

Oracle® Database

2 日でデータ・ウェアハウス・ガイド

11g リリース 1 (11.1)

部品番号 : E05764-01

2007 年 9 月

Oracle Database 2 日でデータ・ウェアハウス・ガイド, 11g リリース 1 (11.1)

部品番号 : E05764-01

原本名 : Oracle Database 2 Day + Data Warehousing Guide 11g Release 1 (11.1)

原本部品番号 : B28314-01

原本著者 : Michelle Bird、Paul Lane

原本協力者 : Mark van de Wiel、Tammy Bednar、Hermann Baer、Kristy Browder Edwards、John Haydu、Sushil Kumar、George Lumpkin、Mughees Minhas、Mark Townsend、Murali Thiagarajan

Copyright © 2007, Oracle. All rights reserved.

制限付権利の説明

このプログラム（ソフトウェアおよびドキュメントを含む）には、オラクル社およびその関連会社に所有権のある情報が含まれています。このプログラムの使用または開示は、オラクル社およびその関連会社との契約に記された制約条件に従うものとします。著作権、特許権およびその他の知的財産権と工業所有権に関する法律により保護されています。

独立して作成された他のソフトウェアとの互換性を得るために必要な場合、もしくは法律によって規定される場合を除き、このプログラムのリバース・エンジニアリング、逆アセンブル、逆コンパイル等は禁止されています。

このドキュメントの情報は、予告なしに変更される場合があります。オラクル社およびその関連会社は、このドキュメントに誤りが無いことの保証は致し兼ねます。これらのプログラムのライセンス契約で許諾されている場合を除き、プログラムを形式、手段（電子的または機械的）、目的に関係なく、複製または転用することはできません。

このプログラムが米国政府機関、もしくは米国政府機関に代わってこのプログラムをライセンスまたは使用する者に提供される場合は、次の注意が適用されます。

U.S. GOVERNMENT RIGHTS

Programs, software, databases, and related documentation and technical data delivered to U.S. Government customers are "commercial computer software" or "commercial technical data" pursuant to the applicable Federal Acquisition Regulation and agency-specific supplemental regulations. As such, use, duplication, disclosure, modification, and adaptation of the Programs, including documentation and technical data, shall be subject to the licensing restrictions set forth in the applicable Oracle license agreement, and, to the extent applicable, the additional rights set forth in FAR 52.227-19, Commercial Computer Software--Restricted Rights (June 1987). Oracle USA, Inc., 500 Oracle Parkway, Redwood City, CA 94065.

このプログラムは、核、航空産業、大量輸送、医療あるいはその他の危険が伴うアプリケーションへの用途を目的としておりません。このプログラムをかかるとして使用する際、上述のアプリケーションを安全に使用するために、適切な安全装置、バックアップ、冗長性 (redundancy)、その他の対策を講じることは使用者の責任となります。万一かかるプログラムの使用に起因して損害が発生いたしましても、オラクル社およびその関連会社は一切責任を負いかねます。

Oracle、JD Edwards、PeopleSoft、Siebel は米国 Oracle Corporation およびその子会社、関連会社の登録商標です。その他の名称は、他社の商標の可能性がありま。

このプログラムは、第三者の Web サイトへリンクし、第三者のコンテンツ、製品、サービスへアクセスすることがあります。オラクル社およびその関連会社は第三者の Web サイトで提供されるコンテンツについては、一切の責任を負いかねます。当該コンテンツの利用は、お客様の責任になります。第三者の製品またはサービスを購入する場合は、第三者と直接の取引となります。オラクル社およびその関連会社は、第三者の製品およびサービスの品質、契約の履行（製品またはサービスの提供、保証義務を含む）に関しては責任を負いかねます。また、第三者との取引により損失や損害が発生いたしましても、オラクル社およびその関連会社は一切の責任を負いかねます。

目次

| | |
|----------------------------------|------|
| はじめに | vii |
| 対象読者 | viii |
| ドキュメントのアクセシビリティについて | viii |
| 関連ドキュメント | viii |
| 表記規則 | ix |
| サポートおよびサービス | ix |
| | |
| 第1部 データ・ウェアハウスの構築 | |
| | |
| 1 データ・ウェアハウスの概要 | |
| このガイドについて | 1-2 |
| このガイドを使用する前に | 1-2 |
| このガイドの対象外 | 1-2 |
| データ・ウェアハウスについて | 1-2 |
| データ・ウェアハウスの主な特性 | 1-3 |
| Oracle データ・ウェアハウスの一般的なタスク | 1-3 |
| このガイドで説明するタスク | 1-3 |
| データ・ウェアハウス管理用ツール | 1-5 |
| | |
| 2 データ・ウェアハウス・システムの設定 | |
| データ・ウェアハウス・システムを設定する一般的な手順 | 2-2 |
| 環境の準備 | 2-2 |
| 適切なハードウェア設定 | 2-2 |
| 必要な CPU 数およびクロック・スピードについて | 2-3 |
| 必要なメモリーについて | 2-3 |
| 必要なディスク数について | 2-3 |
| 十分な I/O 帯域幅を判断する方法について | 2-3 |
| データ・ウェアハウスのハードウェア構成の検証 | 2-4 |
| dd ユーティリティについて | 2-4 |
| 例: dd ユーティリティの使用 | 2-5 |
| Orion ユーティリティについて | 2-5 |
| 例: Orion ユーティリティの使用 | 2-6 |
| データ・ウェアハウスのデータベースの設定 | 2-7 |
| メモリー管理パラメータの設定方法 | 2-7 |
| 例: 初期化パラメータの設定 | 2-7 |
| 重要な他の初期化パラメータの設定 | 2-8 |

| | |
|--|------|
| Oracle Warehouse Builder へのアクセス | 2-9 |
| Oracle Warehouse Builder デモンストレーションのインストール | 2-10 |

3 データ・ソースの識別およびメタデータのインポート

| | |
|---|-----|
| ソースからメタデータをインポートする一般的な手順 | 3-2 |
| Oracle Warehouse Builder のワークスペース、プロジェクトおよびその他のデバイスについて | 3-2 |
| 例：フラット・ファイルからのメタデータのインポート | 3-3 |
| フラット・ファイルの場所の指定 | 3-3 |
| プロジェクトでのモジュールの作成 | 3-4 |
| 「インポート・メタデータ・ウィザード」の起動 | 3-4 |
| 「フラット・ファイル・サンプル・ウィザード」の使用 | 3-5 |
| フラット・ファイル・データのインポート | 3-5 |

4 Oracle Warehouse Builder でのウェアハウスの定義

| | |
|---|------|
| リレーショナル・ターゲット・ウェアハウスを定義する一般的な手順 | 4-2 |
| ウェアハウス・ターゲット・スキーマの指定 | 4-2 |
| Oracle Warehouse Builder のフラット・ファイルのソースについて | 4-3 |
| 演習：ターゲット・モジュールへの外部表の追加 | 4-3 |
| ディメンションについて | 4-4 |
| 演習：ディメンションの理解 | 4-4 |
| レベルについて | 4-5 |
| レベル属性の定義 | 4-6 |
| 階層の定義 | 4-6 |
| ディメンション・ロール | 4-6 |
| レベルの関係 | 4-7 |
| ディメンションの例 | 4-7 |
| 制御行 | 4-8 |
| ディメンションの実装 | 4-8 |
| スター・スキーマ | 4-8 |
| バインド | 4-9 |
| キューブについて | 4-11 |
| キューブの定義 | 4-11 |
| キューブ・メジャー | 4-11 |
| キューブのディメンション | 4-11 |
| キューブの例 | 4-12 |
| キューブの実装 | 4-12 |
| キューブのリレーショナル実装 | 4-12 |
| バインド | 4-13 |

第 II 部 データ・ウェアハウスへのデータのロード

5 ETL ロジックの定義

| | |
|-----------------------|-----|
| マッピングおよび演算子について | 5-2 |
| マッピングを定義する手順 | 5-2 |
| マッピングの作成 | 5-3 |
| 演算子のタイプ | 5-4 |

| | |
|-------------------------------------|------|
| 演算子の追加 | 5-4 |
| ワークスペース・オブジェクトにバインドする演算子の追加 | 5-5 |
| バインドされていない演算子を属性なしで作成 | 5-5 |
| 既存のワークスペース・オブジェクトから選択してバインド | 5-6 |
| 演算子の編集 | 5-6 |
| 演算子の接続 | 5-6 |
| 属性の接続 | 5-7 |
| グループの接続 | 5-7 |
| 例: マッピング・エディタによるステージング領域表の作成 | 5-8 |
| マッピング・プロパティの設定 | 5-9 |
| ターゲット・ロード順序 | 5-9 |
| デフォルトにリセット | 5-10 |
| 演算子、グループおよび属性プロパティの設定 | 5-11 |
| 演算子とワークスペース・オブジェクトの同期化 | 5-11 |
| 演算子の同期化 | 5-12 |
| ワークスペース・オブジェクトから演算子への同期化 | 5-12 |
| ワークスペース・オブジェクトに基づく演算子の同期化 | 5-13 |
| 演算子からワークスペース・オブジェクトへの同期化 | 5-14 |

6 ロード・パフォーマンスの向上

| | |
|---|-----|
| パーティション交換ロードを使用したパフォーマンスの向上 | 6-2 |
| パーティション交換ロードについて | 6-2 |
| パーティション交換ロードのマッピングの構成 | 6-3 |
| ダイレクトおよび間接パーティション交換ロード | 6-4 |
| 間接パーティション交換ロードの使用 | 6-4 |
| 例: ファクト表を公開するためのダイレクト・パーティション交換ロードの使用 | 6-4 |
| パーティション交換ロードの効果的な使用 | 6-5 |
| マッピングでのターゲットの構成 | 6-6 |
| 手順 1: すべてのパーティションの作成 | 6-6 |
| 手順 2: 「LOCAL」 オプションを使用したすべての索引の作成 | 6-7 |
| 手順 3: 「索引を使用」 オプションを使用する主キーまたは一意キー | 6-8 |
| Oracle Warehouse Builder でのパーティション交換ロードの使用の制限 | 6-9 |

7 ターゲット・スキーマへの配布および ETL ロジックの実行

| | |
|----------------------------------|-----|
| 配布について | 7-2 |
| コントロール・センターの概要 | 7-2 |
| 配布の物理詳細の構成 | 7-3 |
| 配布アクション | 7-3 |
| 配布プロセス | 7-4 |
| スケジュールについて | 7-4 |
| スケジュールの定義および使用のプロセス | 7-5 |
| オブジェクトの配布 | 7-6 |
| ETL ジョブの開始 | 7-6 |
| データの表示 | 7-6 |

第 III 部 データ・ウェアハウスの分析およびレポート作成

8 メタデータ依存性の分析

| | |
|---------------------------|-----|
| メタデータ依存性マネージャについて | 8-2 |
| 使用方法のシナリオ | 8-3 |
| 系統および影響分析ダイアグラムについて | 8-3 |
| LIA ダイアグラムの生成 | 8-5 |
| LIA ダイアグラムの表示の変更 | 8-5 |
| LIA ダイアグラムでのグループの使用 | 8-5 |
| オブジェクトの属性の表示 | 8-6 |
| 依存性マネージャでのオブジェクトの変更 | 8-6 |

9 レポートおよび分析の SQL

| | |
|-------------------------------------|------|
| ビジネス問合せに回答する SQL アナリティック機能の使用 | 9-2 |
| ROLLUP 関数を使用したレポートへの合計の追加方法 | 9-3 |
| ROLLUP 関数を使用するとき | 9-3 |
| 例: ROLLUP 関数の使用 | 9-3 |
| CUBE 関数を使用した異なるレベルでの合計の分割方法 | 9-4 |
| CUBE 関数を使用するとき | 9-4 |
| 例: CUBE 関数の使用 | 9-5 |
| GROUPING 関数を使用した小計の追加方法 | 9-5 |
| GROUPING 関数を使用するとき | 9-6 |
| 例: GROUPING 関数の使用 | 9-6 |
| GROUPING SETS 関数を使用した集計の結合方法 | 9-7 |
| GROUPING SETS 関数を使用するとき | 9-7 |
| 例: GROUPING SETS 関数の使用 | 9-7 |
| RANK 関数を使用したランキングの計算方法 | 9-8 |
| RANK 関数を使用するとき | 9-8 |
| 例: RANK 関数の使用 | 9-8 |
| 合計に対する相対的な寄与率の計算方法 | 9-9 |
| 例: 総計に対する相対的な寄与率の計算 | 9-9 |
| ウィンドウ関数を使用した行間計算の実行方法 | 9-11 |
| 例: 行間計算の実行 | 9-11 |
| ウィンドウ関数を使用した移動平均の計算方法 | 9-12 |
| 例: 移動平均の計算 | 9-13 |
| パーティション外部結合を使用したスパースなデータの処理 | 9-13 |
| パーティション外部結合を使用するとき | 9-14 |
| 例: パーティション外部結合の使用 | 9-14 |
| ビジネス問合せを単純化する WITH 句の使用 | 9-14 |
| WITH 句を使用するとき | 9-14 |
| 例: WITH 句の使用 | 9-15 |

第 IV 部 データ・ウェアハウスの管理

10 データ・ウェアハウスのリフレッシュ

| | |
|---------------------------------|------|
| データ・ウェアハウスのリフレッシュについて | 10-2 |
| 例: データ・ウェアハウスのリフレッシュ | 10-2 |
| ローリング・ウィンドウを使用したデータのオフロード | 10-5 |
| 例: ローリング・ウィンドウの使用 | 10-6 |

11 データ・ウェアハウスの操作の最適化

| | |
|--|------|
| システムのオーバーロードの回避 | 11-2 |
| システム・パフォーマンスの監視 | 11-2 |
| パラレル実行パフォーマンスの監視 | 11-2 |
| I/O の監視 | 11-3 |
| データベース・リソース・マネージャの使用 | 11-4 |
| 索引およびマテリアライズド・ビューの使用の最適化 | 11-5 |
| 例: SQL アクセス・アドバイザを使用した索引およびマテリアライズド・ビューの 使用の最適化 | 11-5 |
| 記憶域要件の最適化 | 11-6 |
| データ圧縮による記憶域の改善 | 11-6 |

12 パフォーマンス・ボトルネックの排除

| | |
|---|-------|
| SQL の効率的な実行の検証 | 12-2 |
| オブティマイザ統計の分析 | 12-2 |
| 実行計画の分析 | 12-2 |
| 例: 実行計画の出力の分析 | 12-3 |
| ヒントを使用したデータ・ウェアハウスのパフォーマンスの向上 | 12-4 |
| 例: ヒントを使用したデータ・ウェアハウスのパフォーマンスの向上 | 12-4 |
| アドバイザを使用した SQL パフォーマンスの検証 | 12-5 |
| リソース使用量の最小化によるパフォーマンスの向上 | 12-6 |
| パフォーマンスの向上: パーティション化 | 12-6 |
| パフォーマンスの向上: パーティション・プルーニング | 12-6 |
| パフォーマンスの向上: パーティションワイズ結合 | 12-6 |
| 例: SQL アクセス・アドバイザを使用したパーティション化の評価 | 12-7 |
| パフォーマンスの向上: 間合せのリライトおよびマテリアライズド・ビュー | 12-8 |
| パフォーマンスの向上: 索引 | 12-8 |
| 最適なリソースの使用 | 12-8 |
| パラレル実行でのパフォーマンスの最適化 | 12-9 |
| パラレル実行の動作方法 | 12-9 |
| 並列度の設定 | 12-9 |
| 例: 並列度の設定 | 12-10 |
| 待機イベントについて | 12-10 |

13 データ・ウェアハウスのバックアップおよびリカバリ

| | |
|--|------|
| データ・ウェアハウスのバックアップおよびリカバリの処理方法 | 13-2 |
| バックアップおよびリカバリの計画とベスト・プラクティス | 13-3 |
| ベスト・プラクティス A: ARCHIVELOG モードの使用 | 13-3 |
| 停止時間の許容について | 13-3 |
| ベスト・プラクティス B: Recovery Manager の使用 | 13-4 |
| ベスト・プラクティス C: 読取り専用の表領域の使用 | 13-4 |
| ベスト・プラクティス D: NOLOGGING 操作の計画 | 13-5 |
| 抽出、変換およびロード | 13-6 |
| ETL 計画および NOLOGGING 操作 | 13-6 |
| ブロック変更トラッキング・ファイルのサイジング | 13-7 |
| 増分バックアップ | 13-7 |
| 増分アプローチ | 13-8 |
| ベスト・プラクティス E: すべての表領域の重要度が同等でない場合 | 13-8 |

14 データ・ウェアハウスの操作の最適化

| | |
|-------------------------------------|------|
| データ・ウェアハウスのセキュリティについて | 14-2 |
| データ・ウェアハウスのセキュリティの必要性について | 14-2 |
| データ・ウェアハウスのセキュリティのロールおよび権限の使用 | 14-2 |
| データ・ウェアハウスの仮想プライベート・データベースの使用 | 14-3 |
| 仮想プライベート・データベースの動作方法 | 14-3 |
| Oracle Label Security の概要 | 14-4 |
| Oracle Label Security の動作方法 | 14-4 |
| データ・ウェアハウスによりラベルを利用する方法 | 14-4 |
| データ・ウェアハウスのファイングレイン監査の概要 | 14-5 |
| データ・ウェアハウスの透過的データ暗号化の概要 | 14-5 |

索引

はじめに

ここでは、次の項目について説明します。

- 対象読者
- ドキュメントのアクセシビリティについて
- 関連ドキュメント
- 表記規則
- サポートおよびサービス

対象読者

このガイドは、Oracle Database を使用して日常の一般的なデータ・ウェアハウスのタスクを実行するユーザーを対象としています。コンピュータの基本的な知識があり、『Oracle Database 2 日でデータベース管理者』を参照済であることが主な前提条件です。

このガイドは、特に次のようなユーザーを対象としています。

- データ・ウェアハウスの管理スキルを習得する必要がある Oracle DBA
- データ・ウェアハウスの基礎知識はある程度あるが、Oracle を初めて使用する DBA

ドキュメントのアクセシビリティについて

オラクル社は、障害のあるお客様にもオラクル社の製品、サービスおよびサポート・ドキュメントを簡単にご利用いただけることを目標としています。オラクル社のドキュメントには、ユーザーが障害支援技術を使用して情報を利用できる機能が組み込まれています。HTML 形式のドキュメントで用意されており、障害のあるお客様が簡単にアクセスできるようにマークアップされています。標準規格は改善されつつあります。オラクル社はドキュメントをすべてのお客様がご利用できるように、市場をリードする他の技術ベンダーと積極的に連携して技術的な問題に対応しています。オラクル社のアクセシビリティについての詳細情報は、Oracle Accessibility Program の次の Web サイト <http://www.oracle.com/accessibility/> を参照してください。

ドキュメント内のサンプル・コードのアクセシビリティについて

スクリーン・リーダーは、ドキュメント内のサンプル・コードを正確に読めない場合があります。コード表記規則では閉じ括弧のみを行に記述する必要があります。しかし、一部のスクリーン・リーダーは括弧のみの行を読まない場合があります。

外部 Web サイトのドキュメントのアクセシビリティについて

このドキュメントにはオラクル社およびその関連会社が所有または管理しない Web サイトへのリンクが含まれている場合があります。オラクル社およびその関連会社は、それらの Web サイトのアクセシビリティに関しての評価や言及は行っておりません。

Oracle サポート・サービスへの TTY アクセス

アメリカ国内では、Oracle サポート・サービスへ 24 時間年中無休でテキスト電話 (TTY) アクセスが提供されています。TTY サポートについては、(800)446-2398 にお電話ください。

関連ドキュメント

詳細は、次の Oracle リソースを参照してください。

- 『Oracle Database 2 日でデータベース管理者』
- 『Oracle Database データ・ウェアハウス・ガイド』
- 『Oracle Database 管理者ガイド』

表記規則

このマニュアルでは次の表記規則を使用します。

| 規則 | 意味 |
|---------|--|
| 太字 | 太字は、操作に関連する Graphical User Interface の要素または本文中または用語集で定義されている用語を示します。 |
| イタリック | イタリック体は、特定の値を指定するプレースホルダ変数を示します。 |
| 固定幅フォント | 固定幅フォントは、段落内のコマンド、URL、例内のコード、画面に表示されるテキストまたは入力するテキストを示します。 |

サポートおよびサービス

次の各項に、各サービスに接続するための URL を記載します。

Oracle サポート・サービス

オラクル製品サポートの購入方法、および Oracle サポート・サービスへの連絡方法の詳細は、次の URL を参照してください。

<http://www.oracle.co.jp/support/>

製品マニュアル

製品のマニュアルは、次の URL にあります。

<http://otn.oracle.co.jp/document/>

研修およびトレーニング

研修に関する情報とスケジュールは、次の URL で入手できます。

<http://www.oracle.co.jp/education/>

その他の情報

オラクル製品やサービスに関するその他の情報については、次の URL から参照してください。

<http://www.oracle.co.jp>

<http://otn.oracle.co.jp>

注意： ドキュメント内に記載されている URL や参照ドキュメントには、Oracle Corporation が提供する英語の情報も含まれています。日本語版の情報については、前述の URL を参照してください。

第 I 部

データ・ウェアハウスの構築

ここでは、データ・ウェアハウスの構築方法について説明します。内容は次のとおりです。

- 第 1 章「データ・ウェアハウスの概要」
- 第 2 章「データ・ウェアハウス・システムの設定」
- 第 3 章「データ・ソースの識別およびメタデータのインポート」
- 第 4 章「Oracle Warehouse Builder でのウェアハウスの定義」

データ・ウェアハウスの概要

データ・ウェアハウスの管理、設計および実装が必要なため、ユーザーが Oracle データ・ウェアハウスの操作全体を管理し、効率的なパフォーマンスを維持します。

この章の内容は次のとおりです。

- [このガイドについて](#)
- [データ・ウェアハウスについて](#)
- [Oracle データ・ウェアハウスの一般的なタスク](#)
- [データ・ウェアハウス管理用ツール](#)

このガイドについて

このガイドでは、データ・ウェアハウスを実装して管理するために必要な共通の日常のタスクを実行する方法を説明しています。このガイドの目標は、Oracle Database で使用可能なデータ・ウェアハウス・ソリューションの概要を示すことです。基本的なパフォーマンス監視タスクを実行する方法を含め、データ・ウェアハウスの稼働を続けるために必要な共通の管理タスクおよび設計タスクを行う方法を説明します。

このガイドで使用する主要なインタフェースは、Oracle Enterprise Manager、Oracle Warehouse Builder および SQL*Plus です。

このガイドを使用する前に

このガイドを使用する前に、次のことを実行する必要があります。

- 『Oracle Database 2 日でデータベース管理者』をよく理解すること
- 「データ・ウェアハウス管理用ツール」に示されている必要なツールを手に入れること

このガイドの対象外

このガイドはタスク指向です。タスクが実行される理由と時期を説明することが目的です。必要に応じて、現行のタスクを理解して完了するために必要な概念を説明します。このガイドは、データ・ウェアハウスを Oracle Database に実装するための完全なガイドではありません。機能および使用上の詳細な手順に関する完全な概念の情報は、次の適切なドキュメントを参照してください。

- 『Oracle Database データ・ウェアハウス・ガイド』
- 『Oracle Warehouse Builder ユーザーズ・ガイド』
- 管理タスクについては、『Oracle Database 管理者ガイド』を参照してください。
- OLAP については、『Oracle OLAP リファレンス』を参照してください。
- データ・マイニングについては、『Oracle Data Mining 概要』を参照してください。

また、このガイドでは、Oracle Enterprise Manager で使用可能な機能について使用方法を部分的に説明しますが、Oracle Enterprise Manager の詳細な情報は提供しません。Oracle Enterprise Manager の詳細は、次の資料を参照してください。

- Oracle Enterprise Manager のオンライン・ヘルプ
- 『Oracle Enterprise Manager 概要』

データ・ウェアハウスについて

データ・ウェアハウスは、問合せと分析のために設計されたリレーショナル・データベースまたは多次元データベースです。OLTP システムのドメインであるトランザクション処理用には最適化されません。通常はデータ・ウェアハウスにより、複数のソースから導出された履歴データおよびトランザクション・データが統合されます。データ・ウェアハウスにより、トランザクション・ワークロードから分析ワークロードが分割され、組織が複数のソースからデータを統合できるようにします。

データ・ウェアハウス内のデータには通常、複数月または複数年のデータが格納されており、履歴の分析をサポートします。データ・ウェアハウス内のデータは通常、OLTP アプリケーション、メインフレーム・アプリケーション、または外部データのプロバイダなどのような 1 つ以上のデータ・ソースから抽出、変換およびロード (ETL) の各プロセスを介してロードされます。

データ・ウェアハウスのエンドユーザーは、多くの場合、時間に関連するデータの分析を実行します。例として、昨年の連結売上高、在庫分析、製品別収益および顧客別収益があります。より高度な分析に、傾向分析とデータ・マイニングがあり、既存のデータを使用して傾向または今後を予測します。一般にデータ・ウェアハウスでは、ビジネス・インテリジェンス環境の基盤を提供します。

このガイドでは、スター・スキーマを含むリレーショナル実装について説明しています。多次元データ・ウェアハウスの詳細は、『Oracle Database データ・ウェアハウス・ガイド』を参照してください。

データ・ウェアハウスの主な特性

データ・ウェアハウスの主な特性は次のとおりです。

- データの一部は、単純化およびパフォーマンスの向上のために非正規化されています。
- 大量の履歴データが使用されます。
- 問合せにより大量のデータを取得することがあります。
- 計画済問合せと非定型問合せの両方とも一般的です。
- データのロードは制御されます。

一般的に、データ・ウェアハウスを正常に動作させるためには高いデータ・スループットが重要となります。

Oracle データ・ウェアハウスの一般的なタスク

Oracle データ・ウェアハウス管理者または設計者として、次のタスクを行うことが予想されます。

- データ・ウェアハウスとして使用するための Oracle データベースの構成
- データ・ウェアハウスの設計
- 新規リリースのレベルに合せたデータベースおよびソフトウェアのアップグレードの実行
- スキーマ・オブジェクトの管理（表、索引およびマテリアライズド・ビューなど）
- ユーザーおよびセキュリティの管理
- 抽出、変換およびロード（ETL）の各プロセスに使用するルーチンの開発
- データ・ウェアハウス内のデータに基づいたレポートの作成
- 必要に応じたデータ・ウェアハウスのバックアップおよびリカバリの実行
- データ・ウェアハウスのパフォーマンスの監視と必要に応じた予防処理または修正処理

中小規模のデータ・ウェアハウス環境では、これらのタスクを単独で実行する可能性があります。大企業のような環境の場合、ジョブはデータベース・セキュリティまたはデータベースのチューニングなどの専門を持つ数名の DBA および設計者に分割されます。

このガイドで説明するタスク

このガイドでは、次のタスクについて説明します。

1. データ・ウェアハウスとして使用するための Oracle データベースを構成します。

[第 2 章「データ・ウェアハウス・システムの設定」](#)を参照してください。この項では、このガイド全体の演習で参照されるデモンストレーションにアクセスする方法についても説明されています。

2. データを統合する最初のステップを実行します。

[第 3 章「データ・ソースの識別およびメタデータのインポート」](#)の指示に従います。

3. ウェアハウス内のターゲット・オブジェクトの定義を開始します。
ターゲット・ウェアハウスの外部表、ディメンションおよびキューブを定義する方法については、[第4章「Oracle Warehouse Builder でのウェアハウスの定義」](#)を参照してください。
4. ターゲットにデータを抽出、変換およびロードする計画を定義します。
手順2で指定したソースからデータを抽出し、そのデータを変換し、次に手順3で設計したターゲットにデータをロードする ETL ロジックを定義する方法については、[第5章「ETL ロジックの定義」](#)を参照してください。
5. パフォーマンスの問題および高度な ETL の概念を考慮します。
複合型マッピングおよび ETL パフォーマンスの向上に関する概念については、[第6章「ロード・パフォーマンスの向上」](#)を参照してください。
6. ターゲット・スキーマへ配布し、ETL ロジックを実行します。
マッピングからのコードを使用してターゲット・スキーマを準備する方法、およびその後にそのコードを実行する方法については、[第7章「ターゲット・スキーマへの配布および ETL ロジックの実行」](#)を参照してください。
7. メタデータ依存性を分析します。
[第8章「メタデータ依存性の分析」](#)を参照し、説明されているタスクを完了します。
8. 効率的な SQL を書きます。
[第9章「レポートおよび分析の SQL」](#)を参照し、説明されているタスクを完了します。
この項では、効率的な SQL を書く方法について説明されています。
9. データ・ウェアハウスをリフレッシュします。
[第10章「データ・ウェアハウスのリフレッシュ」](#)を参照し、説明されているタスクを完了します。
10. 操作を最適化します。
[第11章「データ・ウェアハウスの操作の最適化」](#)を参照し、説明されているタスクを完了します。
11. パフォーマンス・ボトルネックを排除します。
[第12章「パフォーマンス・ボトルネックの排除」](#)を参照し、説明されているタスクを完了します。
12. データ・ウェアハウスのバックアップおよびリカバリの基本の一部を確認します。
データ・ウェアハウスをバックアップおよびリカバリする方法に関するいくつかの考慮事項については、[第13章「データ・ウェアハウスのバックアップおよびリカバリ」](#)を参照してください。
13. データ・ウェアハウスのセキュリティの基本の一部を確認します。
安全なデータ・ウェアハウスを作成する方法に関するいくつかの考慮事項については、[第14章「データ・ウェアハウスの操作の最適化」](#)を参照してください。

データ・ウェアハウス管理用ツール

このガイドの目的は、すばやく効率的に Oracle データ・ウェアハウスを作成して管理できるようにすることです。次に目標を達成するために使用する製品、ツールおよびユーティリティの一部を示します。

■ Oracle Universal Installer

Oracle Universal Installer により、Oracle ソフトウェアおよびオプションをインストールします。これにより Database Configuration Assistant (DBCA) が自動的に開始され、データベースがインストールされます。

■ Oracle Enterprise Manager

データベースを管理する主要なツールは Web ベースのインターフェースである Oracle Enterprise Manager です。Oracle ソフトウェアのインストール、データベースの作成またはアップグレード、およびネットワークの構成後、データベースを管理する Oracle Enterprise Manager を使用できます。さらに、Oracle Enterprise Manager ではパフォーマンス・アドバイザのインターフェース、および SQL*Loader、Recovery Manager などの Oracle ユーティリティのインターフェースを使用できます。

■ Oracle Warehouse Builder

データ・ウェアハウスの移入および保持のための主要な製品である Oracle Warehouse Builder によって、単一の製品内での ETL、データ品質管理およびメタデータ管理機能が提供されます。

Oracle Warehouse Builder には、Oracle Database に格納される統合リポジトリが含まれています。Oracle Warehouse Builder は、Oracle Database の機能を使用して、Oracle Database ターゲットへのロードおよび Oracle Database ターゲットのメンテナンスに適したコードを生成します。

■ Database Tuning Pack

Oracle Database Tuning Pack では、データベース全体のチューニング処理を自動化する一連の新しいテクノロジーが提供されるため、データベースの管理コストが大幅に削減され、パフォーマンスと信頼性が向上します。このガイドで使用されている Oracle Database Tuning Pack の主要な機能は、SQL アクセスおよび SQL チューニング・アドバイザです。

データ・ウェアハウス・システムの設定

Oracle Data Warehouse はデータ・ウェアハウスに固有の問合せのタイプおよびデータのサイズを処理するために明示的に構成され、最適化された Oracle Database です。この項では、データ・ウェアハウス環境を最初に構成する方法を説明します。内容は次のとおりです。

- [データ・ウェアハウス・システムを設定する一般的な手順](#)
- [環境の準備](#)
- [データ・ウェアハウスのデータベースの設定](#)
- [Oracle Warehouse Builder へのアクセス](#)

データ・ウェアハウス・システムを設定する一般的な手順

次の手順では、Oracle Database をデータ・ウェアハウスとして使用できるように構成します。その後、Oracle Warehouse Builder をインストールします。Oracle Warehouse Builder は、Oracle データベースを強化し、データ管理計画を立案するためのグラフィカル・ユーザー・インタフェースを備えています。

データ・ウェアハウス・システムをセットアップする手順

1. 2-2 ページの「[環境の準備](#)」の説明に従って、ハードウェアのサイジングおよび構成を行います。
2. Oracle Database ソフトウェアをインストールします。
3. 2-7 ページの「[データ・ウェアハウスのデータベースの設定](#)」の説明に従って、データベースをデータ・ウェアハウスとして使用できるように構成します。
4. Oracle Warehouse Builder ソフトウェアにアクセスします。

Oracle Warehouse Builder は、Oracle Database の Standard Edition および Enterprise Edition にパッケージされているデータ統合製品です。

2-9 ページの「[Oracle Warehouse Builder へのアクセス](#)」の説明に従います。その後は、Oracle Warehouse Builder を使用して一般的なデータ・ウェアハウスのタスクを完了する方法の学習に役立つデモンストレーションをインストールします。

後続の手順

[第3章「データ・ソースの識別およびメタデータのインポート」](#)

環境の準備

データ・ウェアハウスのアーキテクチャの基本的な構成要素は、オンライン・トランザクション処理 (OLTP) システムと同じです。ただし、データのサイズが大規模であるため、個々の基礎単位に合わせて数量を変える必要があります。データ・ウェアハウスのサイジングは、システムに必要なスループットが起点になります。これは次のいずれか、あるいは両方になります。

- ピーク時にシステムにヒットする問合せによりアクセスされるデータ量（許容できる応答時間で）。必要とするスループットを見積もるために既存のアプリケーションからのスループット数およびエクスペリエンスを使用できる場合があります。
- 一定期間内にロードされるデータ量。

一般的に、どの特定の時間でも必要とする最も高いスループットを見積もる必要があります。

ハードウェア・ベンダーによって適切なデータ・ウェアハウス・アプリケーション設定が推奨されるため、サイジングに役立ちます。詳細は、任意のハードウェア・ベンダーにご相談ください。

適切なハードウェア設定

データ・ウェアハウスのパフォーマンスを最大限にするには、適切なサイジングと安定したハードウェア構成が必要です。次の項では、適切な構成に重要なポイントを説明します。

- [必要な CPU 数およびクロック・スピードについて](#)
- [必要なメモリーについて](#)
- [必要なディスク数について](#)
- [十分な I/O 帯域幅を判断する方法について](#)

必要な CPU 数およびクロック・スピードについて

中央演算装置 (CPU) はデータ・ウェアハウスの計算機能に使用されます。データ・ウェアハウスの操作を実行するには、十分な処理能力を持つ CPU が必要です。パラレル操作は同等のシリアル操作よりも CPU に負担がかかります。

必要とする CPU 数のガイドラインとして見積もった最も高いスループットを使用します。大まかな見積りとして、次の公式を使用します。

- $\text{<number of CPUs>} = \text{<maximum throughput in MB/s>} / 200$

つまり、CPU は 1 秒間に約 200MB まで耐えられます。たとえば、システムが 1 秒間に 1200MB の最高スループットを必要とする場合、システムは $\text{<number of CPUs>} = 1200/200 = 6$ CPU を必要とします。1 つのサーバーに 6 つの CPU を持つ設定であれば、このシステムで処理できます。2 ノード・クラスタ・システムには両方のノードに 3 つの CPU を設定できます。

必要なメモリーについて

データ・ウェアハウスのメモリーは大規模なソートなどメモリーに負担をかける処理を行う際に特に重要になります。大部分の間合せは大容量のデータにアクセスするため、データ・キャッシュへのアクセスはデータ・ウェアハウスにとっては重要ではありません。データ・ウェアハウスには OLTP アプリケーションとは異なり重要なメモリー要件がありません。

CPU 数は必要なメモリー容量のよいガイドラインになります。次の単純な公式を使用すると、選択した CPU で必要なメモリー容量を計算できます。

- $\text{<amount of memory in GB>} = 2 * \text{<number of CPUs>}$

たとえば、6 つの CPU を持つシステムが必要とするメモリー容量は $2 * 6 = 12$ GB になります。大部分の一般的なサーバーはこの要件を満たしています。

必要なディスク数について

データ・ウェアハウス環境での一般的な問題は、必要とする最大容量に基づいた記憶領域のサイジングです。記憶領域要件に排他的に基づいたサイジングは、スループットのボトルネックを作成する傾向があります。

必要なディスク・アレイの数を知るために、必要とするスループットを使用します。また、ディスク・アレイが耐えられるスループットを知るために、記憶領域プロバイダの指定を使用します。記憶領域プロバイダは GB/秒で測定され、最初のスループットの見積りは MB/秒に基づいていることに注意してください。平均ディスク制御の最大スループットは 1 秒当たり 2GB で、耐えられるスループットに換算すると $(70\% * 2 \text{ Gbit/s}) / 8 = 180 \text{ MB/s}$ となります。

必要なディスク・アレイ数を決定するには、次の公式を使用します。

- $\text{<number of disk controllers>} = \text{<throughput in MB/s>} / \text{<individual controller throughput in MB/s>}$

たとえば、1 秒間に 1200MB のシステムは、 $1200 / 180 = 7$ disk arrays となり、少なくとも 7 つのディスク・アレイが必要となります。

必要なスループットを維持するために十分な物理ディスクがあるか確認してください。ディスクのスループット数をディスク・ベンダーに尋ねてください。

十分な I/O 帯域幅を判断する方法について

エンドツーエンドの I/O システムは、CPU およびディスクを含んだコンポーネントで構成されます。バランスのよくとれた I/O システムには、I/O システムのすべてのコンポーネントに対して同等の帯域幅が必要です。これらのコンポーネントには次が含まれます。

- サーバーおよび記憶領域間のコネクタであるホスト・バス・アダプタ (HBA)。
- サーバー間およびストレージ・エリア・ネットワーク (SAN) またはネットワーク・アタッチ・ストレージ (NAS) 間のスイッチ。

- ネットワーク接続用のイーサネット・アダプタ (GigE NIC または Infiniband)。クラスタ環境では、I/O スループットのシステムのサイジング時には含めないノード間のインターコネクタ用に追加のプライベート・ポートが必要です。インターコネクタは、ノード間のパラレル実行などのファクタを考慮に入れ、個々にサイジングする必要があります。
- 個別のコンポーネントを接続するワイヤ。

各コンポーネントにはバランスのよくとれた I/O システムにするために十分な I/O 帯域幅が提供される必要があります。見積もった最初のスループットおよびベンダーによるハードウェアの指定は、必要な個別のコンポーネントの量を決定する基準になります。次の表にある変換を使用して、ベンダーによるドット区切りで表された最大スループット数をバイト数で表された耐えられるスループット数に変換します。

表 2-1 スループット・パフォーマンスの変換

| コンポーネント | バイト | 1 秒当たりのバイト数 |
|---------------|------------|-------------|
| HBA | 2G ビット | 200MB |
| 16 ポート・スイッチ | 8 × 2G ビット | 1200MB |
| Fibre Channel | 2G ビット | 200MB |
| GigE NIC | 1G ビット | 80MB |
| Inf-2G ビット | 2G ビット | 160MB |

十分な I/O 帯域幅を確保するために十分なコンポーネントを持つには、ディスクでのデータのレイアウトも成功と失敗を分ける鍵となります。すべてのディスク・アレイに対して十分なスループットを持つシステムを設定し、取得しようとする問合せの全データが 1 つのディスクに格納されている場合、1 つのディスクがボトルネックとなるため、必要なスループットを確保することができません。このような状況を避けるためには、できるかぎり多くのディスク、理想はすべてのディスクへデータをストライプ化します。1MB に対して 256KB のストライプのデータは、マルチブロック読み取り操作および複数のディスクへのデータの振り分け間でのよいバランスを提供します。

自動ストレージ管理の概要 (ASM)

ASM はディスク・グループのディスクに対してデータをストライプ化するために使用できる Oracle Database のコンポーネントです。ASM は全ディスクにおいてバランスのとれたデータを保証します。ディスクは、ASM の実行中に追加または削除でき、すべての使用可能なディスクに対して記憶領域を自動的にリバランスします。ASM はファイル・システムのデータのミラー、およびディスク障害の場合データ損失の回避のために使用できます。ASM のデフォルトのストライプ・サイズは 1MB です。ストライプ・サイズは 128KB まで下げることができます。

ASM を使用しないでストレージ操作を実行することもできますが、失敗する可能性が高くなります。そのため、ASM の使用をお勧めします。

データ・ウェアハウスのハードウェア構成の検証

Oracle Database のインストールの前に、ハードウェアおよびオペレーティング・システムのレベルで設定を検証する必要があります。オペレーティング・システムで必要なスループットおよびパフォーマンスが実行されない場合、Oracle Database はユーザー要件に合わせて実行できなくなります。スループットを検証する 2 つのツールは dd ユーティリティおよび Orion です。

dd ユーティリティについて

UNIX または Linux でオペレーティング・システムのスループットを検証する基本方法は、dd ユーティリティを使用することです。dd ユーティリティは、データ・ブロックをディスクから直接読み取る基本方法で、関連するオーバーヘッドがほとんどないので、dd ユーティリティからの出力は信頼性の高い測定です。Oracle Database は dd ユーティリティが実現する最大スループットの約 90 パーセントに到達します。

例：dd ユーティリティの使用

dd ユーティリティを使用する手順

次は、dd の使用に最も重要なオプションです。

```
bs=BYTES: Read BYTES bytes at a time; use 1 MB
count=BLOCKS: copy only BLOCKS input blocks
if=FILE: read from FILE; set to your device
of=FILE: write to FILE; set to /dev/null to evaluate read performance;
        write to disk would erase all existing data!!!
skip=BLOCKS: skip BLOCKS BYTES-sized blocks at start of input
```

Oracle Database が達成可能な最大スループットを見積もるため、典型的なデータ・ウェアハウス・アプリケーションのワークロード（大規模かつ、ランダムで連続したディスク・アクセスによる）を模倣します。

