

Oracle DBA & Developer Days 2011

日本オラクル、今年最大の技術トレーニングイベント

2011年11月9日(水)～11月11日(金) シェラトン都ホテル東京

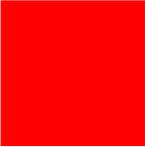


ORACLE®

[B-2] 実践！ Oracle Data Guard の 導入から有効活用までのポイント解説

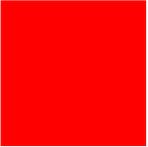
日本オラクル株式会社テクノロジー製品事業統括本部 ソリューション本部
プリンシパルエンジニア 塚井 知之

2011年11月9日



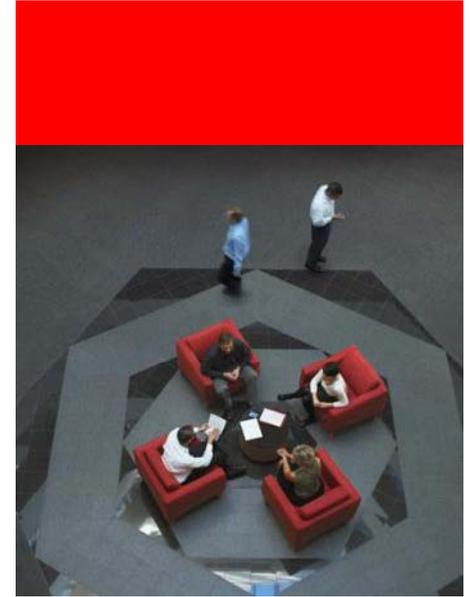
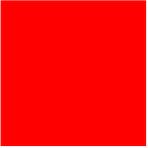
以下の事項は、弊社の一般的な製品の方向性に関する概要を説明するものです。また、情報提供を唯一の目的とするものであり、いかなる契約にも組み込むことはできません。以下の事項は、マテリアルやコード、機能を提供することをコミットメント(確約)するものではないため、購買決定を行う際の判断材料になさらないで下さい。オラクル製品に関して記載されている機能の開発、リリースおよび時期については、弊社の裁量により決定されます。

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。



本日のアジェンダ

- Oracle Data Guard 概要
- Oracle Data Guard 活用のポイント
 - 設計・導入フェーズのポイント
 - 構築フェーズのポイント
 - 運用フェーズのポイント
 - スタンバイ・有効活用のポイント
- まとめ



Oracle Data Guard 概要

Oracle Data Guardの進化

低コストで最大限のデータ保護と業務継続性を提供

待機系の有効活用による投資対効果の向上

メンテナンス停止時間の削減による
更なる高可用性要件への対応

低コストで
災害対策の実現

スタンバイ・
データベース

1996年



- フェイルオーバー機能
- ファイル名の自動変換機能
- OPS(9i以降はRAC)環境のサポート (*1)

1999年



- アーカイブの自動転送
- アーカイブの自動適用
- 複数スタンバイのサポート

2001年



- 同期と非同期転送の提供
- スイッチオーバー機能
- ギャップの自動解決機能
- データ保護モードの実装
- データファイルの自動追加
- ロジカル・スタンバイの提供

2004年



- リアルタイム適用
- ローリング・アップグレード機能
- 自動フェイルオーバー機能
- GUIによる管理機能の強化

2007年



- リアルタイム問合せ (*2)
- 破損ブロックの自動修復 (*2)
- REDO転送の圧縮 (*2)
- スナップショット・スタンバイの提供

2009年



国内外あわせて確認できるだけで
500サイト以上の採用実績

(*1) OPS = Oracle Parallel Server (*2) 利用するにはオプション・ライセンスが必要です
RAC = Real Application Clusters

データベース機能における技術革新

ディザスタ・リカバリに対するストレージ技術の変遷

ストレージ (ハードウェア) による
遠隔地レプリケーションの限界

- ✓ 一般的に同機種でなければ構築できないため、本番サイトと同等の環境となり、完全な2重投資
- ✓ スタンバイ側は業務活用できないため、何も起きず使われないままハードウェアとして寿命を全うするケースがほとんど
- ✓ 複数のバックアップサイトを構築しがたい (有事の際に復旧まではSPoFとなる)
※2次災害=電力不足含む
- ✓ 一般的にソフトウェア実装と比較して正副切り替えのコストが高い (ダウンタイムの長期化)

データベースに、
より全体最適した
ディザスタ・リカバリの実現

データベース・トランザクション認識

待機系の業務利用

通信量の削減

正副切り替え

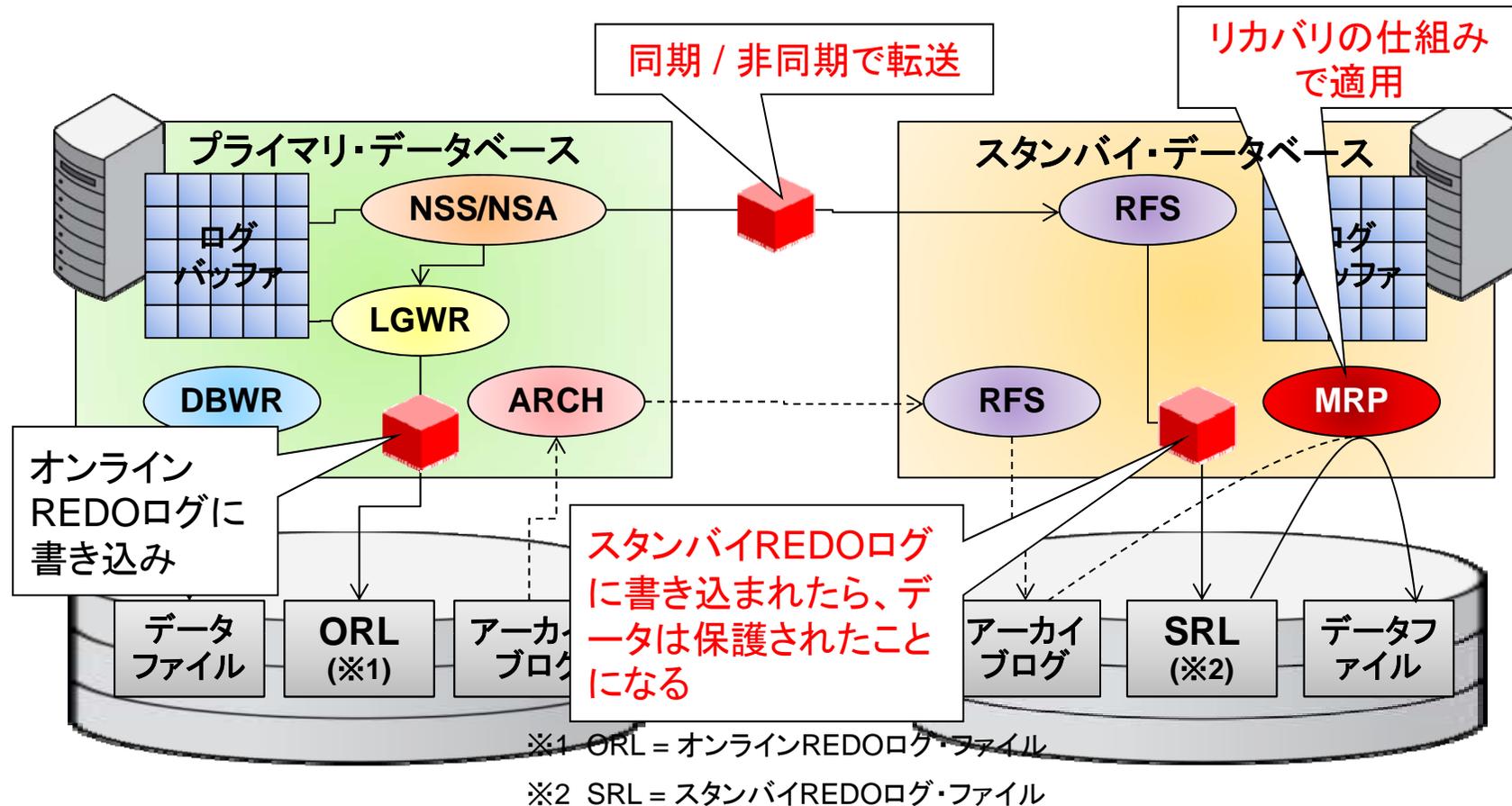
運用簡易化

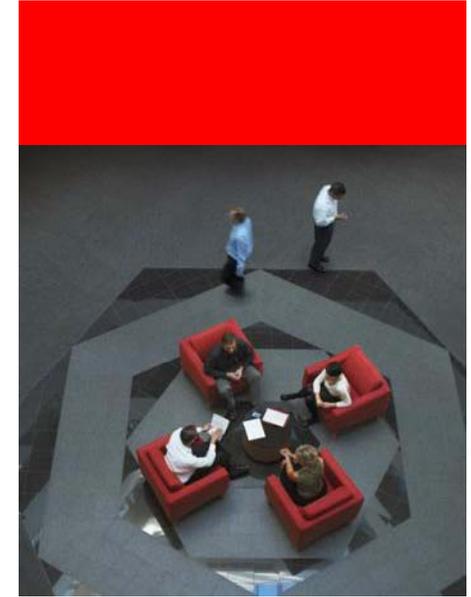
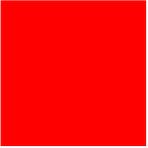
Oracleの
高可用性技術

