

# IBM Information Server概説

## Information as a Serviceの実現に向けた情報統合ソリューション

データウェアハウスの構築やマスターデータ統合、ERP連携などのさまざまな分野において、「情報統合：Information Integration」は企業を悩ませる大きな課題となってきています。そういった背景の中、情報統合を効率化するETL製品を中心に、情報統合プロジェクトをエンド・ツー・エンドで支える製品群が、多くの企業で採用され、情報統合の市場が拡大し続けています。

情報統合分野における潮流で特に注視すべき点は、ITシステムの柔軟性確保という観点から注目されるSOAへの取り組みに対応し、散在する「情報」を必要な形で必要とする上位のビジネスプロセスやアプリケーションに柔軟に提供する新たなソリューションエリアが形成されつつあることです。

これが、インフォメーション・オンデマンドの根幹を支えるInformation as a Serviceのコアとなる実現形態であり、IBM Information Serverという製品群での展開が図られています。



日本アイ・ビー・エム株式会社  
ソフトウェア事業  
インフォメーション・マネジメント  
IODテクニカル・セールス  
ITスペシャリスト

**濱野 正樹** Masaki Hamano

### [プロフィール]

アセンシャル・ソフトウェア株式会社にて、データ統合関連ソフトウェアの技術サポートやプロジェクトに従事。2005年のIBMとの経営統合に伴い現職。

### Article 3

## Overview of IBM Information Server

–Information Integration Solutions to realize Information as a Service (IaaS)–

“Information Integration” has been a major difficulty for companies which deal with projects such as Data Warehouse Implementation, Master Data Integration, and ERP Integration. Against this background, many companies have adopted solutions which provide end-to-end solutions with information integration projects, especially ETL (Extract, Transform, Load) products which increase the efficiency of Information Integration, and this has led to the ongoing expansion of the information integration market.

Against this trend in information integration, close attention should be particularly paid to the fact that a new solutions market sector has been emerging to properly manage dispersed information and to deliver it to business processes or applications in the required manner. The solution adopts the concept of SOA, based on which an increasing number of companies aim to realize a flexible IT infrastructure to keep up with the dynamically and rapidly changing business environment. This solution, named “IBM Information Server” is positioned at the core of “Information as a Service” and its parental concept, “IOD.”

### ① 情報統合におけるInformation Serverの位置付け

インフォメーション・オンデマンド(IOD)を実現するための要素として位置付けられる情報統合の分野では、Information Serverと名付けられた製品群により、新たな製品カテゴリーの創出を目指した展開が進んでいます。これには、例えばWeb Application ServerやPortal Serverなどが、IT(情報技術)システムの構成要素として一般的に受け入れられたように、Information Serverを同様な製品カテゴリーに並ぶものとして位置付けていくという狙いがあり、情報統合の分野における非常に重要な流れだと考えます。

Information Serverは、企業内に散在する多種多様なデータの意味や内容を把握し、それらを整合

性の取れた形に変換・クレンジングして、人・プロセスアプリケーションに提供する基盤となる「情報グリッド」の構築に向けた機能を提供します。なお、この領域におけるテクノロジーの多くは、主にデータウェアハウス構築や情報系システムのソリューションとして発展してきました。

データウェアハウスの構築などにおいては、多種多様なシステムから抽出したデータに対し、複雑な変換処理を行ってデータを集約することが求められます。そういった情報統合や情報管理の効率化を目指した製品の開発に多大な労力が費やされてきたことが、情報統合の分野に特化した機能や製品ラインナップの拡充につながりました。

さらに、多くの企業が実践を目指しているSOA (Service Oriented Architecture: サービス指向アーキテクチャー)における、情報管理に対する課題への対応強化が行われ、それらの製品群がInformation Serverとして用いられるようになりました。

本稿では、このようなソリューションの発展の流れに沿ってIBMの提供する情報統合ソリューションを解説します。

## 2 情報系システムのソリューションとしての発展

Information Serverのコアの製品として位置付けられるWebSphere® DataStage®は、ETL(Extract, Transform, Load)と呼ばれる製品に当たります。当