次の dd コマンドは合計 2GB を読み取る 2 つのデバイスに対してランダムで連続したディスク・アクセスを実行します。スループットは次のコマンドを終了するのに必要な時間によって分けられた 2GB です。

```
dd bs=1048576 count=200 if=/raw/data_1 of=/dev/null &
dd bs=1048576 count=200 skip=200 if=/raw/data_1 of=/dev/null &
dd bs=1048576 count=200 skip=400 if=/raw/data_1 of=/dev/null &
dd bs=1048576 count=200 skip=600 if=/raw/data_1 of=/dev/null &
dd bs=1048576 count=200 skip=800 if=/raw/data_1 of=/dev/null &
dd bs=1048576 count=200 if=/raw/data_2 of=/dev/null &
dd bs=1048576 count=200 skip=200 if=/raw/data_2 of=/dev/null &
dd bs=1048576 count=200 skip=400 if=/raw/data_2 of=/dev/null &
dd bs=1048576 count=200 skip=600 if=/raw/data_2 of=/dev/null &
dd bs=1048576 count=200 skip=800 if=/raw/data_2 of=/dev/null &
```

検証にはデータベースの記憶領域に含む予定のすべての記憶領域装置を含める必要があります。クラスタ環境の構成時に各ノードから dd コマンドを実行する必要があります。

Orion ユーティリティについて

Orion はスループットを測定するためのシステムで、Oracle のようなワークロードを模倣するためにオラクル社が提供するツールです。dd ユーティリティと比較すると、Orion には次の利点があります。

- Orion のシミュレーションはデータベースが生成するワークロードに似ています。
- Orion を使用すると、1 つのシミュレーションで高い信頼性の書取りおよび読取りシミュレーションを実行できます。

オラクル社はデータベースがすでにインストールされている場合でも、最大の達成可能なスループットを検証するために Orion の使用をお勧めします。

次はサポートされる I/O ワークロードのタイプです。

- 小規模およびランダム
- 大規模および連続的
- 大規模およびランダム
- 混合ワークロード

各タイプのワークロードで Orion は 1 秒当たりの MB、1 秒当たりの I/O および I/O 待機時間などのパフォーマンス・メトリクスを計測するための I/O ロードの異なるレベルで検証を実行できます。データ・ウェアハウス・ワークロードは複数のプロセスによって発行された連続する I/O スループットによって一般的に特徴づけられています。構築するシステムのタイプに応じて異なる I/O シミュレーションを実行できます。例は次のとおりです。

- ユーザーまたはアプリケーションがシステムに問い合わせる場合の毎日のワークロード
- ユーザーがシステムにアクセスするまたはしない場合のデータ・ロード
- 索引およびマテリアライズド・ビューの構築
- バックアップ操作

Orion ソフトウェアをダウンロードする場合にブラウザでカーソルを合せるアドレス

<http://www.oracle.com/technology/software/tech/orion/index.html>

Orion はベータ版なのでサポートされていないことに注意してください。

例 : Orion ユーティリティの使用

Orion を起動する手順は、次のとおりです。

```
$ orion -run simple -testname mytest -num_disks 8
```

標準出力は次のとおりです。

```
Orion VERSION 10.2
```

Command line:

```
-run advanced -testname orion14 -matrix point -num_large 4 -size_large 1024  
-num_disks 4 -type seq -num_streamIO 8 -simulate raid0 -cache_size 0 -verbose
```

This maps to this test:

```
Test: orion14  
Small IO size: 8 KB  
Large IO size: 1024 KB  
IO Types: Small Random IOs, Large Sequential Streams  
Number of Concurrent IOs Per Stream: 8  
Force streams to separate disks: No  
Simulated Array Type: RAID 0  
Stripe Depth: 1024 KB  
Write: 0%  
Cache Size: 0 MB  
Duration for each Data Point: 60 seconds  
Small Columns:, 0  
Large Columns:, 4  
Total Data Points: 1
```

```
Name: /dev/vx/rdisk/asm_vol1_1500m    Size: 1572864000  
Name: /dev/vx/rdisk/asm_vol2_1500m    Size: 1573912576  
Name: /dev/vx/rdisk/asm_vol3_1500m    Size: 1573912576  
Name: /dev/vx.rdisk/asm_vol4_1500m    Size: 1573912576  
4 FILES found.
```

```
Maximum Large MBPS=57.30 @ Small=0 and Large=4
```

この例では、特定のワークロードの最大スループットは 1 秒当たり 57.30MB です。

データ・ウェアハウスのデータベースの設定

環境設定および Oracle Database ソフトウェアのインストール後、データベース・パラメータが正しく設定されていることを確認します。設定する必要があるデータベース・パラメータは多くありません。

一般的なガイドラインとして、特に理由がないかぎり、データベース・パラメータは変更しないでください。データ・ウェアハウスの設定には、Oracle Enterprise Manager を使用します。様々なパラメータ設定を表示するには、「データベース」ページに移動して「サーバー」をクリックします。「データベース構成」で「メモリー・パラメータ」または「すべての初期化パラメータ」をクリックします。

メモリー管理パラメータの設定方法

上位レベルには、2つのメモリー・セグメントがあります。

- 共有メモリー: システム・グローバル領域 (SGA) と呼ばれます。Oracle インスタンスで使用されるメモリーです。
- セッションベース・メモリー: プログラム・グローバル領域 (PGA) と呼ばれます。データベース・セッションに使用されるメモリーです。ソートや集計などのデータベース操作の実行に使用されます。

Oracle Database は2つのメモリー領域のメモリー・コンポーネントの配布を自動的にチューニングします。そのため、設定する必要があるのは次のパラメータのみです。

- SGA_TARGET

SGA_TARGET パラメータは共有メモリーに分配するメモリー容量です。データ・ウェアハウスの場合、SGA は PGA によって消費される合計メモリーと比べると比較的小さい容量です。まずは、合計メモリーの 25 パーセントを Oracle Database から SGA へ分配します。SGA の最小値は 100MB です。

- PGA_AGGREGATE_TARGET

PGA_AGGREGATE_TARGET パラメータはすべてのセッションの合計 PGA が消費するターゲット・メモリー容量です。まずは、次の公式を使用して PGA_AGGREGATE_TARGET 値を定義します。

- $PGA_AGGREGATE_TARGET = 3 * SGA_TARGET$ メモリーに合う
PGA_AGGREGATE_TARGET に対して十分な物理メモリーがない場合、PGA_AGGREGATE_TARGET を下げます。

- MEMORY_TARGET および MEMORY_MAX_TARGET

MEMORY_TARGET パラメータによってターゲットのメモリー・サイズを設定でき、関連した初期化パラメータ MEMORY_MAX_TARGET によって最大ターゲット・メモリー容量を設定できます。データベースは、システム・グローバル領域 (SGA) および集計プログラム・グローバル領域 (PGA) 間の要件に従ってメモリーを再分配し、ターゲット・メモリー・サイズに対してチューニングします。ターゲット・メモリー初期化パラメータは動的であるため、データベースの再起動なしにいつでもターゲット・メモリー・サイズを変更できます。最大メモリー・サイズは上限として機能するため、ターゲット・メモリー・サイズを非常に高い値に設定してしまふことはありません。また特定の SGA コンポーネントが簡単に縮小できないこと、または最小サイズで保持される必要があるため、ターゲット・メモリー・サイズを過度に小さく設定してしまふことも防止します。

例: 初期化パラメータの設定

次で示すように、ALTER SYSTEM の発行により初期化パラメータの設定ができます。

```
ALTER SYSTEM SET SGA_TARGET = 1024M;
```

重要な他の初期化パラメータの設定

初めてデータ・ウェアハウスを使用する際は、Database Configuration Assistant (DBCA) の実行時に選択できるデータ・ウェアハウス・テンプレート・データベースが便利です。ただし、次の初期化パラメータを考慮すれば、どのデータベースでもかまいません。

■ COMPATIBLE

COMPATIBLE パラメータは以前のリリースに対してデータベースが持つ互換性のレベルを識別します。最新機能を活用するには、COMPATIBLE パラメータでデータベースのリリース番号に設定します。

■ OPTIMIZER_FEATURES_ENABLE

上位のコストベースのオプティマイザ機能を活用するには、このパラメータがデータベースの現行バージョンの値に設定されているか確認します。

■ DB_BLOCK_SIZE

デフォルト値の 8KB は、データ・ウェアハウスのほとんどの要件に適しています。表圧縮を使用する場合は、より大きなブロック・サイズを指定することを検討してください。

■ DB_FILE_MULTIBLOCK_READ_COUNT

DB_FILE_MULTIBLOCK_READ_COUNT パラメータを使用すると、単一オペレーティング・システムの読取りコールでデータベース・ブロックの読取りができます。データ・ウェアハウスの一般的なワークロードは多くの連続した I/O で構成されているため、多数の小さい I/O ではなく、少数の大きい I/O を活用できるようにします。このパラメータの設定時、ブロック・サイズおよびオペレーティング・システムの最大 I/O サイズを考慮に入れて、次の公式を使用します。

- $DB_FILE_MULTIBLOCK_READ_COUNT * DB_BLOCK_SIZE = <maximum\ operating\ system\ I/O\ size>$

最大のオペレーティング・システムの I/O サイズは 64KB から 1MB の間で変化します。

■ PARALLEL_MAX_SERVERS

PARALLEL_MAX_SERVERS パラメータは、パラレル実行の使用可能な最大プロセス数のリソース制限を設定します。パラレル操作には、操作のあらゆる表に属性を付与する最大並列度 (DOP) として問合せサーバーのプロセス数が最大 2 倍必要です。

Oracle Database はほとんどのシステムにとって十分なデフォルト値を

PARALLEL_MAX_SERVERS パラメータに設定します。PARALLEL_MAX_SERVERS のデフォルト値は次のとおりです。

- $(CPU_COUNT * PARALLEL_THREADS_PER_CPU * (2\ if\ PGA_AGGREGATE_TARGET > 0; otherwise\ 1) * 5)$

上位の DOP 属性を持つ表のパラレル問合せにはこれは十分ではない可能性があります。オラクル社は上位の DOP の問合せを実行するユーザーには、PARALLEL_MAX_SERVERS を次のように設定することをお勧めします。

- $2 * DOP * NUMBER_OF_CONCURRENT_USERS$

たとえば、PARALLEL_MAX_SERVERS パラメータに 64 を設定することで、各問合せは各セットの 8 つの DOP を持つ 2 つのスレーブ・セットを使用することを想定し、同時に 4 つのパラレル問合せを実行できます。

ハードウェア・システムが CPU バインドでも I/O バインドでもない場合、問合せサーバー・プロセスを追加することによって、システムに同時パラレル実行ユーザーの数を増やすことができます。一方システムが CPU バインドまたは I/O バインドになる場合、同時ユーザーの追加はすべてのパフォーマンスに弊害をもたらします。

PARALLEL_MAX_SERVERS パラメータを慎重に設定することが、同時パラレル操作数の制限に効果的な方法です。

- **PARALLEL_ADAPTIVE_MULTI_USER**
TRUE または FALSE になる **PARALLEL_ADAPTIVE_MULTI_USER** パラメータは、現行のワークロードに応じて特定の文の並列度を動的に決定するためのアルゴリズムをサーバーが使用するかどうかを定義します。この機能を活用するには、**PARALLEL_ADAPTIVE_MULTI_USER** を TRUE に設定します。
- **QUERY_REWRITE_ENABLED**
マテリアライズド・ビューに対して問合せのリライトを活用するには、このパラメータを TRUE に設定する必要があります。このパラメータのデフォルトは TRUE です。
- **QUERY_REWRITE_INTEGRITY**
QUERY_REWRITE_INTEGRITY パラメータのデフォルトは ENFORCED です。これは、有効で検証された主キー制約、一意キー制限、および外部キー制限でパラメータ自体が基準になる場合、データベースが完全に最新のマテリアライズド・ビューのみに対して問合せを上書きするということです。

TRUSTED モードでは、オプティマイザはマテリアライズド・ビューのデータが現行であり、ディメンションで宣言された階層関係および RELY 制限が正確であるものとみなします。
- **STAR_TRANSFORMATION_ENABLED**
最適化されたスター型変換を活用するには、このパラメータを TRUE に設定します。

Oracle Warehouse Builder へのアクセス

Oracle Warehouse Builder は、従来のデータ・ウェアハウスを含む様々なタイプのデータ管理計画の設定および配布が可能な柔軟なツールです。

Oracle Warehouse Builder を有効にする手順

1. Oracle Database 11g の Enterprise Edition または Standard Edition のいずれかにアクセスしていることを確認します。

Oracle Database 11g には、Oracle Warehouse Builder サーバー・コンポーネントが事前にインストールされています。また、Oracle Warehouse Builder リポジトリのスキーマも含まれています。
2. Oracle Database 11g にインストールされているデフォルトの Oracle Warehouse Builder スキーマを使用するには、まず、スキーマのロックを解除します。

SYS または SYSDBA ユーザーとして SQL*Plus に接続します。次のコマンドを実行します。

SQL> ALTER USER OWBSYS ACCOUNT UNLOCK;

SQL> ALTER USER OWBSYS IDENTIFIED BY *owbsys_passwd*;
3. Warehouse Builder のデザイン・センターを起動します。

Windows の場合、「スタート」→「プログラム」→「Oracle」→「Warehouse Builder」を選択後、「デザイン・センター」を選択します。

UNIX の場合、*owb_home/owb/bin/unix* に移動し、*owbclient.sh* を実行します。
4. ワークスペースを定義し、ユーザーをワークスペースに割り当てます。

単一の Oracle Warehouse Builder リポジトリでは、関連付けられているオブジェクト上で作業している一連のユーザーに対応している各ワークスペースで、複数のワークスペースを定義できます。たとえば、開発、テストおよび本番などの各環境に対してワークスペースを作成できます。

簡略化するために、MY_WORKSPACE という 1 つのワークスペースを作成し、ユーザーを割り当てます。

「デザイン・センター」ダイアログ・ボックスで、「詳細を表示」をクリックした後、「ワークスペース管理」をクリックします。

「リポジトリ・アシスタント」が表示されます。

プロンプトに従い、「リポジトリ・アシスタント」でデフォルト設定を受け入れ、ワークスペースを作成し、ユーザーをワークスペース所有者として割り当てます。

5. 作成したユーザー名およびパスワードで「デザイン・センター」にログインします。

Oracle Warehouse Builder デモンストレーションのインストール

この後の項では、複数のフラット・ファイル・ソースのデータの統合、データの変換および新規リレーショナル・ターゲットへのロード方法について、演習を使用して説明します。このガイドで提供されている演習を実行するには、Oracle Warehouse Builder デモンストレーションをダウンロードする必要があります。このデモンストレーションでは、製品への理解を深めるために多様な Oracle Warehouse Builder オブジェクトを作成するフラット・ファイル・データおよびスクリプトが用意されています。

このガイドで提供されている Oracle Warehouse Builder の演習を実行する手順

1. デモンストレーションをダウンロードします。

デモンストレーションは、owb_demo.zip と呼ばれる zip ファイル内の一連のファイルで構成されています。次のリンクで取得できます。

http://www.oracle.com/technology/obe/admin/owb10gr2_gs.htm

zip ファイルには、SQL スクリプト、値の形式がカンマ区切りされた 2 つのソース・ファイルおよび Tcl で書かれた 19 個のスクリプトが含まれます。

2. スクリプト owbdemoinit.tcl を編集します。

スクリプト owbdemoinit.tcl によって、他の TCL スクリプトで使用される変数を定義および設定します。次の変数を編集して、コンピュータ環境の値と一致させます。

- set temp space TEMP
- set owbclientpwd workspace_owner
- set sysuser sys
- set syspwd pwd
- set host hostname
- set port portnumber
- set service servicename
- set project owb_project_name
- set owbclient workspace_owner
- set sourcedir drive:/newowbdemo
- set indexspace USERS
- set dataspace USERS
- set snapspace USERS
- set sqlpath drive:/oracle/11.1.0/db_1/BIN
- set sid servicename

3. Warehouse Builder スクリプト・ユーティリティおよび OMB Plus から Tcl スクリプトを実行します。

Windows の場合、「スタート」→「プログラム」→「Oracle」→「Warehouse Builder」を選択後、「OMB Plus」を選択します。

UNIX の場合、`owb_home/owb/bin/unix` に移動し、`OMBPlus.sh` を実行します。

OMB+> プロンプトでは、次のコマンドを入力してスクリプトを含むディレクトリに変換します。

```
cd drive:\\newowbdemo\\
```

次のコマンドを入力して、すべての Tcl スクリプトを必要な順序で実行します。

```
source loadall.tcl
```

4. デザイン・センターを起動し、スクリプト `owbdemoinit.tcl` で指定した資格証明を使用して、ワークスペース所有者としてログインします。
5. デモンストレーションに従い、Oracle Warehouse Builder クライアントが正常に設定されたことを確認します。

デザイン・センターでは、接続エクスプローラ内の右側にある「場所」ノードを展開します。「データベース」の展開後、「Oracle」を展開します。「Oracle」ノードには次の場所が含まれます。

```
OWB_REPOSITORY
```

```
SALES_WH_LOCATION
```

Oracle Warehouse Builder デモンストレーションを正常にインストールすると、デザイン・センターが `EXPENSE_WH` という Oracle モジュールとともに表示されます。

データ・ソースの識別およびメタデータの インポート

一般的に、データ・ウェアハウスのソース・システムは通常、トランザクション処理のアプリケーションです。たとえば、売上分析のデータ・ウェアハウスは通常、現行のオーダー・アクティビティを記録しているオーダー・エントリ・システムからデータを抽出します。

抽出プロセスの設計には問題が発生する場合があります。ソース・システムが複雑で文書化が不完全な場合、抽出するデータの決定は困難です。また、通常ソース・システムは変更できず、そのパフォーマンスまたは可用性も調整されません。これらの問題を処理するには、まずメタデータをインポートしてください。

メタデータはデータ・セットに与えられたオブジェクトのコンテンツを示すデータです。たとえば、表のメタデータは各列のデータ型を示します。

Oracle Database を使用する場合は、メタデータのインポートのためのツールとして Oracle Warehouse Builder をお薦めします。Oracle Warehouse Builder にメタデータをインポートした後、メタデータに注釈を付けて、さらにトランザクション処理のアプリケーションから抽出計画を個別に設計できます。

この項では、Oracle Warehouse Builder を使用してメタデータをインポートする方法を説明します。内容は次のとおりです。

- ソースからメタデータをインポートする一般的な手順
- Oracle Warehouse Builder のワークスペース、プロジェクトおよびその他のデバイスについて
- 例：フラット・ファイルからのメタデータのインポート

ソースからメタデータをインポートする一般的な手順

メタデータをインポートする手順

1. 2-9 ページの「[Oracle Warehouse Builder へのアクセス](#)」の指示に従います。
2. 2-10 ページの「[Oracle Warehouse Builder デモンストレーションのインストール](#)」の説明に従って Oracle Warehouse Builder デモンストレーションをダウンロードおよびインストールします。
3. Oracle Warehouse Builder のプロジェクトを指定します。
詳細は、3-2 ページの「[Oracle Warehouse Builder のワークスペース、プロジェクトおよびその他のデバイスについて](#)」を参照してください。
4. 3-3 ページの「[例:フラット・ファイルからのメタデータのインポート](#)」に従います。
この例では、ソース・ファイルの場所を指定する方法、および「メタデータのインポート」ウィザードを起動する方法を説明します。ファイル、表およびビューなどのデータ・オブジェクトをインポートするプロセスは同じです。したがって、この例を完了すると、Oracle Warehouse Builder にすべてのデータ・オブジェクトをインポートする一般的な方法を理解したことになります。

後続の手順

メタデータをインポートすることによりソースを正しく定義した後は、[第4章「Oracle Warehouse Builder でのウェアハウスの定義](#)」の説明に従って、ターゲット・スキーマを設計します。

この時点では、メタデータのインポートが完了しているのみでデータを抽出していません。このガイドの後続の項でデータの抽出計画を設計および実装します。

Oracle Warehouse Builder のワークスペース、プロジェクトおよびその他のデバイスについて

Oracle Warehouse Builder デモンストレーションをインストールして「デザイン・センター」を起動した後、ワークスペースにログインします。「デザイン・センター」の上部にユーザー名およびワークスペース名が表示されます。

これに関する図と今後のポイントは、次のリンクを参照してください。

http://www.oracle.com/technology/obe/10gr2_owb/owb10gr2_gs/owb/lesson1/start01.htm#

ワークスペースには、関連するプロジェクトで作業する一連のユーザーが含まれます。作成するワークスペースの数を決定するためにセキュリティを考慮することは重要です。共通のモデルは、開発、テストおよび本番に個別のワークスペースを作成することです。このモデルを使用して、開発者などのユーザーによるワークスペースの開発およびテストへのアクセスを可能にしますが、本番ワークスペースへのアクセスは制限されます。

オプションで、ワークスペースをプロジェクトに分割できます。ただし実際は、ワークスペースには通常、アクティブなプロジェクトが1つのみ含まれます。これは、プロジェクトが単純に入れ物であり、セキュリティの実装またはサブジェクト指向のグループ化の確立には適さないためです。セキュリティはワークスペースを介して実装されます。後で説明するように、サブジェクト指向のグループ化の確立はモジュールを介して完了できます。

プロジェクトの目的は、権限に関連するメタデータのセットを含めることです。したがってデータ・ウェアハウスでも、ソースおよびターゲットのすべてのメタデータを同一のプロジェクトに含めます。さらに権限の実装に必要なすべての機能、プロシージャ、変換、マッピングおよびその他のオブジェクトを含めます。プロジェクトには、Oracle Warehouse Builder に作成またはインポートできるオブジェクトの各タイプのノードが含まれます。作成またはインポートできるオブジェクトのタイプの全般を理解するために異なるノードを展開します。

このデモンストレーションでは、プロジェクト・エクスプローラが左側に表示され、2つのプロジェクトが含まれます。MY_PROJECT はデフォルトで、事前にシードされたプロジェクトです。ワークスペース内の単一のアクティブなプロジェクトとして MY_PROJECT を使用できます。このデモンストレーションでは、OWB_DEMO プロジェクトで作業します。

例：フラット・ファイルからのメタデータのインポート

この例では、フラット・ファイルからメタデータをインポートする方法を示します。具体的には、2つのファイル export.csv および expense_categories.csv などの OWB_DEMO プロジェクトにメタデータをインポートして、「ファイル」ノードの下のプロジェクト・エクスプローラに表示することが目的です。

フラット・ファイルからメタデータをインポートする手順

1. 3-3 ページの「[フラット・ファイルの場所の指定](#)」の説明に従って、フラット・ファイルの場所を指定します。
2. 3-4 ページの「[プロジェクトでのモジュールの作成](#)」の説明に従って、OWB_DEMO を編成し、入力フラット・ファイルのメタデータを受け取ります。
3. 3-4 ページの「[インポート・メタデータ・ウィザード](#)」の起動の説明に従って、インポートするファイルを指定します。
4. 3-5 ページの「[フラット・ファイル・サンプル・ウィザード](#)」の使用の説明に従って、メタデータの構成を指定します。
5. 3-5 ページの「[フラット・ファイル・データのインポート](#)」の説明に従って、両方のフラット・ファイルのメタデータをインポートします。

後続の手順

メタデータをインポートすることによりソースを正しく定義した後は、[第4章「Oracle Warehouse Builder でのウェアハウスの定義」](#)の説明に従って、ターゲット・スキーマを設計します。

この時点では、メタデータのインポートが完了しているのみでデータを抽出していません。このガイドの後続の項でデータの抽出計画を設計および実装します。

フラット・ファイルの場所の指定

フラット・ファイルがある場所を指定します。

「デザイン・センター」では、右側の部分が[接続エクスプローラ](#)と呼ばれるエクスプローラで、「場所」と呼ばれるノードが含まれます。「場所」ノードを使用してソース・データがある場所を指定します。

Oracle Warehouse Builder からアクセスできるソースおよびターゲットのタイプの全般を理解するために「場所」ノードおよびその中にあるノードを開きます。

この例では、「ファイル」を右クリックし、「新規」を選択してフラット・ファイルの場所を定義します。

ファイル・システムの場所を作成するダイアログ・ボックスのプロンプトに従います。定義する各場所はコンピュータのファイル・システムの特定のディレクトリに対応します。したがって、ドライブおよびディレクトリに基づいた場所の名前を付けるように考慮します。このデモンストレーションでは、場所に C_NEWOWBDEMO_SOURCEFILES という名前を付けます。

プロジェクトでのモジュールの作成

プロジェクト・エクスプローラで、OWB_DEMO を編成し、入力フラット・ファイルのメタデータを受け取ります。

データ・ウェアハウスの実装では、多数のソース・オブジェクトおよびターゲット・オブジェクトを所有する可能性があります。これらの様々なオブジェクトを編成する方法として、Oracle Warehouse Builder ではモジュールを作成する必要があります。各モジュールは、接続エクスプローラに作成する場所に対応します。

この例では、2つのフラット・ファイルに対して1つの場所のみ所有するので、プロジェクト・エクスプローラに1つのモジュールを作成します。OWB_DEMO の下の「ファイル」ノードを右クリックし、「新規」を選択します。新規のモジュールに SALES_EXPENSES という名前を付けます。この場所には、前のステップで定義した場所、C_NEWOWBDEMO_SOURCEFILES を指定します。

「インポート・メタデータ・ウィザード」の起動

「インポート・メタデータ・ウィザード」を起動します。

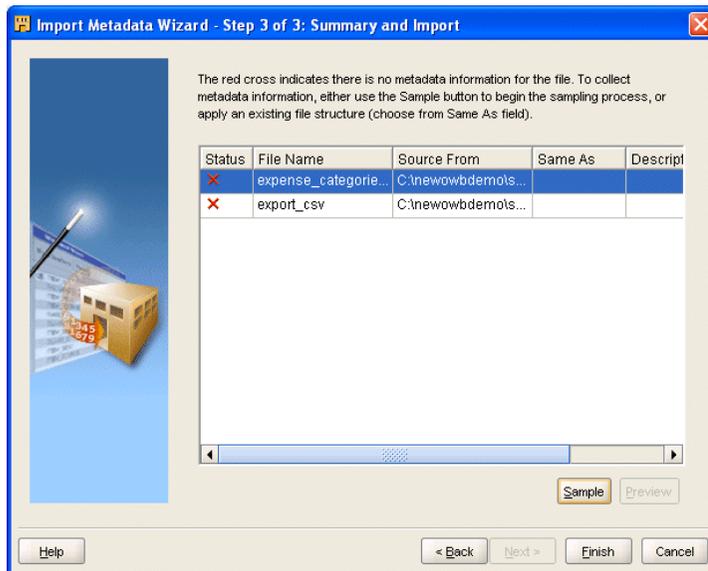
モジュール「SALES_EXPENSES」を右クリックし、「新規」を選択して、「インポート・メタデータ・ウィザード」内のプロンプトに従います。ウィザード内のプロンプトは選択したモジュールのタイプによって異なり、したがってインポートするデータ・オブジェクトによっても異なります。

この例では、2つのフラット・ファイルのインポートを選択しました。「インポート・メタデータ・ウィザード」の「サマリー」ページで、ファイルの1つを選択し、次に「サンプル」を選択して「フラット・ファイル・サンプル・ウィザード」を起動します。

次のステップでは、順番に各ファイルのサンプルを作成し、このページの「終了」を選択してメタデータをインポートします。

図 3-1 に、「インポート・メタデータ・ウィザード」の「サマリーとインポート」ページの「サンプル」を示します。

図 3-1 インポート・メタデータ・ウィザード



「フラット・ファイル・サンプル・ウィザード」の使用

「フラット・ファイル・サンプル・ウィザード」のプロンプトに従って、メタデータの構成を指定します。

サンプルに指定する文字の数に基づいて、ウィザードによりフラット・ファイル・データが読み込まれ、メタデータの構成のための提案が提供されます。サンプルのサイズが小さすぎる場合は、ウィザードによりデータが誤って読み込まれ、無効な提案が作成される可能性があります。それに従って、ウィザードの設定を変更および調整できます。

この例の目的のため、ウィザードによりファイルが区切られないことが正しく決定され、シングル・レコード・タイプが含まれます。またキャラクタ・セットは WE8MSWIN1252 です。「フラット・ファイル」ウィザードにあるすべてのデフォルト設定を受け入れます。

ウィザードによりサンプリングできる様々なタイプのファイルに精通するには、ウィザード・ページのオプションを確認してさらに知識を深めるために「ヘルプ」を選択します。

最初のフラット・ファイルをサンプリングした後、「インポート・メタデータ・ウィザード」のサマリーおよびインポートに関するページに戻り、2番目のファイルをサンプリングします。

前のファイルで行ったように、「フラット・ファイル」ウィザードのデフォルト設定を受け入れます。

フラット・ファイル・データのインポート

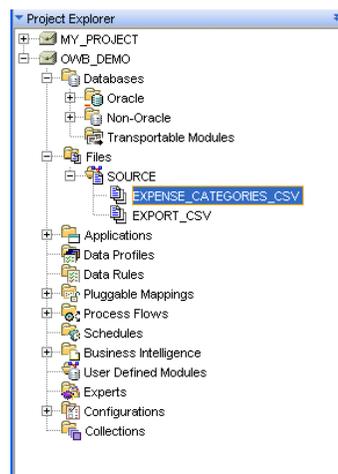
両方のフラット・ファイルのメタデータをインポートします。

サマリーおよびインポートに関するページに再度戻り、「終了」を選択します。

「終了」を選択する場合、ウィザードにより、データのサンプリング時に作成した選択基準に基づいてデータがインポートされます。OWB_DEMO プロジェクトの「ファイル」ノードの下にある SALES_EXPENSES モジュールの下に、2つのカンマ区切りのファイルが表示されます。

図 3-2 に、SOURCE モジュールにインポートされた後の 2つのファイルを示します。

図 3-2 SOURCE フラット・ファイル・モジュール



Oracle Warehouse Builder での ウェアハウスの定義

Oracle Warehouse Builder を使用することにより、リレーショナルまたはディメンションのウェアハウスを設計できます。

Oracle Warehouse Builder では、物理的な実装からディメンションの設計が明示的に区別されます。簡単なクリック操作で、ディメンション・オブジェクトに対してリレーショナル実装または多次元実装のいずれかを選択できます。したがって、リレーショナル・ターゲット・ウェアハウスまたは多次元ウェアハウスとして同じディメンション・オブジェクトを実装できます。

この項では、リレーショナル・ターゲット・ウェアハウスとして実装されるディメンション・モデルを設計する方法を示します。1つのキューブと2つのディメンションで構成される小規模のデータ・ウェアハウスをモデル化します。複合のスノーフレーク・スキーマをモデル化するために Oracle Warehouse Builder を使用できますが、このデモンストレーションの目的のために、2つのディメンションへの外部キー参照がある単一のキューブで構成される1つのスター・スキーマをモデル化します。

この章の内容は次のとおりです。

- [リレーショナル・ターゲット・ウェアハウスを定義する一般的な手順](#)
- [ウェアハウス・ターゲット・スキーマの指定](#)
- [Oracle Warehouse Builder のフラット・ファイルのソースについて](#)
- [ディメンションについて](#)

リレーショナル・ターゲット・ウェアハウスを定義する一般的な手順

リレーショナル・ターゲット・ウェアハウスを定義する手順

1. 4-2 ページの「ウェアハウス・ターゲット・スキーマの指定」の説明に従って、スキーマをウェアハウス・ターゲット・スキーマとして指定します。

2. ウェアハウス・ターゲット・モジュールにソース・オブジェクトおよびターゲット・オブジェクトを定義またはインポートします。

一般的に、ウェアハウス・ターゲット・モジュール内のいずれかのノードを右クリックし、「新規」または「インポート」のどちらかを選択します。Oracle Warehouse Builder により、説明のための適切なウィザードが起動されます。追加情報は、「ヘルプ」をクリックします。

ウェアハウス・ターゲット・モジュールに追加するオブジェクトのタイプはソース・データおよびデータ・ウェアハウスで実行するものの性質に依存します。

このガイドで提供されている演習を続行するには、4-3 ページの「演習：ターゲット・モジュールへの外部表の追加」および 4-4 ページの「演習：ディメンションの理解」を参照してください。

3. 必要に応じて、ソース・オブジェクトおよびターゲット・オブジェクトを構成します。

いくつかのオブジェクトでは追加構成が必要です。ウェアハウス・モジュールに単一のオブジェクトをインポートまたは定義した後、右クリックして「構成」を選択し、設定を確認して必要に応じて変更します。

後続の手順

ターゲット・ウェアハウスを正常に定義した後、次にソース・データを抽出し、必要に応じて変換し、それをウェアハウスにロードする計画を設計します。

ETL 計画の設計については、第 5 章「ETL ロジックの定義」を参照してください。

ウェアハウス・ターゲット・スキーマの指定

従来のデータ・ウェアハウスの実装では、データ・ウェアハウスのターゲットであるターゲット・スキーマは通常 1 つのみです。ただし、Oracle Warehouse Builder では、定義したすべてのユーザーにターゲット・スキーマになる可能性があります。したがって、データ・ウェアハウスのターゲット・スキーマになるスキーマを指定する必要があります。

データ・ウェアハウスのターゲット・スキーマとしてスキーマを指定する手順

1. Oracle Warehouse Builder にスキーマを登録します。

「グローバル・エクスプローラ」パネルで、「セキュリティ」ノードを開きます。

「ユーザー」ノードを右クリックし、「新規」を選択します。

「ユーザーの作成」ダイアログ・ボックスで、「DB ユーザーの作成」を選択し、プロンプトに従います。さらに情報が必要な場合は、「ヘルプ」をクリックします。

このデモンストレーションでは、新規スキーマを作成し、これを EXPENSE_WH と呼びます。

2. 新規スキーマの位置情報を指定します。

接続エクスプローラを右クリックし、「Oracle」ノードの下の「場所」から「新規」を選択します。

「EXPENSE_WH_LOCATION」という場所を作成します。接続をテストするためのオプションを選択します。

- プロジェクト・エクスプローラで、単一のモジュールとスキーマの場所を関連付けます。
3-3 ページの「例:フラット・ファイルからのメタデータのインポート」で、メタデータのインポート元の場所に対応するモジュールを作成しました。同様に、ターゲット・スキーマの場所に対応するモジュールを作成する必要があります。
OWB_DEMO プロジェクトでは、「データベース」ノードを開き、「新規」を選択します。モジュールを作成するウィザードのプロンプトに従います。モジュールのステータスを Warehouse Target として指定することを確認します。
このデモンストレーションでは、モジュールに EXPENSE_WH という名前を付けます。
- 新規のデータ・ウェアハウスのターゲット・スキーマを習熟します。
プロジェクト・エクスプローラで、新しく定義されたウェアハウス・ターゲット・モジュールのためのノードを開きます。モジュール内に定義またはインポートできるタイプのオブジェクトなど、様々なタイプのオブジェクトがノードの下にリストされます。

Oracle Warehouse Builder のフラット・ファイルのソースについて

ターゲット・モジュールに追加するオブジェクトのタイプは、後で設計する ETL ロジックに影響を与える場合があります。データ・ソースがフラット・ファイルに由来する場合は、SQL*Loader コードまたは SQL コードのどちらを生成するか選択できます。コードの各タイプには、固有の利点があります。簡潔に言うと、SQL*Loader は大量のデータの処理により適している一方、SQL ではより広範囲の複合結合および変換が可能です。

Oracle Warehouse Builder で SQL*Loader を使用するには、3-3 ページの「例:フラット・ファイルからのメタデータのインポート」の説明に従ってフラット・ファイルをインポートします。ただし、SQL を使用するには、4-3 ページの「演習:ターゲット・モジュールへの外部表の追加」の説明に従ってウェアハウス・モジュールに外部表を定義する必要があります。

外部表は、Oracle Database (バージョン 9i 以上) 内のデータベース・オブジェクトです。他のデータベース型または 9i リリースより前の Oracle Database では外部表を使用できません。

演習:ターゲット・モジュールへの外部表の追加

外部表は、リレーショナル・フォーマット内のフラット・ファイルからのデータを表し、Oracle Warehouse Builder で通常のソース表のように動作する読み取り専用の表です。作成する外部表は、既存のフラット・ファイル内のシングル・レコード・タイプに対応します。

この演習での目的は、前にインポートした 2 つのフラット・ファイルに必要な外部表を作成することです。2 つのファイルの両方にシングル・レコード・タイプがあるので、各ファイルに対して外部表を 1 つのみ作成する必要があります。

ターゲット・ウェアハウス・モジュールに外部表を追加する手順

- プロジェクト・エクスプローラで、「データベース」ノード、続いて「Oracle」ノードを開きます。
- 外部表を作成するターゲット・モジュールを開きます。
つまり、EXPENSE_WH モジュールが開かれます。
- 「外部表」ノードを右クリックして「新規」を選択します。
Oracle Warehouse Builder で、外部表を作成するウィザードが表示されます。プロンプトに従います。
外部表に EXPENSE_CATEGORIES という名前を付けます。フラット・ファイルを選択するように求められた場合は、EXPENSE_CATEGORIES_CSV を選択します。
- 前のステップを繰り返し、EXPORT_CSV を表す EXPENSE_DATA と呼ばれる外部表を作成します。

- 2つの外部表に対して物理ファイル・システムの詳細を構成します。

モジュールから1つの外部表を右クリックし、「構成」を選択します。「データファイル」ノードを右クリックし、「作成」を選択します。デフォルト名、NEW_DATAFILE_1を受け入れます。外部表がデータを継承するフラット・ファイルの名前を入力します。したがってデータファイル名は、1つの外部表に対しては expense_categories.csv とし、もう1つの外部表に対しては export.csv とします。

ディメンションについて

ディメンションは、データを編成する構造です。一般的に使用されるディメンションの例には、顧客、時間および製品があります。

リレーショナル・ディメンションの場合は、ディメンションの使用によって、問合せパフォーマンスが向上します。これは、ユーザーが既知の階層をドリルダウンしてデータを分析することがよくあるためです。階層の例には、年、四半期、月、日という時間階層があります。Oracle Database では、詳細表ではなくマテリアライズド・ビューからデータを取り出す問合せをリライトして、これらの定義済階層を使用します。

一般的なリレーショナル・ディメンション表には、次の特徴があります。

- ウェアハウス・キーと呼ばれる、値が移入された単一列の主キーがあります。
ウェアハウス・キーは、ディメンションを管理し、ディメンション履歴を保持する技術をサポートし、キューブのサイズを削減します。
- ディメンション・オブジェクトとして明示的に定義された1つ以上の階層があります。階層によって、Oracle サーバーによる問合せリライトの数が最大化されます。
- ディメンションは、スター・スキーマのデータのプライマリ組織単位です。一般的に使用されるディメンションの例は、カスタマ、「製品」および「時間」です。

ディメンションは、一連のレベルと、これらのレベルに定義されている一連の階層で構成されます。ディメンションを作成する際は、次の内容を定義する必要があります。

- **ディメンション属性:** ディメンション・メンバーの特徴を説明する特性です。ディメンション属性には、名前およびデータ型があります。
- **レベル:** データの集計のレベルを定義します。たとえば、PRODUCTS ディメンションには、合計、グループおよび製品のレベルを設定できます。
- **レベル属性:** レベル・メンバーの特徴を説明する特性です。ディメンションの各レベルに、一連のレベル属性があります。
- **階層:** データの編成方法として、順序付けされたレベルまたは一連のデータ値（値ベース階層の場合）を使用する論理構造です。階層構造は一連のレベルの親子関係を説明します。

演習：ディメンションの理解

ディメンションの基本概念および設計を理解するには、この演習で事前定義のディメンションを確認します。

ディメンションの理解を深めるための手順

- データ・オブジェクト・エディタで PRODUCTS ディメンションを開きます。

「プロジェクト・エクスプローラ」パネルで、「OWB_DEMO」、「データベース」、「Oracle」、「SALES_WH」にナビゲートして、「ディメンション」を開きます。「PRODUCTS」をダブルクリックします。

Oracle Warehouse Builder はデータ・オブジェクト・エディタを起動します。データ・オブジェクト・エディタは、様々なデータベースまたはディメンション・オブジェクトの設計、作成および管理が簡単にできる単一インタフェースです。

2. ディメンション属性を監視します。

ディメンション属性は、ディメンション・メンバーの記述特性です。ディメンション属性には、名前とデータ型があり、ディメンション内の1つ以上のレベルに適用できます。これらは、データを格納するためのレベル属性として実装されます。

ディメンション属性のリストには、ディメンション内のあらゆるレベルに必要なすべての属性が含まれる必要があります。

たとえば、PRODUCTS ディメンションには、説明というディメンション属性があります。この属性は、すべてのレベルの合計、グループおよび製品に適用でき、その中には、これらのレベルの各メンバーに対する説明が格納されます。

3. レベルを監視します。

ディメンションのレベルは、データの集計レベルを表します。ディメンションには、少なくとも1つのレベルが含まれている必要があります。ただし、値ベース階層が含まれるディメンションは除きます。各レベルには、レベル属性およびレベル識別子が必要です。

たとえば、PRODUCTS ディメンションには、合計、グループおよび製品のレベルを設定できます。

4. 追加の演習として、次の Web サイトの Oracle By Example を参照して TIME ディメンションおよび通常のディメンションを作成できます。

http://www.oracle.com/technology/obe/10gr2_owb/owb10gr2_gs/owb/lesson3/relational-wh-03.htm

レベルについて

各レベルには、サロゲート識別子とビジネス識別子の2つの識別子が必要です。ディメンションを作成するときは、そのディメンションのサロゲート識別子およびビジネス識別子（複合ビジネス識別子の場合は属性）としてマーク付けされたディメンション属性が、各レベルで実装される必要があります。

サロゲート識別子は、ディメンションのすべてのレベルにまたがって、各レベルのレコードを一意に識別します。この識別子は、単一の属性で構成する必要があります。サロゲート識別子を使用すると、ファクトを最下位ディメンション・レベルのみでなく、どのディメンション・レベルにもフックできます。

リレーショナル実装を使用しているディメンションの場合、サロゲート識別子のデータ型は NUMBER であることが必要です。サロゲート識別子の値は、すべてのディメンション・レベルにまたがって一意である必要があるため、すべてのディメンション・レベルのサロゲート識別子を同じ順序を使用して生成します。

