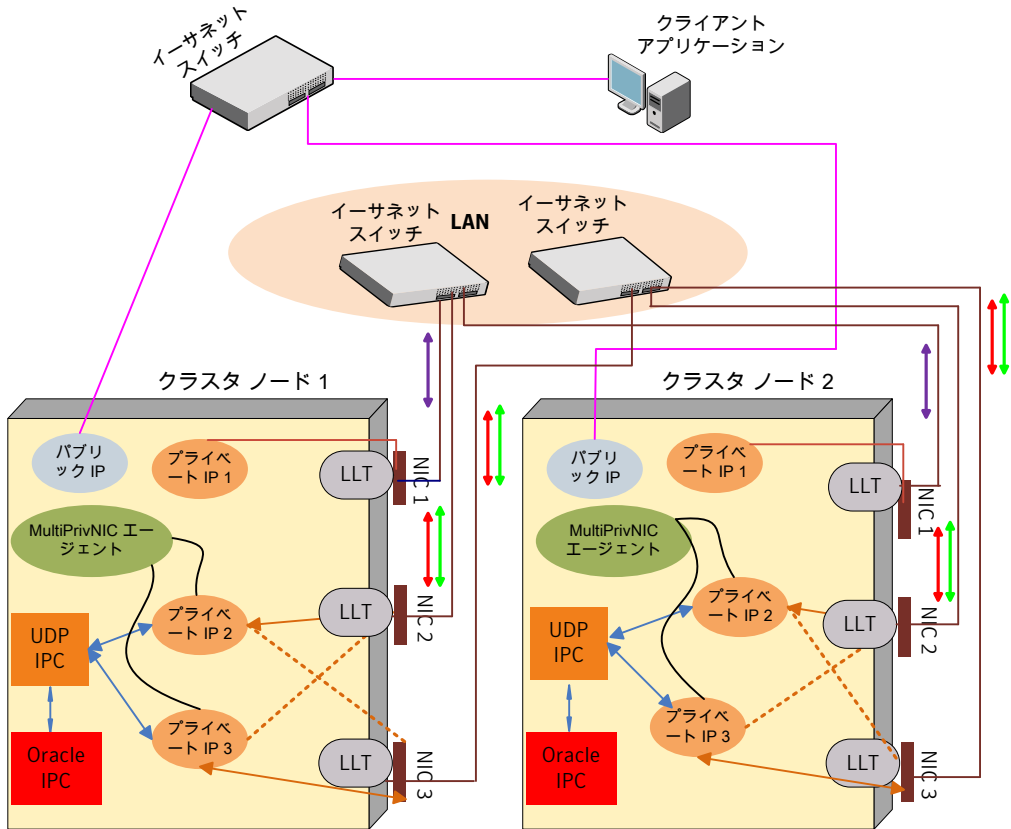
- 1 つの IP アドレスが Oracle Clusterware 通信に使用されます。
- 1 つまたは複数の IP アドレスが、LLT リンクとして設定されている 2 つの物理リンクに分散される Oracle データベース Cache Fusion に使用されます。CFS/CVM/VCS メタデータもこれらのリンクを介して行き来します。

メモ: Oracle データベース Cache Fusion に使用される複数のプライベートネットワーク IP アドレスは、Oracle データベース初期化パラメータ `CLUSTER_INTERCONNECT` を使用して定義されます。`CLUSTER_INTERCONNECT` パラメータが指定されない場合は、Oracle Cache Fusion に使用されるのと同じプライベートネットワーク IP アドレスが Oracle Clusterware 通信に使用されます。

- Oracle Clusterware 通信は、LLT プライベート相互接続リンクのいずれかで行われます。
- リンクに障害が発生した場合、MultiPrivNIC は障害が発生したリンクからその他の使用可能な LLT リンクへとプライベートネットワーク IP アドレスをフェールオーバーします。

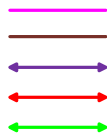
図 A-2 に、UDP IPC および MultiPrivNIC エージェントを含む 2 ノードの SF Oracle RAC クラスタの論理設定を示します (Oracle RAC 10g/Oracle RAC 11g)。

図 A-2 UDP IPC および MultiPrivNIC エージェントを含む SF Oracle RAC クラスタ (Oracle RAC 10g/Oracle RAC 11g)