初はBI( Business Intelligence )ソリューションの一構成要素にすぎないETL製品を中心に、情報統合の分野が形成されてきました。なお、情報系システムにおけるソリューションとしてのETLは決して新しいものではありません。しかしながら、データウェアハウスの構築といった情報統合は、情報の活用に向けて多くの企業が依然として必要とする解決策であるとともに、後段で取り上げるSOAにおける情報管理に必要とされることから、IODにおける重要な適用エリアと考えます。

複数の基幹系システムから抽出したデータの一時保管庫であるODS(Operational Data Store)にデータを格納し、変換処理を行ってデータウェアハウスに格納する典型的なデータウェアハウス構築の例を図1に示します。こういったケースでは、企業内に分散した基幹系システムなどからデータを抽出するとともに、データの正規化/非正規化、コード変換、トランザクションデータとマスターデータのひも付け、データ項目のマッピング、不正データ検出やデータ補正などのさまざまな変換処理を実装し、データウェアハウスにロードする必要があります。抽出・変換・ロードといった各処理の実装からハンドコーディングの作業を排し、GUI(Graphical User Interface)上で部品による開発を行うことで、開発の効率性を向上することがETLの主眼となります。またクレンジング処理として顧客データや法人データの名寄せ処理を組み込むケースも多く、IBMではWebSphere QualityStage™を名

