

# 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®

### 実践!!高可用性システム構築~RAC基本編~

日本オラクル株式会社 基盤技術部  
エンジニア 王小芬

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。

# Agenda

- データベース・サーバーに求められるもの
- RACのアーキテクチャ
- RACのスケールビリティ
- RACの高可用性
- まとめ



# データベース・サーバーに求められるもの

- Active-Standby構成
- Oracle Real Application Clustersとは
- RACの特長

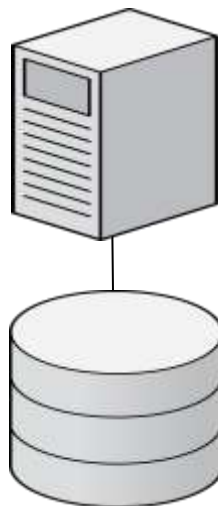
# データベース・サーバーに求められるもの

## データベース・サーバーに求められるものとは??

サーバー障害時の  
ダウンタイムを短くしたい

可用性

サーバー障害時も  
トランザクションの一貫性  
を維持したい



処理できるアクセス・  
ユーザー数を増やしたい

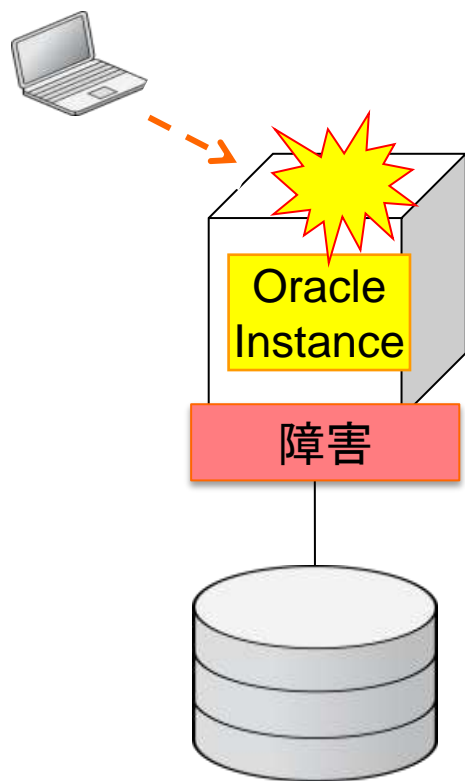
スケーラビリティ

大量データ検索実行の  
並列度を上げたい

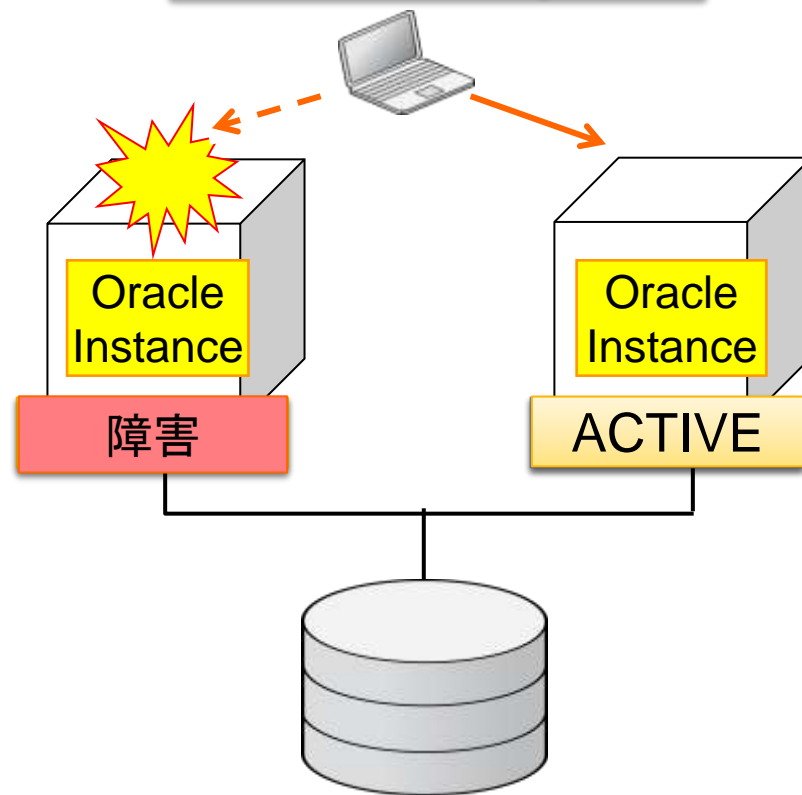
# Active-Standby構成の可用性

- Active-Standby構成は、障害時にstandby機に切りかえるしくみ
- シングル・インスタンスに比べ、高可用性を実現

シングル・インスタンス

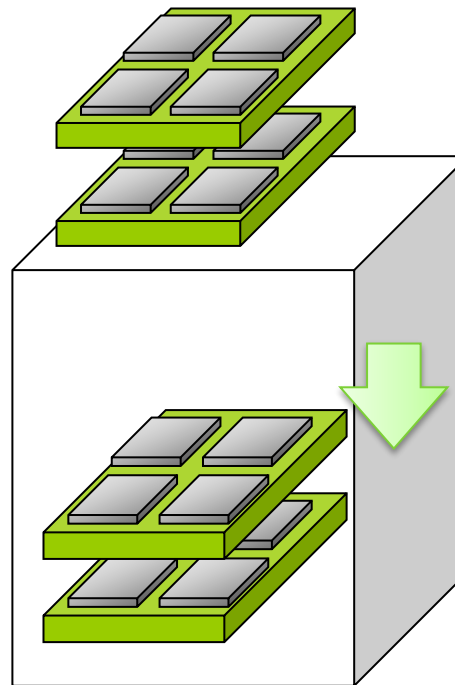


Active-Standby構成



# Active-Standby構成のスケラビリティ

- Active-Standby構成は、スケールアップによる性能向上
  - 1サーバー内でのCPUコア増設

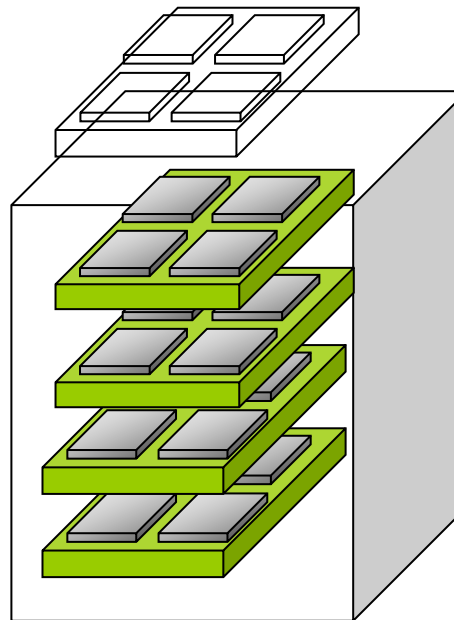


# Active-Standby構成の課題

- Active-Standby構成は、スケールアップによる性能向上
  - 1サーバー内でのCPUコア増設

スケールアップによる性能向上には課題がある

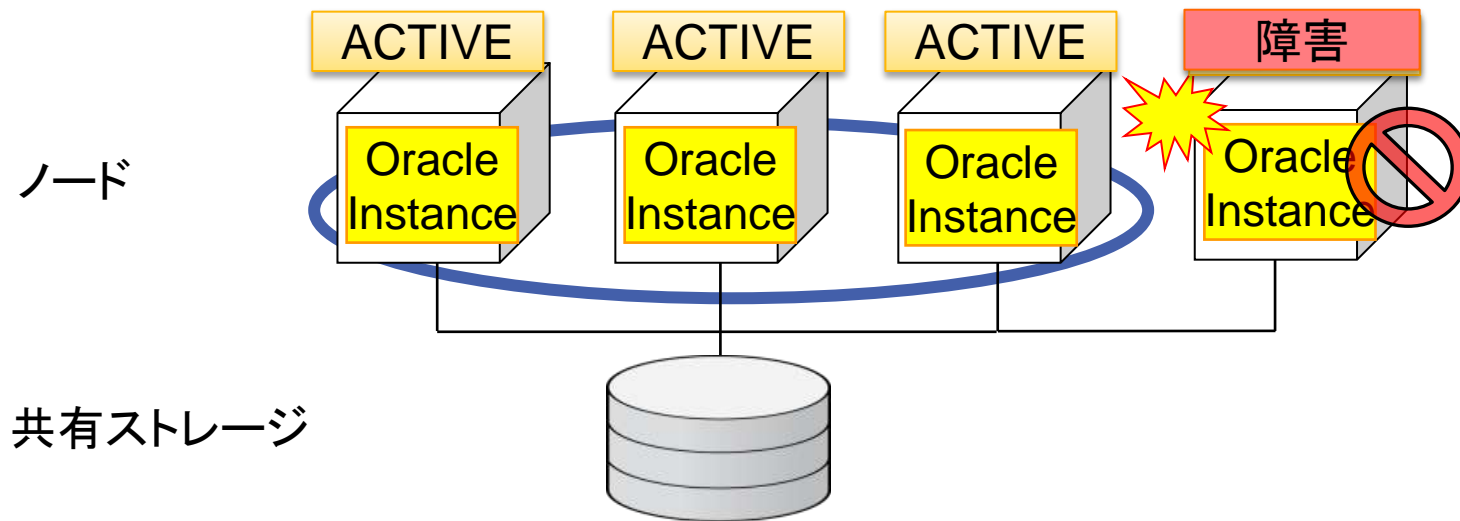
- 1サーバー内のCPUコア搭載数に上限あり
- 多くのCPUコアを搭載できる大型のマシンをあらかじめ用意する必要あり





# Oracle Real Application Clustersとは

- Oracle Real Application Clusters(RAC)は、全ノードがアクティブな、共有ディスク・共有キャッシュ型のクラスタ・データベース
- 全ノードがデータベースの全データにアクセス可能

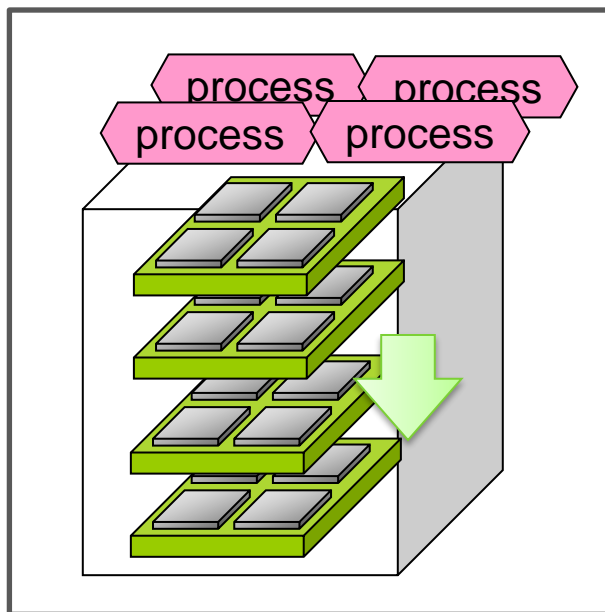


- 特長:
  - ノードの追加による性能向上
  - 障害ノードを切り離してシステムを継続

# 性能を向上させる2つのアプローチ

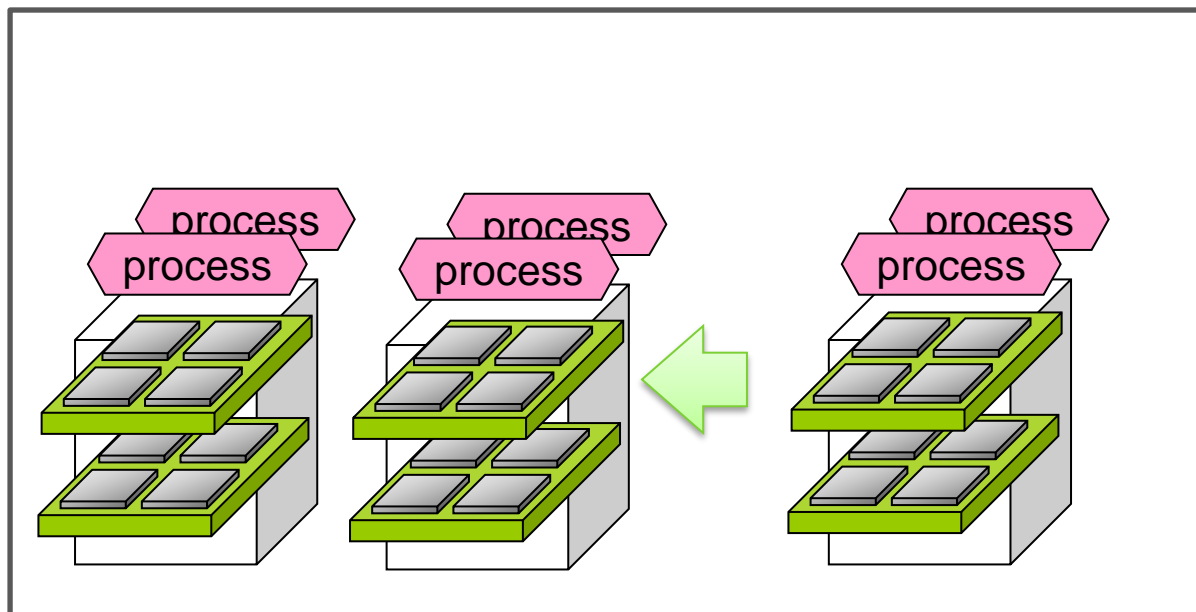
- サーバーの処理性能を高めるためには...
  - CPUコアを増設して同時実行性を向上
- CPUコアを増設するアプローチは2つ: スケールアップ / スケールアウト