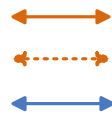


凡例

- パブリックリンク (GigE)
- プライベート相互接続 (GigE)
- Oracle Clusterware ハートビート
- Oracle DB Cache Fusion
- CFS/CVM/VCS メタデータ



- アクティブな接続
- MultiPrivNIC エージェント用のフェールオーバー接続
- Oracle プロセス間通信



UDP IPC および MultiPrivNIC エージェントを含む複数のデータベース用の SF Oracle RAC

この項では、次のシナリオに対して推奨される設定について説明します。

配備シナリオ	負荷分散のために、複数の Oracle RAC 10g または Oracle RAC 11g データベースが複数の LLT リンクで設定されています。
推奨事項	MultiPrivNIC エージェントを使用します。

この項で説明する設定の概要は次のとおりです。

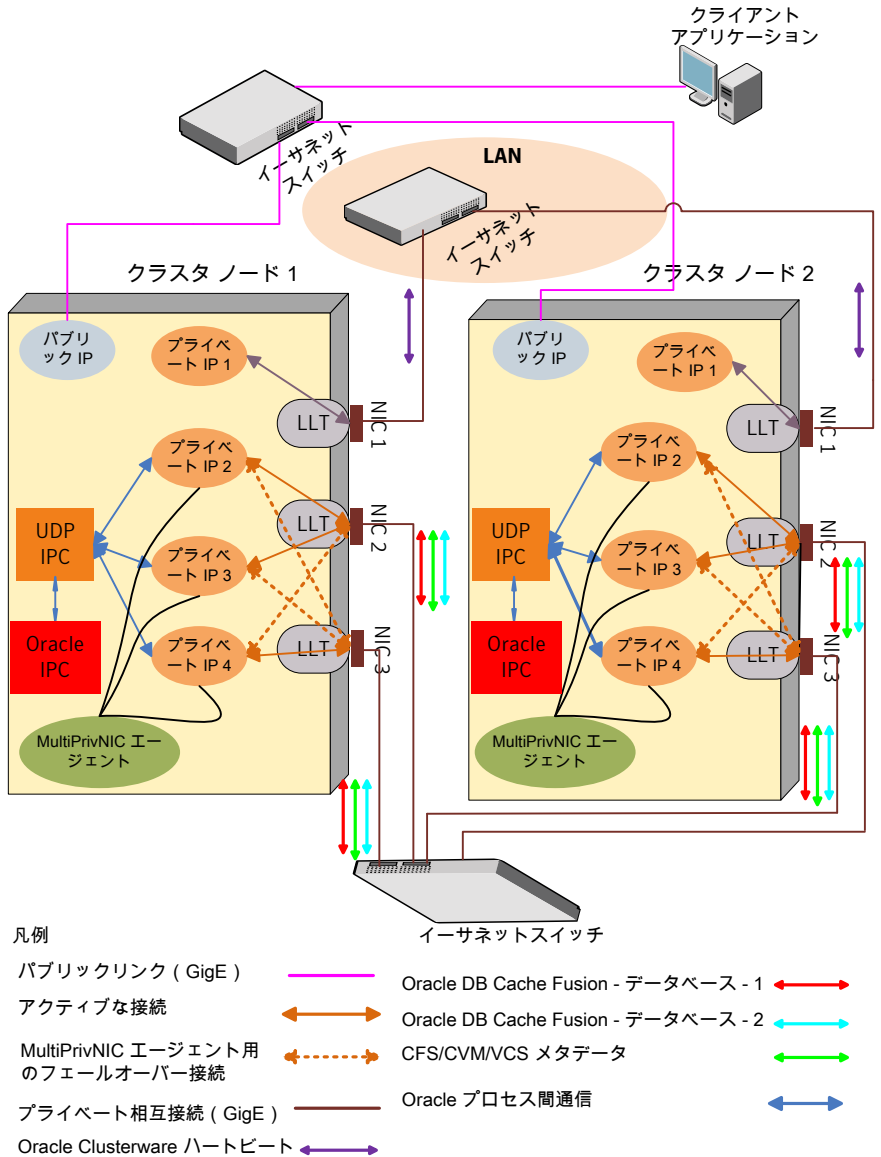
- 1 つの IP アドレスが Oracle Clusterware 通信に使用されます。
- 1 つまたは複数の IP アドレスが、LLT リンクとして設定されている複数の物理リンクに分散される Oracle データベース Cache Fusion に使用されます。CFS/CVM/VCS メタデータもこれらのリンクを介して行き来します。

