

# Oracle Data Integrator : 概要

*Oracle* ホワイト・ペーパー  
2007 年1 月更新

# Oracle Data Integrator : 概要

## 概要

データは企業のもっとも重要な資産の 1 つであり、データ統合は企業における IT システムの根幹をなします。不適切なテクノロジーを使用したデータ統合は、長年にわたって有害な影響を与え、IT 予算だけではなく、企業内の重要なビジネス部門の生産性と対応力にも影響を及ぼします。

Oracle Fusion Middleware の主要コンポーネントである Oracle Data Integrator は、企業の IT インフラストラクチャに強力で信頼性の高いプラットフォームを提供します。以下に示す 5 つの重要な理由から、Oracle Data Integrator の選択が、企業にとってもっとも賢明な決断であることがわかります。

- **パフォーマンス** - 独自のアーキテクチャを利用した業界有数のアプローチ
- **柔軟性** - 複数のデータ統合形式をサポートする単一プラットフォーム
- **生産性** - 再利用性を高める宣言的な設計ツール
- **モジュール性** - 多様性をサポートするホット・プラグブルなナレッジ・モジュール
- **構想** - パフォーマンス主導型のサービス指向アーキテクチャで、プラグブルなバルク・データ統合

"Oracle Data Integrator は、データをゴールドへと変える手助けをしてくれる。"

Overstock.com 社データ・ウェアハウジング部門バイス・プレジデント、  
Jack Garzella

## はじめに

CIO や IT リーダーが抱える数多くのプロジェクトでは、IT インフラストラクチャ全体で、高速かつ正確に、柔軟性を維持してデータを移動することが要求されています。データ統合はテクノロジー（正確には、テクノロジーの集合）であり、個々の要件に応じて、さまざまなアーキテクチャを利用したアプローチでこのデータ移動を実現するものです。

2007 年の IT 最優先事項に関する調査では、多くのミッション・クリティカルなプロジェクトが挙げられています。これらのプロジェクトのすべてにおいて、充実したデータ統合ソリューション・コンポーネントが求められています。

- ビジネス・プロセスの改善プロジェクト
- ビジネス・インテリジェンス・プロジェクト
- 情報管理イニシアチブ

- セキュリティとリスク管理プログラム
- サービス指向アーキテクチャ (SOA) のロールアウト

従来の ETL (抽出、変換、ロード) や、より連携された EII (企業情報統合) を使用してこれらのデータ統合の課題を解決しようとする、結果的に行き詰まる可能性があります。

IBM DataStage や Informatica のような従来のツールを安易に買い続けるのは、もっともコストがかかり、弊害をもたらします。

Oracle Data Integrator のような最新の選択肢を検討して初めて、将来のオープンかつ標準駆動型サービス指向アーキテクチャへの準備を整え、より低コストでより高いパフォーマンスを実現できるのです。

長年、旧式のデータ統合テクノロジー・ベースに投資してきたため、慢性的な根強い問題を見過ごしやすくなっています。次のような問題が含まれます。

- 依然として多くのデータ統合環境で使用されている"ハブ・アンド・スポーク"アーキテクチャは、実装にコストがかかり、IT システム全体の速度を低下させるボトルネックとなることが多くあります。
- 多くのデータ統合プロジェクトは独立プロジェクトとして扱われたままであり、全体的に一貫性を欠いて、サイロ化しています。ほとんどの統合プラットフォームでは、データの待機時間とボリュームに対して複数の体系がサポートされていません。たとえば、1つのプラットフォームで、大容量データのバッチとトリクル・フィードのリアルタイム・データ・フローをサポートできるシステムはほとんどありません。
- トップダウン型設計が多くの統合プラットフォームの標準であるため、開発とメンテナンスにかかるコストは増大し、柔軟性の欠如によって完全なソリューションを市場に出す時間が大幅に遅れています。
- データ統合プロセスと運用システムの厳密な結合によって柔軟性が損なわれたため、更新にかかるコストはさらに増大し、現在のペースの速いビジネス環境に必要な俊敏性が得られません。