ORACLE

Oracle Data Guardのアーキテクチャ

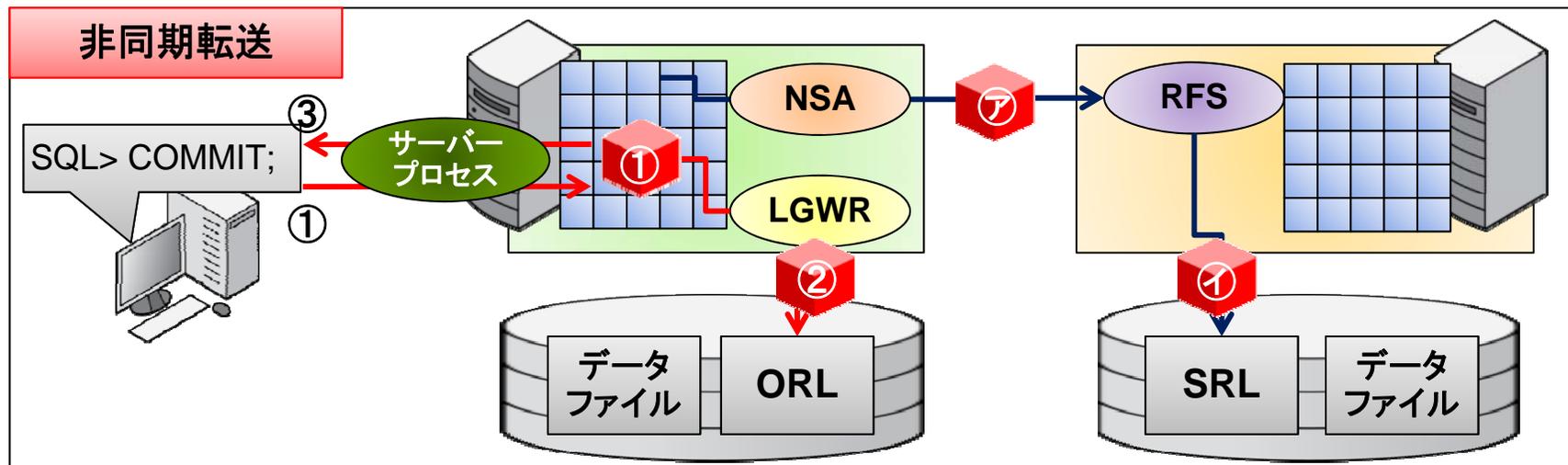
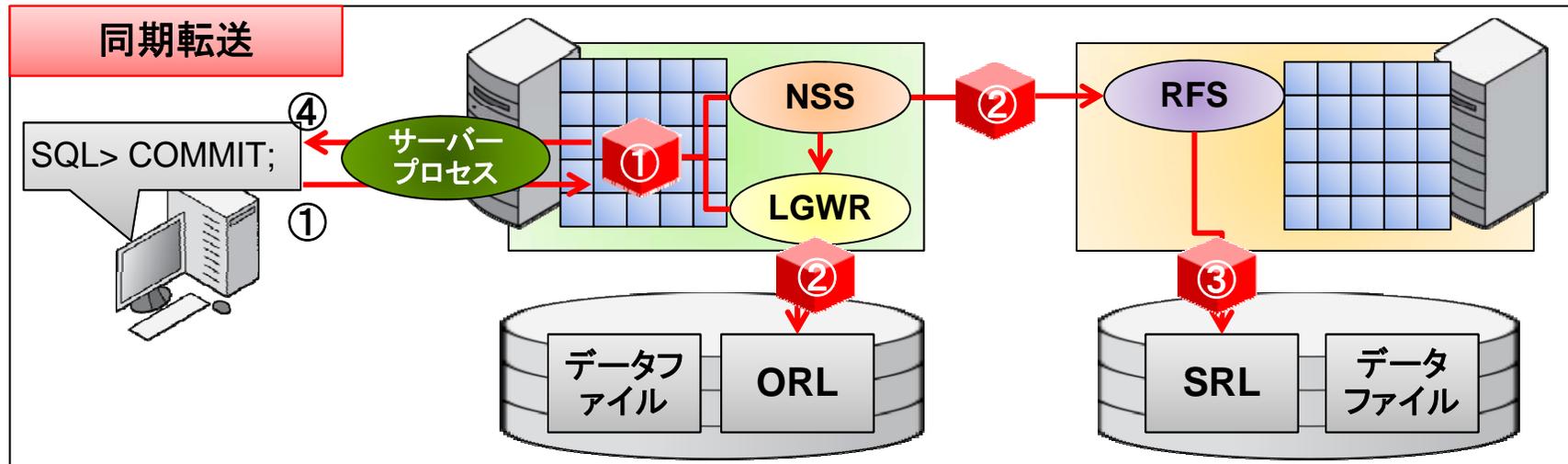
プライマリのログ・バッファからREDOを転送することで
プライマリ・データベースに与える影響を最小化



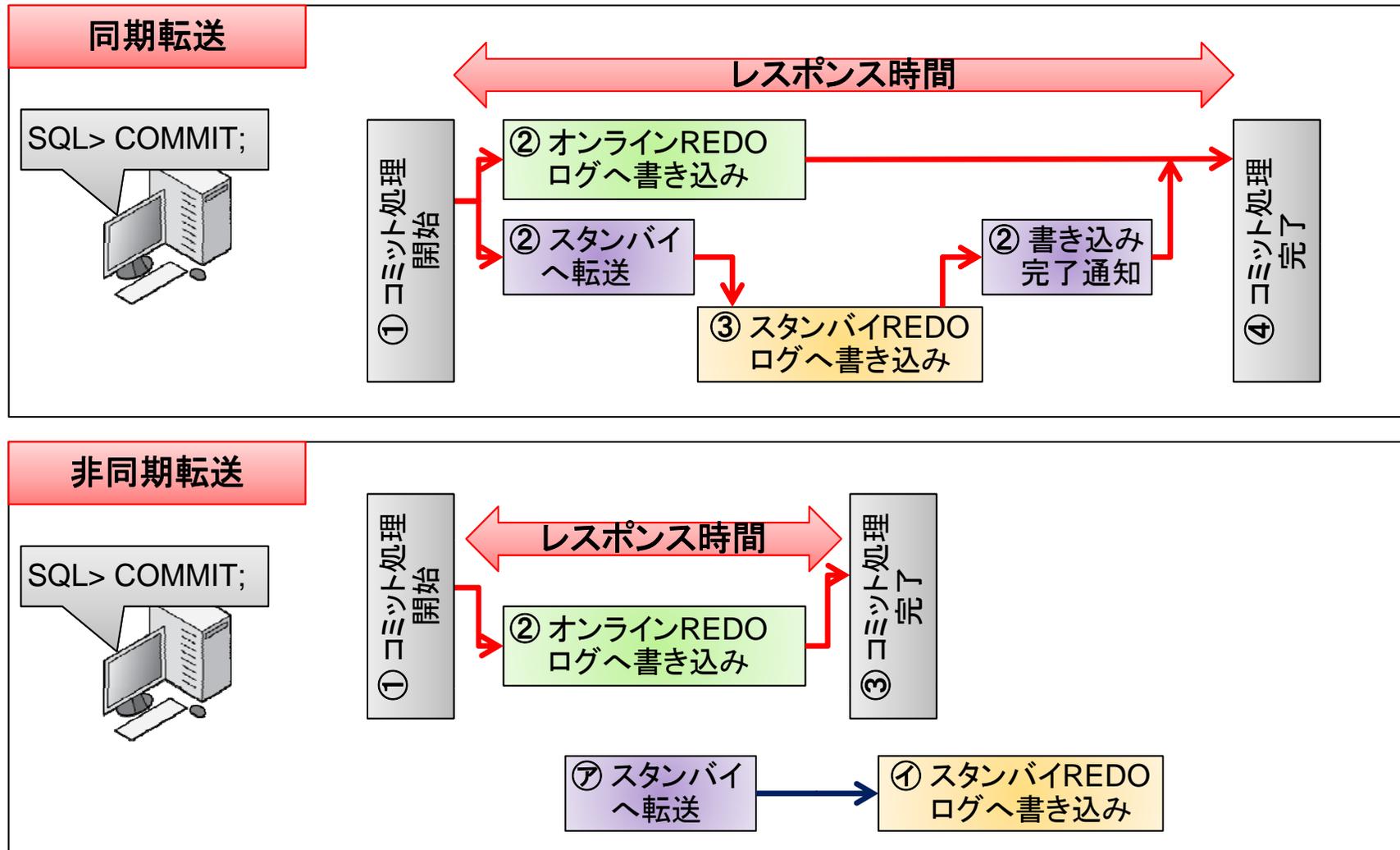


設計・導入フェーズのポイント

同期転送と非同期転送のアーキテクチャ



同期転送と非同期転送の更新処理への影響



同期転送と非同期転送の検討

データ保護の要件と性能要件の選択

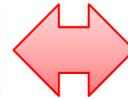
	同期転送	非同期転送
メリット	<ul style="list-style-type: none"> ゼロデータロスを実現することが可能 	<ul style="list-style-type: none"> 性能への影響がほぼない
デメリット	<ul style="list-style-type: none"> 性能への影響を検討する必要がある <p style="text-align: center;">↓</p> <div style="border: 1px solid red; border-radius: 15px; padding: 5px;"> <ul style="list-style-type: none"> アプリケーション特性 ネットワーク遅延 </div>	<ul style="list-style-type: none"> データロスに関する検討が必要 <p style="text-align: center;">↓</p> <div style="border: 1px solid red; border-radius: 15px; padding: 5px;"> <ul style="list-style-type: none"> プライマリがmountできればデータロスは回避可能 トランザクション再実行 </div>

