



Oracle Database Innovation 革新し続ける、
クラウドに最適化されたデータベース
～ データベースクラウドを支える最新テクノロジーの全貌 ～

データベースクラウド環境の 高可用性を実現するOracle MAAの全貌

～ Oracle RAC、ASM、Data Guard、Flashback Technology ～

2015年10月23日
日本オラクル株式会社
クラウド・テクノロジー事業統括 クラウド・テクノロジー製品戦略統括本部
データベースエンジニアリング本部 応用技術グループ
シニアマネジャー
柴田 長



以下の事項は、弊社の一般的な製品の方向性に関する概要を説明するものです。また、情報提供を唯一の目的とするものであり、いかなる契約にも組み込むことはできません。以下の事項は、マテリアルやコード、機能を提供することをコミットメント(確約)するものではないため、購買決定を行う際の判断材料になさらないで下さい。オラクル製品に関して記載されている機能の開発、リリースおよび時期については、弊社の裁量により決定されます。

OracleとJavaは、Oracle Corporation 及びその子会社、関連会社の米国及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。

自己紹介

"しばちょう"こと柴田長(しばたつかさ)と申します。

日本オラクル株式会社
クラウド・テクノロジー事業統括 クラウド・テクノロジー製品戦略統括本部
データベースエンジニアリング 応用技術グループ
シニアマネジャー
柴田 長



Oracle Technology Networkで毎月連載中
「しばちょう先生の試して納得! DBAへの道」

<http://www.oracle.com/technetwork/jp/database/articles/shibacho/index.html>

Program Agenda

- 1 Oracle Maximum Availability Architecture
- 2 Oracle Real Application Clusters
- 3 Oracle Automatic Storage Management
- 4 Oracle Data Guard
- 5 Flashback Technology

A woman with long brown hair and glasses is sitting at a wooden table in a cafe or office setting. She is wearing a brown leather jacket over a blue patterned scarf. She is holding a black mobile phone to her ear with her left hand and looking down at a newspaper or magazine on the table with her right hand. The background is slightly blurred, showing other people and tables in the same environment.

Oracle Maximum Availability Architecture

ITインフラに求められる主要要件

データベース・クラウド環境であっても普遍

1. 高可用性

2. パフォーマンス(拡張性)

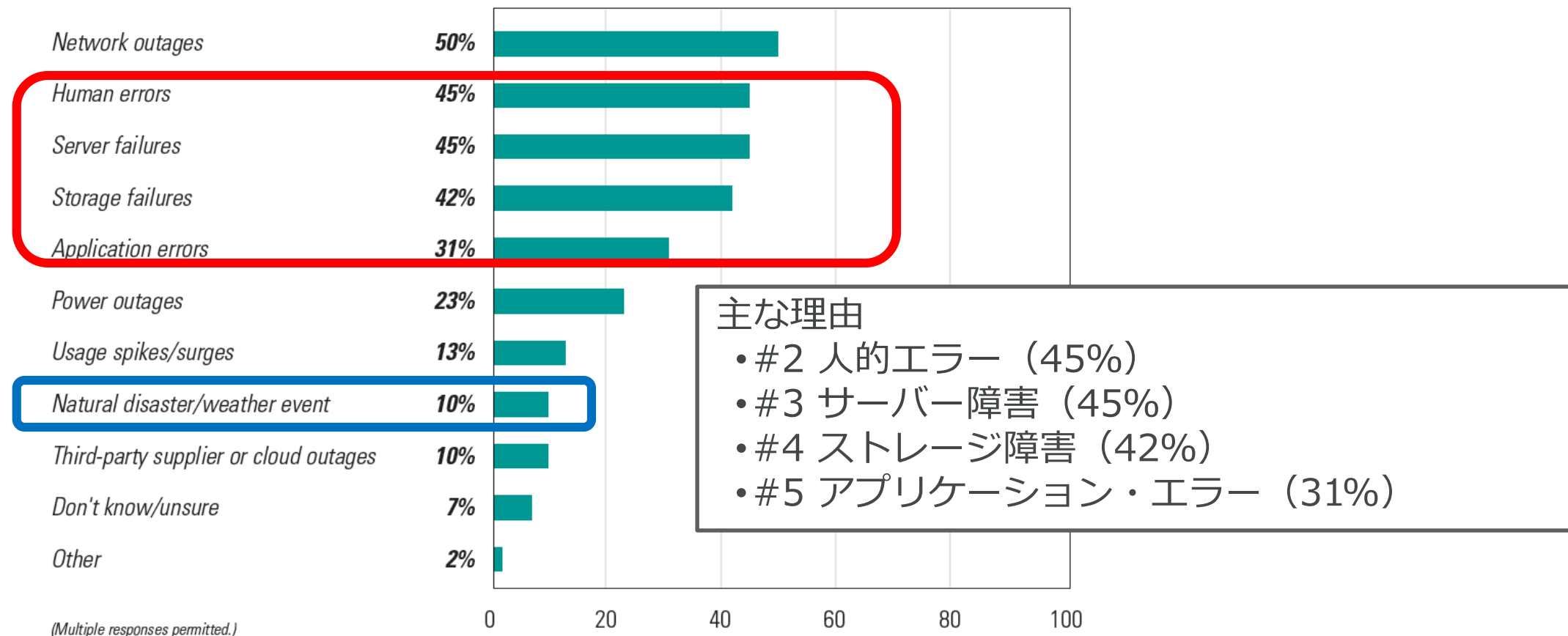
3. 管理性(運用容易性)

4. セキュリティ

- いくら高性能で拡張性が高くても、足回りがガッチリ(高可用性)していなければ、宝の持ち腐れになりかねない

計画”外”停止の主な要因

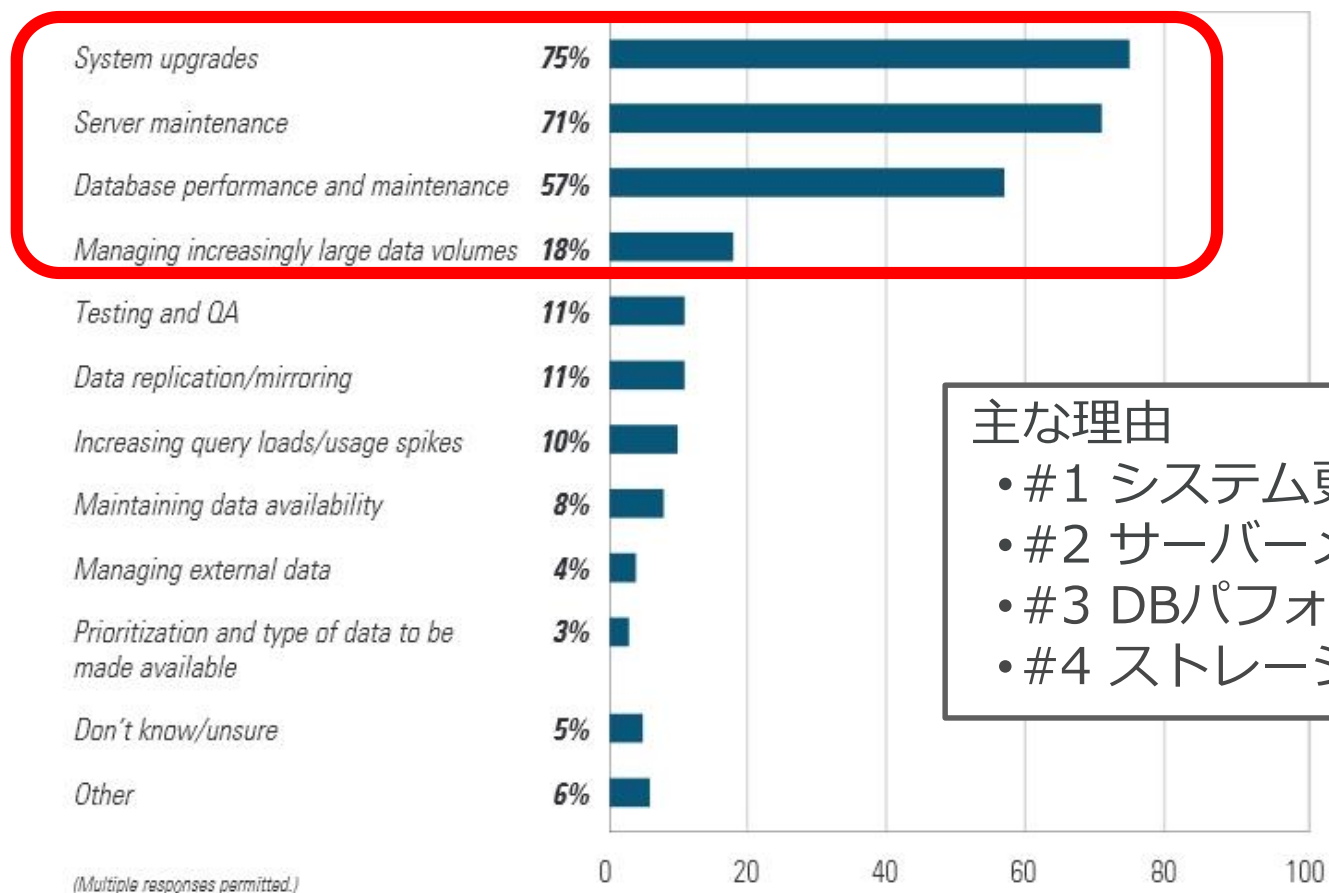
2012 IOUG Database Availability Survey アンケート結果より



計画停止の主要要因

2013 IOUG Database Availability Survey アンケート結果より

Leading Causes of Planned Downtime Over Past Three Years



主な理由

- #1 システム更改 (75%)
- #2 サーバーメンテナンス (71%)
- #3 DBパフォーマンス&メンテナンス (57%)
- #4 ストレージ管理 (18%)

Oracle Maximum Availability Architecture

<http://www.oracle.com/technetwork/jp/content/maa-094615-ja.html>

- Oracle Maximum Availability Architecture (MAA) とは、Oracle 開発チームの**実証済み**高可用性テクノロジーと、顧客の成功事例に基づいたOracleの**ベスト・プラクティスのブループリント**
 - アーキテクチャやベストプラクティス、顧客事例、デモなど詳細情報をWebサイトにて掲載
- MAAの目的
 - 最適な高可用性アーキテクチャの設計から複雑な仕組みを排除すること
 - ハードウェアやOSの影響を受けない
 - サーバーとストレージのコスト削減に利用できる
 - Oracle の新バージョンや新機能に適応できる
 - 計画停止を極小化し、計画外停止を回避、検出および修復するための**ベスト・プラクティス**を提供

Oracle Maximum Availability Architecture

Production Site

RAC

- Scalability
- Server HA

ASM

- Local storage protection

Advanced Security

- Data encryption, protection

Flashback

- Human error correction

Edition-based Redefinition, Online Redefinition, Data Guard, GoldenGate

- Minimal downtime maintenance, upgrades, migrations

Enterprise Manager Cloud Control

- Site Guard, Coordinated Site Failover

Application Continuity

- Application HA

Global Data Services

- Service Failover / Load Balancing



Application Test Suite, Real Application Testing

- Minimal Testing Costs

Active Replica

Active Data Guard

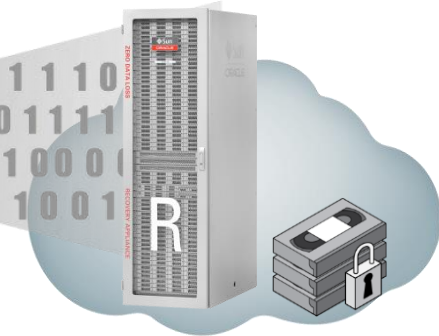
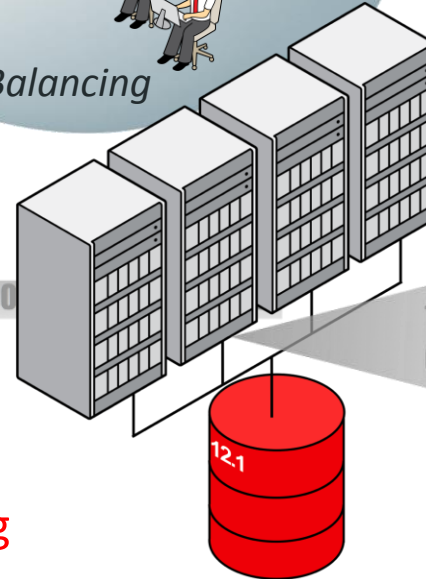
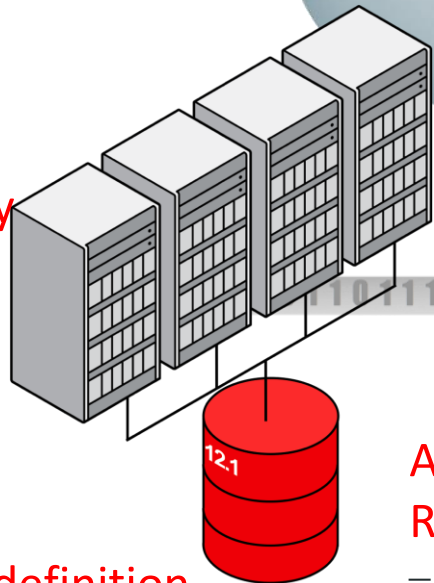
- Data Protection, DR
- Query Offload

GoldenGate

- Active-active replication
- Heterogeneous

RMAN, Oracle Secure Backup, Zero Data Loss Recovery Appliance

- Backup to disk, tape or cloud



Oracle Maximum Availability Architecture

計画停止/計画外停止のダウンタイムを解決

← Oracle Maximum Availability Architecture →

Media and
Storage
Failures

Server Failure

Database,
System,
Cluster,
& Site
Outages

Human or
Application
Error

Flexible
Maintenance
and
Migrations

Oracle Maximum Availability Architecture

計画停止/計画外停止のダウンタイムを解決

← Oracle Maximum Availability Architecture →

Automatic Storage Management (ASM)

Recovery Manager (RMAN),

Oracle Secure Backup

- Storage failure
- Data recovery
- Backups

Oracle RAC

- Instance failure
- Server failure
- Rolling maintenance
- Performance scale-out
- Online instance relocation
- Consolidation

Data Guard

- Database failure
- System failure
- Site failure
- Zero data loss
- Automatic failover
- Corruption protection
- Rolling upgrade
- Read-only offload
- Backup offload

Flashback

- Fast point-in-time recovery
- Granular repair of logical corruptions
 - Transaction
 - Table
 - Database

GoldenGate

- Heterogeneous migrations
- Bi-directional and multi-master replication
- Zero downtime maintenance

Edition Based Redefinition

- Zero downtime application upgrade

停止時間の原因とMAAの対処機能

計画”外”停止 - 1/2

タイプ	障害箇所	対処策/機能	RTO
クラスタ全体の障害	<ul style="list-style-type: none">•クラスタ内の全サーバー停止•インターコネクフル障害•クラスタウェア障害•データベース破損	<ul style="list-style-type: none">•Data Guard によるフェイルオーバー (同一サイト、リモートサイト)	数分
単一ノード障害	<ul style="list-style-type: none">•OS障害•ハードウェア障害•NIC障害•インスタンス障害	<ul style="list-style-type: none">•Data Guardによるフェイルオーバー	数分
		<ul style="list-style-type: none">•RAC/RAC One Node によるフェイルオーバー	数秒
		<ul style="list-style-type: none">•GoldenGate/Streams で複製済みDBへ切替	数分
ストレージ障害	<ul style="list-style-type: none">•ディスク・ドライブ障害•ディスク・コントローラ障害•ストレージ・アレイ障害	<ul style="list-style-type: none">•ASM Mirroring → 自動リバランス	ゼロ
		<ul style="list-style-type: none">•RMAN Backup → Restore + Recovery	数十分～
		<ul style="list-style-type: none">•Data Guard によるフェイルオーバー	数分
		<ul style="list-style-type: none">•GoldenGate/Streams で複製済みDBへ切替	数分

停止時間の原因とMAAの対処機能

計画”外”停止 - 2/2

タイプ	障害箇所	対処策/機能	RTO
データ破損	•HBA障害 •ソフトウェア不具合 •ディスク・コントローラ障害 •ボリュームマネージャーのエラー •OS、デバイスドライバ不具合	•ASM Mirroring	ゼロ
		•DB_BLOCK_CHECKSUM/CHECKINGの設定 → ADGによるAuto Block Media Recovery → RMAN BackupからBlock単位で手動修復	数秒
		•Data Guard によるフェイルオーバー	数分
		•GoldenGate/Streams で複製済みDBへ切替	数分
書込み欠落	•同上	•DB_LOST_WRITE_PROTECT の設定 → Data Guard によるフェイルオーバー	数分
人的エラー	•データベース・オブジェクト削除 •誤った/悪意なデータ変更	•Flashback Technology	数秒～
	•データファイルの削除	•RMAN Backup → Restore + Recovery	数十分～

停止時間の原因とMAAの対処機能

計画停止

タイプ	説明	対処策/機能	RTO
システム変更	•ストレージ移行、ディスク追加	•ASM Rebalancing	ゼロ
	•クラスタ・ノード追加 •各種パラメータ変更 •データベース再配置 •H/Wメンテナンス	•RAC/Cellノード単位でのローリング作業	ゼロ
		•Data Guardでスイッチ・オーバー	数分
		•GoldenGate/Streams で複製済みDBへ切替	数分
	•アップグレード、パッチ適用	•RAC/Cellノード単位でのローリング・アップグレード	ゼロ
		•Data Guardでローリング・アップグレード	数分
		•GoldenGate/Streams で複製済みDBへ切替	数分

**RAC/Cellノード単位でのローリング適用可否は、
パッチの種類に依存(詳細は以降のスライドを参照)**


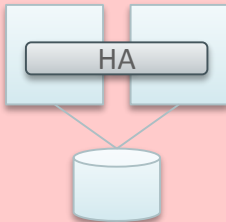
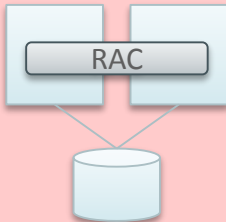
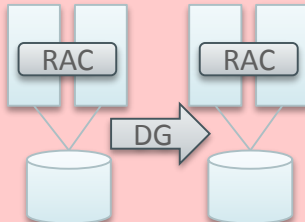
【参考】パッチの種類とリリース頻度

Oracle Database

種類名称	バージョン表記例	リリース頻度	説明
Patch Set Release (PSR)	11.2.0.3.0 → 11.2.0.4.0	1~2年	メジャー・リリースの間で作成され、複数の不具合に対する修正を 統合 したもの。ベース・リリースと過去のパッチ・セットに対して 累積的 なパッチ。
Exadata Bundle Patch (BP)	11.2.0.3.11 → 11.2.0.3.12	1ヶ月	Exadata向け に、PSUを含んだ重要な不具合の修正を統合したもの。PSRに対する 累積 パッチ。
Patch Set Update (PSU)	11.2.0.3.1 → 11.2.0.3.2	3ヶ月	最も重要な修正と最重要セキュリティ修正 (CPU) の内容を含んだ 累積 パッチ。事前公開されたスケジュールに基づき定期的にリリース。
Critical Patch Update (CPU)	N/A	3ヶ月	PSUから セキュリティ修正のみ を抜き出したパッチ。セキュリティ要件だけを満たせば良い環境向け。
個別パッチ、 マージ・パッチ	N/A	適時	ある不具合を修正するために、次のPSU/PSRや新製品のリリースまで待つことができないお客様のために作成される1つ以上の修正を含むパッチ。

パッチの種類と適用方法

計画停止時間の目安 (Oracle Database 11g Release 2~)

Target	Patch Type	RAC Rolling ?	Single	HA	RAC	RAC + DG
						
Database	BP/PSU/CPU	Yes	DB停止後適用 (数分)	F/Oで交互適用 (数分 x 2回)	RACローリング 適用(ゼロ)	
	PSR	No	DB停止後適用 (数十分~数時間)			S/Oで交互適用 (5分未満 x 2回)
Grid Infrastructure (OCW/ASM)	BP/PSU/CPU	Yes	DB停止後適用 (数分~数十分)	F/Oで交互適用 (数分 x 2回)	RACローリング 適用(ゼロ)	
	PSR	Yes	DB停止後適用 (数十分)	F/Oで交互適用 (数分 x 2回)	RACローリング 適用(ゼロ)	
OS	-	Yes	DB停止後適用 (数分~数時間)	F/Oで交互適用 (数分 x 2回)	ローリング適用 (ゼロ)	

A woman with long brown hair and glasses, wearing a brown leather jacket and a blue patterned scarf, is sitting at a wooden table in a cafe. She is holding a black mobile phone to her ear with her left hand and looking down at a newspaper or magazine on the table with her right hand. The background is a blurred cafe interior with other tables and chairs.