## スケールアップ



サーバーにCPUを増設

## スケールアウト

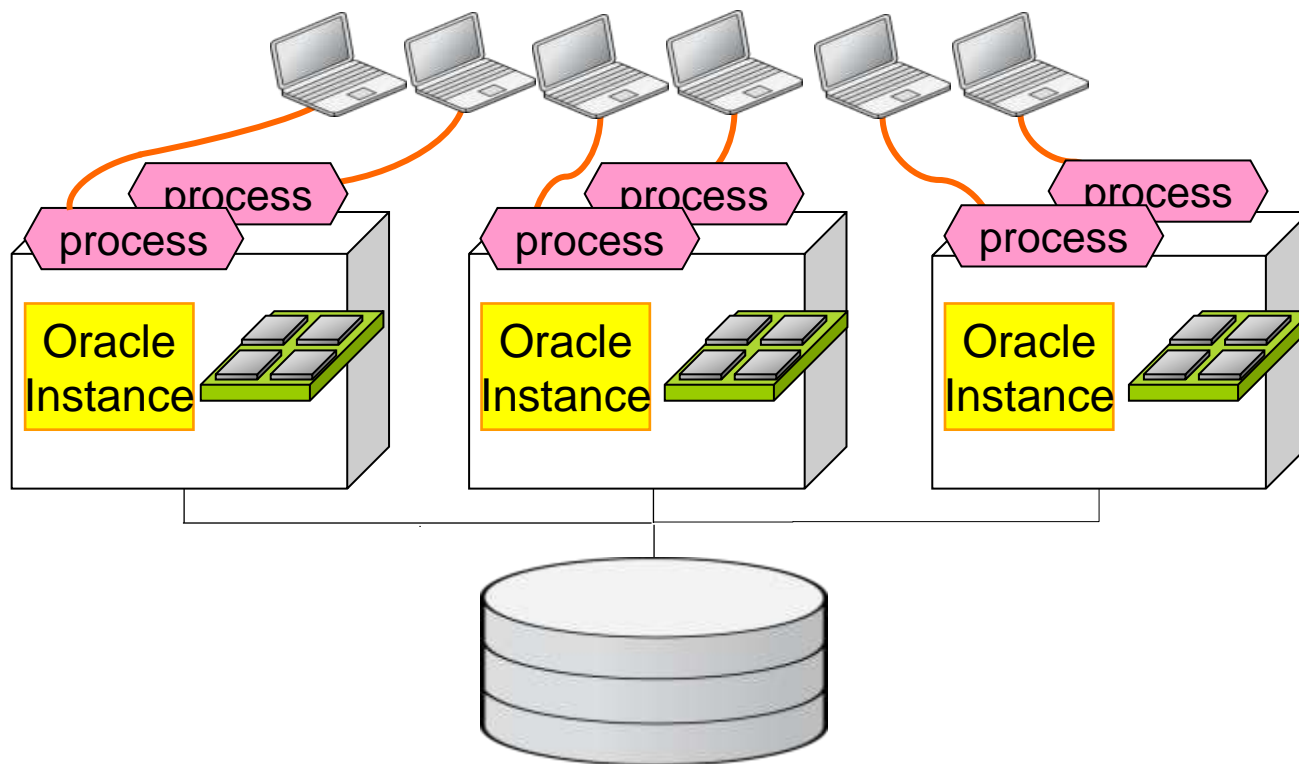


ノードを追加してCPUを増設

# RACのスケラビリティ

## RACの特長

- RACは、全ノードがデータベースの全データにアクセスでき、Active-Activeの構成をとることが可能
- RACは、ノード追加によるスケールアウトで性能向上

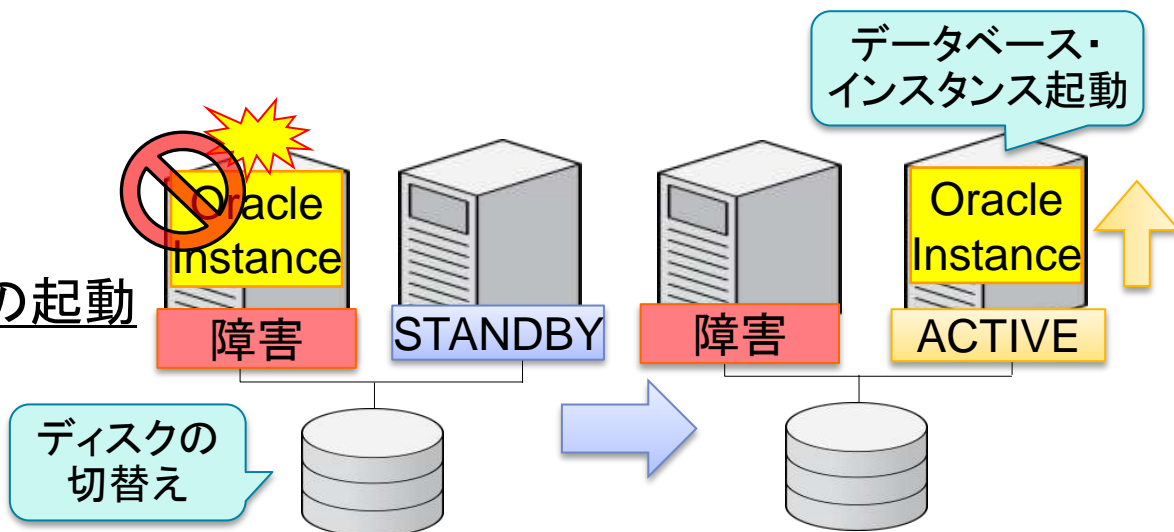


# RACの可用性

## RACの特長

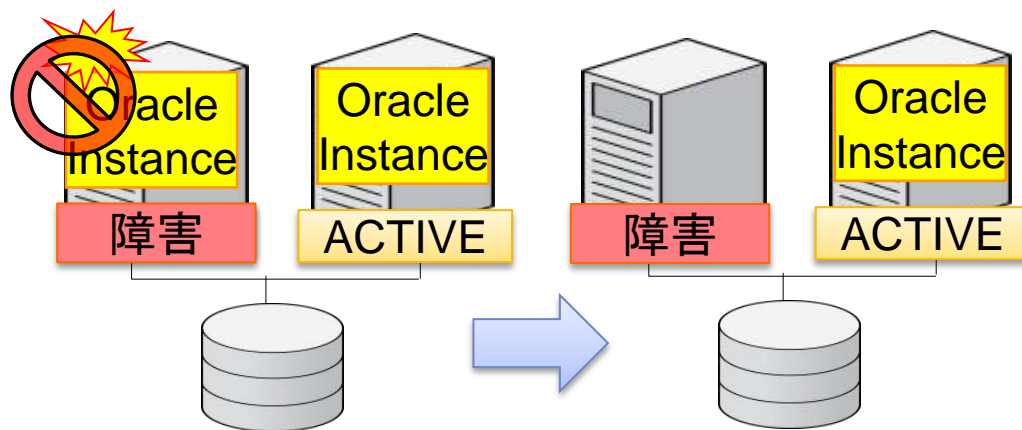
### Active-Standby構成

- 障害ノードの切り離し
- ディスク切替え
- データベース・インスタンスの起動
- リカバリ



### RAC

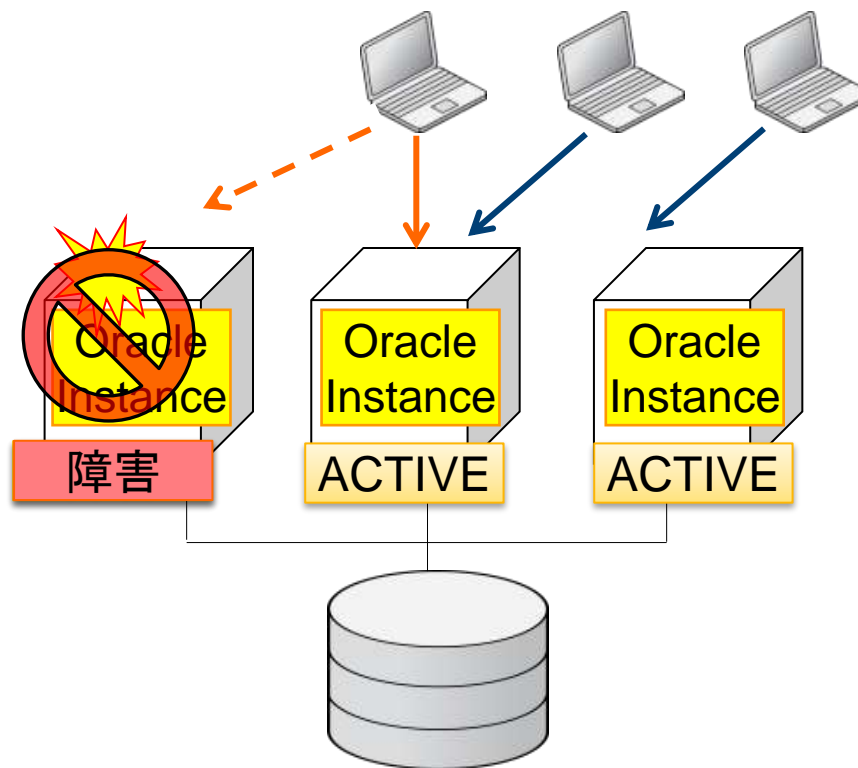
- 障害ノードの切り離し
- リカバリ



# RACの可用性

## RACの特長

- RACは障害ノードを自動的に切り離してシステムを継続
  - 全ノードが、データベースの全データにアクセスできるため、ディスク切替え、データベース・インスタンス再起動は不必要



