

Oracle Data Integrator 12cの新機能概要

ビッグ・データとクラウドの進化

Oracle ホワイト・ペーパー | 2018 年 12 月



目次

目次.....	1
概要.....	4
Oracle Data Integrator 12.2.1.3.1.....	5
Oracle Object Storage と Oracle Object Storage Classic.....	5
自律型データベース：Oracle Autonomous Data Warehouse と Oracle Autonomous Transaction Processing.....	5
Oracle Enterprise Resource Planning (Oracle ERP) Cloud	6
Oracle Data Integrator 12.2.1.3.0.....	7
ビッグ・データ	7
クラウド	8
Oracle Data Integrator 12.2.1.2.6.....	9
ビッグ・データ	9
クラウド	10
ライフサイクル管理	11
開発者の生産性	12
Oracle Data Integrator 12.2.1.1.0.....	13
Hyperion Essbase および Hyperion Planning ナレッジ・モジュール	13
GoldenGate ナレッジ・モジュールにおける統合キャプチャおよび配信サポート	13
キューブおよびディメンションのサポート	13
ビッグ・データ構成ウィザード	14
Oracle Data Integrator 12.2.1.0.0.....	15
ビッグ・データのサポート強化	15
Oracle Data Integrator オブジェクトのライフサイクル管理	15
グローバル Oracle Data Integrator オブジェクト共有のための ODI 交換	16
Oracle 接続性の強化	17
Oracle Enterprise Data Quality との統合の強化	17
複合ファイルの強化	17

複合ファイル、ファイル、LDAP、JMS キューXML、 JMS トピック XML、および XML テクノロジーの強化.....	18
XML および複合 JDBC ドライバ向けの前処理/後処理.....	18
Web サービスのサポートの向上.....	18
インポート/エクスポートおよびリバース・エンジニアリング操作の取消し.....	18
分析関数またはウィンドウ関数のサポート.....	18
Studio/リポジトリに接続しているユーザー・リストの表示機能.....	18
Oracle Data Integrator コンソールの強化.....	18
Oracle Data Integrator 12.1.3.0.1.....	19
おもな機能：ビッグ・データのサポート強化.....	19
Spark および Pig の導入.....	19
Oozie を使用した Oracle Data Integrator ジョブのオーケストレーション	19
Hive ドライバおよびナレッジ・モジュールの強化.....	19
Hadoop 監査ログの取得.....	20
Oracle Data Integrator ファイル・ツールにおける HDFS アクセス.....	20
新規のフラット化コンポーネントおよび不規則コンポーネント	20
Oracle Data Integrator 12.1.3.0.0.....	21
おもな機能：ビッグ・データ・サポートの拡張.....	21
Sqoop の統合	21
HBase 統合.....	21
Hive 追加の最適化.....	21
セキュリティの機能拡張	21
Oracle Data Integrator による米国連邦情報処理標準（FIPS）準拠.....	21
設計時環境と生産性	21
「新規モデルとトポロジ・オブジェクト」ウィザード.....	21
シナリオおよびロード・プランの同時実行に対する制御の向上.....	22
その他の機能.....	22
Oracle Data Integrator XML ドライバの強化	22

JSON のサポート.....	22
マルチスレッド・ターゲット表のロード.....	23
ドキュメントの変更.....	23
Oracle Data Integrator 12.1.2.0.0.....	24
おもな機能：設計時環境と生産性.....	24
設計時環境.....	24
再利用可能なマッピング.....	25
複数ターゲットのサポート.....	25
ステップ・バイ・ステップ・デバッグ.....	26
ナレッジ・モジュール・アーキテクチャ.....	26
実行時パフォーマンスの強化.....	26
セッション・オーバーヘッドの低減と並列処理機能の強化.....	26
Oracle GoldenGate との優れた統合性.....	27
企業規模のデプロイメントを実現する Oracle アーキテクチャ.....	27
Oracle Platform Security Services の統合.....	28
統合された管理と監視.....	28
オラクルの戦略的データ統合プラットフォームへの移行.....	29
Oracle Warehouse Builder の統合.....	29
その他の機能.....	30
XML ドライバの機能拡張.....	30
一意のリポジトリ ID：.....	30
Studio/リポジトリのパフォーマンス強化.....	30
まとめ.....	31

概要

オラクルはデータ統合市場のリーダーであり、Oracle Data Integrator Platform Cloud、Oracle Data Integrator Cloud Service、Oracle Data Integrator、Oracle GoldenGate、Oracle Enterprise Data Quality、Oracle Enterprise Metadata Management、Oracle Stream Analytics など、完全に統合されたもっとも包括的なデータ統合製品を提供しています。オラクルのデータ統合ソリューションは、信頼性の高いタイムリーな異種データに対して、継続的なアクセスを全社的に提供することで、オンプレミスとクラウドで分析データと業務データの両方の統合をサポートします。

Oracle Data Integrator (ODI) は、大量データの高速移動および複雑なデータ変換処理に焦点を合わせたベストオブブリードのデータ統合プラットフォームです。オープンな統合 E-LT アーキテクチャを使用し、ビッグ・データ向けのサポートが拡張された ODI は、エンタープライズ・プラットフォーム間のデータ移動およびデータ変換を高パフォーマンスで提供します。ODI は、オンプレミスまたはクラウド内のいずれであっても、ビッグ・データ管理、サービス指向アーキテクチャ、ビジネス・インテリジェンスといった、データ統合の取組みにおいて重要です。使いやすいユーザー・インタフェースと拡張性に優れたフレームワークによって、生産性が向上し、開発コストが削減され、データ中心アーキテクチャの総所有コストが低減します。

Oracle Data Integrator は、さまざまなテクノロジーの中でも Oracle Cloud、Oracle Autonomous Data Warehouse、Oracle Autonomous Transaction Processing、Oracle ERP Cloud、Oracle Fusion Middleware、Oracle GoldenGate、Oracle Enterprise Data Quality、Oracle Database、Oracle Big Data Appliance、Oracle Exadata と完全に統合されているため、企業を中心にデータを置くことができます。Oracle Data Integrator はオープンかつ標準に準拠しているため、オラクルのアプリケーションだけでなくサード・パーティ・アプリケーションとも連携できます。Oracle Data Integrator の E-LT アーキテクチャにより、異種の RDBMS やビッグ・データ・エンジン向けのネイティブ・コードが生成されます。E-LT アーキテクチャでは、さまざまなソースからデータを抽出し、抽出したデータをターゲットにロードしてから、データベースまたはビッグ・データ・テクノロジーを使用してデータを変換します。データベースとビッグ・データの既存のインフラストラクチャを活かす Oracle Data Integrator により、比類のない効率性がもたらされます。E-LT アーキテクチャは、ネットワーク・トラフィックの軽減と、ターゲット・データを含むサーバーでのデータ変換によって、パフォーマンスを最大化します。

Oracle Data Integrator は、クリティカルなビッグ・データ統合機能も企業にもたらします。さらに、ビッグ・データに対する異種対応を可能にし、もっとも一般的に使用されている Hadoop 標準に至るまでサポートを拡張します。Oracle Data Integrator はデザインと実装の原則が分離されており、マッピングでは、基盤となるビッグ・データ・テクノロジーをシームレスに切り替えできます。こうした手法をとることで、実装の高速化およびビッグ・データ関連開発の効率化が可能となり、手動のコーディングが不要となると同時に、スペシャリストの技能および将来を見据えたビッグ・データへの投資に対する要件が緩和されます。

Oracle Data Integrator はオラクルの戦略的データ統合プラットフォームです。この最先端のデータ統合テクノロジーは、Oracle Data Integrator の 12c バージョンでも進化し、業界他社をはるかに凌駕しています。本書では、Oracle Data Integrator 12c プラットフォームで提供される新機能をいくつか取り上げ、詳しく説明します。

Oracle Data Integrator 12.2.1.3.1

Oracle Data Integrator 12c (12.2.1.3.1) では、以下の領域で新機能が導入されています。

- Oracle Object Storage および Oracle Object Storage Classic との統合
- Oracle Autonomous Data Warehouse Cloud と Oracle Autonomous Transaction Processing のナレッジ・モジュール (KM) を最適化
- Oracle Enterprise Resource Planning (Oracle ERP) Cloud のサポート

これらの主要な投資領域では、Oracle Data Integrator により、技術的な変換と現代化のプロセス全体においてお客様を支えています。

Oracle Object StorageとOracle Object Storage Classic

Oracle Object Storage と Oracle Object Storage Classic は高速で信頼できるセキュアなクラウド・ストレージで、Oracle Cloud Infrastructure (OCI) において Oracle Data Integrator とシームレスに統合できるようになりました。

Data Integrator にはナレッジ・モジュール (KM) と ODI ツールのセットが付属しており、マッピングとパッケージで使用する Oracle Object Storage と Oracle Object Storage Classic に接続し、ローカル・ディレクトリまたは Hadoop 分散ファイル・システム (HDFS) との間でファイル/オブジェクトをアップロード、ダウンロード、削除することができます。