リレーショナル実装の場合、サロゲート識別子は次の目的を果します。

- 子レベルが親レベルとは別の表に格納されている場合、各子レベルのレコードには、親レコードのサロゲート識別子が格納されます。
- ファクト表の各キューブ・レコードには、参照先のディメンション・レコードのサロゲート識別子のみが格納されます。サロゲート識別子を格納することによって、キューブを実装するファクト表のサイズが削減されます。

ビジネス識別子はユーザーが選択する属性リストで構成されます。ビジネス識別子は、特定のレベルで一意であることが必要であり、常にデータ・ソースの自然キーから導出されます。ビジネス識別子はメンバーを一意に識別します。たとえば、製品レベルのビジネス識別子として統一商品コード（UPC）を使用できます。UPC は、各商品に対応する一意のコードです。

ビジネス識別子は、次のように機能します。

- ビジネス期間のレコードを識別します。
- ファクトとディメンション間または2つのレベル間の論理的なリンクを提供します。
- サロゲート・キーを参照可能にします。

ディメンションに子レベルを移入する場合は、その親レベルのビジネス識別子を指定する必要があります。キューブを移入する場合は、そのキューブが参照するディメンション・レベルのビジネス識別子を指定する必要があります。

親識別子は、値ベース階層の親参照に注釈を付けるために使用されます。

たとえば、値ベース階層の EMPLOYEE ディメンションに、ID、FIRST_NAME、LAST_NAME、EMAIL、PHONE、JOB_ID、HIRE_DATE および MANAGER_ID のディメンション属性があるとします。このディメンションでは、ID がサロゲート識別子で、MANAGER_ID が親識別子です。

レベル属性の定義

レベル属性とはレベル・メンバーを説明する特性です。ディメンションの各レベルには一連のレベル属性があります。レベル属性を定義するには、そのレベルで実装するディメンション属性を単純に選択します。レベル属性には個別名とデータ型があります。データ型は、そのレベル属性に実装されているディメンション属性から継承されます。レベル属性の名前は、実装されているディメンション属性の名前とは異なる名前に変更できます。

すべてのレベルは、一連のディメンション属性のサロゲート識別子およびビジネス識別子としてマークされた属性を実装する必要があります。

階層の定義

ディメンション階層は、データの編成方法として、順序付けされたレベルまたは一連のデータ値（値ベース階層の場合）を使用する論理構造です。階層構造は一連のレベルの親子関係を説明します。レベル・ベース階層には少なくとも1つのレベルが必要です。複数の階層の一部を1つのレベルにできます。

たとえば、TIME ディメンションには次の2つの階層が指定されます。

会計階層 : 会計年度 > 会計年度四半期 > 会計年度月 > 会計年度週 > 日

カレンダー階層 : カレンダー年 > カレンダー四半期 > カレンダー月 > 日

すべての階層は、完全な1対nの関係になります。親レベルの1つのレコードは子レベルの複数のレコードに対応します。一方、子レベルの1つのレコードは階層内で1つの親レコードにのみ対応します。

ディメンション・ロール

ディメンション・ロールとはディメンションの別名です。データ・ウェアハウスでは、キューブが同じディメンションを複数回参照する場合、参照する都度そのディメンションを格納する必要はありません。同じディメンションに対する複数回の参照は混乱の原因となります。そこで、ディメンションを参照するたびに別名を作成することで、結合を瞬時に理解できるようになります。このような場合は、同じディメンションがキューブ内の異なるディメンション・ロールを実行します。

たとえば、売上レコードには次の3つの時間値があります。

- 受注記帳時間
- 受注出荷時間
- 受注履行時間

3つの TIME ディメンションを作成し、それらのディメンションにデータを移入するかわりに、ディメンション・ロールを使用できます。1つの TIME ディメンションのモデルを作成し、そのディメンションに3つのロール（受注記帳時間、受注出荷時間、受注履行時間）を作成します。SALES キューブでは、受注時間、出荷時間および受注履行時間のディメンションを参照できます。

ディメンションがデータベースに格納される場合は、1つのディメンションのみが作成され、このディメンションを各ディメンション・ロールが参照します。一方、ディメンションがOLAPカタログに格納される場合は、Oracle Warehouse Builderによって、各ディメンション・ロールに1つのディメンションが作成されます。したがって、TIMEディメンションに3つのロールがある場合、OLAPカタログには3つのディメンションが作成されます。ただし、3つのディメンションはすべて同一の基礎となる表にマッピングされます。これは、OLAPカタログがディメンション・ロールをサポートしていないための対策です。

注意： ディメンション・ロールは、リレーショナル実装を持つディメンションに対してのみ作成できます。

レベルの関係

レベルの関係とは、ディメンション階層内のレベル間のアソシエーションです。レベルの関係は、階層内の親レベルへの参照を格納するレベル属性を使用して実装されます。

たとえば、PRODUCTSディメンションに「合計>グループ>製品」という階層があるとします。Oracle Warehouse Builderでは、2つのレベルの関係（製品からグループ、グループから合計）が作成されます。2つの新しい属性には、このレベルの関係（製品レベルでの関係とグループ・レベルでの関係）が実装されます。これらの属性には、親レベルのサロゲートIDが格納されます。

ディメンションの例

ディメンションの一例として、製品データの編成に使用するPRODUCTSディメンションを紹介します。表4-1に、PRODUCTSディメンションの各レベルと、ディメンション内の各レベルのサロゲート識別子とビジネス識別子をリストします。

表 4-1 PRODUCTS ディメンションのレベルの詳細

| レベル | 属性名 | 識別子 |
|------|-----------|-------|
| 合計 | ID | サロゲート |
| | 名前 | ビジネス |
| | 説明 | |
| グループ | ID | サロゲート |
| | 名前 | ビジネス |
| | 説明 | |
| 製品 | ID | サロゲート |
| | UPC | ビジネス |
| | 名前 | |
| | 説明 | |
| | パッケージ・タイプ | |
| | パッケージ・サイズ | |

PRODUCTS ディメンションの階層は、次のとおりです。

階層 1: 合計 > グループ > 製品

制御行

Oracle Warehouse Builder により、ファクト・データを各レベルのディメンションにリンクできるコントロール行が作成されます。たとえば、TIME ディメンションを2つの異なるキューブで再利用して、予算データを月レベルで記録し、実際のデータを日レベルで記録する場合などです。ディメンションはコントロール行と一緒にロードされるため、追加定義をすることなくこれを実行できます。ディメンション階層の各メンバーは、単一のレコードを使用して表示されます。

すべての制御行には -2 から始まる負のディメンション・キー値があります。上位レベルのレベル値では行が生成され、ファクト表に対して一意でリンクする行として動作できます。このリンクの下位レベルまたは制御行は無効化されます。

「ディメンションの例」で説明されている PRODUCTS ディメンションを考慮します。4つの製品カテゴリを含む表からデータをこのディメンションにロードします。Oracle Warehouse Builder は表 4-2 に示されるディメンションに制御行を挿入します。これらの行によりどのディメンション・レベルでもキューブにリンクできます。表にはすべてのディメンション属性値が含まれるわけではありません。

表 4-2 PRODUCTS ディメンションに作成された制御行

| ディメンション・キー | 合計名 | カテゴリ名 | 製品名 |
|------------|-------|----------|-----|
| -3 | TOTAL | | |
| -9 | TOTAL | ハードウェア | |
| -10 | TOTAL | ソフトウェア | |
| -11 | TOTAL | エレクトロニクス | |
| -12 | TOTAL | 周辺機器 | |

ディメンション内の行の実数を取得するには、NULL 行を除いた WHERE 句を含む行の数を数えます。たとえば、「製品」上の数を取得するには、「製品」内の NULL 行を除いた WHERE 句を含む行の数を数えます。

ディメンションの実装

ディメンションの実装とは、ディメンションおよびそのデータを物理的に格納する方法を指定することです。Oracle Warehouse Builder によって、ディメンション・オブジェクトに対して複数のタイプの実装（多次元実装を含む）を行うことができます。ただし、このガイドでは、リレーショナル実装についてのみ説明します。

スター・スキーマ

スター・スキーマ実装では、Oracle Warehouse Builder によって単一の表にディメンション・データが格納されます。同じ表またはビューに複数のディメンション・レベルのデータが格納されるため、表にディメンション・キーの列を指定する必要があります。ディメンション・キーの列は、ディメンションの主キーです。この列はキューブの外部キー参照も形成します。

各レベルには、ディメンション属性のサブセットが実装されます。デフォルトでは、レベル属性の名前はディメンション属性の名前と同じです。すべてのレベルのデータが同じ表に格納されることで生じる名前の衝突を回避するため、Oracle Warehouse Builder では、スター表の名前の指定に、次のガイドラインが使用されます。

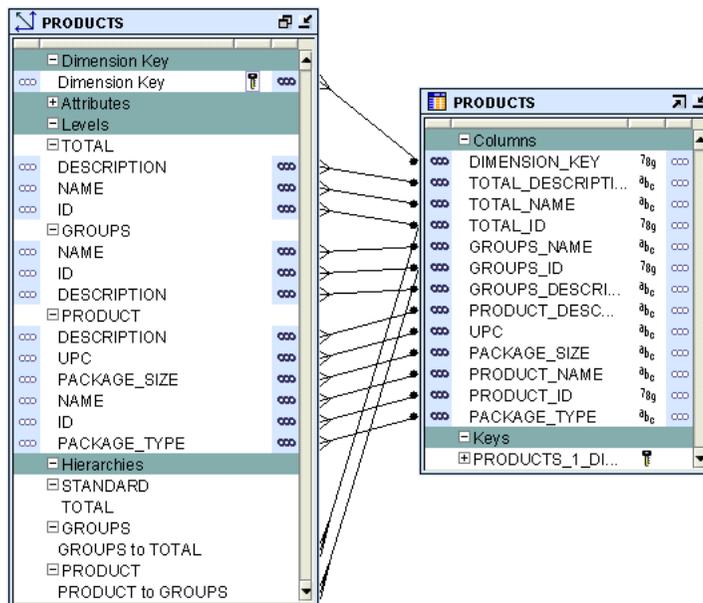
- レベル属性の名前が一意でない場合は、Oracle Warehouse Builder によってレベル名に接頭辞が付けられます。
- レベル属性の名前が一意の場合は、接頭辞は使用されません。

注意： 接頭辞が使用されないようにするには、「ディメンションの作成」ウィザードまたはデータ・オブジェクト・エディタで、レベル属性の名前を明示的に変更する必要があります。

たとえば、スター・スキーマを使用して PRODUCTS ディメンションを実装する場合、Oracle Warehouse Builder は単一の表を使用し、ディメンションにすべてのレベルを実装します。

図 4-1 は、PRODUCTS ディメンションのスター・スキーマ実装を示しています。すべてのレベルの属性が、PRODUCTS という 1 つの表のそれぞれの列にマッピングされています。DIMENSION_KEY という列に、ディメンションのサロゲート ID が格納されています。この列は表の主キーです。

図 4-1 PRODUCTS ディメンションのスター・スキーマ実装



バインド

バインドを実行する場合は、各属性のデータおよびディメンション内のレベルの関係を格納するデータベース列を指定します。ディメンションに対して自動バインドまたは手動バインドを実行できます。

自動バインド 自動バインドを実行すると、Oracle Warehouse Builder によって、ディメンション・オブジェクトの属性がそのデータを格納するデータベース列にバインドされます。自動バインドを初めて実行するときは、ディメンション・データの格納に使用される表も作成されます。

すでにバインド済のディメンションに対して自動バインドを実行すると、Oracle Warehouse Builder では次のルールが使用されます。

- ディメンションの実装方法が同じ場合は、Oracle Warehouse Builder によって既存の実装オブジェクトにディメンション・オブジェクトが再バインドされます。
たとえば、スター・スキーマ実装方法を使用して PRODUCTS ディメンションを作成し、自動バインドを実行したとします。ディメンション・データは Products と呼ばれる表に格納されます。後日、ディメンション定義を変更しましたが、実装方法はスターのままです。この状況で PRODUCTS ディメンションを自動バインドすると、Oracle Warehouse Builder によって同じ実装表に PRODUCTS ディメンションの属性が再バインドされます。
- ディメンションの実装方法を変更すると、Oracle Warehouse Builder によって、古い実装オブジェクトが削除され、一連の新しい実装表が作成されます。古い実装オブジェクトを保持する場合は、自動バインドを実行する前に、ディメンション・オブジェクトをアンバインドする必要があります。実装方法の詳細は、4-8 ページの「スター・スキーマ」を参照してください。

たとえば、スター・スキーマ実装方法を使用して PRODUCTS ディメンションを作成し、それを実装表にバインドしたとします。その後、このディメンションを編集し、実装方法をスノーフレイクに変更しました。変更した PRODUCTS ディメンションの自動バインドを実行すると、Oracle Warehouse Builder によって、ディメンション・データが格納されていた表は削除され、新しい実装表が作成されて、ディメンション属性および関係が新しい実装表にバインドされます。

自動バインドを実行する手順

1. プロジェクト・エクスプローラで、ディメンションを右クリックして「**エディタを開く**」を選択します。

このディメンションのデータ・オブジェクト・エディタが表示されます。

2. 「ディメンシヨナル」タブで、「ディメンション」ノードを右クリックして「**バインド**」を選択します。

あるいは、キャンバスで「ディメンション」ノードを選択し、「**オブジェクト**」メニューから「**バインド**」を選択します。

「バインド」オプションが有効になっていない場合は、ディメンションがリレーシヨナル・ディメンションであるかどうか、また「**手動**」オプションが「**記憶域**」タブの「**実装**」セクションで設定されていないかどうかをチェックします。

4-8 ページの「**スター・スキーマ**」に説明されているように、自動バインドによって実装設定が使用されます。

手動バインド 通常は、手動バインドを使用して既存の表をディメンションにバインドします。自動バインドまたは再バインドが不要な場合は、手動バインドを使用します。

ディメンションの手動バインドを実行する手順

1. ディメンション・データの格納に使用する実装オブジェクト（表またはビュー）を作成します。

リレーシヨナル・ディメンションの場合は、ディメンションのサロゲート識別子のロードに使用する順序を作成します。既存の順序を使用することもできます。

2. プロジェクト・エクスプローラで、ディメンションを右クリックして「**エディタを開く**」を選択します。

そのディメンションのデータ・オブジェクト・エディタが表示されます。キャンバスでは「ディメンシヨナル」タブがアクティブになります。

3. ディメンションを右クリックして「**詳細の表示**」を選択します。

Oracle Warehouse Builder によって、ディメンションと同じ名前の新しいタブが表示されます。

4. パレットから、実装オブジェクトを表す演算子をキャンバスにドラッグ・アンド・ドロップします。

Oracle Warehouse Builder によって、「新規または既存の <オブジェクト> を追加」ダイアログ・ボックスが表示されます。たとえば、ディメンション・データが表に格納されている場合は、パレットからキャンバスに表演算子をドラッグ・アンド・ドロップすると、「新規または既存の表を追加」ダイアログ・ボックスが表示されます。

5. 「**既存の <オブジェクト> を選択**」オプションを選択し、選択ツリーに表示されるオブジェクトのリストからデータ・オブジェクトを選択します。

6. 「**OK**」をクリックします。

追加したオブジェクトを表すノードがキャンバスに表示されます。

7. ディメンション・データの格納に複数のデータ・オブジェクトが使用される場合は、各データ・オブジェクトについてステップ 4 から 6 を実行します。

8. ディメンションの各レベルの属性を、そのデータを格納する列にマッピングします。そのためには、ディメンション属性の上でマウスを押したまま、属性値を格納する列にドラッグ・アンド・ドロップします。

レベルの関係も、そのデータを格納するデータベース列にマッピングします。

たとえば、4-7 ページの「[ディメンションの例](#)」に記載されている PRODUCTS ディメンションの場合、PRODUCTS ディメンションのグループ・レベルの名前属性は、Products_tab 表の Group_name 属性に格納されます。名前属性の上でマウスを押したまま、Products_tab 表の Group_name 属性にドラッグ・アンド・ドロップします。

キューブについて

キューブには、メジャーおよび1つ以上のディメンションへのリンクが含まれます。キューブの軸にはディメンション・メンバーが組み込まれ、キューブの本体にはメジャーの値が格納されます。ほとんどのメジャーは加法的なメジャーです。たとえば、売上データは、キューブ・エッジに時間、製品および顧客のディメンション値が指定され、キューブ本体にバリュー売上とドル売上のメジャーから値が格納されるキューブに編成できます。

キューブは、外部キー制約によってディメンション表にリンクされます。これらの制約は、データ整合性が最も重要とされるデータ・ウェアハウス環境には不可欠です。この制約によって、データ・ウェアハウスを使用する日常業務での参照整合性が順守されます。

データ分析アプリケーションでは、通常、多くのディメンションにわたってデータが集計されます。これによって、データの変則的または特異なパターンを確認できます。キューブの使用は、この種の操作を実行する最も効率的な方法です。リレーショナル実装で、ウェアハウス・キーが設定されているディメンションを設計する場合、キューブの行の長さは一般的に短くなります。これは、ウェアハウス・キーの長さが対応する本来のデータより短いからです。その結果、キューブ・データに必要な記憶領域は少なくなります。

一般的なキューブには、次の要素が含まれます。

- 一連の外部キー参照列、またはデータ・リストの場合は人工キーまたは一連のウェアハウス・キー列に定義された主キー。キューブがデータ・リストの場合、外部キー参照列ではキューブの各行は一意に識別されません。
- 表を対応するディメンションにリンクする一連の外部キー参照列。

キューブの定義

キューブは、一連のディメンションにわたって定義されたメジャーの集合で構成されます。キューブを作成するには、次の要素を定義する必要があります。

- [キューブ・メジャー](#)
- [キューブのディメンション](#)

キューブ・メジャー

メジャーとは、調査および分析可能なデータで、通常は加法的な数値になります。メジャーの例としては、売上、原価および収益があります。キューブには、1つ以上のディメンションが必要です。メジャーの集計も可能です。ただし、集計できるのは数値メジャーのみです。

キューブのディメンション

キューブは一連のディメンションで定義されます。キューブはディメンション内で最下位レベル以外のレベルを参照できます。

純粋なリレーショナル実装を使用するキューブでは、ディメンション・ロールによって、複数回同じディメンションを再使用できます。ディメンション・ロールの詳細は、4-6 ページの「[ディメンション・ロール](#)」を参照してください。

キューブを有効にする前に、キューブが参照するディメンションすべてが有効になっていることを確認してください。

ディメンション参照を定義する場合に指定する要素

- キューブが参照するディメンションとディメンション内のレベル。
リレーショナル実装を使用するキューブでは、ディメンションの中間レベルを参照できません。Oracle Warehouse Builder では、レベルの非サロゲート識別子（ビジネス・キーなど）への参照をサポートしています。
- リレーショナル実装を使用するディメンションの場合は、各ディメンションのディメンション・ロール。ディメンション・ロールは、キューブ内でディメンション参照が実行するロールを示します。ディメンション・ロールの指定はオプションです。

キューブの例

SALES キューブには、集計された売上データが格納されます。このキューブには Value_sales および Dollar_sales という 2 つのメジャーが含まれています。

- Value_sales: 売上額を売上数量に換算して格納します。
- Dollar_sales: 売上額を格納します。

表 4-3 に、SALES キューブのディメンションを示します。この表には、キューブが参照するディメンションの名前とディメンションのレベルがリストされています。

表 4-3 SALES キューブのディメンション

| ディメンション名 | レベル名 |
|-----------|------|
| Products | 製品 |
| Customers | 顧客 |
| Times | 日 |

キューブの実装

キューブを実装する際は、キューブの物理的な記憶域の詳細を指定します。ディメンションの場合と同様に、Oracle Warehouse Builder によって、リレーショナル形式または多次元形式でキューブを実装できます。このガイドでは、リレーショナル実装について説明します。

キューブのリレーショナル実装

キューブ・データの格納に使用するデータベース・オブジェクトは、ファクト表と呼ばれます。キューブは 1 つのファクト表のみを使用して実装する必要があります。ファクト表にはキューブのメジャーおよびディメンション参照の列が含まれます。

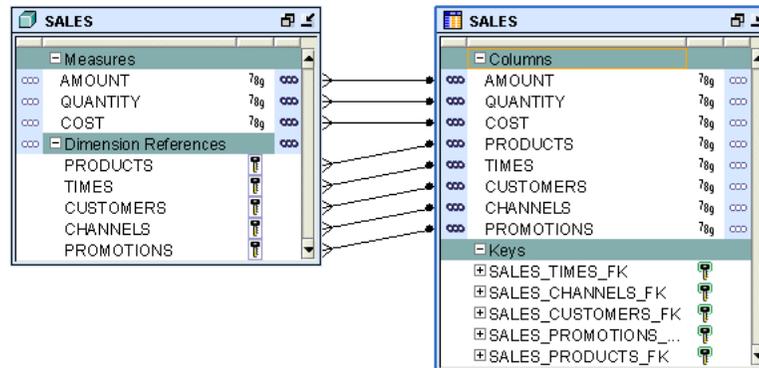
キューブを実装する手順

- キューブ・データを格納する表またはマテリアライズド・ビューを選択します。
- 各メジャーについて、メジャー・データを格納する列を選択します。
- 各ディメンション参照について、ディメンション参照を格納する列を選択します。

各ディメンション参照は、ファクト表の列およびファクト表からディメンション表への外部キー（オプション）に対応します。ファクト表からディメンション表への 1 対 n の関係が順守される必要があります。

図 4-2 に、SALES キューブのリレーショナル実装のバインドを示します。SALES キューブのデータは SALES という表に格納されます。

図 4-2 SALES キューブの実装



バインド

バインドを実行する場合は、キューブのメジャーとディメンション参照の各データを格納するデータベース列を指定します。キューブに対して自動バインドまたは手動バインドを実行できます。

自動バインド 自動バインドを実行すると、Oracle Warehouse Builder によって、キューブ・データを格納する表が作成され、キューブのメジャーおよび参照がデータベースの列にバインドされます。自動バインドを実行する手順の詳細は、4-9 ページの「[自動バインド](#)」を参照してください。

キューブの自動バインドを実行する場合は、キューブが参照するディメンションを自動バインドした後に、キューブを自動バインドします。キューブが最後に自動バインドされた後に、そのキューブによって参照されるディメンションが自動バインドされた場合は、キューブを配布できません。

たとえば、TIMES ディメンションと PRODUCTS ディメンションを参照する SALES キューブを作成し、そのキューブの自動バインドを実行したとします。その後、PRODUCTS ディメンションの定義を変更しました。ここで、再度 SALES キューブを自動バインドしようとすると、Oracle Warehouse Builder によってエラーが生成されます。最初に PRODUCTS ディメンションを自動バインドした後に、キューブを自動バインドする必要があります。

手動バインド 手動バインドでは、最初にキューブ・データを格納する表またはビューを作成し、次にそのデータを格納するデータベースの列にキューブの参照およびメジャーをマッピングする必要があります。あるいは、既存のデータベース表またはビューを使用して、キューブ・データを格納することもできます。

キューブの手動バインドを実行する手順

1. キューブ・データを格納する表またはビューを作成します。
2. プロジェクト・エクスプローラで、キューブを右クリックして「**エディタを開く**」を選択します。

そのキューブのデータ・オブジェクト・エディタが開きます。キャンバスでは「ディメンショナル」タブがアクティブになります。

3. キューブを右クリックして「**詳細の表示**」を選択します。

Oracle Warehouse Builder によって、キューブと同じ名前の新しいタブが表示されます。

4. パレットから、実装オブジェクトを表す演算子をキャンバスにドラッグ・アンド・ドロップします。

Oracle Warehouse Builder によって、「新規または既存の <オブジェクト> を追加」ダイアログ・ボックスが表示されます。たとえば、キューブ・データが表に格納されている場合は、パレットからキャンバスに表演算子をドラッグ・アンド・ドロップすると、「新規または既存の表を追加」ダイアログ・ボックスが表示されます。

5. 「既存の <オブジェクト> を選択」オプションを選択し、選択ツリーに表示されるオブジェクトのリストからデータ・オブジェクトを選択します。
6. 「OK」をクリックします。

追加したオブジェクトを表すノードがキャンバスに表示されます。

7. キューブのメジャーとディメンション参照をキューブ・データが格納される列にマッピングします。そのためには、メジャーまたはディメンション参照をマウス・ボタンで押したまま、そのメジャーまたはディメンション参照を格納するデータ・オブジェクト属性にドラッグ・アンド・ドロップします。

第II部

データ・ウェアハウスへのデータのロード

ここでは、データ・ウェアハウスへのデータのロード方法について説明します。内容は次のとおりです。

- [第5章「ETL ロジックの定義」](#)
- [第6章「ロード・パフォーマンスの向上」](#)
- [第7章「ターゲット・スキーマへの配布および ETL ロジックの実行」](#)

ETL ロジックの定義

Oracle Warehouse Builder でデータ・オブジェクトの定義を作成およびインポートした後は、データをソースからターゲットに移動する抽出、変換およびロード (ETL) の各操作を設計できます。Oracle Warehouse Builder では、マッピングを使用してこれらの操作を設計します。

この項では、マッピングの作成、編集および使用方法に関する次のトピックについて説明します。

- [マッピングおよび演算子について \(5-2 ページ\)](#)
- [マッピングを定義する手順 \(5-2 ページ\)](#)
- [マッピングの作成 \(5-3 ページ\)](#)
- [演算子の追加 \(5-4 ページ\)](#)
- [演算子の編集 \(5-6 ページ\)](#)
- [演算子の接続 \(5-6 ページ\)](#)
- [マッピング・プロパティの設定 \(5-9 ページ\)](#)
- [演算子、グループおよび属性プロパティの設定 \(5-11 ページ\)](#)
- [演算子とワークスペース・オブジェクトの同期化 \(5-11 ページ\)](#)

マッピングおよび演算子について

マッピングとはデータをソースから抽出し、変換してターゲットにロードする一連の操作を表します。これによりデータ・フローおよびデータで実行される操作がビジュアル表示されます。マッピングを Oracle Warehouse Builder で設計する場合は、マッピング・エディタ・インタフェースを使用します。

マッピングの基本となる設計要素は演算子です。演算子を使用して、データ・フローでソースおよびターゲットを表現します。また、演算子を使用して、ソースからターゲットへのデータの変換方法を定義することもできます。ソースとして選択する演算子は、マッピングの設計方法に影響を与えます。Oracle Warehouse Builder では、選択された演算子に基づいて、次のいずれかのマッピング生成言語にマッピングが割り当てられます。

- PL/SQL
- SQL*Loader
- ABAP

各コード言語では、マッピングを設計する際、特定のルールに従う必要があります。

このガイドでは、PL/SQL マッピングを定義する方法を説明します。他のタイプのマッピングを定義するには、『Oracle Warehouse Builder ユーザーズ・ガイド』を参照してください。PL/SQL マッピングを定義する基本ルールは、PL/SQL マッピングには、フラット・ファイル演算子または SAP/R3 ソース以外のすべてのタイプのソース演算子を含めることができます。

マッピングを定義する手順

マッピングを定義する場合に参照する項

1. [マッピングの作成](#) (5-3 ページ)
2. [演算子の追加](#) (5-4 ページ)
3. [演算子の編集](#) (5-6 ページ)
4. [演算子の接続](#) (5-6 ページ)
5. [マッピング・プロパティの設定](#) (5-9 ページ)
6. [演算子、グループおよび属性プロパティの設定](#) (5-11 ページ)
7. Oracle Warehouse Builder のオンライン・ヘルプマッピング構成のリファレンス
8. マッピング・デザインに問題がない場合、ツールバーの「生成」アイコンを選択してコードを生成します。

後続の手順

マッピングを設計してマッピングのコードを生成した後、プロセス・フローを作成するか、直接配布を続行して実行できます。

プロセス・フローを使用して、マッピングを相互に関連付けます。たとえば、あるマッピングが完了すると電子メール通知がトリガーされて別のマッピングが起動するように、プロセス・フローを設計できます。

マッピングおよび作成した関連するプロセス・フローを配布し、マッピングを実行します。

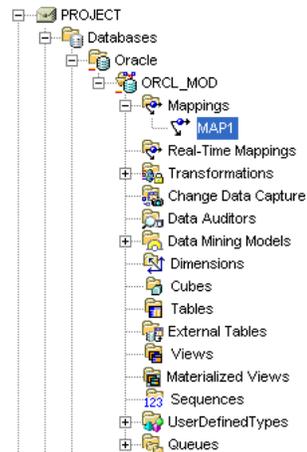
マッピングの作成

マッピングを作成する手順

1. プロジェクト・エクスプローラの「マッピング」ノードにナビゲートします。このノードは、ウェアハウス・ターゲット・モジュールの下、「データベース」フォルダの下の「Oracle」フォルダにあります。

図 5-1 「プロジェクト・エクスプローラの「マッピング」ノード」に、MAP1 というマップが格納されている「マッピング」ノードを示します。この例では、ウェアハウス・ターゲットの名前は ORCL_MOD です。

図 5-1 プロジェクト・エクスプローラの「マッピング」ノード



2. 「マッピング」を右クリックして「新規」を選択します。
「マッピングの作成」ダイアログ・ボックスが表示されます。
3. 新しいマッピングの名前と説明（オプション）を入力します。
マッピング名の指定と説明の記述に関するルールを表示するには、「ヘルプ」を選択します。
4. 「OK」をクリックします。
マッピングの定義が保存され、その名前がプロジェクト・エクスプローラ内に挿入されます。また、そのマッピングのマッピング・エディタが表示され、タイトル・バーにマッピングの名前が表示されます。

以前に作成したマッピングを開く手順

1. プロジェクト・エクスプローラでウェアハウス・ターゲット・モジュールを「データベース」フォルダに指定して、次に「Oracle データベース」フォルダに指定します。
2. 「マッピング」ノードを開きます。
3. 次のいずれかの方法で、マッピング・エディタを開きます。
 - マッピングをダブルクリックします。
 - マッピングを選択し、「編集」メニューから「エディタを開く」を選択します。
 - マッピングを選択し、[Ctrl] キーを押しながら [O] キーを押します。
 - マッピングを右クリックし、「エディタを開く」を選択します。

Oracle Warehouse Builder によって、マッピング・エディタが表示されます。

演算子のタイプ

マッピングを設計するときは、マッピング・エディタ・パレットから演算子を選択し、キャンバスにドラッグできます。

- **Oracle ソース / ターゲット演算子**: これらの演算子を使用して、Oracle 表、ビュー、マテリアライズド・ビューなどの Oracle Database オブジェクトを示します。
- **リモートおよび Oracle 以外のソース演算子とターゲット演算子**: これらの演算子の使用には、特別な要件があります。
- **データ・フロー演算子**: データ・フロー演算子を使用して、データを変換します。
- **事前 / 事後処理演算子**: マッピングの実行前または後に処理を実行するには、事前 / 事後処理演算子を使用します。マッピングに対する値を指定するには、マッピング・パラメータ演算子を使用します。
- **プラグgable・マッピング演算子**: プラガブル・マッピング演算子は、マッピング演算子の再利用可能なグループで、単一の演算子として機能します。

演算子の追加

演算子をマッピングに追加する手順は、選択する演算子のタイプによって異なります。これは、一部の演算子はワークスペース・オブジェクトにバインドされ、その他の演算子はバインドされないためです。一般的に、データ・ソースまたはターゲット演算子を追加すると、Oracle Warehouse Builder によって、Oracle Warehouse Builder ワークスペースのそのオブジェクトのバージョンと、マッピング・エディタ用の個別のバージョンが維持されます。たとえば、表演算子をマッピングに追加した場合は、Oracle Warehouse Builder によってその表のコピーがワークスペースで維持されます。個々のバージョンは、1 つにバインドされます。つまり、マッピングのバージョンがワークスペースのバージョンにバインドされます。

2 つのバージョンを区別するために、この項では、ワークスペース内のオブジェクトを総称的にワークスペース・オブジェクトと呼び、具体的には、ワークスペース表、ワークスペース・ビューのように表現します。また、マッピングの演算子は表演算子、ビュー演算子のように表現します。したがって、ディメンションをマッピングに追加する場合、マッピングのディメンションはディメンション演算子、ワークスペースのディメンションはワークスペース・ディメンションと表されます。

Oracle Warehouse Builder では、ユーザーがこれらのオブジェクトの変更定義を同期化できるように、一部の演算子に個別のワークスペース・オブジェクトが保持されます。たとえば、ワークスペース表の新規メタデータ定義を再インポートした場合は、これらの変更内容をマッピングの表演算子に伝播する必要があります。反対に、マッピングの表演算子を変更した場合は、これらの変更内容を関連するワークスペース表に伝播する必要があります。これらのタスクは、同期化と呼ばれるプロセスによって実現します。Oracle Warehouse Builder では、自動的に同期化またはマッピング・エディタから手動で同期化できます。

演算子をマッピングに追加する手順

1. マッピング・エディタを開きます。
2. 「マッピング」メニューから「追加」を選択し、演算子を選択します。または、ツールボックスからマッピング・エディタのキャンバスに演算子のアイコンをドラッグ・アンド・ドロップします。

ワークスペース・オブジェクトにバインドできる演算子を選択すると、マッピング・エディタに「マッピング < 演算子名 > の追加」ダイアログ・ボックスが表示されます。このダイアログ・ボックスの使用の詳細は、「ヘルプ」を選択してください。

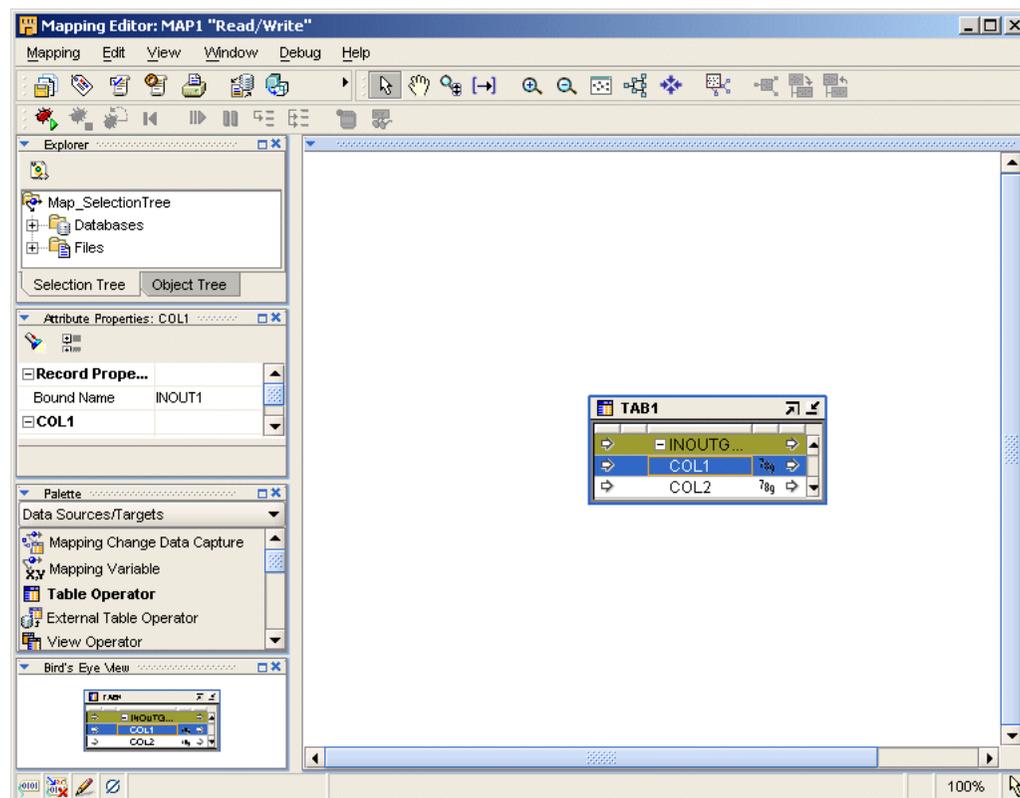
ワークスペース・オブジェクトにバインドできない演算子を選択した場合は、演算子を作成するためのウィザードまたはダイアログ・ボックスが表示されます。

3. Oracle Warehouse Builder に表示される指示に従い、「OK」をクリックします。

マッピング・エディタでは、キャンバスに最大化された演算子が表示されます。左上の隅に演算子名が表示されます。各属性名およびデータ型を表示できます。

演算子を最小化するには、右上隅の矢印をクリックします。マッピング・エディタのキャンバスに演算子がアイコンとして表示されます。

図 5-2 マッピング・エディタに表示された表演算子のソース



ワークスペース・オブジェクトにバインドする演算子の追加

ワークスペース・オブジェクトにバインドできる演算子を追加すると、マッピング・エディタに「マッピング < 演算子名 > の追加」ダイアログ・ボックスが表示されます。次のオプションから1つ選択します。

- バインドされていない演算子を属性なしで作成
- 既存のワークスペース・オブジェクトから選択してバインド

バインドされていない演算子を属性なしで作成

このオプションは、マッピング・エディタを使用して、新規ステージング領域表や新規ターゲット表など、新しいワークスペース・オブジェクトを定義する場合に使用します。

「バインドされていない演算子を属性なしで作成」を選択した後、新しいオブジェクトの名前を入力します。属性のない演算子がキャンバスに表示されます。

これで、演算子の属性を追加および定義できるようになりました (5-6 ページの「演算子の編集」を参照)。次に、ターゲット・モジュールに新規ワークスペース・オブジェクトを作成するために、演算子を右クリックして「作成とバインド」を選択します。

このオプションをマッピング設計で使用する方法の例は、5-8 ページの「例: マッピング・エディタによるステージング領域表の作成」を参照してください。

既存のワークスペース・オブジェクトから選択してバインド

このオプションは、ワークスペース内の定義済またはインポート済のオブジェクトに基づいて、演算子を追加する場合に使用します。

接頭辞を入力してオブジェクトを検索するか、選択したモジュール内のオブジェクトの表示リストからオブジェクトを選択します。

複数のアイテムを選択するには、[Ctrl] キーを押しながら各アイテムをクリックします。連続したアイテムのグループを選択するには、選択範囲の最初のオブジェクトをクリックし、[Shift] キーを押しながら最後のオブジェクトをクリックします。

演算子は、マッピングと同じモジュール内のワークスペース・オブジェクトまたは別のモジュールのワークスペース・オブジェクトに基づいて追加できます。別のモジュールのワークスペース・オブジェクトを選択すると、マッピング・エディタによってコネクタが作成されます（コネクタが存在していない場合）。このコネクタによって、マッピングのロケーションとワークスペース・オブジェクトのロケーション間のデータの移動パスが確立されます。

演算子の編集

各演算子には、エディタが関連付けられています。その演算子エディタを使用して、演算子、グループおよび属性に関する一般的な情報と構造的な情報を指定します。演算子エディタでは、グループおよび属性を追加または削除したり、名前を変更できます。演算子の名前も変更できます。

演算子の編集は、ロード・プロパティや条件付き動作の割当てとは異なります。ロード・プロパティや条件付き動作を指定するには、プロパティ・ウィンドウ（5-11 ページの「[演算子、グループおよび属性プロパティの設定](#)」を参照）を使用します。

演算子、グループまたは属性を編集する手順

1. マッピング・エディタのキャンバスから演算子を選択します。

あるいは、演算子内のグループまたは属性を選択します。

2. 選択したアイテムを右クリックし、「**詳細をオープン**」を選択します。

マッピング・エディタに演算子エディタが表示され、「名前」タブ、「グループ」タブ、および各タイプの演算子のグループの「入力 / 出力」タブが表示されます。

いくつかの演算子には追加のタブが含まれます。たとえば、Match Merge 演算子には一致ルールおよびマージ・ルールを定義するタブが含まれます。

3. 各タブに表示されるプロンプトに従い、「OK」をクリックして終了します。

タブの実行方法については、「ヘルプ」を参照してください。

演算子の接続

マッピング・ソース演算子、データを変換する演算子およびターゲット演算子を選択した後は、これらの演算子を接続できます。データ・フローの接続により、ソースから演算子を通じてターゲットに至るデータ・フローの様子が視覚的に表されます。

演算子は、次のいずれかの方法で接続できます。

- **属性の接続** : 1 度に個々の演算子属性を相互に接続します。
- **グループの接続** : 2 つのグループ間ですべての属性を接続するための条件を定義します。
- **演算子ウィザードの使用** : ピボット演算子、Name and Address 演算子などの演算子は、ウィザードを使用してデータ・フロー接続を定義できます。

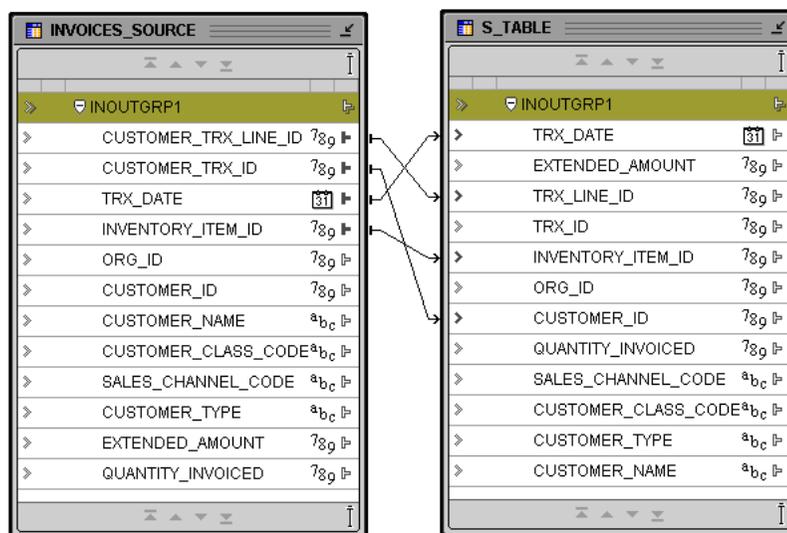
属性の接続

一方の演算子の単一の出力属性から、他方の演算子の単一の入力属性に線を引きます。

属性を接続する手順

1. マウス・ボタンをクリックしたまま、ポインタを出力属性の上に置きます。
 2. 出力属性からデータのフロー先となる入力属性にマウスをドラッグします。
マウスをドラッグするとき、接続を示す行がマッピング・エディタ・キャンパスに表示されます。
 3. 入力属性の上でマウス・ボタンを放します。
 4. ステップ1から3を繰り返し、必要なデータ・フロー接続をすべて作成します。
- 図 5-3 に、属性が接続されたマッピングを示します。