Oracle Real Application Clusters

Oracle Real Application Clusters (RAC)

Oracle9i より提供されているクラスタ技術

- 複数ノードで構成する共有ディスク/共有キャッシュ型のデータベース
 - 全ノードが全データに直接アクセス可能
 - Cache Fusion Technologyで、キャッシュ・データの一貫性を維持

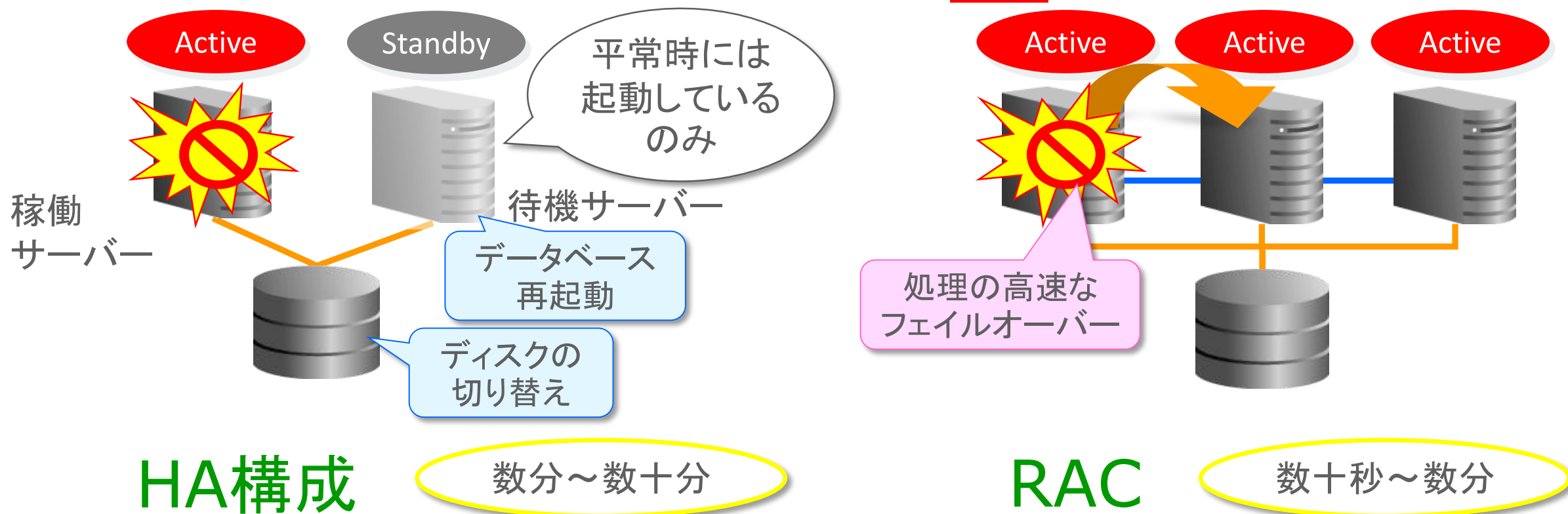
• 主な特徴

拡張性	✓ 負荷の増減に応じた処理性能の最適化が可能
可用性	✓ 高速なフェイルオーバーを実現し、システム障害時のダウンタイムを最小化
投資コスト	✓ 最低限必要な構成で導入でき、初期コストを抑えることが可能 ✓ リソースの有効活用により 最適な投資コストを実現

RACによる高可用性の実現

高速なフェイルオーバー

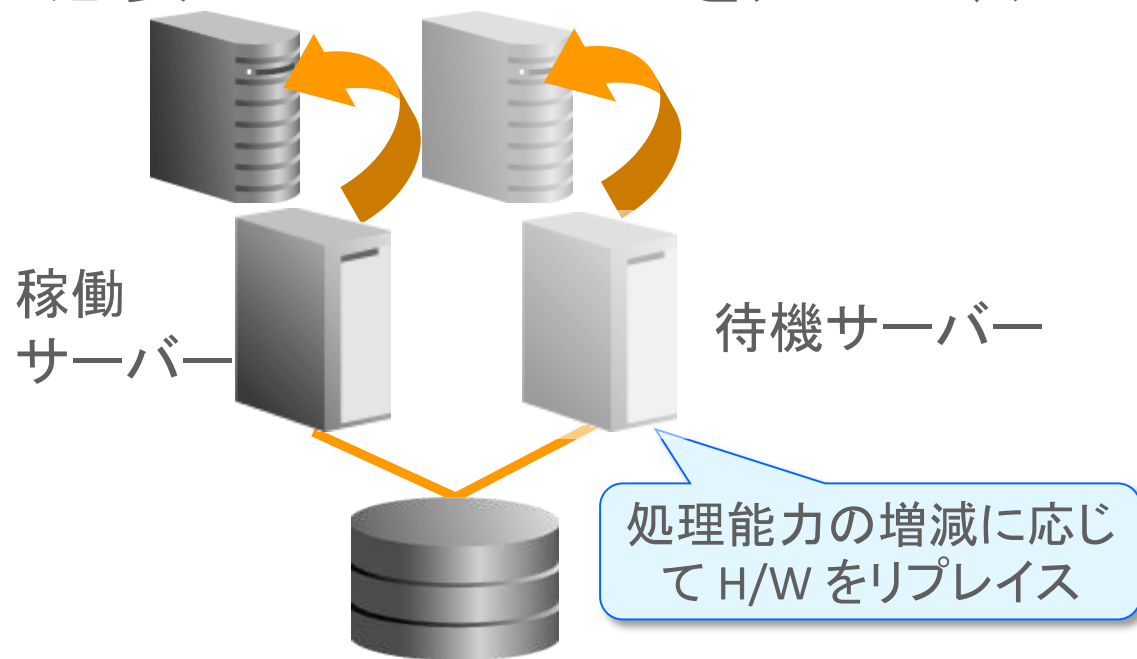
- 障害が発生した場合、リカバリや切り替えを高速に実施



RACによる高拡張性の実現

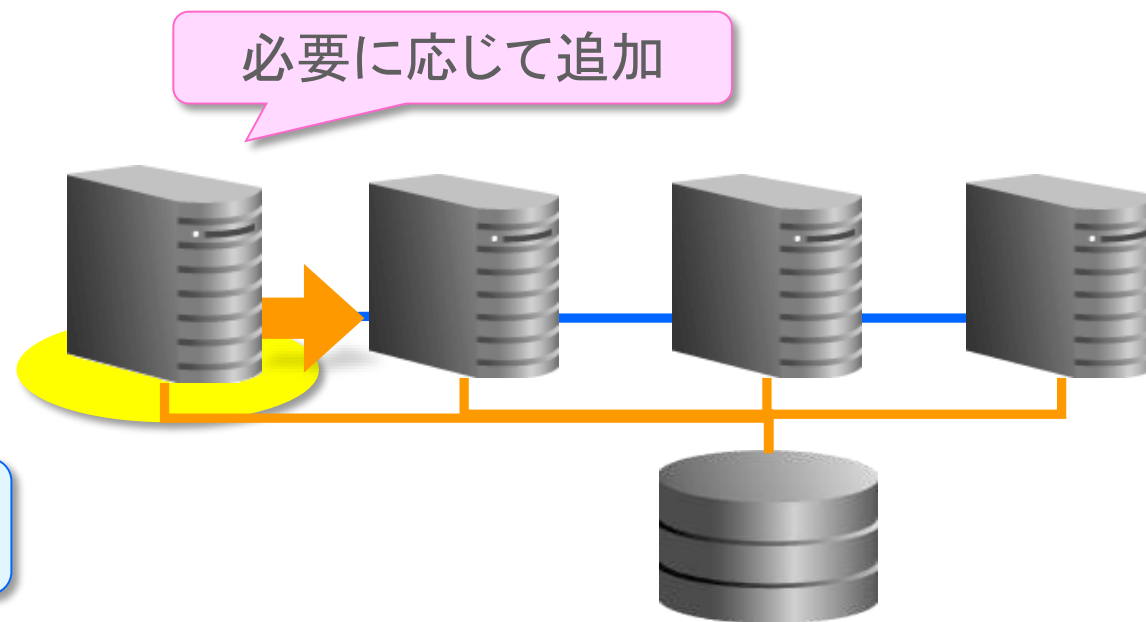
サーバー追加によるスケーラビリティの向上

- 必要に応じてサーバを追加し、処理能力の拡張が可能



HA構成

入れかえ

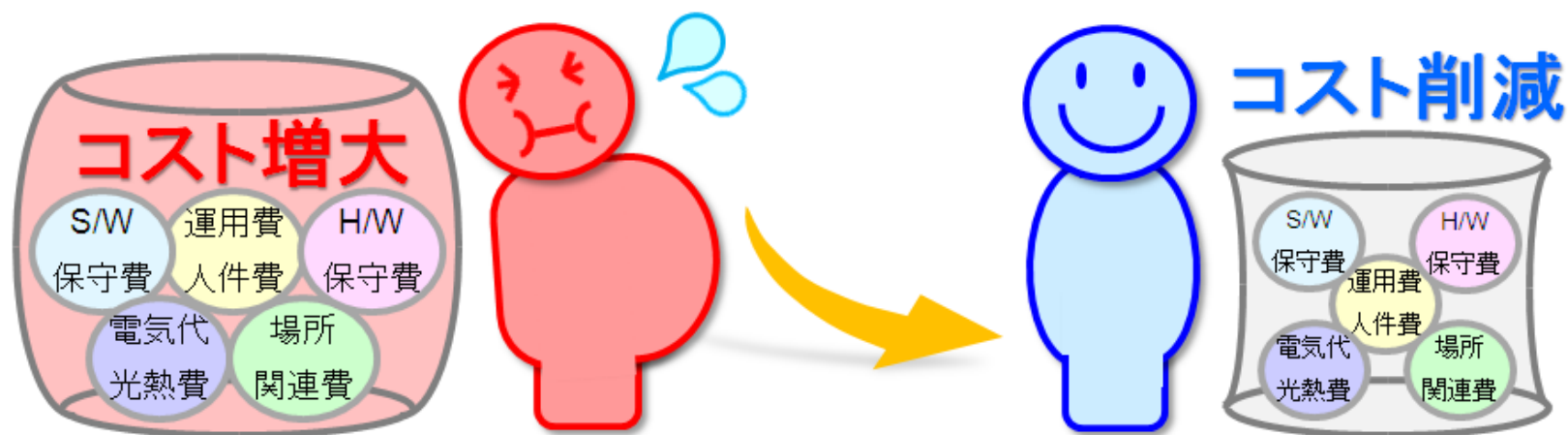


RAC

買ったし

RACによる最適な投資コストの実現

- 必要に応じてサーバ追加による性能拡張が可能のため導入時は最小限の構成でシステム構築が可能
- 全サーバーで処理を行うので待機用途のサーバは不要
- システム統合でリソースを共有化し、遊休リソースをなくしサーバを集約
 - ✓ 集約することで運用管理も簡易化されコストも削減可能



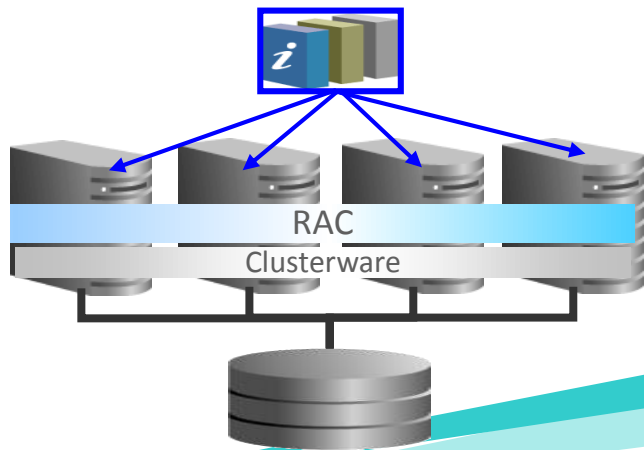
RAC/ASMの進化のステージ

クラウド・コンピューティングを支える統合システム基盤

Oracle 9i ~

単一サービス基盤

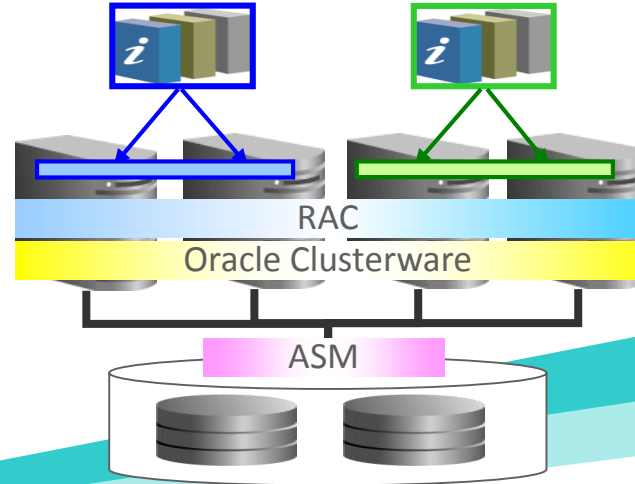
可用性と拡張性



Oracle 10g ~

複数サービス基盤
共用データベース

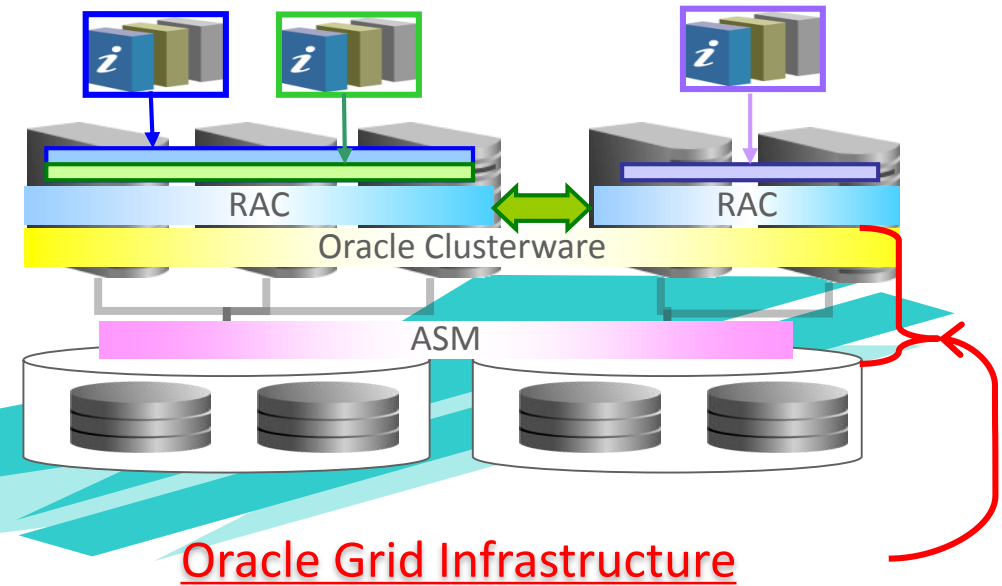
データベース・グリッド



Oracle 11g R2 ~

複数のグリッドを束ねた
共用インフラストラクチャ

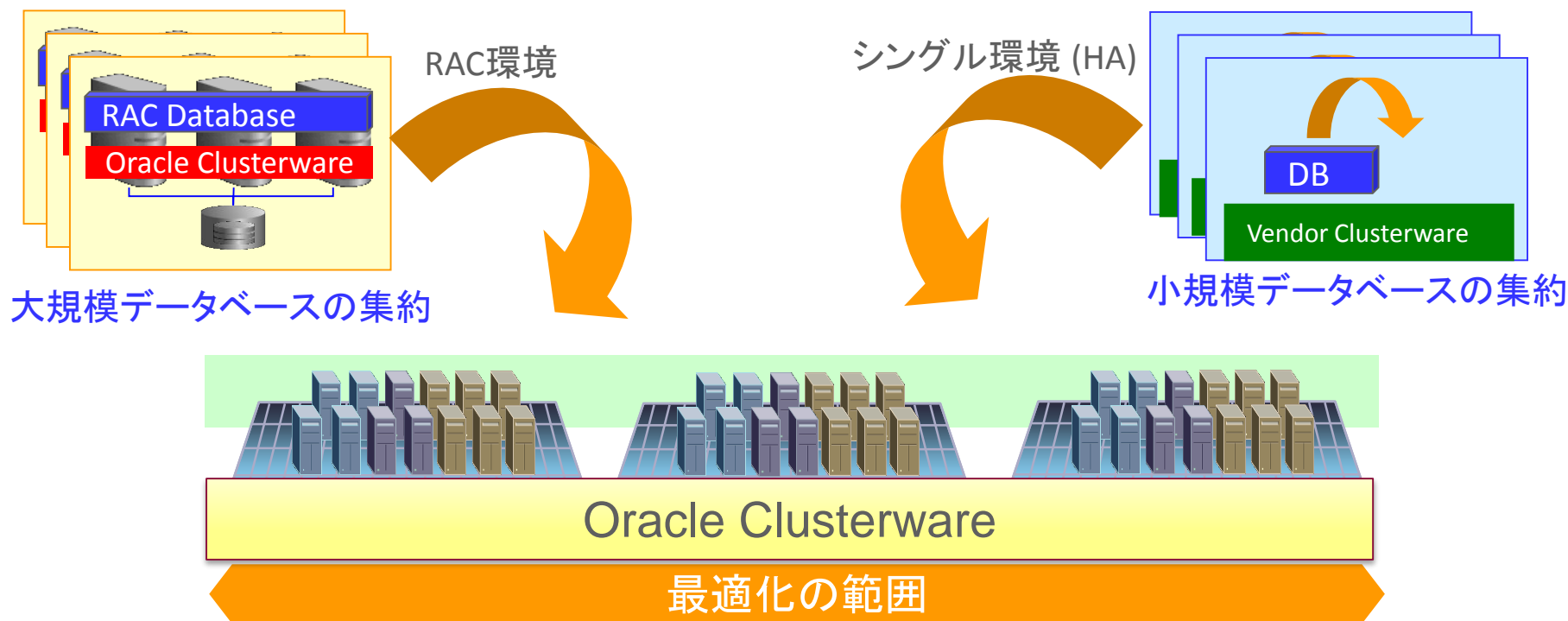
データセンター・グリッド



ポリシーベース管理

Oracle Real Application Clusters 11g Release 2～

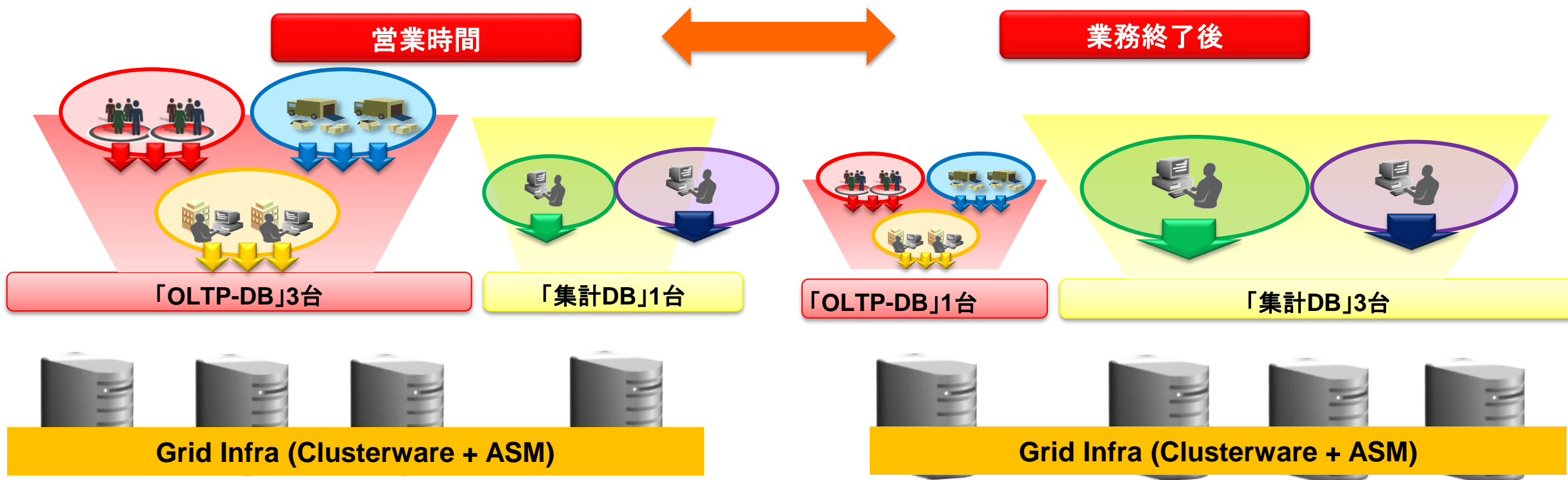
- より大規模なリソースを集約し、柔軟にリソースを配置できる動的なインフラを提供 ⇒ インフラ全体の最適化