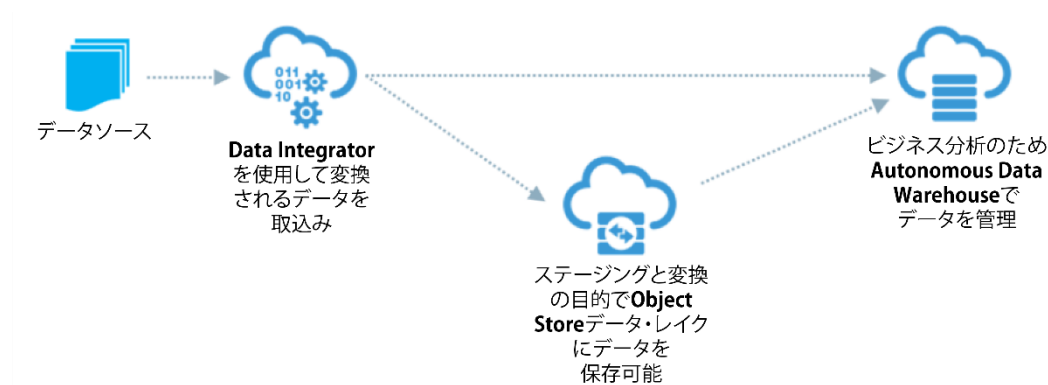
自律型データベース：Oracle Autonomous Data WarehouseとOracle Autonomous Transaction Processing

Oracle Data Integrator には、次の Oracle 自律型データベースによって認定される最適化されたロードおよび統合ナレッジ・モジュール (KM) が同梱されるようになりました。

- Oracle Autonomous Data Warehouse Cloud (Oracle ADWC)
- Oracle Autonomous Transaction Processing (Oracle ATP)

Oracle Data Integrator により、Oracle ADWC と Oracle ATP がシームレスに統合されます。Data Integrator を自律型データベースと統合することにより、さまざまなワークロードに合わせてチューニングおよび最適化された完全管理環境において、Oracle Data Integrator と Oracle データベースのパフォーマンスを最大限に引き出すことができます。

Oracle ADWC と Oracle ATP の両方でナレッジ・モジュールの同じセットを使用し、Oracle Object Storage と Oracle Object Storage Classic の新しいネイティブ統合も利用します。



図：Oracle Data IntegratorとAutonomous Data Warehouse Cloud

Oracle ADWC または Oracle ATP にデータを直接ロードできることに加えて、Oracle Data Integrator のユーザーは、Oracle Autonomous Data Warehouse と Oracle Object Storage の間のネイティブ統合による利点を活用して、Oracle ADWC や Oracle ATP に極めて高速にデータを転送することができます。その後 Data Integrator では、Oracle Autonomous Data Warehouse Cloud と Oracle Autonomous Transaction Processing のロード・プロセス全体を自動化することができます。

Oracle Enterprise Resource Planning (Oracle ERP) Cloud

このリリースには、Oracle Enterprise Resource Planning (Oracle ERP) Cloud の新しいテクノロジーおよびナレッジ・モジュールが追加されています。Oracle ERP は、財務、プロジェクト管理、調達、リスク管理、およびサイズ、業界、または地理的条件に関係なくすべてのビジネスにおいて重要な毎日の他の主要な活動のためのクラウド・アプリケーション・スイートです。

Oracle Data Integrator は、Oracle Enterprise Resource Planning Cloud とシームレスに統合され、データウェアハウス、データ・マート、またはデータ・レイクに Oracle ERP のデータを統合する組織を支援します。このネイティブ統合により、Data Integrator ユーザーはデータを Oracle ERP Cloud にロードすることもできます。

Oracle Data Integrator 12.2.1.3.0

Oracle Data Integrator 12c (12.2.1.3.0) では、以下の領域が強化されています。

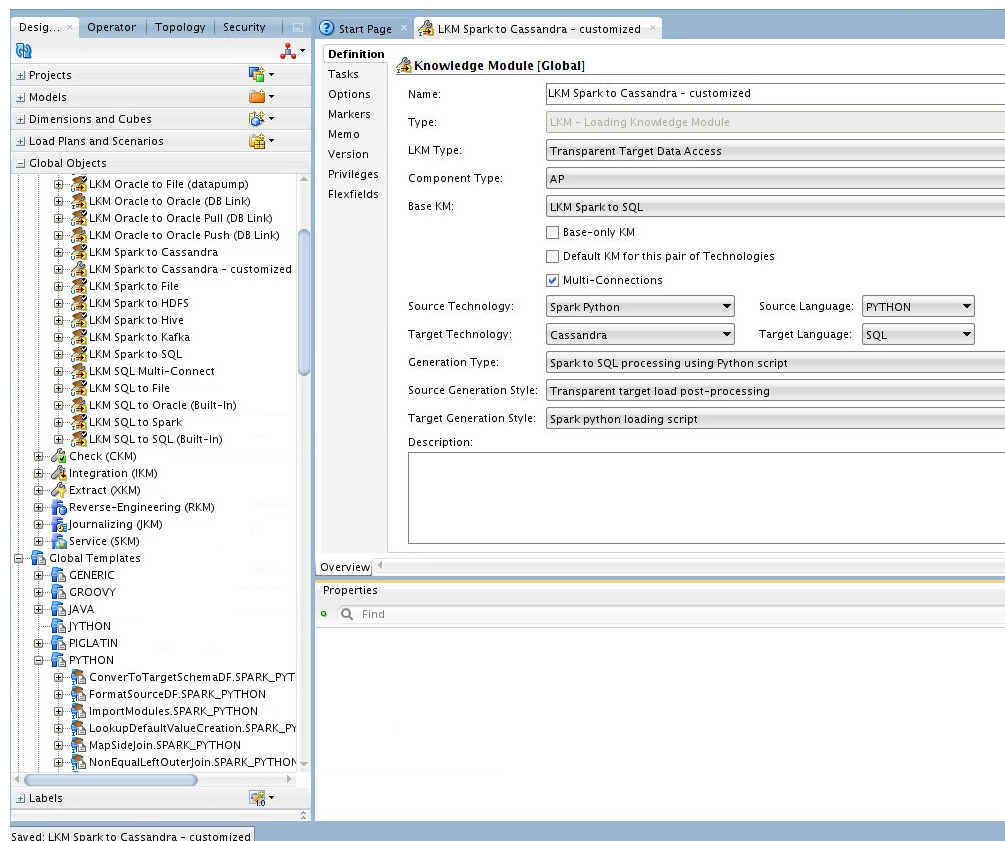
- ビッグ・データ
- クラウド

クラウドおよびビッグ・データは相変わらず主要な投資領域であり、Oracle Data Integrator は、技術を駆使した変換と最新化のプロセスを通じてお客様を支えています。


ビッグ・データ

ビッグ・データは、以下の進化点も含め、Oracle Data Integrator によって発展し続けます。

- Spark ナレッジ・モジュールの機能強化：本リリースでは、手作業によるスクリプトにひけをとらない、高パフォーマンスで読み取りしやすい Spark コードの生成がフォーカスされています。さらに、Spark ナレッジ・モジュールは、Dataframes をはじめとする Apache Spark 2.x の最新機能を活用しています。適用可能な場面では SparkSQL も活用されており、Spark 上で実行される ODI プロセスが高速化されます。
- ナレッジ・モジュール・エディタにおける Spark ナレッジ・モジュールのサポート：Spark ナレッジ・モジュールがナレッジ・モジュール・エディタにおいて完全にサポートされ、特定の要件に合わせたカスタマイズが可能です。



図：ナレッジ・モジュール・エディタにおけるSPARKナレッジ・モジュール

- 
- Hadoop 複合タイプの強化：ODI 12.2.1.2.6 では、Hive および HDFS（バッチ・ボード）テクノロジー向けに複合タイプのサポートが導入されました。ODI 12.2.1.3.0 では、HDFS だけでなく Apache Kafka においても複合タイプをサポートすることで Spark Streaming 向けにこの機能が強化されています。
 - ビッグ・データ構成ウィザード：Cloudera's Distribution including Apache Hadoop（CDH）の最新リリース向け新テンプレートにより、ビッグ・データ構成ウィザードが改善されています。

クラウド

クラウドでは以下の点が改善されています。

- Salesforce.com での認定：Oracle Data Integrator は、Salesforce.com によって完全認定されており、このテクノロジー向けのすぐに使用可能な JDBC ドライバが追加されています。