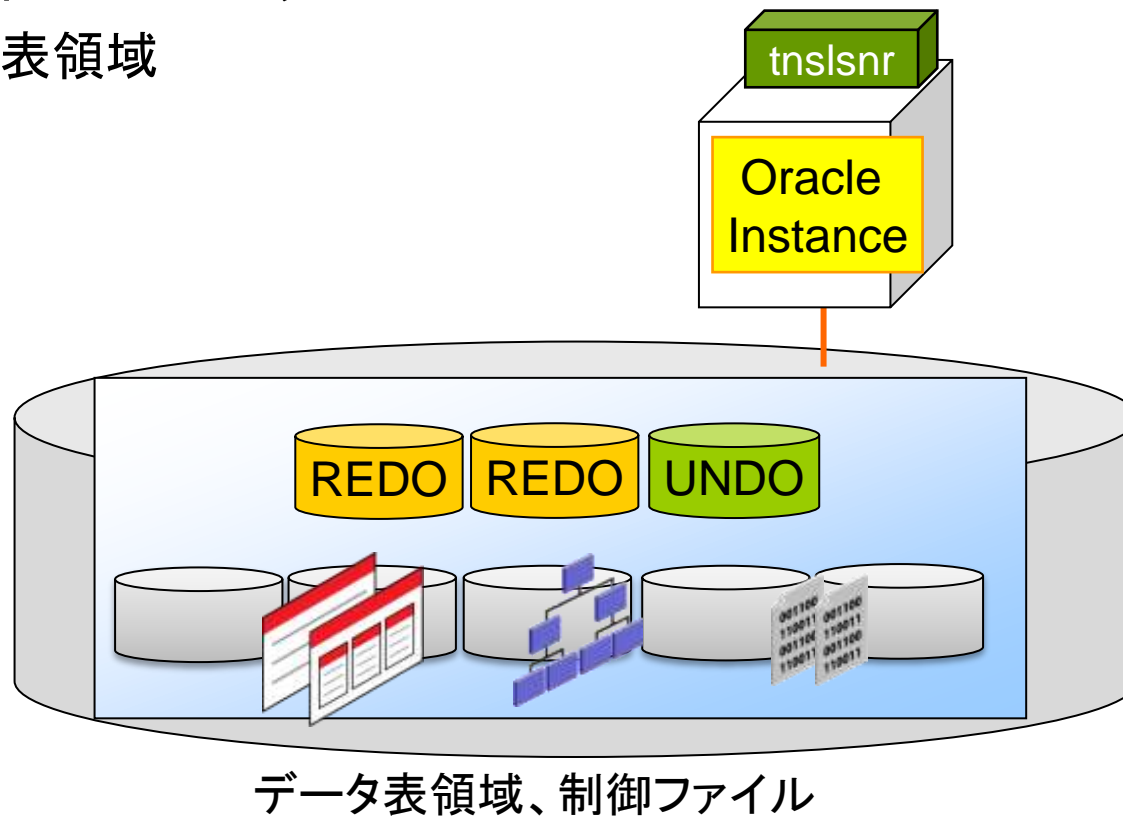


# RACのアーキテクチャ

- データベース構造
- ハードウェア構成
- ソフトウェア構成

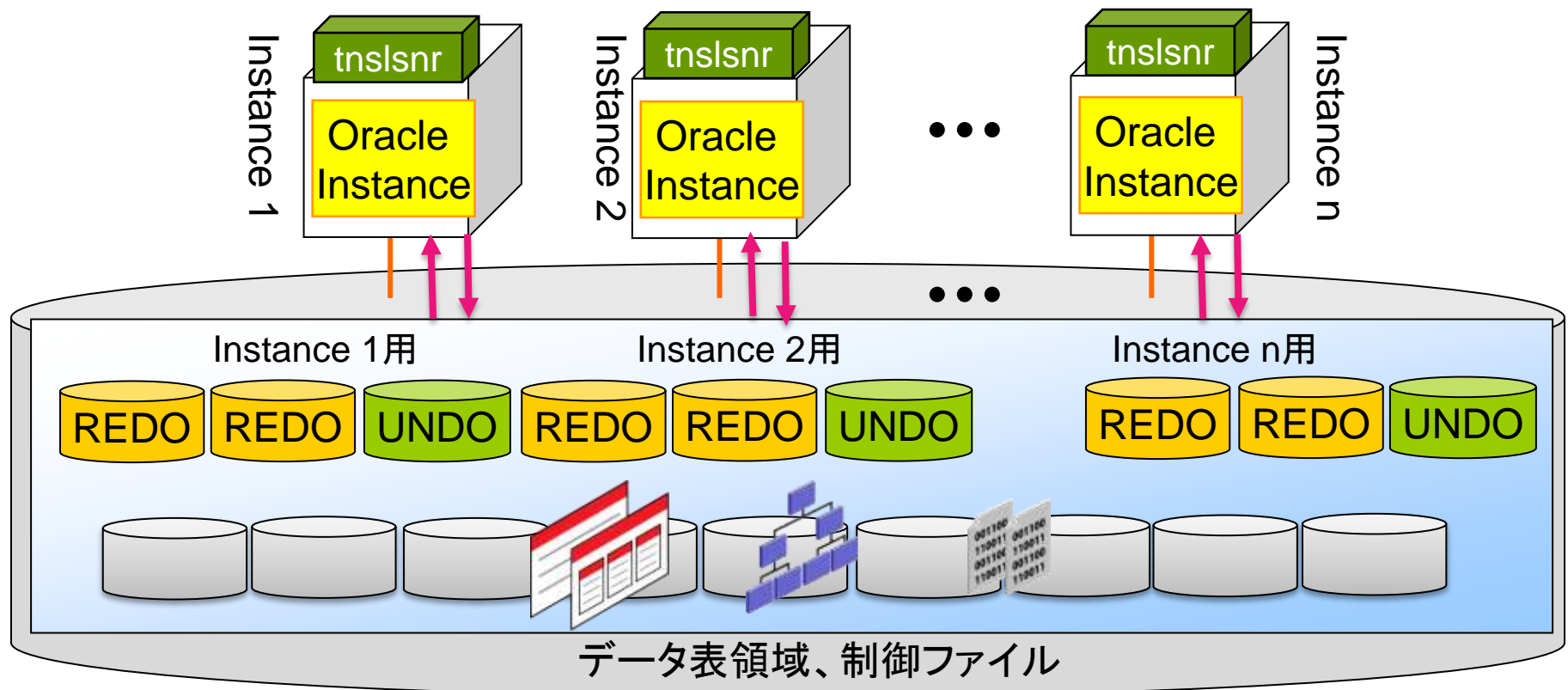
# シングル・インスタンスのデータベース構造

- シングル・インスタンスのデータベース構造
  - データ表領域
  - 制御ファイル
  - オンラインREDOログ
  - UNDO表領域



# RACのデータベース構造

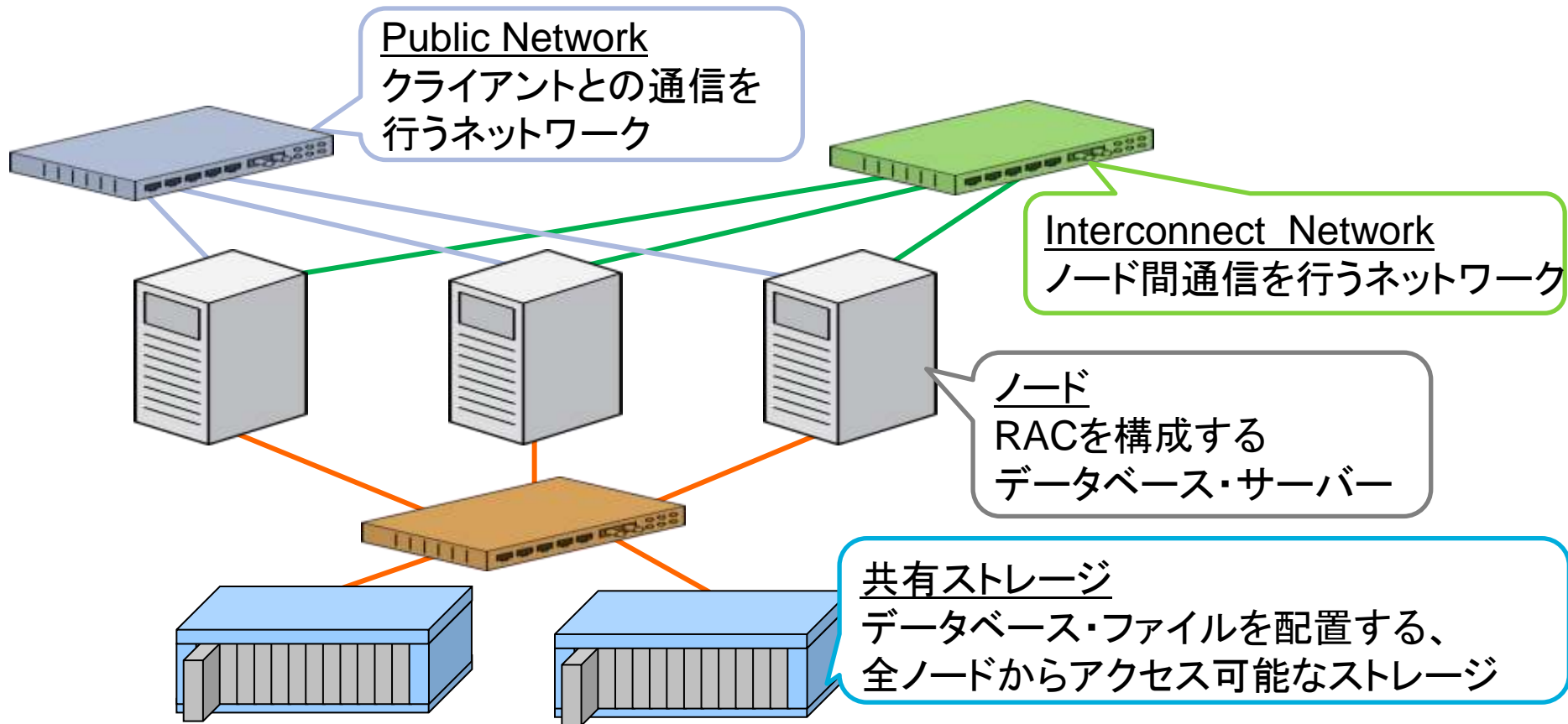
- RACのデータベース構造は、基本的にシングル・インスタンスと同様
  - 全ノードが共有ストレージ上の全データベース・ファイルをオープン
- Oracle インスタンスの分だけREDOログとUNDO表領域を追加





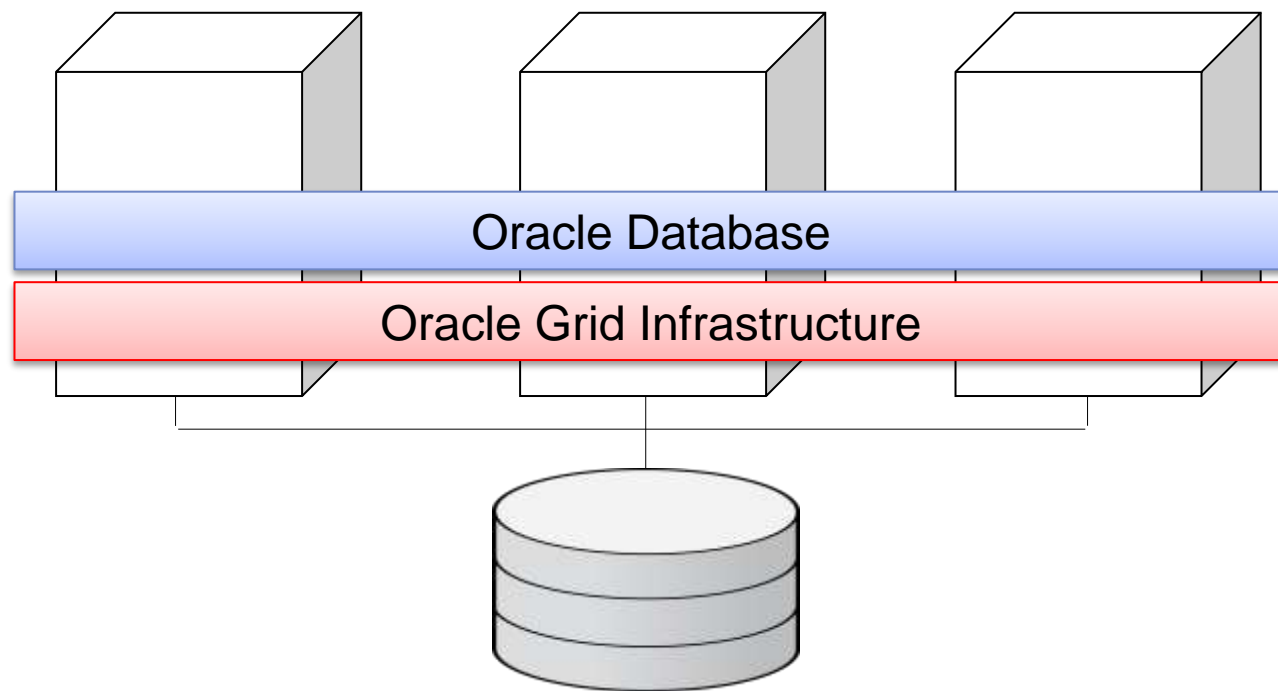
# RACのハードウェア構成

- RACを構築するために必要なハードウェア:
  - ノード、共有ストレージ、インターコネクト・ネットワーク



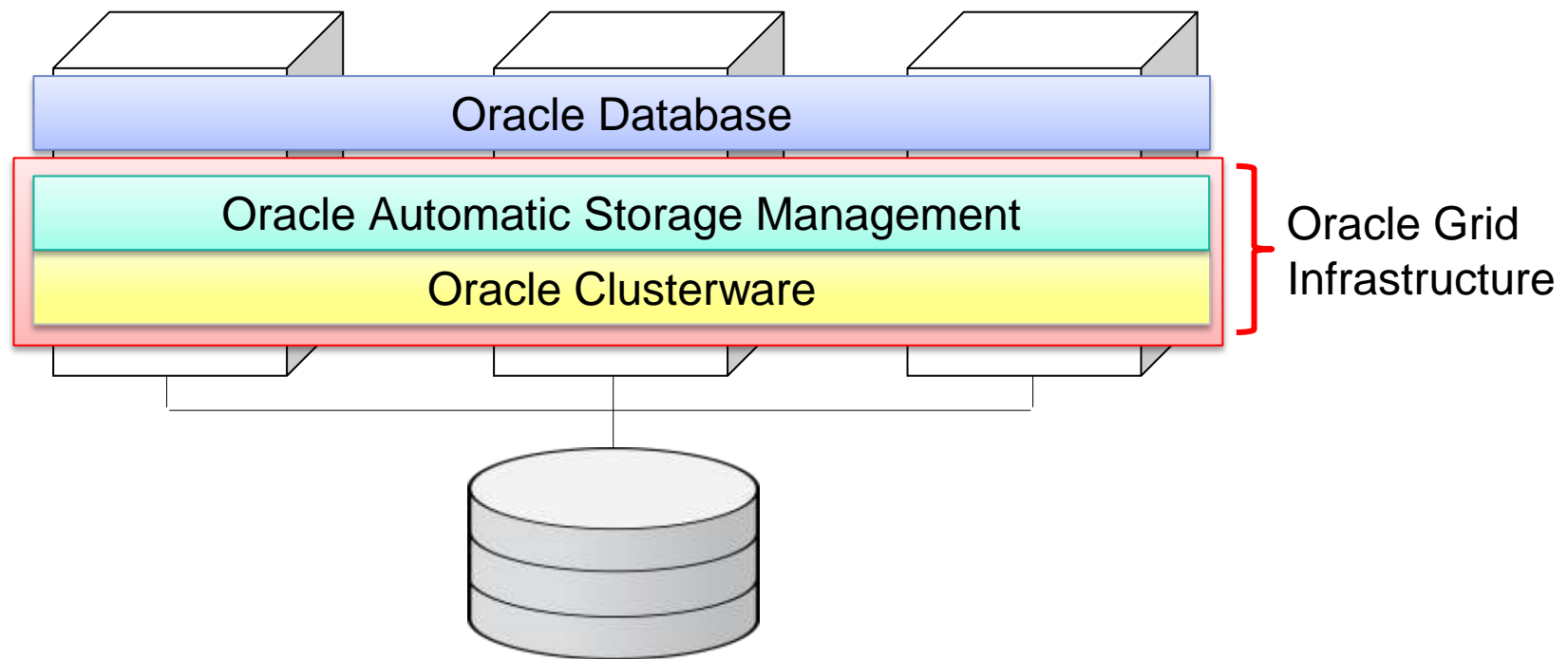
# RACのソフトウェア構成

- Oracle Database 11g Release 2 でRACを構築するために必要なソフトウェア：
  1. Oracle Grid Infrastructure
  2. Oracle Database



# RACのソフトウェア構成

- Oracle Grid Infrastructureは2つのソフトウェアを包含
  - Oracle Clusterware:  
クラスタ・メンバーシップとプロセスを管理するOracleのソフトウェア
  - Oracle Automatic Storage Management (ASM):  
ストレージを管理するOracleのソフトウェア



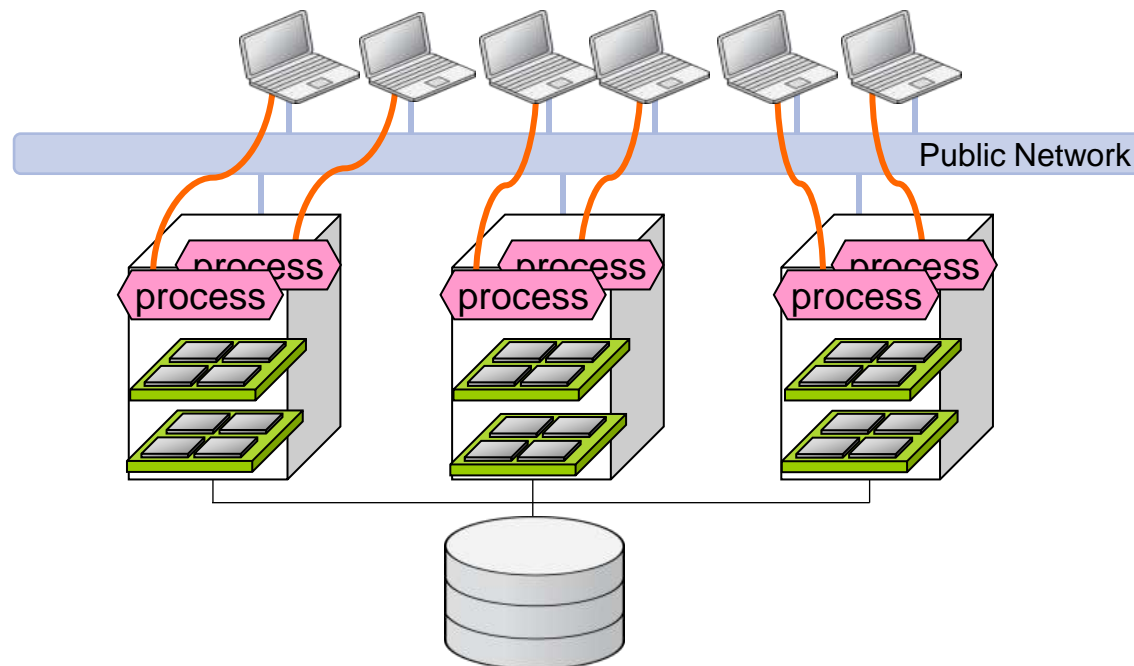


# RACのスケラビリティ

- RACがスケールアウトする理由
- RACのスケールアウトによる効果
- RACにおけるパラレル・クエリー

# RACのスケラビリティ

- RACがスケールアウトする理由
  - CPUコア増設による同時実行性の向上
  - シングル・インスタンスと同様のデータアクセス方式
    - シングル・インスタンスと同様の接続方式
    - どのノードに接続しても同じSQLを実行可能
    - メモリを基本とするデータアクセス方式