ポリシーベース管理

柔軟なリソース制御が可能に

- 時間帯(負荷)に応じて、各データベースに割り当てるサーバー台数を変更し、サーバー・リソースを効率的に活用



Oracle Real Application Clustersの高可用性

サーバー障害時の継続稼働に必要な機能

1. 障害サーバーの切り離し

- クラスタ・ノード・メンバーシップ管理
- インスタンス・リカバリ

2. 自身の健全性を維持

- Oracle Clusterware 関連プロセスの監視
- RAC データベース 関連プロセスの監視

3. 正常サーバーへの接続

- サービス名を用いたRAC データベースへの接続
- 接続時のフェイルオーバーと仮想IPアドレス
- FANイベントとFCF

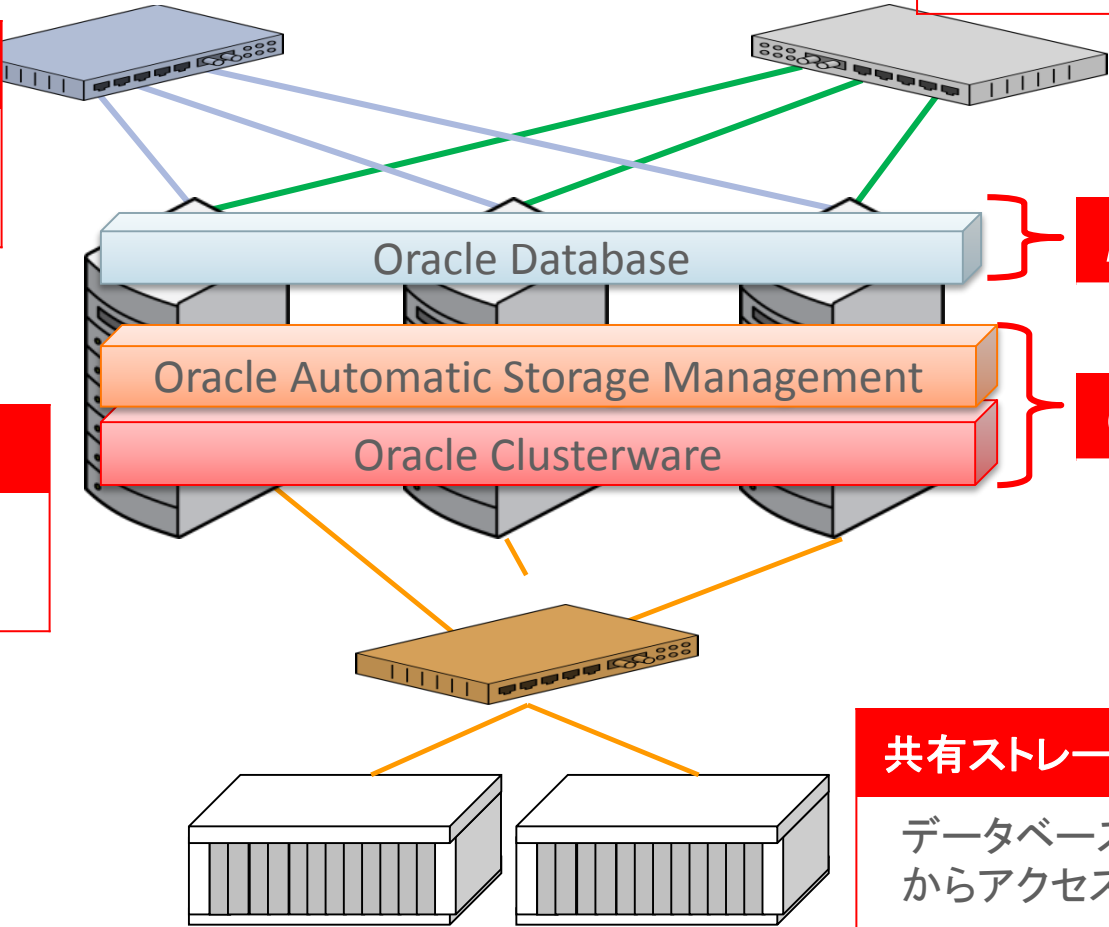
Oracle Real Application Clusters

必要なH/W & s/w構成図

パブリック・ネットワーク
クライアントとの通信を行うネットワーク

インターコネクト・ネットワーク
ノード間通信を行うネットワーク

ノード
RACを構成するデータベース・サーバー



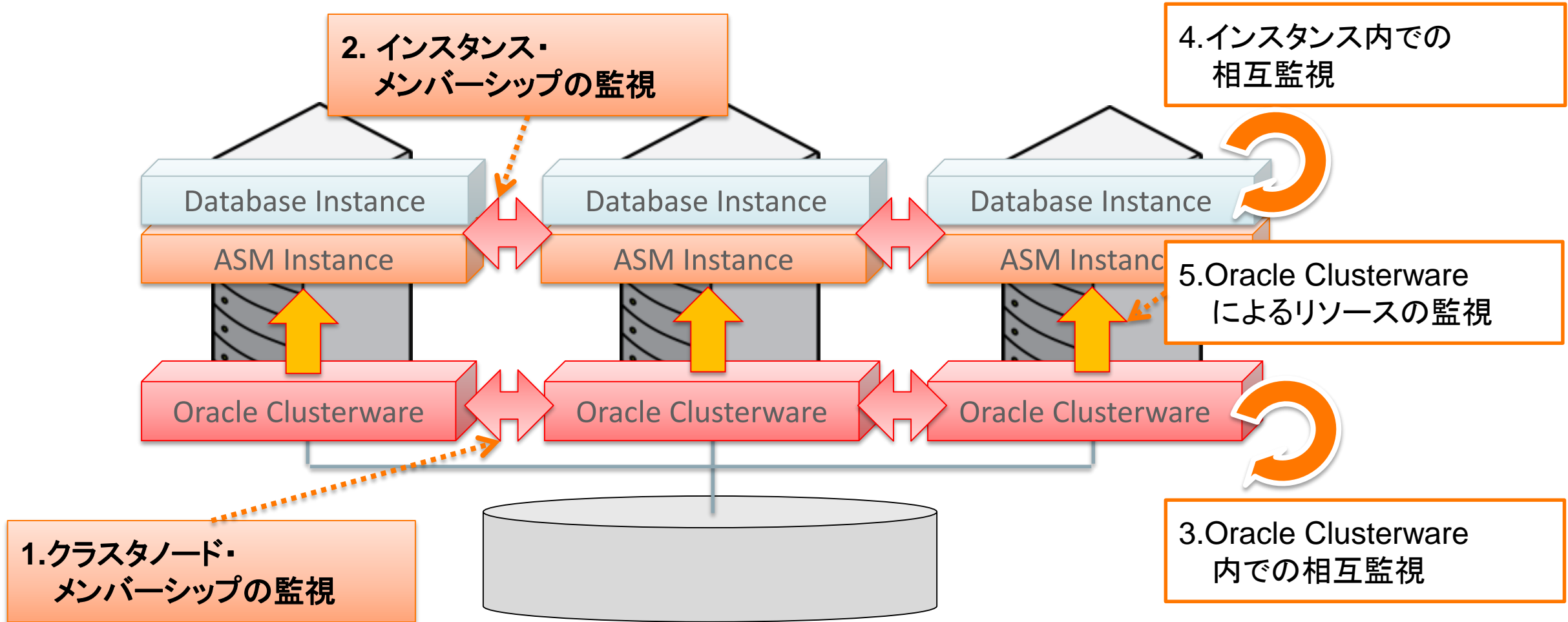
Database

Grid Infrastructure

共有ストレージ
データベース・ファイルを配置する全ノードからアクセス可能なストレージ

Oracle Real Application Clusters

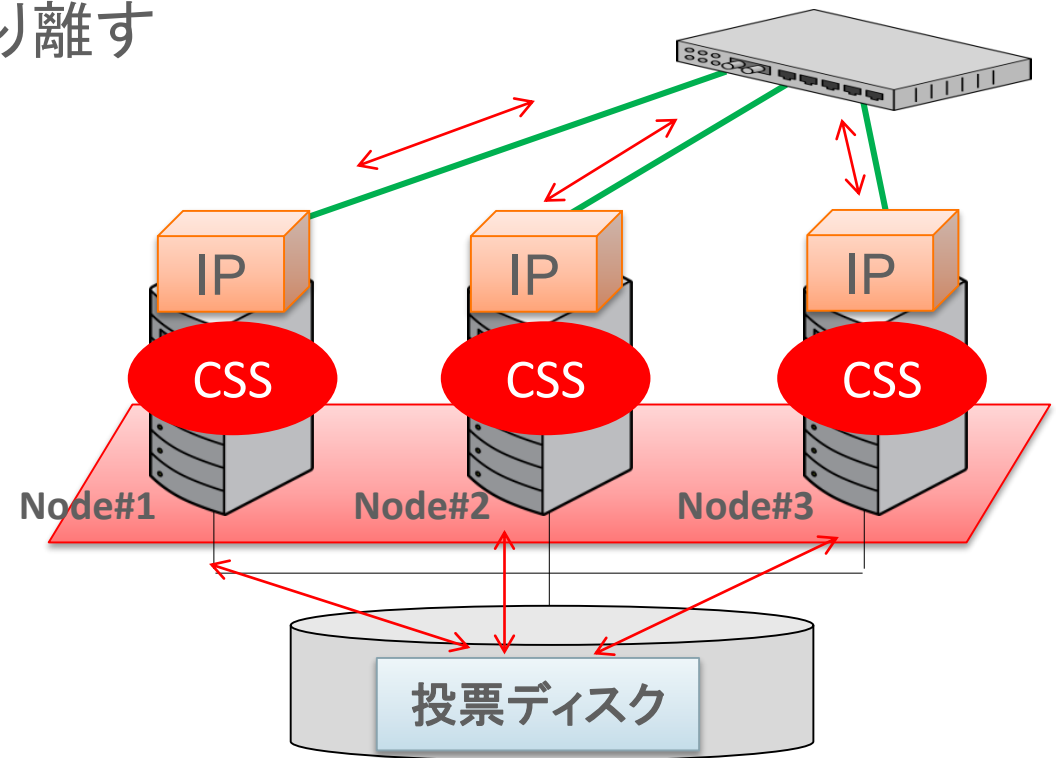
可用性を維持するために行っている監視



【参考】クラスタ・ノードのメンバーシップ管理

Cluster Synchronization Services (CSS) によるクラスタノードの異常検知

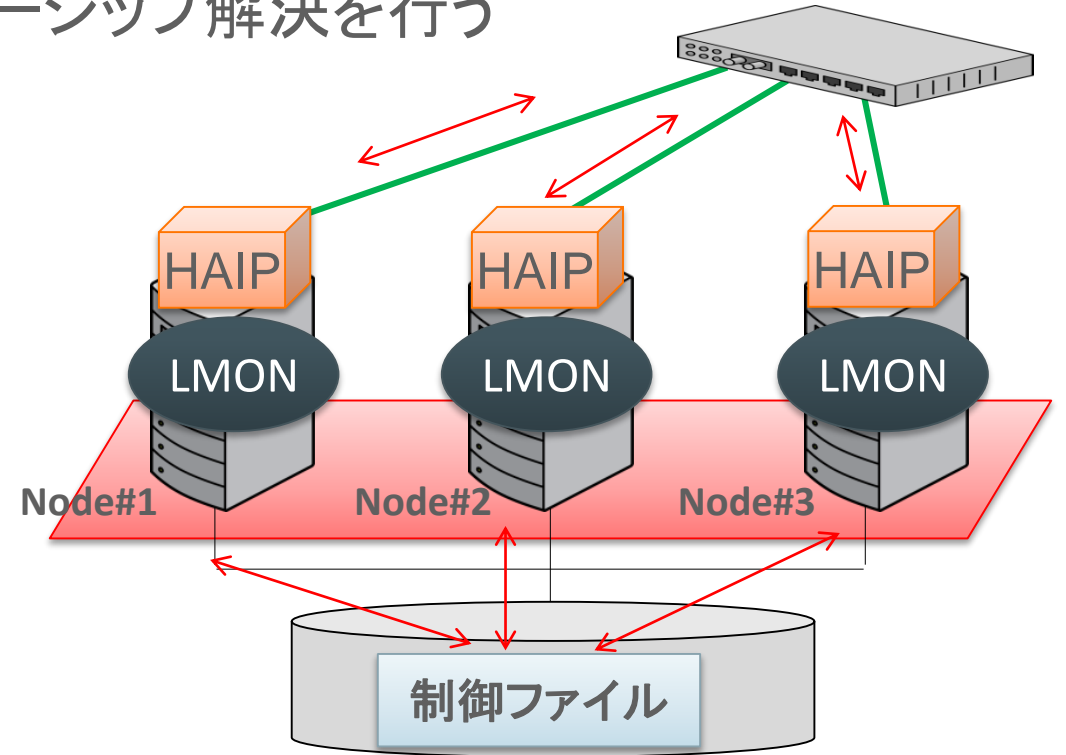
- 各ノードの CSS がクラスタ・ノード構成の一貫性に責任を持つ
 - 相互に生存監視を実施し、異常ノードを切り離す
- CSSは2種類のハートビートをしている
 - ネットワーク・ハートビート(1秒毎)
 - 死活監視のため
 - misscount (30sec) 間ハートビートがない場合はメンバーシップの変更
 - ディスク・ハートビート
 - スプリットブレイン解決時のため



【参考】インスタンス・メンバーシップ管理

LMONによるハートビート

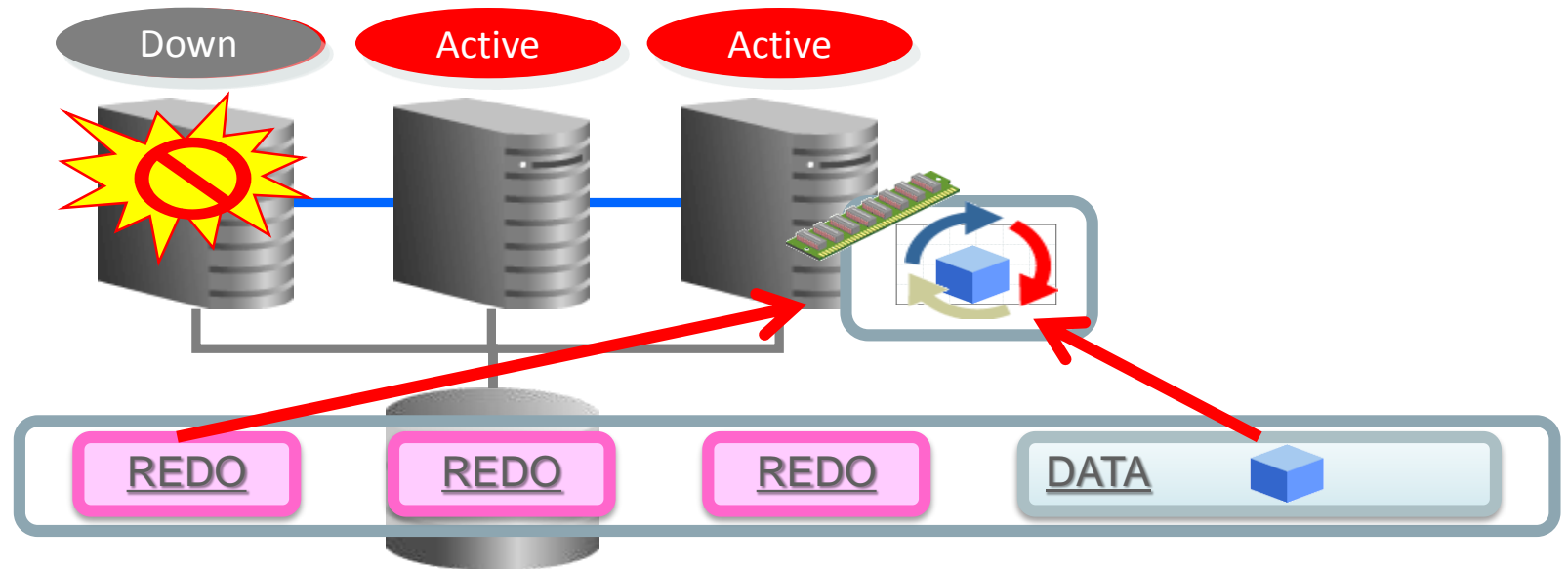
- 各ノードの LMON がメンバーシップを監視
 - インスタンスや回線障害の検知時にはメンバーシップ解決を行う
- 2種類のハートビートを行っている
 - ネットワーク・ハートビート
 - cluster_interconnect で指定されたネットワーク(11.2.0.2からは通常HAIP)
 - 一定時間ハートビートがない場合はタイムアウトし、メンバーシップの解決が行われる
 - ディスク・ハートビート(制御ファイル)