変更

- 同期転送と非同期転送の切り替えはオンラインで変更可能
- トランザクションの重要度に合わせて、日中と夜間(バッチ)で定期的に変更することも可能

アプリケーション特性と同期転送の検討

Data Guardの転送はREDOデータのみ



検索処理への影響はなし



REDOを多く生成する処理が多いと影響を受けやすい



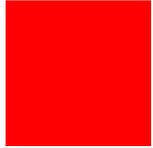
現行システムの性能情報から影響を予測

Top 5 Timed Events		Avg %Total	
~~~~~		wait	Call
Event	Waits	Time (s)	Time
-----	-----	-----	-----
CPU time		31	61.4
reliable message	52	3	6.8
control file sequential read	5,561	2	4.8
log file sync	1,236	2	4.6
log file sequential read	50	2	3.1

STATSPACKレポートの例

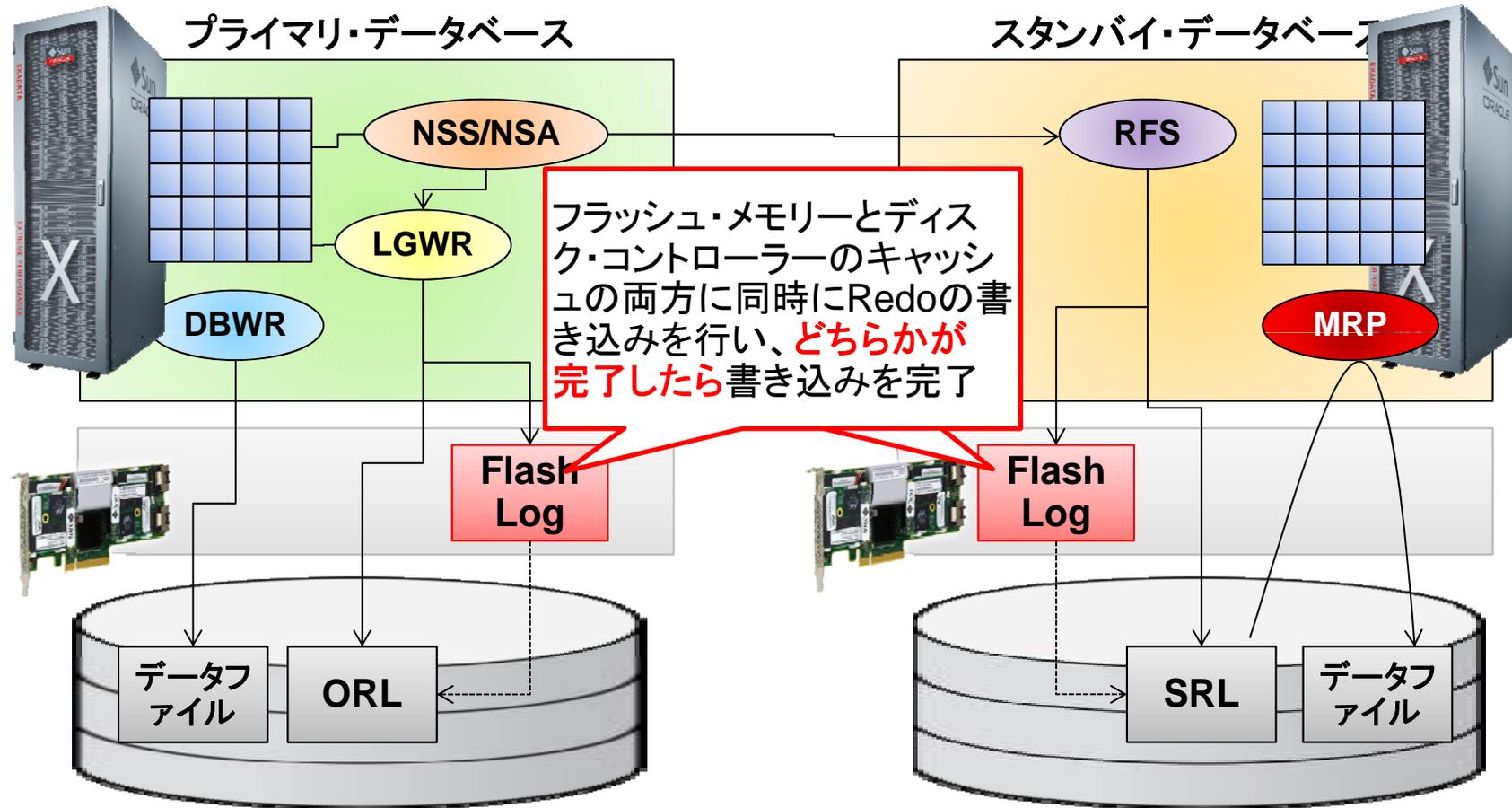
Top 5 Timed Events					
Event	Waits	Time(s)	Avg Wait(ms)	% Total Call Time	Wait Class
db file sequential read	44,150,491	323,066	7	57.6	User I/O
CPU time		190,828		34.0	
db file scattered read	3,657,924	35,654	10	6.4	User I/O
read by other session	3,685,512	18,828	5	3.4	User I/O
log file sync	3,542,011	11,210	3	2.0	Commit

AWRレポートの例



**Screen Only**

# [参考] Exadata Smart Flash Log を利用 するREDO転送



※ 本機能は Exadata Storage Server 11.2.2.4 以降で利用することができます