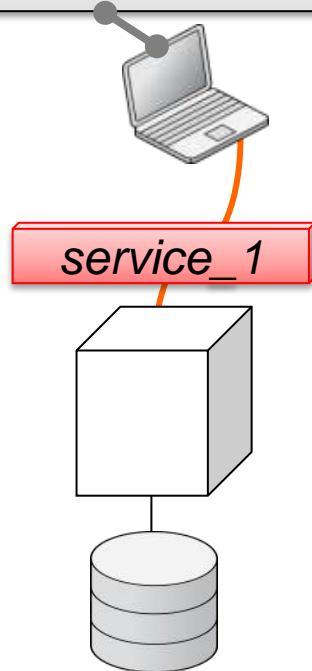


# シングル・インスタンスと同様の接続方式

- RACへの接続方法は、基本的にシングル・インスタンスと同様
- Oracleクライアントはサービス名を指定して接続
  - サービスは、RACのノードを仮想化

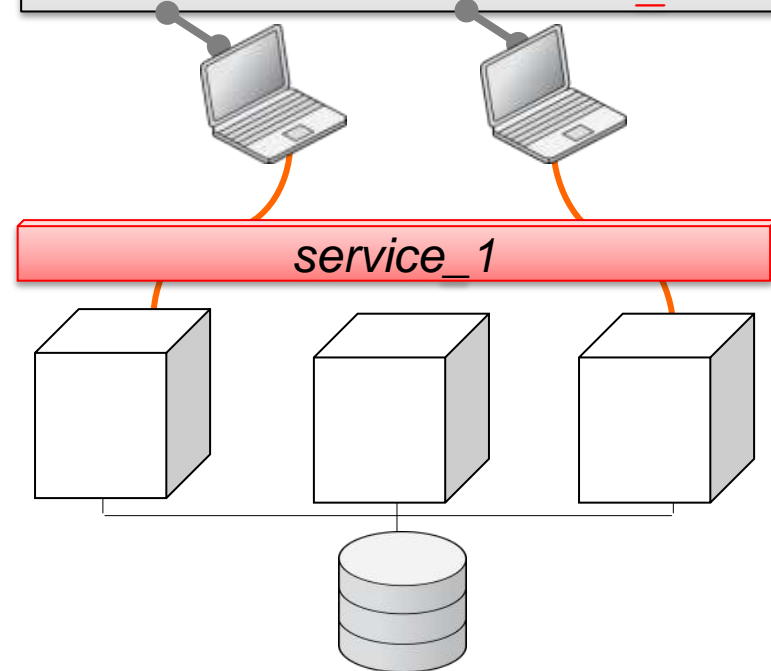
## シングル・インスタンス

```
$ sqlplus username/password@  
hostname:1521/service_1
```



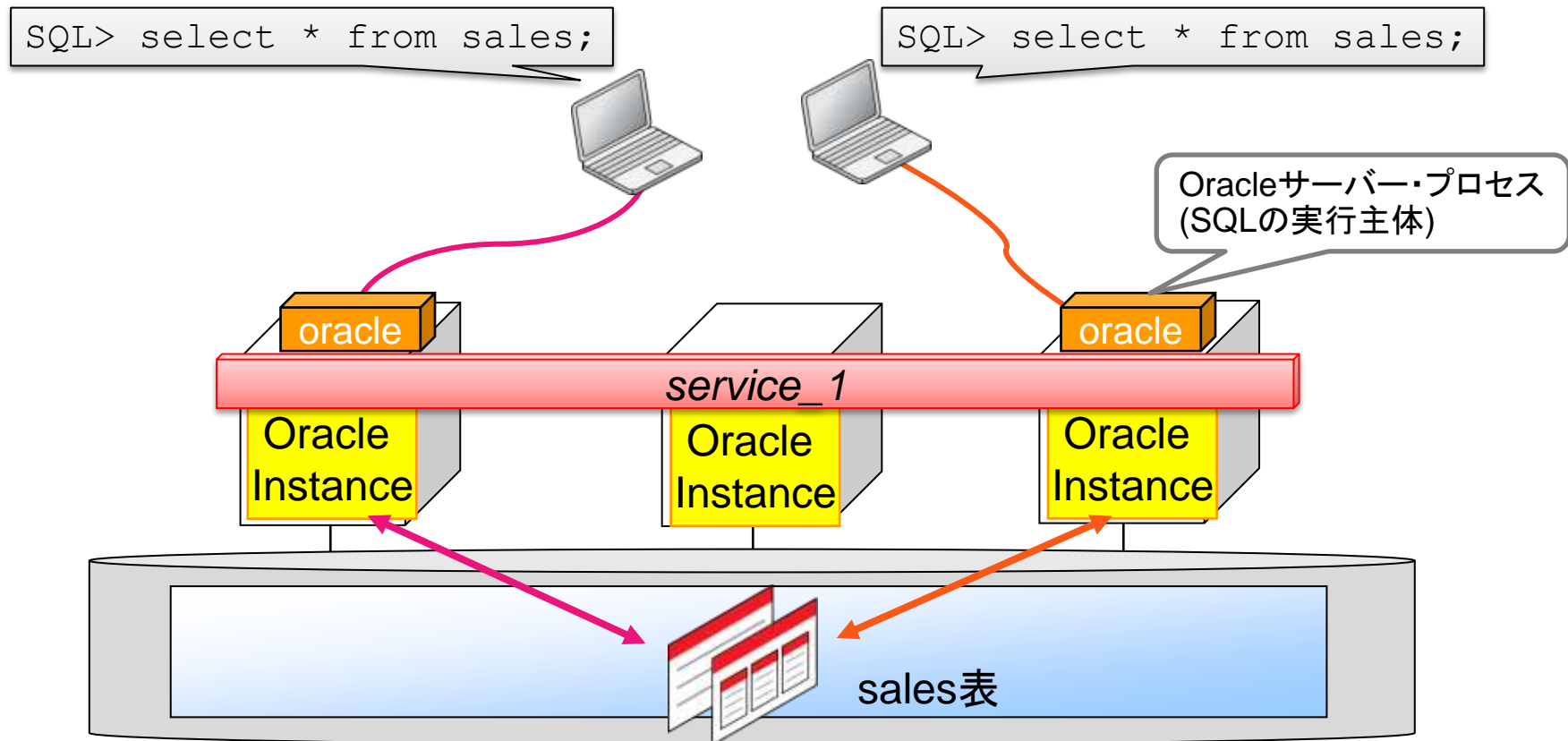
## RAC

```
$ sqlplus username/password@  
hostname:1521/service_1
```



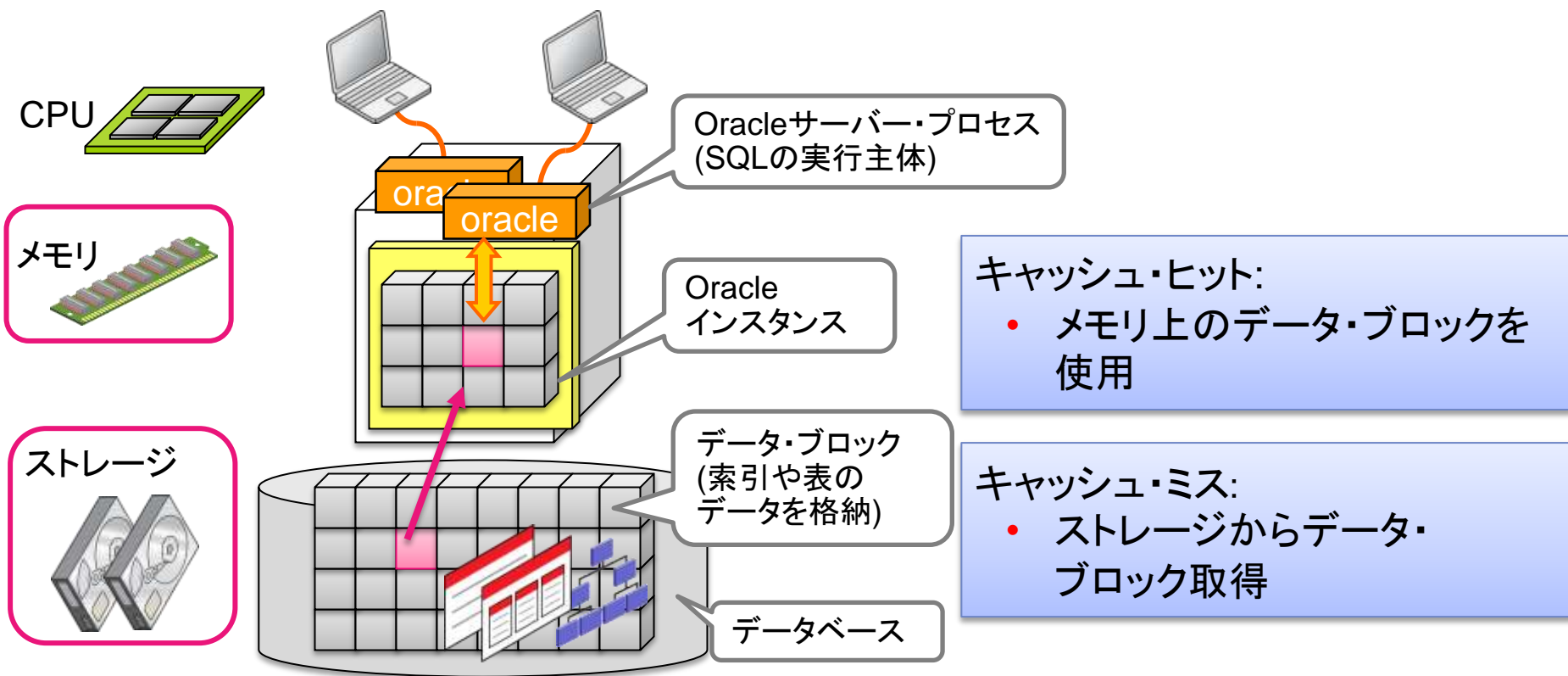
# どのノードに接続しても同じSQLを実行可能

- RACでは全ノードがデータベースの全データにアクセスできるため、どのノードに接続しても同じSQLを実行可能



# シングル・インスタンスのデータアクセス方式

- シングル・インスタンスもRACも、ストレージ (低速) に格納されたデータ・ブロックをメモリ (高速) にキャッシュして処理を行う
  - データ・ブロック: Oracle DatabaseのI/Oの最小単位、メモリの置換単位
  - バッファ・キャッシュ: 物理I/Oを最適化するためにデータ・ブロックのコピーを格納するメモリ領域