Oracle Data Integrator 12.2.1.2.6

4 つの主要な領域、すなわちビッグ・データ、クラウド、ライフサイクル管理、および開発者の生産性における機能強化についてさらに紹介します。

ビッグ・データ

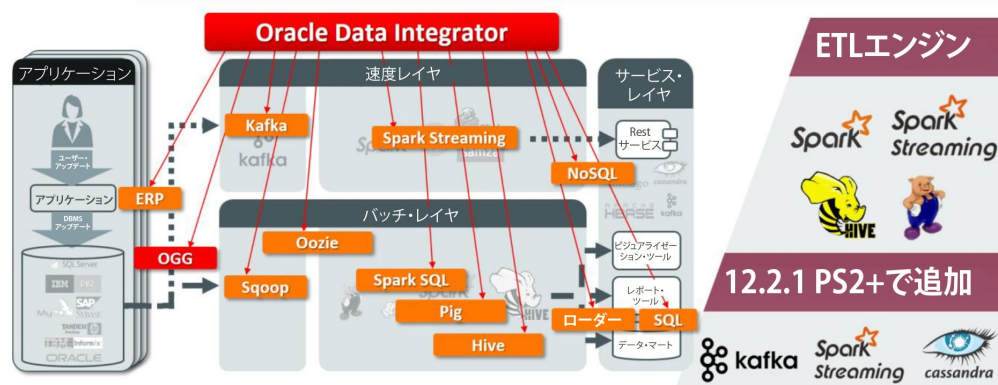
Spark Streaming のサポート：Oracle Data Integrator では Spark Streaming のサポートが追加されており、エンドユーザーがコードを 1 行も書かなくても、ビッグ・データのストリーミング・ジョブを簡単に作成できます。Spark Streaming はもちろん、データ処理に関していえば、Data Integrator は Hive、Pig、およびバッチ Spark をすでにサポートしています。マッピングの論理設計と物理設計を独自に分離したことで、Oracle Data Integrator は市場で提供されている唯一のデータ統合ツールとなっており、開発者は一般的なビジネス・ロジックでマッピングを設計する柔軟性を獲得し、さまざまなデータ処理テクノロジー（Hive、Spark、Spark Streaming など）向けに必要なコードを生成できます。この独自の機能は、将来においても有効性が発揮されるデータ統合プロセスを実現する一助ともなります。

Apache Kafka および Apache Cassandra のサポート：Apache Kafka および Cassandra は、ソースおよびターゲットの両方として、Oracle Data Integrator の最新バージョンで認定されています。

Hadoop 複合タイプおよび保管形式：本リリースでは、市場をリードする Hadoop に対するサポートがさらに拡張され、Avro、Parquet、JSON など、さまざまな形式で保管されるデータへのネイティブ・アクセスが可能になっています。さらに、Array、Struct、Map など、マッピングにおいて複合タイプまたはネストされたタイプを利用する新機能も追加されています。

ビッグ・データ構成ウィザードの強化：ODI 12.2.1.1.0 で導入されたビッグ・データ構成ウィザードが改善され、Kafka や Cassandra といった新しい Hadoop テクノロジーに対応しました。さらには、Kerberos で保護される Hadoop クラスタを使用して Oracle Data Integrator を構成できるようになりました。

ビッグ・データ変換のためのODI



図：ビッグ・データのためのODI

クラウド

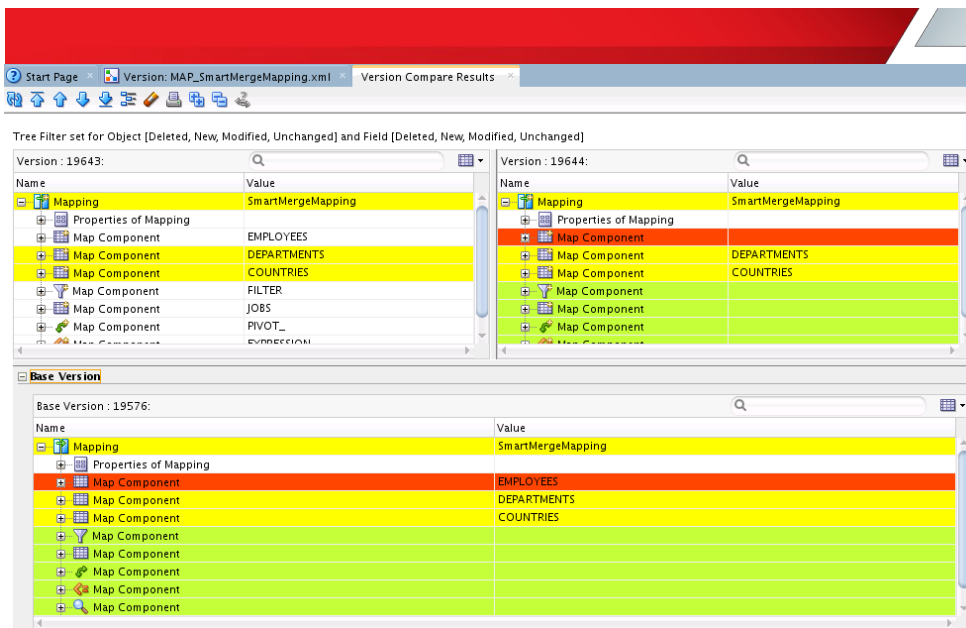
RESTful サービスのサポート：Oracle Data Integrator で RESTful サービスを起動できるようになりました。RESTful サービスの接続性、リソース URI、メソッドおよびパラメータは、他のデータソースの接続性と同様、トポロジ構成の中に構成できます。数多くのパラメータがサポートされており、広く普及する RESTful サービスを最大限の柔軟性でサポートします。大きなペイロードのアップロードまたはダウンロードのために、データのチャンク化およびページ区切りもサポートされています。

Business Intelligence Cloud Service (BICS) のナレッジ・モジュール：特別な設定なしで、Business Intelligence Cloud Service が Oracle Data Integrator においてサポートされるようになりました。Business Intelligence Cloud Service の接続性のトポロジにおける定義、メタデータのリバース・エンジニアリング、そして BICS へのデータのロードを、他のターゲット・データ・サーバーと同じように実行できます。

ライフサイクル管理

Git のサポートおよびライフサイクル管理機能の向上：Oracle Data Integrator では、Apache Subversion に加え、外部のバージョン管理システムとして Git のサポートが追加されました。サポートされたバージョン管理システムのいずれかを使用することで、統合されたユーザー・エクスペリエンスを得られます。ライフサイクル管理の向上ニーズに対し、多くの最新オペレーションが追加されています。保留中の変更すべてを表示したり、オブジェクトのセットに対応するバージョンを 1 か所から作成したりできます。複数バージョンやデプロイメント・アーカイブを作成しながらシナリオを再生成することが可能で、シナリオを常に現行のオブジェクト・バージョンに符合させることができます。ブランチまたはタグからオブジェクトを選択して移入できます。SDK API も強化され、継続的な統合に必要なさまざまなオペレーションがサポートされています。

強化されたマージ機能：3 方向マージおよびオブジェクト変更検出に基づく、変更の自動マージが行われるようマージ機能が強化されました。マージ・ユーザー・インタフェースでの、ソースまたはターゲットからのプロパティまたはオブジェクトの選択が新たにサポートされ、競合解消が簡素化されています。



図：強化されたマージ機能



開発者の生産性

強化されたナレッジ・モジュール・フレームワーク：ナレッジ・モジュール開発フレームワークに革新的な新機能が追加され、柔軟性を最大限に、メンテナンスを最小限にします。あるナレッジ・モジュールから別のナレッジ・モジュールへと手順を継承でき、オブジェクト指向プログラミング言語のように手順をオーバーライドできます。最新のテンプレート言語と構文が導入され、生成されたコードに対する管理性が向上します。さらに、構文ハイライト、オートコンプリート、コード・ブロックの折りたたみなど、他にも生産性向上に寄与する機能が追加されています。

Oracle Data Integrator 12.2.1.1.0

Hyperion EssbaseおよびHyperion Planningナレッジ・モジュール

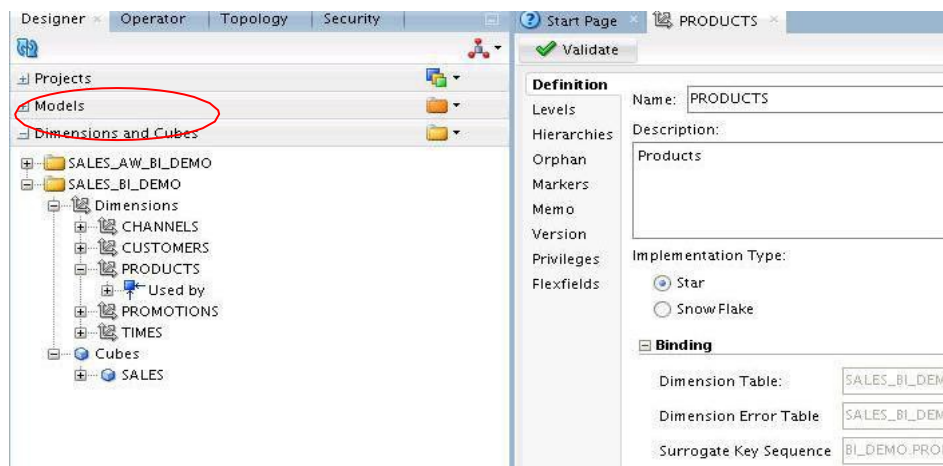
Hyperion Essbase および Hyperion Planning ナレッジ・モジュールが Oracle Data Integrator で即時利用できるになり、これら Hyperion アプリケーションの最新バージョン（11.1.2.4）がサポートされます。