ORACLE

# REDO転送に必要な帯域幅

REDO転送に必要な帯域幅の試算

• **帯域幅 = REDO Size [MB/秒] ÷ 0.7**

- 0.7はREDOブロックのヘッダーやTCPパケットサイズなどを考慮した安全係数
- 初期構築時やギャップ解決時も考慮

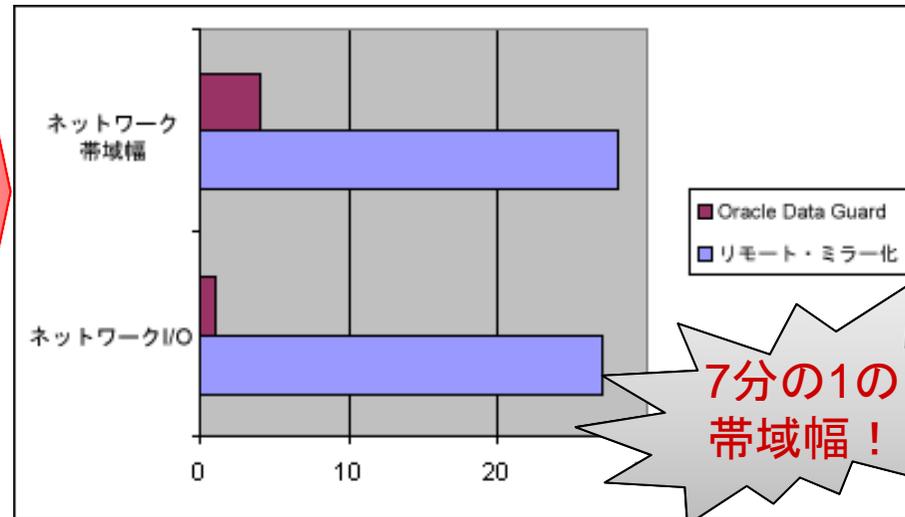
Oracle Corporation メール  
システムの内部分析の例

- REDO転送に必要な帯域幅は、データベース関連ファイル全体のおよそ**7分の1**
- つまり、ストレージミラーに比べ必要な帯域幅は**7分の1**

参考URL:

<http://www.oracle.com/technetwork/jp/database/dataguardremotemirroring-084059-ja.html>

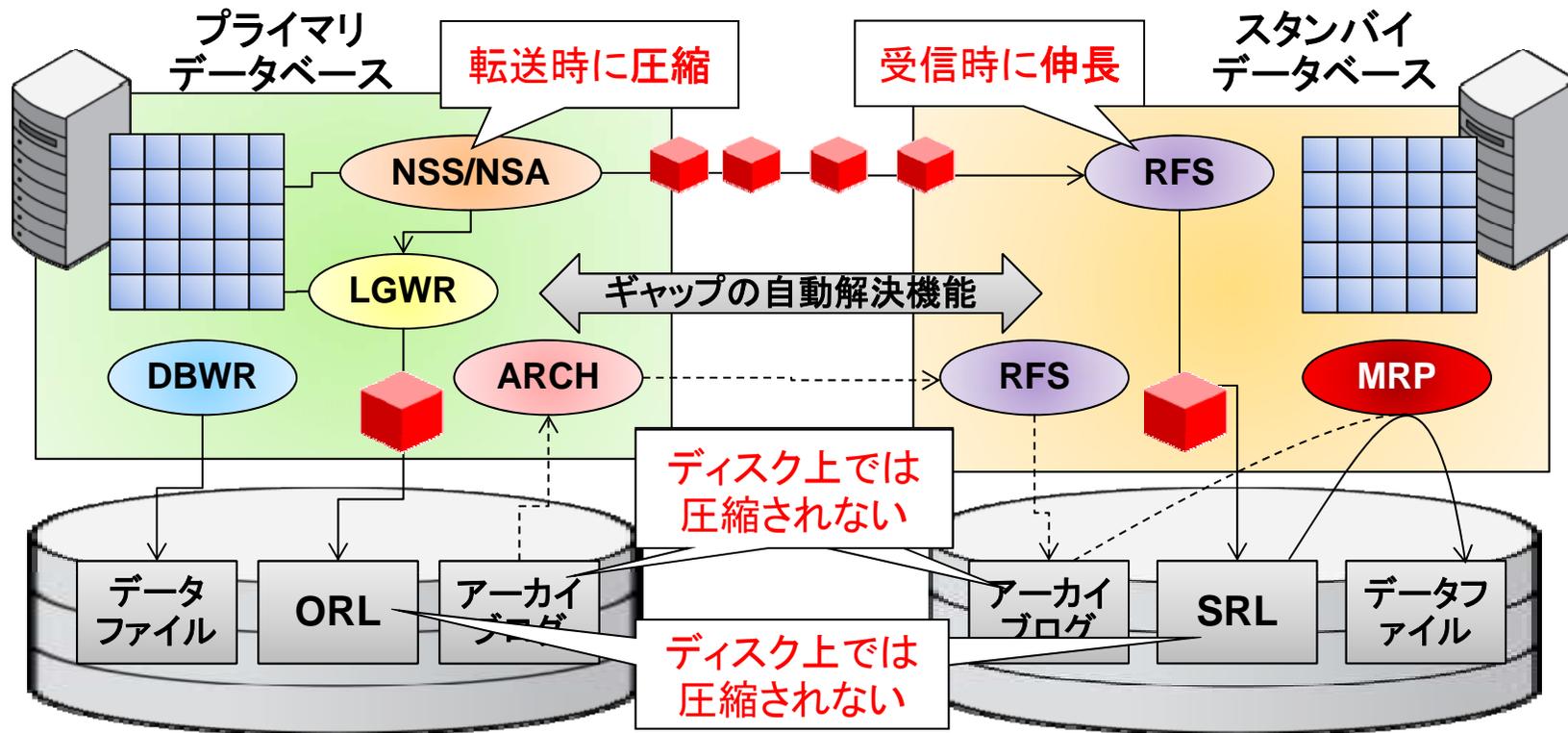
ネットワーク帯域コスト削減を実現



ORACLE

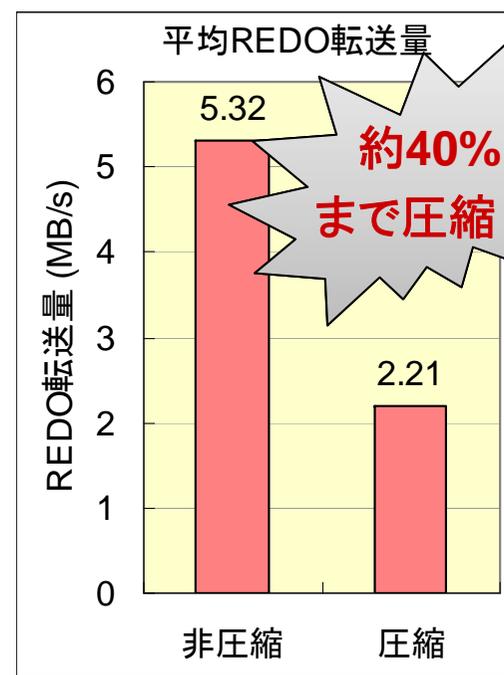
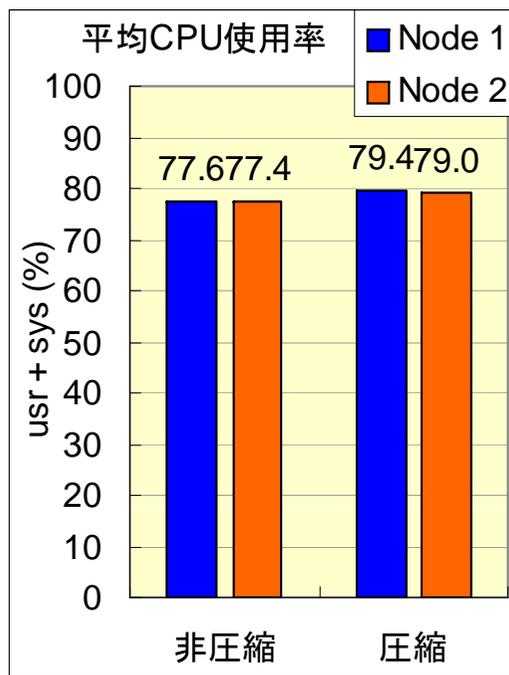
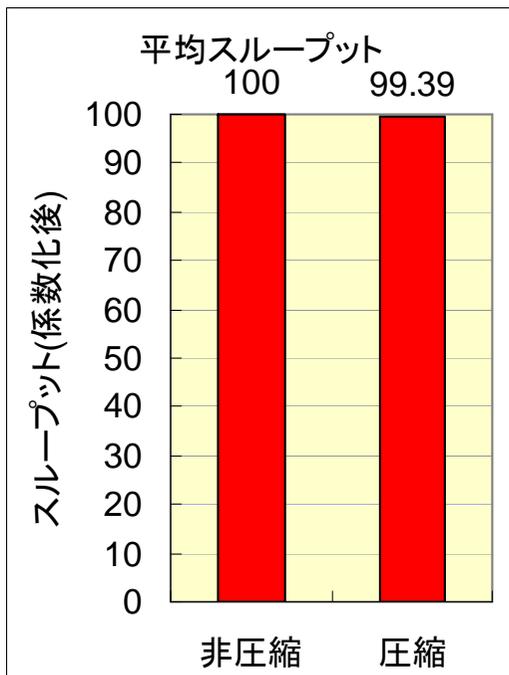
# REDO転送の圧縮の動作

- Advanced Compression Option で提供される機能 → 通信コストの削減
- 転送時に圧縮、受信時に伸長
- ごくわずかにCPUリソースの追加消費が発生



# REDO圧縮の効果(11g R1)

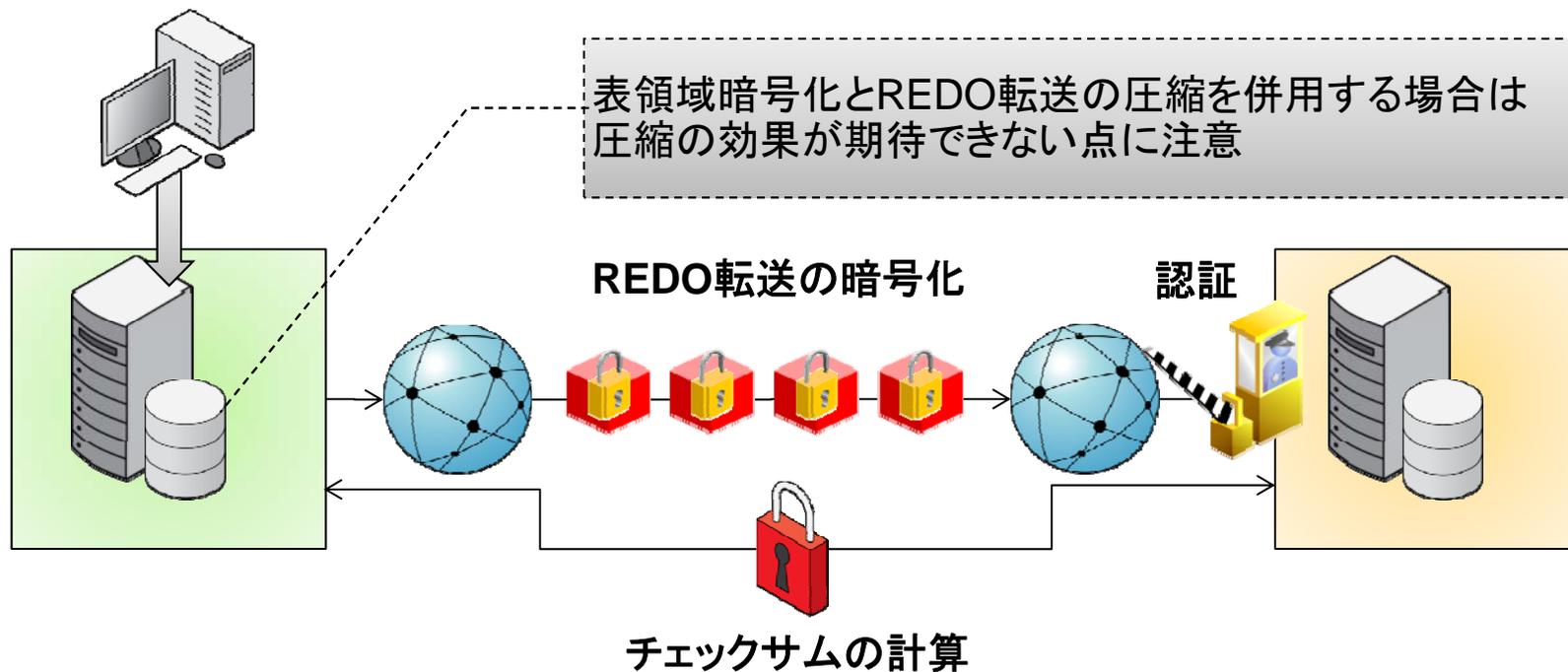
- 2ノード構成のReal Application Clusters(RAC)環境で実施
- 圧縮によってREDO転送のサイズが**40%程度まで減少**
- スループット / CPUへのオーバーヘッドはごくわずか