Oracle Data Integrator を使用すると、従来の IBM や Informatica によるデータ統合ソリューションとは異なり、組織はデータ統合イニシアチブのコストを削減して複雑さを緩和するとともに、データ変換の迅速化と精度の向上が図れます。Oracle Data Integrator は、データ・ウェアハウジングから、ビジネス・インテリジェンス、マスター・データ管理まで、また、リアルタイムやサービス指向アーキテクチャ・データ・サービスの配置といったすべての統合ニーズをサポートします。

### 3つの"適切さ"

適切なデータを、適切な場所で、そして適切なタイミングで取得することが、データ統合戦略を成功させるためのキーポイントです。

- **適切なデータ**：データは、その用途にふさわしいだけでなく、正確で信頼できるものでなければなりません。

"われわれが最初に驚いたのは、そのスピードです。Oracle Data Integratorのおかげで、ETLの開発が加速しました。きわめて迅速に実装と習熟が実現でき、数日でソリューションを完全に運用できたのです。"

iBasis 社プロジェクト・マネージャ、  
Leon Shaigorodsky

- **適切な場所**：総合的な情報のエコシステムは複数の運用システムや分析システムで構成されており、これらすべては、その場所に関係なく、他システムのデータを活用する必要があります。
- **適切なタイミング**：データはすぐに陳腐化します。タイムリーにデータが取得できない意思決定支援システムは役に立ちません。締め切り時間までに注文情報を取得できない発送アプリケーションは、効率的とは言えません。データの用途にふさわしい待機時間で、適切な時間内にデータを取得することは、今日のビジネスが直面するもっとも重要な課題の1つです。

## 高性能統合への革新的なアプローチ

Oracle Data Integrator は、これら 3 つの"適切さ"を実現するために必要なすべてのソリューション要素を提供します。

データベース、ファイル、パッケージ・アプリケーションなど、どのようなソース・システムやターゲット・システムにも接続できる機能によって、データがデータ統合プロセスから隔離されたまま存在するというののないようにします。また、データの整合性を維持する組込み機能によって、処理データに間違いがないことだけでなく、一貫性があり基準に準拠していることが保証されます。

適切なタイミングは、統合プロセスのパフォーマンスに関連します。Oracle Data Integrator は、高性能な ELT アーキテクチャをネイティブに実装する、市場で唯一のデータ統合プラットフォームです。ELT アプローチは、中間の ETL 変換サーバーを介してすべてのデータを移動するのではなく、ターゲット・データベース・エンジンの能力を活用して変換を実行します。これによって、パフォーマンスの飛躍的な向上と、総所有コストの大幅な削減が実現されます。データは、ソースからターゲットへ直接移動され、他のツールと比べて何倍もの速さで変換されます。

オラクルのネイティブ ELT アーキテクチャは、Informatica や IBM WebSphere によって導入された"プッシュダウン最適化"とは異なり、単なる結果論ではありません。プッシュダウン最適化と呼ばれるこれらの変換の多くは、依然として ETL エンジンの内部で実行されますが、エンジンを介してネットワーク経由でデータを物理的に転送する必要があります。プッシュダウン最適化によって、Informatica や IBM の ETL プロセスのパフォーマンスは向上する場合がありますが、オラクルのアーキテクチャは独自のパフォーマンス・クラスにあります。

適切なタイミングについてのビジネス・ニーズを満たすその他の主な機能には、Oracle Data Integrator の高度な機能が含まれます。これらには、更新されたレコードのみを転送する Changed Data Capture や、すべての統合プロセスの待機時間をバッチからニア・リアルタイム、そしてリアルタイムまで、シームレスに変更する機能があります。