メモ: Oracle データベース Cache Fusion に使用される複数のプライベートネットワーク IP アドレスは、Oracle データベース初期化パラメータ `CLUSTER_INTERCONNECT` を使用して定義されます。`CLUSTER_INTERCONNECT` パラメータが指定されない場合は、Oracle Cache Fusion に使用されるのと同じプライベートネットワーク IP アドレスが Oracle Clusterware 通信に使用されます。

- Oracle Clusterware 通信は、LLT プライベート相互接続リンクのいずれかで行われます。
- リンクに障害が発生した場合、MultiPrivNIC は障害が発生したリンクからその他の使用可能な LLT リンクへとプライベートネットワーク IP アドレスをフェールオーバーします。

図 A-3 に、UDP IPC および MultiPrivNIC エージェントを含む複数のデータベース用の 2 ノードの SF Oracle RAC クラスタの論理設定を示します (Oracle RAC 10g/Oracle RAC 11g)。

図 A-3 UDP IPC および MultiPrivNIC エージェントを含む複数のデータベース用の SF Oracle RAC クラスタ(Oracle RAC 10g/Oracle RAC 11g)



分離された Oracle トラフィック および MultiPrivNIC エージェントを含む SF Oracle RAC クラスタ

この項では、次のシナリオに対して推奨される設定について説明します。

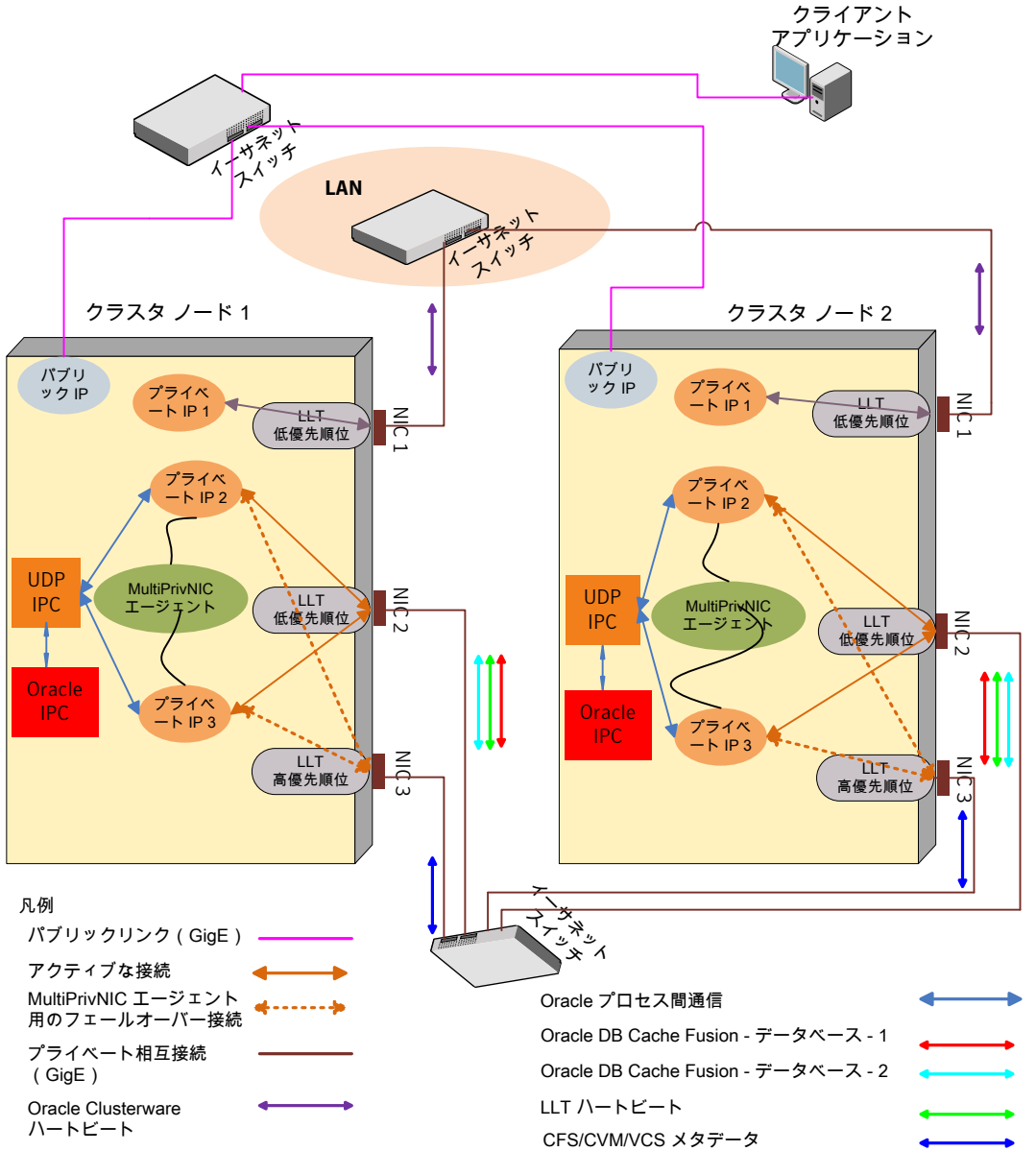
配備シナリオ	Oracle RAC 10g または Oracle RAC 11g データベース Cache Fusion トラフィックが CFS/CVM/VCS メタデータから分離されます。
推奨事項	MultiPrivNIC エージェントを使用します。