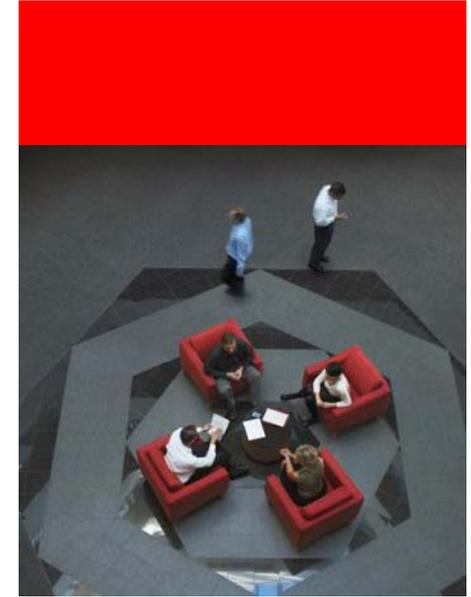
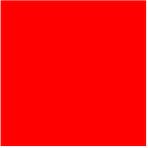


# REDO転送の暗号化

ネットワーク上でREDOデータを改竄される可能性を防止

- REDO転送には、認証ネットワークセッションを使用
- REDO転送の整合性チェックサムを計算 (Advanced Security Option)
- REDOデータを暗号化して送信 (Advanced Security Option)





# 構築フェーズのポイント

# スタンバイ・データベースの作成ステップ例

## プライマリを停止せずにスタンバイを作成することが可能

### プライマリ・データベース

1. 強制ロギングの有効化(推奨)
2. REDO転送の認証の構成
3. 初期化パラメーターの設定
4. REDO転送のためのネットワークの構成
5. アーカイブ・ログ・モードの有効化
6. フル・バックアップの取得
12. スタンバイ・REDOログ・ファイルの作成

○ = データベースの再起動が必要な操作

### スタンバイ・データベース

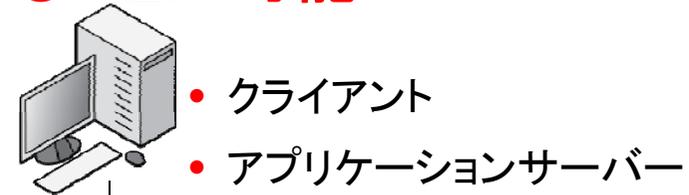
7. 初期化パラメーターの設定
8. REDO転送のためのネットワークの構成
9. フル・バックアップのリストア
10. スタンバイ・REDOログ・ファイルの作成
11. 管理リカバリモードの開始

2. パスワード・ファイルを利用するには再起動が必要 (ただし、DBCAで作成された場合はデフォルトで有効)
5. アーカイブ・ログ・モードへの変更はmount状態で実行可能

# REDO転送のネットワーク経路

宛先の指定を明示的に指定することが可能

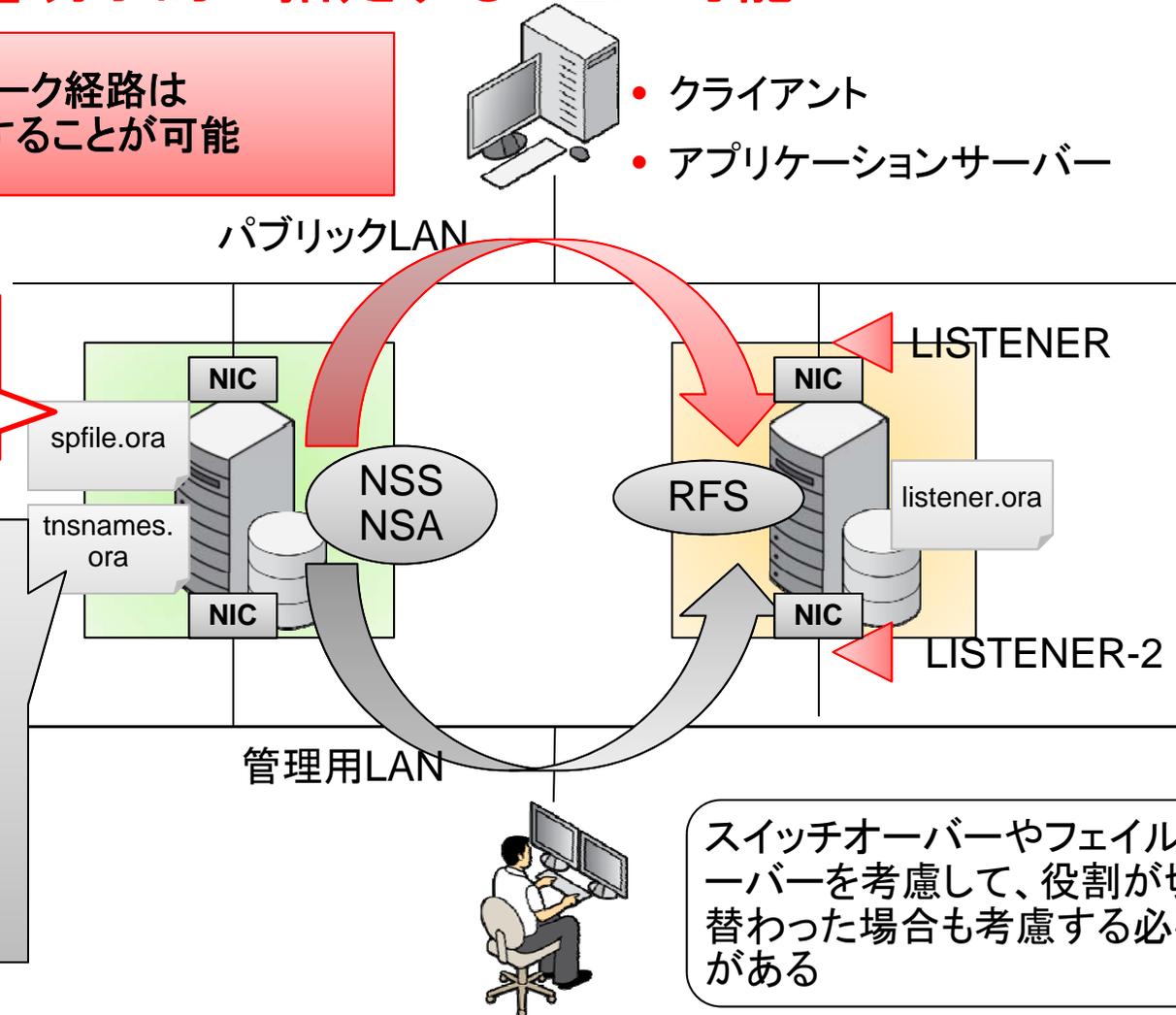
REDO転送のネットワーク経路は  
REDOの宛先で指定することが可能



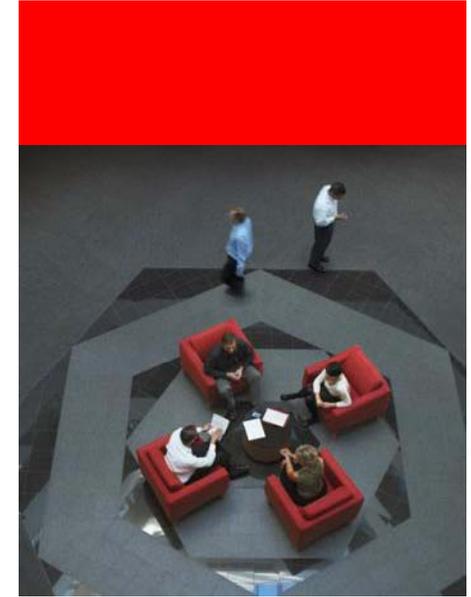
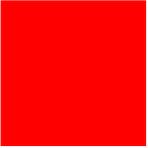
宛先の指定