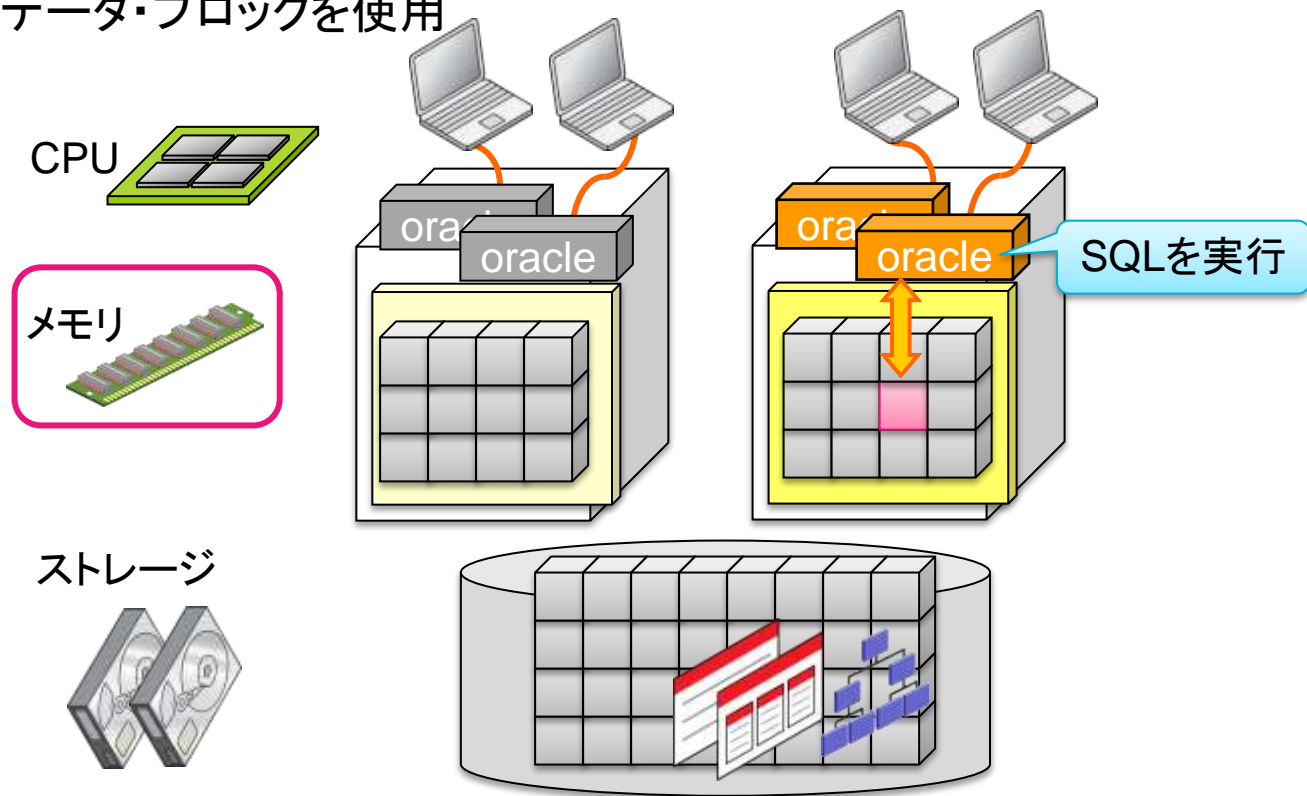




# RACのデータアクセス方式

## キャッシュ・ヒット

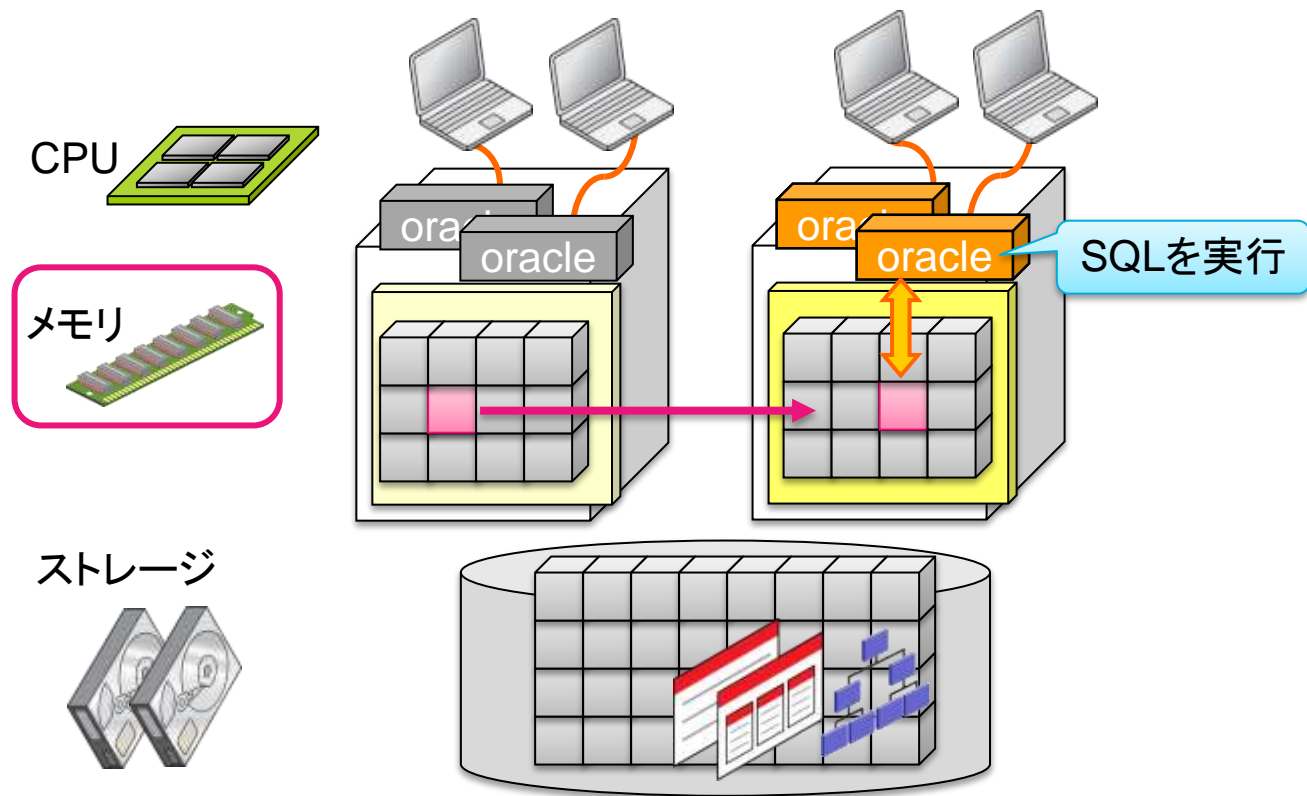
- RACもシングル・インスタンス同様、メモリ(高速)を基本とするデータアクセス方式
- 複数ノードのデータの一貫性を自動的に維持 → **Cache Fusion**
- キャッシュ・ヒット(ローカル・ノードにデータ・ブロックがキャッシュされている)
  - メモリ上のデータ・ブロックを使用



# RACのデータアクセス方式

## キャッシュ・ミス

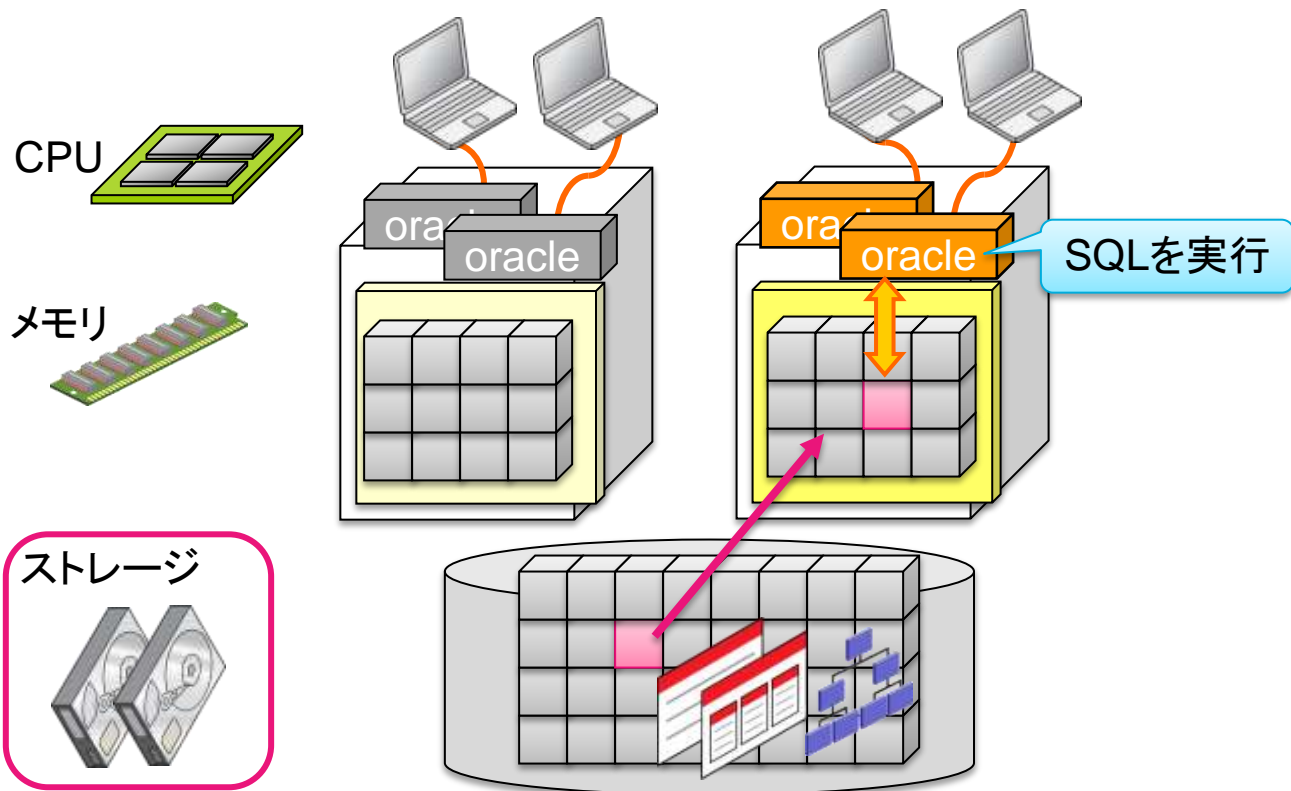
- キャッシュ・ミス(ローカル・ノードにデータ・ブロックがキャッシュされていない)
  - リモート・ノードのメモリ上からデータ・ブロックを転送
  - ストレージからデータ・ブロックを取得
- データの一貫性は自動的に維持



# RACのデータアクセス方式

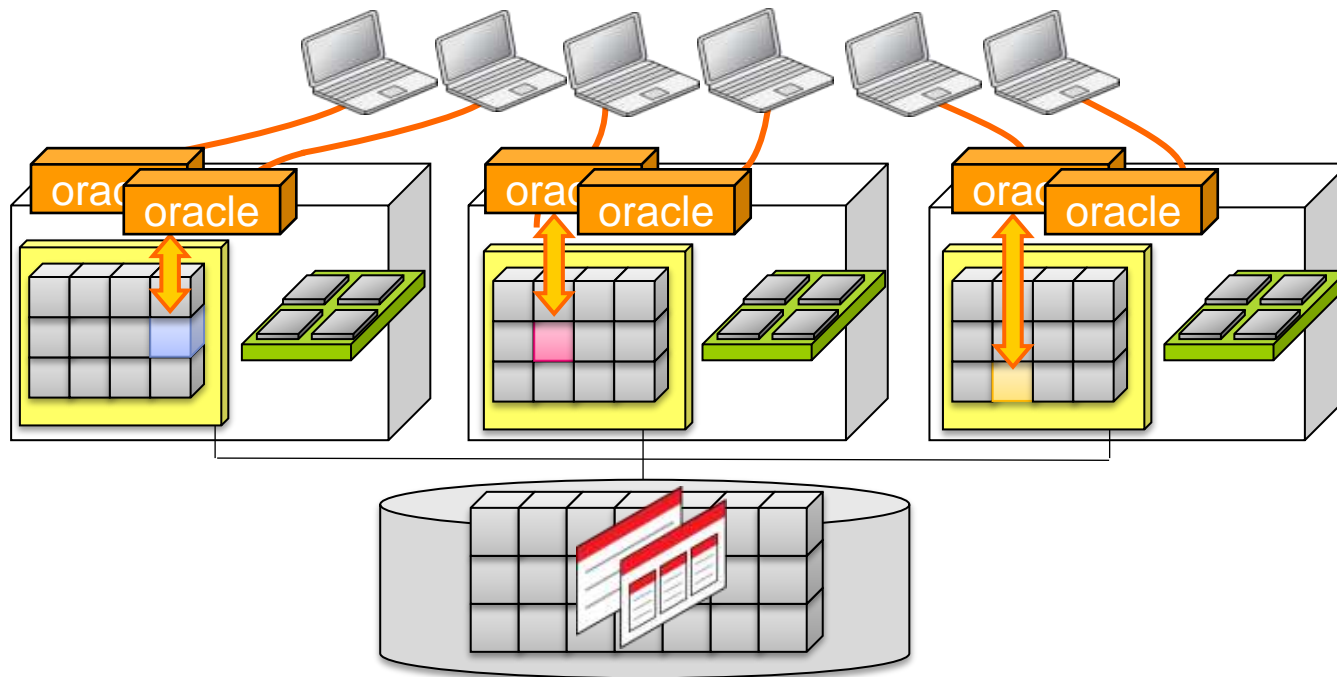
## キャッシュ・ミス

- キャッシュ・ミス(ローカル・ノードにデータ・ブロックがキャッシュされていない)
  - リモート・ノードのメモリ上からデータ・ブロックを転送
  - ストレージからデータ・ブロックを取得
- データの一貫性は自動的に維持



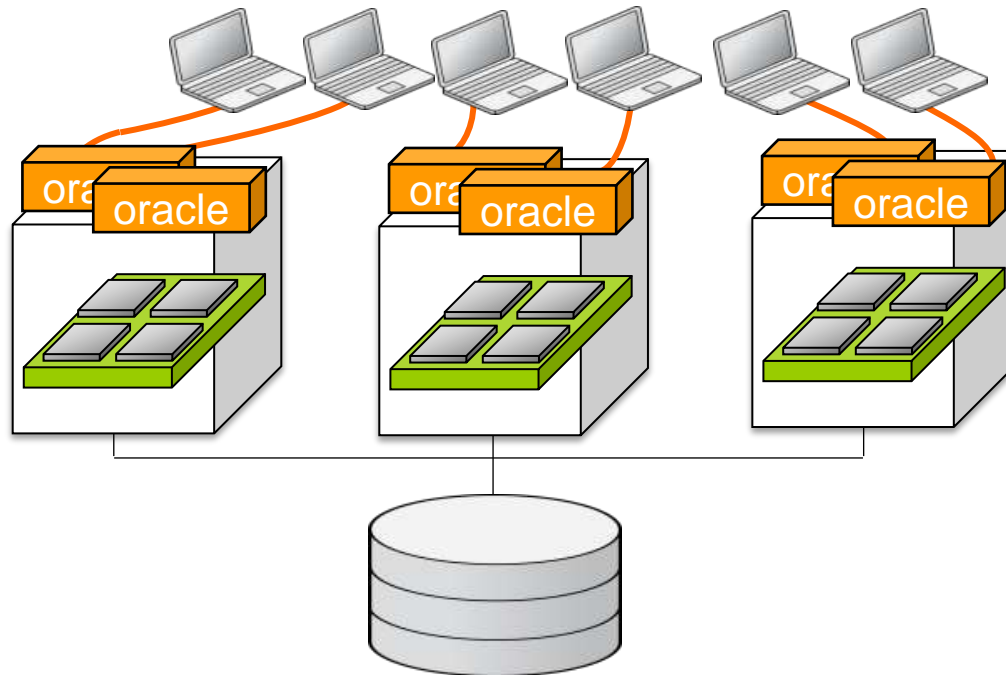
# RACはスケールアウトで性能向上

- CPUコア増設による同時実行性の向上
- シングル・インスタンスと同様のデータアクセス方式
  - シングル・インスタンスと同様の接続方式
  - どのノードに接続しても同じSQLを実行可能
  - メモリを基本とするデータアクセス