インスタンス・ダウン時のデータ・ブロックの自動リカバリ

Oracle Real Application Clusters

- 障害が発生したインスタンスのキャッシュ上にしか存在しなかったデータ・ブロックは、生存する他インスタンスが自動的にリカバリを実施
 - 正常インスタンスが障害インスタンスのREDOログを読み込み自動リカバリ
 - 障害インスタンスの再起動を待つ必要はない



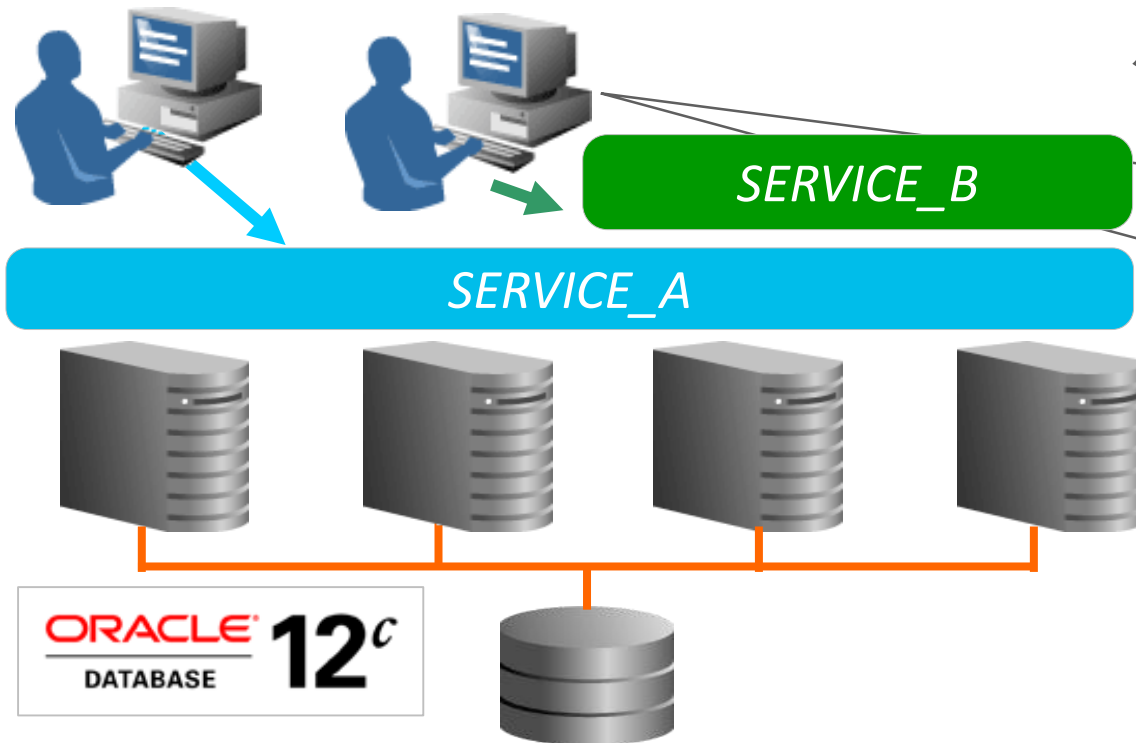
クラスタ・データベースへの接続に求められること

アプリケーション・レイヤー視点

- クラスタとしての整合性を取ったり、ノード内の健全性を維持する動作をしていても、アプリケーションから正常なサーバーに接続できなければシステムとして動作しない
- 接続に求められること
 1. インスタンスを意識せずに接続ができること
 2. 接続リクエストが正常なノードに割り振られる
 3. 障害ノードに対する無効な接続は破棄されること

サービスを使用してクライアントから接続

- Oracle クライアントは「**サービス**」に対して接続します
 - Oracle クライアントは RAC の **ノード数**や **インスタンス名**は知りません



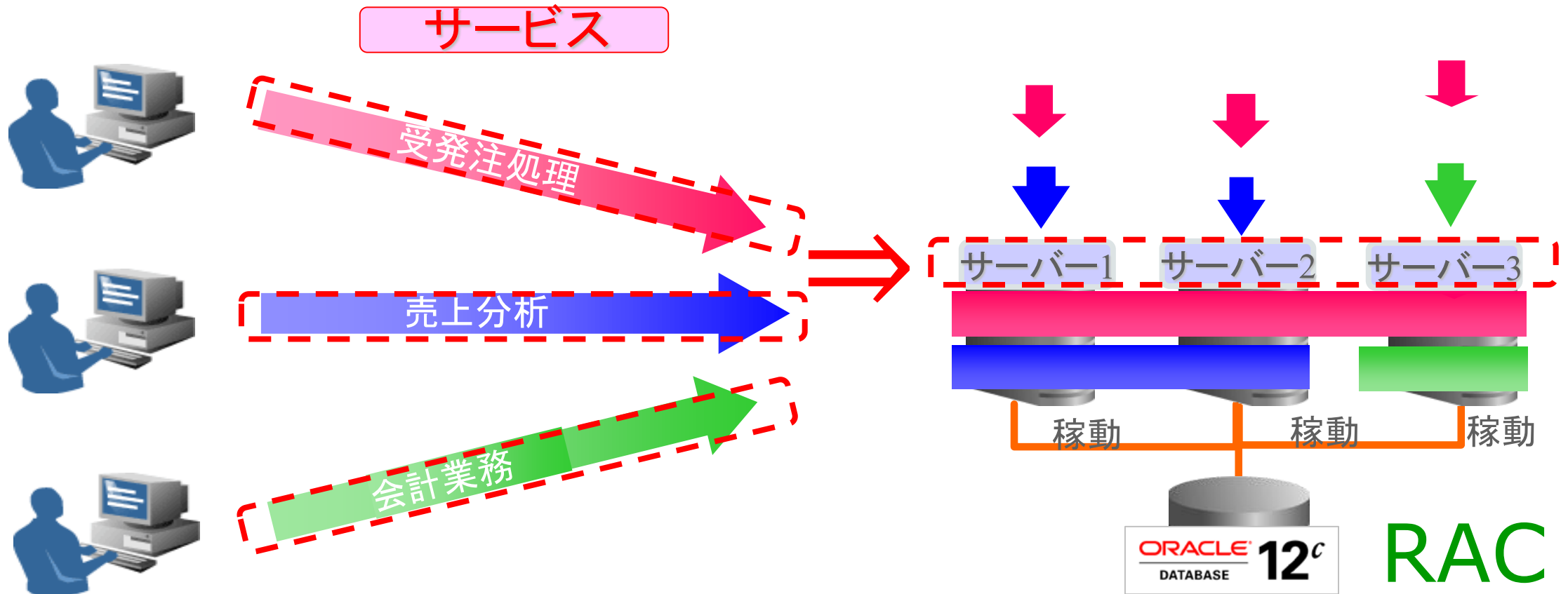
クライアントは「SERVICE_NAME」を指定して接続する

```
(DESCRIPTION =  
(ADDRESS=(PROTOCOL=TCP)  
(HOST=scanname) (PORT = port ) )  
(CONNECT_DATA=  
(SERVICE_NAME=SERVICE_A ) ) )
```

接続するサービス名と
SCANリスナーが稼動している
アドレス情報 (スキャン名、port 番号等) を記述

サービスを利用したワークロード管理

- ユーザーが意識しなくても各サーバへのワークロード(負荷)を均等分散



Oracle Real Application Clustersの高可用性

サーバー障害時の継続稼働に必要な機能

1. 障害サーバーの切り離し

- クラスタ・ノード・メンバーシップ管理
- インスタンス・リカバリ

2. 自身の健全性を維持

- Oracle Clusterware 関連プロセスの監視
- RAC データベース 関連プロセスの監視

3. 正常サーバーへの接続

- サービス名を用いたRAC データベースへの接続
- 接続時のフェイルオーバーと仮想IPアドレス
- FANイベントとFCF

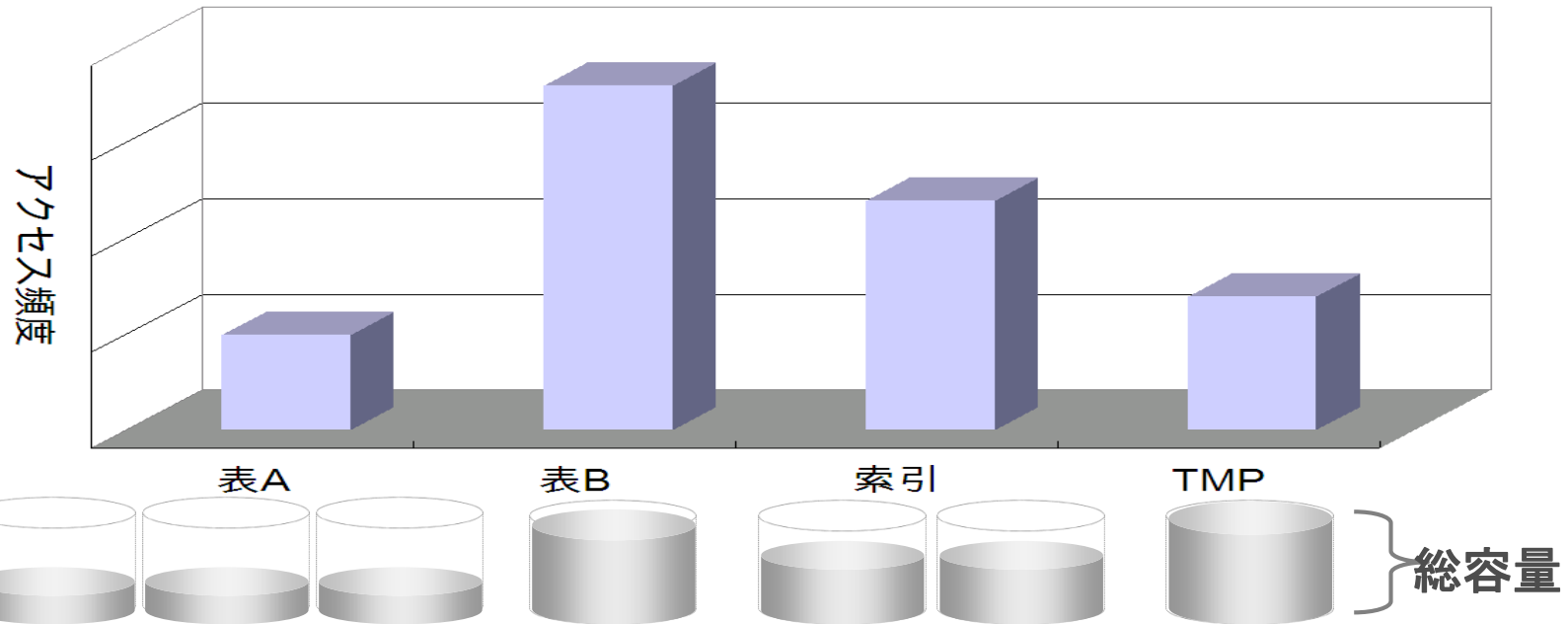
A woman with long brown hair and glasses is sitting at a wooden table in a cafe or office setting. She is wearing a brown leather jacket over a blue patterned scarf. She is holding a black mobile phone to her ear with her left hand and looking down at a newspaper or magazine on the table with her right hand. The background is slightly blurred, showing other people and tables.

Oracle Automatic Storage Management

データベース・ストレージ管理の課題

複雑なディスク管理

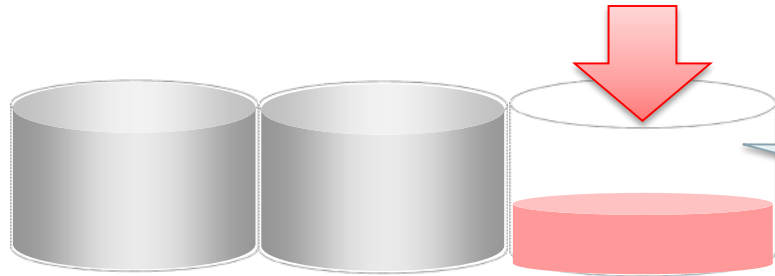
- 業務の複雑化により、従来の個別最適化を目指す運用は困難な傾向
– 必要なディスク性能を事前見積もり？ 偏りの最適化？



データベース・ストレージ管理の課題

ディスク追加時の課題

- 領域不足/性能劣化の改善の為、ディスク追加

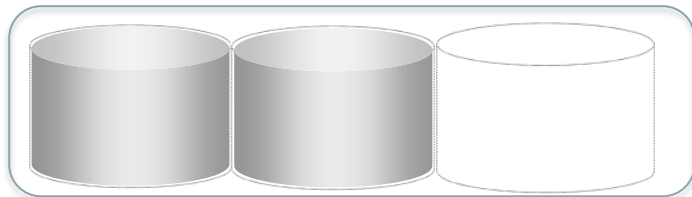


ホットスポットが発生

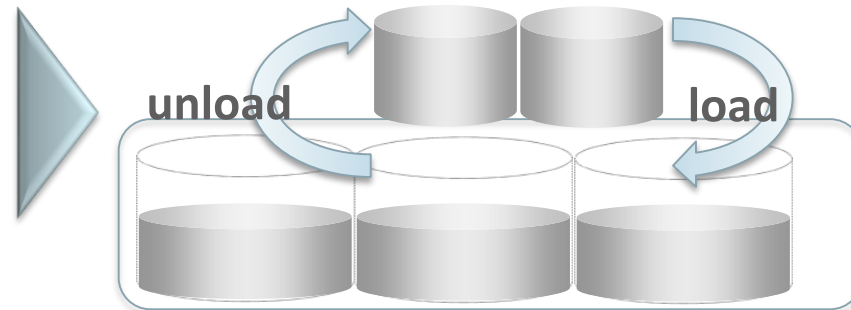
ディスクは3本あるが、最新データが入っている1本のみアクセスが集中

- ホットスポット回避のためには既存データの再配置が必要

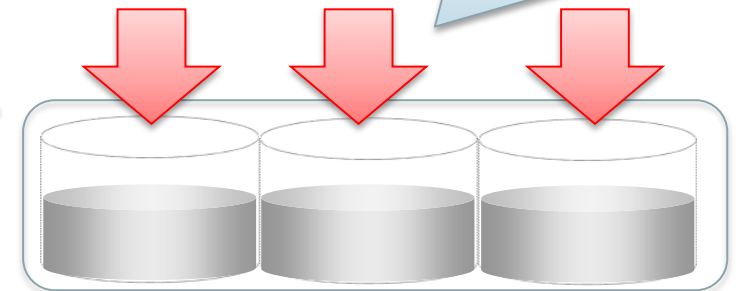
ディスクを追加し
RAIDグループを再構成



データ再配置



全ディスクに対してI/O



Oracle によるストレージ仮想化

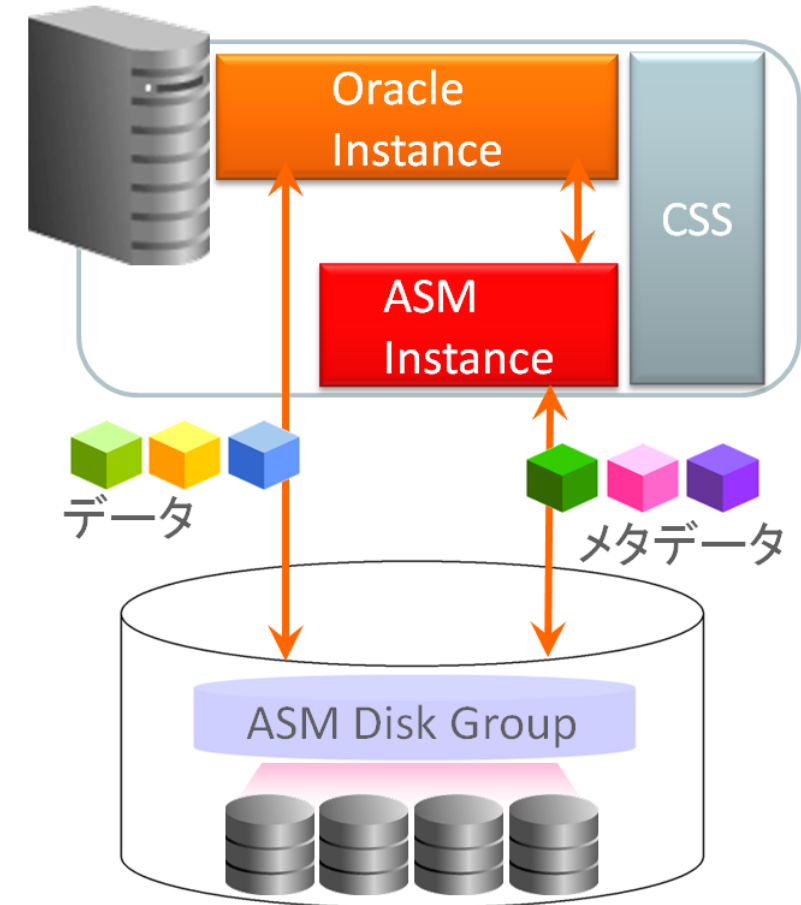
Oracle Automatic Storage Management (ASM)

- Oracle 10gより提供されているディスク構成の仮想化技術
 - Oracleデータベースに対してボリューム・マネージャ兼ファイルシステム
 - Oracle Databaseにフラットなディスク・プールを提供 + ディスク管理工数を大幅削減
 - 複数ディスク・アレイにまたがってディスクを仮想化、ディスク追加/削除時にデータを透過的に再配分
 - エディション(EE/SE)に関係なく、シングル環境、クラスタ環境共に使用可
 - 11g Release2より、ASMクラスタファイルシステム(ACFS)が実装
- **11g 以降では約9割以上がASMを採用**