この項で説明する設定の概要は次のとおりです。

- 1 つの IP アドレスが Oracle Clusterware 通信に使用されます。この IP アドレスは、優先度の低い LLT リンクに設定されます。
- 1 つまたは複数の IP アドレスが、優先度の低い LLT リンクとして設定されている 2 つの物理リンクに分散される Oracle データベース Cache Fusion に使用されます。LLT ハートビートもこのリンクを介して行き来します。
- Oracle Clusterware 通信は、優先度の低い LLT プライベート相互接続リンクのいずれかで行われます。
- CFS/CVM/VCS メタデータは、優先度の高い LLT リンクとして設定されている 3 つ目の物理リンクを介して行き来します。
- 優先度の低い LLT リンクのいずれかで障害が発生した場合、MultiPrivNIC エージェントは、障害が発生した優先度の低いリンクからその他の使用可能な優先度の低い LLT リンクへとプライベートネットワーク IP アドレスをフェールオーバーします。
- 優先度の低い LLT リンクの両方で障害が発生した場合、MultiPrivNIC エージェントは、障害が発生した優先度の低いリンクからその他の使用可能な優先度の高い LLT リンクへとプライベートネットワーク IP アドレスをフェールオーバーするため、Oracle データベース Cache Fusion トラフィックと CFS/CVM/VCS メタデータがマージされます。

図 A-4 に、VCS/CVM/CFS トラフィックから分離された Oracle トラフィックを含む 2 ノードの SF Oracle RAC クラスタの論理設定を示します (Oracle RAC 10g/Oracle RAC 11g)。

図 A-4 分離された Oracle トラフィック および MultiPrivNIC エージェントを含む SF Oracle RAC クラスタ (Oracle RAC 10g/Oracle RAC 11g)