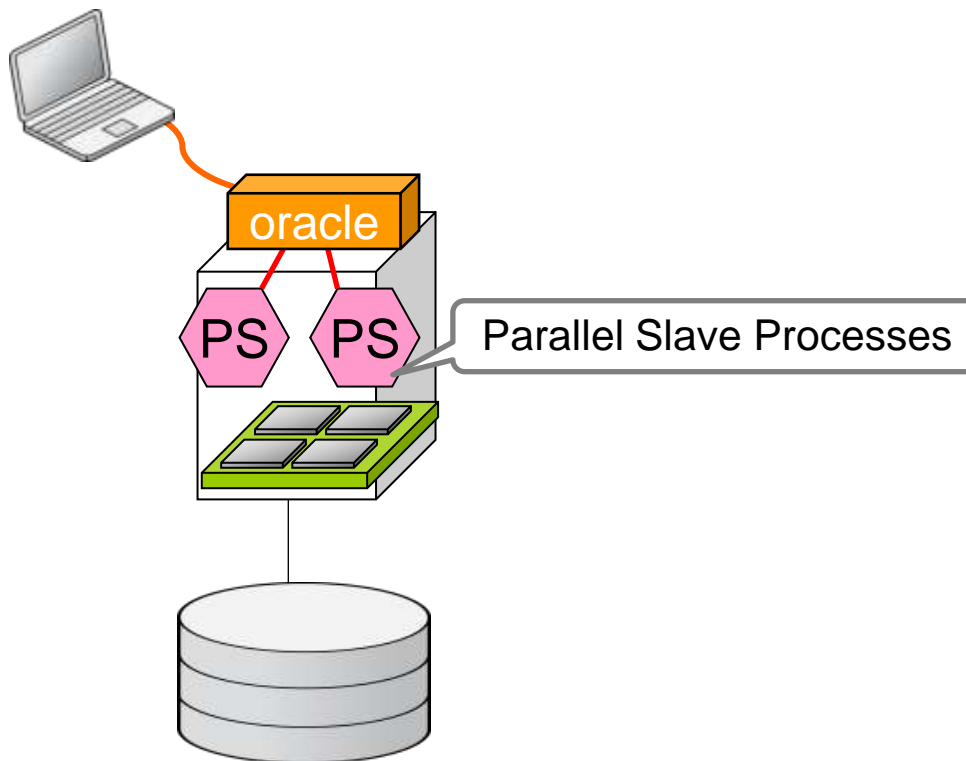
# RACのスケールアウトによる効果

- RACは、複数ノードのCPUを利用して同時実行性を向上
  - 処理できるアクセス・ユーザー数の増加



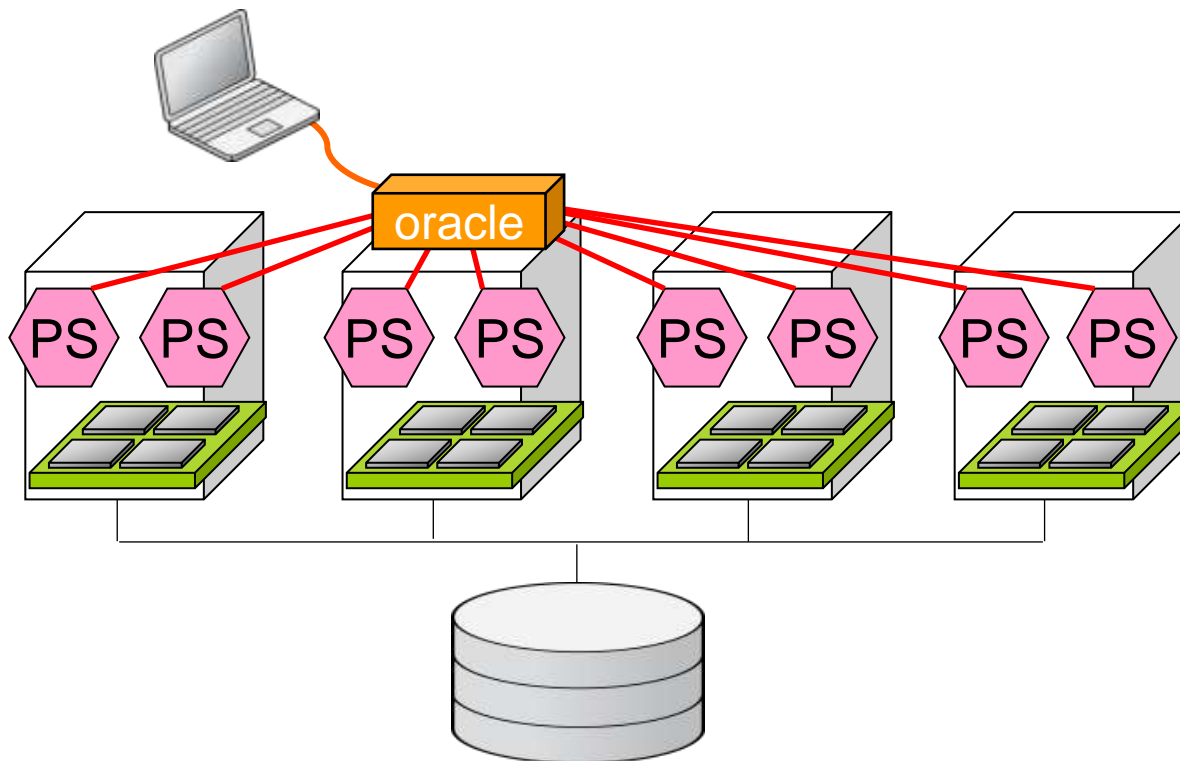
# RACのスケールアウトによる効果

- RACは、複数ノードのCPUを利用して同時実行性を向上
  - 処理できるアクセス・ユーザー数の増加
  - 1つのSQL実行における並列度の向上



# RACにおけるパラレル・クエリー

- RACは複数ノードのCPUを利用して、1つのSQL実行における並列度を上げることが可能
- 1つのSQLの実行時間を短縮することが可能



# RACのスケラビリティまとめ

- RACはスケールアウトで性能向上
  - CPUコア増設による同時実行性の向上
  - シングル・インスタンスと同様のデータアクセス方式
- 複数ノードのデータの一貫性を自動的に維持
  - Cache Fusion
  
- RACは優れたスケラビリティを実現
  - 処理できるアクセス・ユーザー数を増加
  - 1つのSQL実行における並列度の向上



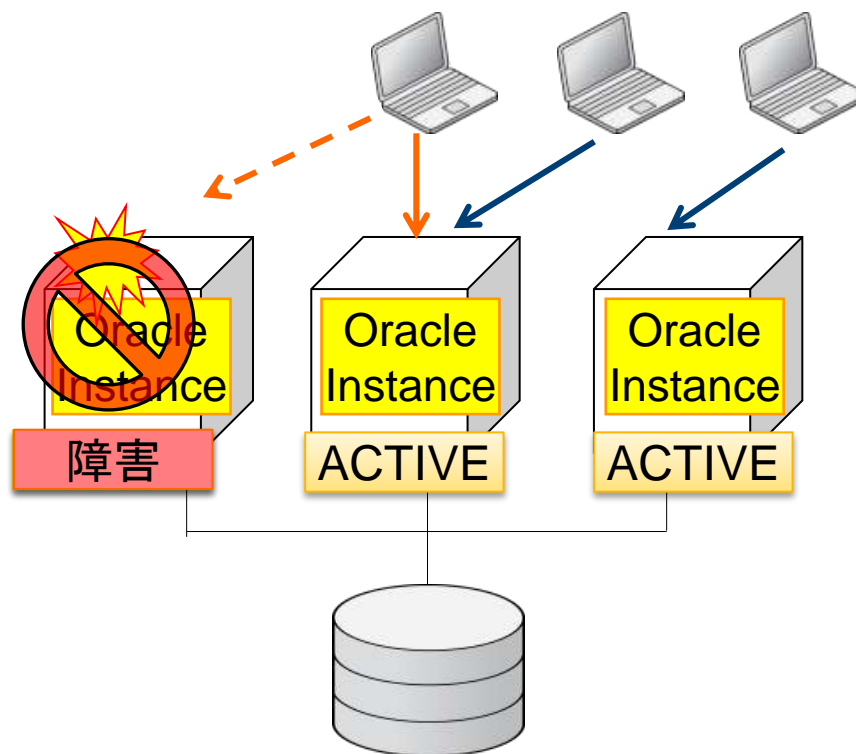


# RACの高可用性

- RACのインスタンス・リカバリのしくみ
- Oracle Clusterwareの役割
- RAC高可用性のしくみ

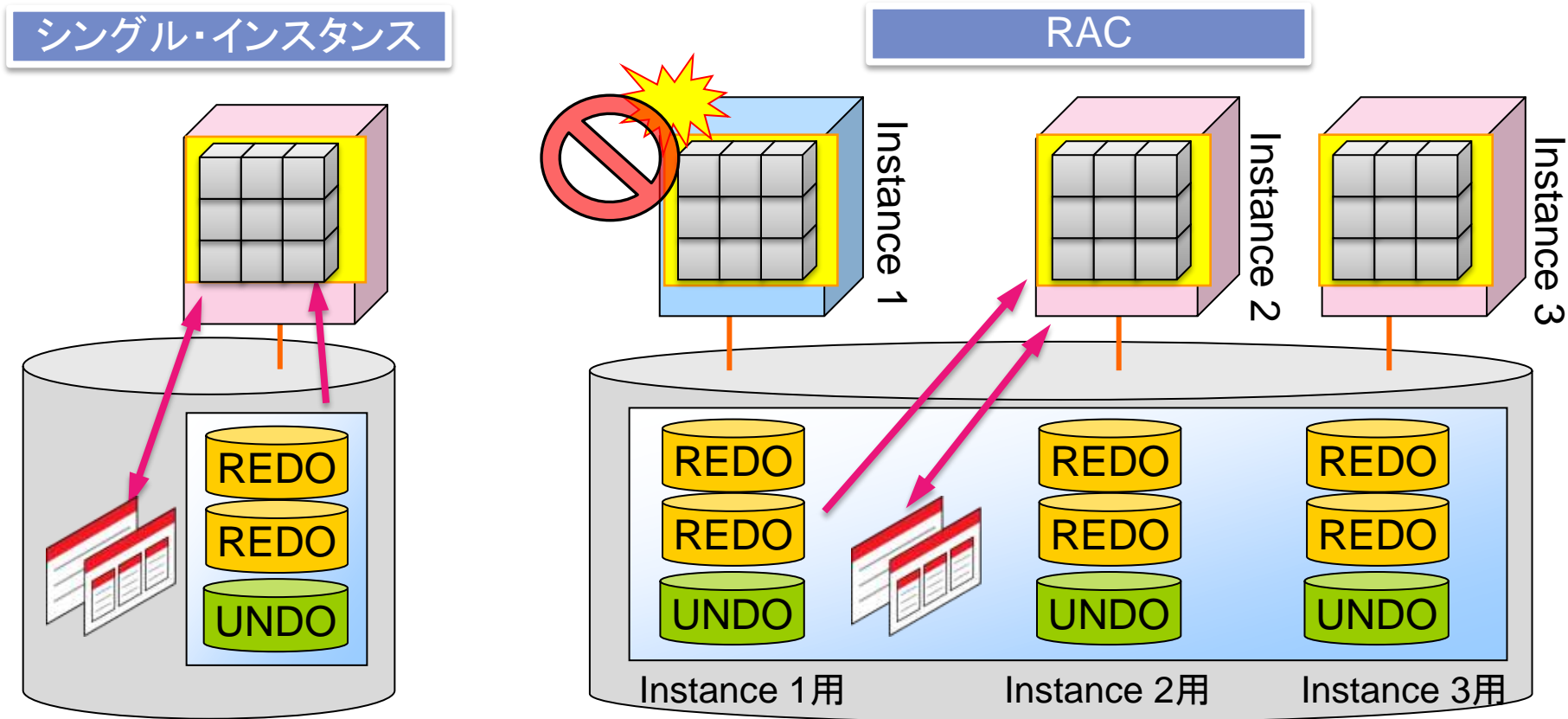
# RACの高可用性

- RACは障害ノードを自動的に切り離してシステムを継続
  - 障害ノード以外のノードへの接続は維持
  - 障害発生時も、トランザクションの一貫性を自動的に維持
- インスタンス・リカバリ