Oracle Automatic Storage Management

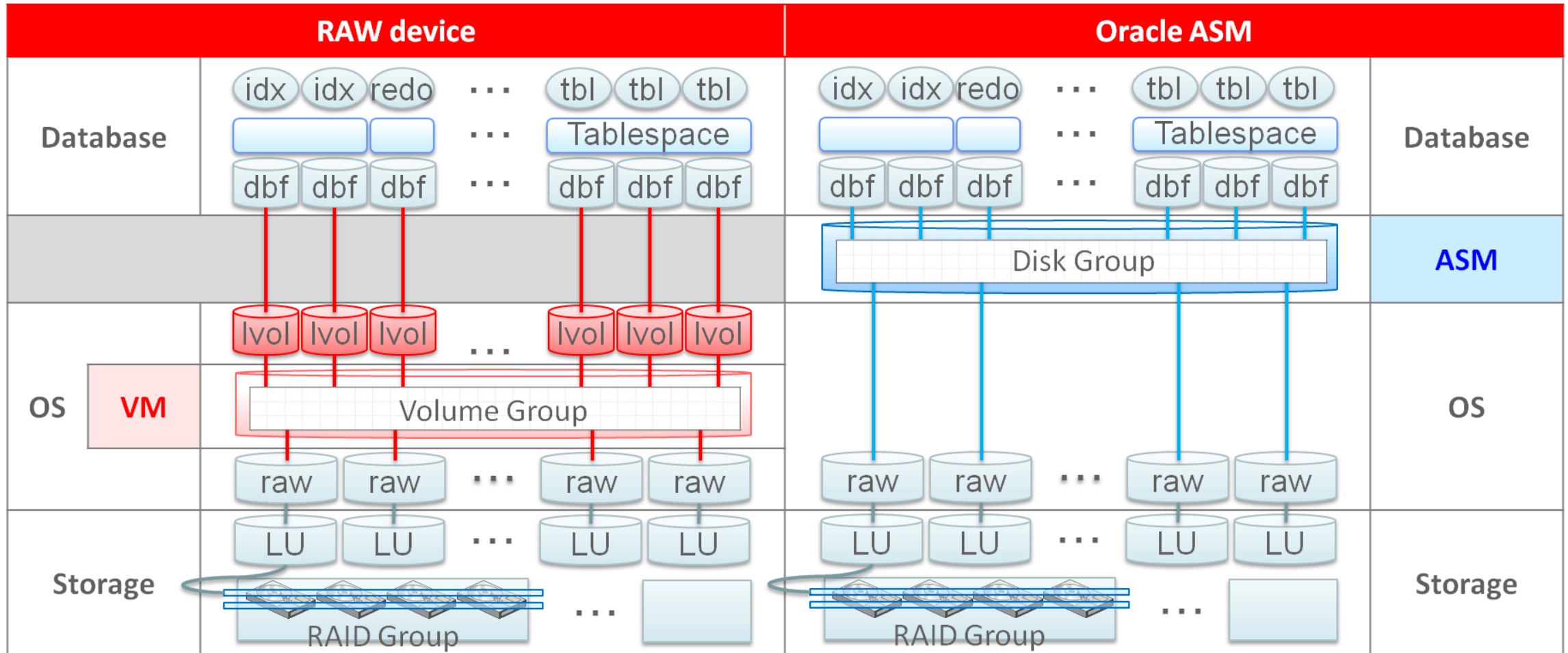
アーキテクチャ

- ASMインスタンス
 - ASM Diskgroupを管理するメモリとプロセス群
- Cluster Synchronization Services
 - Oracle Clusterwareのメンバシップ管理サービスを利用
 - DBインスタンスとASMインスタンスの存在通知する
- ASM Diskgroup
 - Oracleインスタンスに対する仮想化ストレージプール
- ASM Disk
 - ASM Diskgroupを構成する個々のDisk(Logical Unit)



【参考】RAW deviceとOracle ASMの比較

スタック構成図



ASMによる全体最適化機能

- **ストライピング**

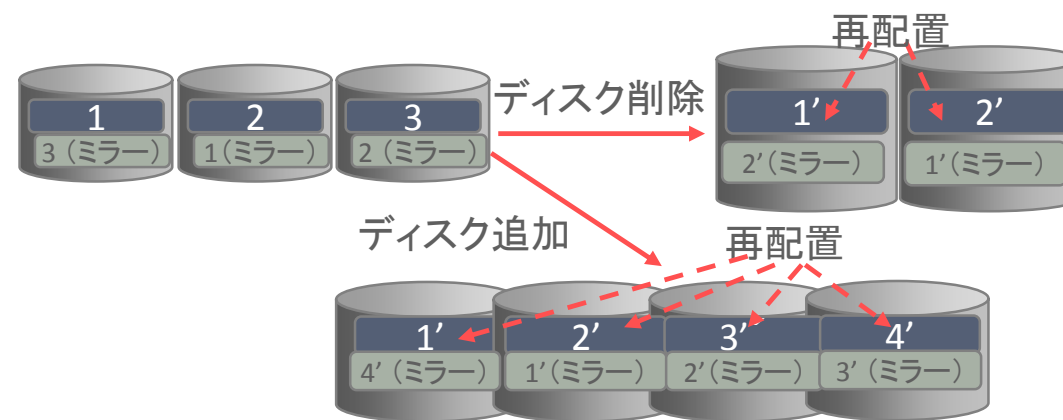
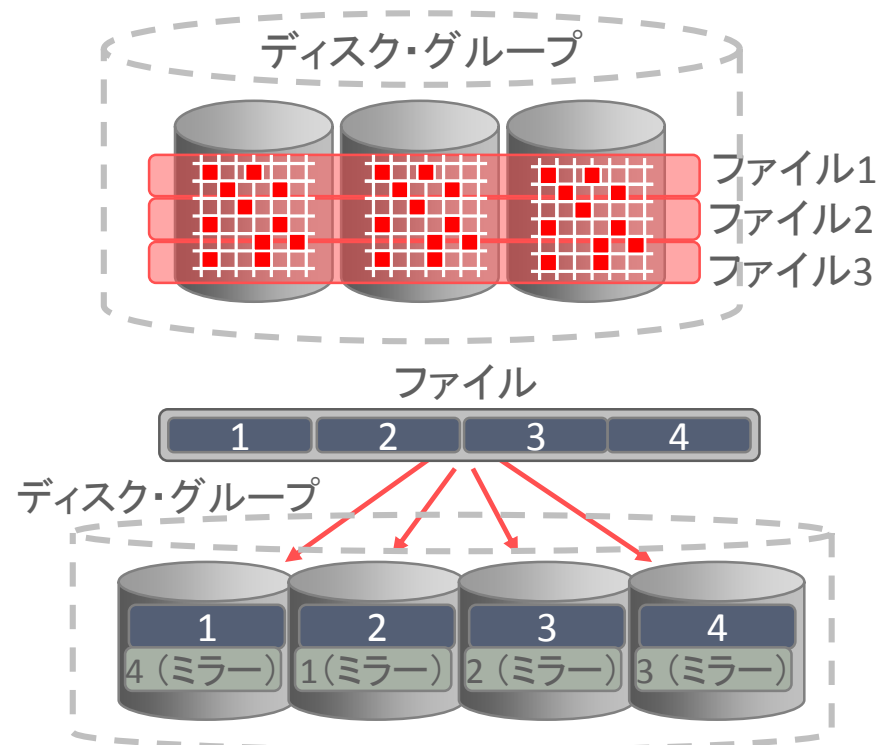
- ディスク・グループ内の、全てのディスクでストライピング(ホットスポットが発生しない)
- 性能の維持

- **ミラーリング**

- ファイルタイプに応じて、Oracleレベルでミラーリング(2重化/3重化/ミラー無し)
- 可用性の担保

- **動的リバランシング**

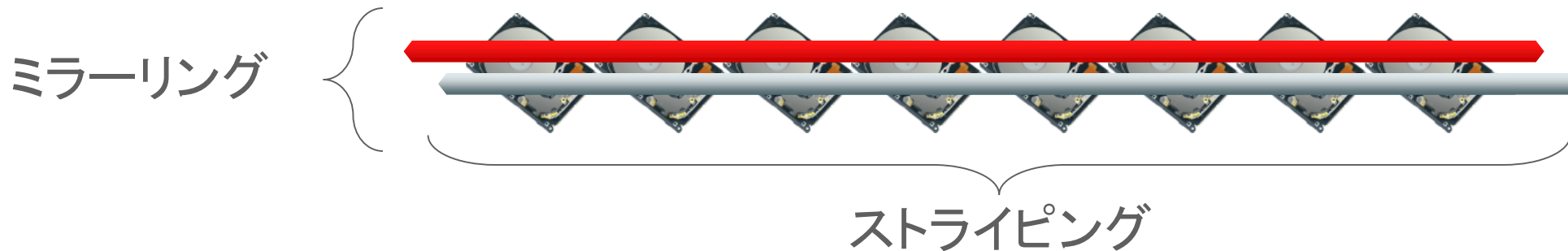
- ディスクの追加/削除時に自動的にデータを再配置
- 拡張性



Oracle ASMの基本思想

Stripe And Mirror Everything (S. A. M. E)

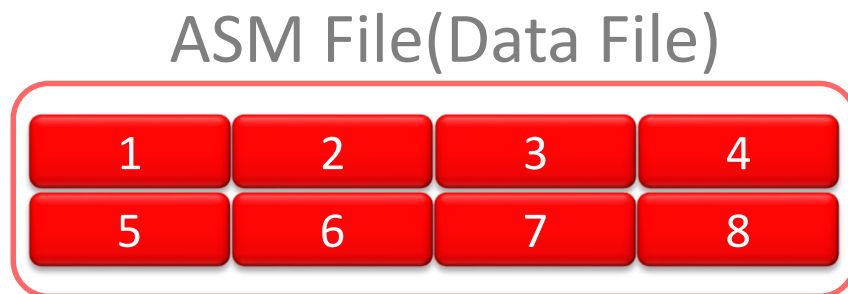
- 「全てのDiskの均等利用を目的に、データをストライプして全てのDisk上に分散配置し、ミラーリングも行う」という設計手法
 - I/O性能の確保: 全てのDiskのI/O帯域をフル活用
 - 可用性を確保: ミラーリングの採用
 - 設計の簡素化: 物理的なDisk構成を隠蔽し、特別な設計は不要



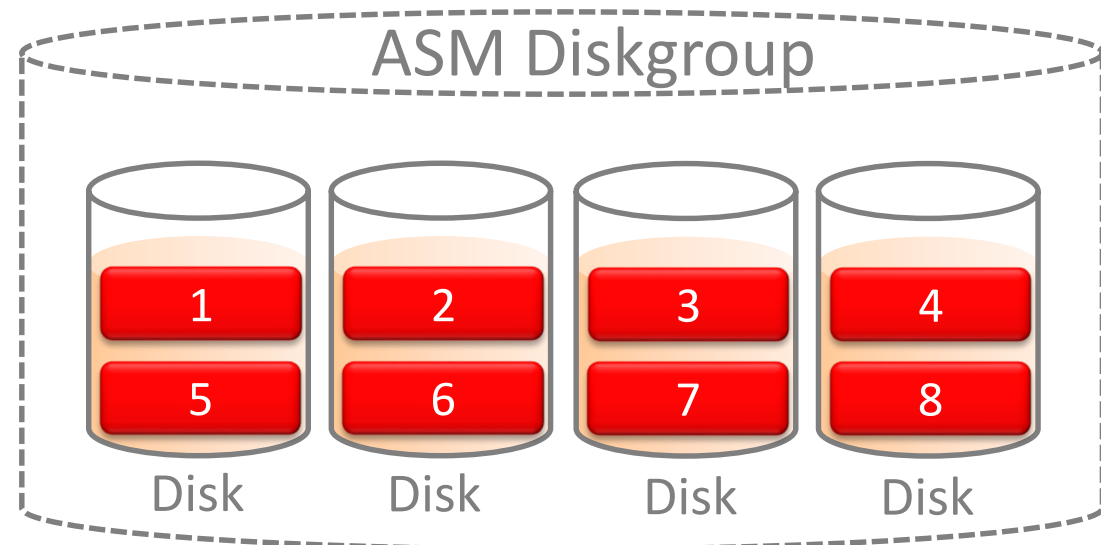
Oracle ASMによるストライピング

ASM File(データファイル)の分散配置例

- ASM Diskgroupに含まれる全てのASM Diskに対して、ASM File(Data File)をFile Extent(Allocation Unit:=AU)単位に分割して配置



 File Extent (AU)



Oracle ASMによるミラーリング

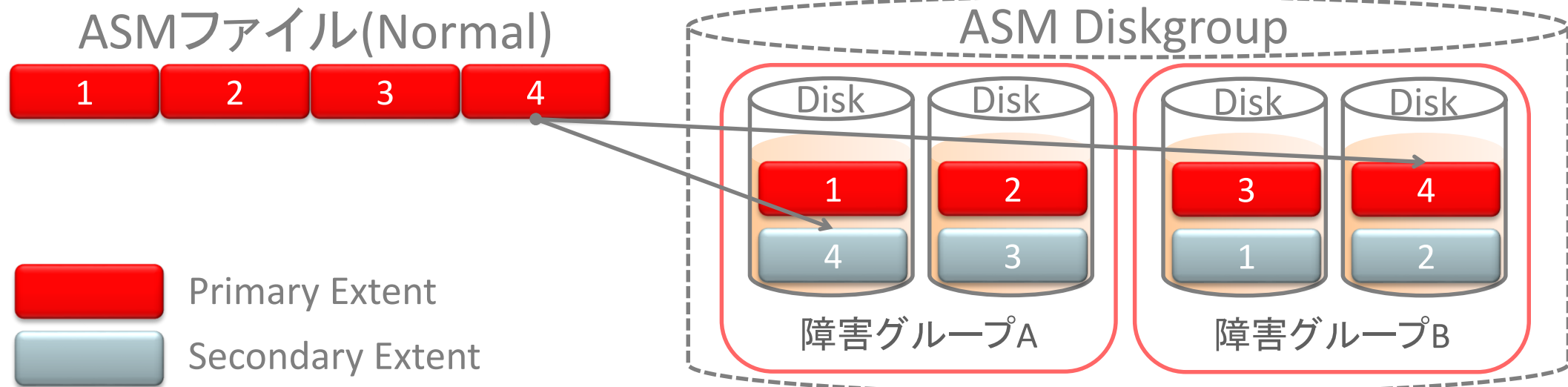
Oracle Clientに透過的、かつ自動的にBlockを修復

- ASM Diskgroupに、ミラーリング (External/Normal/High) の設定が可能
 - Normal / High Redundancy時
 - 読み取り処理時にI/Oエラーを検知した場合
 - セカンダリから読み取り、**不良ブロックは自動修復**
 - Oracle Clientに対して透過的 (ORAエラーは戻らない)
 - 書き込み処理時にI/Oエラーを検知した場合
 - 障害Diskを**自動でオフライン化**
 - Oracle Clientに対して透過的 (ORAエラーは戻らない)
 - 障害Diskの復旧時、高速ミラー再同期により生存Disk側から必要最小限の差分データを同期
 - 復旧できない場合、ASM Diskgroupから切り離し (自動リバランスが発生)

Oracle ASMによるミラーリング

Normal Redundancy時のミラーリングと障害グループ例

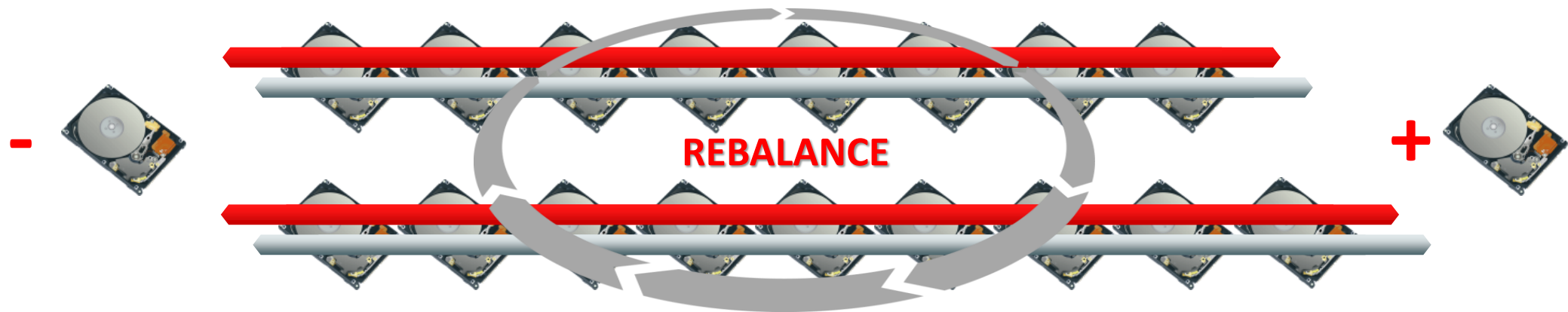
- 異なる障害グループに属するASM Disk間で保持
- 通常、リソース(電源等)を共有している単位(筐体/コントローラー)で設定



Oracle ASMのリバランス(データ再配置)

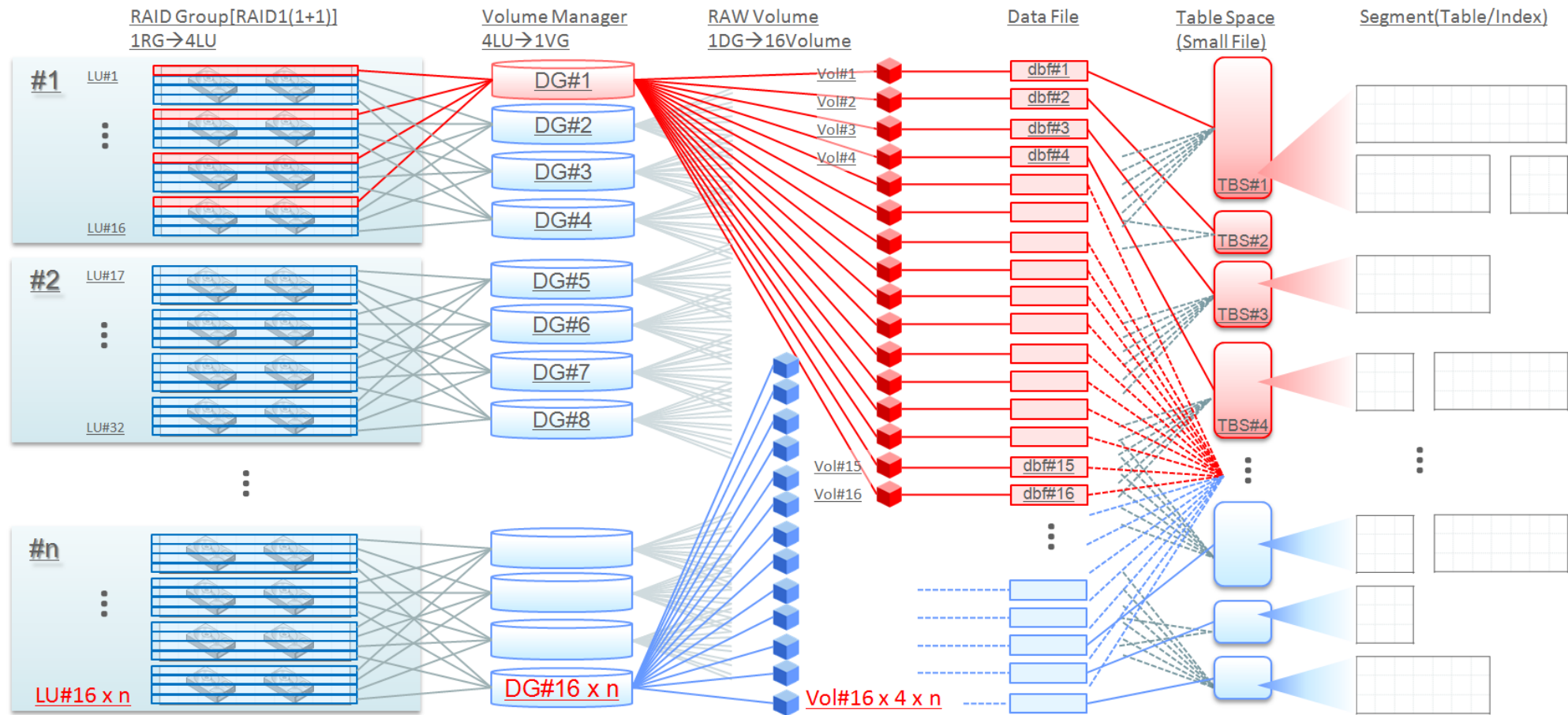
データベース無停止でリバランスが可能

- ASM Disを追加／削除(故障)した際、データの再配置を実施
 - メタデータ(配置状況)を元に、**ASM File単位**で全てのDiskに均等配置されるように**最小限のExtent(AU)の移動**で実現
 - 多重度(リバランス強度)の設定や計画実行で、業務影響を制御可能