図 5-3 マッピングで接続した演算子



属性を接続するときは、次のルールに留意します。

- 同じ入力属性には2度接続できません。
- 同じ演算子内の属性には接続できません。
- 入力専用の属性から接続することも、出力専用の属性に接続することもできません。
- 確立されているカーディナリティに矛盾する方法では演算子を接続できません。かわりに、ジョイナ演算子を使用します。

グループの接続

グループを接続すると、属性が自動的にコピーされるか、詳細情報を入力するように指示されます。

演算子グループを既存の属性を持たないターゲット・グループに接続すると、属性が自動的にコピーされ、各属性が接続されます。これは、「例: マッピング・エディタによるステージング領域表の作成」に示すようなマッピングを設計する場合に便利です。

例：マッピング・エディタによるステージング領域表の作成

マッピング・エディタでバインドされていない演算子を使用すると、ステージング領域表をすばやく作成できます。

次の手順は、既存のソース表に基づいてステージング表を作成する方法を示しています。この手順は、ビュー、マテリアライズド・ビュー、フラット・ファイルおよび変換の作成にも使用できます。

ソース表をステージング表にマップする手順

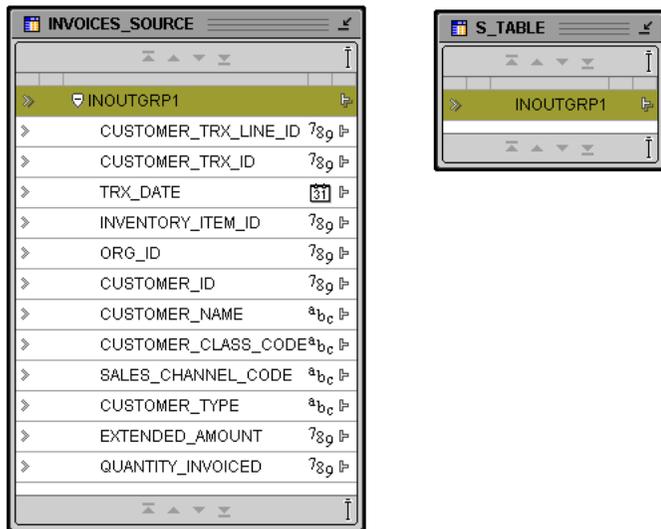
1. マッピング・エディタで、ソース表を追加します。

メニュー・バーから「マッピング」、「追加」、「データ・ソース/ターゲット」の順に選択します。「データ・ソース/ターゲット」メニューで、「演算子」を選択します。

2. 「演算子の追加」ダイアログ・ボックスを使用して、マッピング内のソース表の演算子を選択およびバインドします。「演算子の追加」ダイアログ・ボックスから、「バインドされていない演算子を属性なしで作成」を選択します。

マッピングは、[図 5-4](#) と類似していて、1つのソース表と属性のない1つのステージング領域表があります。

図 5-4 バインドされていないステージング表（属性なし）とソース表



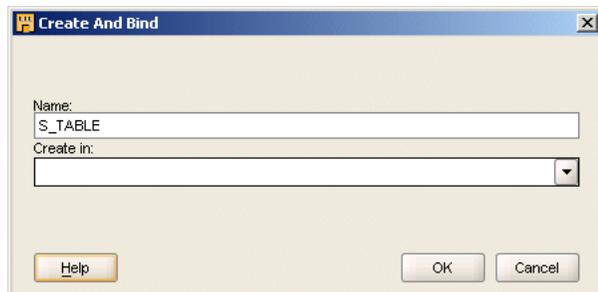
3. ソース演算子のグループにマウス・ポインタを置き、マウス・ボタンを押したままにします。
4. ステージング領域表グループまでマウスをドラッグします。

Oracle Warehouse Builder によって、ソース属性がステージング領域表にコピーされ、2つの演算子が接続されます。

- マッピング・エディタでは、マッピングに追加したバインドされていない表を選択します。右クリックして「作成とバインド」を選択します。

図 5-5 に示すダイアログ・ボックスが表示されます。

図 5-5 「作成とバインド」ダイアログ・ボックス



- 「作成場所」に、表を作成するターゲット・モジュールを指定します。

Oracle Warehouse Builder によって、指定したターゲット・モジュールに新しい表が作成されます。

マッピング・プロパティの設定

マッピング・キャンバスで空白を選択すると、エディタの左側にある「プロパティ・インスペクタ」にマッピング・プロパティが表示されます。マッピングについて次のプロパティを設定できます。

- [ターゲット・ロード順序](#) (5-9 ページ)

ターゲット・ロード順序

マッピングに1つのターゲットのみが含まれている場合、あるいはマッピングが SQL*Loader または ABAP マッピングの場合は、ターゲット・ロード順序が適用されません。デフォルトの設定を受け入れてマッピングの設計を続行してください。

複数のターゲットがある PL/SQL マッピングを設計すると、Oracle Warehouse Builder により、ターゲットをロードするデフォルトの順序が計算されます。ターゲット間に外部キー関連を定義すると、Oracle Warehouse Builder により、親の次に子をロードするデフォルトの順序が作成されます。外部キー関係を作成しない場合、またはターゲット表に再帰的な関連がある場合は、Oracle Warehouse Builder により、無作為な順序がデフォルトとして割り当てられます。

デフォルトのロード順序は、「ターゲット・ロード順序」プロパティを設定することで上書きできます。ターゲットの順序を変更する際に間違えた変更を実行した場合は、「デフォルトにリセット」オプションを選択することで、デフォルトの順序に戻ります。

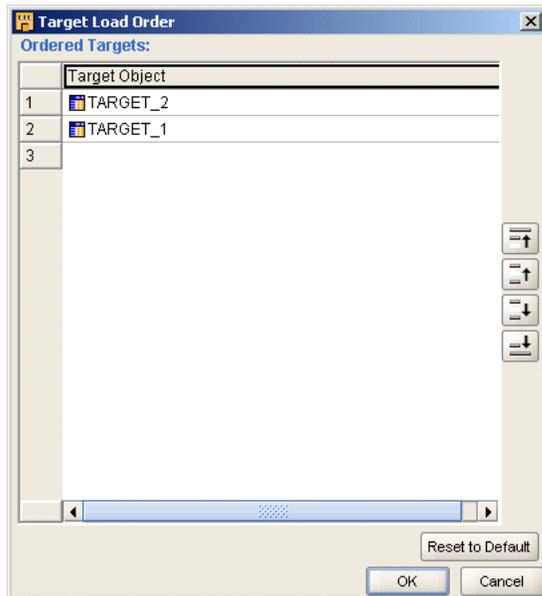
複数のターゲットのロード順序を指定する手順

- マッピング・キャンバスの空白をクリックすると、プロパティ・ウィンドウの右上隅にマッピング・プロパティが表示されます。

- 「ターゲットのロード順序のマップ」プロパティに進み、右側にある省略記号ボタンをクリックします。

図 5-6 に示すように、Oracle Warehouse Builder によって、「ターゲット・ロード順序」ダイアログ・ボックスが表示されます。このダイアログ・ボックスでは、TARGET2 ロードが TARGET1 の前に表示されています。

図 5-6 「ターゲット・ロード順序」ダイアログ・ボックス



- ロード順序を変更するには、ターゲットを選択し、右側のシャトル・ボタンを使用してリスト内でターゲットを上または下に移動します。

デフォルトにリセット

「デフォルトにリセット」ボタンは、ターゲット・ロード順序を再計算する場合に使用します。ターゲットの順序変更を誤った場合、または順序を割り当てた後に当初の順序が無効になるようなマッピングの設計変更があった場合は、再計算が必要になります。

演算子、グループおよび属性プロパティの設定

キャンバスでオブジェクトを選択すると、エディタの左側にあるプロパティ・インスペクタに、関連するプロパティが表示されます。

図 5-7 に、表演算子のプロパティ・インスペクタを示します。

図 5-7 表演算子のプロパティ・インスペクタ

次のタイプのプロパティを表示および設定できます。

- **演算子のプロパティ**: 演算子全体に影響を与えるプロパティです。設定できるプロパティは演算子のタイプに依存します。
- **グループのプロパティ**: これは属性のグループに影響を与えるプロパティです。ほとんどの演算子には、グループのプロパティが設定されていません。グループ・プロパティが設定される演算子の例としては、スプリッタ演算子やデブリケータ解除演算子があります。
- **属性のプロパティ**: これはソース演算子とターゲット演算子の属性に関するプロパティです。属性のプロパティの例としては、データ型、精度、スケールがあります。

演算子とワークスペース・オブジェクトの同期化

マッピングに使用する多くの演算子に対応する定義は、Oracle Warehouse Builder ワークスペースにあります。このことは、表演算子やビュー演算子など、ソース演算子とターゲット演算子に当てはまります。また、順序演算子や変換演算子など、定義が複数のマッピングにまたがって使用される他の演算子にも当てはまります。このような演算子を変更するときは、変更内容をワークスペース・オブジェクトに伝播する必要があります。

変更の伝播方向は、次の選択肢から決定できます。

ワークスペース・オブジェクトから演算子への同期化: 本番環境でマッピングの使用を開始した後、ソースまたはターゲットに対して ETL 設計に影響する変更が行われる場合があります。通常、このような変更を管理する最善の方法は、Oracle Warehouse Builder のオンライン・ヘルプで説明されているように、Oracle Warehouse Builder の依存性マネージャを使用することです。依存性マネージャを使用すると、変更による影響が自動的に評価され、影響を受けるすべてのマッピングがまとめて同期化されます。また、5-12 ページの「ワークスペース・オブジェクトから演算子への同期化」で説明されているように、マッピング・エディタで、オブジェクトを手動で同期化することもできます。

演算子からワークスペース・オブジェクトへの同期化: マッピングの演算子を変更する場合は、変更内容に対応するワークスペース定義に伝播させます。たとえば、インポートしてマッピングに使用したソースの場合、属性に複合物理名が付いている場合があります。

同期化がリフレッシュとは異なることに注意してください。リフレッシュ・コマンドでは、マルチユーザー環境で他のユーザーが行った変更にあわせて、確実に最新の状態が保たれます。同期化では、対応するワークスペース・オブジェクトと演算子が照合されます。

演算子の同期化

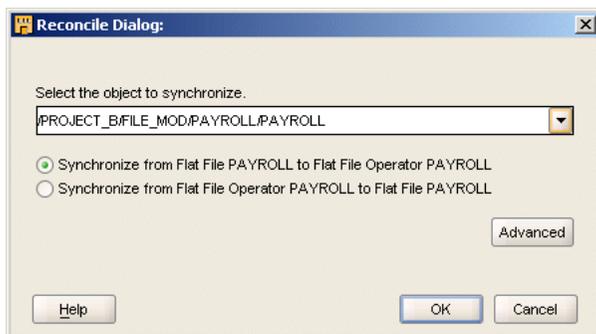
同期化する場合は、演算子を1つ選択し、指定したワークスペース・オブジェクトの定義とその演算子を同期化します。

演算子を同期化する手順

1. マッピング・エディタのキャンバスで演算子を選択します。
2. 「編集」メニューから、「同期化」を選択するか、演算子のヘッダーを右クリックして「同期化」を選択します。

図 5-8 に示す「演算子の同期化」ダイアログ・ボックスが表示されます。

図 5-8 演算子の同期化



3. デフォルトでは、Oracle Warehouse Builder により、選択した演算子とワークスペース内の関連するオブジェクトを同期化するためのオプションが選択されます。デフォルトを受け入れるか、リスト・ボックスから別のワークスペース・オブジェクトを選択します。

この手順では、「ワークスペース・オブジェクトから演算子に同期」、または「演算子からワークスペース・オブジェクトに同期」のいずれかを選択します。

4. (オプション)「拡張」をクリックして「一致方針」を設定します。
「一致方針」の使用方法を参照するには、「ヘルプ」を選択します。
5. 「OK」をクリックします。

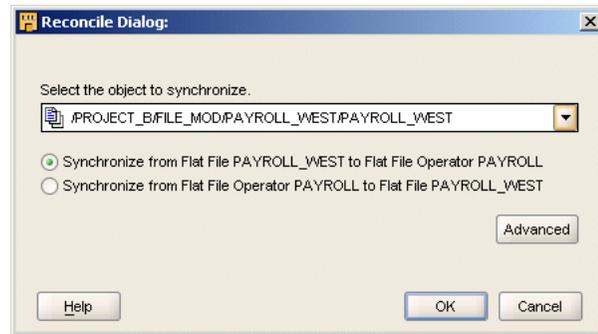
ワークスペース・オブジェクトから演算子への同期化

マッピング・エディタでは、次のような理由により、ワークスペース・オブジェクトから演算子への同期化を実行できます。

- **手動による変更の伝播:** ワークスペース・オブジェクトの変更を関連する演算子に伝播します。ワークスペース・オブジェクトの変更には、構造的な変更、属性名の変更、属性のデータ型の変更などがあります。ワークスペース・オブジェクトの変更を複数のマッピングに自動的に伝播する方法は、Oracle Warehouse Builder のオンライン・ヘルプを参照してください。
- **演算子と新規ワークスペース・オブジェクトの同期化:** たとえば、マッピングをデータ・ウェアハウスの特定のバージョンから新しいバージョンに移行し、各バージョンの別のオブジェクト定義を維持する場合、演算子と新規のワークスペース・オブジェクトを関連付けることができます。

図 5-9 に、フラット・ファイル演算子と新規ワークスペース・オブジェクトとの同期化の例を示します。

図 5-9 異なるワークスペース・オブジェクトからの同期化



- **表によるマッピングのプロトタイプ作成**: 設計環境で作業している場合は、表を使用した ETL ロジックの設計を選択できます。ただし、本番環境では、他のワークスペース・オブジェクト・タイプ（ビュー、マテリアライズド・ビュー、キューブなど）に基づいたマッピングが可能です。

ワークスペース・オブジェクトに基づく演算子の同期化

表 5-1 に、演算子および同期化できるワークスペース・オブジェクトのタイプを示します。

表 5-1 ワークスペース・オブジェクトと同期化される演算子

| 同期対象: 演算子 | 同期基準: ワークスペース・オブジェクトのタイプ |
|-----------------|---|
| キューブ演算子 | 表、ビュー、マテリアライズド・ビュー、フラット・ファイル、ディメンションおよびキューブ |
| ディメンション演算子 | 表、外部表、ビュー、マテリアライズド・ビュー、フラット・ファイル、ディメンションおよびキューブ |
| 外部表演算子 | 表、外部表、ビュー、マテリアライズド・ビュー、フラット・ファイル、ディメンションおよびキューブ |
| フラット・ファイル演算子 | 表、外部表、ビュー、マテリアライズド・ビュー、フラット・ファイル、ディメンションおよびキューブ |
| キー参照演算子 | 表のみ |
| マテリアライズド・ビュー演算子 | 表、外部表、ビュー、マテリアライズド・ビュー、ファイル、ディメンションおよびキューブ |
| マッピング後プロセス演算子 | 変換のみ |
| マッピング前プロセス演算子 | 変換のみ |
| 順序演算子 | 順序のみ |
| 表演算子 | 表、外部表、ビュー、マテリアライズド・ビュー、フラット・ファイル、ディメンションおよびキューブ |
| 変換演算子 | 変換のみ |
| ビュー演算子 | 表、外部表、ビュー、マテリアライズド・ビュー、ファイル、ディメンションおよびキューブ |

外部表演算子から同期化するとき、Oracle Warehouse Builder は、関連付けられたフラット・ファイルではなく、ワークスペース外部表のみに基づいて演算子を更新することに注意してください。

演算子からワークスペース・オブジェクトへの同期化

マッピングの演算子を変更した場合は、それらの変更をワークスペース・オブジェクトに伝播する必要があります。同期化によって、表、ビュー、マテリアライズド・ビュー、変換およびフラット・ファイルの各演算子の変更を伝播できます。

次のような理由により、演算子からワークスペース・オブジェクトへの同期化を実行します。

- **変更の伝播:** 演算子の変更を関連するワークスペース・オブジェクトに伝播します。演算子または属性のビジネス名を変更すると、Oracle Warehouse Builder によってそのビジネス名の最初の 30 文字がバウンド名として伝播されます。
- **ワークスペース・オブジェクトの置換:** 同期化により、既存のワークスペース・オブジェクトを置換します。

演算子からの同期化は、他の演算子とワークスペース・オブジェクトとの依存関係には影響を与えません。

6

ロード・パフォーマンスの向上

この項は、パフォーマンス要件を満たす ETL ロジックを作成するためのガイドとして使用します。

パーティション交換ロードを使用したパフォーマンスの向上

データ・パーティションを使用すると、ターゲット・システムでデータをロードまたは削除するときのパフォーマンスが向上します。この方法は、パーティション交換ロード (PEL) と呼ばれています。

PEL は、比較的少量のデータを、大量の履歴データを含むターゲットにロードする際に使用してください。ターゲットにできるものとしては、データ・ウェアハウスの表、ディメンション、キューブなどがあります。

パーティション交換ロードについて

ターゲット・システムのパーティションを操作すると、パーティション交換ロード (PEL) を使用して、即時にデータを追加または削除できます。表が空のパーティションで交換されると、新規データが追加されます。

パーティション交換ロードを使用すると、新規データをパーティションとしてターゲット表と交換し、ロードできます。たとえば、新規データを保持する表は、パーティションのアイデンティティをターゲット表から引き継ぎ、このパーティションはソース表のアイデンティティを引き継ぎます。この交換プロセスは、実際のデータの移動を伴わない DDL 操作です。

図 6-1 パーティション交換ロードの概要

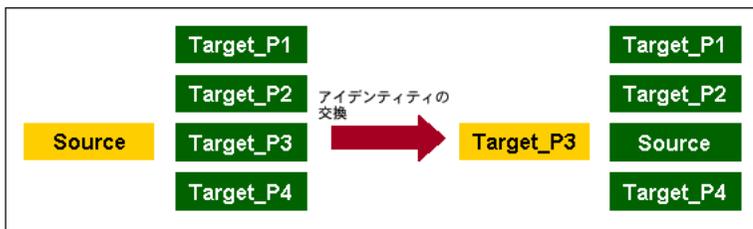


図 6-1 では、ソース表 Source のデータが、4つのパーティション (Target_P1、Target_P2、Target_P3 および Target_P4) で構成されたターゲット表に挿入されます。新規データを Target_P3 にロードする必要がある場合、パーティション交換操作では、実際のデータを移動せずに、データ・オブジェクトの名前のみを交換します。交換後は、以前の Source という名前が Target_P3 に変更され、この Target_P3 という名前が Source になります。ターゲット表には、Target_P1、Target_P2、Target_P3 および Target_P4 の4つのパーティションが含まれたままです。Oracle9i で使用可能なパーティション交換操作は、データを移動することなくロード・プロセスを完了します。

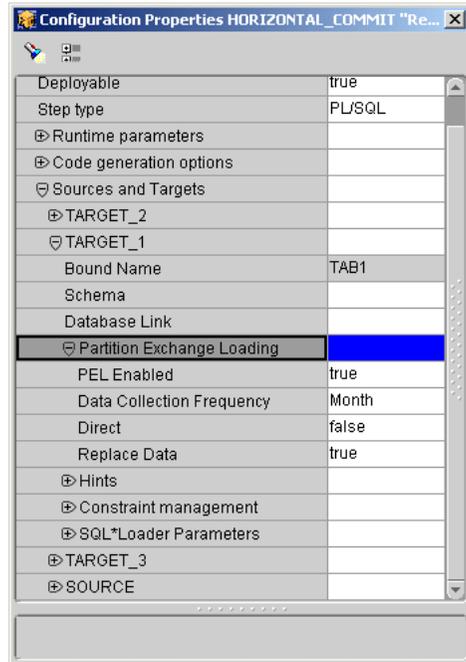
パーティション交換ロードのマッピングの構成

パーティション交換ロードのマッピングを構成する手順

1. プロジェクト・エクスプローラで、マッピングを右クリックして「構成」を選択します。

Oracle Warehouse Builder によって「構成プロパティ」ウィンドウが表示されます。

図 6-2 PEL 構成プロパティ



2. デフォルトでは、パーティション交換ロードはすべてのマッピングに対して無効です。パーティション交換ロードを使用するには、「**PEL 有効**」を選択します。
3. 「**データ・コレクション頻度**」を使用して、マッピングを実行するたびに収集する新規データの量を指定します。このパラメータを設定して、年、四半期、月、日、時間、分ごとのデータの収集を指定します。これによって、パーティション数が決定されます。
4. 収集したデータをステージングする一時表を作成する場合は、パーティション交換を実行する前に、「**ダイレクト**」を選択します。このパラメータを選択しないと、一時表は作成されないまま、ソース表がパーティションとしてターゲット表に直接スワップされます。
5. 「**データの置換**」を選択すると、Oracle Warehouse Builder によって、ターゲット・パーティションにある既存のデータが新しく収集したデータで置換されます。このパラメータを選択しないと、ターゲット・パーティションには、既存のデータが保持されます。新規データは、空ではないパーティションに挿入されます。このパラメータは、ローカル・パーティションに影響を与え、ターゲット表のパーティションの削除またはスワップに使用できます。「TRUNCATE/INSERT」プロパティは、表レベルで設定できます。

ダイレクトおよび間接パーティション交換ロード

Oracle Warehouse Builder を使用して、パーティション交換によってターゲットをロードする場合は、ターゲットを間接的または直接的にロードできます。

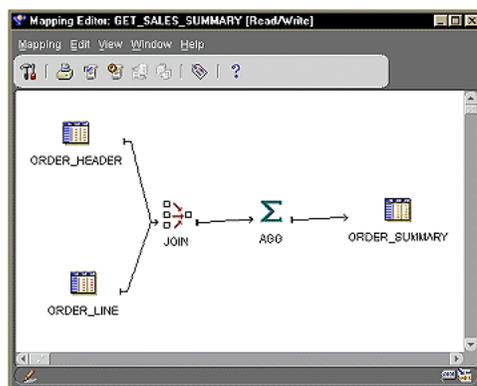
- **間接パーティション交換ロード**: デフォルトでは、パーティション交換プロセスを開始する前に、Oracle Warehouse Builder によってソース・データをステージングする一時表が作成され、維持されます。たとえば、マッピングにリモート・ソースまたは複数ソースの結合が含まれている場合は、間接 PEL を使用します。
- **ダイレクト・パーティション交換ロード**: ターゲット構造に一致するマッピングのソースを設計します。たとえば、マッピングでダイレクト PEL を使用すると、以前実行したマッピングでロードしたファクト表をすぐに公開できます。

間接パーティション交換ロードの使用

パーティション交換ロードを使用して設計したマッピングにリモート・ソースまたは複数ソースの結合が含まれている場合、パーティション交換を続行する前に、ソース処理が実行され、データがステージングされる必要があります。したがって、このようなマッピングは、ダイレクト PEL を False に設定した状態で構成します。ソース処理の結果を格納する一時表は、Oracle Warehouse Builder によって透過的に作成され、維持されます。この表は、PEL の実行後に Oracle Warehouse Builder によって削除されます。

図 6-3 は、2つのソースを結合し、集計を実行するマッピングを示しています。ORDER_SUMMARY 表にロードされたすべての新規データが常に同じパーティションにロードされる場合は、このマッピングで間接 PEL を使用して、ロードのパフォーマンスを改善できます。この場合、Oracle Warehouse Builder によって、アグリゲータと ORDER_SUMMARY の間に一時表が透過的に作成されます。

図 6-3 複数ソースでのマッピング



Oracle Warehouse Builder では、ターゲット表と同じ構造（列、索引、制約が同じ）を使用して、一時表が作成されます。この一時表は、パフォーマンス向上のために、Oracle Warehouse Builder によってパラレル直接パス・ロード INSERT を使用してロードされます。INSERT の後、Oracle Warehouse Builder では、一時表の索引と制約がパラレルで作成されます。

例：ファクト表を公開するためのダイレクト・パーティション交換ロードの使用

ソース表がローカルで、データの品質が高い場合は、ダイレクト PEL を使用します。マッピングは、ソースとターゲットが同じデータベースにあり、完全に同じ構造であるように設計する必要があります。ソースとターゲットの索引、制約、列数、列タイプおよび列の長さは、すべて同じである必要があります。

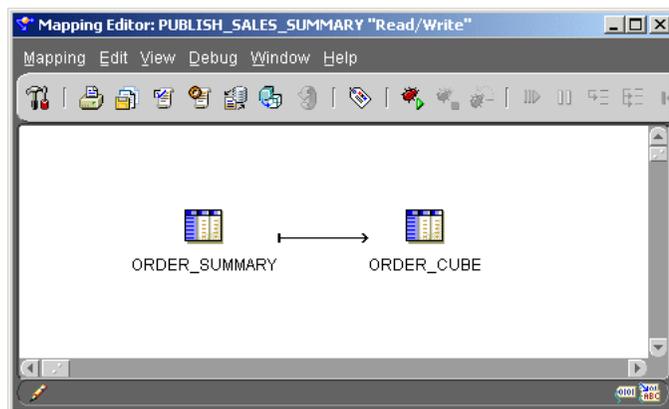
たとえば、図 6-3 と同じマッピングがあり、データをターゲットにロードするタイミングをより細かく制御するとしてします。データ量によってはロードに数時間がかかり、現状では、ターゲット表がいつ更新されるかも正確にはわかりません。

ダイレクト PEL を使用してデータを即時にターゲットにロードする手順

1. ソース・データを結合し、必要に応じてデータを変換し、確実に検証してステージング表にロードするマッピングを 1 つ設計します。このマッピングは PEL を使用するよう構成しないでください。

ステージング表を、別のマッピングでロードする最終ターゲットの構造に正確に一致するように設計します。たとえば、[図 6-3](#)にあるステージング表の ORDER_SUMMARY は、[図 6-4](#)にある最終ターゲット ORDER_CUBE と同じ構造である必要があります。

2. [図 6-4](#)に示すように、ステージング表から最終ターゲットにデータをロードする 2 番目のマッピングを作成します。このマッピングはダイレクト PEL を使用するよう構成してください。

図 6-4 Publish_Sales_Summary マッピング

3. Oracle Warehouse Builder のプロセス・フロー・エディタまたは Oracle Workflow を使用して、最初のマッピングの完了後に 2 番目のマッピングを起動します。

パーティション交換ロードの効果的な使用

次の条件が True の場合は、スケーラビリティの高いロード・パフォーマンスのためにパーティション交換ロードを効果的に使用できます。

- **表のパーティション化と表領域:** ターゲット表は、1 つの DATE 列ごとにレンジ・パーティション化される必要があります。すべてのパーティションは同じ表領域内に作成される必要があります。すべての表は同じ表領域内に作成されます。
- **既存の履歴データ:** ターゲット表には、大量の履歴データが格納される必要があります。PEL の使用例には、ターゲットが OLTP データベースや Web ログ・ファイルからデータを毎日収集するクリック・ストリーム・アプリケーションがあります。新規データは変換され、履歴データが格納されているターゲットにロードされます。
- **新規データ:** すべての新規データは、ターゲット表の同じパーティションにロードされる必要があります。たとえば、ターゲット表が日別に区切られている場合、日別のデータは 1 つのパーティションにロードされる必要があります。
- **ロード頻度:** ロード頻度は、データ収集頻度以下である必要があります。
- **グローバル索引なし:** ターゲット表にはグローバル索引は含めないでください。

マッピングでのターゲットの構成

パーティション交換ロードに対してマッピングでのターゲットを構成する手順

- 手順 1: すべてのパーティションの作成
- 手順 2: 「LOCAL」 オプションを使用したすべての索引の作成
- 手順 3: 「索引を使用」 オプションを使用する主キーまたは一意キー

手順 1: すべてのパーティションの作成

パーティションは、実行中に自動的に作成されません。すべてのパーティションは、パーティション交換ロードを使用する前に、すべてのパーティションを作成する必要があります。

たとえば、新しいデータ収集の頻度として「月」を選択する場合、新しいデータの各月に対して必要なパーティションをすべて作成する必要があります。データ・オブジェクト・エディタを使用して表、ディメンションまたはキューブのパーティションを作成します。

図 6-5 では、表 ORDER_SUMMARY に対するプロパティ・インスペクタ・ウィンドウが表示されます。

パーティション交換ロードを使用するには、すべてのパーティションがネーミング規則に従って命名される必要があります。たとえば、2002 年 5 月のデータを保持するパーティションの場合、そのパーティション名は Y2002_Q2_M05 という形式にする必要があります。

PEL でパーティションを識別するには、次のいずれかの形式の名前をパーティションに指定する必要があります。

Ydddd

Ydddd_Qd

Ydddd_Qd_Mdd

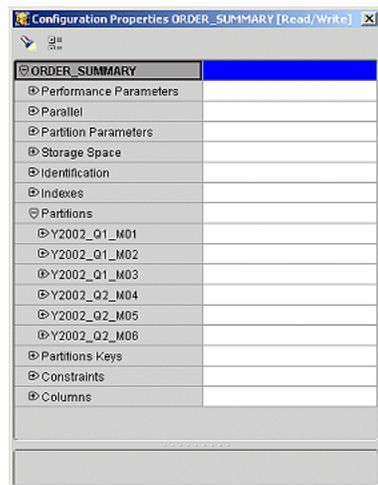
Ydddd_Qd_Mdd_Ddd

Ydddd_Qd_Mdd_Ddd_Hdd

Ydddd_Qd_Mdd_Ddd_Hdd_Mdd

d には、10 進数の数字が入ります。すべての文字は、大文字にする必要があります。小文字は認識されません。

図 6-5 表 ORDER_SUMMARY の構成プロパティ



各パーティションを正しく命名すると、Oracle Warehouse Builder によって各パーティションの「上限値」プロパティが自動的に計算されます。そうでない場合は、Oracle Warehouse Builder で DDL 文が作成されるように、上限値を各パーティションに対して手動で構成する必要があります。次に、Oracle Warehouse Builder によって作成された DDL 文の例を示します。

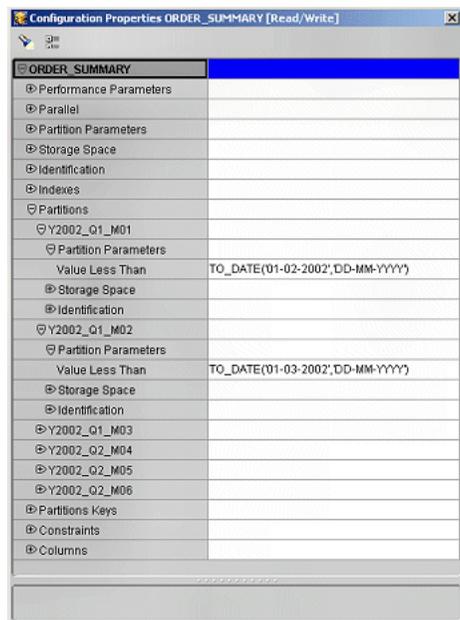
```

. . .
PARTITION A_PARTITION_NAME
VALUES LESS THAN (TO_DATE('01-06-2002', 'DD-MM-YYYY')),
. . .

```

図 6-6 では、「上限値」パラメータの自動的に生成された構成値が表示されます。

図 6-6 自動生成された「上限値」設定



手順 2: 「LOCAL」オプションを使用したすべての索引の作成

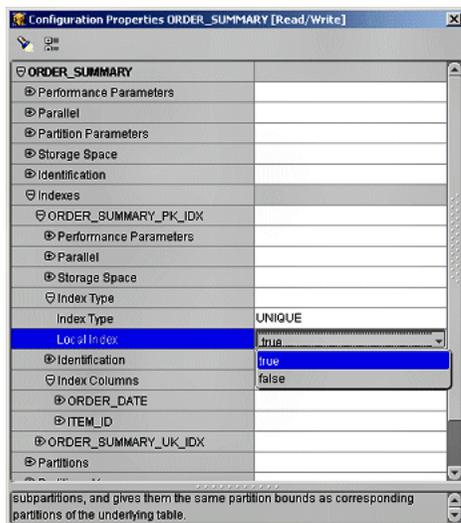
図 6-7 では、ORDER_SUMMARY 表に追加された索引 (ORDER_SUMMARY_PK_IDX) を示しています。この索引には、2つの列、ORDER_DATE と ITEM_ID があります。次のように構成します。

- 「索引タイプ」パラメータを UNIQUE に設定します。
- 「ローカル索引」パラメータを true に設定します。

これで、Oracle Warehouse Builder で、表 ORDER_SUMMARY に一意のローカル索引を作成する DDL 文を作成できます。

ローカル索引を使用すると、PEL のパフォーマンスを最大限に活用できます。ローカル索引では、すべての索引が表と同じ方法でパーティション化されている必要があります。PEL を使用して一時表をターゲット表にスワップすると、索引セグメントのアイデンティティもスワップされます。

図 6-7 ローカル索引としての索引の構成



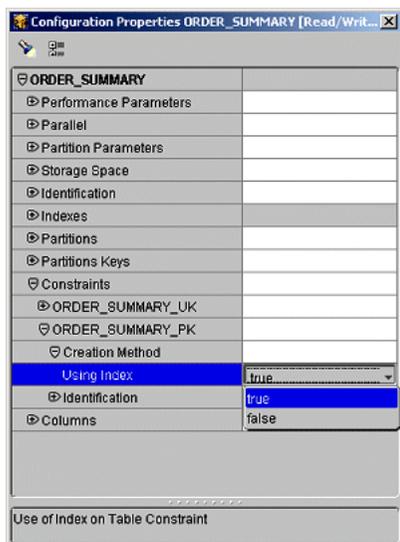
ローカル索引として索引が作成されている場合、Oracle サーバーでは、パーティション・キー列を索引の最初の列にする必要があります。前述の例では、パーティション・キーは ORDER_DATE で、索引 ORDER_SUMMARY_PK_IDX の最初の列になります。

手順 3: 「索引を使用」オプションを使用する主キーまたは一意キー

この手順では、主キー制約および一意キー制約がすべて、「索引を使用」オプションで作成されるように指定する必要があります。図 6-8 では、ORDER_SUMMARY 表の主キー制約 ORDER_SUMMARY_PK が、「索引を使用」オプションで指定されている例を示しています。

「索引を使用」オプションを使用すると、表に制約を追加した際に索引が自動的に作成されなくなります。サーバーは、制約と同じリストを持つ既存の索引を検索します。したがって、主キー制約または一意キー制約は、ユーザー定義の一意のローカル索引で補足する必要があります。制約 ORDER_SUMMARY_PK に必要な索引は、「手順 2: 「LOCAL」オプションを使用したすべての索引の作成」で作成した ORDER_SUMMARY_PK_IDX です。

図 6-8 「索引を使用」オプションによる制約の指定



Oracle Warehouse Builder でのパーティション交換ロードの使用の制限

Oracle Warehouse Builder での PEL の使用には、次の制限があります。

- **日付パーティション・キーは1つのみ**：許容される DATE データ型のパーティション・キー列は1つのみです。数値パーティション・キーは、Oracle Warehouse Builder にはサポートされていません。
- **一般的なカレンダーのみ**：現在の PEL メソッドは、世界中で採用されている一般的なカレンダーのみをサポートしています。ユーザー定義の会計年度別や四半期別のビジネス・カレンダーは、現時点ではサポートされていません。
- **すべての日付パーティションが同じ表領域内にあること**：ターゲット（表、ディメンションまたはキューブ）のすべてのパーティションは、同じ表領域内に作成される必要があります。
- **すべての索引パーティションが同じ表領域内にあること**：ターゲット（表、ディメンションまたはキューブ）のすべての索引は、同じ表領域内に作成される必要があります。ただし、索引とデータでは、別の表領域を使用できます。

ターゲット・スキーマへの配布および ETL ロジックの実行

この章の内容は次のとおりです。

- [配布について](#)
- [スケジュールについて](#)
- [スケジュールの定義および使用のプロセス](#)

配布について

配布は、Oracle Warehouse Builder リポジトリ内の論理オブジェクトから、ターゲット・ロケーション内に物理オブジェクトを作成するプロセスです。

マッピングまたはプロセス・フローの配布には、次の手順が含まれます。

- 必要に応じて、PL/SQL、SQL*Loader または ABAP スクリプトを生成します。
- スクリプトをデザイン・センターからコントロール・センターにコピーします。また、SQL*Loader 制御ファイルもコントロール・センターにコピーします。
- 新しいコネクタを配布します。つまり、ロケーション間にデータベース・リンクとデータベース・ディレクトリを作成します。

マッピングまたはプロセス・フローの配布の後、スクリプトを明示的に開始する必要があります。

配布できるのは、ユーザーに COMPILE 権限があるオブジェクトのみです。デフォルトでは、リポジトリ内のすべてのオブジェクトに対してこの権限があります。ただし、リポジトリの所有者が異なるセキュリティ・ポリシーを設定している場合があります。

デザイン・センターのナビゲーション・ツリーから直接配布するか、またはコントロール・センター・マネージャを使用して配布できます。

注意： オブジェクトは常に Oracle Warehouse Builder を使用して保守してください。配布した物理オブジェクトを SQL で手動で変更しないでください。そうしないと、論理オブジェクトと物理オブジェクトの同期がとれなくなり、予測できない結果を引き起こす可能性があります。

注意： オブジェクトを配布するたびに、Oracle Warehouse Builder では、すべての設計オブジェクトに対するすべての変更がリポジトリに自動的に保存されます。「プリファレンス」ダイアログ・ボックスの「コミットを要求」を選択すると、警告メッセージの表示を選択できます。

コントロール・センターの概要

コントロール・センターには、すべての配布に関する詳細情報が格納されます。次の情報が含まれており、オブジェクト単位またはジョブ単位でアクセスできます。

- 各オブジェクトの現在の配布ステータス
- 各オブジェクトに対するすべての配布試行の履歴
- 各マッピングおよびプロセス・フローに対するすべての ETL 開始試行の履歴
- すべての配布ジョブからのメッセージの完全ログ

コントロール・センターは、ターゲット・ロケーションと同じデータベース内にスキーマとして実装されています。各リポジトリにはデフォルトのコントロール・センターがあり、これは、リポジトリ所有者のスキーマ内に作成されています。たとえば、REP_OWNER リポジトリの所有者には、REP_OWNER というスキーマがあり、このスキーマには、デザイン・センターとデフォルトのコントロール・センターの両方からのメタデータが格納されます。

ローカル・システムに配布するには、デフォルトのコントロール・センターを使用できます。または、別のシステムに配布するために、追加のコントロール・センターを作成できます。一度に1つのコントロール・センターのみアクティブにできます。

コントロール・センター・マネージャでは、配布のすべての側面を表示および管理できる総合的な配布コンソールが提供されます。ここから、アクティブなコントロール・センターに格納されている情報にアクセスできます。

リポジトリ・ブラウザを使用して、配布データ・レポートにアクセスすることもできます (Oracle Warehouse Builder のオンライン・ヘルプを参照)。

新規コントロール・センターを作成する手順

1. 接続エクスプローラで、「コントロール・センター」を右クリックして「新規」を選択します。

「コントロール・センターの作成」ダイアログ・ボックスが表示されます。

2. このダイアログ・ボックスを完了します。詳細は、「ヘルプ」ボタンをクリックしてください。

コントロール・センターは、構成の作成ウィザードを使用して作成することもできます。

コントロール・センターをアクティブにする手順

1. コントロール・センターを使用するように、構成を作成または編集します。

「新規構成の作成」を参照してください。

2. その構成をアクティブ化します。

配布の物理詳細の構成

Oracle Warehouse Builder では、オブジェクトの論理設計が配布の物理詳細から分離されています。この分離は、構成パラメータに物理詳細を格納することによって実現します。構成と呼ばれるオブジェクトには、すべての構成設定が格納されます。各配布ロケーションには、それぞれのオブジェクト・パラメータに異なる設定を指定して、異なる構成を作成できます。

配布前に、ターゲット・オブジェクト、マッピングおよびモジュールの構成を確認してください。

オブジェクトが配布可能になる条件は、次のとおりです。

- そのターゲット・ロケーションが完全に定義され、有効であり、オブジェクトのモジュールに対して選択されている必要があります。
- その配布可能パラメータが選択されている（デフォルト）必要があります。
- 検証および生成がエラーなしに実行されている必要があります。

配布アクション

デザイン・センターで定義した新規オブジェクトは、ただちにコントロール・センター・マネージャにリストされます。各オブジェクトには表示可能なデフォルトの配布アクションがあります。このデフォルトは以前のアクションによって設定され、オブジェクトのタイプによって異なります。デフォルトは、コントロール・センター・マネージャで別の配布アクションを選択することで上書きできます。

次の配布アクションがあります。

- **作成**: ターゲット・ロケーションにオブジェクトを作成します。指定した名前のオブジェクトがすでに存在している場合は、エラーが発生する可能性があります。
- **アップグレード**: 可能な場合、データを喪失することなくオブジェクトを変更します。アップグレードは、元に戻して再実行できます。このアクションは、スケジュールなど、一部のオブジェクト・タイプには使用できません。
- **削除**: ターゲット・ロケーションからオブジェクトを削除します。
- **置換**: オブジェクトを削除して再作成します。このアクションは、アップグレードより高速ですが、データはすべて削除されます。

配布プロセス

データ・システムのライフサイクル期間中は、通常、配布プロセスで次の手順を実行します。

1. オブジェクト設定および使用するコントロール・センターが指定された構成を選択します。
2. オブジェクトをターゲット・ロケーションに配布します。個別に配布、1 つずつ順に配布、すべて一度に配布のいずれも可能です。
3. 配布の結果を確認します。オブジェクトの配布に失敗した場合は、問題を修正し、再試行してください。
4. ETL プロセスを開始します。
5. ユーザーの要求やソース・データの変更などにあわせて、ターゲット・オブジェクトの設計を変更します。
6. 変更したオブジェクトの配布アクションを「アップグレード」または「置換」に設定します。
7. これらの手順を繰り返します。