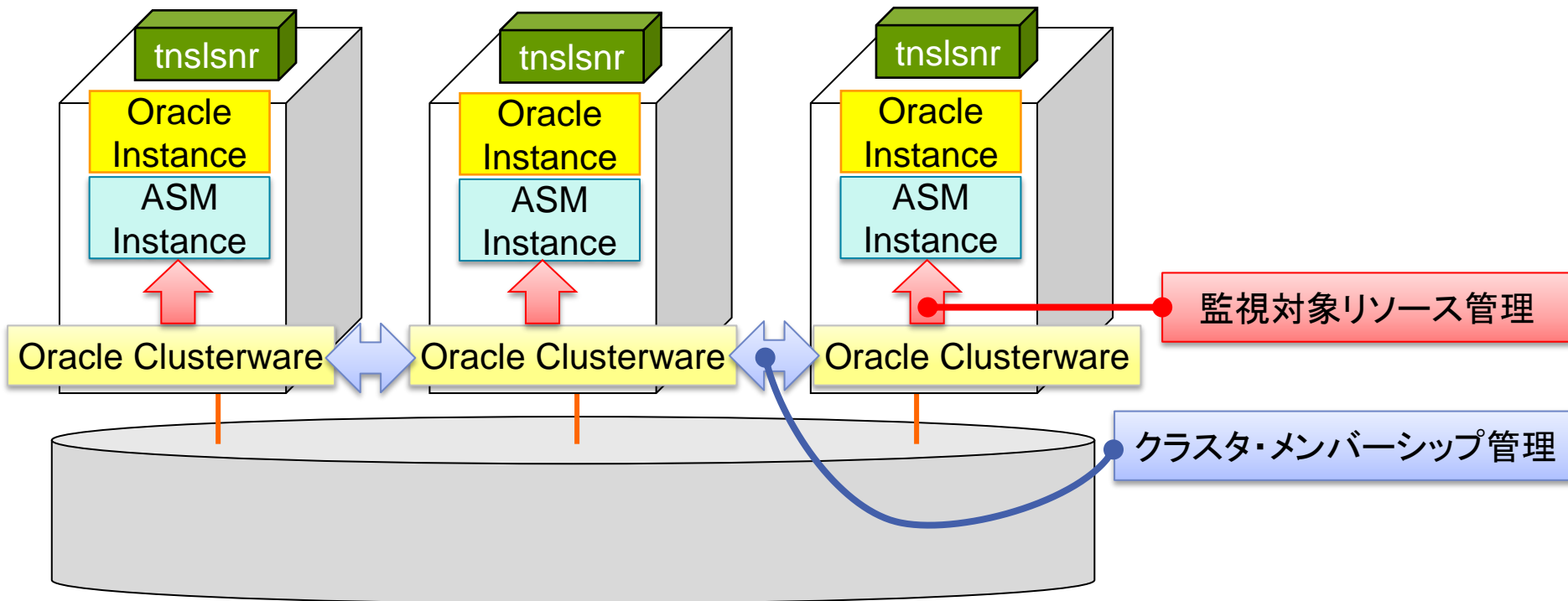
# RACのインスタンス・リカバリのしくみ

- シングル・インスタンス同様、REDOログを読込みインスタンス・リカバリを行う
- RACでは、正常ノードのデータベース・インスタンスが障害ノードのREDOログを読込みインスタンス・リカバリを行う



# Oracle Clusterwareの役割

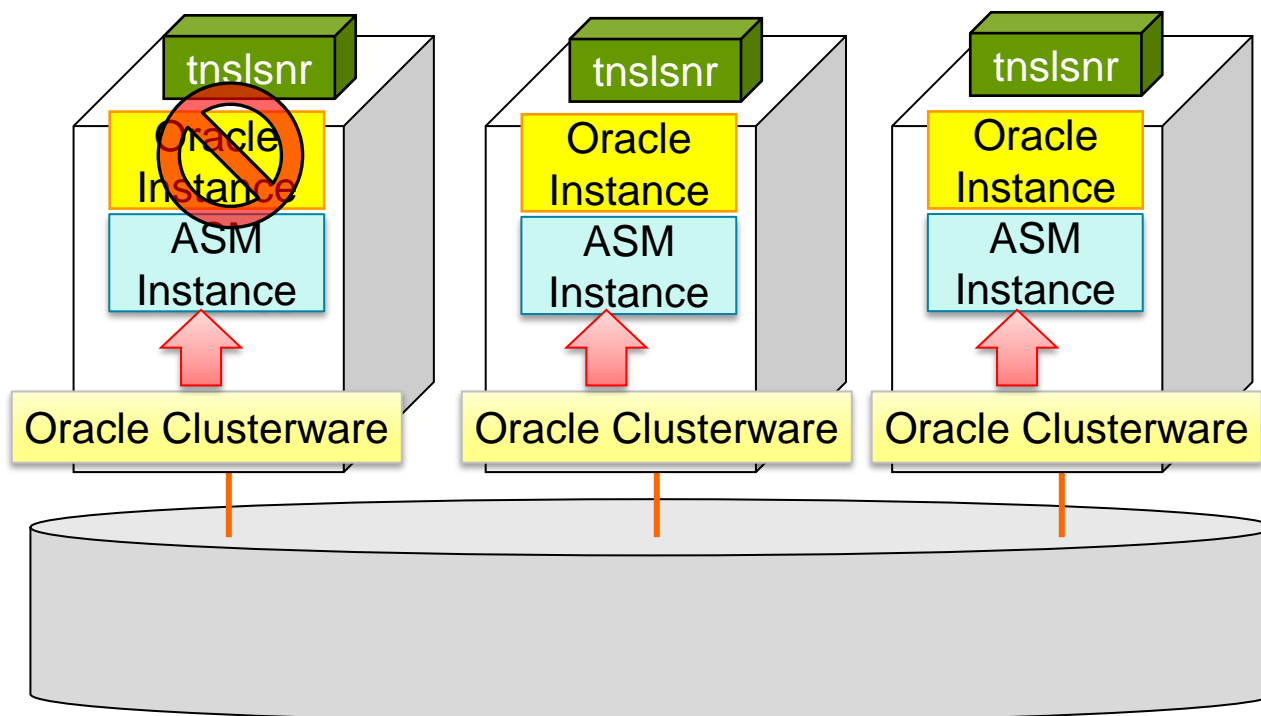
- 監視対象リソースの管理
  - リスナーやOracleインスタンス等の監視対象となるプロセスの起動/停止
- クラスタ・メンバーシップの管理
  - 各ノードが正常であることを確認



# 監視対象リソースの管理

## RAC高可用性のしくみ

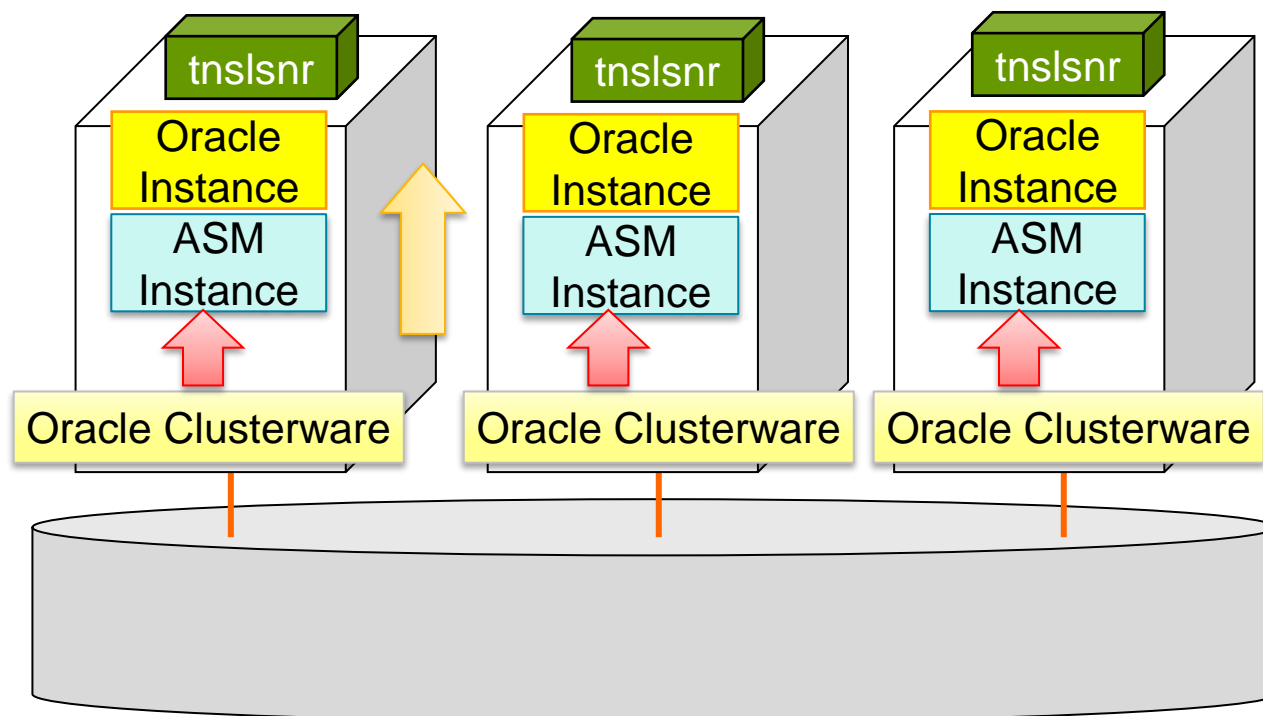
- 監視対象リソースの起動 / 停止の管理:
  - 障害でリソースが異常終了した場合、リソースを自動的に再起動



# 監視対象リソースの管理

## RAC高可用性のしくみ

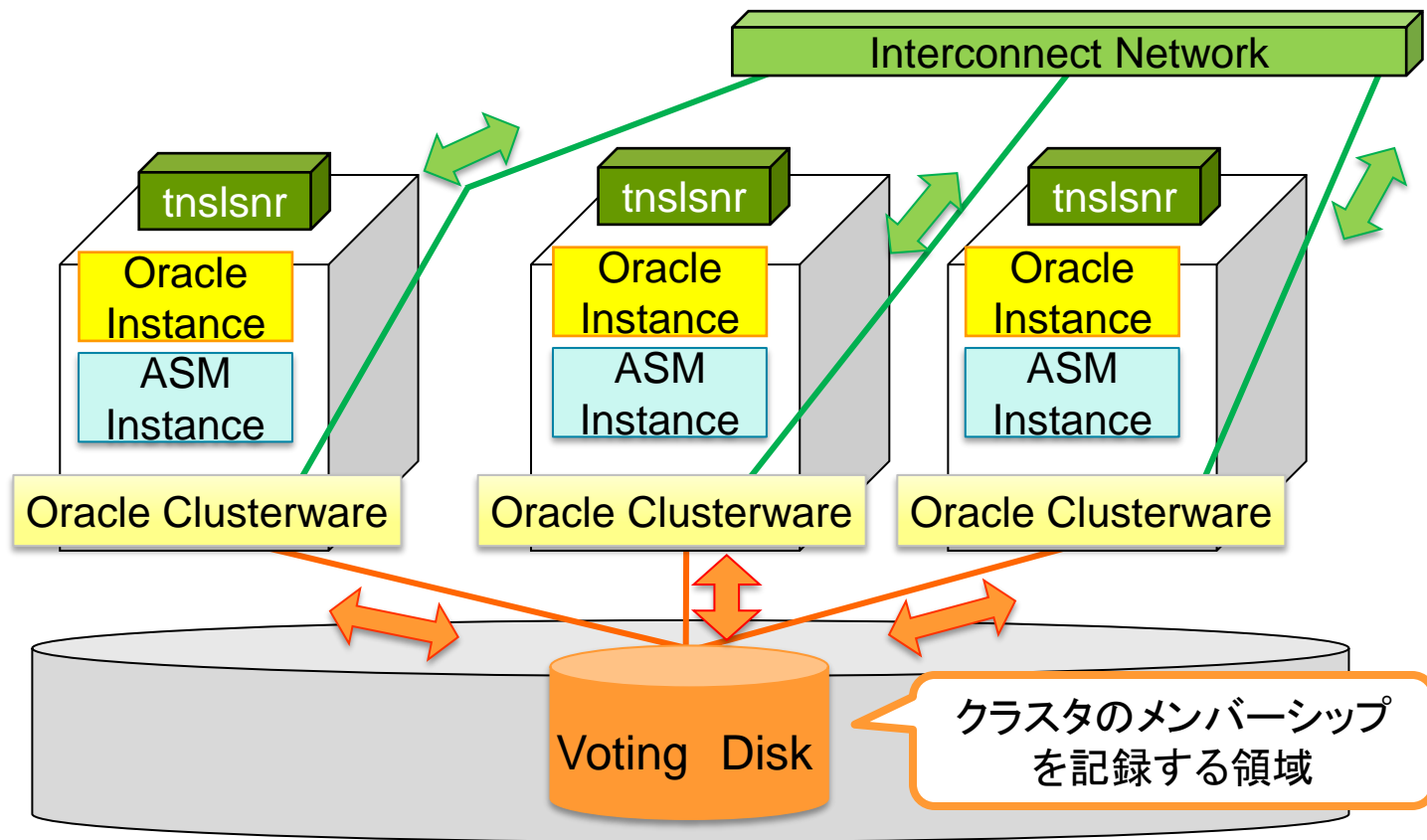
- 監視対象リソースの起動 / 停止の管理:
  - 障害でリソースが異常終了した場合、リソースを自動的に再起動



# クラスタ・メンバーシップの管理

## RAC高可用性のしくみ

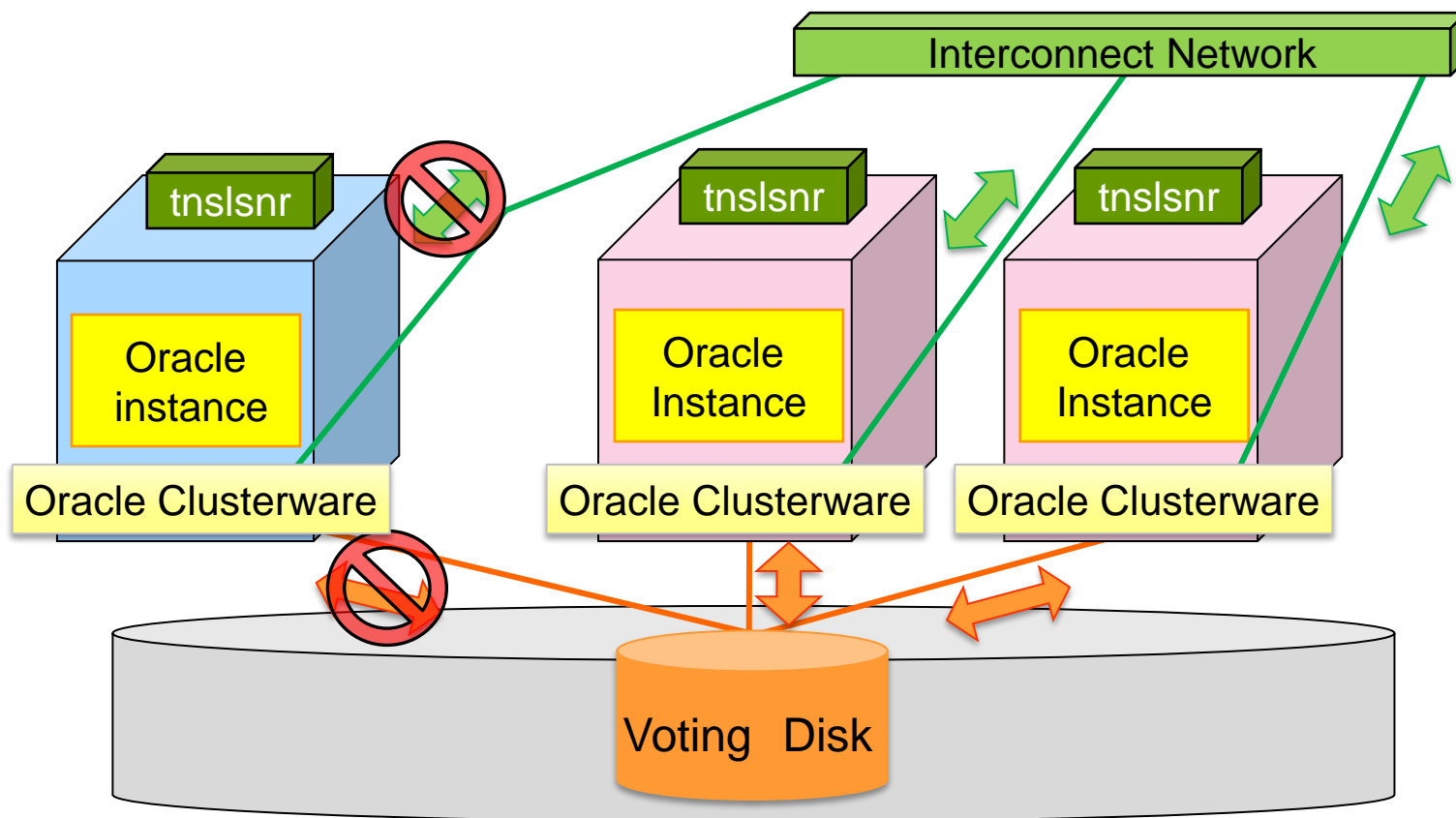
- ノード間の生存確認(ハートビート)の経路:
  - インターコネク特・ネットワーク上
  - 共有ディスク上