NIC 結合、VCS IPC、および PrivNIC エージェントを含む SF Oracle RAC クラスタ

この項では、次のシナリオに対して推奨される設定について説明します。

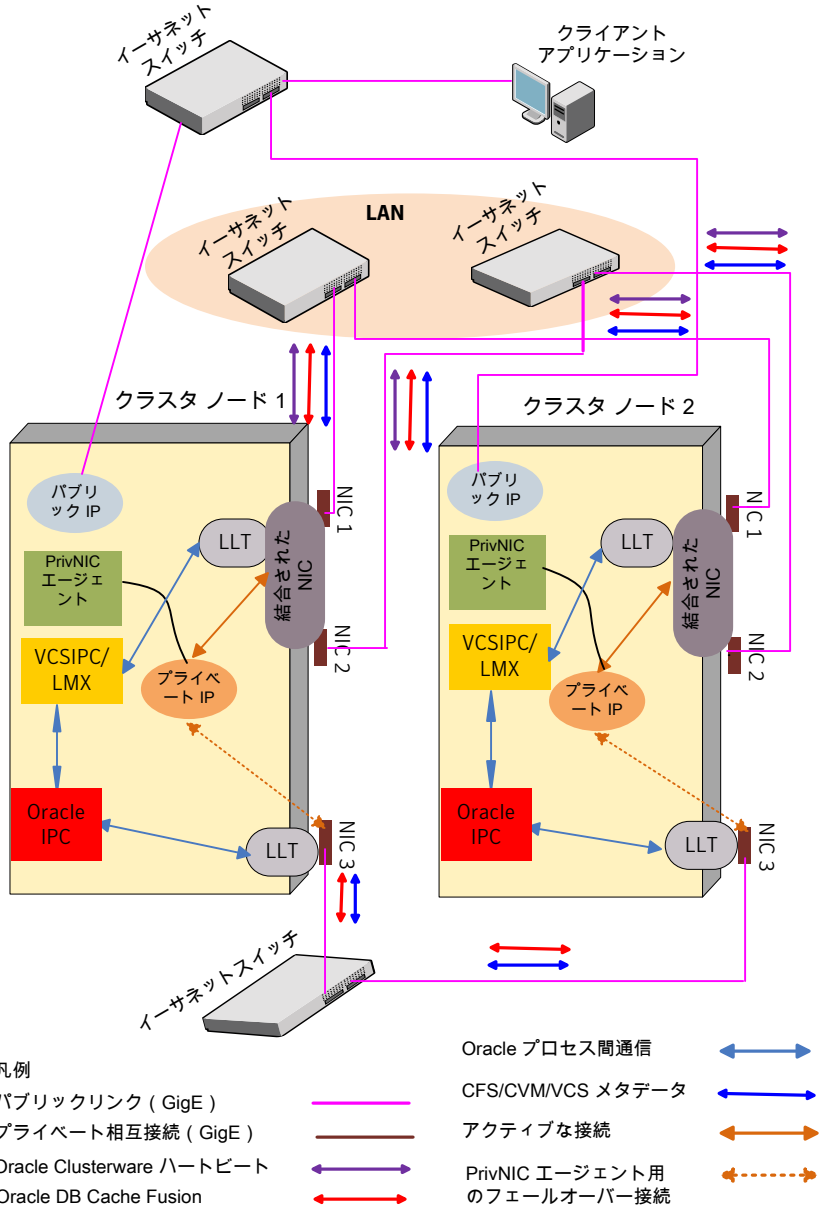
配備シナリオ	LMX/LLT 上で VCS IPC を使用して Oracle RAC 10g データベース Cache Fusion トラフィックを分散するために、VCS IPC 結合された NIC インターフェースが別の NIC インターフェースと共に使用されます。
推奨事項	PrivNIC エージェントを使用します。

この項で説明する設定の概要は次のとおりです。

- Oracle は、Cache Fusion のために LMX/LLT 上で VCS IPC を使用します。
- Oracle データベース Cache Fusion と、LLT リンクとして設定されている結合された NIC インターフェース上で行われる Oracle Clusterware 通信に、1 つの共通プライベートネットワーク IP アドレスが使用されます。
- Oracle Clusterware 通信は、その他の LLT プライベート相互接続で行われます。
- Oracle データベース Cache Fusion および CFS/CVM/VCS メタデータは、3 つの物理リンク上で LLT を介して行き来します。
- 結合された NIC インターフェースに障害が発生した場合、PrivNIC は障害が発生したリンクからその他の使用可能な LLT リンクへとプライベートネットワーク IP アドレスをフェールオーバーします。

[図 A-5](#) に、NIC 結合、VCS IPC、および PrivNIC エージェントを含む 2 ノードの SF Oracle RAC クラスタの論理設定を示します (Oracle RAC 10g)。