```
log_archive_dest_N =  
'SERVICE = <tns名> ...
```

```
<tns名> =  
(DESCRIPTION =  
  (ADDRESS =  
    (PROTOCOL = TCP)  
    (HOST = <hostname>)  
    (PORT = 1521))  
  (CONNECT_DATA =  
    (SERVER =  
      DEDICATED)  
    (SID = stby1))  
)
```



スイッチオーバーやフェイルオーバーを考慮して、役割が切り替わった場合も考慮する必要がある



# 運用フェーズのポイント

# GUIを使ったData Guardの管理 Oracle Enterprise Managerを活用

- Enterprise Manager では、GUIベースにより Data Guard の各種機能を操作可能
- 運用担当者の作業を簡素化

The screenshot shows the Oracle Enterprise Manager interface for Grid Control 11g. The main content area displays the 'Data Guard' configuration for a standby database. Key sections include:

- 概要 (Summary):** Shows 'Data Guardステータス' as '標準' (Standard) and 'ファスト・スタートフェイルオーバー' as '無効' (Disabled).
- プライマリ・データベース (Primary Database):** Lists 'stbpolice' as the primary database with 'Data Guardステータス' as '標準'.
- スタンバイ・データベース (Standby Database):** Lists 'stbpolice' as the standby database with 'Data Guardステータス' as '標準'.
- パフォーマンス (Performance):** Includes a table for 'リアルタイム同合わせ' (Real-time synchronization) with columns for '名前' (Name), 'ホスト' (Host), 'Data Guardステータス' (Status), 'ロール' (Role), 'リアルタイム同合わせ' (Real-time synchronization), and '最終変更' (Last change).

Data Guard パフォーマンス画面

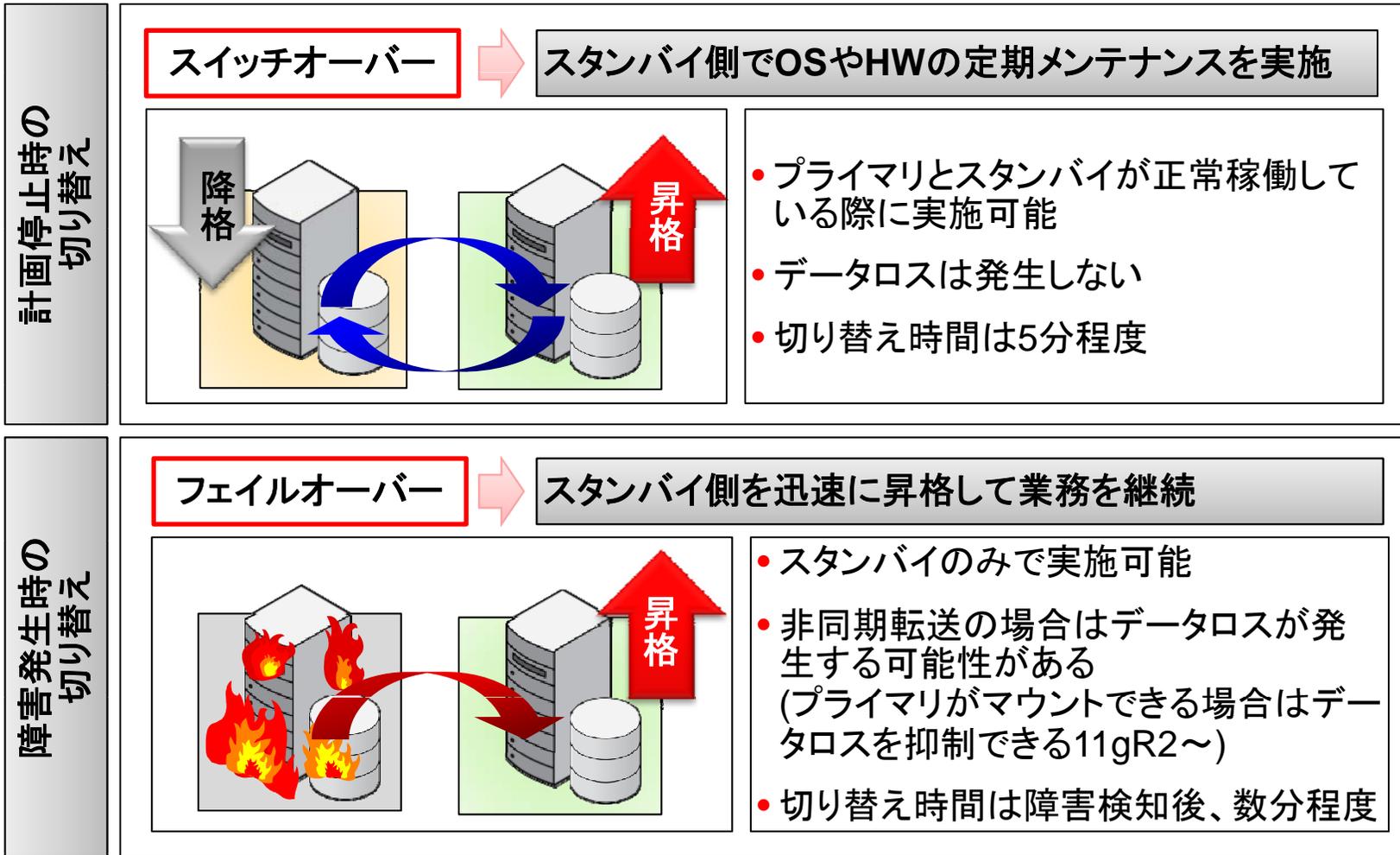
The screenshot shows the Oracle Enterprise Manager interface for Grid Control 10g, displaying performance metrics for Data Guard. Key sections include:

- Chicago_02 - REDO生成率 (REDO Generation Rate):** A line graph showing REDO generation rate in KB/sec over time. The current rate is 1 KB/sec.
- Boston_02 - タイム・ラグ (Time Lag):** A line graph showing transport lag and apply lag in seconds. The current transport lag is 0.0 seconds and the current apply lag is 0.0 seconds.