## 要求に応じて発展する単一プラットフォーム

データ統合は、バッチ ETL のみに限られたものではなく、サービス指向アーキテクチャ・データ・サービスだけでもありません。また、イベント・ドリブンのリアルタイム・データ・フローだけでもありません。データ統合は、上記すべて、およびさらに多くのユースケースの組合せを含みます。多くの場合、データ統合プロジェクトは、データ・ウェアハウスの生成や、2 つの運用システム間でのリアルタイムのデータ交換などの、ある特定の要件から開始されます。しかし、統合要件は新たな方向に増大していきます。初期のプロジェクトの範囲に関係なく、

"われわれは、マニュアル・コーディングされた ETL スクリプトへの依存を軽減し、テラデータ・ウェアハウスの能力を活用してデータ変換を行う、データ統合ツールを必要としていました。Oracle Data Integrator を利用したことで、開発労力を抑えて、大量の異種データをデータ・ウェアハウスに統合する標準化プロセスを配置できました。"

Sabre Holdings 社バイス・プレジデント、  
Dan Hazel

統合要件に合わせて拡張できる統合プラットフォームを選択することは、運用を維持し、IT 投資を最大限に活用するのにきわめて重要です。

## 柔軟性を実現するキーポイント

Oracle Data Integrator は、データ、イベント、およびサービスという複数の統合形式をサポートします。さまざまなデータベース間や、エンタープライズ・リソース・プランニングやカスタマ・リレーションシップ・マネジメントなどのパッケージ・アプリケーション間、およびフラット・ファイル間でデータを直接移動できます。また、Enterprise Service Bus やメッセージ指向のミドルウェアにデータやイベントを提供することや、データ・サービスや変換サービスを提供してサービス指向アーキテクチャの一部をなすこともできます。

Oracle Data Integrator は、月単位や週単位、日単位のバッチから、時間単位、より頻繁なバッチ、ニア・リアルタイムやリアルタイムまで、すべての待機時間をサポートします。事前に定義したスケジュールや、新規データのイベントによって統合プロセスを起動します。

Oracle Data Integrator は、月単位や週単位、日単位のバッチから、時間単位、より頻繁なバッチ、ニア・リアルタイムやリアルタイムまで、すべての待機時間をサポートします。

市場にあふれている多くの"統合されていない"統合プラットフォーム（異種のテクノロジーを次々と寄せ集めて構築したもの）とは異なり、Oracle Data Integrator は、共通のフレームワーク内で、これらすべての統合形式や待機時間をサポートする単一プラットフォームです。開発のアプローチと方法は、プラットフォーム全体に組み込まれているため、共有コンポーネントを Oracle Data Integrator と Oracle Fusion Middleware で再利用しやすくなっています。

Oracle Data Integrator を利用すると、妥協することなく、すべてのデータ統合要件に対応できます。Oracle Fusion Middleware のその他の製品と組み合わせて使用すると、Oracle Data Integrator はエンドツーエンドの IT アーキテクチャの中心コンポーネントとなり、企業全体で共有できるトップレベルのデータ・サービスを提供します。

## 生産性の向上と統合リソースの削減

統合プロセスの設計フェーズは、IT インフラストラクチャ全体でもっとも重要なフェーズです。不適切な統合設計は、システム全体のパフォーマンスを大幅に低下させます。さらに、設計フェーズ自体が時間とコストのかかるプロセスになることも多くあります。

Oracle Data Integrator で導入された宣言的な設計は、従来の統合プロセス設計に関する多くの課題を解消します。この方法では、ビジネス・ユーザーが設計に参加できるため、開発者のワークロードが削減され、開発とメンテナンスにかかる時間が大幅に短縮されます。

## 設計とメンテナンスにおける生産性の向上要因

Informatica や IBM などによる、ETL 設計への従来のアプローチでは、開発者がデータ・フロー全体を設計してから、このフローのノードに変換を位置付け、コードをマッピングする必要があります。これは、プロセスに時間がかかるだけでなく、重大なパフォーマンスの問題を誘発する可能性があります。1 つの変換の配置ミスが、システム全体のパフォーマンスに大きな影響を与えるからです。