# クラスタ・メンバーシップの管理

## RAC高可用性のしくみ

- ハート・ビートに失敗すると、自動的にその障害ノードを切り離す

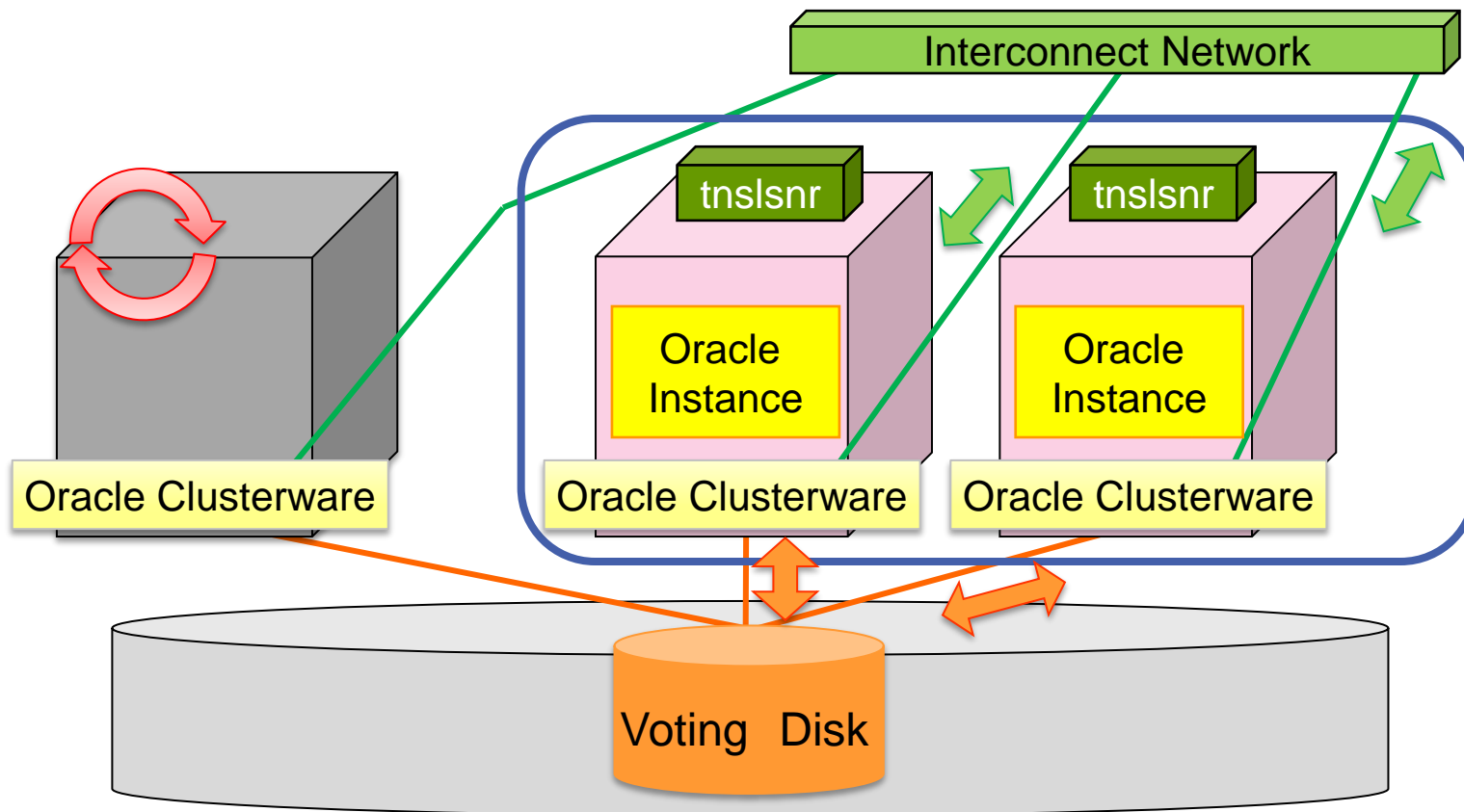




# クラスタ・メンバーシップの管理

## RAC高可用性のしくみ

- 障害ノードのOSを再起動させることで、クラスタから切り離し
- 生存ノードでシステムを継続
  - 生存ノードがインスタンス・リカバリを行う



# RAC高可用性のまとめ

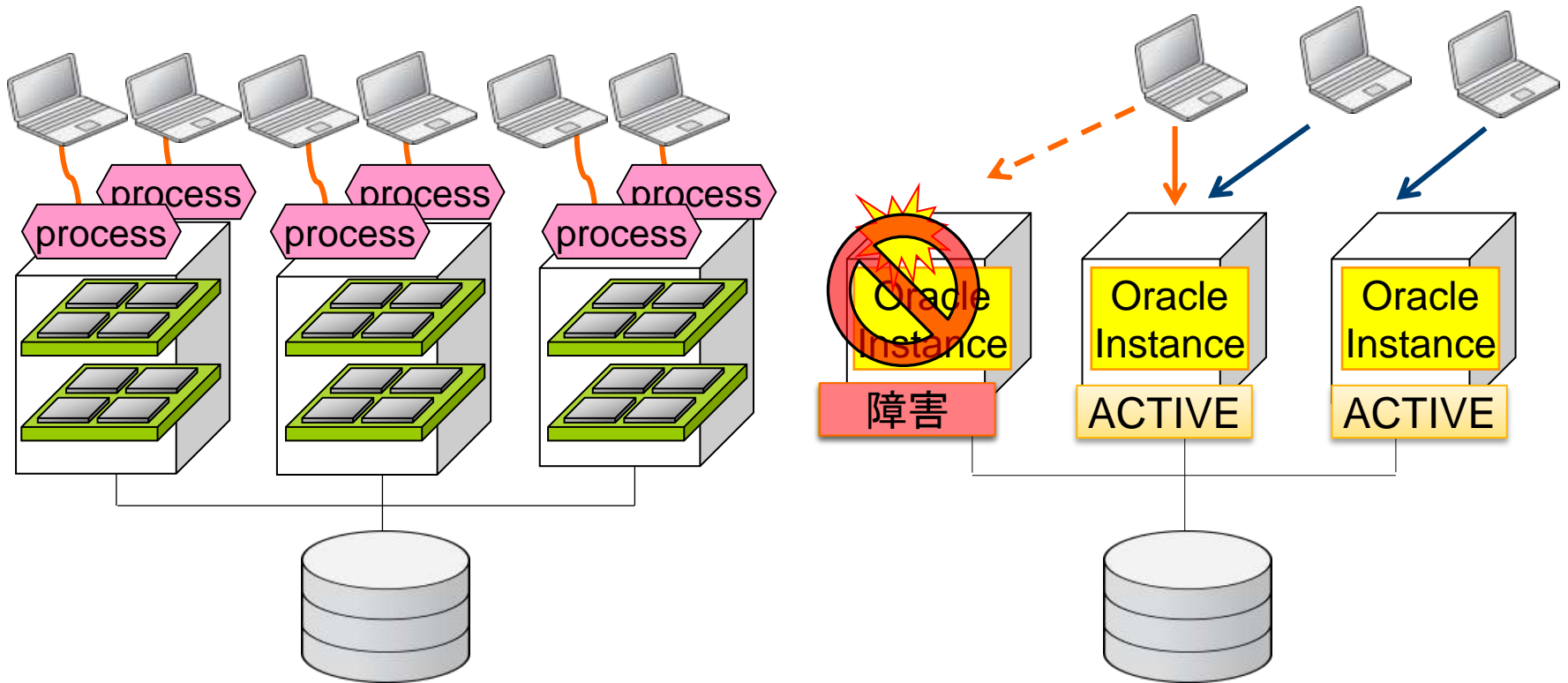
- RACは障害ノードを自動的に切り離してシステムを継続
  - 正常ノードがインスタンス・リカバリを行う
- Oracle Clusterware
  - 監視対象リソースの管理
  - クラスタ・メンバーシップの管理
- RACは高可用性を実現
  - ✓ サーバー障害時のダウンタイムが短い
  - ✓ トランザクションの一貫性が維持される

# まとめ



# Oracle Real Application Clusters

- RACは最適なデータベース・サーバー環境
  - スケーラビリティ → 複数ノードのCPUを利用して同時実行性を向上
  - 高可用性 → 障害ノードを切り離してシステムを継続



# OTNセミナーオンデマンド

コンテンツに対する  
ご意見・ご感想を是非お寄せください。

OTNオンデマンド 感想



[http://blogs.oracle.com/oracle4engineer/entry/otn\\_ondemand\\_questionnaire](http://blogs.oracle.com/oracle4engineer/entry/otn_ondemand_questionnaire)

上記に簡単なアンケート入力フォームをご用意しております。

セミナー講師/資料作成者にフィードバックし、  
コンテンツのより一層の改善に役立てさせていただきます。

是非ご協力をよろしくお願いいたします。

# OTNセミナーオンデマンド

日本オラクルのエンジニアが作成したセミナー資料・動画ダウンロードサイト

## 掲載コンテンツカテゴリ(一部抜粋)

Database 基礎

Database 現場テクニック

Database スペシャリストが語る

Java

WebLogic Server/アプリケーション・グリッド

EPM/BI 技術情報

サーバー

ストレージ



超入門! Oracle データベースって何  
再生時間: 60分

100以上のコンテンツをログイン不要でダウンロードし放題

データベースからハードウェアまで充実のラインナップ

毎月、旬なトピックの新作コンテンツが続々登場

## 例えばこんな使い方

- 製品概要を効率的につかむ
- 基礎を体系的に学ぶ/学ばせる
- 時間や場所を選ばず(オンデマンド)に受講
- スマートフォンで通勤中にも受講可能



毎月チェック!



コンテンツ一覧 はこちら

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

新作&おすすめコンテンツ情報 はこちら

<http://oracletech.jp/seminar/recommended/000073.html>

OTNオンデマンド



# オラクルエンジニア通信

オラクル製品に関わるエンジニアの方のための技術情報サイト

## オラクルエンジニア通信 - 技術資料、マニュアル、セミナー

Oracleエンジニアのための技術情報サイト by Oracle Japan

新着情報を知りたい

技術資料を探したい

セミナーを受けたい

**About**

Oracleエンジニアの方がスキルアップしていただくために、厳選した情報をお届けしています

技術資料		インストールガイド・設定チュートリアルetc. 欲しい資料への最短ルート
アクセスランキング		他のエンジニアは何を見ているのか？人気資料のランキングは毎月更新
特集テーマ Pick UP		性能管理やチューニングなど月間テーマを掘り下げて詳細にご説明
技術コラム		SQLスクリプト、索引メンテナンスetc. 当たり前運用/機能が見違える!?

<http://blogs.oracle.com/oracle4engineer/>

オラクルエンジニア通信



製品/技術  
情報



Oracle Databaseっていくら？オプション機能も見積れる簡単ツールが大活躍

セミナー



基礎から最新技術までお勧めセミナーで自分にあった学習方法が見つかる

スキルアップ



ORACLE MASTER ! 試験頻出分野の模擬問題と解説を好評連載中

Viva!  
Developer



全国で活躍しているエンジニアにスポットライト。きらりと輝くスキルと視点を盗もう

<http://oracletech.jp/>

oracletech





あなたにいちばん近いオラクル



# Oracle Direct

まずはお問合せください

Oracle Direct



システムの検討・構築から運用まで、ITプロジェクト全般の相談窓口としてご支援いたします。  
システム構成やライセンス/購入方法などお気軽にお問い合わせ下さい。

## Web問い合わせフォーム

専用お問い合わせフォームにてご相談内容を承ります。  
[http://www.oracle.co.jp/inq\\_pl/INQUIRY/quest?rid=28](http://www.oracle.co.jp/inq_pl/INQUIRY/quest?rid=28)

※フォームの入力にはログインが必要となります。  
※こちらから詳細確認のお電話を差し上げる場合がありますので  
ご登録の連絡先が最新のものになっているかご確認下さい。

## フリーダイヤル

0120-155-096

※月曜～金曜  
9:00～12:00、13:00～18:00  
(祝日および年末年始除く)

ORACLE

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

**ORACLE®**