GoldenGateナレッジ・モジュールにおける統合キャプチャおよび配信サポート

Oracle データベース向け GoldenGate ジャーナル化ナレッジ・モジュール（JKM）がアップデートされ、統合キャプチャと配信に対するサポートが追加されました。アップデートされた機能によりパフォーマンスが改善され、スケーラビリティとロードバランシングが向上します。

キューブおよびディメンションのサポート

Core ETL-ELT の拡張が行われ、ODI はキューブおよびディメンションという 2 種類のディメンション・オブジェクトをサポートします。キューブ・オブジェクトおよびディメンション・オブジェクトをマッピングにおいて直接作成、使用することができるようになり、ディメンション・オブジェクトのロードを自動化する付属のパターンにより、開発者の生産性が向上します。これにより、Type 2 Slowly Changing Dimensions も改善され、最新の Type 3 Slowly Changing Dimensions が ODI でサポートされます。



図：ODI STUDIOでのキューブおよびディメンション



ビッグ・データ構成ウィザード

最新のビッグ・データ構成ウィザードが ODI Studio Gallery で利用できるようになり、Hadoop、Hive、Spark、Pig、Oozie などの Hadoop テクノロジー向けのトポロジ・オブジェクトを構成する単一のエントリ・ポイントとなります。

Oracle Data Integrator 12.2.1.0.0

おもな機能：ライフサイクル管理および強化されたビッグ・データのサポート

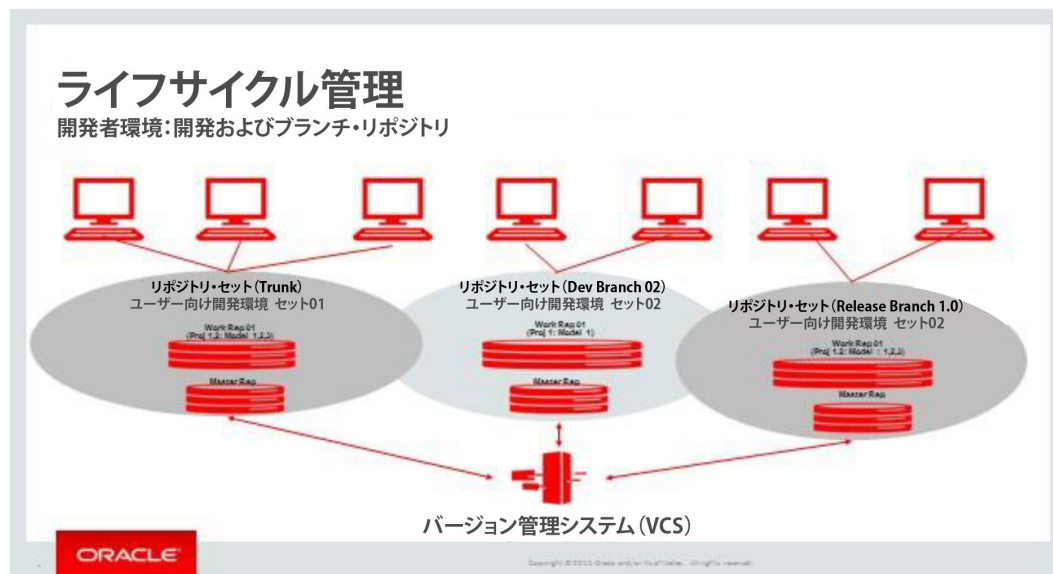
ビッグ・データのサポート強化

Oracle Data Integrator のこれまでのリリースでは、Apache Spark、Apache Pig、および Oozie を使用したオーケストレーションがサポートされ、Hadoop プロジェクトの水準を高めるために Oracle Data Integration が展開するビッグ・データの最新のアップデートと機能がとりまとめられていました。本リリースでは、Oozie ワークフロー生成のためにタスク実行モードとセッション実行モードをユーザーが選択できるようにすることで、前述の機能を強化しています。ユーザーは、Oozie ワークフロー生成において、タスク実行モードまたはセッション実行モードを選択できます。新しいセッション・モードでは、パッケージのトランザクション、スクリプト、およびループがサポートされます。Oracle Data Integrator は、実行されたオブジェクトに基づいて正しいモードを自動的に選択します。モードは手動で選択することもできます。

Oracle Data Integrator オブジェクトのライフサイクル管理

Oracle Data Integrator は、リリース管理機能を導入し、開発環境とデプロイ環境を区別しています。開発環境からデプロイ用アーティファクトを作成できます。デプロイ用アーティファクトは、テストのため Quality Assurance 環境でデプロイし、その後本番環境に提供できます。デプロイ用アーティファクトは、Oracle Data Integrator Studio を使用して作成するか、またはコマンドラインから作成できます。

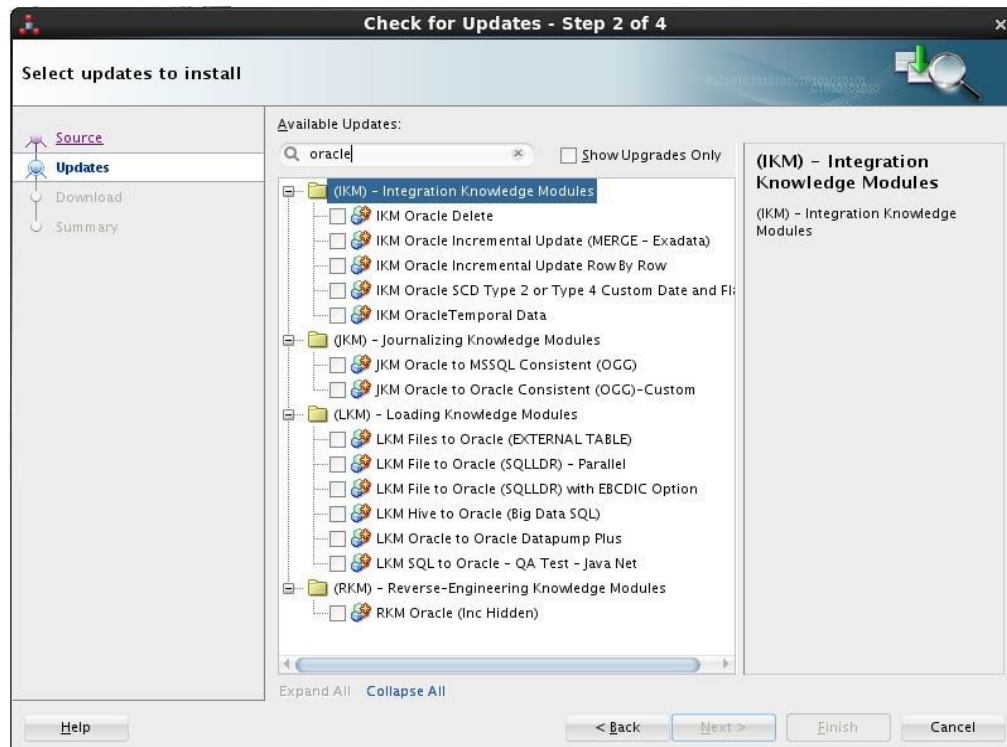
Oracle Data Integrator は Subversion と統合されており、Subversion 内の Oracle Data Integrator オブジェクトを管理できます。この統合を利用して、タグを作成し、Oracle Data Integrator オブジェクトの各バージョンのスナップショットを作成できます。分散した場所からの並行開発、または複数リリース向け並行開発のためにブランチを作成することもできます。



図：ライフサイクル管理

グローバルOracle Data Integratorオブジェクト共有のためのODI交換

Oracle Data Integrator は、公式またはサード・パーティの更新センターを通じ、Oracle または別の Oracle Data Integrator ユーザーがグローバルな ODI オブジェクトを参照、ダウンロード、インストールできる、コミュニティのような場を提供します。この機能は、グローバル・ナレッジ・モジュール、グローバル・ユーザー関数およびマッピング・コンポーネントで利用できます。Oracle Data Integrator Studio の Help メニューにあるメニュー項目 Check for Updates を使用することで更新センターに接続し、前述のグローバル Oracle Data Integrator オブジェクトを取得できます。



図：グローバルOracle Data Integratorオブジェクト

Oracle接続性の強化

Oracle Data Integrator のナレッジ・モジュールのレパートリーは継続して改善されています。パーティション交換ロードを実行するナレッジ・モジュールを使用でき、必要に応じてパーティションを交換できます。外部表を使用するロード・ナレッジ・モジュールも改良されており、一度に複数のファイルをロードできます。Data Pump ナレッジ・モジュールも改良されています。

Oracle Enterprise Data Qualityとの統合の強化

新しい Oracle Enterprise Data Quality (Oracle EDQ) テクノロジーがトポロジで利用可能となり、Oracle EDQ 用のデータ・サーバー、物理スキーマ、論理スキーマを作成できます。さらに、コンテキストおよび論理スキーマを通じて EDQ データ・サーバーをサポートするよう OdiEnterpriseDataQuality ツールが改良されています。