注意： Oracle Warehouse Builder では、配布前に、リポジトリに対するすべての変更が自動的に保存されます。

スケジュールについて

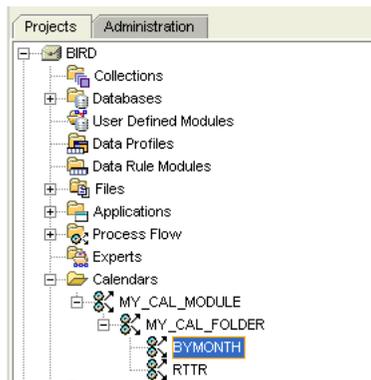
スケジュールを使用して、Oracle Warehouse Builder で設計した操作の実行時期および実行頻度を計画します。スケジュールは、Oracle Database リリース 10g 以上で実行するマッピングおよびプロセス・フローに適用できます。

Oracle Warehouse Builder を使用する開発フェーズでは、マッピングとプロセス・フローをスケジュールリングするかわりに、「コントロール・センター」からすぐに起動および停止します。

スケジュールは、1 度実行するように定義するか、またはユーザー・インタフェースで定義する間隔に基づいて繰り返し実行するように定義できます。定義する各スケジュールに対して、Oracle Warehouse Builder により、iCal カレンダー標準に準拠するコードが生成されます。

図 7-1 に示すように、プロジェクト・エクスプローラでは、スケジュールはプロジェクトのコンテキストで定義され「スケジュール」ノードの下にスケジュール・モジュールに含まれます。

図 7-1 プロジェクト・エクスプローラ上のスケジュール



Oracle Warehouse Builder では、新規に作成するプロジェクトごとにデフォルトのスケジュール・モジュール MY_CAL_MODULE が作成されます。デフォルト・モジュールの下にスケジュールを作成するか、「スケジュール」ノードを右クリックして「新規」を選択し、新規モジュールを作成します。

Oracle Workflow への Oracle Warehouse Builder スケジュールの配布

Oracle Warehouse Builder スケジュールを Oracle Workflow に正常に配布するには、『Oracle Warehouse Builder インストレーションおよび管理ガイド』で説明するように、正しいバージョンの Oracle Workflow にアクセスします。

ターゲット・ロケーションが 'CREATE SYNONYM' システム権限を持っていることを確認します。「評価ロケーション」が指定されているか、配布ロケーションが「コントロール・センター」スキーマと異なるデータベース・インスタンスを参照している場合、配布ロケーションには 'CREATE DATABASE LINK' システム権限が必要です。

Oracle Warehouse Builder でジョブをスケジュールするためのオプション

この項では、Oracle Warehouse Builder ETL プロセスをスケジュールするときに優先される方法を示します。

スケジュールの定義および使用のプロセス

1. スケジュールを含むモジュールを作成するには、「スケジュール」ノードを右クリックして「新規」を選択します。
2. スケジュールを作成するには、スケジュール・モジュールを右クリックして「新規」を選択します。

Oracle Warehouse Builder により、「スケジュール」ウィザードが起動します。

3. **名前と説明** ページで、24 文字以内でスケジュールの名前を入力します。

Oracle Warehouse Builder オブジェクトでは、ほとんどの場合、物理名は 1 から 30 文字までの英数字、およびビジネス名は 1 から 2000 文字までの英数字を使用できるというルールがあります。

4. 「スケジュール」ウィザードの指示に従ってください。

ウィザードを使用して、「開始日」、「終了日」、「開始時間」、「終了時間」、「頻度単位」、および「繰り返し」の値を指定します。ウィザードを完了すると、Oracle Warehouse Builder により、選択したスケジュール・モジュールの下にスケジュールが保存されます。

5. 「プロジェクト・エクスプローラ」で、ウィザードを使用して作成したスケジュールを右クリックして「エディタを開く」を選択します。

Oracle Warehouse Builder にスケジュール・エディタが表示されます。選択内容をレビューして計算済実行時間の一覧を表示します。複合スケジュールの場合、「句」の値を入力できます。

6. マッピングまたはプロセス・フローにスケジュールを適用するには、「プロジェクト・エクスプローラ」のオブジェクトの上で右クリックし、「構成」を選択します。「参照カレンダー」フィールドでは、省略記号ボタンを選択して既存のスケジュールを表示します。

スケジュールするマッピングおよびプロセス・フローでは、物理名は 25 文字以内、ビジネス名は 1995 文字以内である必要があります。この制限により、Oracle Warehouse Builder では、配布および実行に必要な接尾辞 `_job` およびその他の内部文字をマッピング名に追加できるようになります。

7. スケジュールを配布します。

マッピングを配布する場合、たとえば関連するマッピングおよびプロセス・フローおよびターゲット・データ・オブジェクトも配布する必要があります。同様に、関連するスケジュールもすべて配布する必要があります。

関連するオブジェクトとともに正しく配布された場合、ターゲット・スキーマは作成したスケジュールに基づいてマッピングまたはプロセス・フローを実行します。

オブジェクトの配布

配布は、生成したコードを使用して、メタデータからターゲット・ロケーションに物理オブジェクトを作成するプロセスです。オブジェクトは、プロジェクト・エクスプローラから配布することも、コントロール・センター・マネージャを使用して配布することもできます。

Oracle Warehouse Builder では、オブジェクトを自動的に検証して生成します。

プロジェクト・エクスプローラからの配布は、デフォルト・アクションに制限されています。デフォルト・アクションは、「作成」、「置換」、「削除」または「更新」に設定されている可能性があります。このデフォルト・アクションを上書きするには、配布プロセスを完全に制御するコントロール・センター・マネージャを使用します。

プロジェクト・エクスプローラから配布する手順

該当するオブジェクトを選択して、ツールバーの「配布」アイコンをクリックします。

ステータス・メッセージが、「デザイン・センター」ウィンドウの下部に表示されます。配布が完了したことを通知する場合は、配布前にプリファレンスで「**配布完了メッセージの表示**」を選択します。

コントロール・センター・マネージャを開く手順

1. プロジェクトを開きます。
2. 「ツール」メニューから「**コントロール・センター・マネージャ**」を選択します。

ETL ジョブの開始

ETL は、データをそのソース・ロケーションから抽出し、マッピングで定義したとおりに変換し、ターゲット・オブジェクトにロードするプロセスです。ETL を開始するときは、データをジョブとして Oracle Warehouse Builder ジョブ・キューに送信します。このジョブは、すぐに開始できます。また、Oracle Enterprise Manager のスケジューラなど、スケジューラを使用する場合は、スケジュールされた時間に開始できます。ETL は、配布と同様に、プロジェクト・エクスプローラから開始することも、コントロール・センター・マネージャを使用して開始することもできます。SQL スクリプトを実行する Oracle Warehouse Builder 外部のツールをして、ETL を開始することもできます。

プロジェクト・エクスプローラから ETL を開始する手順

マッピングまたはプロセス・フローを選択して、「設計」メニューから「**開始**」を選択します。

データの表示

ETL を完了した後は、Oracle Warehouse Builder 内のデータ・オブジェクトを確認し、結果が期待どおりであることを簡単に検証できます。

データを表示する手順

プロジェクト・エクスプローラで、オブジェクトを右クリックして「**データ**」を選択します。「**データ・ビューア**」が開き、オブジェクトの内容が表示されます。

第 III 部

データ・ウェアハウスの分析および レポート作成

ここでは、データ・ウェアハウスの管理方法について説明します。内容は次のとおりです。

- [第 8 章「メタデータ依存性の分析」](#)
- [第 9 章「レポートおよび分析の SQL」](#)

メタデータ依存性の分析

メタデータ依存性マネージャを使用すると、Oracle Warehouse Builder リポジトリ内のオブジェクト定義またはメタデータに対して加えられた変更による影響を検出および解決できます。

この章の内容は次のとおりです。

- [メタデータ依存性マネージャについて](#)
- [使用方法のシナリオ](#)
- [系統および影響分析ダイアグラムについて](#)
- [LIA ダイアグラムの生成](#)
- [LIA ダイアグラムの表示の変更](#)
- [依存性マネージャでのオブジェクトの変更](#)

メタデータ依存性マネージャについて

メタデータ依存性マネージャでは、あらゆるデータ・オブジェクトの系統ダイアグラムおよび影響ダイアグラムが作成されます。系統ダイアグラムは、データ・ソースに戻るオブジェクトのプロセス・フローをトレースし、そのパスに存在するすべてのオブジェクトを表示します。影響ダイアグラムは、選択したオブジェクトから導出されるすべてのオブジェクトを識別します。

このタイプの情報は、多くの場面で役に立ちます。たとえば、これらのダイアグラムを次の場合に使用できます。

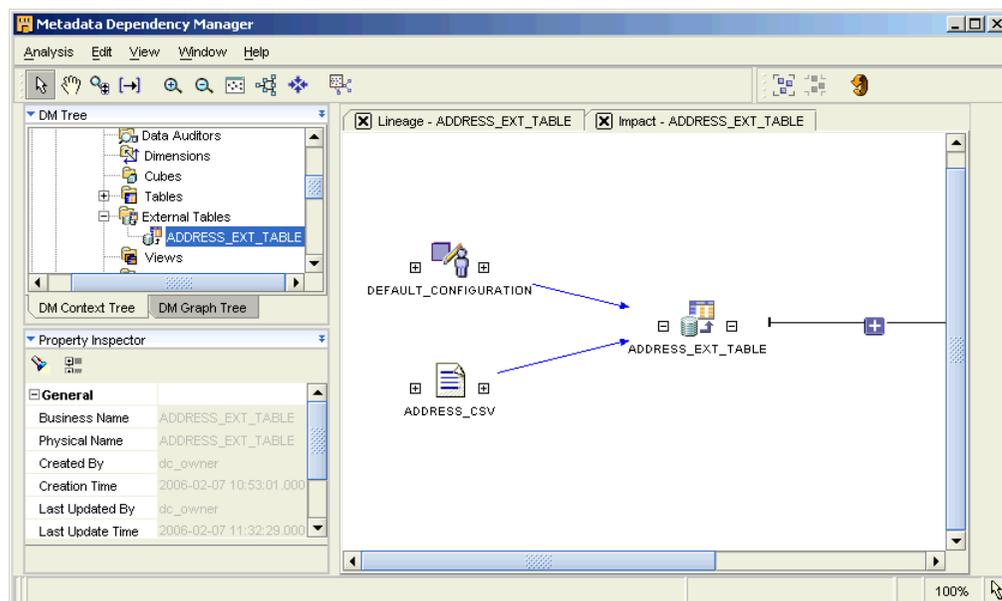
- 法的な規制または財務データのトレースを必要とするその他の監査に準拠する。
- ソース・データの変更に対応するためにシステム設計を変更する。
- レポートおよび分析の新しい要件に対応するためにシステム設計を変更する。
- プロジェクトを通して使用されるプラグガブル・マップの設計変更が及ぼす影響を評価する。

依存性マネージャを使用すると、将来にわたる変更が及ぼす影響をプレビューしながらプロジェクトを計画できます。ソース・システムに対する変更を計画している場合は、依存性マネージャを使用して、その変更がウェアハウス設計に与える影響を評価できます。または、すでに変更が行われている場合は、ETL 設計の更新とデータ・ウェアハウスの再構築に必要な時間を見積もることができます。

図 8-1 に、メタデータ依存性マネージャを示します。Oracle Warehouse Builder の他のウィンドウと同じように、依存性マネージャには、メニュー、ツールバー、ナビゲータ、プロパティ・インスペクタおよびキャンバスがあります。キャンバスには 1 つ以上のダイアグラムが表示されます。

これらのコンポーネントに固有の情報を参照するには、「ヘルプ」メニューの「トピック」を選択してください。

図 8-1 メタデータ依存性マネージャ



使用方法のシナリオ

データ・ウェアハウスの設計と構築が完了した後に、ソース・データベースが変更されることがよくあります。このような変更は、これに付随して Oracle Warehouse Builder 設計またはメタデータに対する変更の必要性や影響があることを意味します。Oracle Warehouse Builder を使用すると、変更済のソース定義をリポジトリに再インポートできます。ただし、元のウェアハウス設計は、再インポートした定義に対しては有効ではなくなります。そのため、設計を変更して不整合を修正する必要があります。

まず最初に、ソースの変更によってウェアハウス設計がどのように影響を受けるのかを見極める必要があります。次に、そのソースに依存しているすべての設計オブジェクトについてメタデータを同期し、影響を受けるすべてのオブジェクトを更新してソースの変更を反映します。このプロセスの完了後に、更新済の設計を再配布してデータ・ウェアハウスを再構築し、データを同期します。

このシナリオでは、会社は CUSTOMERS という名前のフラット・ファイルからすべての統計レポートを取得しています。これまでは、追加のパラメータのレポートを生成するにはファイル定義を変更する必要がありました。統計分析データベースは最近 Oracle9i から Oracle Database 10g に移行されました。

この会社の設計担当者は、まず最初に、設計リポジトリ内のこのファイルの変更済メタデータ定義を、そのデータを基にしているすべての設計オブジェクトおよびマッピングに同期する必要があります。CUSTOMERS フラット・ファイルは ETL 設計の複数のマッピングの一部であるため、この定義に対する変更はこれらのすべてを無効にします。CUSTOMERS フラット・ファイルの影響を受けるすべての領域を手動で追跡し、定義を更新することは、エラーが起りやすいプロセスです。

依存性マネージャは、データ・ソースに対する変更によって影響を受けるすべてのオブジェクトを識別することを可能にします。関連する各設計オブジェクトが確実に更新され、ETL 設計は有効に保たれます。このプロセスの完了後に、影響を受けた表および更新済の ETL 設計オブジェクトをデータ・ウェアハウスに再配布できます。

システムおよび影響分析ダイアグラムについて

システムおよび影響分析 (LIA) ダイアグラムは、Oracle Warehouse Builder で管理するオブジェクトの関係を示します。これらの関係はマッピングとプロセス・フローで構築されます。特定のオブジェクトのシステムダイアグラムはそのソース・オブジェクトを示し、影響ダイアグラムはそのターゲットを示します。

システムと影響は互いのミラー・イメージです。オブジェクト A がオブジェクト B のシステムダイアグラムの一部である場合、オブジェクト B はオブジェクト A の影響ダイアグラムの一部です。ダイアグラムを左から右に参照すると影響がわかり、右から左に参照するとシステムがわかります。

たとえば、データをファイルから抽出して外部表としての表にロードするマッピングがあるとします。この関係は次のようになります。

```
flat_file > external_table > table
```

図 8-2 に、ADDRESS_EXT_TABLE という名前の外部表のシステムダイアグラムを示します。ADDRESS_CSV はフラット・ファイルであり、ADDRESS_EXT_TABLE のシステムの一部です。つまり、ADDRESS_CSV に対する変更は ADDRESS_EXT_TABLE に影響を与えます。

図 8-2 ADDRESS_EXT_TABLE の系統分析ダイアグラム

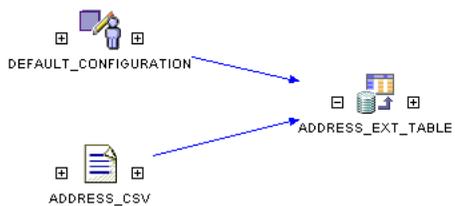


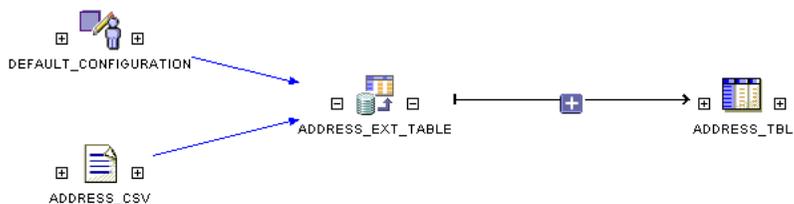
図 8-3 に、ADDRESS_EXT_TABLE の影響ダイアグラムを示します。これには ADDRESS_TBL が含まれています。ADDRESS_EXT_TABLE 表に加えられた変更は、ADDRESS_TBL に影響を与えます。ADDRESS_EXT_TABLE は ADDRESS_TBL の系統の一部です。

図 8-3 ADDRESS_EXT_TABLE の影響分析ダイアグラム



オブジェクト・アイコンの両側にあるいずれかのプラス記号 (+) (図 8-4 に示します) をクリックすると、オブジェクトの系統と影響の両方を簡単に表示できます。

図 8-4 系統および影響分析ダイアグラム



LIA ダイアグラムの生成

最初にメタデータ依存性マネージャを開くと、デザイン・センターのプロジェクト・エクスプローラから LIA ダイアグラムを生成できます。

デザイン・センターからダイアグラムを生成する手順

1. 分析対象のオブジェクトが表示されるまで、プロジェクト・エクスプローラを展開表示します。
2. オブジェクトを右クリックし、ポップアップ・メニューから「**系統**」または「**影響**」を選択します。

ダイアグラムが表示されたメタデータ依存性マネージャが開きます。

「系統」および「影響」コマンドは、「ビュー」メニューからも使用できます。

メタデータ依存性マネージャからダイアグラムを作成する手順

1. デザイン・センターの「ツール」メニューから、「**メタデータ依存性マネージャ**」を選択します。
メタデータ依存性マネージャが表示されます。
2. 分析対象のオブジェクトが表示されるまで、DM コンテキスト・ツリーを展開表示します。
3. オブジェクトを右クリックし、ポップアップ・メニューから「**系統の表示**」または「**影響の表示**」を選択します。

キャンバスにダイアグラムが表示されます。

LIA ダイアグラムの表示の変更

オブジェクトおよびダイアグラム・タイプの初期選択によって、最初の開始点およびそのダイアグラムのオブジェクトからのブランチの方向が決まります。次の手順で LIA ダイアグラムを変更できます。

- ダイアグラムに別のオブジェクトをドラッグ・アンド・ドロップします。
- オブジェクト・アイコンの隣のプラス記号 (+) およびマイナス (-) 記号をクリックして、ダイアグラムのブランチを展開表示または縮小表示します。
- 8-5 ページの「LIA ダイアグラムでのグループの使用」の説明に従って、グループ・ツールを使用してダイアグラムのセクションを単一のアイコンに縮小表示します。
- 8-6 ページの「オブジェクトの属性の表示」の説明に従って、オブジェクトをダブルクリックしてその属性を表示します。

LIA ダイアグラムでのグループの使用

グループを使用すると、複雑なダイアグラムでオブジェクトを編成できるため、オブジェクトを簡単に探して編集できます。ダイアグラム内のオブジェクトの数を減らすと、対象オブジェクトをより簡単に見つけることができます。

グループを作成する手順

1. ボックスで囲んだ複数のオブジェクトをドラッグ・アンド・ドロップして、オブジェクトのグループを選択します。
2. 「選択したオブジェクトのグループ化」ツールをクリックします。
「選択したデータ・オブジェクトのグループ化」ダイアログ・ボックスが表示されます。
3. グループの名前を入力します。

選択したオブジェクトが単一のアイコンに縮小表示されます。

グループ内の個々のオブジェクトを表示するには、フォルダ・アイコンをダブルクリックします。これらのオブジェクトは、グループ化されていないオブジェクトと同じように操作できます。

オブジェクトのグループを解除するには、グループを選択して「選択したオブジェクトのグループ解除」ツールをクリックします。

オブジェクトの属性の表示

ダイアグラムのオブジェクト・アイコンを展開表示して、属性を検査できます。図 8-5 に、2つの展開表示されたアイコンを示します。アイコンの列属性はマッピングによって接続されています。

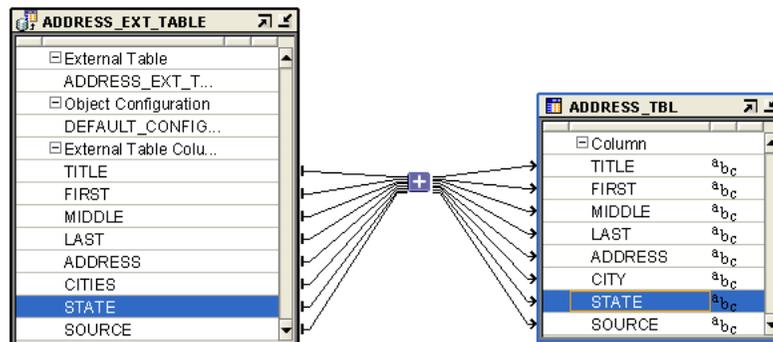
アイコンを展開表示するには、ダブルクリックします。アイコンに戻すには、右上角の下矢印をクリックします。

属性の LIA ダイアグラムを生成する手順

1. オブジェクトの LIA ダイアグラムを生成します。
2. アイコンをダブルクリックして属性を表示します。
3. 属性を右クリックして「系統の表示」または「影響の表示」を選択します。
 選択した属性の系統または影響パスにある属性が別の色で強調表示されます。

この詳細情報は、鑑査または変更の伝播を計画する際に使用できます。

図 8-5 LIA ダイアグラムで展開表示されたアイコン



依存性マネージャでのオブジェクトの変更

LIA ダイアグラムは、1つ以上のオブジェクトの変更によって無効化された可能性があるすべてのオブジェクトを識別します。これにより、影響があったオブジェクトを検査し、必要に応じて変更できます。

これらの変更は、依存性マネージャで次の方法を使用して行うことができます。

- オブジェクト・エディタ: オブジェクトに標準のエディタを開いて、拡張性の高い編集機能を提供します。
- 変更の伝播ツール: 名前、データ型、または、単一の列または属性の精度の変更など、基本的な編集機能を提供します。

オブジェクト・エディタを開く手順

ダイアグラムのオブジェクト・アイコンを右クリックし、ポップアップ・メニューから「**エディタを開く**」を選択します。

変更の伝播ツールを使用する手順

1. ダイアグラムのオブジェクト・アイコンをダブルクリックします。

アイコンが展開されて属性が表示されます。

2. 変更対象の属性を右クリックし、ポップアップ・メニューから「**変更の伝播**」を選択します。

「変更の伝播」ダイアログ・ボックスが表示されます。

3. 「新規の値」列の属性を変更し、「伝播」ボックスを選択します。
4. 「OK」をクリックします。

レポートおよび分析の SQL

この項では、ビジネス問合せから導出された効果的なビジネス・レポートの作成方法を説明します。この項の内容は次のとおりです。

- ビジネス問合せに応答する SQL アナリティック機能の使用
- パーティション外部結合を使用したスパースなデータの処理
- ビジネス問合せを単純化する WITH 句の使用

ビジネス問合せに回答する SQL アナリティック機能の使用

Oracle Database では、集計および分析 SQL 関数のファミリの導入によって SQL の分析処理機能を強化しています。これらの関数によって、ランキング、パーセンタイルおよび移動平均の計算など、次の問合せに回答できます。

- 上位 10 製品の国別売上は何ですか。
- 在庫のある製品の毎週の移動平均はどうですか。
- 総売上高の何パーセントが第 4 四半期に発生していますか。
- 第 4 四半期の平均割引は年間の平均割引よりどれくらい高いですか。
- 国内の精油所の 20 パーセントが閉鎖された場合、既存の精油所の収益性ランキングはどうなりますか。

集計関数は、合計の異なるタイプを導出でき、追加計算のこれらの合計を使用できるデータ・ウェアハウスの基礎となる部分です。データ・ウェアハウスの集計パフォーマンスを向上させるには、Oracle Database は GROUP BY 列に対するいくつかの拡張を提供します。CUBE、ROLLUP、GROUPING および GROUPING SETS 関数により、より早く簡単に問合せおよびレポート作成ができます。ROLLUP 関数は、最も個人的で詳細なものから総計まで、集計レベルを上げながら SUM、COUNT、MAX、MIN および AVG などの集計を計算します。CUBE 関数は ROLLUP に似た拡張で、単一の文で集計可能なすべての組合せを計算できます。

分析関数は行のグループに基づいて集計値を計算します。これらの関数は各グループに対して複数の行を戻すという点で集計関数とは異なります。この行のグループはウィンドウと呼ばれます。このウィンドウを使用すると、移動平均や累積合計などを計算できます。行ウィンドウは各行に対して定義されます。このウィンドウは現行の行の計算を実行するために使用する行の範囲を決定します。ウィンドウ・サイズは時間などの論理間隔、または行の物理的な行の数に基づくことができます。一部の関数はウィンドウとのみ使用され、ウィンドウ関数として参照されます。

パフォーマンスを向上させるために、集計および分析関数をパラレルで実行できます。つまり、複数のプロセスでこれらの関数すべてを同時に実行できます。これらの機能によって計算、分析およびレポート作成がより簡単で効果的になり、データ・ウェアハウスのパフォーマンス、スケーラビリティおよび簡易性が向上します。

高度な SQL および PL/SQL 機能を活用でき、Oracle Database はビジネス問合せを SQL に変更します。この項ではこれらの高度な機能を説明します。この項の内容は次のとおりです。

- [ROLLUP 関数を使用したレポートへの合計の追加方法](#)
- [CUBE 関数を使用した異なるレベルでの合計の分割方法](#)
- [GROUPING 関数を使用した小計の追加方法](#)
- [GROUPING SETS 関数を使用した集計の結合方法](#)
- [RANK 関数を使用したランキングの計算方法](#)
- [合計に対する相対的な寄与率の計算方法](#)
- [ウィンドウ関数を使用した行間計算の実行方法](#)
- [ウィンドウ関数を使用した移動平均の計算方法](#)

ROLLUP 関数を使用したレポートへの合計の追加方法

ROLLUP 関数によって、SELECT で総計やディメンションの特定のグループを横断する複数の小計レベルを計算できます。ROLLUP 関数は GROUP BY 句の単純な拡張であるため、この構文は非常に簡単に使用できます。ROLLUP 関数は非常に効果的で、問合せにかかるオーバーヘッドは最小限に抑えられます。ROLLUP 関数のアクションは簡単です。これは、最も詳細なレベルから総計まで、ROLLUP 関数で指定されたグループ・リストに従ってロールアップする小計を作成するためです。ROLLUP は、その引数として、グループ化列の順序付けリストをとりません。最初に、GROUP BY 句で指定された標準の集計値を計算します。次に、グループ化列のリストを右から左に移動しながら、順番に高いレベルの小計を作成します。最後に、総計を作成します。

ROLLUP 関数を使用するとき

タスクに小計が含まれる場合、特に小計が時間または地理などの階層ディメンションに基づいている場合、ROLLUP 関数は有効です。また、ROLLUP 関数によって、マテリアライズド・ビューのメンテナンスが簡素化および高速化できます。

例 : ROLLUP 関数の使用

ビジネス・レポートの準備時に共通して要求されるのは、収入額順で、異なる製品カテゴリにわたる四半期の総売上高の検索です。次の問合せはこれを達成し、後に実行するより複雑な問合せの構築の開始点に使用されます。

```
SELECT t.calendar_quarter_desc quarter
, p.prod_category category
, TO_CHAR(SUM(s.amount_sold), 'L999G999G990D00') revenue
FROM times t
, products p
, sales s
WHERE t.time_id = s.time_id
AND p.prod_id = s.prod_id
AND s.time_id BETWEEN TO_DATE('01-JAN-2001', 'dd-MON-yyyy')
AND TO_DATE('31-DEC-2001', 'dd-MON-yyyy')
GROUP BY t.calendar_quarter_desc, p.prod_category
ORDER BY t.calendar_quarter_desc
, SUM(s.amount_sold);
```

| QUARTER | CATEGORY | REVENUE |
|---------|-----------------------------|----------------|
| 2001-01 | Software/Other | \$860,819.81 |
| 2001-01 | Electronics | \$1,239,287.71 |
| 2001-01 | Hardware | \$1,301,343.45 |
| 2001-01 | Photo | \$1,370,706.38 |
| 2001-01 | Peripherals and Accessories | \$1,774,940.09 |
| 2001-02 | Software/Other | \$872,157.38 |
| 2001-02 | Electronics | \$1,144,187.90 |
| 2001-02 | Hardware | \$1,557,059.59 |
| 2001-02 | Photo | \$1,563,475.51 |
| 2001-02 | Peripherals and Accessories | \$1,785,588.01 |
| 2001-03 | Software/Other | \$877,630.85 |
| 2001-03 | Electronics | \$1,017,536.82 |
| 2001-03 | Photo | \$1,607,315.63 |
| 2001-03 | Hardware | \$1,651,454.29 |
| 2001-03 | Peripherals and Accessories | \$2,042,061.04 |
| 2001-04 | Software/Other | \$943,296.36 |
| 2001-04 | Hardware | \$1,174,512.68 |
| 2001-04 | Electronics | \$1,303,838.52 |
| 2001-04 | Photo | \$1,792,131.39 |
| 2001-04 | Peripherals and Accessories | \$2,257,118.57 |

この問合せは便利ですが、同じレポート上で異なるカテゴリの総計を確認する必要がある可能性があります。次の例は、元の問題に総計を追加するための ROLLUP 関数の使用方法を説明しています。

```
SELECT t.calendar_quarter_desc quarter
, p.prod_category category
, TO_CHAR(SUM(s.amount_sold), 'L999G999G990D00') revenue
FROM times t
, products p
, sales s
WHERE t.time_id = s.time_id
AND p.prod_id = s.prod_id
AND s.time_id BETWEEN TO_DATE('01-JAN-2001', 'dd-MON-yyyy')
AND TO_DATE('31-DEC-2001', 'dd-MON-yyyy')
GROUP BY ROLLUP(t.calendar_quarter_desc, p.prod_category)
ORDER BY t.calendar_quarter_desc
, SUM(s.amount_sold);
```

| QUARTER | CATEGORY | REVENUE |
|---------|-----------------------------|-----------------|
| 2001-01 | Software/Other | \$860,819.81 |
| 2001-01 | Electronics | \$1,239,287.71 |
| 2001-01 | Hardware | \$1,301,343.45 |
| 2001-01 | Photo | \$1,370,706.38 |
| 2001-01 | Peripherals and Accessories | \$1,774,940.09 |
| 2001-01 | | \$6,547,097.44 |
| 2001-02 | Software/Other | \$872,157.38 |
| 2001-02 | Electronics | \$1,144,187.90 |
| 2001-02 | Hardware | \$1,557,059.59 |
| 2001-02 | Photo | \$1,563,475.51 |
| 2001-02 | Peripherals and Accessories | \$1,785,588.01 |
| 2001-02 | | \$6,922,468.39 |
| 2001-03 | Software/Other | \$877,630.85 |
| 2001-03 | Electronics | \$1,017,536.82 |
| 2001-03 | Photo | \$1,607,315.63 |
| 2001-03 | Hardware | \$1,651,454.29 |
| 2001-03 | Peripherals and Accessories | \$2,042,061.04 |
| 2001-03 | | \$7,195,998.63 |
| 2001-04 | Software/Other | \$943,296.36 |
| 2001-04 | Hardware | \$1,174,512.68 |
| 2001-04 | Electronics | \$1,303,838.52 |
| 2001-04 | Photo | \$1,792,131.39 |
| 2001-04 | Peripherals and Accessories | \$2,257,118.57 |
| 2001-04 | | \$7,470,897.52 |
| | | \$28,136,461.98 |

CUBE 関数を使用した異なるレベルでの合計の分割方法

CUBE 関数は指定されたグループ化列の集合を取り、それらを取り得るすべての組合せに対して小計を作成します。多次元分析によって、CUBE 関数は指定されたディメンションを持つデータ・キューブに対して計算されるすべての小計を生成します。CUBE(time, region, department) を指定した場合、結果セットには同等の ROLLUP 関数および追加の組合せに含まれるすべての値が含まれます。

CUBE 関数を使用するとき

複数の表にわたるレポートを必要とする状況では、CUBE 関数の使用を検討してください。**複数の表にわたるレポート**に必要なデータは、CUBE 関数を使用する単一の SELECT 文で生成できます。ROLLUP と同様に、CUBE 関数もマテリアライズド・ビューの生成に便利です。CUBE 関数を含む問合せがパラレルで実行されるより、マテリアライズド・ビューを移入した方がより高速になることに注意してください。

例 : CUBE 関数の使用

四半期の総額のみでなく、選択した期間における異なる製品カテゴリの総額を入手する場合、次の例のように、CUBE 関数を使用してこれらの計算ができます。

CUBE 関数を使用する手順

```
SELECT t.calendar_quarter_desc quarter
, p.prod_category category
, TO_CHAR(SUM(s.amount_sold), 'L999G999G990D00') revenue
FROM times t
, products p
, sales s
WHERE t.time_id = s.time_id
AND p.prod_id = s.prod_id
AND s.time_id BETWEEN TO_DATE('01-JAN-2001', 'dd-MON-yyyy')
AND TO_DATE('31-DEC-2001', 'dd-MON-yyyy')
GROUP BY CUBE(t.calendar_quarter_desc, p.prod_category)
ORDER BY t.calendar_quarter_desc
, SUM(s.amount_sold);
```

| QUARTER | CATEGORY | REVENUE |
|---------|-----------------------------|-----------------|
| 2001-01 | Software/Other | \$860,819.81 |
| 2001-01 | Electronics | \$1,239,287.71 |
| 2001-01 | Hardware | \$1,301,343.45 |
| 2001-01 | Photo | \$1,370,706.38 |
| 2001-01 | Peripherals and Accessories | \$1,774,940.09 |
| 2001-01 | | \$6,547,097.44 |
| 2001-02 | Software/Other | \$872,157.38 |
| 2001-02 | Electronics | \$1,144,187.90 |
| 2001-02 | Hardware | \$1,557,059.59 |
| 2001-02 | Photo | \$1,563,475.51 |
| 2001-02 | Peripherals and Accessories | \$1,785,588.01 |
| 2001-02 | | \$6,922,468.39 |
| 2001-03 | Software/Other | \$877,630.85 |
| 2001-03 | Electronics | \$1,017,536.82 |
| 2001-03 | Photo | \$1,607,315.63 |
| 2001-03 | Hardware | \$1,651,454.29 |
| 2001-03 | Peripherals and Accessories | \$2,042,061.04 |
| 2001-03 | | \$7,195,998.63 |
| 2001-04 | Software/Other | \$943,296.36 |
| 2001-04 | Hardware | \$1,174,512.68 |
| 2001-04 | Electronics | \$1,303,838.52 |
| 2001-04 | Photo | \$1,792,131.39 |
| 2001-04 | Peripherals and Accessories | \$2,257,118.57 |
| 2001-04 | | \$7,470,897.52 |
| | Software/Other | \$3,553,904.40 |
| | Electronics | \$4,704,850.95 |
| | Hardware | \$5,684,370.01 |
| | Photo | \$6,333,628.91 |
| | Peripherals and Accessories | \$7,859,707.71 |
| | | \$28,136,461.98 |

GROUPING 関数を使用した小計の追加方法

ROLLUP および CUBE 関数を使用する際には、2つの課題があります。第一に、どの結果セット行が小計であるかをプログラム上でどのように判断し、指定された小計の正確な小計レベルをどのように探し出すかということです。合計に対する割合を計算する場合に小計を使用する必要があるため、どの行が求める小計であるかを判断する簡単な方法が必要です。第2に、格納される NULL 値および ROLLUP、または CUBE 関数によって作成される NULL 値の両方が問合せ結果に含まれる場合、どう処理するかという問題です。この2つをどのように区別するかが問題になります。

この問題は GROUPING 関数で処理します。単一の列を引数として使用し、ROLLUP または CUBE 関数により作成された NULL 値が発生した場合、GROUPING 関数は 1 を戻します。つまり、NULL 値が小計の行であることを示す場合、GROUPING は 1 を戻します。格納された NULL 値を含むその他のタイプの値では 0 を戻します。

GROUPING 関数を使用するとき

NULL 値または ROLLUP または CUBE 操作で作成された NULL 値を処理する必要がある場合、GROUPING 関数を使用します。NULL 値を使用する 1 つの理由に、NULL フィールドに説明を置くことです。たとえば、番号が総計を示すテキストなどです。

例 : GROUPING 関数の使用

値が総計を示している場合にわかりにくいため、レポート内の説明の列がさらに必要になる場合があります。GROUPING 関数を使用すると、次の例のように、問合せ結果に総計を示すラベルを挿入することができます。

GROUPING 関数を使用する手順

```
SELECT DECODE (GROUPING (t.calendar_quarter_desc)
              , 0, t.calendar_quarter_desc
              , 1, 'TOTAL'
              ) quarter
, DECODE (GROUPING (p.prod_category)
          , 0, p.prod_category
          , 1, 'TOTAL'
          ) category
, TO_CHAR (SUM (s.amount_sold), 'L999G999G990D00') revenue
FROM times t
, products p
, sales s
WHERE t.time_id = s.time_id
AND p.prod_id = s.prod_id
AND s.time_id BETWEEN TO_DATE ('01-JAN-2001', 'dd-MON-yyyy')
AND TO_DATE ('31-DEC-2001', 'dd-MON-yyyy')
GROUP BY CUBE (t.calendar_quarter_desc, p.prod_category)
ORDER BY t.calendar_quarter_desc
, SUM (s.amount_sold);
```

| QUARTER | CATEGORY | REVENUE |
|---------|-----------------------------|----------------|
| 2001-01 | Software/Other | \$860,819.81 |
| 2001-01 | Electronics | \$1,239,287.71 |
| 2001-01 | Hardware | \$1,301,343.45 |
| 2001-01 | Photo | \$1,370,706.38 |
| 2001-01 | Peripherals and Accessories | \$1,774,940.09 |
| 2001-01 | TOTAL | \$6,547,097.44 |
| 2001-02 | Software/Other | \$872,157.38 |
| 2001-02 | Electronics | \$1,144,187.90 |
| 2001-02 | Hardware | \$1,557,059.59 |
| 2001-02 | Photo | \$1,563,475.51 |
| 2001-02 | Peripherals and Accessories | \$1,785,588.01 |
| 2001-02 | TOTAL | \$6,922,468.39 |
| 2001-03 | Software/Other | \$877,630.85 |
| 2001-03 | Electronics | \$1,017,536.82 |
| 2001-03 | Photo | \$1,607,315.63 |
| 2001-03 | Hardware | \$1,651,454.29 |
| 2001-03 | Peripherals and Accessories | \$2,042,061.04 |
| 2001-03 | TOTAL | \$7,195,998.63 |
| 2001-04 | Software/Other | \$943,296.36 |
| 2001-04 | Hardware | \$1,174,512.68 |
| 2001-04 | Electronics | \$1,303,838.52 |
| 2001-04 | Photo | \$1,792,131.39 |

| | | |
|---------|-----------------------------|-----------------|
| 2001-04 | Peripherals and Accessories | \$2,257,118.57 |
| 2001-04 | TOTAL | \$7,470,897.52 |
| TOTAL | Software/Other | \$3,553,904.40 |
| TOTAL | Electronics | \$4,704,850.95 |
| TOTAL | Hardware | \$5,684,370.01 |
| TOTAL | Photo | \$6,333,628.91 |
| TOTAL | Peripherals and Accessories | \$7,859,707.71 |
| TOTAL | TOTAL | \$28,136,461.98 |

GROUPING SETS 関数を使用した集計の結合方法

GROUP BY 句内で GROUPING SETS 関数を使用して、作成するグループの集合を選択的に指定できます。これにより、CUBE の全データを計算せずに複数のディメンションにおよぼ正確な指定ができ、すべてのディメンションの総計は必要ありません。

GROUPING SETS 関数を使用するとき

データ・キューブ内の特定の小計が必要ですべての小計が入手できない場合、GROUPING SETS 関数を使用します。

例 : GROUPING SETS 関数の使用

販売チャンネルに基づいた営業番号の総計を参照する場合、チャンネル・クラスごとの総計を取得するため別の問合せを追加するかわりに、次の例のように、GROUPING SETS 関数を使用できます。

GROUPING SETS 関数を使用する手順

```
SELECT DECODE(GROUPING(t.calendar_quarter_desc)
              , 0, t.calendar_quarter_desc
              , 1, 'TOTAL'
              ) quarter
, DECODE(GROUPING(c.channel_class)
         , 0, c.channel_class
         , 1, '--all--'
         ) channel
, DECODE(GROUPING(p.prod_category)
         , 0, p.prod_category
         , 1, 'TOTAL'
         ) category
, TO_CHAR(SUM(s.amount_sold), 'L999G999G990D00') revenue
FROM times t
, products p
, channels c
, sales s
WHERE t.time_id = s.time_id
AND p.prod_id = s.prod_id
AND c.channel_id = s.channel_id
AND s.time_id BETWEEN TO_DATE('01-JAN-2001', 'dd-MON-yyyy')
AND TO_DATE('31-DEC-2001', 'dd-MON-yyyy')
GROUP BY GROUPING SETS(c.channel_class,
                       CUBE(t.calendar_quarter_desc, p.prod_category))
ORDER BY t.calendar_quarter_desc
, SUM(s.amount_sold);
```

| QUARTER | CHANNEL | CATEGORY | REVENUE |
|---------|---------|-----------------------------|----------------|
| 2001-01 | --all-- | Software/Other | \$860,819.81 |
| 2001-01 | --all-- | Electronics | \$1,239,287.71 |
| 2001-01 | --all-- | Hardware | \$1,301,343.45 |
| 2001-01 | --all-- | Photo | \$1,370,706.38 |
| 2001-01 | --all-- | Peripherals and Accessories | \$1,774,940.09 |
| 2001-01 | --all-- | TOTAL | \$6,547,097.44 |

| | | | |
|---------|----------|-----------------------------|-----------------|
| 2001-02 | --all-- | Software/Other | \$872,157.38 |
| 2001-02 | --all-- | Electronics | \$1,144,187.90 |
| 2001-02 | --all-- | Hardware | \$1,557,059.59 |
| 2001-02 | --all-- | Photo | \$1,563,475.51 |
| 2001-02 | --all-- | Peripherals and Accessories | \$1,785,588.01 |
| 2001-02 | --all-- | TOTAL | \$6,922,468.39 |
| 2001-03 | --all-- | Software/Other | \$877,630.85 |
| 2001-03 | --all-- | Electronics | \$1,017,536.82 |
| 2001-03 | --all-- | Photo | \$1,607,315.63 |
| 2001-03 | --all-- | Hardware | \$1,651,454.29 |
| 2001-03 | --all-- | Peripherals and Accessories | \$2,042,061.04 |
| 2001-03 | --all-- | TOTAL | \$7,195,998.63 |
| 2001-04 | --all-- | Software/Other | \$943,296.36 |
| 2001-04 | --all-- | Hardware | \$1,174,512.68 |
| 2001-04 | --all-- | Electronics | \$1,303,838.52 |
| 2001-04 | --all-- | Photo | \$1,792,131.39 |
| 2001-04 | --all-- | Peripherals and Accessories | \$2,257,118.57 |
| 2001-04 | --all-- | TOTAL | \$7,470,897.52 |
| TOTAL | --all-- | Software/Other | \$3,553,904.40 |
| TOTAL | --all-- | Electronics | \$4,704,850.95 |
| TOTAL | --all-- | Hardware | \$5,684,370.01 |
| TOTAL | --all-- | Photo | \$6,333,628.91 |
| TOTAL | Indirect | TOTAL | \$6,709,496.66 |
| TOTAL | --all-- | Peripherals and Accessories | \$7,859,707.71 |
| TOTAL | Others | TOTAL | \$8,038,529.96 |
| TOTAL | Direct | TOTAL | \$13,388,435.36 |
| TOTAL | --all-- | TOTAL | \$28,136,461.98 |

RANK 関数を使用したランキングの計算方法

ビジネス情報の処理には、複雑なランキング、小計、移動平均およびリード / ラグ比較を含む高度な計算が必要です。これらの集計および分析タスクはビジネス・インテリジェンス問合せに不可欠です。ウィンドウ関数の使用によって解決します。

RANK 関数を使用するとき

複雑な問合せを実行したり問合せ結果を分析する場合には、RANK 関数を使用します。

例 : RANK 関数の使用

四半期の収入番号のランクを示す追加の列を確認する場合、次の例のように、RANK 関数を使用します。

RANK 関数を使用する手順

```
SELECT DECODE(GROUPING(t.calendar_quarter_desc)
              , 0, t.calendar_quarter_desc
              , 1, 'TOTAL'
              ) quarter
, DECODE(GROUPING(t.calendar_quarter_desc) + GROUPING(p.prod_category)
        , 0, RANK() OVER (PARTITION BY t.calendar_quarter_desc
                          ORDER BY SUM(s.amount_sold))
        , 1, null
        ) ranking
, DECODE(GROUPING(c.channel_class)
        , 0, c.channel_class
        , 1, '--all--'
        ) channel
, DECODE(GROUPING(p.prod_category)
        , 0, p.prod_category
        , 1, 'TOTAL'
        ) category
, TO_CHAR(SUM(s.amount_sold), 'L999G999G990D00') revenue
```

```

FROM times t
, products p
, channels c
, sales s
WHERE t.time_id = s.time_id
AND p.prod_id = s.prod_id
AND c.channel_id = s.channel_id
AND s.time_id BETWEEN TO_DATE('01-JAN-2001','dd-MON-yyyy')
AND TO_DATE('31-DEC-2001','dd-MON-yyyy')
GROUP BY GROUPING SETS(c.channel_class,
CUBE(t.calendar_quarter_desc, p.prod_category))
ORDER BY t.calendar_quarter_desc
, SUM(s.amount_sold);

```

| QUARTER | RANKING | CHANNEL | CATEGORY | REVENUE |
|---------|---------|----------|-----------------------------|-----------------|
| 2001-01 | 1 | --all-- | Software/Other | \$860,819.81 |
| 2001-01 | 2 | --all-- | Electronics | \$1,239,287.71 |
| 2001-01 | 3 | --all-- | Hardware | \$1,301,343.45 |
| 2001-01 | 4 | --all-- | Photo | \$1,370,706.38 |
| 2001-01 | 5 | --all-- | Peripherals and Accessories | \$1,774,940.09 |
| 2001-01 | | --all-- | TOTAL | \$6,547,097.44 |
| 2001-02 | 1 | --all-- | Software/Other | \$872,157.38 |
| 2001-02 | 2 | --all-- | Electronics | \$1,144,187.90 |
| 2001-02 | 3 | --all-- | Hardware | \$1,557,059.59 |
| 2001-02 | 4 | --all-- | Photo | \$1,563,475.51 |
| 2001-02 | 5 | --all-- | Peripherals and Accessories | \$1,785,588.01 |
| 2001-02 | | --all-- | TOTAL | \$6,922,468.39 |
| 2001-03 | 1 | --all-- | Software/Other | \$877,630.85 |
| 2001-03 | 2 | --all-- | Electronics | \$1,017,536.82 |
| 2001-03 | 3 | --all-- | Photo | \$1,607,315.63 |
| 2001-03 | 4 | --all-- | Hardware | \$1,651,454.29 |
| 2001-03 | 5 | --all-- | Peripherals and Accessories | \$2,042,061.04 |
| 2001-03 | | --all-- | TOTAL | \$7,195,998.63 |
| 2001-04 | 1 | --all-- | Software/Other | \$943,296.36 |
| 2001-04 | 2 | --all-- | Hardware | \$1,174,512.68 |
| 2001-04 | 3 | --all-- | Electronics | \$1,303,838.52 |
| 2001-04 | 4 | --all-- | Photo | \$1,792,131.39 |
| 2001-04 | 5 | --all-- | Peripherals and Accessories | \$2,257,118.57 |
| 2001-04 | | --all-- | TOTAL | \$7,470,897.52 |
| TOTAL | | --all-- | Software/Other | \$3,553,904.40 |
| TOTAL | | --all-- | Electronics | \$4,704,850.95 |
| TOTAL | | --all-- | Hardware | \$5,684,370.01 |
| TOTAL | | --all-- | Photo | \$6,333,628.91 |
| TOTAL | | Indirect | TOTAL | \$6,709,496.66 |
| TOTAL | | --all-- | Peripherals and Accessories | \$7,859,707.71 |
| TOTAL | | Others | TOTAL | \$8,038,529.96 |
| TOTAL | | Direct | TOTAL | \$13,388,435.36 |
| TOTAL | | --all-- | TOTAL | \$28,136,461.98 |

In this example, the PARTITION BY clause defines the boundaries for the RANK function.