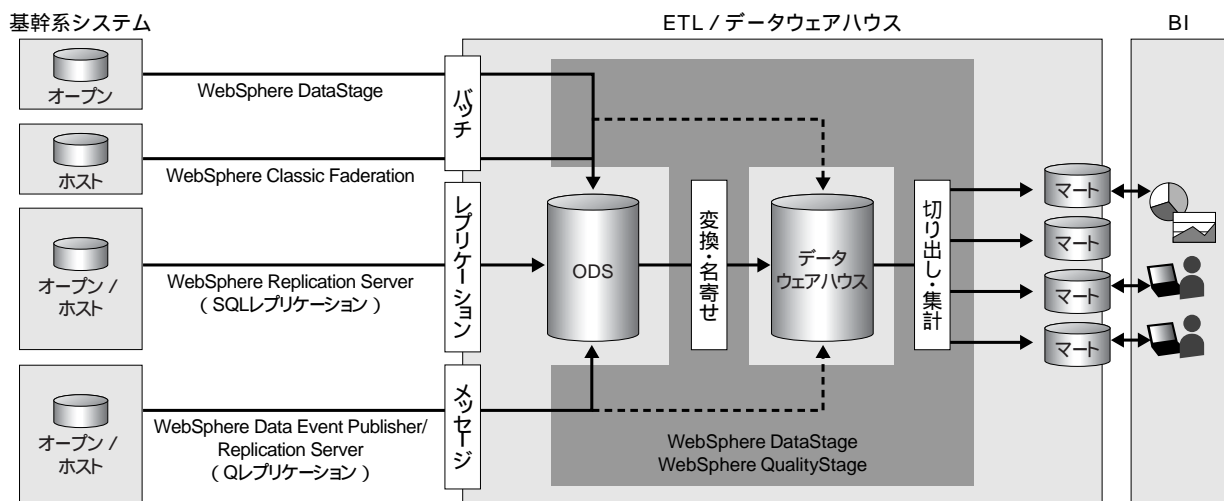


図1. データウェアハウスの構成例

寄せツールとして提供しています。

さらに情報統合基盤の構築を総合的に実現するという観点から、従来はデータベースの一機能やアダプターと位置付けられてきた製品も、上記製品と並びInformation Serverにラインナップされています。データの抽出に当たってはそのリアルタイム性の要件などから、一般的に「バッチ」「レプリケーション」「メッセージ」の三つの方式の採用が考えられます。WebSphere DataStageはリレーショナルデータベースやERPシステム(Enterprise Resource Planning: 統合基幹業務システム)などからバッチ方式でデータの抽出を行う機能を提供していますが、WebSphere Classic Federationとの組み合わせにより、ホスト上のIMS™(Information Management System)などの階層型データベースやVSAM(Virtual Storage Access Method: 仮想記憶アクセス方式)などからもデータの抽出が可能となります。また、更新差分データのみを抽出する仕組みとしてWebSphere Replication ServerによるSQL(Structured Query Language: 標準照会言語)レプリケーションを活用する実装や、よりリアルタイム性を重視したメッセージ方式の実装として、MQ(Message Queue)ベースのQレプリケーションやWebSphere Data Event Publisherとの組み合わせも考えられます。

### 3 情報統合プロジェクトを支える製品群の拡充

ここまで紹介してきたETLなどの製品は、IT基盤として適用される製品ですが、これらに加えて図2に示したように、情報統合のプロジェクトの「分析」「設計」「開発」「テスト」といったライフサイクルをサポートする製品も拡充され、Information Serverにラインナップされています。また、プロジェクトにおいて各フェーズで役割の異なる担当者が、以下に紹介する製品群を使い分けることを想定し、製品間でのメタデータを共有し、エンド・ツー・エンドの整合性と一貫性の確保を可能にしています。

まず「分析」や「設計」のフェーズを支える製品としてRDA(Rational® Data Architect)やIA(WebSphere Information Analyzer)があります。「どの」データを

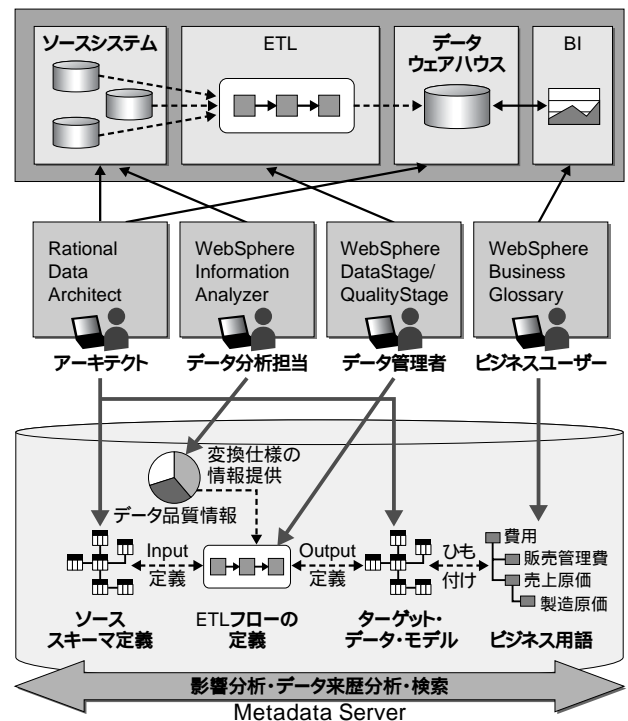


図2. 情報統合のライフサイクルを支える製品群

「どのように」抽出・変換する必要があるのか定義するには、データ抽出の対象となるソースシステム上のデータ構造や品質の把握が必要となります。RDAはデータベースやDDL(Data Definition Language)などからリバースエンジニアリングを行い、「どの」データ項目を抽出すべきかの判断に役立つソースシステムのデータ構造の全体図を提供できます。また、RDAはターゲットの論理モデル・物理モデルの設計を行うユーザーインターフェースを提供します。それに加えて、ソースデータを「どのように」抽出・変換してこのターゲットに格納すべきかの判断材料となるデータ品質情報を、IAによりレポート出力できます。

例えば、「性別」という項目には「M/F」「Mr./Ms.」「1/2」などの形式でデータが格納されていることが想定されますが、複数のデータソースから「性別」の項目を集約し表記形式を統一するには、実際にデータを分析しなければ、どのような変換処理を行うべきか判断できません。IAは実データもしくはサンプリングデータから、どのような表記形式のデータが、どの程度のパーセンテージで含まれているかというドメイン分布に加え、データの欠落、冗長性、関数依存関係などのデータ品質を分析できます。