複合ファイルの強化

Oracle Data Integrator の Native Format ビルダー・ユーティリティが Oracle Data Integrator Studio に同梱されるようになり、Oracle Data Integrator のユーザー・インターフェースから nXSD ファイルを作成できるようになりました。

複合ファイル、ファイル、LDAP、JMSキューXML、JMSトピックXML、およびXMLテクノロジーの強化

ユーザビリティを強化するため、Oracle Data Integrator の JDBC プロパティ（複合ファイル、ファイル、LDAP、JMS キューXML、および XML テクノロジー）が、デフォルト値（該当する場合）およびプロパティの説明とともにデータ・サーバー・レベルで表示されるようになりました。

XMLおよび複合JDBCドライバ向けの前処理/後処理

Oracle Data Integrator では、XML および複合ファイル・ドライバにデータをフィードする方法をカスタマイズできます。これにより、Oracle Data Integrator を使用して外部エンドポイントからデータを受け取る場合、または外部エンドポイントにデータを書き出す場合のデータ処理に追加可能な、中間処理段階がサポートされます。さらには、ODI XML または複合ファイル JDBC ドライバを使用するデータ・サーバー構成の一部として、中間処理段階の複雑な構成もサポートします。

Webサービスのサポートの向上

Oracle Data Integrator は、トポロジで利用可能な最新の SOAP Web Service テクノロジーを提供し、Web Service 向けのデータ・サーバー、物理スキーマおよび論理スキーマを作成可能にします。Oracle Web Service Management ポリシーが Web Service データ・サーバーに添付できるようになりました。さらに、OdilnvokeWebService ツールが強化され、コンテキストおよび論理スキーマを通じて Web Service データ・サーバーをサポートするようになりました。

インポート/エクスポートおよびリバース・エンジニアリング操作の取消し

Oracle Data Integrator は、長時間実行される可能性のあるインポート/エクスポートおよびリバース・エンジニアリング操作を取り消せるようになりました。

分析関数またはウィンドウ関数のサポート

Oracle Data Integrator は、即時利用可能な分析関数またはウィンドウ関数をマッピング・レベルでサポートします。分析関数（PERCENT_RANK、LAST、FIRST、LAG など）を、どのコンポーネントにおいても対応付けの式のレベルで使用できます。

Studio/リポジトリに接続しているユーザー・リストの表示機能

メニュー項目「ユーザー・アクティビティの確認」が「セキュリティ」メニューに追加されました。これにより、「ユーザー接続」ダイアログでユーザー・アクティビティ記録を表示、除外、保存できます。これは、Oracle Data Integration Studio および Oracle Data Integrator コンソールの両方で利用できます。

Oracle Data Integratorコンソールの強化

Oracle Data Integrator コンソールの全体的なルック・アンド・フィールが改善されました。さらには、ユーザーやプロファイルの作成などのセキュリティ・タスクを Oracle Data Integrator コンソールで実行できるようになりました。リリース管理作業を Oracle Data Integrator コンソールから実行でき、トポロジ・タスクに関連する機能が強化されています。

Oracle Data Integrator 12.1.3.0.1

おもな機能：ビッグ・データのサポート強化

SparkおよびPigの導入

Oracle Data Integrator には Spark または Pig を使用したマッピングの実行が導入されています。Oracle Data Integrator では、実装言語に依存しない論理設計を通じてマッピングを定義できます。ユーザーは、Hadoop ベースの変換において、生成される変換コードとして Hive、Spark、または Pig から選択することができ、環境とユースケースに基づいて最適な実装を選択できます。

- **Spark** : Oracle Data Integrator マッピングでは PySpark を生成でき、Python 言語で Spark プログラミング・モデルが公開されます。Apache Spark は大規模データ処理向けの変換エンジンです。Apache Spark により、大規模データセットの高速なインメモリ処理が可能になります。カスタム PySpark コードは、ユーザー定義関数またはテーブル・ファンクション・コンポーネントによって追加可能です。
- **Pig** : Oracle Data Integrator マッピングでは、変換言語および実行エンジンとして Pig Latin を利用できます。Apache Pig は、Hadoop において大規模データセットを分析するためのプラットフォームで、データ分析プログラムを表現するための高度な言語である Pig Latin を使用します。Pig 変換は、ローカル・モードまたはマップ削減モードのいずれかで実行できます。カスタム Pig コードは、ユーザー定義関数またはテーブル・ファンクション・コンポーネントによって追加可能です。


Oozieを使用したOracle Data Integratorジョブのオーケストレーション

Oracle Data Integrator は、ジョブ（例：マッピング、パッケージ、シナリオ、プロシージャ）のオーケストレーション・エンジンとして、従来の Oracle Data Integrator Agent および Apache Oozie の両方をサポートします。Apache Oozie により、オーケストレーション用の Oracle Data Integrator エージェントをインストールしなくても、Hadoop インフラストラクチャ上でのネイティブ実行が可能です。

Oozie ツールを利用して、ジョブのスケジュール設定、管理、および監視ができます。Oracle Data Integrator は、Oozie のネイティブ・アクションを使用して、Hadoop のプロセスおよび条件付きブランチ処理ロジックを実行します。

Hiveドライバおよびナレッジ・モジュールの強化

Oracle Data Integrator には、全面的な JDBC 準拠やパフォーマンス向上など、数多くの利点をもたらす WebLogic Hive JDBC ドライバが含まれています。この新しいドライバを活かすため、すべての Hive ナレッジ・モジュールが再作成されています。そして、ソースからのロードが主要な目的であるナレッジ・モジュールは、ロード・ナレッジ・モジュールとして提供されるようになり、1 回のマッピングで他のロード・ナレッジ・モジュールと組み合わせることができます。新クラスのロード・ナレッジ・モジュールである“_direct load_”も、中間ステージングなしにターゲットをロードできます。さらに、テーブル・ファンクション・コンポーネントが、Hive 構成メンバーをサポートするよう拡張されています。



新規/更新済み Hive コンポーネント・スタイル・ナレッジ・モジュール：

- LKM SQL to Hive SQOOP
- LKM SQL to File SQOOPDirect
- LKM SQL to HBase SQOOPDirect
- LKM HBase to SQL SQOOP
- LKM File to SQL SQOOP
- LKM Hive to SQL SQOOP
- LKM File to Hive LOADDATA
- LKM File to Hive LOAD DATA Direct
- LKM HBase to HiveHBASE-SERDE
- LKM Hive to HBase Incremental Update HBASE-SERDE Direct
- LKM Hive Append
- LKM Hive to File Direct
- LKM Hive to OracleOLH-OSCH
- LKM File to Oracle OLH-OSCH Direct
- LKM Hive to OracleOLH-OSCH
- LKM Hive to Oracle OLH-OSCH Direct
- XKM Hive Sort

Hadoop 監査ログの取得

Oracle Data Integrator は、Oozie、Pig および他のタスク実行のため Hadoop 監査ログの結果を Operator タスクに統合します。このログ結果には Map-Reduce 統計が表示され、ネイティブ Web コンソールに表示される完全な Hadoop 統計へのリンクが提供されます。

Oracle Data Integrator ファイル・ツールにおける HDFS アクセス

パッケージとプロシージャで利用される Oracle Data Integrator のファイル・ベース・ツールが強化され、HDFS ファイル処理が追加されています。これには、ファイルのコピー、移動、追加および削除、ファイル変更の検出、フォルダの管理、FTP を使用した HDFS へのファイルの直接転送が含まれます。

新規のフラット化コンポーネントおよび不規則コンポーネント

Oracle Data Integrator では、マッピング用にフラット化コンポーネントが導入されており、複雑な下位構造体を属性フラット・リストの一部として処理できます。

さらなる処理のため Key-Value リストを名前付き属性に変換する不規則コンポーネントも導入されています。

Oracle Data Integrator 12.1.3.0.0

おもな機能：ビッグ・データ・サポートの拡張

Sqoopの統合

Oracle Data Integrator では、Apache Sqoop の使用による、以下のソースおよびターゲットのロードを導入しています。

- リレーショナル・データベースから HDFS、Hive、HBase へ（ナレッジ・モジュール IKM File- Hive to SQL（SQOOP）使用）
- HDFS および Hive からリレーショナル・データベースへ（ナレッジ・モジュール IKM SQL to Hive- HBase-File（SQOOP）使用）

Sqoop を使用すると、Hadoop Map-Reduce プロセスにおいて、並行 JDBC 接続を使用したメカニズムのロードおよびアンロードが可能となります。

HBase統合

Oracle Data Integrator では、新しいテクノロジーと以下のナレッジ・モジュールを介して Apache HBase がサポートされています。

- LKM HBase to Hive（HBase-SerDe）
- IKM Hive to HBase Incremental Update（HBase-SerDe）
- RKM HBase

Hive追加の最適化

Hive に書き込むナレッジ・モジュールが Hive 0.8+をサポートし、既存のデータを新規追加ファイルにコピーするのではなく、既存のデータファイルにデータを追加できるようになりました。