従来型のRAW device構成例

目的はストライピングによるI/O性能の向上



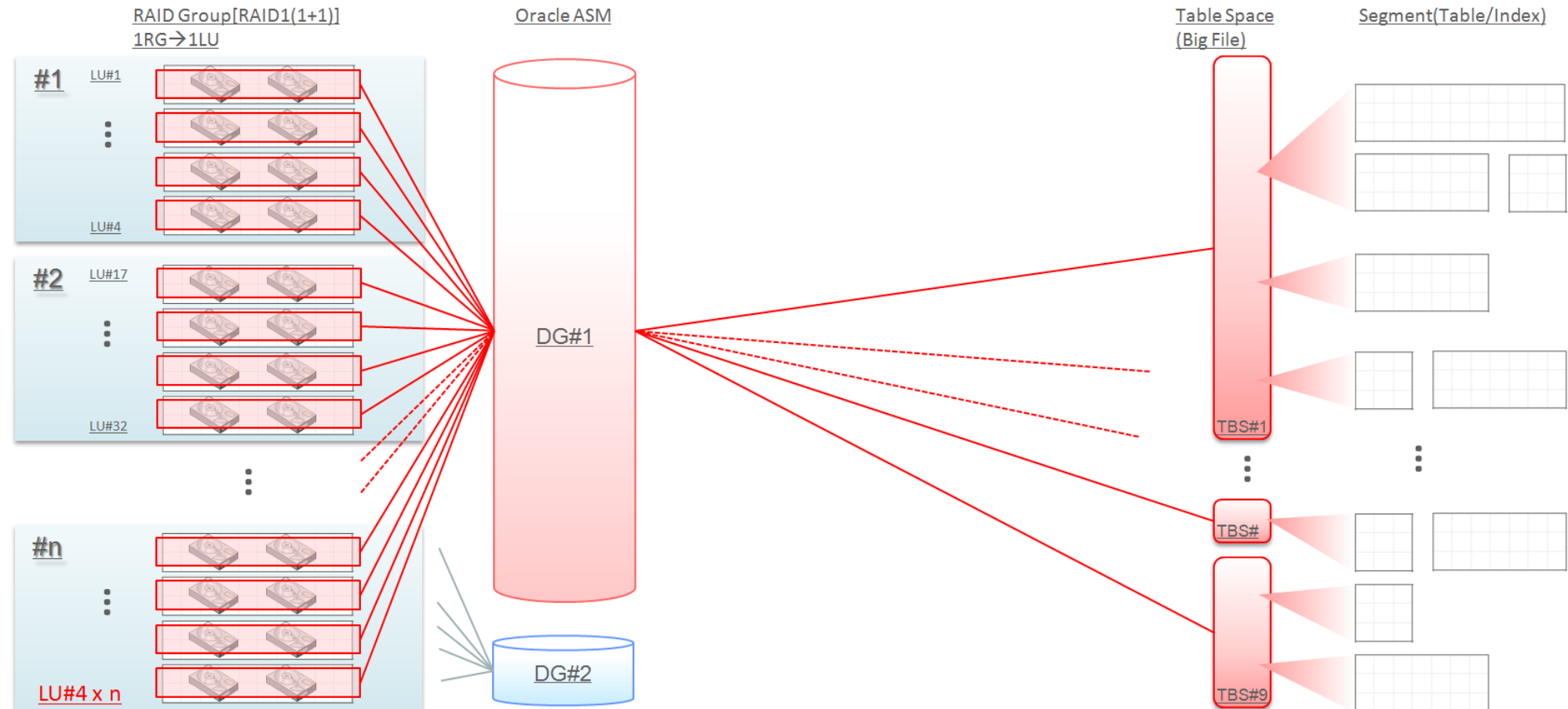
従来のRAWデバイス構成の課題

運用の複雑化

- 表領域が非常に細かく分割されている
 - 空き領域が表領域毎に独立している為、無駄な空き領域が増大
 - 監視対象(表領域)が多く、頻繁に領域不足に陥り、運用工数が増大
 - データ・ファイル数が多く、SQLの性能劣化やミス・オペレーションを誘発
 - 管理レイヤー数が多い為、運用オペレーションの複雑化
- データ・ファイル追加時に、既存データをリバランスしていない
 - 空き領域が新規ボリュームにのみ存在する為、新たにINSERTされるレコードがそのボリュームに集中することで、ボトルネックが発生し易い
 - 既存レコードは既存ボリューム内に格納されている為、性能改善効果は無し

Oracle ASMの構成例

Simple is the BEST



Oracle ASMによる運用管理の簡素化

従来構成の課題を解決

- オペレーションの簡素化
 - 表領域拡張やDisk追加の手順が簡素化し、運用オペミスのリスクが減少
- 管理対象オブジェクトの削減
 - ASM Diskgroupの容量内で表領域を自由に拡張可能であり、従来のVolumeやRAWデバイス(データファイル)を意識する必要なし
 - ストライピングでI/Oが均等化することで、表領域を細かく分割してI/O競合を回避する必要なし。表領域の総数を大幅に削減可能
- データ再配置の工数不要
 - Disk追加時に自動的に既存データの再配置(リバランシング)を実施

ASMによるデータベースの物理設計の簡易化

DBA のストレージ管理の効率化





Oracle Data Guard

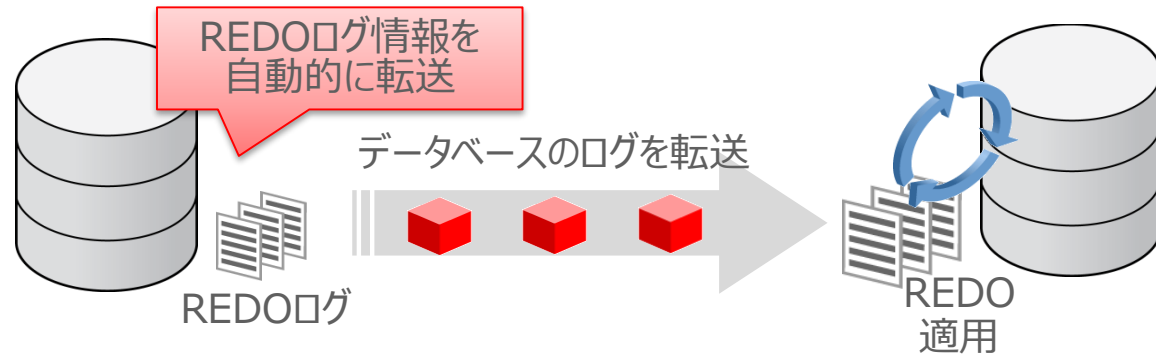
Oracle Data Guard

リアルタイム・データベース複製、災害対策

■仕組み

プライマリ・データベース

スタンバイ・データベース



■転送モード

	同期転送 (SYNC)	非同期転送 (ASYNC)
データ保護	プライマリ DBでの更新はスタンバイ DBへの転送完了後に確定	プライマリ DBでの更新はスタンバイ DBへの転送未完了でも確定
性能への影響	スタンバイ DBへの転送時間に依存してプライマリ DBの更新処理が待機	プライマリ DBへの更新処理はスタンバイ DBへの転送を待機しない

特徴：

- ① データ誤差無し
- ② 高速なデータ同期、ネットワーク帯域小
- ③ トランザクションの順次性保障

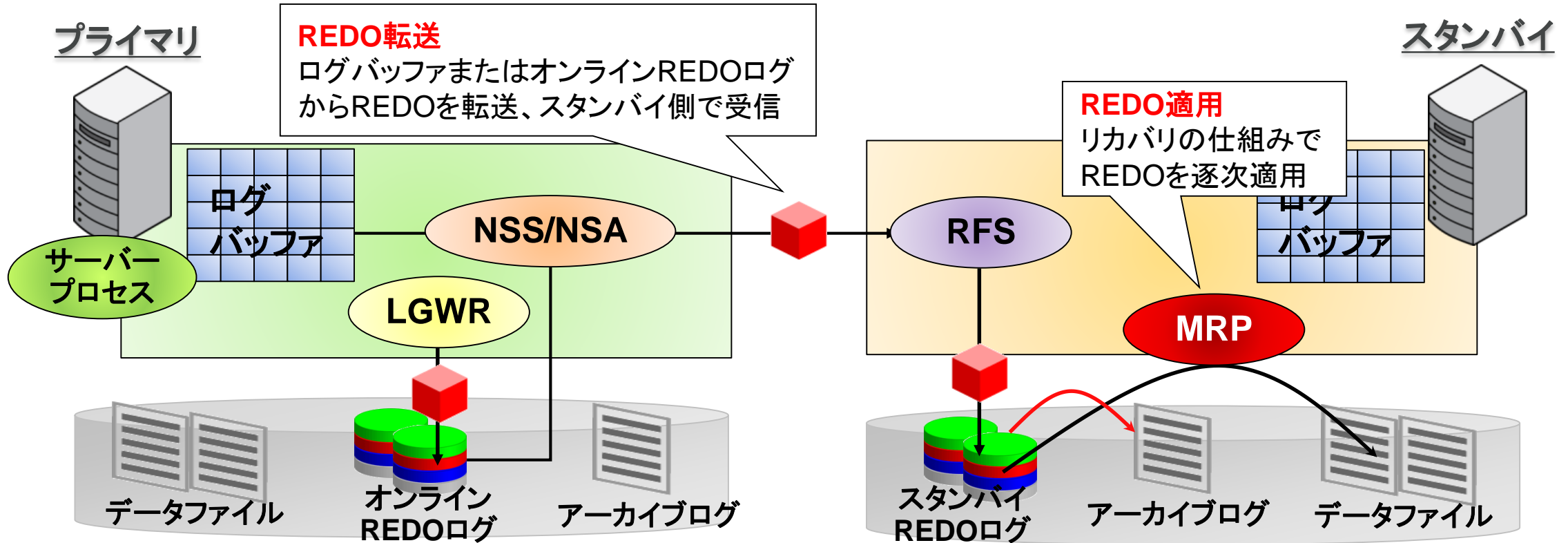
用途：

- 本番データベースのコピーを作成し、データを保護
- 災害対策/データ保護、移行/アップグレード
- (ADG) 検索、バックアップ、データ破損対策

※ADG : Active Data Guard Option

Data Guard

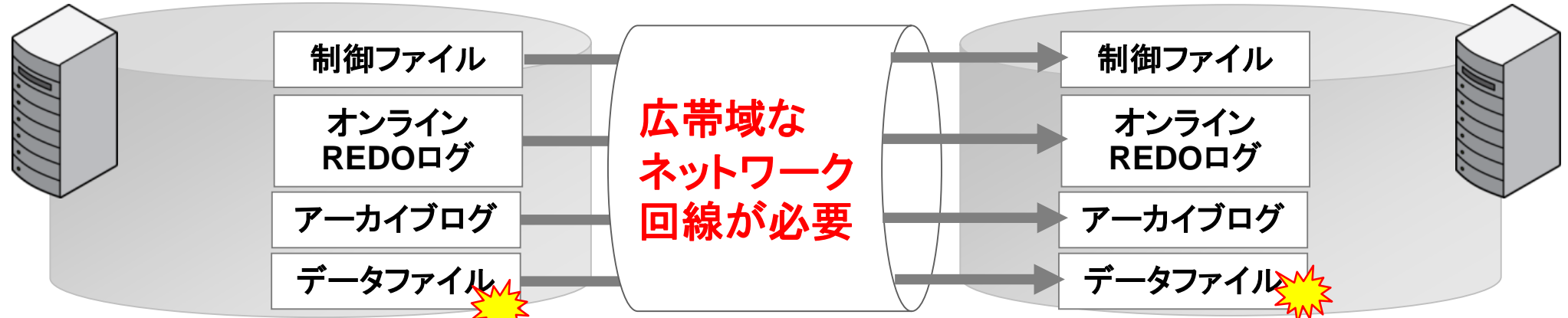
アーキテクチャ



データファイルはデータ・ブロックレベルで等しいが、
データファイルをコピーしているわけではない

ストレージのリモートミラーと Data Guard

ストレージの
リモートミラー



プライマリのデータ破損がそのままスタンバイに反映される

Data Guard



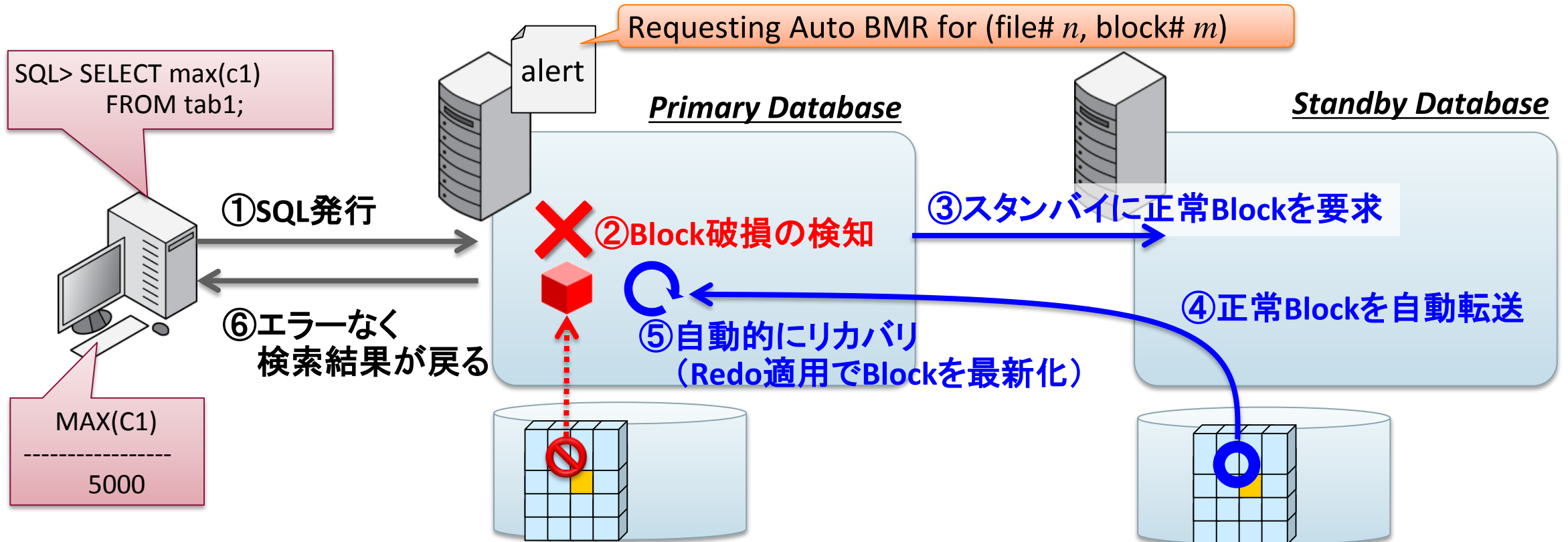
プライマリのデータ破損はスタンバイに反映されない

(参考)ストレージミラーとの比較

比較項目		Oracle Active Data Guard		ストレージミラー	
オンライン処理時の挙動	転送方式	○	DBトランザクション単位	○	ストレージ変更ブロック単位
	転送モード	○	同期 / 非同期	○	同期 / 非同期
	転送対象	○	REDOログのみ	×	REDOログ, datafileなど全てのDB関連ファイルが対象
	転送される処理	○	DML / DDLのみ	×	SELECT処理を含むほぼ全てのSQL
	転送確認	○	<ul style="list-style-type: none"> Enterprise Managerで確認可能 転送データが不正の場合、転送時に確認可能 	×	スタンバイ側が正常に起動するかは、起動してみるまでわからない
災害時の挙動	切り替え時間	○	短時間で起動可能	×	Disk MountからDB起動/リカバリ処理など長時間の作業が発生
	切り替え処理内容	○	SQLのみで切り替え可能	×	OSコマンドからDB処理まで様々な処理が必要
	切り替え手順	○	Platinum / Gold levelで同一手順で切り替え	×	Platinum / Gold levelで違う切り替え手順を確立する必要あり
構築/保守運用コスト	スタンバイ側の利用	○	検索用など様々な利用用途	×	同期を一時停止しなければいけない
	H/Wの制限	○	特になし	×	同一機種のDisk / Storage Mirror機能が必要
	ライセンスコスト	○	DB EEで標準サポート / Active Data Guardオプション(11g)のみ追加コスト	×	DB以外にStorage Mirror利用ライセンスが必要

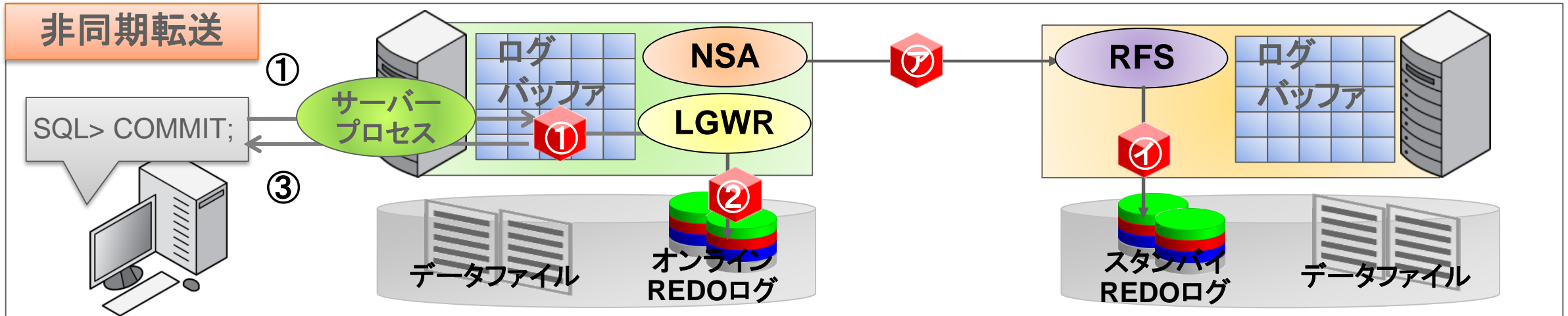
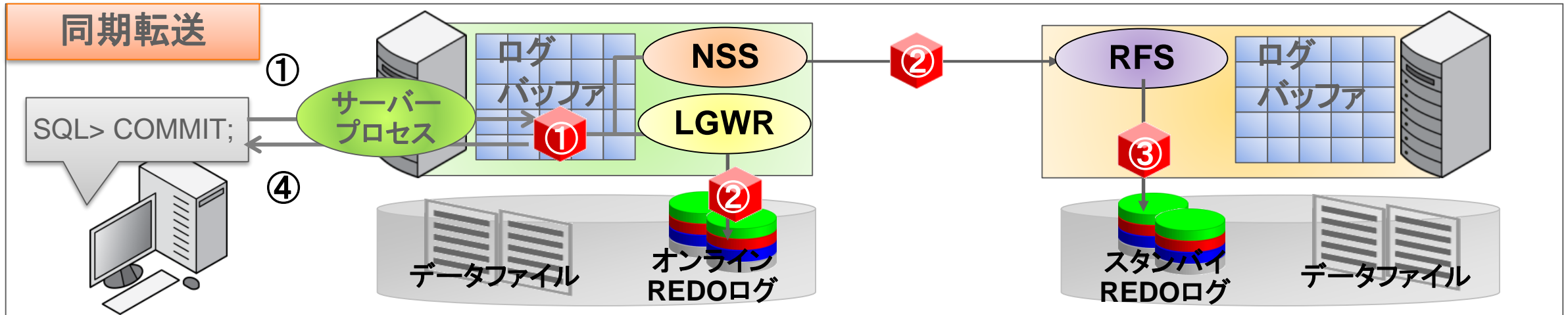
Automatic Block Media Recovery