その上で「開発」や「テスト」のフェーズを支える製品としては、既に紹介したWebSphere DataStageやQualityStageがあります。これらの製品で整合性の取れた品質の高いデータをターゲットに格納するための変換・クレンジング処理を実装する形になります。RDAで作成されたデータモデルやIAのデータ品質情報は、統合のメタデータ・リポジトリであるMetadata Serverに格納されており、作成済みのデータモデルをメタデータとしてDataStageのETLプログラムにロードし、テーブル定義を部品として使用したり、クレンジング処理の実装に当たって各データ項目のデータ品質情報を参照することなどが可能です。各フェーズのツールで作成・管理されるメタデータをツールごとに管理した場合、メタデータに不整合が発生しプログラム品質や管理効率の低下を招くことが考えられますが、Metadata Serverの統合メタデータ管理により不整合を排除し、効率性を向上させることができます。

またWebSphere Business Glossaryはビジネスユーザーが理解可能な「ビジネス・メタデータ」のディクショナリー管理を行うWebインターフェースを提供します。例えばデータベースのテーブル項目名として使用されているテクニカル・メタデータのAR (Accounts Payable) ビジネス・メタデータである「売掛金」という用語にひも付けて管理することで、ビジネスユーザーとITユーザー間のコミュニケーションの円滑化を図ることができます。また、ETL処理などの実行情報はオペレーショナル・メタデータとして管理されています。この実行情報を基に、例えば「AP」というデータウェアハウス上のテーブル項目に格納されている情報は、システム上の「ACCOUNTS\_PAYABLE」という項目から抽出され、A、B、Cという三つのETL処理を経て格納されたという「データ来歴」を分析できます。こういったトレーサビリティの確保により、ビジネスユーザーからの仕様変更要求やデータ不備に対する追跡と修正の要求などに対し、ITユーザーはより迅速なフィードバックと対応が可能になります。

#### 4 IODの実現に向けたIaaSへの展開

データウェアハウス構築への適用を中心にInformation Serverの製品群を紹介してきました。これらの製品は情報系システムのソリューションを超えて、例えばインフラの合理化や、企業統合に伴うシステム統合、マスターデータ統合、ERP連携などのさまざまな情報統合のエリアにも適用されてきました

主にメッセージングによるアプリケーション連携のための製品としてEAI (Enterprise Application Integration) が多くの企業で受け入れられてきました。ここまで見てきたように、ETLは分散するデータの統合にかかわる課題を総合的に扱うソリューションを提供するようになったことと、より広範なエリアで適用されるようになったことに伴い、図3に示したように、アプリケーションが扱うデータを統合・連携させる基盤ソフトウェアとして位置付けられました。このような製品群は一般的にEDI (Enterprise Data Integration) という名称で呼ばれてきました。

これに対して、主に異種混合のデータベースの活用という観点から、データの所在場所を意識せず、あたかも一つのデータベースを扱っているように異種混合のデータソースにアクセスさせるフェデレーションの技術が発展してきました。このようなエリアはEII (Enterprise Information Integration) という名称で呼ばれ、IBMはWebSphere Federation Serverという製品を提供しています。Information Serverはこれらの技術を統合プラットフォームで実現してSOAへ適用範囲を拡大し、IaaS (Information as a Service) という考え方に基づいて情報のサービス化を行います。