セキュリティの機能拡張

Oracle Data Integratorによる米国連邦情報処理標準（FIPS）準拠

Oracle Data Integrator では、ナレッジ・モジュール、プロシージャ、シナリオ、アクションおよびパスワードを暗号化するための標準暗号化アルゴリズムとして Advanced Encryption Standard（AES）を利用しています。要件を満たすよう暗号化アルゴリズムと鍵の長さを構成できます。リポジトリ・エクスポートに含まれるパスワードおよびその他の機密情報は暗号化され、パスワードで保護されます。

設計時環境と生産性

「新規モデルとトポロジ・オブジェクト」ウィザード

デザイナー・ナビゲータの「新規モデルおよびトポロジ・オブジェクトの作成」ダイアログでは、新規モデルを作成し、そのモデルを新規または既存のトポロジ・オブジェクトに関連付けることができます（作業リポジトリに接続されている場合）。詳細なオプションが必要な場合を除き、このダイアログを使用することで、トポロジ・エディタを使用しなくてもトポロジ・オブジェクトを作成できます。

Create New Model and Topology Objects

Enter all details to create a new model and optionally any new data servers, contexts and schemas

Name:

Technology:

Logical Schema:

Context:

Data Server

Name:

User:

Password:

Driver List:

Driver:

URL:

Physical Schema

Name:

Schema (Schema):

Schema (Work Schema):

図：「新規モデルおよびトポロジ・オブジェクトの作成」ウィザード

シナリオおよびロード・プランの同時実行に対する制御の向上

制御の向上により、シナリオまたはロード・プランにおける同時実行を制限したり、強制的に同時実行を待機、または強制的に同時実行して実行エラーを発生させたりできます。

その他の機能


Oracle Data Integrator XMLドライバの強化

XMLスキーマのサポートについて以下の拡張機能が追加されています。

- 再帰：Oracle Data Integrator は、XML スキーマ内の再帰をサポートします。
- any、anyType、anyAttribute：これらの型によって定義されるデータは、元のドキュメントから XML マークアップによって文字列型の列に格納されます。
- メタデータのアノテーションは、使用する表名、列名、型、長さ、精度を ODI XML ドライバに指示するため、XML スキーマ内に追加できます。

JSONのサポート

Oracle Data Integrator 複合ファイル・ドライバでは、JSON 形式でのファイルの読み取りと書き込みが可能となりました。JSON 構造体は、nXSD スキーマを通じて定義されます。



マルチスレッド・ターゲット表のロード

Oracle Data Integrator は、複数のパラレル接続を使用してターゲット表をロードする機能を提供します。この機能は、トポロジ・ナビゲータのデータ・サーバー・レベルで利用できる「ターゲットの並列度」プロパティによって制御します。

ドキュメントの変更

これまで『Oracle Data Integrator 開発者ガイド』で提供されていた情報があらためて整理されました。Oracle Data Integrator のドキュメント・ライブラリに以下の新しいガイドが追加されています。

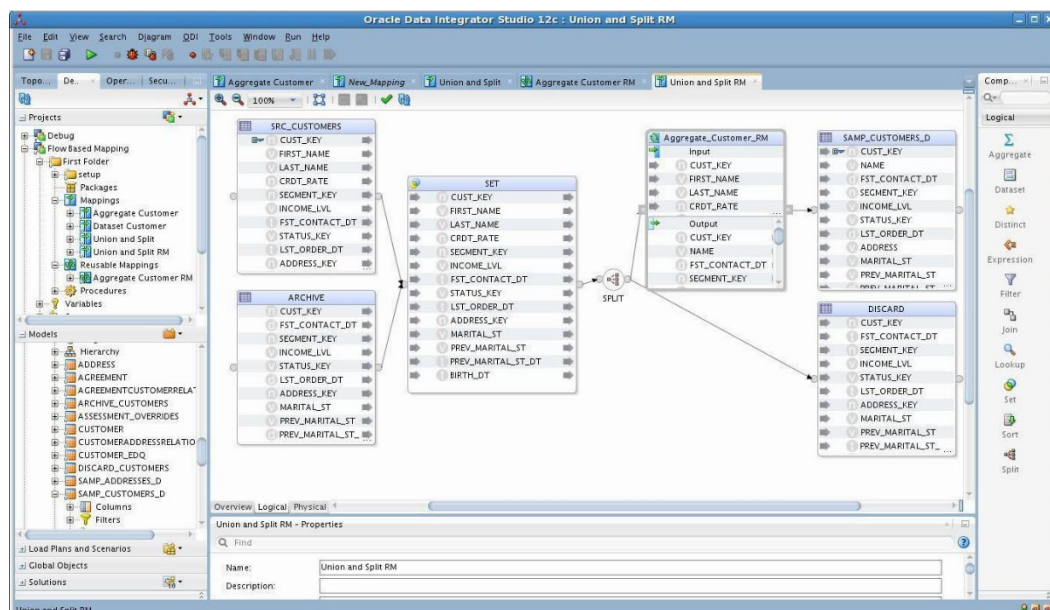
- Oracle Data Integrator の理解
- Oracle Data Integrator の管理
- Oracle Data Integrator ツール・リファレンス・ガイド

Oracle Data Integrator 12.1.2.0.0

おもな機能：設計時環境と生産性

設計時環境

Oracle Data Integrator 12c では新しいフローベースの宣言的ユーザー・インタフェースを採用し、生産性を大幅に向上させました。この改良されたユーザー環境は、シンプルながらも強力で包括的です。このリリースでは、ユーザーの使い勝手と生産性を強化させるために、Oracle Data Integrator Studio のクライアントが全面的に再設計されました。新しいユーザー・インタフェースでは、これまでの宣言的設計モデル手法とフローベースの方法が融合されています。そのため、一般的なデータ統合の設計やデプロイメントのユースケースが簡素化され、実装に要する時間が短縮されます。新しいユーザー・インタフェースでは、宣言的アプローチの分かりやすさと使いやすさに、構成可能なフローの柔軟性と拡張性が融合されています。ソースとターゲットのデータ形式とデータ統合プロセスを記述するのはデータ統合の設計担当者です。ビジネス・ユーザーや開発者は実行方法ではなくその内容に専念できます。ソースをターゲットに接続するのはマッピング（Oracle Data Integrator 11g のインタフェースの概念を継承する機能）で、このとき使用するフローを構成するコンポーネントには、結合、フィルタ、集計、設定、分割などがあります。さまざまなソース・システムとターゲット・システム間にこれらのプロセスを実装するために必要なコードは、Oracle Data Integrator Enterprise Edition が生成、デプロイ、管理します。



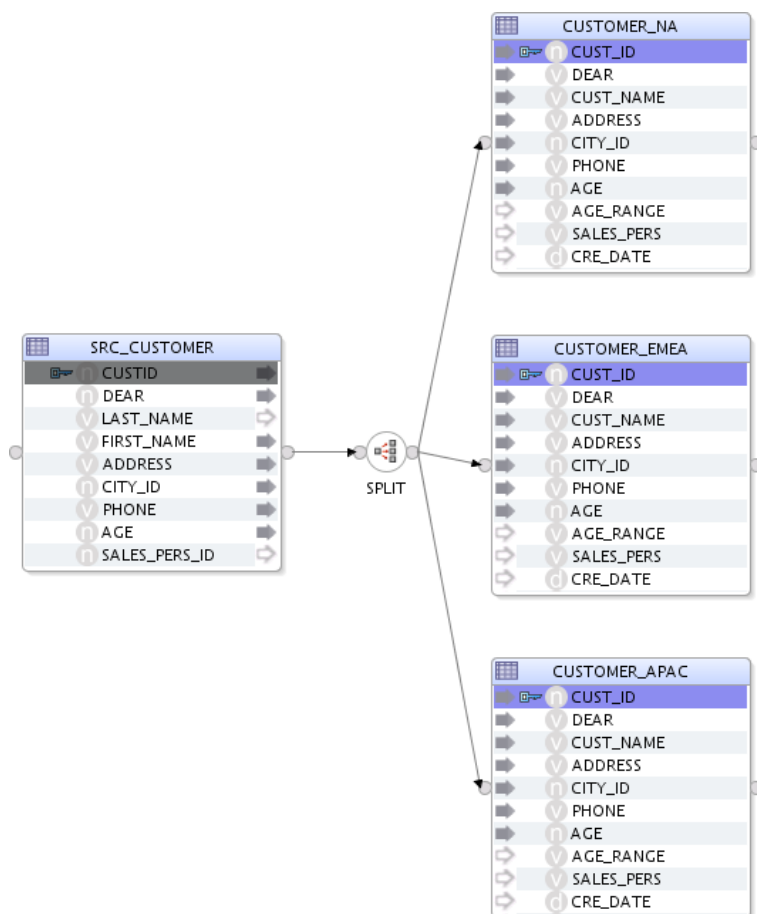
図：Oracle Data Integratorのフローベースの宣言的ユーザー・インタフェース