- Boston_02 - 適用頻度 (Apply Frequency):** A line graph showing apply frequency in KB/sec over time. The current apply frequency is 0.66 KB/sec.

対象データベース一覧(プライマリ/スタンバイ)

# スイッチオーバーとフェイルオーバー

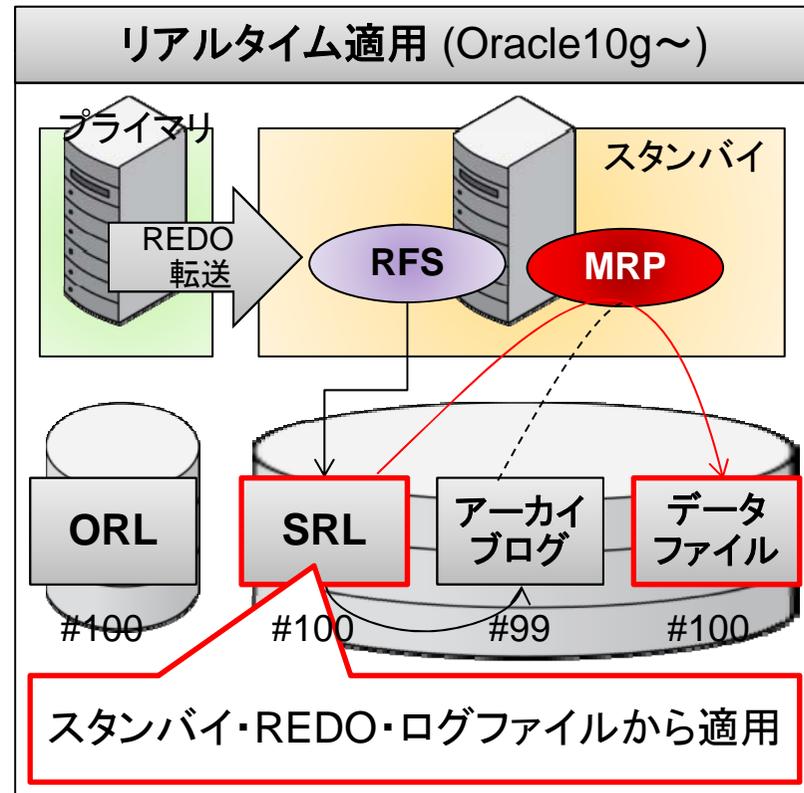
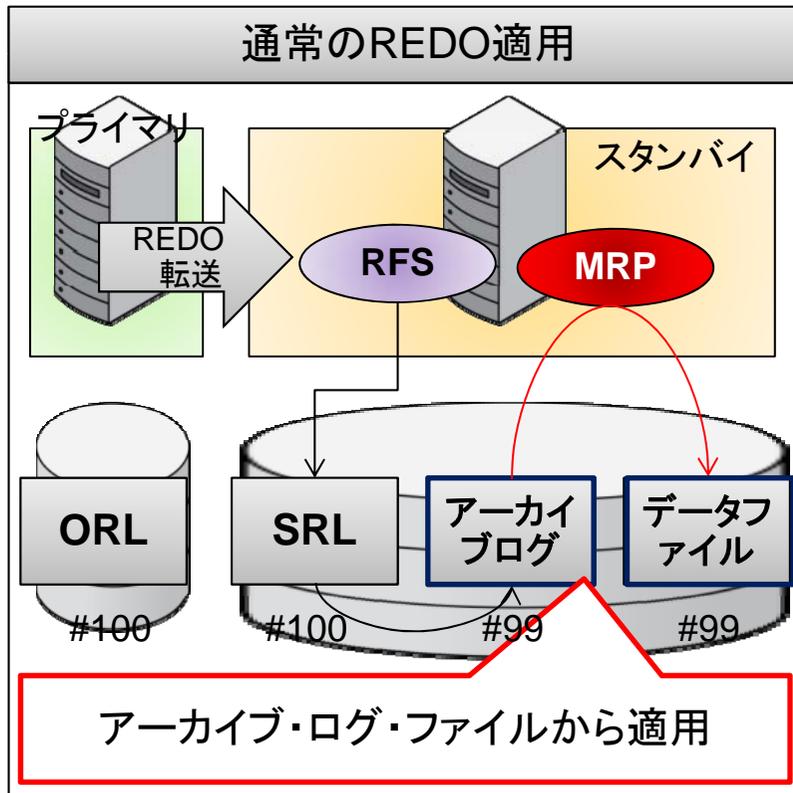
プライマリとスタンバイの役割を切り替えて停止時間を削減



# リアルタイム適用

受信後すぐに適用することで、常に最新状態までリカバリ済み

- スイッチオーバー/フェイルオーバーの高速化
- スタンバイ側で最新の情報を検索することが可能 (Active Data Guard Optionが必要)



# フェイルオーバー後のスタンバイ再作成

## 迅速に耐障害性を確保

スタンバイ・データベースへフェイルオーバー後に、旧・本番データベースを新たにスタンバイとして追加して、**迅速に耐障害性を向上させる**

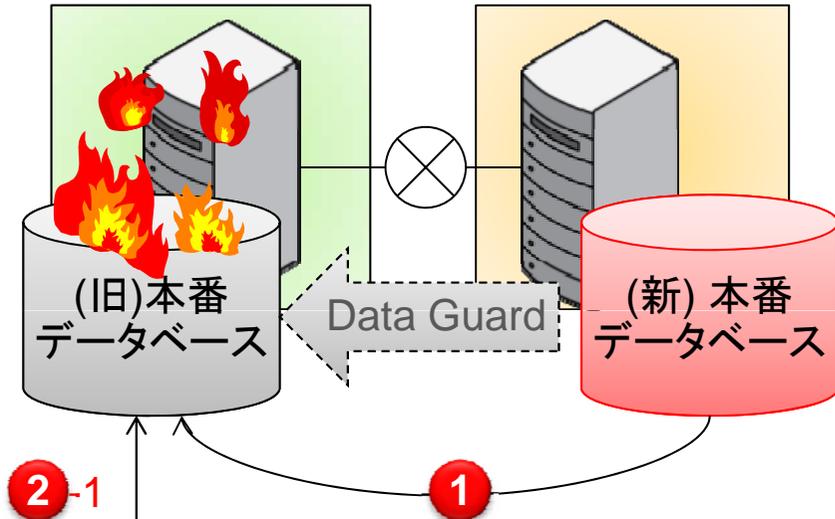
- 1 新・本番データベースから再作成する
- 2 旧・本番データベースの過去の状態から再作成する

### 2-1 フラッシュバック・データベースを利用

DBサーバーの物理的な障害(CPUやメモリ、ネットワークの障害など)だった場合は、旧本番DBでフラッシュバックすることによって新スタンバイとして構築

### 2-2 過去のバックアップを利用

ストレージの障害だった場合は、過去のバックアップを利用して新スタンバイとして構築 ()

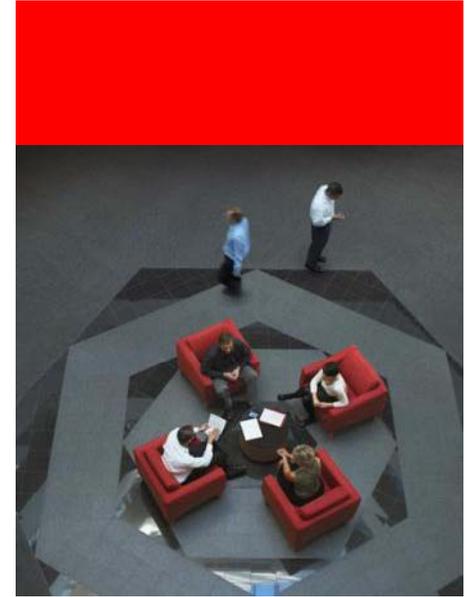
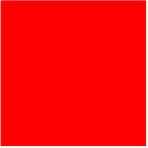


### 1 新・本番データベースから再作成

新・本番データベースで新たに取得したフルバックアップを旧・本番環境に転送し、新スタンバイとして構築

### 検討のポイント

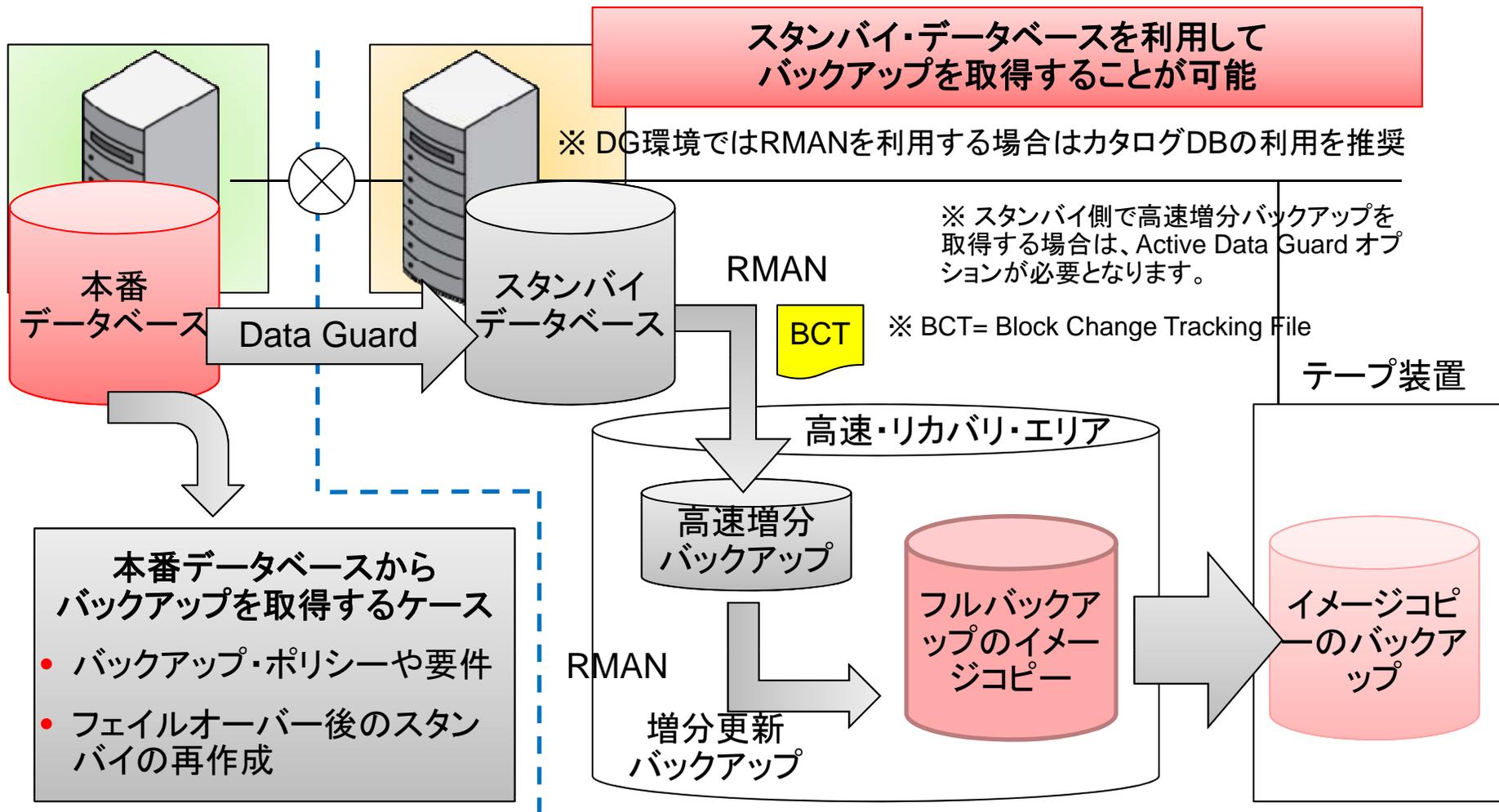
- ネットワーク帯域への影響
- 転送時間



# スタンバイ有効活用のポイント

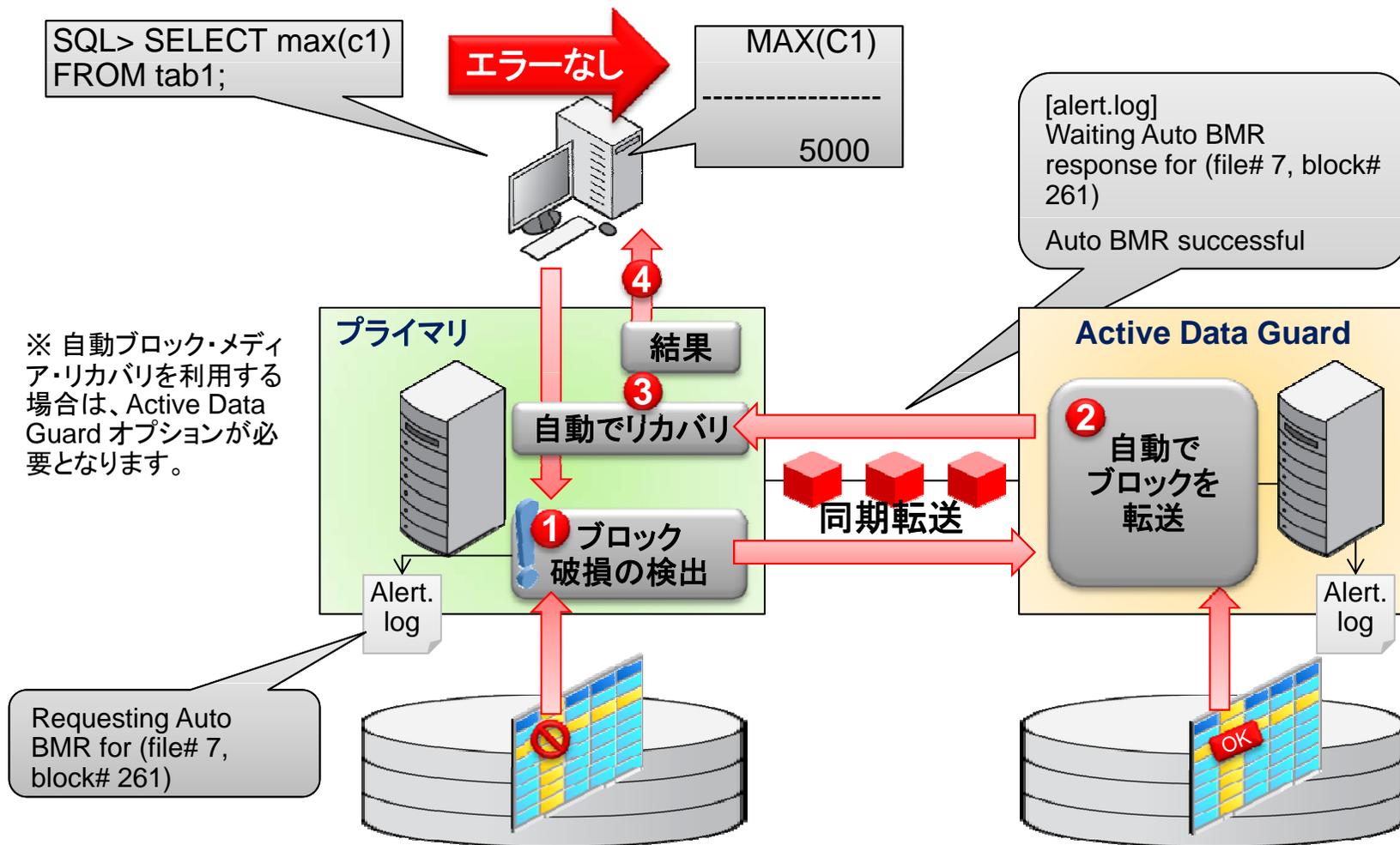
# スタンバイからのバックアップ運用イメージ

## プライマリ・データベースの負荷を軽減



# 自動ブロック・メディア・リカバリ

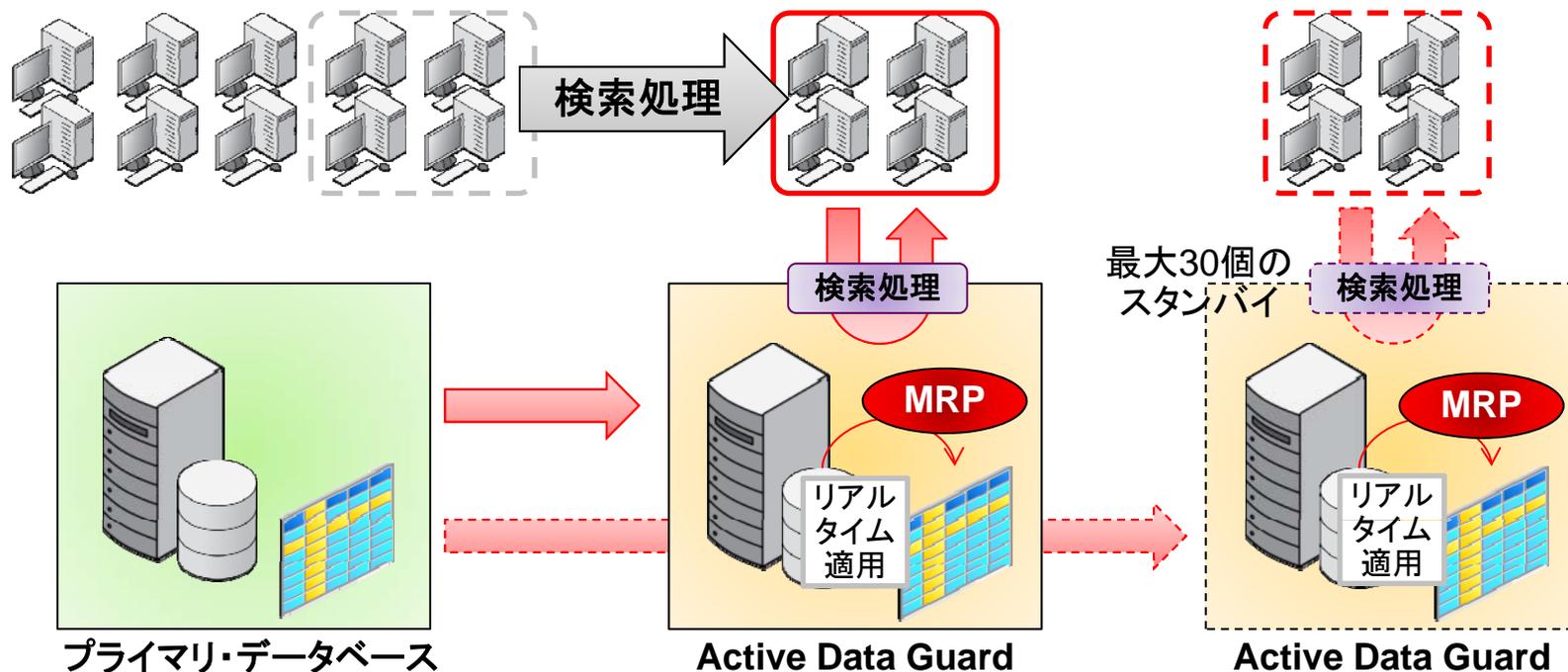
## Active Data Guardによる透過的な自動修復機能

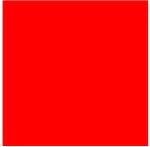


# 検索処理のオフロード(リアルタイム問合せ)

災害対策機能を維持して、最新の結果をスタンバイ側で検索

- スタンバイ側での検索は読み取り一貫性を保証
- 1つのプライマリ・データベースに対して、最大30個のスタンバイを作成可能
- 検索可能な許容可能な適用ラグ(時間)を指定することでデータの品質を保証 (同期転送)





# リアルタイム問合せを利用する際のポイント

- DBLINKを使って更新処理をプライマリにオフロード
- 利用できない処理
  - 順序/一時表/ストアド・プロシージャなど
- 組み合わせて利用する際のベストプラクティス
  - BIEE
  - TopLink
  - EBS Reporting
- リアルタイム問合せのパフォーマンス・チューニング
  - STATSPACKを利用する方法 → Note#454848.1
  - リアルタイムSQL監視 (11.2.0.2)



詳細は

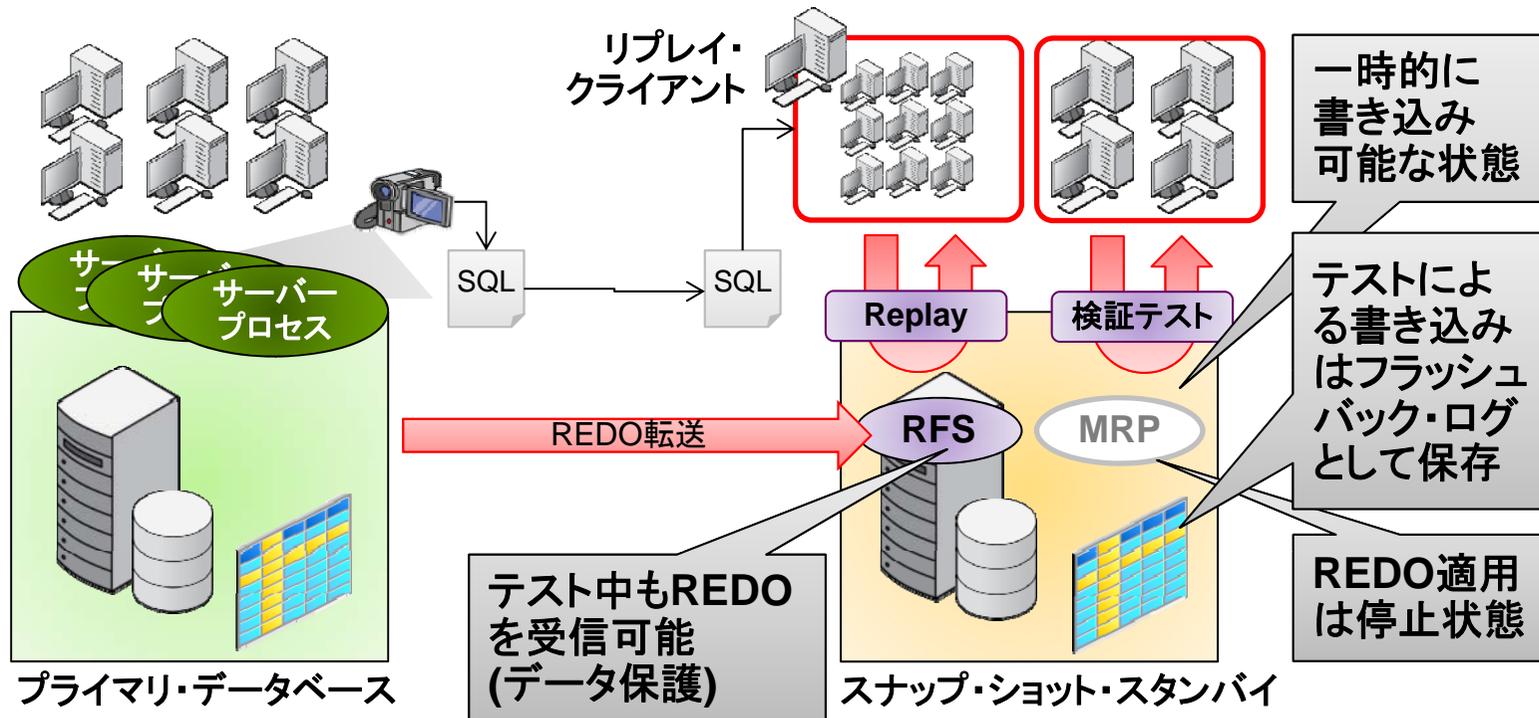
MAAのホワイトペーパーを参照のこと