Active Data Guardによる透過的なBlock修復（逆向きも有効）



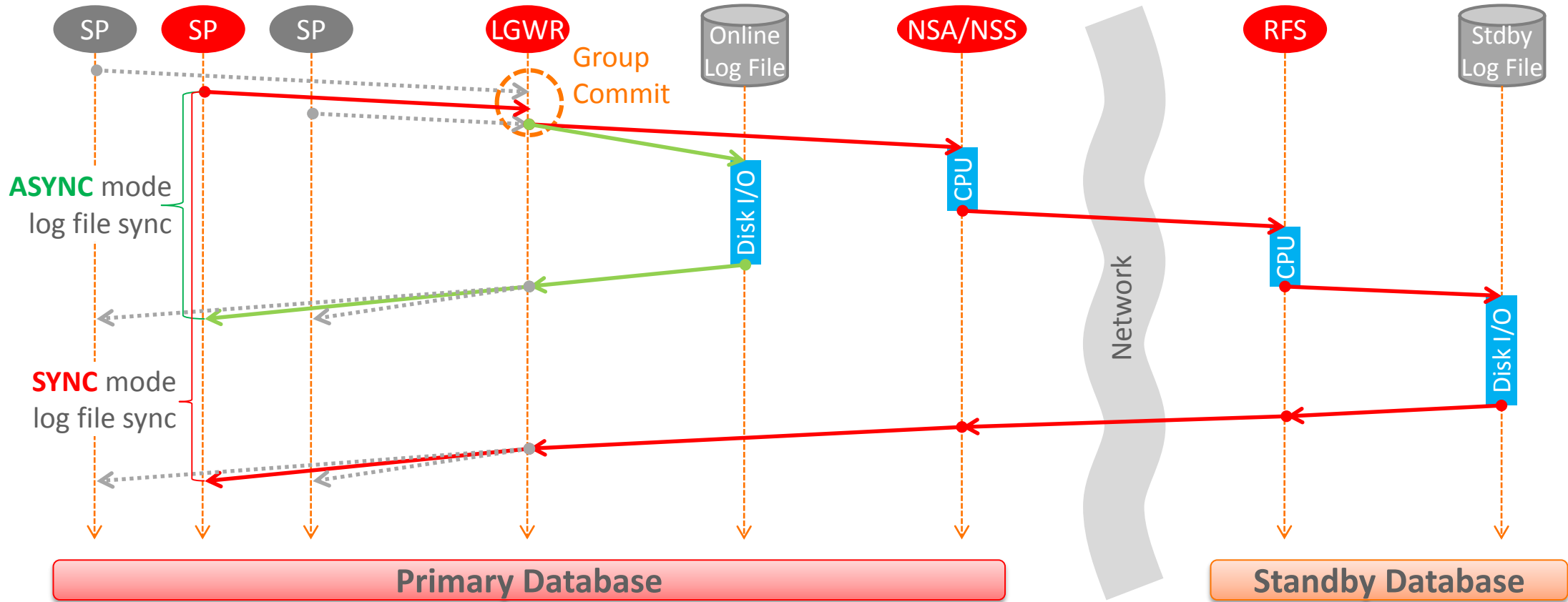
同期転送と非同期転送

アーキテクチャ



Oracle Data Guard

Sequence Diagram of Redo Transfer



同期転送と非同期転送

データ保護の要件と性能要件の選択

※ データロス:
未転送のままプライマリに残った
REDOをスタンバイに適用できない

- 同期転送と非同期転送の切替はオンラインで変更可能

	同期転送	非同期転送
メリット	ゼロ・データロスを実現可能	性能への影響がほぼない
デメリット	性能への影響を検討する必要がある	データロスに関する検討が必要



- アプリケーション特性 (REDO生成量)
- ネットワーク帯域、遅延
- スタンバイREDOログファイルのI/O性能

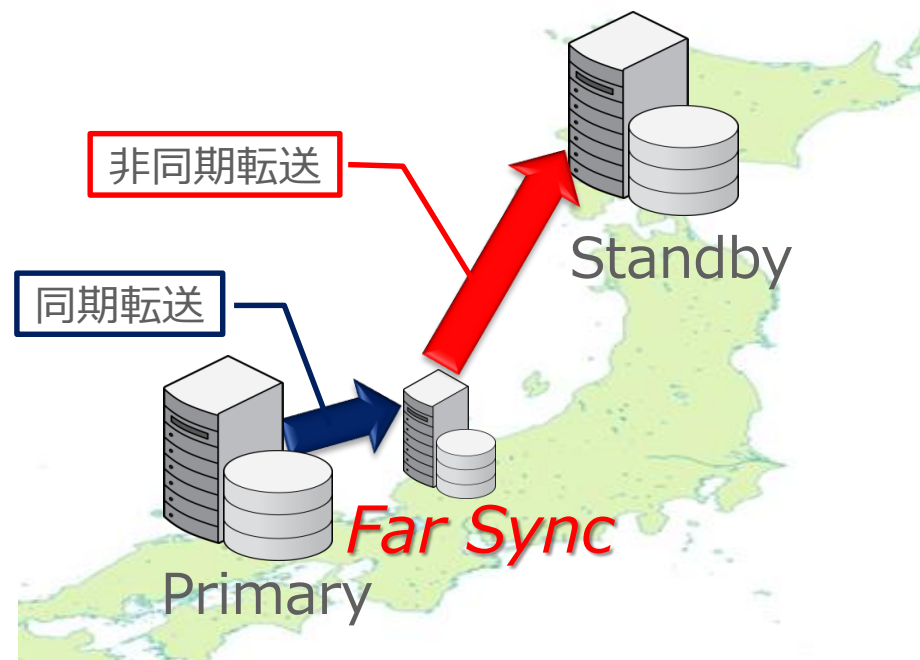


- プライマリがmountできれば
全てのRedoを転送出来る為、
データロスは回避可能
- トランザクション再実行

遠隔地スタンバイ構成での同期転送

Far Sync (Oracle Database 12c Release 1~)

- 遠隔地スタンバイで実現が難しかった「ゼロ・データロス」のスタンバイ構成を実現



- 同期転送のオーバーヘッド軽減**
 - ✓ 近距離の遠隔同期インスタンスまでの同期転送
- ゼロ・データロスの実現**
 - ✓ Primary DB 停止時にも、必要な REDO データは遠隔同期インスタンスへ転送済み
- 最小限のファイル構成**
 - ✓ 遠隔同期インスタンスは制御ファイルと REDO ログファイルのみから構成
- シームレスなロール変換**
 - ✓ 遠隔同期インスタンスを意識せずスイッチオーバーの実行が可能

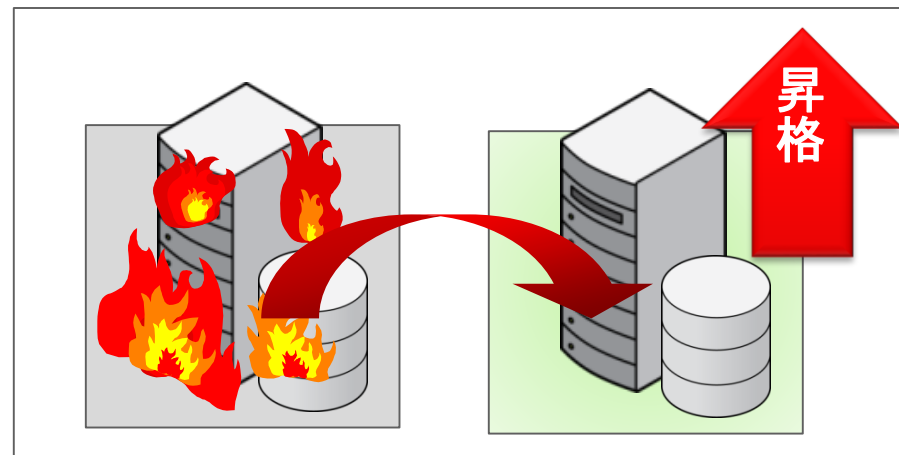
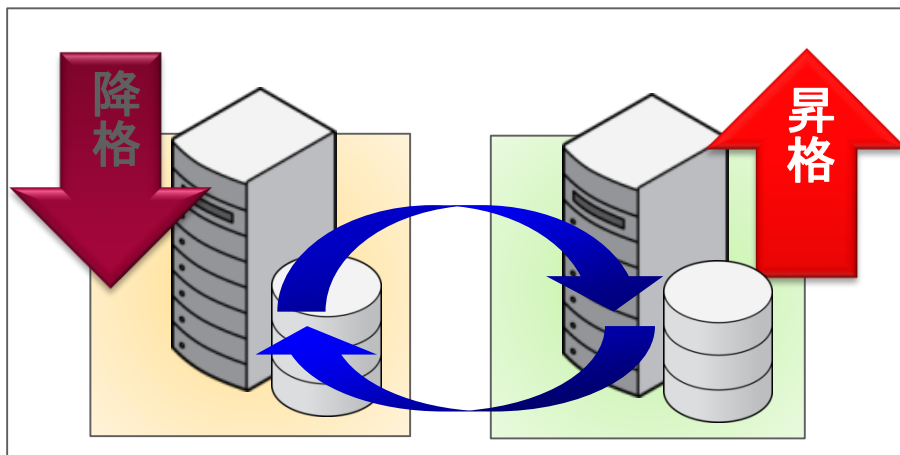
ゼロ・データロスのData Guardの切替操作

- スイッチオーバー

- 計画停止用途
- データロスなしを保証

- フェイルオーバー

- 計画外停止用途
- 同期転送ならデータロスなし
- 非同期転送ならデータロスあり
(未転送データ分)

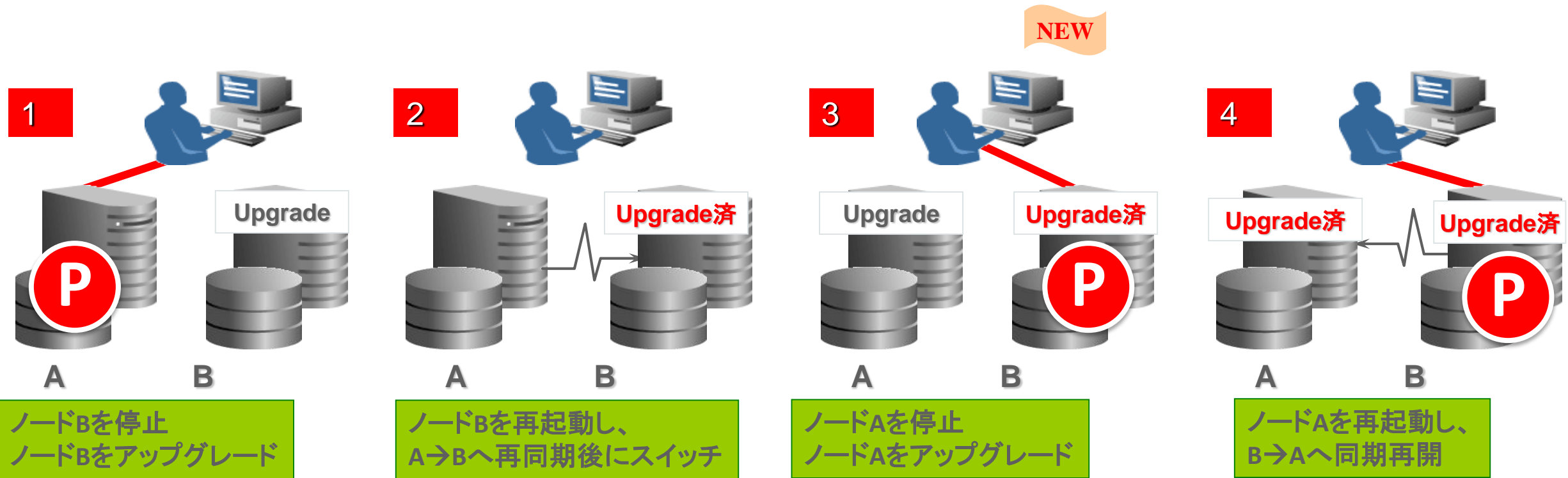


どちらも、サーバー / OS / ストレージ構成に依存しない手順
SQL、またはOracle Enterprise Managerの1クリックで実行可能

Data Guardのスイッチ・オーバーの活用

ローリング・アップグレード

- 計画停止時間を極小化して、パッチ適用やアップグレードの実施が可能



※必要に応じて、スイッチ・オーバー

A woman with long brown hair and glasses, wearing a brown leather jacket and a blue patterned scarf, is sitting at a wooden table in a cafe. She is holding a black mobile phone to her ear with her left hand and looking down at a newspaper on the table with her right hand. The background is a blurred cafe interior with other tables and chairs, and a person is visible sitting at a table in the distance. The overall lighting is soft and natural.

Flashback Technology

Flashback Technology

もしもの時の救世主！！

- ユーザー・エラーからの早急かつ容易な復旧が可能

- DBのバックアップ全体のリストア不要⇒変更されたブロックのみをリストア、DBを特定時点まで戻す
- 過去データの参照が可能！⇒ **不正なデータ改竄防止に効果**

■Flashback機能による復旧イメージ



■Flashback機能の種類

機能名	使用する領域	利用するデータ	概要	機能カテゴリ	指定方法	使用可能なエディション
Flashback Query	UNDO	UNDOデータ	テーブル単位で過去時点のデータを表示	参照	SCN TIMESTAMP	SE/EE
Flashback Version Query	UNDO	UNDOデータ	指定した時間間隔のすべての変更履歴を行単位で表示	参照	SCN TIMESTAMP	SE/EE
Flashback Transaction Query	UNDO	UNDOデータ サブメンタルログ	一定期間に行われたトランザクションの変更履歴とUNDO_SQL文を表示	参照	XID	EE
Flashback Data Archive	FDA領域	履歴表データ	一定期間の履歴データを保持し、テーブル単位で遠い過去のデータを表示	参照	SCN TIMESTAMP	EE
Flashback Transaction	UNDO	UNDOデータ サブメンタルログ	トランザクション単位での操作取り消し	リカバリ	XID	EE
Flashback Table	UNDO	UNDOデータ	テーブル単位で過去の時点へ復旧	リカバリ	リストアポイント SCN TIMESTAMP	EE
Flashback Drop	RECYCLEBIN	RECYCLEBIN	Dropしたテーブルを復旧	リカバリ	テーブル指定	SE/EE
Flashback Database	高速リカバリ領域	Flashback log	データベース全体を過去の特定の時点に復旧	リカバリ	リストアポイント SCN TIMESTAMP	EE

Flashback Database

データベース全体を指定された過去の時点の状態へ

- 活用例

- 人的ミス(データ削除や不適切な更新処理の実行等)からの[迅速な復旧](#)
- Primary側でLost Writeが発生した為に、Data GuardでFail-Over実行後、旧Primaryを新Standby環境として迅速に復旧させる場合

- Oracle Database独自のロギング・メカニズム (Flashback Log)

- データ更新時、自動的に更新前ブロック・イメージを高速リカバリ領域に保存
- Flashback Log(更新前ブロック・イメージ)でDB全体の復旧を実現
 - リストア/リカバリ不要、Export/Import処理よりも高速

- Flashback Databaseでリカバリ可能なオペレーション

- DML処理、TRUNCATE、スキーマ・ユーザーの削除(DROP USER)

Flashback Database によるリカバリ

リカバリに要する時間のイメージ

Point-in-Timeリカバリの
場合

バックアップファイルのリストア

ロールフォワード

Flashback Databaseに
よるリカバリの場合

Flashback Log
の適用

時間

【参考】Flashback Database

有効化の設定

- Flashback Logの取得を有効化(選択可能)
 - Flashback LogモードをON
 - **SQL> ALTER DATABASE FLASHBACK ON;**
 - 保証付きリストア・ポイントを作成
 - **SQL> CREATE RESTORE POINT <GRPName>**

保証付きリストア・ポイントは、復旧可能な時点をGRP作成時点に限定し、Flashback Logの生成量を大幅に抑制可能

GUARANTEE FLASHBACK DATABASE;

- その他の設定
 - Archive Logモード
 - 高速リカバリ領域(DB_RECOVERY_FILE_DEST / DB_RECOVERY_FILE_DEST_SIZE)
 - Flashback Log保持期間(DB_FLASHBACK_RETENTION_TARGET)

【参考】Flashback Database

復旧の実行例

1. MOUNTモードで起動 (OPEN状態では実行不可)

2. Flashback Databaseの実行

```
SQL> FLASHBACK DATABASE TO <?>;
```

3. 読み取り専用でOPENして、データ確認

```
SQL> alter database open READ ONLY;
```

4. RESETLOGSでOPEN

```
SQL> shutdown immediate;
```

```
SQL> startup MOUNT;
```

```
SQL> alter database open RESETLOGS;
```

✓SCN指定

```
SCN 100
```

✓現時点から1時間前指定

```
TIMESTAMP (SYSDATE - 1/24)
```

✓絶対日時指定

```
TIMESTAMP (TO_TIMESTAMP (
  '2012/11/20 12:00:00',
  'YYYY/MM/DD HH24:MI:SS'))
```

✓リストア・ポイント指定

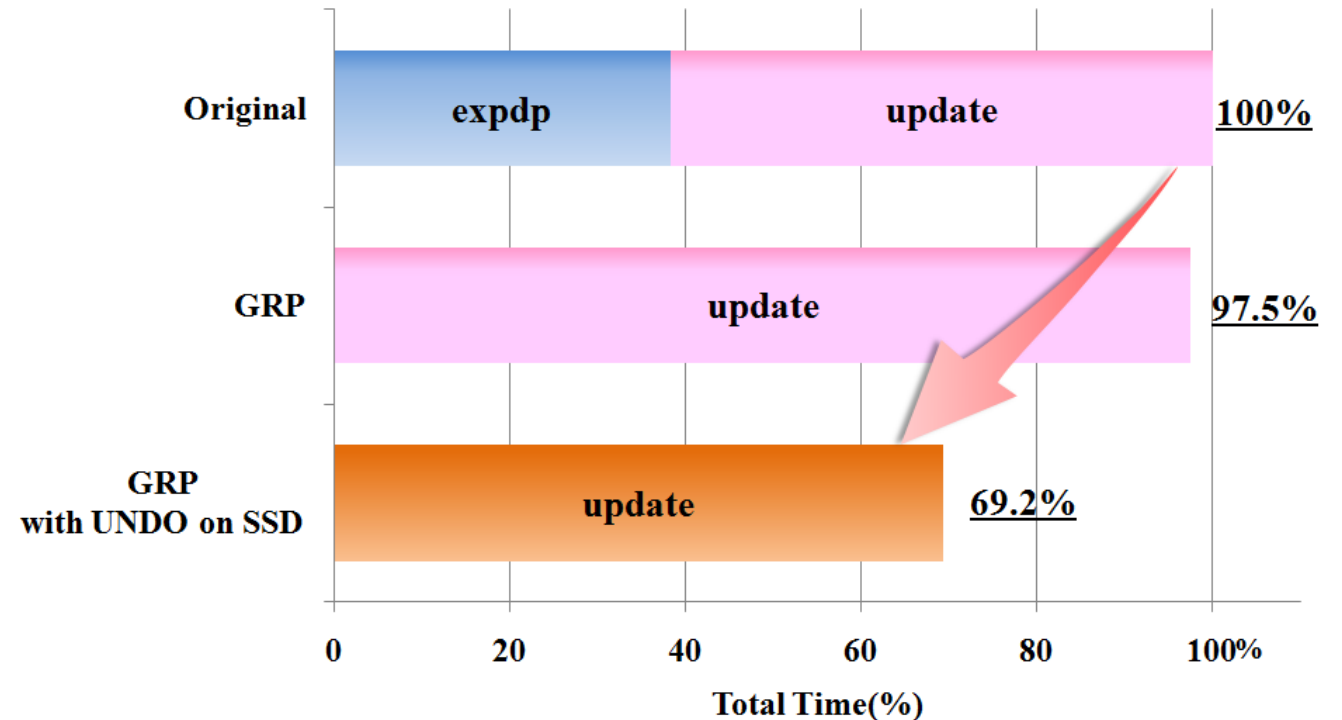
```
RESTORE POINT <RPName>
```

Flashback Database

Flashback Loggingのオーバーヘッドとチューニング

- Flashback Loggingの有り無しで、全件update処理時間が増加する傾向
 - UNDOの物理読み込みがボトルネック
- UNDO表領域をSSD上に配置することで、高速化を実現