合計に対する相対的な寄与率の計算方法

共通のビジネス・インテリジェンス要求は、特定の時間間隔に基づく総収入に対する各製品カテゴリの寄与率を計算します。

例：総計に対する相対的な寄与率の計算

四半期ベースで収入番号の差異を取得します。次の例のように、PARTITION BY 製品カテゴリを持つウィンドウ関数を使用して実行できます。

総計に対する相対的な寄与率を計算する手順

```
SELECT DECODE(GROUPING(t.calendar_quarter_desc)
              , 0, t.calendar_quarter_desc
              , 1, 'TOTAL'
              ) quarter
, DECODE(GROUPING(t.calendar_quarter_desc) + GROUPING(p.prod_category)
        , 0, RANK() OVER (PARTITION BY t.calendar_quarter_desc
                          ORDER BY SUM(s.amount_sold))
        , 1, null
        ) RANKING
, DECODE(GROUPING(c.channel_class)
        , 0, c.channel_class
        , 1, '--all--'
        ) channel
, DECODE(GROUPING(p.prod_category)
        , 0, p.prod_category
        , 1, 'TOTAL'
        ) category
, TO_CHAR(SUM(s.amount_sold), 'L999G999G990D00') revenue
, TO_CHAR(100 * RATIO_TO_REPORT(SUM(s.amount_sold))
OVER (PARTITION BY (TO_CHAR(GROUPING(p.prod_category) ||
t.calendar_quarter_desc)), '990D0') percent
FROM times t
, products p
, channels c
, sales s
WHERE t.time_id = s.time_id
AND p.prod_id = s.prod_id
AND c.channel_id = s.channel_id
AND s.time_id BETWEEN TO_DATE('01-JAN-2001', 'dd-MON-yyyy')
AND TO_DATE('31-DEC-2001', 'dd-MON-yyyy')
GROUP BY GROUPING SETS(c.channel_class,
CUBE(t.calendar_quarter_desc, p.prod_category))
ORDER BY t.calendar_quarter_desc
, SUM(s.amount_sold);
```

| QUARTER | RANKING | CHANNEL | CATEGORY | REVENUE | PERC |
|---------|---------|---------|----------------|----------------|-------|
| 2001-01 | 1 | --all-- | Software/Other | \$860,819.81 | 13.1 |
| 2001-01 | 2 | --all-- | Electronics | \$1,239,287.71 | 18.9 |
| 2001-01 | 3 | --all-- | Hardware | \$1,301,343.45 | 19.9 |
| 2001-01 | 4 | --all-- | Photo | \$1,370,706.38 | 20.9 |
| 2001-01 | 5 | --all-- | Peripherals | \$1,774,940.09 | 27.1 |
| 2001-01 | | --all-- | TOTAL | \$6,547,097.44 | 100.0 |
| 2001-02 | 1 | --all-- | Software/Other | \$872,157.38 | 12.6 |
| 2001-02 | 2 | --all-- | Electronics | \$1,144,187.90 | 16.5 |
| 2001-02 | 3 | --all-- | Hardware | \$1,557,059.59 | 22.5 |
| 2001-02 | 4 | --all-- | Photo | \$1,563,475.51 | 22.6 |
| 2001-02 | 5 | --all-- | Peripherals | \$1,785,588.01 | 25.8 |
| 2001-02 | | --all-- | TOTAL | \$6,922,468.39 | 100.0 |
| 2001-03 | 1 | --all-- | Software/Other | \$877,630.85 | 12.2 |
| 2001-03 | 2 | --all-- | Electronics | \$1,017,536.82 | 14.1 |
| 2001-03 | 3 | --all-- | Photo | \$1,607,315.63 | 22.3 |
| 2001-03 | 4 | --all-- | Hardware | \$1,651,454.29 | 22.9 |
| 2001-03 | 5 | --all-- | Peripherals | \$2,042,061.04 | 28.4 |
| 2001-03 | | --all-- | TOTAL | \$7,195,998.63 | 100.0 |
| 2001-04 | 1 | --all-- | Software/Other | \$943,296.36 | 12.6 |
| 2001-04 | 2 | --all-- | Hardware | \$1,174,512.68 | 15.7 |
| 2001-04 | 3 | --all-- | Electronics | \$1,303,838.52 | 17.5 |
| 2001-04 | 4 | --all-- | Photo | \$1,792,131.39 | 24.0 |
| 2001-04 | 5 | --all-- | Peripherals | \$2,257,118.57 | 30.2 |
| 2001-04 | | --all-- | TOTAL | \$7,470,897.52 | 100.0 |
| TOTAL | | --all-- | Software/Other | \$3,553,904.40 | 12.6 |

| | | | | |
|-------|----------|-------------|-----------------|------|
| TOTAL | --all-- | Electronics | \$4,704,850.95 | 16.7 |
| TOTAL | --all-- | Hardware | \$5,684,370.01 | 20.2 |
| TOTAL | --all-- | Photo | \$6,333,628.91 | 22.5 |
| TOTAL | Indirect | TOTAL | \$6,709,496.66 | 11.9 |
| TOTAL | --all-- | Peripherals | \$7,859,707.71 | 27.9 |
| TOTAL | Others | TOTAL | \$8,038,529.96 | 14.3 |
| TOTAL | Direct | TOTAL | \$13,388,435.36 | 23.8 |
| TOTAL | --all-- | TOTAL | \$28,136,461.98 | 50.0 |

"Peripherals" was used instead of "Peripherals and Accessories" to save space.

ウィンドウ関数を使用した行間計算の実行方法

ビジネス・インテリジェンスに共通する問題は、特定の結果がどのように他の結果と関連しているかということです。単一問合せでこれを行うには、ウィンドウ関数を使用して単一の文で行間計算を実行できます。

例：行間計算の実行

各四半期の総収入に対するそれぞれの製品カテゴリの寄与率を調べる場合、次の例のように、ウィンドウ関数 `RATIO_TO_REPORT` を使用できます。四半期ごとの `RATIO_TO_REPORT` から総計を除外するには、`GROUPING(p.prod_category)` を持つ連結を使用する必要がありますことに注意してください。

行間計算を行う手順

```
SELECT DECODE(GROUPING(t.calendar_quarter_desc)
              , 0, t.calendar_quarter_desc
              , 1, 'TOTAL'
              ) quarter
, DECODE(GROUPING(t.calendar_quarter_desc) + GROUPING(p.prod_category)
        , 0, RANK() OVER (PARTITION BY t.calendar_quarter_desc
                        ORDER BY SUM(s.amount_sold)
                        , 1, null
                        ) RANKING
        , DECODE(GROUPING(c.channel_class)
                  , 0, c.channel_class
                  , 1, '--all--'
                  ) channel
        , DECODE(GROUPING(p.prod_category)
                  , 0, p.prod_category
                  , 1, 'TOTAL'
                  ) category
        , TO_CHAR(SUM(s.amount_sold), 'L999G999G990D00') revenue
        , TO_CHAR(100 * RATIO_TO_REPORT(SUM(s.amount_sold)
OVER (PARTITION BY (TO_CHAR(GROUPING(p.prod_category) ||
t.calendar_quarter_desc)), '990D0') percent
        , DECODE(GROUPING(t.calendar_quarter_desc) + GROUPING(p.prod_category)
                  , 0, TO_CHAR(SUM(s.amount_sold) - LAG(SUM(s.amount_sold), 1)
OVER (PARTITION BY p.prod_category
      ORDER BY t.calendar_quarter_desc), 'L999G990D00')
                  , 1, null
                  ) q_q_diff
FROM times t
, products p
, channels c
, sales s
WHERE t.time_id = s.time_id
AND p.prod_id = s.prod_id
AND c.channel_id = s.channel_id
AND s.time_id BETWEEN TO_DATE('01-JAN-2001', 'dd-MON-yyyy')
AND TO_DATE('31-DEC-2001', 'dd-MON-yyyy')
GROUP BY GROUPING SETS(c.channel_class,
```

```

CUBE(t.calendar_quarter_desc, p.prod_category))
ORDER BY t.calendar_quarter_desc
, SUM(s.amount_sold);

```

| QUARTER | RANKING | CHANNEL | CATEGORY | REVENUE | PERC | Q_Q_DIFF |
|---------|---------|----------|----------------|-----------------|-------|---------------|
| 2001-01 | 1 | --all-- | Software/Other | \$860,819.81 | 13.1 | |
| 2001-01 | 2 | --all-- | Electronics | \$1,239,287.71 | 18.9 | |
| 2001-01 | 3 | --all-- | Hardware | \$1,301,343.45 | 19.9 | |
| 2001-01 | 4 | --all-- | Photo | \$1,370,706.38 | 20.9 | |
| 2001-01 | 5 | --all-- | Peripherals | \$1,774,940.09 | 27.1 | |
| 2001-01 | | --all-- | TOTAL | \$6,547,097.44 | 100.0 | |
| 2001-02 | 1 | --all-- | Software/Other | \$872,157.38 | 12.6 | \$11,337.57 |
| 2001-02 | 2 | --all-- | Electronics | \$1,144,187.90 | 16.5 | -\$95,099.81 |
| 2001-02 | 3 | --all-- | Hardware | \$1,557,059.59 | 22.5 | \$255,716.14 |
| 2001-02 | 4 | --all-- | Photo | \$1,563,475.51 | 22.6 | \$192,769.13 |
| 2001-02 | 5 | --all-- | Peripherals | \$1,785,588.01 | 25.8 | \$10,647.92 |
| 2001-02 | | --all-- | TOTAL | \$6,922,468.39 | 100.0 | |
| 2001-03 | 1 | --all-- | Software/Other | \$877,630.85 | 12.2 | \$5,473.47 |
| 2001-03 | 2 | --all-- | Electronics | \$1,017,536.82 | 14.1 | -\$126,651.08 |
| 2001-03 | 3 | --all-- | Photo | \$1,607,315.63 | 22.3 | \$43,840.12 |
| 2001-03 | 4 | --all-- | Hardware | \$1,651,454.29 | 22.9 | \$94,394.70 |
| 2001-03 | 5 | --all-- | Peripherals | \$2,042,061.04 | 28.4 | \$256,473.03 |
| 2001-03 | | --all-- | TOTAL | \$7,195,998.63 | 100.0 | |
| 2001-04 | 1 | --all-- | Software/Other | \$943,296.36 | 12.6 | \$65,665.51 |
| 2001-04 | 2 | --all-- | Hardware | \$1,174,512.68 | 15.7 | -\$476,941.61 |
| 2001-04 | 3 | --all-- | Electronics | \$1,303,838.52 | 17.5 | \$286,301.70 |
| 2001-04 | 4 | --all-- | Photo | \$1,792,131.39 | 24.0 | \$184,815.76 |
| 2001-04 | 5 | --all-- | Peripherals | \$2,257,118.57 | 30.2 | \$215,057.53 |
| 2001-04 | | --all-- | TOTAL | \$7,470,897.52 | 100.0 | |
| TOTAL | | --all-- | Software/Other | \$3,553,904.40 | 12.6 | |
| TOTAL | | --all-- | Electronics | \$4,704,850.95 | 16.7 | |
| TOTAL | | --all-- | Hardware | \$5,684,370.01 | 20.2 | |
| TOTAL | | --all-- | Photo | \$6,333,628.91 | 22.5 | |
| TOTAL | | Indirect | TOTAL | \$6,709,496.66 | 11.9 | |
| TOTAL | | --all-- | Peripherals | \$7,859,707.71 | 27.9 | |
| TOTAL | | Others | TOTAL | \$8,038,529.96 | 14.3 | |
| TOTAL | | Direct | TOTAL | \$13,388,435.36 | 23.8 | |
| TOTAL | | --all-- | TOTAL | \$28,136,461.98 | 50.0 | |

"Peripherals" was used instead of "Peripherals and Accessories" to save space.

ウィンドウ関数を使用した移動平均の計算方法

ウィンドウ関数を使用して移動集計を作成できます。移動関数は物理行の数に基づくか、論理時間間隔である可能性があります。ウィンドウ関数は PARTITION キーワードを使用し、パーティションの各行に対してデータのウィンドウの変動を定義できます。このウィンドウでは現行の行の計算を実行するために使用する行の範囲を決定します。ウィンドウ・サイズは行の物理的な数または時間などの論理間隔に基づきます。ウィンドウには開始行および終了行が含まれます。この定義によっては、ウィンドウは片方の行または両方の行へ移動する場合があります。たとえば、累積 SUM 関数を定義したウィンドウには、パーティションの最初の行に固定された開始行が含まれ、終了行は開始点からパーティションの最後の行までスライドします。反対に、移動平均を定義したウィンドウには開始点と終了点の変動が含まれるため、定数の物理的範囲または論理的範囲が維持されます。

ウィンドウ関数は一般的に、移動および累積バージョンの SUM、AVERAGE、COUNT、MAX、MIN、および他の多くの関数の計算に使用されます。また、問合せの SELECT 句、ORDER BY 句でのみ使用できます。ウィンドウ関数には、ウィンドウで最初の値を戻す FIRST_VALUE 関数、およびウィンドウで最終値を戻す LAST_VALUE 関数が含まれます。これらの関数によって自己結合なしに表内の複数行にアクセスできます。

例：移動平均の計算

次の例は、論理的な時間間隔を使用した、各製品の製品収入の7日間の変動平均を取得する問合せを示しています。

変動平均を計算する手順は、次のとおりです。

```
SELECT time_id
, prod_name
, TO_CHAR(revenue, 'L999G990D00') revenue
, TO_CHAR(AVG(revenue) OVER (PARTITION BY prod_name ORDER BY time_id
RANGE INTERVAL '7' DAY PRECEDING), 'L999G990D00') mv_7day_avg
FROM
( SELECT s.time_id, p.prod_name, SUM(s.amount_sold) revenue
FROM products p
, sales s
WHERE p.prod_id = s.prod_id
AND s.time_id BETWEEN TO_DATE('25-JUN-2001', 'dd-MON-yyyy')
AND TO_DATE('16-JUL-2001', 'dd-MON-yyyy')
AND p.prod_name LIKE '%Memory%'
AND p.prod_category = 'Photo'
GROUP BY s.time_id, p.prod_name
)
ORDER BY time_id, prod_name;
```

| TIME_ID | PROD_NAME | REVENUE | MV_7DAY_AVG |
|-----------|-------------------|------------|-------------|
| 26-JUN-01 | 256MB Memory Card | \$560.15 | \$560.15 |
| 30-JUN-01 | 256MB Memory Card | \$844.00 | \$702.08 |
| 02-JUL-01 | 128MB Memory Card | \$3,283.74 | \$3,283.74 |
| 02-JUL-01 | 256MB Memory Card | \$3,903.32 | \$1,769.16 |
| 03-JUL-01 | 256MB Memory Card | \$699.37 | \$1,501.71 |
| 08-JUL-01 | 128MB Memory Card | \$3,283.74 | \$3,283.74 |
| 08-JUL-01 | 256MB Memory Card | \$3,903.32 | \$2,835.34 |
| 10-JUL-01 | 256MB Memory Card | \$138.82 | \$1,580.50 |

パーティション外部結合を使用したスパースなデータの処理

データは通常、スパースな形式で格納されています。つまり、ディメンション値の特定の組合せで値が存在しない場合、**ファクト表**（売上などの重要なファクトを含むデータ・ウェアハウスの表）には行が存在していません。しかし、ファクト・データが存在しない場合でも、ディメンション値のすべての組合せの行を表示し、データを**稠密な**形式で表示する必要がある場合があります。たとえば、製品が特定の期間販売されていない場合でも、その期間の売上値をゼロとして製品の横に表示する場合などです。さらに、データが時間ディメンションに沿って稠密であれば、時系列の計算を簡単に実行できます。これは、稠密なデータが期間ごとに一定数の行を占めているため、物理オフセットを指定した分析ウィンドウ関数の使用が単純化されるためです。

データの**稠密化**は、スパース・データを稠密な形式に変換するプロセスです。スパース性の問題を解決するために、パーティション外部結合を使用して時系列または他のディメンションとのギャップを埋めることができます。この結合は問合せで定義した各論理パーティションに外部結合を適用することで、従来の外部結合の構文を拡張したものです。Oracle Database は、PARTITION BY 句で指定した式に基づいて、問合せの行を論理的にパーティション化します。パーティション外部結合の結果は、結合のもう一方の表を含む論理的にパーティション化された表にある各パーティションの外部結合の UNION 操作です。時間ディメンションのみではなく他のディメンションのギャップを埋めるためにも、このタイプの結合を使用してください。

パーティション外部結合を使用するとき

結果セットの欠落した行を埋める場合または時系列計算を処理する場合、パーティション外部結合を使用します。

例：パーティション外部結合の使用

特定の製品の数週間の売上を表示する必要がある場合があります。この例では、フォト・カテゴリのメモリー・カードを使用します。これらの製品は頻繁に売れるものではなく、数週間売れないこともあるためです。簡単に比較するため、次の例のように、パーティション外部結合を使用してデータを稠密にする必要があります。

パーティション外部結合を使用する手順は、次のとおりです。

```
SELECT tim.week_ending_day
, rev.prod_name product
, nvl(SUM(rev.amount_sold),0) revenue
FROM (SELECT p.prod_name, s.time_id, s.amount_sold
      FROM products p
      , sales s
      WHERE s.prod_id = p.prod_id
            AND p.prod_category = 'Photo'
            AND p.prod_name LIKE '%Memory%'
            AND s.time_id BETWEEN TO_DATE('25-JUN-2001','dd-MON-yyyy')
                               AND TO_DATE('16-JUL-2001','dd-MON-yyyy')
      ) rev
PARTITION BY (prod_name)
RIGHT OUTER JOIN (SELECT time_id, week_ending_day FROM times
                 WHERE week_ending_day
                       BETWEEN TO_DATE('01-JUL-2001','dd-MON-yyyy')
                       AND TO_DATE('16-JUL-2001','dd-MON-yyyy')
                 ) tim
ON (rev.time_id = tim.time_id)
GROUP BY tim.week_ending_day
, rev.prod_name
ORDER BY tim.week_ending_day
, rev.prod_name;
```

| WEEK_ENDI | PRODUCT | REVENUE |
|-----------|-------------------|---------|
| 01-JUL-01 | 128MB Memory Card | 0 |
| 01-JUL-01 | 256MB Memory Card | 1404.15 |
| 08-JUL-01 | 128MB Memory Card | 6567.48 |
| 08-JUL-01 | 256MB Memory Card | 8506.01 |
| 15-JUL-01 | 128MB Memory Card | 0 |
| 15-JUL-01 | 256MB Memory Card | 138.82 |

ビジネス問合せを単純化する WITH 句の使用

異なるタイプの結合およびウィンドウ関数を活用し、多くの表にアクセスする問合せは、非常に複雑になります。WITH 句を使用すると、問合せの増分的な構築によりこの複雑さを解消できます。複雑な問合せで問合せブロックが複数回発生する場合、SELECT 文の同じ問合せブロックを再使用できます。Oracle Database は問合せブロックの結果を取得して、ユーザーの一時表領域に結果を格納します。

WITH 句を使用するとき

同じ問合せブロックに対して複数の参照が問合せに含まれ、結合および集計がある場合、WITH 句を使用します。

例 : WITH 句の使用

2001年7月締め最初の3週間のフォト・カテゴリのメモリー・カード製品の売上を比較すると想定します。次の問合せはこの期間に売れなかった一部の製品を考慮し、前の週と関連する売上が増加または減少を戻します。最終的に、問合せは特定の週におけるメモリー・カードの売上に対する寄与率の割合を取得します。WITH 句の使用のために、問合せの個々のセクションの複雑さは解消されます。

WITH 句を使用する手順は、次のとおりです。

```
WITH sales_numbers AS
( SELECT s.prod_id, s.amount_sold, t.week_ending_day
  FROM sales s
    , times t
    , products p
  WHERE s.time_id = t.time_id
  AND s.prod_id = p.prod_id
  AND p.prod_category = 'Photo'
  AND p.prod_name LIKE '%Memory%'
  AND t.week_ending_day BETWEEN TO_DATE('01-JUL-2001','dd-MON-yyyy')
                           AND TO_DATE('16-JUL-2001','dd-MON-yyyy')
)
, product_revenue AS
( SELECT p.prod_name product, s.week_ending_day, SUM(s.amount_sold) revenue
  FROM products p
    LEFT OUTER JOIN (SELECT prod_id, amount_sold, week_ending_day
                     FROM sales_numbers) s
    ON (s.prod_id = p.prod_id)
  WHERE p.prod_category = 'Photo'
  AND p.prod_name LIKE '%Memory%'
  GROUP BY p.prod_name, s.week_ending_day
)
, weeks AS
( SELECT distinct week_ending_day week FROM times WHERE week_ending_day
  BETWEEN TO_DATE('01-JUL-2001','dd-MON-yyyy')
  AND TO_DATE('16-JUL-2001','dd-MON-yyyy')
)
, complete_product_revenue AS
( SELECT w.week, pr.product, nvl(pr.revenue,0) revenue
  FROM product_revenue pr
    PARTITION BY (product)
  RIGHT OUTER JOIN weeks w
    ON (w.week = pr.week_ending_day)
)
SELECT week
, product
, TO_CHAR(revenue, 'L999G990D00') revenue
, TO_CHAR(revenue - lag(revenue,1) OVER (PARTITION BY product
  ORDER BY week), 'L999G990D00') w_w_diff
, TO_CHAR(100 * RATIO_TO_REPORT(revenue) OVER (PARTITION BY week), '990D0') percentage
FROM complete_product_revenue
ORDER BY week, product;
```

| WEEK | PRODUCT | REVENUE | W_W_DIFF | PERCENT |
|-----------|-------------------|------------|-------------|---------|
| 01-JUL-01 | 128MB Memory Card | \$0.00 | | 0.0 |
| 01-JUL-01 | 256MB Memory Card | \$1,404.15 | | 100.0 |
| 01-JUL-01 | 64MB Memory Card | \$0.00 | | 0.0 |
| 08-JUL-01 | 128MB Memory Card | \$6,567.48 | \$6,567.48 | 43.6 |
| 08-JUL-01 | 256MB Memory Card | \$8,506.01 | \$7,101.86 | 56.4 |
| 08-JUL-01 | 64MB Memory Card | \$0.00 | \$0.00 | 0.0 |
| 15-JUL-01 | 128MB Memory Card | \$0.00 | -\$6,567.48 | 0.0 |
| 15-JUL-01 | 256MB Memory Card | \$138.82 | -\$8,367.19 | 100.0 |
| 15-JUL-01 | 64MB Memory Card | \$0.00 | \$0.00 | 0.0 |

第 IV 部

データ・ウェアハウスの管理

ここでは、データ・ウェアハウスのメンテナンス方法について説明します。内容は次のとおりです。

- [第 10 章「データ・ウェアハウスのリフレッシュ」](#)
- [第 11 章「データ・ウェアハウスの操作の最適化」](#)
- [第 12 章「パフォーマンス・ボトルネックの排除」](#)
- [第 13 章「データ・ウェアハウスのバックアップおよびリカバリ」](#)
- [第 14 章「データ・ウェアハウスの操作の最適化」](#)

データ・ウェアハウスのリフレッシュ

この章の内容は次のとおりです。

- データ・ウェアハウスのリフレッシュについて
- ローリング・ウィンドウを使用したデータのオフロード

データ・ウェアハウスのリフレッシュについて

データ・ウェアハウスを定期的に更新し、現在の情報を導出します。データの更新プロセスは、**リフレッシュ処理**と呼ばれます。

抽出、変換およびロード (ETL) は、スケジュールに基づいて実行され、元のソース・システムへの変更が反映されます。この手順の間、新規の更新されたデータを本番データのウェアハウス・スキーマに物理的に挿入し、ユーザーがこのデータを使用できるようにするために、他のすべての必要な手順 (索引の作成、制限の検証、バックアップ・コピーの作成) に従います。このデータをすべてデータ・ウェアハウスにロードすると、マテリアライズド・ビューを更新して最新のデータを反映する必要があります。

データ・ウェアハウスのパーティショニング・スキームは、データ・ウェアハウスのロード・プロセスでの操作のリフレッシュの効率性を決定する際に重要です。ロード・プロセスは、データ・ウェアハウス表のパーティショニング・スキームの選択時に考慮されます。

多くのデータ・ウェアハウスでは、定期的な新規データがロードされます。たとえば、日、週末または月ごとに新規データがデータ・ウェアハウスに取り込まれます。週末または月末にロードされたデータは通常、その週末はその月のトランザクションに対応します。この一般的なシナリオでは、データ・ウェアハウスは時間ごとにロードされます。ここで、データ・ウェアハウス表をデータ列上でパーティション化することをお勧めします。データ・ウェアハウスの例として、新規データが毎月、売上表にロードされると想定します。さらに、sales 表は毎月、パーティション化されるものとします。これらの手順では、ロード・プロセスによる新しい月 (2006 年第 1 四半期) の sales 表へのデータの追加処理方法を示します。

例: データ・ウェアハウスのリフレッシュ

多くの問合せでは、日付で問合せが制限されている products、customers および sales 表の一部の列が要求されます。マテリアライズド・ビューを利用すると、3つの表に対する問合せの大部分が高速化されます。マテリアライズド・ビューが作成される事前作成表を使用して、sales 表のパーティション計画を持つ動機のマテリアライズド・ビューのパーティション計画を選択します。

次に、マテリアライズド・ビューのリフレッシュに関する例を示します。ここでは、パーティション交換ロード操作を使用します。この例は sh スキーマの sales 表に基づいています。

マテリアライズド・ビューをリフレッシュする手順

1. マテリアライズド・ビューに基づいた表を作成します。

```
CREATE TABLE sales_prod_cust_mv
( time_id DATE
, prod_id NUMBER
, prod_name VARCHAR2(50)
, cust_id NUMBER
, cust_first_name VARCHAR2(20)
, cust_last_name VARCHAR2(40)
, amount_sold NUMBER
, quantity_sold NUMBER
)
PARTITION BY RANGE (time_id)
( PARTITION p1999 VALUES LESS THAN (TO_DATE('01-JAN-2000','DD-MON-YYYY'))
, PARTITION p2000 VALUES LESS THAN (TO_DATE('01-JAN-2001','DD-MON-YYYY'))
, PARTITION p2001h1 VALUES LESS THAN (TO_DATE('01-JUL-2001','DD-MON-YYYY'))
, PARTITION p2001h2 VALUES LESS THAN (TO_DATE('01-JAN-2002','DD-MON-YYYY'))
, PARTITION p2001q1 VALUES LESS THAN (TO_DATE('01-APR-2002','DD-MON-YYYY'))
, PARTITION p2002q2 VALUES LESS THAN (TO_DATE('01-JUL-2002','DD-MON-YYYY'))
, PARTITION p2002q3 VALUES LESS THAN (TO_DATE('01-OCT-2002','DD-MON-YYYY'))
, PARTITION p2002q4 VALUES LESS THAN (TO_DATE('01-JAN-2003','DD-MON-YYYY'))
, PARTITION p2003q1 VALUES LESS THAN (TO_DATE('01-APR-2003','DD-MON-YYYY'))
, PARTITION p2003q2 VALUES LESS THAN (TO_DATE('01-JUL-2003','DD-MON-YYYY'))
, PARTITION p2003q3 VALUES LESS THAN (TO_DATE('01-OCT-2003','DD-MON-YYYY'))
```

```

, PARTITION p2003q4 VALUES LESS THAN (TO_DATE('01-JAN-2004','DD-MON-YYYY'))
, PARTITION p2004q1 VALUES LESS THAN (TO_DATE('01-APR-2004','DD-MON-YYYY'))
, PARTITION p2004q2 VALUES LESS THAN (TO_DATE('01-JUL-2004','DD-MON-YYYY'))
, PARTITION p2004q3 VALUES LESS THAN (TO_DATE('01-OCT-2004','DD-MON-YYYY'))
, PARTITION p2004q4 VALUES LESS THAN (TO_DATE('01-JAN-2005','DD-MON-YYYY'))
, PARTITION p2005q1 VALUES LESS THAN (TO_DATE('01-APR-2005','DD-MON-YYYY'))
, PARTITION p2005q2 VALUES LESS THAN (TO_DATE('01-JUL-2005','DD-MON-YYYY'))
, PARTITION p2005q3 VALUES LESS THAN (TO_DATE('01-OCT-2005','DD-MON-YYYY'))
, PARTITION p2005q4 VALUES LESS THAN (TO_DATE('01-JAN-2006','DD-MON-YYYY'))
, PARTITION p2006q1 VALUES LESS THAN (TO_DATE('01-APR-2006','DD-MON-YYYY'))
) PARALLEL COMPRESS;

```

2. sales 表の初期表をロードします。

```

ALTER SESSION ENABLE PARALLEL DML;
INSERT /*+ PARALLEL smv */ INTO sales_prod_cust_mv smv
SELECT /*+ PARALLEL s PARALLEL c */ s.time_id
, s.prod_id
, p.prod_name
, s.cust_id
, cust_first_name
, c.cust_last_name
, SUM(s.amount_sold)
, SUM(s.quantity_sold)
FROM sales s
, products p
, customers c
WHERE s.cust_id = c.cust_id
AND s.prod_id = p.prod_id
GROUP BY s.time_id
, s.prod_id
, p.prod_name
, s.cust_id
, c.cust_first_name
, c.cust_last_name;
COMMIT;

```

3. マテリアライズド・ビューを作成します。

```

CREATE MATERIALIZED VIEW sales_prod_cust_mv
ON PREBUILT TABLE
ENABLE QUERY REWRITE
AS SELECT s.time_id
, s.prod_id
, p.prod_name
, s.cust_id
, c.cust_first_name
, c.cust_last_name
, SUM(s.amount_sold) amount_sold
, SUM(s.quantity_sold) quantity_sold
FROM sales s
, products p
, customers c
WHERE s.cust_id = c.cust_id
AND s.prod_id = p.prod_id
GROUP BY s.time_id
, s.prod_id
, p.prod_name
, s.cust_id
, c.cust_first_name
, c.cust_last_name;

```

4. 新規パーティションで交換される個別の表をロードします。

```
CREATE TABLE sales_q1_2006 PARALLEL COMPRESS
AS SELECT * FROM sales
WHERE 0 = 1;

/* This would be the regular ETL job */

ALTER SESSION ENABLE PARALLEL DML;

INSERT /* PARALLEL qs */ INTO sales_q1_2006 qs
SELECT /* PARALLEL s */ prod_id
, cust_id
, add_months(time_id,3)
, channel_id
, promo_id
, quantity_sold
, amount_sold
FROM sales PARTITION(sales_q4_2005) s;

COMMIT;

CREATE BITMAP INDEX bmp_indx_prod_id ON sales_q1_2006 (prod_id);
CREATE BITMAP INDEX bmp_indx_cust_id ON sales_q1_2006 (cust_id);
CREATE BITMAP INDEX bmp_indx_time_id ON sales_q1_2006 (time_id);
CREATE BITMAP INDEX bmp_indx_channel_id ON sales_q1_2006 (channel_id);
CREATE BITMAP INDEX bmp_indx_promo_id ON sales_q1_2006 (promo_id);

ALTER TABLE sales_q1_2006 ADD CONSTRAINT sales_q_prod_fk
FOREIGN KEY (prod_id) REFERENCES products(prod_id) ENABLE NOVALIDATE;

ALTER TABLE sales_q1_2006 ADD CONSTRAINT sales_q_cust_fk
FOREIGN KEY (cust_id) REFERENCES customers(cust_id) ENABLE NOVALIDATE;

ALTER TABLE sales_q1_2006 ADD CONSTRAINT sales_q_time_fk
FOREIGN KEY (time_id) REFERENCES times(time_id) ENABLE NOVALIDATE;

ALTER TABLE sales_q1_2006 ADD CONSTRAINT sales_q_channel_fk
FOREIGN KEY (channel_id) REFERENCES channels(channel_id) ENABLE NOVALIDATE;

ALTER TABLE sales_q1_2006 ADD CONSTRAINT sales_q_promo_fk
FOREIGN KEY (promo_id) REFERENCES promotions(promo_id) ENABLE NOVALIDATE;

BEGIN
  DBMS_STATS.GATHER_TABLE_STATS('SH','SALES_Q1_2006');
END;
/
```

5. マテリアライズド・ビューのパーティションで交換される個別の表を作成およびロードします。

```
CREATE TABLE sales_mv_q1_2006 PARALLEL COMPRESS
AS SELECT * FROM sales_prod_cust_mv
WHERE 1 = 0;

ALTER SESSION ENABLE PARALLEL DML;

INSERT /*+ PARALLEL smv */ INTO sales_mv_q1_2006 smv
SELECT /*+ PARALLEL s PARALLEL c */ s.time_id
, s.prod_id
, p.prod_name
```

```

, s.cust_id
, cust_first_name
, c.cust_last_name
, SUM(s.amount_sold)
, SUM(s.quantity_sold)
FROM sales_q1_2006 s
, products p
, customers c
WHERE s.cust_id = c.cust_id
AND s.prod_id = p.prod_id
GROUP BY s.time_id
, s.prod_id
, p.prod_name
, s.cust_id
, c.cust_first_name
, c.cust_last_name;

COMMIT;

```

6. 統計を収集します。

```

BEGIN
  DBMS_STATS.GATHER_TABLE_STATS('SH', 'SALES_MV_Q1_2006');
END;

```

7. パーティションを交換します。

```

ALTER TABLE sales
EXCHANGE PARTITION sales_q1_2006
WITH TABLE sales_q1_2006
INCLUDING INDEXES WITHOUT VALIDATION;

ALTER TABLE sales_prod_cust_mv
EXCHANGE PARTITION p2006q1
WITH TABLE sales_mv_q1_2006
INCLUDING INDEXES WITHOUT VALIDATION;

```

8. マテリアライズド・ビューが再度フレッシュになったことがデータベースに通知されます。

```

ALTER MATERIALIZED VIEW sales_prod_cust_mv CONSIDER FRESH;

```

このシナリオでは、事前作成表を使用し、この制約は RELY 制約ではないため、問合せのリライト機能は query_rewrite_integrity を STALE_TOLERATED に設定した場合にのみ有効です。

ローリング・ウィンドウを使用したデータのオフロード

データを削除したりアーカイブする際にローリング・ウィンドウを使用すると特に効果的です。たとえば、最新の 36 か月の売上データをデータ・ウェアハウスに格納する場合にローリング・ウィンドウを使用します。新規パーティションは、新しい月の sales 表にそれぞれ追加され、古いパーティションは sales 表から削除される場合があります。この方法では、ウェアハウスに常に 36 か月分のデータが保持されます。

例：ローリング・ウィンドウの使用

次に sh スキーマの sales 表のローリング・ウィンドウの例を示します。

ローリング・ウィンドウを使用する手順

1. 2005 年 12 月の売上を追加します。

```
ALTER TABLE sales  
ADD PARTITION sales_12_2005 VALUES LESS THAN ('01-JAN-2006');
```

既存の索引を再作成する必要があります。

2. 1999 年のパーティションを削除します。

```
ALTER TABLE sales  
DROP PARTITION sales_1999;
```

データ・ウェアハウスの操作の最適化

この項では、データ・ウェアハウスのパフォーマンスを最適化する方法を説明します。また次の項が含まれます。

- システムのオーバーロードの回避
- 索引およびマテリアライズド・ビューの使用の最適化
- 記憶域要件の最適化

システムのオーバーロードの回避

この項では、システムのオーバーロードを識別して回避する方法を説明します。一般的に、『Oracle Database 2 日でパフォーマンス・チューニング・ガイド』の説明に従って、自動診断機能である自動データベース診断モニター（ADDM）を使用してデータベースでのパフォーマンスの問題を識別する必要があります。この項では、システムにおけるパフォーマンスの問題を回避する別の方法を説明します。また次の項が含まれます。

- システム・パフォーマンスの監視
- データベース・リソース・マネージャの使用

システム・パフォーマンスの監視

この項では、重要なメトリックを定期的に監視することによりシステムのオーバーロードを回避する方法の詳細を提供します。Oracle Enterprise Manager の「データベース・パフォーマンス」ページを介してこれらのメトリックを監視できます。この項の内容は次のとおりです。

- パラレル実行パフォーマンスの監視
- I/O の監視

パラレル実行パフォーマンスの監視

この項では、パラレル実行パフォーマンスを監視する方法を説明します。多くのパラレル文がダウングレードされているとします。これは、パフォーマンスの問題を示します。想定よりも低い並列度で文が実行されると大幅に時間がかかり、文がダウングレードされたかどうかによって実行時間が異なります。パラレル文がダウングレードする原因には次のようなものが考えられます。