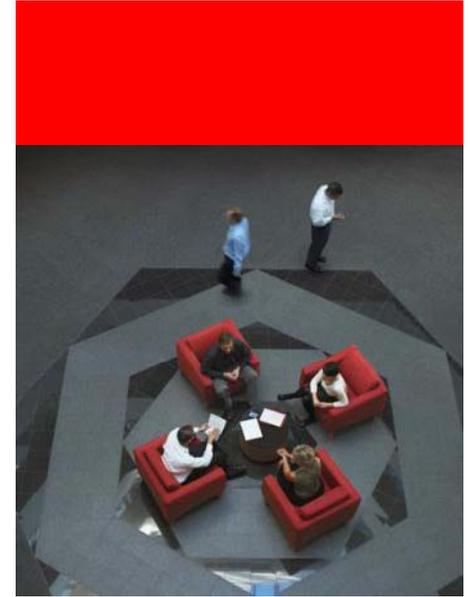
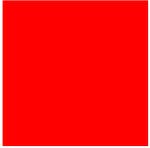
<http://www.oracle.com/technetwork/database/features/availability/maa-wp-11gr1-activedataguard-1-128199.pdf>

# スナップショット・スタンバイ

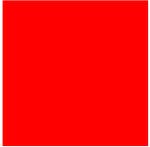
## スタンバイを一時的にテスト・データベースとして活用

- フラッシュバック・データベースの機能を利用 → テスト終了後は再びスタンバイへ
- スナップショット・スタンバイ中も、REDOを受信可能なためデータ保護要件は確保
- RAT(Real Application Testing)を合わせて利用することで、迅速に検証環境を提供





最後に



## 本日、お話できなかったポイント

- Network転送のチューニング
  - SDUのサイズの計算式
  - RECV_BUF_SIZE / SEND_BUF_SIZE
  - TCP.NODELAY = YES
- log_bufferは8MB以上
- パラレル・アーカイブ転送のプロセス数の計算式
- SRLの計算式
- など



詳細は

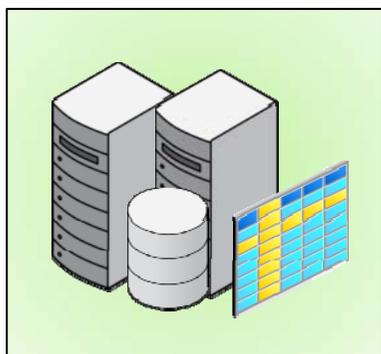
### High Availability Best Practice (8章)と

<http://www.oracle.com/technetwork/database/focus-areas/availability/ha-best-practices-112-452463.pdf>

# 本日、ご紹介した有償オプション

## Real Application Clusters Option

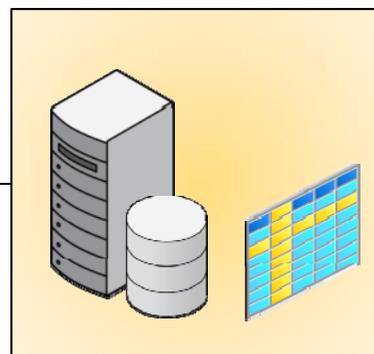
- サーバー障害に対する可用性の向上
- 計画停止時間の削減



プライマリ・データベース

## Active Data Guard Option

- リアルタイム問合せ
- 自動メディアブロック・リカバリ
- スタンバイ側での高速増分バックアップ



スタンバイ・データベース

## Advanced Compression Option

- REDO転送の圧縮

## Real Application Testing Option

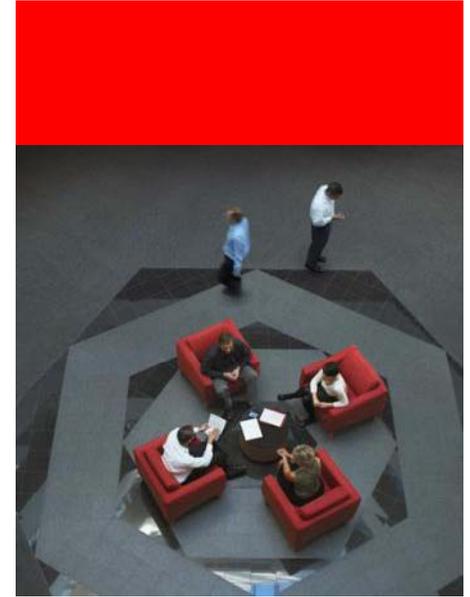
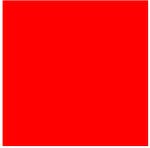
- 本番ワークロードの実行

## Advanced Security Option

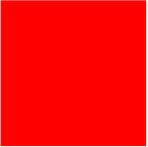
- REDO転送の暗号化

## EM Diagnostic Pack/Tuning Pack

- 可用性に関する監視と性能チューニング



# まとめ



# まとめ

- 高度なデータ保護要件への対応
  - 本番データベースの計画停止および計画外停止時間を削減データベースへの変更を同期もしくは非同期で低いオーバーヘッドで送信
  - スタンバイ側で受信した変更は、即時に適用し常に最新状態にリカバリ済み
- 低コストで災害対策サイトの構築
  - ストレージミラーに比べて、変更ログのみを転送するため低帯域のネットワークで構築することで**通信コストを抑える**ことが可能
  - ログ転送/ログ適用/ロールの推移などの包括的な機能を提供することで**管理コストの削減**が可能
  - スタンバイサイトを有効に利用することで**遊休リソースの維持コストを削減**することが可能



# **Hardware and Software Engineered to Work Together**

ORACLE

ORACLE®