図 A-5 NIC 結合、VCS IPC、および PrivNIC エージェントを含む SF Oracle RAC クラスタ



NIC 結合、UDP IPC、および PrivNIC エージェントを含む SF Oracle RAC クラスタ

この項では、次のシナリオに対して推奨される設定について説明します。

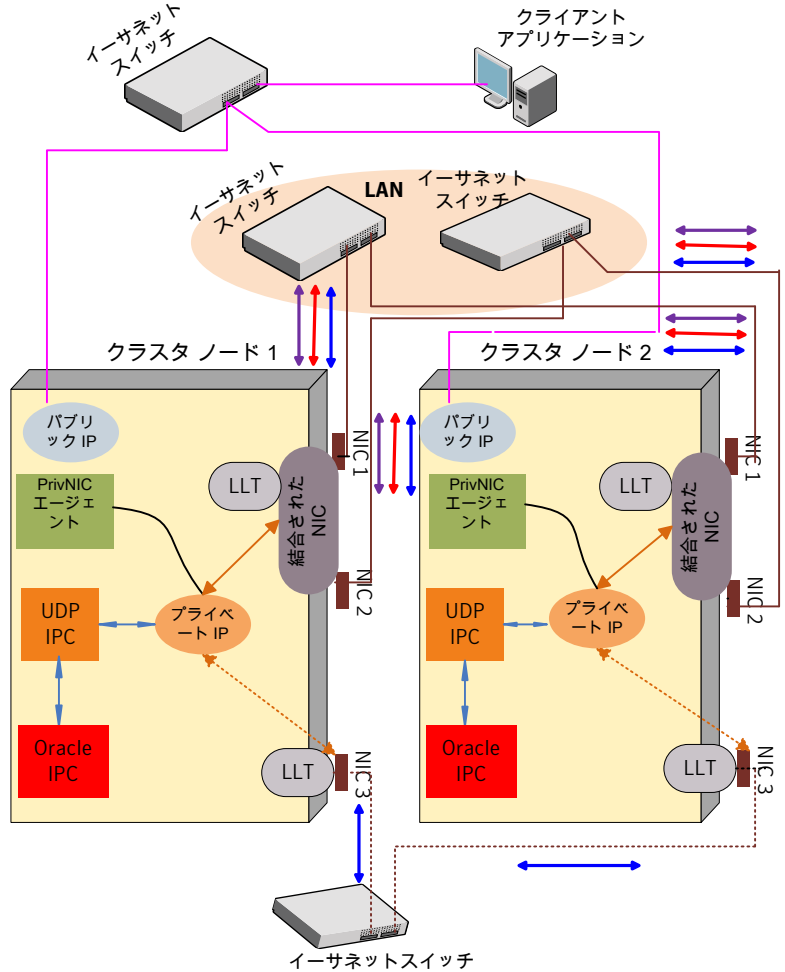
配備シナリオ	Oracle RAC 10g または Oracle RAC 11g と UDP IPC が、Oracle データベース Cache Fusion トラフィックの分散のために結合された NIC インターフェイスを使用するように設定されています。 2 つ目のリンクはスタンバイリンクとして設定されます。
推奨事項	PrivNIC エージェントを使用します。

この項で説明する設定の概要は次のとおりです。

- Oracle Clusterware 通信と、結合された NIC インターフェースの基盤となる 2 つの物理リンクに分散される Oracle データベース Cache Fusion に、共通の IP アドレスが使用されます。結合された NIC インターフェースは単一の LLT リンクとして設定されます。
- CFS/CVM/VCS メタデータも結合されたリンクを介して行き来します。
- 結合されたリンクに障害が発生した場合、PrivNIC は障害が発生したリンクから使用可能なスタンバイ LLT リンクへとプライベートネットワーク IP アドレスをフェールオーバーします。

図 A-6 に、NIC 結合、UDP IPC、および PrivNIC エージェントを含む 2 ノードの SF Oracle RAC クラスタの論理設定を示します。

図 A-6 NIC 結合、UDP IPC、および PrivNIC エージェントを含む SF Oracle RAC クラスタ



凡例

- パブリックリンク (GigE)
- プライベート相互接続 (GigE)
- Oracle Clusterware ハートビート
- Oracle DB Cache Fusion
- スタンバイプライベート相互接続 (GigE)

- イーサネットスイッチ
- Oracle プロセス間通信
- CFS/CVM/VCS メタデータ アクティブな接続
- PrivNIC エージェント用のフェールオーバー接続

38 | 付録 A 推奨される SF Oracle RAC 配備シナリオ
NIC 結合、UDP IPC、および PrivNIC エージェントを含む SF Oracle RAC クラスター