図4に示したように、Information Serverは、シス

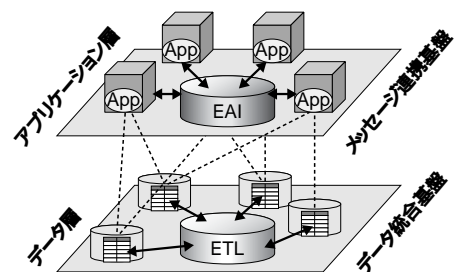


図3. ETLとEAIのすみ分け

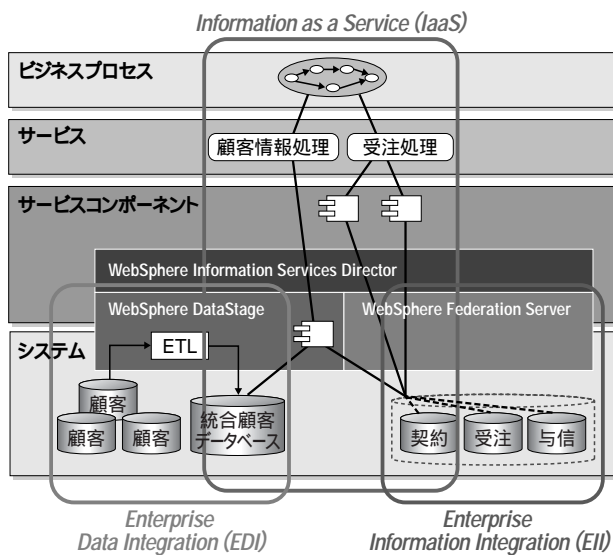


図4. SOAにおける情報サービスの位置付け

システムレイヤーに位置する既存IT資産の情報の所在や、プラットフォームの種別などを仮想化する情報サービス化のレイヤーを提供します。これにより上位のレイヤーに配置されるサービスコンポーネントと使用する情報との疎結合化を確保するとともに、情報の活用に向けた変換ロジックの集中管理と再利用化を促進できます。

例えば受注のプロセスを考えた場合、契約・受注・与信・在庫などの管理システムとの連携が想定されます。これらのシステムが異種のプラットフォームやデータソースに格納されていると、情報の管理ポイントが分散するとともに、情報アクセスの形態はプラットフォームの種類などに依存した密結合とならざるを得ない場合があります。このようなケースではWebSphere Federation Serverにより複数のデータソースを仮想化して、一つのデータソースであるかのように見せて対応することが可能です。

また例えば、顧客情報処理のサービスをつくるケースを想定すると、オペレーターが入力した顧客住所情報の国土地理協会・町字ファイルに基づいた表記形式の標準化や、登録顧客情報が既存の顧客情報と重複する場合のワーニング出力など、クレンジングを行う要件が発生することがあります。このようなサービスコンポーネントの汎用的なプログラムでの実装は容易ではありませんが、こういった変換・クレンジング処理はDataStageとQualityStageの組み合わせで効

率的な実装を行うことができ、その実装プログラムをWebSphere Information Services Director経由でサービスコンポーネントとして公開・共有化し、効率的な情報の活用と集中管理を実現できます。

ただし、例えば顧客マスターが複数のシステムで管理されている場合など、SOAの実装やサービスの再利用性について課題があるときには、従来のEDIのアプローチで統合顧客マスターの作成を検討する必要があるかもしれません。また、複数の顧客データベースをあたかも一つのデータベースであるかのように仮想化するEIIのアプローチが、ワークロードの観点から最適かもしれません。いずれのアプローチを採ることができるのもInformation Serverの大きな価値といえるでしょう。

ITユーザーと業務を担当するビジネスユーザーとの協調を促進して、劇的に変化するビジネス環境に俊敏に対応していくことがSOAのビジネス上の価値であることを考えると、テクニカルな情報属性をビジネス用語に視覚化するメタデータ管理も、非常に重要な要素として位置付けられます。

以上のように、情報統合の分野で発展してきた製品群は、Information ServerとしてSOAにおける価値を提供する形で活用が進んでいます。

[ 参考文献 ]

- [ 1 ] Philip Howard: SOA and Information Services, Bloor Research (2006.3).
- [ 2 ] Norbert Bieberstein, Sanjay Bose, Marc Fiammante Keith Jones, Rawn Shah(株)テックバイザー・ジェイビー 訳, 清水敏正監修: SOA実践ガイドブック, ISBN4-7981-1241-0 (2006.11.15).