- 最初の並列度が高く、低くする必要がある場合。
- 使用可能なパラレル・サーバーが十分になく、システムがオーバーロードしている場合。

パラレル実行を監視する手順

1. データベースのホームページで「パフォーマンス」をクリックします。
「パフォーマンス」ページが表示されます。
2. ページを下へスクロールします。リンクのリストの下で、「PQ」をクリックします。
「PQ」ページが表示されます。次の項目に対してパラレル問合せのパフォーマンス特性が表示されます。
 - パラレル・セッション
 - パラレル・スレーブ
 - DML および DDL のパラレル化
 - シリアライズおよび文のダウングレード

図 11-1 パラレル実行の監視



I/O の監視

この項では、I/O パフォーマンスを監視する方法を説明します。システムのスループットがシステム構成に基づいて予想していたものより大幅に低く（第2章「データ・ウェアハウス・システムの設定」を参照）、ユーザーがパフォーマンスに関して不満がある場合は、パフォーマンスに問題があると考えられます。適切に構成されたシステムで通常のデータ・ウェアハウス・ワークロードを処理する場合、単一の I/O ブロックに対して、おおむね良好な I/O と、待ち時間が比較的少ない（30 分以下）が予想されます。

I/O パフォーマンス監視する手順

1. データベースのホームページで「パフォーマンス」をクリックします。

「パフォーマンス」ページが表示されます。

2. ページを下へスクロールします。リンクのリストの下で、「I/O」をクリックします。

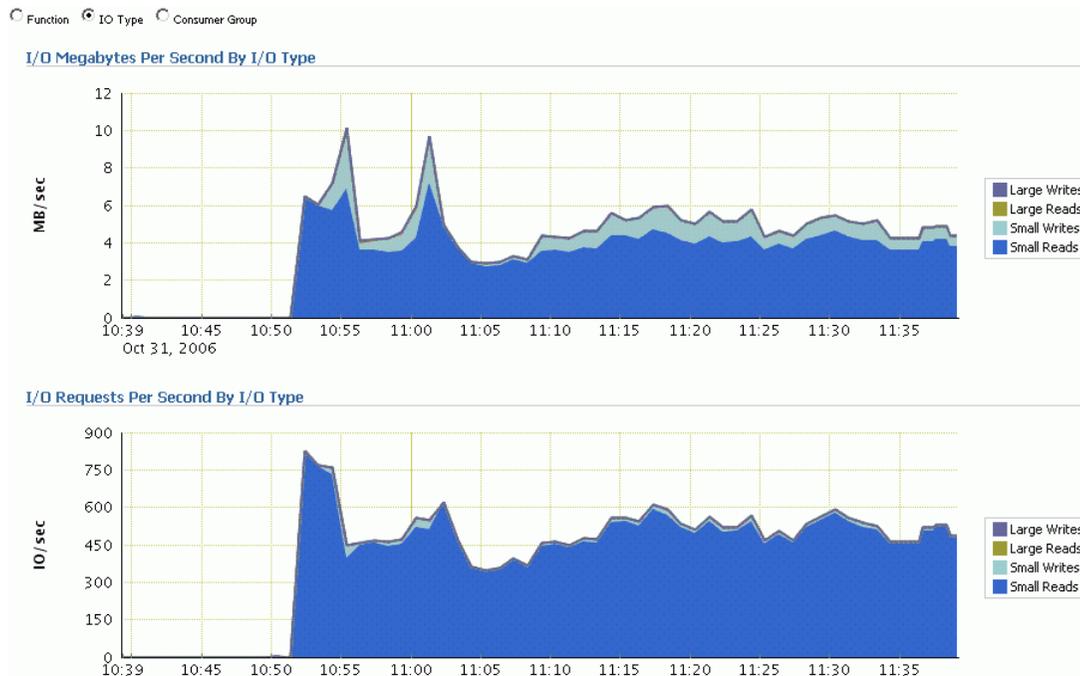
I/O ページが表示され、「ファンクションごとの I/O MB/秒」と「ファンクションごとの I/O リクエスト数/秒」が表示されます。

3. 読取り / 書込み操作に関する詳細は、「I/O タイプ」を選択します。

次の項目に対して I/O 詳細が表示されます。

- 大規模の書込み
- 大規模の読取り
- 小規模の書込み
- 小規模の読取り

図 11-2 I/O の監視



データベース・リソース・マネージャの使用

データベース・リソース・マネージャには、Oracle システム内の作業に優先度を設定する機能があります。たとえば、優先度の高いジョブのあるユーザーは、オンライン作業のレスポンス時間を最短に抑えるためにリソースを取得する一方、バッチ・ジョブやレポートなどの優先度の低いジョブのあるユーザーに対しては、レスポンス時間が遅くなります。このように優先度を割り当てることにより、リソースをより細かく制御できます。また、コンシューマ・グループに対して、コンシューマ・グループの自動切替え、最大アクティブ・セッションの制御、問合せ実行時間の見積りおよび UNDO プール割当て制限などの機能があります。

各コンシューマ・グループの同時にアクティブなセッションの最大数を指定できます。この制限に到達した場合、データベース・リソース・マネージャにより、すべての後続のリクエストが並べられ、既存のアクティブ・セッションが完了した後にのみそれらが実行されます。

データベース・リソース・マネージャは Oracle Database の一部で、データベース内部の異なるプロセスを区別できます。結果として、データベース・リソース・マネージャにより、データベース内部で実行されている個別の操作に優先度を割り当てることができます。

データベース・リソース・マネージャを使用して、次のことを実行できます。

- システムにかかる負荷およびユーザー数に関係なく、処理リソースの最低量をユーザーに保証します。
- CPU 時間の割合を異なるユーザーおよびアプリケーションに割り当てることにより、使用可能な処理リソースを配分します。データ・ウェアハウスでは、バッチ・ジョブよりもリレーショナル・オンライン分析処理 (ROLAP) アプリケーションに対する割当ての方が多くなります。
- 管理者が定義した基準に基づいて、ユーザーを 1 つのグループから他のグループに自動的に切り替えます。あるユーザー・グループのメンバーが、指定時間より長くかかるセッションを作成した場合、そのセッションはリソース要件の異なる他のユーザー・グループに自動的に切り替えられます。
- 特定のリソース割当て方法を使用するようにインスタンスを構成します。この方法は、インスタンスを停止して再起動する必要なく、日中設定から夜間設定などのように動的に変更できます。

索引およびマテリアライズド・ビューの使用の最適化

索引およびマテリアライズド・ビューを適切に使用することにより、データ・ウェアハウスのパフォーマンスを向上させることができます。SQL アクセス・アドバイザーの主な利点は、現行のワークロードを推奨事項の基準として使用できることです。

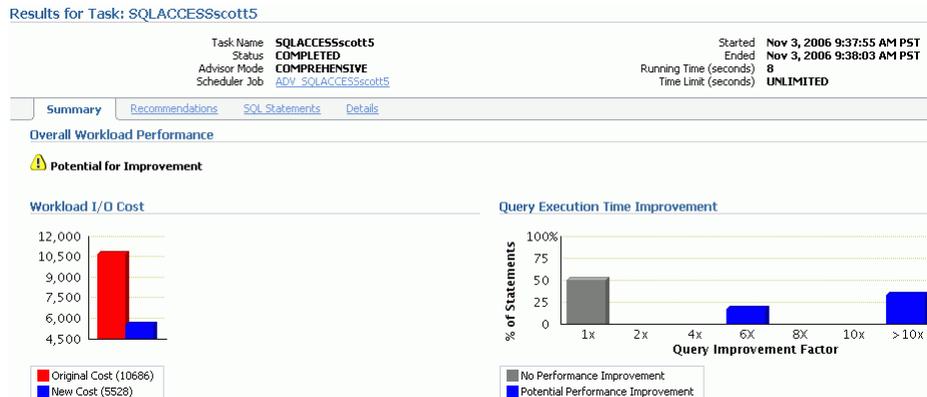
例：SQL アクセス・アドバイザーを使用した索引およびマテリアライズド・ビューの使用の最適化

この例では、特定の索引およびマテリアライズド・ビューを利用しているシステムで実行されているワークロードを所有すると仮定します。

索引およびマテリアライズド・ビューを最適化する手順

1. 「セントラル・アドバイザー」ページから、「SQL アドバイザ」をクリックします。
「アドバイザー」ページが表示されます。
2. 「アドバイザー」ページから、「SQL アクセス・アドバイザー」をクリックします。
「SQL アクセス・アドバイザー」ページが表示されます。
3. 「デフォルト・オプションを使用」を選択して、「続行」をクリックします。
「ワークロード・ソース」ページが表示されます。
4. 「既存の SQL チューニング・セットを使用します。」で、ワークロード・ソースを選択します。に移動し、「選択」をクリックします。次に「次へ」をクリックします。
「推奨オプション」ページが表示されます。
5. 「索引」、「マテリアライズド・ビュー」および「包括モード」を選択します。「次へ」をクリックします。
「スケジュール」ページが表示されます。
6. 「発行」をクリックします。
「推奨」ページが表示されます。
7. 「名前」フィールドに名前を入力し、開始する必要がある場合は「即時」を選択します。「次へ」をクリックします。
「確認」ページが表示されます。
8. 「発行」をクリックします。
「確認」ページが表示されます。
9. タスク名を選択し、「結果の表示」をクリックします。
「タスクの結果」ページが表示されます。このページには、可能なコスト改善が表示されません。追加情報は、「推奨」、「SQL 文」または「詳細」で表示できます。

図 11-3 推奨される改善



記憶域要件の最適化

データベース・ブロック内の重複値を排除してアーカイブされたデータを圧縮することにより、記憶域要件を削減できます。圧縮可能なデータベース・オブジェクトには、表およびマテリアライズド・ビューが含まれます。パーティション表では、パーティションの一部を圧縮するかまたはすべてのパーティションを圧縮するかを選択できます。圧縮属性は、表領域、表または表のパーティションに対して宣言できます。表領域レベルで宣言された場合、その表領域で作成されたすべての表はデフォルトで圧縮されます。表（またはパーティションまたは表領域）の圧縮属性は変更可能で、その変更はその表に移動する新規データのみにも適用されます。この結果、単一の表またはパーティションにはいくつかの圧縮されたブロックといくつかの通常のブロックが含まれます。これにより、圧縮後もデータのサイズが増加しないことが保証されます。圧縮によってブロックのサイズが増加する場合は、圧縮はそのブロックには適用されません。

データ圧縮による記憶域の改善

いくつかのパーティションを圧縮するか、またはパーティション化されたヒープ構成表を作成できます。これを実行するには、完全なパーティション化された表を圧縮対象の表として定義するか、またはパーティション・レベルごとに定義するか、いずれかの定義を行います。特定の宣言のないパーティションは表定義から属性を継承し、表レベルの指定がない場合は表領域定義から継承します。

パーティションを圧縮するかまたは未圧縮のままにするかについての決定は、パーティション化されていない表と同じルールに従います。ただし、範囲の権限およびデータを論理的に個別のパーティションに区切るコンポジット・パーティション化のため、パーティション表などは、主に読取り専用のデータ（パーティション）の圧縮する部分として適切な候補です。たとえばこれは、古いデータが使用不可になる前の中間の段階としてのすべてのローリング・ウィンドウ操作に役立ちます。データを圧縮することにより、より多くの古いデータをオンラインで保持でき、追加の記憶域の使用量の負荷を最小化できます。

また、既存の未圧縮の表パーティションを後で変更したり、新規の圧縮パーティションおよび未圧縮パーティションを追加したり、MERGE PARTITION、SPLIT PARTITION または MOVE PARTITION などのデータ移動が必要なパーティション・メンテナンス操作の一部として圧縮属性を変更できます。パーティションにはデータを含めることができ、または空にすることもできます。

一部または完全に圧縮されたパーティション表のアクセスおよびメンテナンスは、完全に未圧縮のパーティション表と同じです。また、完全に未圧縮のパーティション表に適用されたすべてのルールも、一部または完全に圧縮されたパーティション表に対して有効です。

データ圧縮を使用する手順

次の例では、1つの圧縮パーティション、`costs_old`を使用してレンジ・パーティション化された表を作成します。表の圧縮属性および他のすべてのパーティションは表領域レベルから継承されます。

```
CREATE TABLE costs_demo (  
  prod_id    NUMBER(6),    time_id    DATE,  
  unit_cost  NUMBER(10,2), unit_price NUMBER(10,2))  
PARTITION BY RANGE (time_id)  
  (PARTITION costs_old  
    VALUES LESS THAN (TO_DATE('01-JAN-2003', 'DD-MON-YYYY')) COMPRESS,  
   PARTITION costs_q1_2003  
    VALUES LESS THAN (TO_DATE('01-APR-2003', 'DD-MON-YYYY')),  
   PARTITION costs_q2_2003  
    VALUES LESS THAN (TO_DATE('01-JUN-2003', 'DD-MON-YYYY')),  
   PARTITION costs_recent VALUES LESS THAN (MAXVALUE));
```

パフォーマンス・ボトルネックの排除

この項では、パフォーマンスの問題を識別して削減する方法を説明します。内容は次のとおりです。

- SQL の効率的な実行の検証
- リソース使用量の最小化によるパフォーマンスの向上
- 最適なリソースの使用

SQL の効率的な実行の検証

システムの正常な実行を保証するために重要なことは、パフォーマンスの問題を排除することです。この項では、これらのボトルネックを検索して排除するいくつかの方法を説明します。内容は次のとおりです。

- [オブティマイザ統計の分析](#)
- [実行計画の分析](#)
- [ヒントを使用したデータ・ウェアハウスのパフォーマンスの向上](#)
- [アドバイザを使用した SQL パフォーマンスの検証](#)

オブティマイザ統計の分析

オブティマイザ統計は、データベース内のデータベースおよびオブジェクトについてより詳細に説明するデータが収集されたものです。これらの統計はデータ・ディクショナリに格納され、問合せオブティマイザにより使用されて各 SQL 文に最適な実行計画を選択します。オブティマイザ統計には、次が含まれます。

- 表統計（行数、ブロックおよび平均の行長さ）
- 列統計（列内の固有の値の数、列内の NULL 値の数およびデータ配分）
- 索引統計（リーフ・ブロックの数、レベルおよびクラスタリング・ファクタ）
- システム統計（CPU と I/O のパフォーマンスおよび使用率）

オブティマイザ統計は、データ・ディクショナリに格納されます。オブティマイザ統計は、次に類似するデータ・ディクショナリ・ビューを使用して参照できます。

```
SELECT * FROM DBA_SCHEDULER_JOBS WHERE JOB_NAME = 'GATHER_STATS_JOB';
```

データベース内のオブジェクトは常に変更されるので、これらのデータベース・オブジェクトを正確に示すために統計を定期的に更新する必要があります。統計は、Oracle Database により自動的に維持されるか、または DBMS_STATS パッケージを使用して手動でオブティマイザ統計を維持することもできます。

実行計画の分析

SQL 文を実行するには、Oracle Database では多数のステップを実行する必要があります。これらの各ステップでは、データベースから物理的にデータ行を取得するか、文を発行するユーザー用になんらかの方法でそれらのデータを用意しておきます。Oracle Database が文の実行に使用するステップの組合せは、実行計画と呼ばれます。実行計画には、文がアクセスする各表のアクセス・パスおよび適切な結合方法を使用した表の順序（結合順序）が含まれます。

EXPLAIN PLAN 文を使用して、オブティマイザが SQL 文用に選択した実行計画を調査できます。その文が発行されると、オブティマイザにより実行計画が選択され、計画を説明するデータがデータベース表に挿入されます。単純に EXPLAIN PLAN 文を発行し出力表を問い合わせることで実行できます。

EXPLAIN PLAN 文を使用する一般的なガイドラインは次のとおりです。

- SQL スクリプトである UTLXPLAN.SQL を使用して、スキーマ内にある PLAN_TABLE というサンプルの出力表を作成します。
- SQL 文より前に EXPLAIN PLAN FOR 句を含めます。
- EXPLAIN PLAN 文を発行した後、Oracle Database により提供されたスクリプトの 1 つまたはパッケージの 1 つを使用して最新の計画の表の出力を表示します。
- EXPLAIN PLAN 出力内の実行順序は、最も右側にインデントされた行から始まります。2 つの行が等しくインデントされている場合は、通常、上位の行が最初に実行されます。

例：実行計画の出力の分析

次の文では、2つの EXPLAIN PLAN 文の出力を示し、1つは動的プルーニングを使用し、もう1つは静的プルーニングを使用します。

EXPLAIN PLAN の出力を分析する手順

```
EXPLAIN PLAN FOR
SELECT p.prod_name
, c.channel_desc
, SUM(s.amount_sold) revenue
FROM products p
, channels c
, sales s
WHERE s.prod_id = p.prod_id
AND s.channel_id = c.channel_id
AND s.time_id BETWEEN '01-12-2001' AND '31-12-2001'
GROUP BY p.prod_name
, c.channel_desc;
SELECT * FROM TABLE(DBMS_XPLAN.DISPLAY);
```

WITHOUT TO_DATE

| Id | Operation | Name | Rows | Bytes | Cost | Time | Pstart | Pstop |
|-----|-----------------------------------|----------------|------|-------|---------|----------|--------|-------|
| 0 | SELECT STATEMENT | | 252 | 15876 | 305 (1) | 00:00:06 | | |
| 1 | HASH GROUP BY | | 252 | 15876 | 305 (1) | 00:00:06 | | |
| *2 | FILTER | | | | | | | |
| *3 | HASH JOIN | | 2255 | 138K | 304 (1) | 00:00:06 | | |
| 4 | TABLE ACCESS FULL | PRODUCTS | 72 | 2160 | 2 (0) | 00:00:01 | | |
| 5 | MERGE JOIN | | 2286 | 75438 | 302 (1) | 00:00:06 | | |
| 6 | TABLE ACCESS BY INDEX ROWID | CHANNELS | 5 | 65 | 2 (0) | 00:00:01 | | |
| 7 | INDEX FULL SCAN | CHANNELS_PK | 5 | | 1 (0) | 00:00:01 | | |
| *8 | SORT JOIN | | 2286 | 45720 | 299 (1) | 00:00:06 | | |
| 9 | PARTITION RANGE ITERATOR | | 2286 | 45720 | 298 (0) | 00:00:06 | KEY | KEY |
| 10 | TABLE ACCESS BY LOCAL INDEX ROWID | SALES | 2286 | 45720 | 298 (0) | 00:00:06 | KEY | KEY |
| 11 | BITMAP CONVERSION TO ROWIDS | | | | | | | |
| *12 | BITMAP INDEX RANGE SCAN | SALES_TIME_BIX | | | | | KEY | KEY |

Predicate Information (identified by operation id):

```
2 - filter("TO_DATE('01-12-2001') <= TO_DATE('31-12-2001')")
3 - access("S"."PROD_ID"="P"."PROD_ID")
8 - access("S"."CHANNEL_ID"="C"."CHANNEL_ID")
   filter("S"."CHANNEL_ID"="C"."CHANNEL_ID")
12 - access("S"."TIME_ID" >= '01-12-2001' AND "S"."TIME_ID" <= '31-12-2001')
```

Note the values of KEY KEY for Pstart and Pstop.

WITH TO_DATE

| Id | Operation | Name | Rows | Bytes | Cost (%CPU) | Time | Pstart | Pstop |
|----|------------------------|----------|-------|-------|-------------|----------|--------|-------|
| 0 | SELECT STATEMENT | | 252 | 15876 | 31 (20) | 00:00:01 | | |
| 1 | HASH GROUP BY | | 252 | 15876 | 31 (20) | 00:00:01 | | |
| *2 | HASH JOIN | | 21717 | 1336K | 28 (11) | 00:00:01 | | |
| 3 | TABLE ACCESS FULL | PRODUCTS | 72 | 2160 | 2 (0) | 00:00:01 | | |
| *4 | HASH JOIN | | 21717 | 699K | 26 (12) | 00:00:01 | | |
| 5 | TABLE ACCESS FULL | CHANNELS | 5 | 65 | 3 (0) | 00:00:01 | | |
| 6 | PARTITION RANGE SINGLE | | 21717 | 424K | 22 (10) | 00:00:01 | 20 | 20 |
| *7 | TABLE ACCESS FULL | SALES | 21717 | 424K | 22 (10) | 00:00:01 | 20 | 20 |

Predicate Information (identified by operation id):

```
2 - access("S"."PROD_ID"="P"."PROD_ID")
4 - access("S"."CHANNEL_ID"="C"."CHANNEL_ID")
7 - filter("S"."TIME_ID">=TO_DATE('2001-12-01 00:00:00', 'yyyy-mm-dd hh24:mi:ss') AND
          "S"."TIME_ID"<=TO_DATE('2001-12-31 00:00:00', 'yyyy-mm-dd hh24:mi:ss'))
```

Note the values of 20 20 for Pstart and Pstop.

最初の実行計画は、Pstart および Pstop それぞれの KEY KEY 値を使用して動的ブルーニングを表示します。動的ブルーニングは、実行時間にデータベースによりアクセスするパーティションを特定する必要があることを意味します。静的ブルーニングの場合は、解析時間にデータベースにより、より効果的な実行のためにアクセスするパーティションを把握します。

明示的な日付変換を使用して実行計画を頻繁に改善できます。明示的な日付変換の使用は、最適なパーティション・ブルーニングおよび索引の使用のためのベスト・プラクティスです。

ヒントを使用したデータ・ウェアハウスのパフォーマンスの向上

決定事項に関するヒントは通常オプティマイザによって作成されます。オプティマイザによって管理されていないデータについての情報をアプリケーション開発者として保持している場合があります。ヒントによってメカニズムが提供され、特定の基準に基づく問合せ実行プランを選択するためにオプティマイザに通知されます。

たとえば、ユーザーが特定の問合せに対して特定の索引が選択可能であることを把握している場合があります。この情報に基づいて、オプティマイザが選択するより、より効率的な実行プランが選択できる場合があります。このような場合、ヒントを使用して、オプティマイザに通知し、最適な実行プランを使用します。

デフォルトでは、Oracle Warehouse Builder には、一般的なデータ・ロードを最適化するヒントが含まれます。詳細は、『Oracle Warehouse Builder ユーザーズ・ガイド』を参照してください。

例：ヒントを使用したデータ・ウェアハウスのパフォーマンスの向上

システムがアイドル状態である間に昨年の売上表のサマリーを迅速に実行することを想定します。この場合は、次の文を発行できます。

データ・ウェアハウスのパフォーマンスを向上させるためにヒントを使用する手順

```
SELECT /*+ PARALLEL(s,16) */ SUM(amount_sold)
FROM sales s
WHERE s.time_id BETWEEN TO_DATE('01-JAN-2005','DD-MON-YYYY')
      AND TO_DATE('31-DEC-2005','DD-MON-YYYY');
```

データ・ウェアハウスでヒントを使用するもう 1 つの一般的な方法は、圧縮の使用によるレコードの効率的なロードを保証することです。次の SQL に示すように、APPEND ヒントを使用します。

```
...
INSERT /* +APPEND */ INTO my_materialized_view
...
```

アドバイザーを使用した SQL パフォーマンスの検証

SQL チューニング・アドバイザーおよび SQL アクセス・アドバイザーを使用すると、特定の SQL 文または SQL 文のセットを調べるアドバイス・モードで問合せ最適化を起動でき、SQL 文の効率を向上させる推奨事項が提供されます。SQL チューニング・アドバイザーおよび SQL アクセス・アドバイザーは SQL プロファイルの作成、SQL 文の再構築、付加的な索引またはマテリアライズド・ビューの作成、および最適化統計のリフレッシュなどの様々なタイプの推奨事項を作成できます。さらに、Oracle Enterprise Manager を使用すると、数回のマウス・クリックでこれらの推奨事項を受け入れ、実装できます。

SQL アクセス・アドバイザーは、主に索引およびマテリアライズド・ビューの追加および削除などのスキーマ変更の推奨事項を作成する場合に使用します。また、パーティション計画も推奨します。SQL チューニング・アドバイザーは、SQL プロファイルの作成、SQL 文の再構築などの他のタイプの推奨事項の作成に使用します。新しい索引を作成してパフォーマンスが大幅に向上できる場合、SQL チューニング・アドバイザーは索引の作成を推奨する可能性があります。ただし、これらの推奨事項は、典型的な SQL 文のセットを含んだ SQL ワークロードを使用して SQL アクセス・アドバイザーを実行し、検証する必要があります。

例：SQL チューニング・アドバイザーを使用した SQL パフォーマンスの検証

SQL チューニング・アドバイザーを使用して、単一または複数の SQL 文をチューニングできます。複数の SQL 文をチューニングする場合、SQL チューニング・アドバイザーは SQL 文間の相互依存を認識しないことに注意してください。かわりに、多数の SQL 文に対して SQL チューニング・アドバイザーを実行すると有効です。

SQL チューニング・アドバイザーを実行して SQL パフォーマンスを検証する手順は、次のとおりです。

1. 「セントラル・アドバイザー」ページに移動して、「SQL アドバイザ」をクリックします。
「SQL アドバイザ」ページが表示されます。
2. 「SQL チューニング・アドバイザーのスケジュール」をクリックします。
「SQL チューニング・アドバイザーのスケジュール」ページが表示されます。推奨される名前は「名前」フィールドに表示され、変更が可能です。その後「包括」を選択して実行する包括的な分析を取得します。「スケジュール」の「即時」を選択します。適切な SQL チューニング・セットを選択した後、「OK」をクリックします。
3. 「処理中」ページが表示されます。その後、「推奨」ページにパフォーマンス向上の推奨事項が表示されます。「推奨の表示」をクリックします。
「推奨」ページが表示されます。

図 12-1 SQL チューニング・アドバイザー：推奨

Only one recommendation should be implemented.

SQL Text

```
select sum(s.amount_sold) from sales_copy s, customers c where s.cust_id = c.cust_id and c.cust_last_name = 'Sexton'
```

Select Recommendation
 Original Explain Plan (Annotated)

Implement

| Select | Type | Findings | Recommendations | Rationale | New Benefit Explain (%) Plan |
|--------|-------|---|--|--|------------------------------|
| 1 | Index | The execution plan of this statement can be improved by creating one or more indices. | Consider running the Access Advisor to improve the physical schema design or creating the recommended index. SH.CUSTOMERS("CUST_LAST_NAME") SH.SALES_COPY("CUST_ID") | Creating the recommended indices significantly improves the execution plan of this statement. However, it might be preferable to run "Access Advisor" using a representative SQL workload as opposed to a single statement. This will allow to get comprehensive index recommendations which takes into account index maintenance overhead and additional space consumption. | 100.0 |

4. 推奨事項は、「実装」をクリックして実装できる索引を作成することです。または、必要に応じて SQL アクセス・アドバイザーも実行できます。

リソース使用量の最小化によるパフォーマンスの向上

リソースの使用量を最小化できるので、次の機能を使用してデータ・ウェアハウスのパフォーマンスを向上させることができます。

- パフォーマンスの向上:パーティション化
- パフォーマンスの向上:問合せのリライトおよびマテリアライズド・ビュー
- パフォーマンスの向上:索引

パフォーマンスの向上:パーティション化

データ・ウェアハウスにはサイズの大きな表が含まれていることが多く、これらのサイズの大きな表を管理する技術およびこれらの表全体に良質な問合せのパフォーマンスを提供する技術の両方が必要になります。この項では、パーティション化とこれらの要件に対処するための主要な方法を説明します。データ・ウェアハウスの問合せのパフォーマンスに関連する2つの機能は、パーティション・プルーニングとパーティションワイズ結合です。

パフォーマンスの向上:パーティション・プルーニング

パーティション・プルーニングは、データ・ウェアハウスに必須のパフォーマンス機能です。パーティション・プルーニングでは、最適化により SQL 文の FROM 句および WHERE 句が分析され、パーティション・アクセス・リストの作成時に不必要なパーティションが排除されます。これにより、Oracle Database を使用して SQL 文に関連するパーティションでのみ操作を実行できます。レンジ・パーティション化列またはリスト・パーティション化列にある範囲述語、LIKE 述語、等価述語および IN-リスト述語を使用する場合とハッシュ・パーティション化列にある等価述語および IN-リストを使用する場合は、Oracle Database によりパーティションがプルーニングされます。

パーティション・プルーニングによりディスクから取得するデータ量を大幅に削減し処理時間の使用を短縮できるので、問合せのパフォーマンスおよびリソース使用量が向上します。グローバルなパーティション索引を使用して異なる列の索引および表をパーティション化する場合は、基礎となる表が排除できない場合でもパーティション・プルーニングにより索引パーティションが排除されます。

実際の SQL 文に応じて、Oracle Database により静的プルーニングおよび動的プルーニングが使用されます。静的プルーニングは事前にアクセスされたパーティションの情報とともにコンパイル時に発生し、一方動的プルーニングは実行時に発生し、これは文によってアクセスされる正確なパーティションを事前に把握していることを意味します。静的プルーニングのサンプルの使用例は、パーティション・キー列に固定リテラルがある WHERE 条件を含む SQL 文です。動的プルーニングの例は、WHERE 条件内の演算子または関数の使用です。

パーティション・プルーニングは、プルーニングが発生するオブジェクトの統計に影響するので、文の実行計画にも影響します。

パフォーマンスの向上:パーティションワイズ結合

パーティションワイズ結合では、結合がパラレルで実行される場合にパラレル実行サーバー間で交換されるデータの量を最小化することにより問合せのレスポンス時間が削減されます。これによりレスポンス時間が大幅に削減され、CPU およびメモリー・リソース両方の使用量を改善します。Oracle Real Application Clusters 環境では、パーティションワイズ結合は相互接続のデータ通信を回避するか、少なくとも制限します。このことは大規模な結合操作に対する良質なスケーラビリティを実現するために重要です。

パーティションワイズ結合では、全体的または部分的です。Oracle Database により使用する結合のタイプが決定されます。

例：SQL アクセス・アドバイザーを使用したパーティション化の評価

データ・ウェアハウス環境では、常にパーティション化を考慮する必要があります。

パーティション化を評価する手順は、次のとおりです。

- 「セントラル・アドバイザー」ページで、「SQL アドバイザ」をクリックします。
「SQL アドバイザ」ページが表示されます。
- 「SQL アクセス・アドバイザー」をクリックします。
「SQL アクセス・アドバイザー」ページが表示されます。
- 「初期オプション」から、「デフォルト・オプションを使用」を選択し、「続行」をクリックします。
- 「ワークロード・ソース」から、「現在と最近の SQL アクティビティ」を選択し、「次へ」をクリックします。
「推奨オプション」ページが表示されます。
- 「パーティショニング」を選択し、次に「包括モード」を選択して、「次へ」をクリックします。
「スケジュール」ページが表示されます。
- 「タスク名」フィールドに SQLACCEStest1 を入力し、「次へ」をクリックします。
「確認」ページが表示されます。「発行」をクリックします。

図 12-2 パーティション化の評価

SQL Access Advisor: Review
Database database Cancel Show SQL Back Step 4 of 4 Submit

Please review the SQL Access Advisor options and values you have selected.

Task Name **SQLACCESS_partition**
Task Description **SQL Access Advisor**
Scheduled Start Time **Run Immediately**

Options

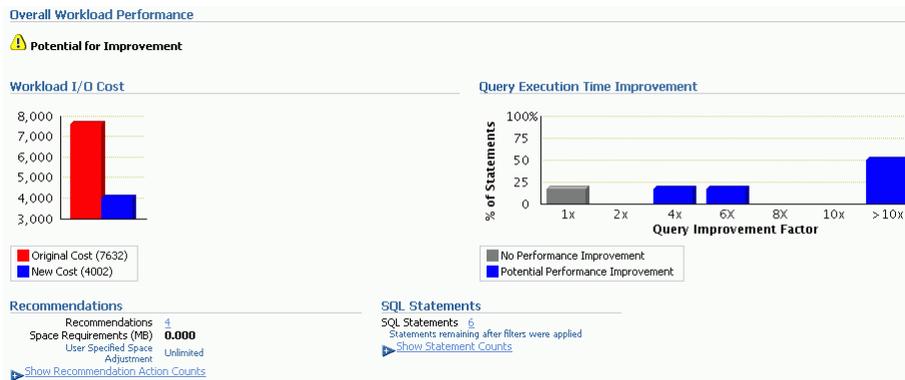
Show All Options

| Modified Option | Value | Description |
|--|---------------------------|---|
| <input checked="" type="checkbox"/> Analysis Scope | Partitions on Base Tables | The type of recommendations that are allowed |
| <input checked="" type="checkbox"/> Workload SQL Limit | 25 | Specifies the number of SQL statements to be analyzed |

Cancel Show SQL Back Step 4 of 4 Submit

- 「発行」をクリックします。
「確認」ページが表示されます。
- タスクを選択し、「結果の表示」をクリックします。「タスクの結果」ページが表示され、パーティション化の結果として可能な向上が示されます。

図 12-3 パーティション化の結果



パフォーマンスの向上 : 問合せのリライトおよびマテリアライズド・ビュー

データ・ウェアハウスでは、マテリアライズド・ビューを使用して売上の合計などの集計データを計算して格納できます。また、それらを使用して集計を加えてまたは集計を除いて結合を計算でき、コストの高い計算と同様、頻繁に実行されるサイズの大きな表間のコストの高い結合にも便利です。マテリアライズド・ビューにより、大規模な結合または問合せを処理する前にサマリーが作成されたデータを計算して格納するので、大規模または重要なクラスの間合せのためのコストの高い結合および集計に関連付けられたオーバーヘッドは排除されます。これらの環境でマテリアライズド・ビューは多くの場合、サマリーと呼ばれます。

マテリアライズド・ビューを作成して保持する主な利点の1つは、表またはビューで表されるSQL文を、詳細表で定義される1つ以上のマテリアライズド・ビューにアクセスする文に変換する問合せのリライト機能を利用できることです。変換はユーザーおよびアプリケーションに対して透過的で、SQL文内のマテリアライズド・ビューを使用したり参照する必要はありません。問合せのリライト機能は透過的なので、索引と同様にアプリケーション・コード内のSQLを無効にすることなく、マテリアライズド・ビューを追加または削除できます。

基礎となる表に大量のデータが含まれる場合、必要な集計の計算またはこれらの表間の結合の計算は、コストが高く時間がかかる処理になります。このような場合、問合せにより応答が返されるまで数分または数時間かかります。マテリアライズド・ビューにはすでに計算された集計と結合が含まれているので、Oracle Databaseによって強力な問合せのリライト・プロセスが使用され、マテリアライズド・ビューを使用した問合せに迅速に応答します。

パフォーマンスの向上 : 索引

ビットマップ索引はデータ・ウェアハウス環境において広く使用されています。データ・ウェアハウス環境では通常、大量のデータおよび非定型問合せがありますが、DML トランザクションが同時に発生することは稀です。索引が表内のデータのサイズの数倍である場合があるため、サイズの大きな表に従来のBツリーを完全に索引付けすると、ディスク領域に関してコストが非常に高くなる可能性があります。ビットマップ索引は通常、表内の索引付けされたデータのサイズのほんの一部にすぎません。このようなアプリケーションの場合、ビットマップ索引には次のような特性があります。

- 大規模な非定型問合せに対するレスポンス時間の削減
- 他の索引の方法と比較した場合の記憶域要件の減少
- 比較的CPUの数が少ないハードウェアまたはメモリー量が少ないハードウェア上での大幅なパフォーマンスの向上
- パラレルDMLおよびロード中の効率的なメンテナンス

最適ナリソースの使用

可能な場合にパラレルで操作を実行すると、システムでのリソースの使用を最大化できます。特定の時点でリソースによる制限がない場合、データベース操作をより速く実行できます。操作はCPUリソース、I/O容量、メモリーまたは（クラスタ内の）相互接続のトラフィックによって制限される場合があります。データベース操作のパフォーマンスを向上させるには、パフォーマンスの問題に焦点を当て、その問題を排除する必要があります（問題が他のリソースに移行する可能性を考慮し、排除する必要があります）。Oracle Databaseでは、使用可能なリソースの使用を最大化する機能を提供し、また不要なリソースの使用も回避します。

パラレル実行でのパフォーマンスの最適化

パラレル実行を使用すると、通常ディジョン・サポート・システム (DSS) およびデータ・ウェアハウスに関連付けられているサイズの大きなデータベース上で、データ集中型の操作のレスポンス時間を大幅に削減できます。また、特定のタイプのオンライン・トランザクション処理 (OLTP) およびハイブリッド・システム上で、パラレル実行の実装もできます。パラレル実行はパラレル化とも呼ばれます。パラレル化のタスクの概念を簡潔に説明すると、1つのプロセスで問合せに関するすべての処理を実行するのではなく、多くのプロセスが同時に各処理を実行します。たとえば、1年間ですべての四半期を1つのプロセスで処理せずに、四半期それぞれを4つのプロセスで処理します。これを使用するとパフォーマンスの大幅な向上が見込めます。パラレル実行では、次のプロセスを向上できます。

- サイズの大きな表のスキャン、結合またはパーティション化された索引のスキャンを要求する問合せ
- 大規模な索引の作成
- サイズの大きな表の作成 (マテリアライズド・ビューを含む)
- 大量データの挿入、更新、マージおよび操作の削除

また、パラレル実行を使用して、Oracle データベース内のオブジェクト型にアクセスできます。たとえば、パラレル実行により、ラージ・オブジェクト (LOB) にアクセスできます。

パラレル実行は、次のすべての特性を持つシステム上で有効です。

- 対称型マルチプロセッサ (SMP)、クラスタ、または大規模なパラレル・システム
- 十分な I/O 帯域幅
- 稼働中でない CPU または断続的に使用されている CPU (CPU の使用率が通常 30% 未満のシステムなど)
- ソート、ハッシュおよび I/O バッファなどの追加のメモリー集中処理をサポートする十分なメモリー

これらの特性を持たないシステムでは、パラレル実行を使用しても大幅にパフォーマンスが向上しない場合があります。実際に、大容量を使用しているシステムまたは I/O 帯域幅の少ないシステム上では、システムのパフォーマンスが低下する可能性があります。

パラレル実行の動作方法

パラレル実行では、SQL 文の実行タスクが複数の小さな単位に分割され、各ユニットは個別のプロセスで実行されます。また、受信するデータ (表、索引、パーティション) も **グラニュル** と呼ばれる単位に分類されます。問合せを並行して実行するユーザー・シャドウ・プロセスは、パラレル実行のコーディネータまたは問合せコーディネータとしての役割を持ちます。問合せコーディネータの役割は次のとおりです。

- 問合せの解析および並列度の判別
- 1セットまたは2セットのスレーブ (スレッドまたはプロセス) の割当て
- 問合せの制御およびパラレル問合せスレーブへの手順の送信
- パラレル問合せスレーブによるスキャンが必要な表または索引の判別
- ユーザーに対する最終的な出力の作成

並列度の設定

パラレル実行コーディネータでは、2つ以上のインスタンスのパラレル実行サーバーで SQL 文を処理する場合があります。単一の操作に関連付けられているパラレル実行サーバーの数は、並列度と呼ばれます。

単一の操作は、ORDER BY 操作または索引なしの列表に対して結合を実行する全表スキャンなどの SQL 文の一部です。

並列度 (DOP) を指定する方法は次のとおりです。

- PARALLEL ヒントを持つ文レベルでの指定
- ALTER SESSION FORCE PARALLEL 文を発行してセッション・レベルでの指定
- 表の定義内で表レベルでの指定
- 索引の定義内で索引レベルでの指定

例：並列度の設定

表で DOP を 4 に設定する場合を想定します。

並列度を設定する手順

次の文を発行します。

```
ALTER TABLE orders PARALLEL 4;
```

待機イベントについて

待機イベントはサーバー・プロセスによって増分した統計で、サーバー・プロセスは、イベントが完了し、処理の続行が可能となるまで待機する必要があることを示します。セッションの待機理由は様々で、大量の入力に対する待機、ディスクへの書込み操作などのサービスを完了するためのオペレーティング・システムに対する待機、またはロックやラッチを待機する時間が含まれます。

各セッションがリソースを待機する間、有効な動作は実行されません。待機イベントの多くが問題の原因となります。待機イベント・データにより、ラッチ競合、バッファ競合および I/O 競合などのパフォーマンスに影響を与える可能性のある問題の症状が明らかになります。

データ・ウェアハウスの バックアップおよびリカバリ

この項で説明する、データ・ウェアハウスのバックアップおよびリカバリに関する考慮事項には、次の内容が含まれます。

- [データ・ウェアハウスのバックアップおよびリカバリの処理方法](#)
- [バックアップおよびリカバリの計画とベスト・プラクティス](#)

データ・ウェアハウスのバックアップおよびリカバリの処理方法

バックアップおよびリカバリは、管理者にとって最も重要なタスクであり、データ・ウェアハウスについても同様のことが言えます。ただし、データ・ウェアハウスはデータベースのサイズが大きいため、管理者にとってはバックアップおよびリカバリに関する新たな課題が浮上することになります。

データ・ウェアハウスは、無数のリソースからデータが取得されるという点において特徴的で、最終的にデータベースに挿入される前に変換されますが、これは主にサイズが非常に大きいためです。大規模なデータ・ウェアハウスのリカバリの管理は大変なタスクで、従来の OLTP のバックアップ計画およびリカバリ計画では十分に要件を満たしません。

データ・ウェアハウスと OLTP システムの相違点は次のとおりです。

- 通常、データ・ウェアハウスのサイズははるかに大規模です。
- データ・ウェアハウスは、運用システムとは異なる可用性の要件を備えています。ビジネスにおける決断でも、データ・ウェアハウスからの情報に依存し、たとえばサービス・デスクが操作できないような状況では、事態はさらに悪化します。また、データ・ウェアハウスのサイズのため、一定水準のデータ・ウェアハウスの可用性を保証するには非常に高いコストがかかります。
- データ・ウェアハウスは、通常 ETL（抽出、変換およびロード）と呼ばれるより優れた制御のプロセスを介して移入されます。この結果、データ・ウェアハウス内の更新は正確に把握され、データ・ソースからの再生成が可能です。
- 単一のデータ・ウェアハウスには通常、多くの履歴データが格納され、大抵の場合は変更されるオブジェクトではありません。変更されないデータは、1 回のみバックアップする必要があります。

システム設計の一部としてバックアップ計画を計画し、バックアップする対象およびバックアップする頻度を考慮する必要があります。バックアップ設計で最も重要な変数は、バックアップまたはリカバリを実行する必要があるリソースの量とリカバリ時間目標（システム全体またはシステムの一部の使用不可能な状態の許容時間）です。