結果として、この従来のアプローチを使用して容認可能なパフォーマンスを得るには、オペレーティング・システムとデータベースのチューニングに精通した、高度なスキルを持つ統合開発者が必要になります。これらのアプローチでは、開発者がまったく必要のない作業を繰り返して行わなければならない。

たとえば、3つのソース・テーブルから1つのターゲット・テーブルにデータを挿入するとします。従来の ETL アプローチでは、データ・フローの各手順をすべて設計する必要があります。(1) 3つのソース・テーブルを接続し、(2) 1番目のソースをソートし、(3) 2番目のソースからデータのフィルタリングと検索を行い、(4) 検索の結果の変換とステージングを実行し、(5) 3番目のソースを読み取って結合し、(6) 結合の結果の変換とステージングを実行し、(7) 最後にターゲットへ書き込みます。この時間のかかる手順は、別のソースから別のターゲットへデータを挿入するたびに繰り返す必要があります。

"とてつもなくパワフルな開発環境。  
(…) われわれの開発チームは、生産性が少なくとも 50%向上しました。"

ABN AMRO Interfinance 社 IT ディレクタ、  
Leon Schurgers

Oracle Data Integrator の宣言的な設計アプローチを使用すると、3つの手順だけで同じ結果を得られます。(1) 3つのソースを1つに連結し、(2) ソース・フィールドとターゲット・フィールドとのマッピングを宣言し、(3) データ・フローを生成する組込みテンプレートを選択します。ここでは、従来のアプローチが必要な中間手順について心配する必要はまったくありません。中間手順は、Oracle Data Integrator のナレッジ・モジュールによって自動的に生成されます。

したがって、開発者は、*何*を変換するのか - 宣言的なルール - についてのみ集中すればよく、*どのように*変換するのか - プロセスの技術手順 - について考える必要はありません。別のソースから別のターゲットにデータを挿入する場合、ソースやターゲットの数に関係なく、これらの3つの手順を完了するだけです。

さらに、別のソース・テーブルを追加して上記の例を更新するとします。従来のアプローチには、次の追加手順が必要です。(1) 新しいソース・テーブルを接続し、(2) プロセス内の最適な配置場所を見つけ出し、(3) 読み取りとソートを実行し、(4) 適切なステージング結果に結合し、(5) 新しい結果の変換とステージングを行い、(6) この新しい結果を既存のプロセスに取り込みます。

しかし、宣言的な設計の場合、(1) 追加のソースを連結し、(2) ソース・フィールドとターゲット・フィールドとの新規マッピングを宣言するだけで良いのです。Oracle Data Integrator のナレッジ・モジュールによって、プロセス内の新規テーブルの最適な挿入場所が自動的に見つけ出され、適切な内部フローが生成されます。

Oracle Data Integrator を使用すると、開発チームは次のメリットが得られます。

- マッピングへの専念による、習熟期間の短縮
- 少ない手順での変換の実装による、設計の迅速化
- メンテナンス・コストの大幅な削減

動作の変更は Oracle Data Integrator のナレッジ・モジュール内に限定して保管されているため、変更による付加的影響は、Oracle Data Integrator 内のすべての変換に自動的に継承および伝播されます。

## 迅速な立ち上げと将来の豊富な選択肢

前述のとおり、Oracle Data Integrator の独自の生産性向上は、ナレッジ・モジュールと呼ばれる、データ統合のモジュール化に対する業界初のアプローチからもたらされます。ナレッジ・モジュールによって、Oracle Data Integrator は、他に例を見ないほどの柔軟性、生産性、そしてモジュール性が得られ、そのまま総所有コストの削減につながるのです。

## 再構成可能なナレッジ・モジュール層

Oracle Data Integrator のナレッジ・モジュールは、製品の統合アーキテクチャの中心にあります。ナレッジ・モジュールによって、柔軟かつ拡張可能なモジュール型統合プロセスが実現できます。ナレッジ・モジュールは、データ・フローの実装とコード生成の両方を定義するマスター・テンプレートです。