Flashback LogではUNDOの更新前ブロックも保持する必要がある為、通常時には発生しないUNDOの物理読み込みが追加される



【White Paper】 進化したバックアップ／リカバリを実現するFlashback Database活用のベストプラクティス

<http://www.oracle.com/jp/gridcenter/partner/nssol/wp-fbdb-gridcent-nssol-v2-484389-ja.pdf>

Oracle MAAで救われた実例

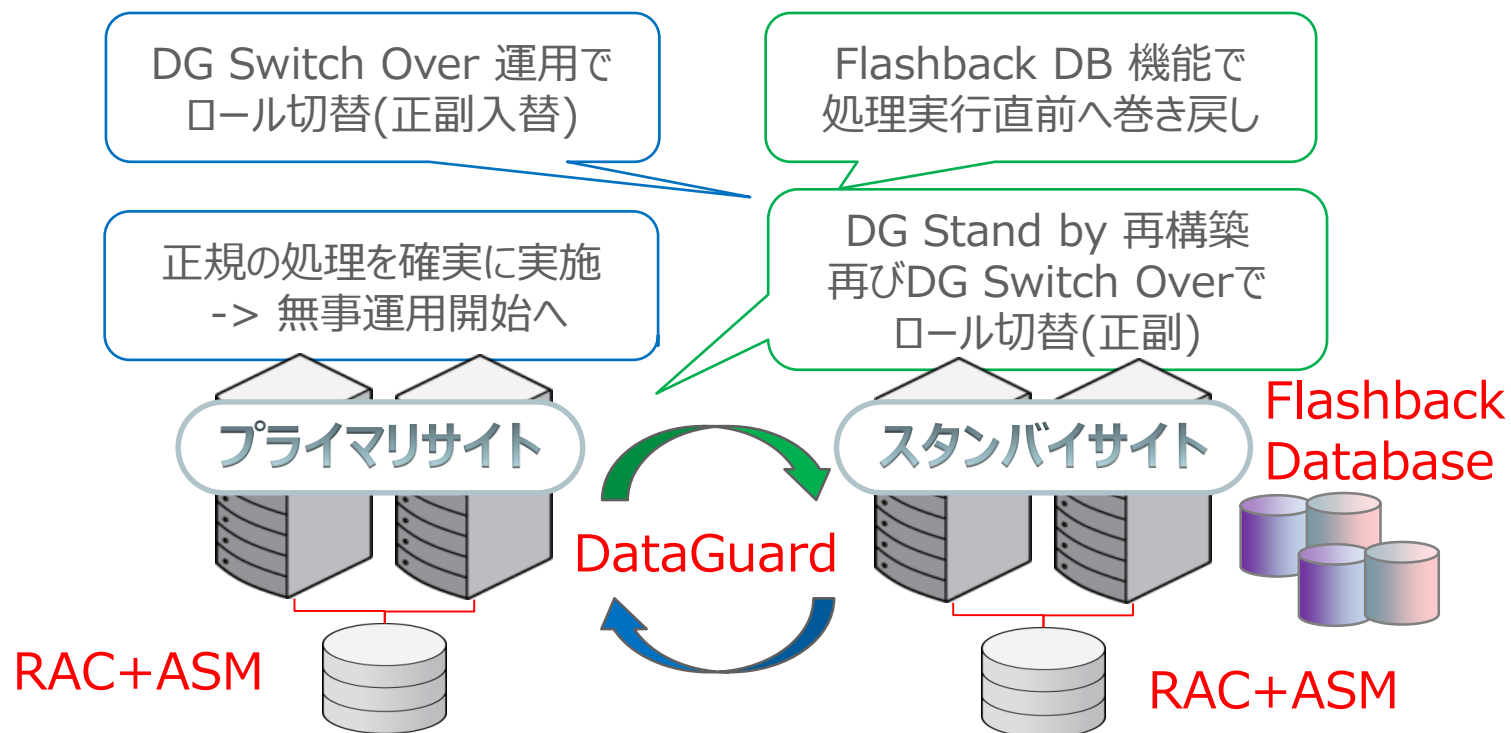
カットオーバー直前のデータ移行処理の実行順序を誤り、論理データ破壊！
稼働前なので、バックアップはデータベース作成直後のものしか取得していない！！

- https://blogs.oracle.com/dbjp/entry/exadata_000193

RAC/ASM構成 (Exadata)

+ **Data Guard** による DR環境

プライマリ・データベースの性能を重視
しスタンバイ・データベースのみで
Flashback Log を確保



A woman with long brown hair and glasses is sitting at a wooden table in a bright, modern office or cafe. She is wearing a brown leather jacket over a blue patterned scarf. She is holding a black mobile phone to her ear with her left hand and looking down at a newspaper or magazine on the table with her right hand. The background is slightly blurred, showing other people and large windows.

Summary

Oracle **M**aximum **A**vailable **A**rchitecture



Oracle Database が提供する高可用性ソリューション(MAA)は可用性を高めるだけでは留まらず、性能および管理性の向上も実現します。

- Oracle **R**ead **A**pplication **C**lusters
- Oracle **A**utomatic **S**torage **M**anagement
- Oracle **D**ata **G**uard
- **Flashback** Technology

Oracle Maximum Availability Architecture

Production Site

RAC

- Scalability
- Server HA

ASM

- Local storage protection

Advanced Security

- Data encryption, protection

Flashback

- Human error correction

Edition-based Redefinition, Online Redefinition, Data Guard, GoldenGate

- Minimal downtime maintenance, upgrades, migrations

Enterprise Manager Cloud Control

- Site Guard, Coordinated Site Failover

Application Continuity

- Application HA

Global Data Services

- Service Failover / Load Balancing



Application Test Suite, Real Application Testing

- Minimal Testing Costs

Active Replica

Active Data Guard

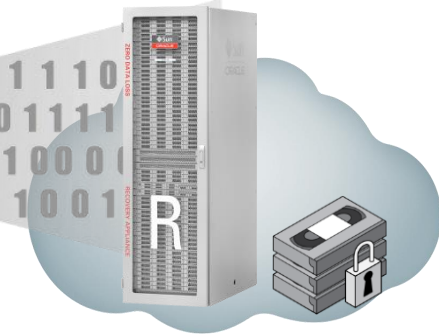
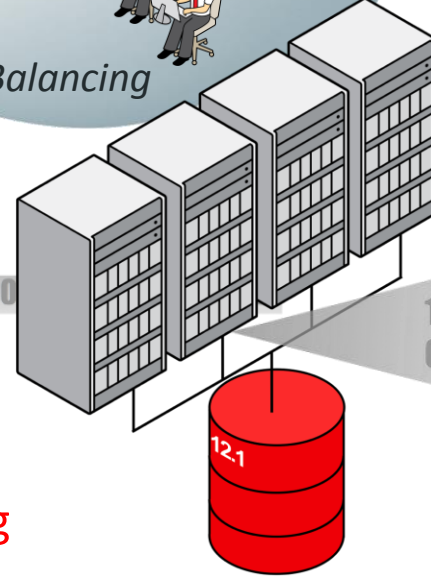
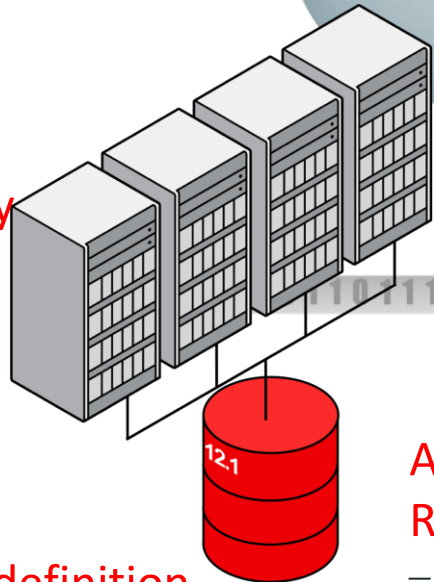
- Data Protection, DR
- Query Offload

GoldenGate

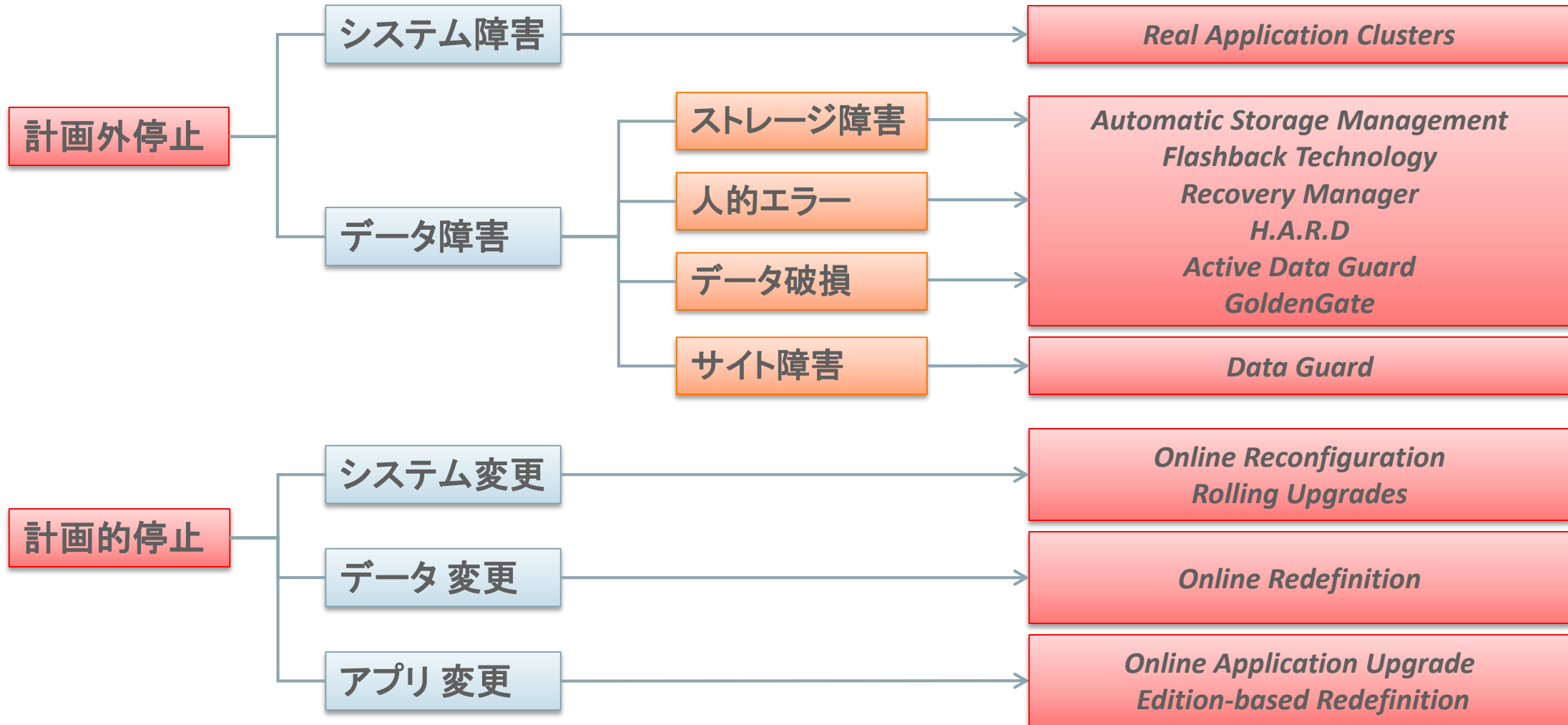
- Active-active replication
- Heterogeneous

RMAN, Oracle Secure Backup, Zero Data Loss Recovery Appliance

- Backup to disk, tape or cloud



Oracle MAAの高可用性機能



Oracle MAA: サービス・レベル区分

計画外停止及び計画停止に対するサービス・レベル

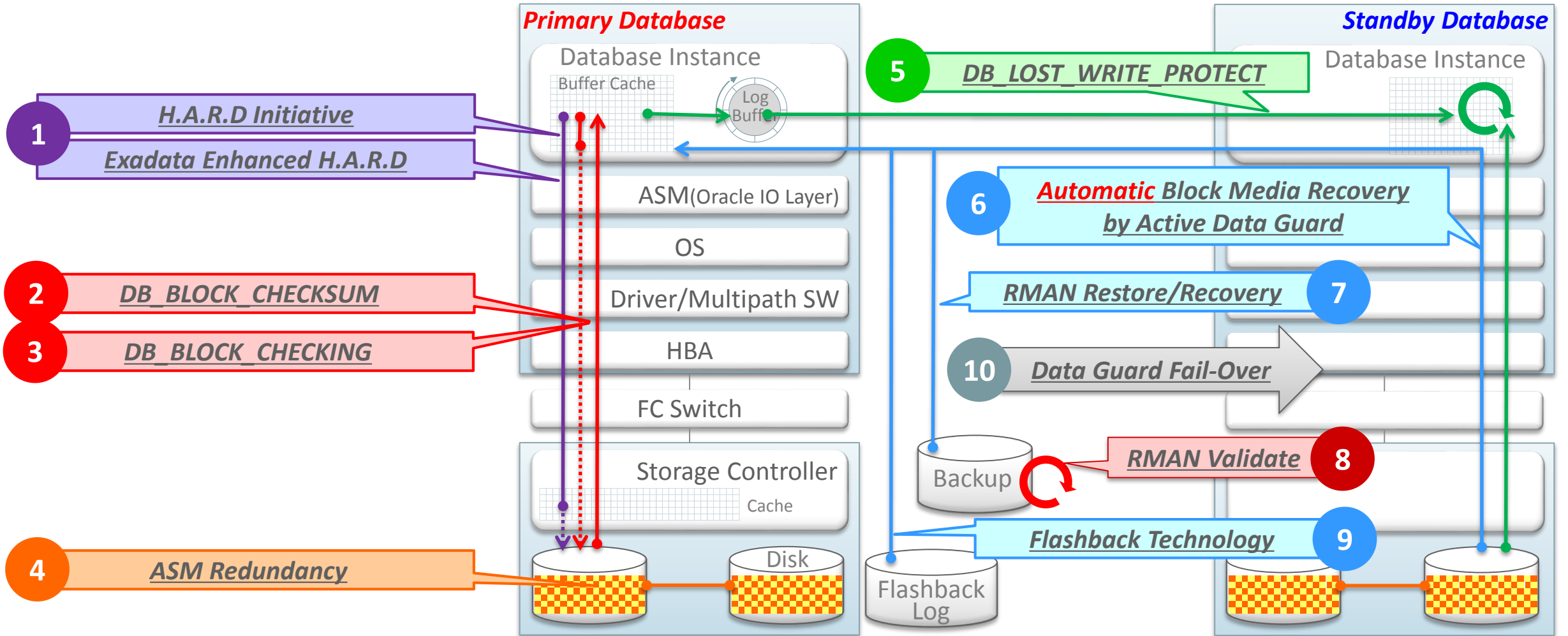
停止クラス 高可用性層	計画外停止(ローカル・サイト)	計画メンテナンス	データ保護	リカバリ不能なローカル停止および障害 時リカバリ
Platinum	Platinum対応アプリケーションでは <u>アプリケーションの停止なし</u>	<u>アプリケーションの停止なし</u>	包括的なランタイム検証と 手動チェックの組合せ	Platinum対応アプリケーションではアプリ ケーションの停止なし、処理中のトランザク ションを維持、データ損失ゼロ
Gold	包括的な高可用性と <u>障害時リカバリ</u>	すべてがローリングまたは オンライン	包括的なランタイム検証と 手動チェックの組合せ	リアルタイム・フェイルオーバー、ゼロまたはゼ ロに近いデータ損失
Silver	自動フェイルオーバーを含む 高可用性	一部ローリング、 一部オンライン、 一部オフライン	基本的なランタイム検証と 手動チェックの組合せ	バックアップからのリストア、最後のバックアッ プ以降に生成されたデータを失う可能性
Bronze	単一インスタンス、リカバリ可能な インスタンスおよびサーバー障害での 自動再起動	一部オンライン、 大部分オフライン	基本的なランタイム検証と 手動チェックの組合せ	バックアップからのリストア、最後のバックアッ プ以降に生成されたデータを失う可能性

Oracle MAA: サービス・レベル区分

本セッションでご紹介した機能のマッピング

		Data Guard GoldenGate	RAC	ASM	Flashback
PLATINUM	重要アプリケーションに対して無停止 データロス：ゼロ	✓	✓	✓	✓
GOLD	包括的な高可用性と災害からの保護 データロス：ゼロもしくはほぼゼロ	✓	✓	✓	✓
SILVER	ローカルサイト障害に対する高可用性 データロス：最新バックアップ時点まで保護		✓ RAC One Node	✓	✓
BRONZE	シングル・インスタンス、基本的なサービス再起動 データロス：最新バックアップ時点まで保護			✓	✓

Oracle MAA の Data Protection



より詳細な技術資料

- しばちょう先生の試して納得！ DBAへの道
 - <http://www.oracle.com/technetwork/jp/database/articles/shibacho/index.html>
- 【Oracle DBA & Developer Day 2014】
しばちょう先生による特別講義！ RMANバックアップの運用と高速化チューニング
 - http://www.oracle.com/webfolder/s/delivery_production/docs/FY15h1/doc8/B2-2-print.pdf
- 【Oracle DBA & Developer Day 2013】
高可用性ベスト・プラクティスによるデータ破壊対策完全版
 - <http://www.oracle.com/webfolder/technetwork/jp/ondemand/ddd2013/C-2.pdf>
- 【Oracle DBA & Developer Day 2012】
高可用性システムに適した管理性と性能を向上させる ASM と RMANの魅力
 - http://otndnld.oracle.co.jp/ondemand/ddd/PDF/MA-4_print_c.pdf

Oracle Real Application Clusters/Oracle Clusterware の高可用性機能

 - http://otndnld.oracle.co.jp/ondemand/ddd/PDF/MA-2_print_c.pdf

Oracle Data Guard / Oracle GoldenGate 高可用性のための実践Tips

 - http://otndnld.oracle.co.jp/ondemand/ddd/PDF/MA-3_print_r.pdf

ご質問・ご相談等ございましたら、
お気軽にお問い合わせください。

あなたにいちばん近いオラクル

Oracle Direct
0120-155-096

(平日9:00-12:00 / 13:00-18:00)

<http://www.oracle.com/jp/direct/index.html>

Oracle Direct

検索

各種無償支援サービスもごございます。

Hardware and Software **Engineered to Work Together**

ORACLE®