バックアップおよびリカバリ計画を計画する場合は、NOLOGGING 操作を考慮する必要があります。従来のリカバリは、バックアップをリストアしてアーカイブ・ログから変更を適用するものですが、NOLOGGING 操作には適用されません。さらに、NOLOGGING 操作により操作されたデータに依存する後続の操作は失敗します。バックアップおよびリカバリ計画を設計する場合は、NOLOGGING 操作を考慮する必要があります。

NOLOGGING 操作が行われている場合は、バックアップを作成することはありません。

次の 2 つの計画のうち 1 つを計画するか、または 2 つの計画の組合せを計画します。

- ETL 計画。リカバリできないトランザクションを含まないバックアップをリカバリし、バックアップと失敗の間で行われた ETL をリプレイします。
- 増分バックアップ計画。その他のリカバリできないトランザクションが行われた後、すぐにバックアップを実行します。変更済データ・ブロックに基づいた増分バックアップを有効にするトラッキング・ファイル機能が提供されます。Recovery Manager によりトラッキング・ファイル機能が利用されます。

バックアップおよびリカバリの計画とベスト・プラクティス

バックアップおよびリカバリ計画を考案することは、大変なタスクです。また、数百 GB の保護対象のデータを所有して失敗した場合にリカバリが必要である場合は、計画は非常に複雑になります。

次のベスト・プラクティスにより、ウェアハウスのバックアップおよびリカバリ計画の実装をサポートします。

- **ベスト・プラクティス A: ARCHIVELOG モードの使用**
- **ベスト・プラクティス B: Recovery Manager の使用**
- **ベスト・プラクティス C: 読取り専用の表領域の使用**
- **ベスト・プラクティス D: NOLOGGING 操作の計画**
- **ベスト・プラクティス E: すべての表領域の重要度が同等でない場合**

ベスト・プラクティス A: ARCHIVELOG モードの使用

アーカイブ REDO ログはデータベースへの変更の記録を構成するため、データがまったく失われない場合のリカバリに非常に重要です。Oracle Database は、次の 2 つのモードのいずれかで実行されます。

- **ARCHIVELOG:** 再利用される前に満杯になったオンラインの REDO ログ・ファイルがアーカイブされます。
- **NOARCHIVELOG:** 再利用される前に満杯になったオンラインの REDO ログ・ファイルがアーカイブされません。

ARCHIVELOG モードでデータベースを実行すると、次の利点があります。

- データベースはインスタンス障害およびメディア障害の両方から完全にリカバリされます。
- データベースを開いて使用可能な状態で、ユーザーはバックアップを実行できます。
- Oracle Database により、多重化されたアーカイブ・ログによるそのアーカイブ・ログ上の単一点障害の回避がサポートされます。
- ユーザーは、表領域の Point-in-Time リカバリ (TSPITR) の実行機能のようなより多くのリカバリ・オプションを使用できます。
- アーカイブ REDO ログは、プライマリ・データベースの完全なコピーであるフィジカル・スタンバイ・データベースに転送および適用できます。
- データベースはインスタンス障害およびメディア障害の両方から完全にリカバリされます。

NOARCHIVELOG モードでデータベースを実行すると、次の結果が得られます。

- 停止後、データベースが完全に閉じられた状態で、ユーザーはデータベースのバックアップのみ実行できます。
- 最後のバックアップ後はすべてのトランザクションの損失の原因となるため、通常、メディア・リカバリ・オプションのみがデータベース全体をリストアします。

停止時間の許容について

データベースが開いている間、または閉じている間でも Oracle データベースのバックアップは作成できます。データベースの計画停止時間は業務を混乱させ、特に複数のタイムゾーンにいるユーザーを最長 24 時間サポートする世界規模の企業では顕著です。このような場合は、データベースへの割込みを最小化するようにバックアップ計画を設計することが重要です。

業種によっては、停止時間に余裕がある企業もあります。ビジネス計画全体で、停止時間は最小限またはまったくない状態が必要とされる場合は、バックアップ計画にオンライン・バックアップを実装する必要があります。データベースは、バックアップのために停止する必要がありません。オンライン・バックアップでは、データベースを ARCHIVELOG モードにする必要があります。

基本的に、ARCHIVELOG モードを使用しない理由などありません。どんなデータ・ウェアハウス（つまりすべての重要なデータベース）にとっても ARCHIVELOG モードの使用が必要となります。特に、データ・ウェアハウスのサイズ（およびデータ・ウェアハウスをバックアップする時間）を考慮すると、データ・ウェアハウスが NOARCHIVELOG モードを使用する場合に必要なデータ・ウェアハウスのオフライン・バックアップの作成は、通常、実行可能ではありません。

もちろん、サイズの大きいデータ・ウェアハウスは大規模なデータの変更が発生する可能性があります。大規模なデータの変更があると大量のログ・ファイルが生成されます。大量のアーカイブ・ログ・ファイルの管理に適応するために、Recovery Manager により、アーカイブ済のログ・ファイルを圧縮するオプションが提供されます。これにより、リカバリのためのより迅速なアクセシビリティを実現するために、より多くのアーカイブ・ログをディスクに保存できます。

要約すると、ベスト・プラクティスは、データベースをアーカイブ・ログ・モードにしてオンライン・バックアップのオプションと Point-in-Time リカバリのオプションを使用可能にすることです。

ベスト・プラクティス B: Recovery Manager の使用

Recovery Manager を採用する理由は多数あります。Recovery Manager をバックアップおよびリカバリ計画に統合する理由のいくつかは、次の項目が提供されるからです。

- 拡張性の高いレポート
- 増分バックアップ
- 停止時間のないバックアップ
- バックアップおよびリストアの検証
- バックアップおよびリストアの最適化
- メディア・マネージャとの簡単な統合
- ブロック・メディア・リカバリ
- アーカイブ・ログの検証および管理
- 破損ブロックの検出

ベスト・プラクティス C: 読取り専用の表領域の使用

データ・ウェアハウスが直面する最大の問題の 1 つは、一般的にデータ・ウェアハウスが大規模であるということです。強力なバックアップのハードウェアを使用しても、バックアップに数時間かかります。したがって、バックアップするデータの量の最小化を考慮することは、バックアップのパフォーマンスの向上のために重要なことの 1 つです。読取り専用の表領域は、データ・ウェアハウス内にバックアップするデータの量を削減するための最も簡単なメカニズムです。

読取り専用の表領域の利点は、データを 1 回のみバックアップすればよいという点です。つまり、データ・ウェアハウスに 5 年間の履歴データが含まれる場合は、データの最初の 4 年間は読取り専用にできます。理論上は、データベースの通常のバックアップでデータの 20% のみがバックアップされます。これにより、データ・ウェアハウスのバックアップに必要な時間を大幅に削減できます。

ほとんどのデータ・ウェアハウスでは、時間でレンジ・パーティション化された表にデータが格納されます。通常のデータ・ウェアハウスでは、一般的に 30 日間から 1 年間までの範囲の期間で、データはアクティブです。この期間中、履歴データを更新でき、また変更できます（たとえば、小売業者が購入日後の 30 日間までは返品を受け入れる場合は、売上データの記録をこの期間中に変更できます）。ただし、一度データが特定の日付に到達すると、多くの場合データは静的であると認識されます。

パーティション化を利用すると、ユーザーは読取り専用のデータの静的な部分を作成できます。**Recovery Manager**により、読取り専用のパーティションまたは表ではなく、読取り専用の表領域がサポートされます。読取り専用の表領域を利用してバックアップ・ウィンドウを削減するには、読取り専用の表領域の定数データ・パーティションの格納計画を考案する必要があります。ローリング・ウィンドウを実装する2つの計画は次のとおりです。

- データ全体が静的になったら、読取り / 書込み表領域から読取り専用表領域にパーティションを移動する定期的なプロセスを実装します。

この場合のベスト・プラクティスは、データベースを ARCHIVELOG モードにしてオンライン・バックアップおよび Point-in-Time リカバリのオプションを使用可能にすることです。

- 各表領域に少数のパーティションを持つ一連の表領域を作成し、その表領域内のデータが年数を重ねるにつれて定期的に1つの表領域を読取り / 書込み両用から読取り専用に変更します。

1つ考慮する必要があるのは、データのバックアップはリカバリ・プロセスの半分にすぎないということです。テープのシステムを構成する場合、4時間でデータ・ウェアハウスの読取り / 書込み両用の部分をバックアップでき、データベースの80%が読取り専用であるときに完全なリカバリが必要な場合は、必然的にテープのシステムではデータベースをリカバリするために20時間かかります。

要約すると、ベスト・プラクティスは、静的表および静的パーティションを読取り専用の表領域に配置することです。読取り専用の表領域を1回のみバックアップする必要があります。

ベスト・プラクティス D: NOLOGGING 操作の計画

一般的に、データ・ウェアハウスの最優先事項はパフォーマンスです。データ・ウェアハウスにより、オンラインのユーザーに快適な問合せのパフォーマンスが提供される必要があるだけでなく、大量のデータを最短時間内にロードするために、ETL プロセス中、データ・ウェアハウスは効率的である必要があります。

データ・ウェアハウスにより利用される共通の最適化の1つは、NOLOGGING モードを使用して大量データ操作を実行することです。NOLOGGING モードをサポートするデータベース操作は、直接パス・ロードおよび直接パス・インサート、索引の作成および表の作成です。1つの操作が NOLOGGING モードで実行される場合は、データは REDO ログに書き込まれません（より正確に言うと、サイズの小さいメタデータのセットのみ REDO ログに書き込まれます）。このモードは、データ・ウェアハウスの広範囲で使用され、大量データ操作のパフォーマンスを50%まで改善できます。

ただし、リカバリをサポートするために必要なデータがログ・ファイルに書き込まれることはないため、そのかわりに、従来のバックアップ・メカニズムを使用して NOLOGGING 操作をリカバリできません。さらに、NOLOGGING 操作が発生するデータに対する後続の操作は、たとえこれらの操作が NOLOGGING 操作を使用していなかったとしてもリカバリされません。NOLOGGING 操作により提供されるパフォーマンスの向上のため、一般的に ETL プロセス中は、データ・ウェアハウスによる NOLOGGING モードの使用をお勧めします。

バックアップおよびリカバリ計画を考案する場合は、NOLOGGING 操作の存在を考慮する必要があります。データベースが NOLOGGING 操作に依存する場合、ログ・ファイルにより NOLOGGING 操作はリカバリできないため、従来のリカバリ計画（最新のテープのバックアップからリカバリし、アーカイブ・ログ・ファイルを適用する）は適用されません。

覚えておく必要がある第1の原則は、NOLOGGING 操作が発生しているときにバックアップを作成しないことです。現在この原則を規定していないため、DBA は NOLOGGING 操作がバックアップ操作に重複しないようなバックアップ・ジョブおよび ETL ジョブをスケジュールする必要があります。

NOLOGGING 操作が存在するバックアップおよびリカバリには、2つのアプローチがあります。ETL バックアップまたは増分バックアップです。データ・ウェアハウス内で NOLOGGING 操作を使用していない場合は、次のオプションのいずれも選択する必要はありません。アーカイブ・ログを使用してデータ・ウェアハウスをリカバリできます。ただし次のオプションは、リカバリの場合のアーカイブ・ログ・ベースのアプローチにおいていくつかのパフォーマンスの面で優れています。

- 抽出、変換およびロード
- 増分バックアップ

抽出、変換およびロード

ETL プロセスでは、データをデータにロード（再ロード）するためにいくつかの Oracle の機能またはツール、およびメソッドの組合せを使用します。これらの機能またはツールの構成要素は次のとおりです。

- トランスポータブル表領域。Oracle トランスポータブル表領域機能により、ユーザーは Oracle データベース間で表領域を迅速に移動させることができます。この方法は、データベース間で大量データを移動するのに最も効率的です。Oracle Database により、プラットフォーム間で表領域を転送できる機能が提供されます。ソース・プラットフォームとターゲット・プラットフォームが異なるエンディアンネスである場合、Recovery Manager によりターゲット・フォーマットに転送されている表領域が変換されます。
- SQL*Loader。SQL*Loader により、外部プラットフォーム・ファイルから1つの Oracle データベース上の複数の表にデータがロードされます。SQL*Loader には、データファイル内のデータの形式にほとんど制限のない強力なデータ解析エンジンがあります。
- データ・ポンプ（エクスポート / インポート）。Oracle Database により、Oracle Data Pump テクノロジーが提供され、これにより1つのデータベースから他のデータベースにデータおよびメタデータを高速で移動することが可能です。このテクノロジーは、データ移動ユーティリティであるデータ・ポンプ・エクスポートおよびデータ・ポンプ・インポートの基本となります。
- 外部表。外部表機能とは、既存の SQL*Loader 機能を補完するものです。この機能により、データベース内の表にあるデータと同様に、外部ソースのデータにアクセスできるようになります。

ETL 計画および NOLOGGING 操作

アプローチの1つは、定期的にデータベースのバックアップをして、その週全体の ETL プロセスを再作成するために必要なデータファイルを保存することです。リカバリが必要な場合、データ・ウェアハウスは最新のバックアップからリカバリされます。次に、従来のリカバリのシナリオで行われていたようにアーカイブ REDO ログを適用してロールフォワードを行うかわりに、データ・ウェアハウスは ETL プロセスをリプレイすることでロールフォワードを行います。この事例では、ETL プロセスは簡単にリプレイできることを前提とし、通常、各 ETL プロセスの抽出ファイルのセットの保存も含まれます（たとえば不正なデータ送信の修復の必要があるかどうかを識別できるように、データ・ウェアハウスの多くはすでにベスト・プラクティスとしてこれを行っています）。

このアプローチのサンプル実装は、毎週末データ・ウェアハウスのバックアップを作成し、毎晩の ETL プロセスをサポートするために必要なファイルを保存することです。したがって、データベースをリカバリするためには最大7日間分の ETL プロセスが再適用される必要があります。データ・ウェアハウス管理者は、テープからのリカバリ・スピードおよび前の ETL 実行からのパフォーマンス・データに基づいて簡単に時間の長さを計画してデータ・ウェアハウスをリカバリできます。

基本的にデータ・ウェアハウス管理者は、NOLOGGING 操作を介して ETL プロセスでのより高いパフォーマンスを獲得しており、その際は多少複雑であり自動化されていないリカバリ・プロセスを通過する必要があります。データ・ウェアハウス管理者の多くは、これを価値のある代償だと理解しています。

このアプローチの弱点は、データ・ウェアハウスで発生した関連する変更のすべてを追跡する負荷がデータ・ウェアハウス管理者にかかるということです。このアプローチでは、ETL プロセスの範囲に入らない変更は取得されません。たとえば、いくつかのデータ・ウェアハウスでは、エンドユーザーが所有する表およびデータ構造を作成します。それらの変更はリカバリ時に失われます。この制限はエンドユーザーに伝達される必要があります。かわりに、管理者がエンドユーザーに個別の表領域にあるすべてのプライベート・データベース・オブジェクトを作成するように指示し、リカバリの間、DBA は ETL プロセスをリプレイするアプローチを使用して残りのデータベースをリカバリする一方で従来のリカバリを使用してこの表領域をリカバリすることもできます。

要約すると、ベスト・プラクティスは、リカバリできない (NOLOGGING) トランザクションを含まないバックアップをリストアすることです。次に ETL プロセスをリプレイしてデータを再ロードします。

ブロック変更トラッキング・ファイルのサイジング

ブロック変更トラッキング・ファイルのサイズは次の項目に比例します。

- データベース・サイズ (バイト)。ブロック変更トラッキング・ファイルには、データベース内のデータファイル・ブロックを表すデータが含まれます。データはおおよそデータベースのサイズ全体の 250000 分の 1 です。
- 有効なスレッドの数。すべての Real Application Clusters (RAC) インスタンスは同一のブロック変更トラッキング・ファイルにアクセスできますが、インスタンスにより、ロックまたはノード間のブロック・スワッピングなしでトラッキング・ファイルの異なる領域が更新されます。個別のインスタンスに対してではなく、データベース全体に対してブロック変更トラッキング・ファイルを有効化できます。
- 変更済ブロックのメタデータ。ブロック変更トラッキング・ファイルは最後のバックアップ以来の変更に加えて、前のバックアップ間の変更に記録を保持しています。トラッキング・ファイルは、最大で 8 個のバックアップの変更履歴を保持します。トラッキング・ファイルに 8 個のバックアップの変更履歴が含まれる場合は、Oracle データベースにより最も古い変更履歴情報が上書きされます。

たとえばスレッドを 1 つのみ所有し、Recovery Manager リポジトリに 8 個のバックアップを保持する 500GB のデータベースの場合は、20MB のブロック変更トラッキング・ファイルが必要です。

$((\text{Threads} * 2) + \text{number of old backups}) * (\text{database size in bytes})$

----- = 20MB

250000

増分バックアップ

NOLOGGING 操作が存在する、より自動化されたバックアップおよびリカバリ計画では、Recovery Manager の増分バックアップ機能を利用します。Recovery Manager が最初にリリースされて以来、増分バックアップは Recovery Manager の一部を構成しています。増分バックアップには、前のバックアップ後に変更されたブロックのみをバックアップする機能があります。データファイルの増分バックアップでは、データファイル内のすべての使用済ブロックのバックアップを必要とするのではなく、基本的にブロックごとのデータの変更が取得されます。結果として取得されるバックアップのセットは、データファイル内のすべてのブロックが変更されないのであれば、通常すべてのデータファイルのバックアップより小さくて効率的です。

Oracle Database には、変更トラッキング・ファイル機能を実装して増分を高速化する機能があります。ブロック変更トラッキングを有効にする場合は、Oracle Database により、すべてのデータベースの変更の物理的な場所が追跡されます。Recovery Manager により自動的に変更トラッキング・ファイルが使用され、増分バックアップ中に読み込む必要のあるブロックが決定されて、バックアップするブロックに直接アクセスします。

増分アプローチ

このアプローチを使用する通常のバックアップおよびリカバリ計画は、毎週末データ・ウェアハウスをバックアップし、ETL プロセスが完了した後は毎晩データ・ウェアハウスの増分バックアップを取得します。従来のバックアップと同様に、増分バックアップと NOLOGGING 操作を同時に実行しないように注意してください。データ・ウェアハウスをリカバリするために、データベースのバックアップはリストアされ、毎晩の増分バックアップは再適用されます。アーカイブ・ログでは NOLOGGING 操作は取得されませんが、NOLOGGING 操作からのデータは増分バックアップに存在します。さらに、前のアプローチとは異なり、このバックアップおよびリカバリ計画は、Recovery Manager を使用して完全に管理できます。

ETL をリプレイするアプローチおよび増分バックアップのアプローチの両方とも、多くの NOLOGGING 操作で構成されているワークロードであるデータベースの効率的で安全なバックアップおよびリカバリのソリューションとしてお勧めします。バックアップおよびリカバリ計画において NOLOGGING 操作を考慮することが最も重要です。

要約すると、ベスト・プラクティスは、NOLOGGING 操作のためオブジェクトをリカバリできない状態しておくダイレクト・ロードの後、ブロック変更トラッキング機能を実装して増分バックアップを作成することです。

ベスト・プラクティス E: すべての表領域の重要度が同等でない場合

データベース内の各表領域を同様に処理することが最も単純なバックアップおよびリカバリのシナリオですが、オラクル社は、DBA が必要に応じて各表領域に対してバックアップおよびリカバリのシナリオを考案できるような柔軟性を提供しています。

バックアップおよびリカバリという観点で見ると、データ・ウェアハウス内のすべての表領域が同等に重要だということはありません。DBA はこの情報を利用して、より効率的なバックアップおよびリカバリ計画を必要に応じて考案できます。バックアップおよびリカバリの基本粒度は表領域なので、異なる表領域には異なるバックアップおよびリカバリ計画が存在する可能性があります。最も基本的なレベルでは、一時表領域がバックアップされることはありません (Recovery Manager により規定されるルール)。

さらに、いくつかのデータ・ウェアハウスでは、明示的な一時表領域ではないにもかかわらず、エンドユーザーが一時表領域と増分の結果を格納するスクラッチ領域専用の表領域であるために、基本的に一時表領域として機能している表領域があります。ビジネス要件によっては、これらの表領域をバックアップおよびリストアする必要はありません。かわりに、これらの表領域が失われた場合は、エンドユーザーは所有するデータ・オブジェクトを再作成します。

データ・ウェアハウスの多くは、その他のデータより重要な一部データを持っています。たとえば、データ・ウェアハウス内の売上データは非常に重要で、リカバリの段階では、このデータはできるだけ早くオンラインにする必要があります。しかし、同一のデータ・ウェアハウス内では、企業の Web サイトからのクリック・ストリーム・データを格納している表の重要度はかなり低くなります。ビジネスを継続する上で、このデータが数日間オフラインになること、またはデータベース・ファイル損失時にクリック・ストリーム・データが数日間失われることは許容範囲内です。このシナリオでは、売上データを含む表領域は頻繁にバックアップする必要があり、一方でクリック・ストリーム・データを含む表領域は毎週 1 回または 2 回バックアップする必要があります。

データ・ウェアハウスの操作の最適化

この項では、データ・ウェアハウスのセキュリティの準備の考慮事項について説明します。次のトピックが含まれます。

- データ・ウェアハウスのセキュリティについて
- データ・ウェアハウスのセキュリティのロールおよび権限の使用
- データ・ウェアハウスの仮想プライベート・データベースの使用
- Oracle Label Security の概要
- データ・ウェアハウスのファイングレイン監査の概要
- データ・ウェアハウスの透過的データ暗号化の概要

データ・ウェアハウスのセキュリティについて

データ・ウェアハウジングは、セキュリティに対して固有の課題を投げかけます。業務データ・ウェアハウスは、巨大なシステムであることが多く、多様なセキュリティ上のニーズに応えるとともに、数多くのユーザーに対応しています。また、データ・ウェアハウスは、柔軟性のある、強力なセキュリティ・インフラストラクチャを必要としますが、そのセキュリティ機能は、厳密なパフォーマンスとスケーラビリティなど条件付きの環境で動作する必要があります。

データ・ウェアハウスのセキュリティの必要性について

よく知られたセキュリティに関する多くの基本要件は、他のシステムと同等にデータ・ウェアハウスにも適用されます。アプリケーションでは、不正ユーザーによるデータへのアクセスまたは修正を防止する必要があります。また、アプリケーションおよび基礎データは、簡単にハッカーが盗用できるものであってはなりません。データは、正当なユーザーが必要に応じてデータを使用できるようにする必要があります。さらに、ユーザーが実行するアクティビティのレコードをシステムで保持する必要があります。

ウェアハウスには複数のソースから統合されたデータが含まれるため、情報を盗もうとする側から見ると、企業で最も利益を得られるターゲットとなる可能性があるため、セキュリティの要件はデータ・ウェアハウスでは特に重要です。さらに、強固なセキュリティ・インフラストラクチャにより、有効性が大幅に改善され、データ・ウェアハウス環境のコストが削減されます。

データ・ウェアハウスのセキュリティの一般的な顧客シナリオには次のような事例があります。

- ある企業は、多くの部署や子会社で広く使用される企業データ・ウェアハウスを管理しています。その企業では、各部署の従業員には自分の部署に関連するデータのみを表示し、本社の従業員には全体像が表示されるセキュリティ・インフラストラクチャを必要としています。
- ある企業のデータ・ウェアハウスには、個人情報が格納されています。プライバシーに関連する法では、そのような個人情報の使用が規制される場合があります。データ・ウェアハウスでは、このような法律に沿うようにデータを管理する必要があります。
- ある企業は、データ・ウェアハウスのデータをクライアントに販売します。クライアントには、購入または申し込んだデータのみが表示されるようにし、他のクライアントのデータの表示は許可しません。

データ・ウェアハウスのセキュリティのロールおよび権限の使用

システム権限、オブジェクト権限およびロールは、基本的なデータベース・セキュリティを提供します。また、ユーザーによるデータへのアクセスを制御し、ユーザーが実行可能な SQL 文の種類を制限するように設計されています。ロールは異なるレベルのデータベース・アクセスを作成するために使用できる権限のグループです。たとえば、表およびプログラムを作成できるように、アプリケーション開発者に対してロールを作成できます。

必要な権限を所有する場合にかぎり、他のユーザーに権限およびロールを付与できます。ロールおよび権限の付与は、管理者レベルで開始します。データベースの作成で管理ユーザー SYS により、すべてのシステム権限および事前定義された Oracle ロールの作成と付与が行われます。ユーザー SYS は、他のユーザーに権限およびロールを付与でき、これらを付与されたユーザーに対して、他のユーザーに特定の権限を付与する権限を付与することもできます。明示的に付与された権限がない場合、ユーザーはデータベース内のどの情報にもアクセスできません。

ロールおよび権限はデータのセキュリティを規定し、データ・ウェアハウスにとってこれらのロールおよび権限の使用は不可欠です。これは多くのアプリケーションおよびツールを介してユーザーがデータにアクセスするためです。

データ・ウェアハウスの仮想プライベート・データベースの使用

仮想プライベート・データベース (VPD) によって、表、ビューまたはシノニムで直接、粒度の細かいレベルにセキュリティを規定できます。これは、セキュリティ・ポリシーが表、ビューまたはシノニムに直接連結されており、ユーザーがデータにアクセスする際にセキュリティをバイパスする方法がない場合に自動的に適用されるためです。述語を使用する動的に付与されている SQL 文によって、VPD はデータへのアクセスを低いレベルに制限し、セキュリティ・ポリシーをデータベース・オブジェクトに関連付けます。また、完全なデータ分離の保証があれば、複数のユーザーが単一のデータベース・サーバーに格納された重要なデータへ直接セキュア・アクセスできます。VPD を使用すると、銀行顧客には顧客自身のアカウント履歴のみを見せ、複数の企業（競争企業同士）のデータを扱う企業の場合は、同じデータ・ウェアハウスからデータを扱っている企業にそれぞれのデータのみを見せることができます。さらに、VPD はセキュリティ関連列への制御アクセスを提供するので、従業員は自身の給与のみを確認できるようになります。

VPD は透過アプリケーションです。セキュリティはデータベース層で規定され、データベース内でのデータ・アクセスを制御するためのアプリケーション固有のロジックを考慮します。標準のアプリケーションおよびカスタム・ビルドのアプリケーションは、アプリケーション・コードの単一行の変更を行うことなく、ファイングレイン・アクセス制御を活用できます。

Enterprise 内では、VPD によりアプリケーションの配布にかかる所有コストが削減され、データにアクセスするすべてのアプリケーション内ではなくウェアハウス内にセキュリティが構築されます。セキュリティは、ユーザーによるデータへのアクセスに関係なくデータベースによって規定されているためにより強固になりましたが、非定型の間合せのツールまたは新しいレポート・ライターを介してデータへアクセスするユーザーはセキュリティを経由できません。多くのエンドユーザー・ツールと同様、様々なアプリケーションをサポートする Enterprise Data Warehouse において、仮想プライベート・データベースは主要なテクノロジーです。

仮想プライベート・データベースの動作方法

仮想プライベート・データベースは表、ビューまたはシノニムを使用するセキュリティ・ポリシーとの関連付けにより有効になります。管理者は PL/SQL DBMS_RLS パッケージを使用して、ポリシー・ファンクションとデータベース・オブジェクトをバインドします。連結したセキュリティ・ポリシーを使用するオブジェクトへの直接アクセスまたは間接アクセスにより、データベースはポリシーを実装したファンクションを参照します。ポリシー・ファンクションは、データベースがユーザーの SQL 文に追加する述語 (WHERE 句) を戻します。そのため、ユーザーによるデータへのアクセスの修正は透過的で動的です。

アプリケーション・コンテキストによって、アクセス条件は組織、サブスクライバ番号、アカウント番号または地位など、データベース管理者が重要と考える事実上すべての属性に基づくことができます。たとえば、ユーザーが顧客、営業担当またはマーケティング・アナリストかどうか関係なく、売上データのウェアハウスは顧客番号に基づいたアクセスを規定できます。この場合、顧客は Web 上で注文履歴を確認でき (顧客自身の注文のみ)、営業担当は複数の注文を見ることができます (自身の企業の顧客の注文のみ)。また、アナリストの場合、過去の 2 四半期におけるすべての売上を分析できます。

アプリケーション・コンテキストは、特定のオブジェクトでファイングレイン・アクセス制御ポリシーに適用する可能性のあるセキュアなデータ・キャッシュとして機能します。ユーザーがデータベースにログインすると、Oracle Database はユーザー・セッションのキャッシュ情報にアプリケーション・コンテキストを設定します。アプリケーション・コンテキストの情報は、特定のアプリケーションに関連した情報に基づいて開発者によって定義されます。たとえば、地域別売上データを問い合わせるレポート・アプリケーションは、ユーザーの地位および部門に関するアクセス制御に基づくことができます。この場合、アプリケーションはユーザーのログイン時に各ユーザーのアプリケーション・コンテキストを最初に設定し、各ユーザーの地位および部門の従業員と部署に関する表から問い合わせられたデータを使用してコンテキストを移入できます。地域別売上表で VPD ポリシーを実装しているパッケージは、ユーザーの地位および部門の各間合せを移入するためにこのアプリケーション・コンテキストを参照します。アプリケーション・コンテキストはパフォーマンスの劣化を引き起こす可能性のある実行中の副間合せを不要とみなします。

Oracle Label Security の概要

Oracle Database のセキュリティ・オプションである Oracle Label Security はラベルベースのアクセス制御を規定するために仮想プライベート・データベース (VPD) を拡張します。Oracle Label Security は完全な VPD 対応アプリケーションで、ラベル付きデータ管理によって VPD を拡大します。セキュアなデータ・ウェアハウスの配布を簡略化し、基本的なセキュリティを提供します。

ラベルベースのアクセス制御を使用すると、表の行に機密性ラベルが割り当てられ、これらのラベルに基づいてデータへのアクセスが制御されて、適切なセキュリティ・ラベルがマークされたデータを確認できます。たとえば、組織が企業の機密情報とパートナに関する情報を区別する場合です。または、特定の主要なパートナと共有できる機密情報、および会計部門や営業部門などの特定の企業内部のグループのみがアクセス可能な機密情報などがある場合です。ラベル・データの管理機能は、適切なデータ・アクセス・レベルで適切な人に適切な情報を提供するために組織にとって非常に有利です。

Oracle Label Security の動作方法

Oracle Label Security は、ラベルのコレクション、ユーザーの認可およびセキュリティ強制オプションであるポリシーを使用します。一度作成されると、ポリシーはすべてのアプリケーション・スキーマまたは特定のアプリケーション表に適用できます。Oracle Label Security は単一のデータ・ウェアハウス内で複数のポリシー定義をサポートします。ラベル定義、ユーザーの認可および強制オプションはポリシーごとに定義されます。たとえば、マーケティングのポリシーにはマーケティングのみのラベルやマネージャおよび上級副社長などのラベルが含まれます。

Oracle Label Security は行に含まれるラベル、各データベース・セッションと関連付けられたラベル、およびセッションに割り当てられた Oracle Label Security 権限に基づくデータベース表の列へのアクセスを調整します。これにより、ユーザーに標準のデータベース・システムおよびオブジェクト権限が付与された後にアプリケーション表でアクセス調整が行われます。たとえば、ユーザーに表に対する SELECT 権限があるとします。ユーザーが表で SELECT 文を実行すると、Oracle Label Security は選択した行を評価し、ユーザーに割り当てられた権限およびアクセス・ラベルに基づいて選択した行にアクセスできるかどうかを決定します。また、Oracle Label Security は UPDATE 文、DELETE 文および INSERT 文に対してセキュリティ・チェックも行います。ラベルは表およびマテリアライズド・ビューに適用できます。マテリアライズド・ビューによりパフォーマンスが増加し、ラベルによりセキュリティが増加するために、データ・ウェアハウス環境で目的の柔軟性、スピードおよびスケーラビリティを実現します。

データ・ウェアハウスによりラベルを利用する方法

Oracle Label Security を使用すると、複数のソースから利便性、管理性および集中管理によって、1 つの大きなシステムに情報を統合できます。このセキュリティ・オプションはアプリケーション自体で行われるので、PL/SQL プログラミングを行う必要がありません。これにより、情報の統合およびデータ自体でのセキュリティの制限によるリスクの最小化が可能になり、行レベルでのデータへのアクセス制御によって精度の高いアクセス・セキュリティが提供されます。

データ・ウェアハウスのファイングレイン監査の概要

ファイングレイン監査により、コンテンツに基づいたデータ・アクセスの監視ができます。監査するレコードの列および条件を指定でき、指定列を使用した接続で使用した DML 文の特定のタイプに対する監査の制限を条件に含むことができます。また、監査イベントの発生時におけるルーチン名を指定できます。このルーチンは管理者に通知またはアラートするか、エラーおよび異常を処理できます。たとえば、中央税務当局で使用すると、納税申告へのアクセスを追跡し従業員による詮索を防止できます。特定の表で SELECT 文を発行した特定のユーザーを把握するだけでは十分ではなく、日常業務に必要ではない情報にアクセスしようとする際のさらに細かいレベルでの監査が強固なセキュリティに必要です。この場合、列または行の SELECT 文には仕事に無関係な情報が含まれます。ファイングレイン監査はこの機能を提供します。

ファイングレイン監査は DBMS_FGA パッケージまたはデータベース・トリガーを使用してユーザー・アプリケーションで実装できます。

データ・ウェアハウスの透過的データ暗号化の概要

透過的データ暗号化によって、オペレーティング・システムのファイルに格納されるデータとして、データベース列にある機密データを暗号化できます。データベース外部のセキュリティ・モジュールで暗号キーの管理および記憶領域をセキュアに保ちます。既存のアプリケーションに暗号化したルーチンを埋め込む必要がなく、コストおよび暗号化の複雑さを大幅に削減しました。いくつかの単純なコマンドを使用することで、機密のアプリケーション・データは暗号化されます。

多くの暗号化ソリューションは、アプリケーション・コード内の暗号化ファンクションに対する特定のコールを必要とします。これには高いコストが必要で、アプリケーションの広範囲にわたる理解、およびソフトウェアの書込みと保持の技術が要求されます。一般的に、ほとんどの組織には既存のアプリケーションを修正して暗号化ルーチンに対するコールを作成する時間や専門家が不足しています。透過的データ暗号化は、Oracle Database に埋め込まれた暗号化により暗号化問題を処理します。これは SQL*Loader ではなく直接パス・ロードを使用して行われることに注意してください。

SQL を介して実行されるアプリケーション・ロジックは変更することなく実行を継続します。言い換えれば、アプリケーションはアプリケーション表にデータを挿入するために同じ構文を使用でき、Oracle Database はディスクへ情報を書き込む前にデータを自動的に暗号化します。後続の SELECT オプションには透過的に復号化されたデータが含まれ、アプリケーションは通常の作業を継続します。

索引

A

ARCHIVELOG モード
 ベスト・プラクティス, 13-3
ASM, 2-4

C

CPU, 2-3
CUBE 関数, 9-4

D

dd ユーティリティ, 2-4

E

ETL
 ランタイム・パフォーマンスの改善, 6-1

G

GROUPING SETS 関数, 9-7
GROUPING 関数, 9-5

I

I/O, 2-3, 11-3
 監視, 11-3

N

NOLOGGING
 ベスト・プラクティス, 13-5

O

Oracle Label Security, 14-4
 データ・ウェアハウス, 14-3
Orion ユーティリティ, 2-5

R

RANK 関数, 9-8
Recovery Manager
 ベスト・プラクティス, 13-4
ROLLUP 関数, 9-3

S

SQL アクセス・アドバイザー, 12-5
 索引の最適化, 11-5
 パーティション化の評価, 12-7
SQL チューニング・アドバイザー, 12-5

V

VPD
 動作, 14-3

W

WITH 句, 9-14

い

一般的なタスク
 データ・ウェアハウス, 1-3
移動平均の計算, 9-12

う

ウィザード
 「外部表の作成」ウィザード, 4-3

え

影響分析ダイアグラム, 8-3, 8-6
演算子
 接続, 5-6
 編集, 5-6
 リポジトリ・オブジェクトとの同期化, 5-11
演算子, マッピング, 5-2
 接続, 5-6
 タイプ, 5-4
 追加, 5-4
 編集, 5-6
演算子の属性, 5-7
演算子のプロパティ
 設定, 5-11

お

オペティマイザ統計, 12-2

か

階層

説明, 4-6

外部表

作成用ウィザード, 4-3

新規定義の作成, 4-3

定義, 4-3

「フラット・ファイル」も参照

「外部表の作成」ウィザード, 4-3

仮想プライベート・データベース

データ・ウェアハウス, 14-3

き

記憶域

最適化, 11-6

キューブ, 4-11

ディメンション, 4-11

メジャー, 4-11

リレーショナル実装, 4-12

例, 4-12

行間計算

実行, 9-11

く

クエリー・リライト, 2-8, 2-9, 4-4, 10-5, 12-8

グループ

LIA ダイアグラム, 8-5

接続, 5-7

プロパティの設定, 5-11

グループ・プロパティ, 5-11

け

系統ダイアグラム, 8-3, 8-6

権限

データ・ウェアハウスのセキュリティ, 14-2

さ

索引, 12-8

SQL アクセス・アドバイザーを使用した最適化, 11-5

最適化, 11-5

作成

マッピング, 5-1

し

実行計画, 12-2

分析, 12-2

実装

スター・スキーマ, 4-8

リレーショナル・キューブ, 4-12

自動ストレージ管理 (ASM), 2-4

自動バインド・ルール, 4-9

初期化パラメータ

設定, 2-8

す

スケジューリング

説明, 7-4

スケジュール

作成, 7-5

使用, 7-5

定義, 7-5

ストレージ管理

手動, 2-4

スパース性, 9-13

せ

セキュリティ

データ・ウェアハウス, 14-1

接続

演算子, 5-6

グループ, 5-7

属性, 5-7

設定

マッピング・プロパティ, 5-9

そ

相対的な寄与率の計算, 9-9

属性

接続, 5-7

プロパティの設定, 5-11

レベル属性, 4-6

属性のプロパティ, 設定, 5-11

た

ダイアグラム

影響分析, 8-3, 8-6

系統, 8-3, 8-6

待機イベント, 12-10

ち

稠密化, 9-13

つ

追加

マッピング演算子, 5-4

ツール

データ・ウェアハウス, 1-5

て

定義

スケジュール, 7-5

マッピング, 5-1

ディスク

要件, 2-3

ディメンション

階層, 4-6

実装, 4-8

スター・スキーマ実装, 4-8

ディメンション・ロール, 4-6

例, 4-7

- レベル属性, 4-6
- レベルの関係, 4-7
- データ圧縮, 11-6
- データ・ウェアハウス
 - 一般的なタスク, 1-3
 - 主な特性, 1-3
 - 設定, 2-7
 - セットアップ, 2-2
 - ツール, 1-5
- データ・ウェアハウスのセキュリティ, 14-1
- データのオフロード
 - ローリング・ウィンドウ, 10-5

と

- 透過的データ暗号化
 - データ・ウェアハウス, 14-5
- 同期化
 - 演算子トリポジットリ・オブジェクト, 5-11
- 統計
 - オブティマイザ, 12-2

は

- パーティション化, 12-6
- パーティション外部結合, 9-13
- パーティション交換ロード (PEL), 6-2
 - 制限, 6-6, 6-9
 - 説明, 6-2
 - ターゲットの構成, 6-6
 - パフォーマンスの考慮事項, 6-5
 - マッピング, 6-3
- パーティション・プルーニング, 12-6
- パーティションワイズ結合, 12-6
- ハードウェア構成, 2-2
 - 確認, 2-4
- バインド
 - 自動バインド, ルール, 4-9
- バックアップ, 13-2
- バックアップトリカバリ, 13-2
 - ベスト・プラクティス, 13-3
- パフォーマンス
 - SQL アクセス・アドバイザの使用, 12-5
 - SQL チューニング・アドバイザの使用, 12-5
 - クエリー・リライトの使用, 12-8
 - ヒントの使用, 12-4
- パラメータ
 - 初期化, 2-8
 - メモリー管理, 2-7
- パラレル問合せ, 11-2, 12-9
 - 監視, 11-2
- パラレル実行, 12-9
 - 動作, 12-9

ひ

- 必要なメモリー, 2-3
- 表
 - 「外部表」も参照
- ヒント, 12-4

ふ

- ファイングレイン監査
 - データ・ウェアハウス, 14-5
- フラット・ファイル
 - 「外部表」も参照
- プロパティ
 - 設定, 5-11
 - マッピング, 5-9
- 分析 SQL, 9-2
- 分析機能, 9-2

へ

- 並列度 (DOP), 12-9
- ベスト・プラクティス
 - ARCHIVELOG モード, 13-3
 - NOLOGGING, 13-5
 - Recovery Manager, 13-4
 - バックアップトリカバリ, 13-3
 - 表領域の相違, 13-8
 - 読取り専用表領域, 13-4

ま

- マッピング
 - 作成, 5-1
 - 説明, 5-2
 - 定義, 5-1
 - パーティション交換ロード, 6-3
 - プロパティの設定, 5-9
- マッピング演算子
 - 接続, 5-6
 - 説明, 5-2
 - タイプ, 5-4
 - 追加, 5-4
 - 編集, 5-6
 - リポジットリ・オブジェクトとの同期化, 5-11
- マッピング演算子, 設定, 5-11
- マテリアライズド・ビュー, 12-8

め

- メタデータ
 - 依存性, 8-1
 - メタデータ依存性, ダイアグラム, 8-2
 - メタデータ依存性マネージャ, 8-2
- メモリー, 2-3
- メモリー管理
 - パラメータ, 2-7

よ

- 読取り専用表領域
 - ベスト・プラクティス, 13-4

ら

- ランタイム・パフォーマンスの改善, 6-1

り

リソース使用量, 12-6

最小化, 12-6

リソース・マネージャ, 11-4

リフレッシュ, 10-2

データ・ウェアハウス, 10-2

リライト

問合せ, 2-8, 2-9

ろ

ローリング・ウィンドウ, 10-5

ロール

ディメンション・ロール, 4-6

データ・ウェアハウスのセキュリティ, 14-2