ナレッジ・モジュールは、変換ルールに関係なくデータ・フローを生成するため、汎用で使用できます。同時に、ナレッジ・モジュールはきわめて具体的です。生成するコードや実装する統合戦略は、特定のテクノロジーやアプリケーション・プラットフォームに対して細かく調整されています。

Oracle Data Integrator は包括的なナレッジ・モジュールのライブラリを提供しており、これらをカスタマイズして既存のベスト・プラクティスを実装することもできます。ナレッジ・モジュールをカスタマイズした例には、特定のデータベース・インスタンスに対するパフォーマンス・チューニング、企業標準への準拠とデータ・ポリシーの実施、既存の実行フローへの業界ノウハウの盛り込みなどがあります。

それぞれのナレッジ・モジュールは特定のテクノロジーに対して最適化されているため、適切なナレッジ・モジュールを選ぶだけで、一連のプロセスをある環境から別の環境へ移植できます。ナレッジ・モジュールは Oracle Data Integrator をホット・プラグ可能にする基盤をなすものであり、Oracle Data Integrator が、パフォーマンスのオーバーヘッドや別の ETL サーバーなしで、可能な限り豊富なプラットフォームやアプリケーションをサポートすることを可能にします。

ナレッジ・モジュールは宣言的な設計を実現する基盤コンポーネントであり、設計から、開発、実装、メンテナンスまでのエンドツーエンドのデータ統合プロセスを迅速化します。

## オラクルのデータ統合構想

要約すると、Oracle Data Integrator は、パフォーマンスと生産性の向上を目的とする独自のデータ統合プラットフォームであり、高い柔軟性とモジュール性を実現します。競合他社による従来の ETL アプローチと比較すると、Oracle Data Integrator は以下の領域で卓越した能力を発揮します。

- **パフォーマンス**：業界有数の独自のアプローチである高性能 ELT アーキテクチャによる、優れたパフォーマンス
- **生産性**：宣言的な設計による、再利用の促進と開発およびメンテナンスの迅速化

ナレッジ・モジュールは宣言的な設計を実現する基盤コンポーネントであり、設計から、開発、実装、メンテナンスまでの全般的なデータ統合プロセスを迅速化します。

- **柔軟性**：複数のデータ統合形式と待機時間をネイティブにサポートする単一プラットフォーム
- **モジュール性**：アプリケーションとデータベース・プラットフォームの多様性をサポートするホット・プラグ可能なナレッジ・モジュール

オラクルの顧客に対するコミットメントは、つねに、最高の価値をもたらす最高品質のエンタープライズ・ソフトウェアを提供することにあります。Oracle Data Integrator は、包括的なデータ統合ソリューションを提供するための基礎となるテクノロジーであり、ホット・プラグ機能と多様な IT 環境への幅広いサポートに対する Oracle Fusion Middleware のコミットメントを際立たせるものです。

IT エコシステムで使用されているデータベースやアプリケーションに関係なく、Oracle Data Integrator ソリューションを最適化して、最高性能の一括処理やリアルタイムの変換を実現できます。オラクルの構想は、今後何年にもわたって企業のビジネス価値を向上し続ける、次世代の堅牢なサービス指向アーキテクチャの中から、これらの機能を組み合わせて実現することにあります。



Oracle Data Integrator:概要  
2007年1月更新  
Yves de Montcheuil

Oracle Corporation  
World Headquarters  
500 Oracle Parkway  
Redwood Shores, CA 94065  
U.S.A.

海外からのお問合せ窓口：  
電話：+1.650.506.7000  
ファクシミリ：+1.650.506.7200  
oracle.com

Copyright © 2007, Oracle. 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、ここに記載される内容は予告なく変更されることがあります。  
本文書は一切間違いがないことを保証するものではなく、さらに、口述による明示または法律による黙示を問わず、特定の目的に対する商品性もしくは適合性についての黙示的な保証を含み、いかなる他の保証や条件も提供するものではありません。オラクル社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。本文書はオラクル社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、JD Edwards、PeopleSoft および Siebel は、オラクル社およびその子会社、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。