再利用可能なマッピング

開発時にマッピング・ロジックをシームレスに再利用できるようになったことで、パラダイムも強化されました。そのため、開発者はより簡単かつ効率的にソリューションを完成できます。再利用可能なマッピングを使用してフロー・セクションをカプセル化しておく、複数のマッピングで再利用できます。再利用可能なマッピングには、カプセル化されたフローに接続するための入/出力シグネチャを設定でき、再利用可能なマッピングの内部にカプセル化されているソースとターゲットを含めることができます。

複数ターゲットのサポート

Oracle Data Integrator のフローベースの宣言的ユーザー・インターフェースでは、単一フローの一部として複数のターゲットをロードするマッピングを作成できます。ターゲットが 1 つであっても複数であっても、ターゲットをロードする順序を指定できます。また、オプションで分割コンポーネントを使用すると、1 つ以上の条件に基づいて、行のルーティング先となるターゲットを変更することができます。

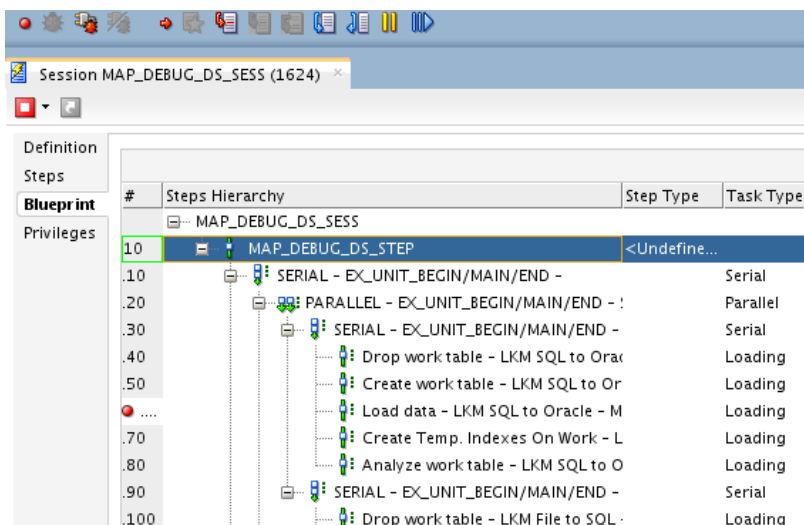


図：複数ターゲットのマッピング

ステップ・バイ・ステップ・デバッグ

Oracle Data Integrator にはステップ・バイ・ステップ・デバッグが導入されており、マッピング、パッケージ、プロシージャ、およびシナリオをステップ・バイ・ステップ方式でデバッグできます。ユーザーは、これらのオブジェクト内でタスクの実行状況を手動で詳しく調査したり、あらかじめ定義しておいた場所で実行が中断されるようにブレークポイントを設定したりできます。

デバッグ・セッション中に変数の値の確認や変更ができ、使用しているソースおよびターゲットのデータの問い合わせもできます。これには、コミットされていないトランザクションの内容も含まれるため、内部状況をより詳しく把握できます。



図：デバッガ－実行フローの制御

ナレッジ・モジュール・アーキテクチャ

本リリースでは、Oracle Data Integrator 11g 以降で利用できるテンプレート式ナレッジ・モジュールに加え、コンポーネント式ナレッジ・モジュールと呼ばれる新しい方式のナレッジ・モジュールが導入されました。この新しい方式のナレッジ・モジュールは拡張性に富んだコンポーネント・フレームワークを備えているため、変換の順序を宣言できるなど、マッピングの設計全体が向上します。また、モジュール同士を結合できるため、再利用性も向上し、コードとデータの重複を回避できるだけでなく、データベース・リンクを使用した Oracle の接続性も改善されます。

実行時パフォーマンスの強化

セッション・オーバーヘッドの低減と並列処理機能の強化

Oracle Data Integrator 12c では、パフォーマンス強化を目的に実行時の実行環境が改良されました。セッションの実行計画をキャッシュできるブループリント機能が導入されるなど、セッション実行時のオーバーヘッドを削減するさまざまな変更が行われてきました。また、並列処理機能も改良されたため、もともと高かった、Oracle Data Integrator アーキテクチャのパフォーマンスがさらに向上します。この効果は、複数のソースを並列してステージング・エリアにロードすることによりもたらされます。ロード処理の並列度はマップの物理ビューでカスタマイズできます。一時データベース・オブジェクトに一意の名前を使用できるオプションもあり、これを使用すると同じマッピングの平行実行が可能です。

Oracle GoldenGateとの優れた統合性

高パフォーマンスの E-LT 機能に Oracle GoldenGate が統合されたため、データウェアハウスへのリアルタイム・データのロードと変換のスピード・アップと効率化を図れるようになりました。ソース・システムやバッチ・ウィンドウの依存関係に影響を与えずに、リアルタイム・データウェアハウス・ソリューションを簡単に構成してデプロイできます。

Oracle GoldenGate は、Oracle Data Integrator 内のチェンジ・データ・キャプチャ (CDC) フレームワークのソースとして統合されますが、この統合性が次の部分で改良されました。

- Oracle GoldenGate のソース・システムとターゲット・システムは、Oracle Data Integrator のトポロジでデータ・サーバーとして構成されます。抽出プロセスとレプリケート・プロセスは、物理スキーマと論理スキーマによって表現されます。トポロジ内ではこのように表現されるため、複数コンテキストの構成を、それぞれのコンテキストに関する基本理念に従って分離できます。
- Oracle GoldenGate のパラメータのほとんどは、物理スキーマの構成時に抽出プロセスとレプリケート・プロセスに追加できます。ユーザー・インタフェースでは、リストからパラメータを選択できます。そのため、生成後の Oracle GoldenGate パラメータ・ファイルに変更を加える必要性が最小限に抑えられます。
- Oracle Data Integrator の 1 つのマッピングを、ジャーナル化したチェンジ・データ・キャプチャのロードとターゲットのバルク・ロードに使用できます。これは、Oracle GoldenGate ジャーナル化ナレッジ・モジュールによって可能となります。GoldenGate ジャーナル化ナレッジ・モジュールは、Oracle GoldenGate レプリケーション・ターゲットとは対照的に、ソース・モデルを使用するだけでなく、デプロイ仕様の一部としてマッピングでのジャーナル化構成を使用します。ジャーナル化されたロードおよびバルク・ロードのための 1 つのマッピングにおいて、複数のデプロイ仕様を使用できます。
- Oracle GoldenGate のパラメータ・ファイルは、JAgent テクノロジーを使用して自動的に Oracle GoldenGate のソース・インスタンスおよびターゲット・インスタンスにデプロイして開始できるようになりました。

企業規模のデプロイメントを実現するOracleアーキテクチャ

WebLogic 管理フレームワークを使用したスタンドアロン・エージェントの管理

Oracle Data Integrator のスタンドアロン・エージェントは WebLogic 管理フレームワークを通じて管理されます。これには次のメリットがあります。

- 構成ウィザードからユーザー・インタフェースを使用して構成できる
- 複数の構成を別々のドメインで維持できる
- エージェントの制御と自動再起動にノード・マネージャを使用できる

Oracle Platform Security Servicesの統合

Oracle Data Integrator では、リソースへのアクセスの制御に Oracle Platform Security Services の認可モデルを使用できます。エンタープライズ・ロールを Oracle Data Integrator ロールにマップすることで、さまざまなツールのエンタープライズ・ユーザーを認可できます。

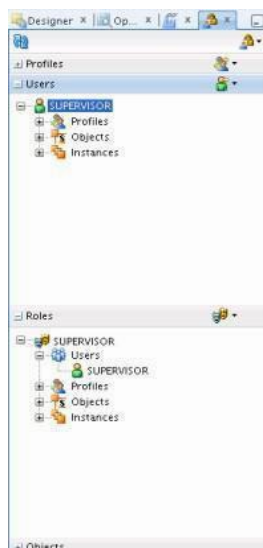


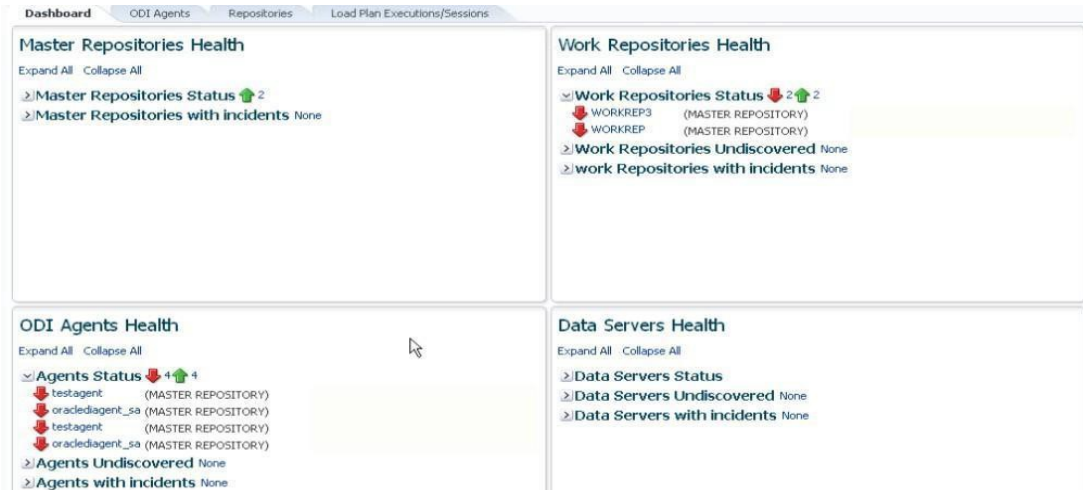
図 - Oracle Platform Security Servicesの統合

統合された管理と監視

オラクルが導入した Management Pack for Oracle Data Integrator は、Oracle Enterprise Manager Cloud Control の高度な管理機能を利用して Oracle Data Integrator 環境にトップダウン型の統合ソリューションを提供する製品です。Management Pack for Oracle Data Integrator は Oracle Data Integrator インフラストラクチャ全体に対する統合型のビューを提供することで、ユーザーが Oracle Enterprise Manager Cloud Control から一元的にすべてのコンポーネントを監視および管理できるようにします。このパックが提供するおもな機能は次のとおりです。

- 複数の Oracle Data Integrator ドメインを集中管理
- 追加設定なしでの Oracle Data Integrator コンポーネントの可用性およびパフォーマンスの監視、履歴データへのアクセス、ログの追跡、潜在的な問題に関する通知の受取り
- Oracle Data Integrator セッション・アクティビティのエンド・ツー・エンドの追跡、実行統計のレビュー、特定のステップまたはタスクから Oracle データベース・アクティビティの詳細レポートへのドリルダウン
- 堅牢でスケーラブルなアラート機能による品質保証契約（SLA）の管理
- Oracle Data Integrator のスタンドアロン・エージェントと JEE エージェントに対するリアルタイムと履歴の詳細パフォーマンス統計の取得
- Oracle Enterprise Manager Cloud Control フレームワークを使用した、Oracle Data Integrator と各種コンポーネント（データベースやその他の Oracle Fusion Middleware 製品）間での依存関係の自動的な検出とモデル化

Management Pack for Oracle Data Integrator は、11g（11.1.1.7.0 以降）バージョンと 12c バージョンの両方の Oracle Data Integrator をサポートします。



図：Management Pack for Oracle Data Integratorのホームページ

オラクルの戦略的データ統合プラットフォームへの移行

Oracle Warehouse Builderの統合

Oracle Data Integrator では、Oracle Warehouse Builder との相互運用性が向上したため、ユーザーはこれを基盤としてオラクルの戦略的プラットフォーム Oracle Data Integrator への移行を開始できます。Oracle Warehouse Builder（OWB）のジョブは、OdiStartOwbJob ツールを使用して Oracle Data Integrator で実行できるようになりました。Oracle Warehouse Builder のリポジトリは、トポロジでデータ・サーバーとして構成されます。そのため、Oracle Warehouse Builder のすべてのジョブの詳しい実行内容が Operator、Oracle Data Integrator コンソール、および Oracle Enterprise Manager にセッションとして表示されます。

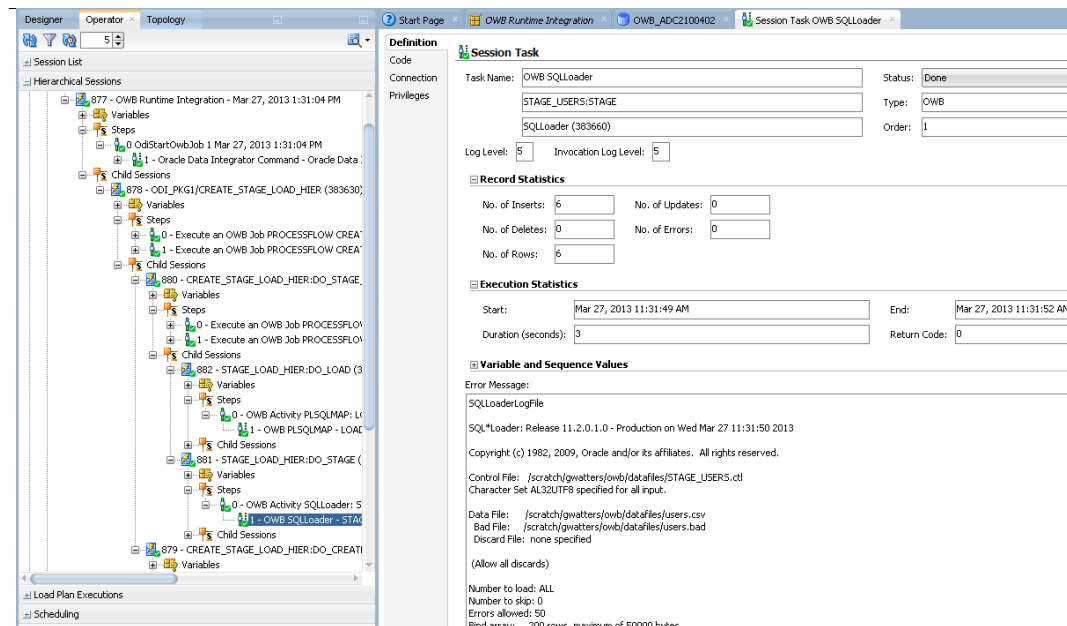


図 - Oracle Warehouse Builderの統合 - 実行時

その他の機能

XMLドライバの機能拡張

次の XML スキーマ構成メンバーがサポート対象に加われました。

- List および Union : List または Union に基づく要素は、VARCHAR 列にマップされます。
- 置換グループ : 置換グループに基づく要素からは、置換グループのすべてのタイプに対応する表がそれぞれ作成されます。
- 混合コンテンツ : 混合コンテンツを伴う要素は VARCHAR 列にマップされ、その要素のテキストとマークアップ・コンテンツがこの列に含まれます。
- アノテーション : XML スキーマのアノテーションの内容は、表のメタデータに格納されます。

一意のリポジトリ ID :

マスター・リポジトリおよび作業リポジトリでは、GUID の表記規則を満たす一意の ID が使用されるようになりました。そのため、アーティファクトをインポートする際の競合が回避され、特定のプロジェクトや組織で使用する複数のリポジトリの管理と統合が容易になります。

Studio/リポジトリのパフォーマンス強化

Studio/リポジトリのパフォーマンスが強化され、リモート・リポジトリへのアクセスにおける無駄が削減され、データベース・アクセスが最適化されています。

まとめ

Oracle Data Integrator 12c のリリースに伴い、オラクルは複数の新たな機能拡張を行いました。例としては、フローベースの宣言的ユーザー・インタフェースの再設計、再利用可能なマッピング、実行時パフォーマンスの向上、Oracle GoldenGate 統合の改良、Cloud、ビッグ・データ、XML の機能拡張、Oracle Warehouse Builder の統合などがあります。Oracle Data Integrator 12c リリースでは、オラクルの戦略的データ統合プラットフォームの改良を継続しながらも、宣言的設計、ナレッジ・モジュール、ホット・プラグ機能、E-LT アーキテクチャといった重要な製品差別化要素を維持しています。



Oracle Corporation, World Headquarters

500 Oracle Parkway
Redwood Shores, CA 94065, USA

海外からのお問い合わせ窓口

電話：+1.650.506.7000
ファクシミリ：+1.650.506.7200

CONNECT WITH US



blogs.oracle.com/oracle

facebook.com/oracle

twitter.com/oracle

oracle.com

Copyright © 2018, Oracle and/or its affiliates. All rights reserved. 本文書は情報提供のみを目的として提供されており、ここに記載されている内容は予告なく変更されることがあります。本文書は、その内容に誤りがないことを保証するものではなく、また、口頭による明示的保証や法律による黙示的保証を含め、商品性ないし特定目的適合性に関する黙示的保証および条件などのいかなる保証および条件も提供するものではありません。オラクルは本文書に関するいかなる法的責任も明確に否認し、本文書によって直接的または間接的に確立される契約義務はないものとします。本文書はオラクルの書面による許可を前もって得ることなく、いかなる目的のためにも、電子または印刷を含むいかなる形式や手段によっても再作成または送信することはできません。

Oracle および Java は Oracle およびその子会社、関連会社の登録商標です。その他の名称はそれぞれの会社の商標です。

Intel および Intel Xeon は Intel Corporation の商標または登録商標です。すべての SPARC 商標はライセンスに基づいて使用される SPARC International, Inc. の商標または登録商標です。AMD、Opteron、AMD ロゴおよび AMD Opteron ロゴは、Advanced Micro Devices の商標または登録商標です。UNIX は、The Open Group の登録商標です。0116

ホワイト・ペーパー：Oracle Data Integrator 12c の新機能 2018 年 12 月
著者：Oracle



Oracle is committed to developing practices and products that help protect the environment