

**XML Data Integration**  
Feature Story

特集

# XMLデータ統合

## XML DB とデータ統合環境

XMLが最初に市場に登場してから10年以上が経過した。その間、XMLに対する評価は大きく変化してきた。複数媒体への対応が可能な文書作成用言語のSGMLをもとに開発されたため、当初はドキュメント分野の技術として見られる傾向にあったが、現在では柔軟なデータの管理と統合を可能にする基盤技術として位置づけられている。

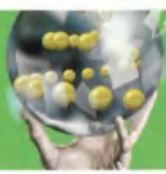
早期からその技術的優位性を理解していたベンダーやユーザー企業はそれほど多くなく、XMLに対するユーザー側の理解度や市場での認知度の低さと、それに伴う技術者や対応製品の少なさが、XML技術の導入・普及を妨げる大きな要因になっていたが、なかでも最大の障壁は、XMLを解釈するパフォーマンスにあった。

2000年頃から、柔軟なデータ管理を実現するXMLデータベースも登場したが、処理性能は依然として大きな課題だった。Webアプリケーションが普及するにつれてWebサービスがアプリケーション統合の中核技術として認識され始めると、XMLは再評価され、仮想化技術の進展やプロセッサ性能の向上によるプラットフォームの性能向上とも相まって、データ管理・統合を実現する有力な技術基盤になった。

いまやXMLは、非構造化データの急増と複雑化したシステム環境への対応をはじめ、情報の戦略的利用、アプリケーションのマッシュアップ型開発、市場や環境の変化に対する即応性など、現在の企業で必要とされる多くの課題に対して、XMLは効果的なソリューションを提供すると見られている。今回の特集では、これまでその実力が過小評価されてきたXMLについて、現在の市場概要を紹介し、現在市場に提供されている主なXMLデータベース製品とデータ連携ツールについて紹介する。

第1部 市場概要.....P26

第2部 主要製品動向.....P28



# 急成長するXMLデータベース市場 柔軟なデータ構造が企業の課題を吸収

XMLデータベースが普及を加速しそうな兆しを見せ始めている。これまでデータの検索・抽出といったデータベース管理システム(DBMS)の基本要件でパフォーマンスの低さが指摘されてきたが、技術の進展と製品の成熟、プラットフォームの性能向上によって克服されつつある。この数年で企業内に急増した非構造化データに対する管理ニーズの高まりが、XMLデータベースの需要を押し上げている。連携ツールの登場により、XMLデータベース自体の適用領域が、従来の電子商取引におけるデータ交換から、社内情報資産の効率管理・活用や社内外システムとのデータ連携へと拡大し始めた。

## 情報系のデータベース

これまで企業は、サーバ統合やアプリケーション統合、仮想化技術など利用してITシステムのインフラを簡素化し、複雑性を低減させてきた。これらのシステムは、ほとんどがWebアプリケーションを中心に利用される情報系のプラットフォームで、処理量が増加するにつれ、比較的安価なサーバを増設するスケールアウトの手法によって、処理性能が向上されてきた。

このため、管理に必要なサーバが増加し続け、運用・管理に費やす手間とコストが増大していた。また、各部門ごとに設置されるファイル・サーバなどもシステム環境を複雑化する要因の1つになっており、これらをシンプルな環境に統合できれば、それだけで運用・管理コストは低減できた。

これに対して、企業の人事・給与や財務・管理などの業務アプリケーションで利用されるプラットフォームには、高い拡張性と可用性、信頼性を備えた基幹系サーバと呼ばれる大型のサーバが導

入され、ERP(経営資源管理)やSCM(サプライ・チェーン管理)、CRM(顧客関係管理)といった業務パッケージとリレーショナル・データベースによって、日々の定型業務を処理していた。

基幹系システムの領域では、性能とともに可用性と安定性が重視され、いったん稼働し始めると、新しく登場した技術をすぐに導入したり、構成要素の変更やシステム改変が行なわれることはほとんどない。こうした形態は、製造業や流通業、金融業などの分野で一般的だが、コンテンツ/情報提供サービスやインターネット接続サービスを中核事業としている業態では、情報系システムがそのまま基幹系システムになるケースもある。

新製品や新サービスの市場投入などで営業成績や販売売上、顧客数が増減することはあっても、従業員の数や管理部門の業務量が大きく増減することはないため、基幹系システムで利用されるデータベース管理システムには、構造化されたトランザクション処理データ(定型データ)を、いかに高速に格納・

検索・抽出できるかが求められてきた。現在のリレーショナル・データベースは、柔軟な拡張性やデータ構造を備えつつあるが、その基本は定型データを高速・適切で安定して管理できるかにある。

その一方で、情報系システムのデータベースには、高い柔軟性が求められる。極論すれば、リレーショナル・データベースで処理しないデータ、処理できないデータすべてに対応する必要がある。情報系システムでは、データベースにも、アプリケーションとプラットフォームと同様の変化への対応力と高度な柔軟性が求められるが、これがXMLデータベースの需要を喚起する要因の1つになっている。

## 非構造化データへの対応

現在企業では、電子メールや画像データ、音声データ、ワープロ文書、表計算シート、営業資料などの非構造化(非定型)データが急増しており、デスクトップPCのハードディスクや部門のファイル・サーバに格納され、個人や部門

## 国内XMLデータベース市場規模推移



出典:富士キメラ総研「2006/パッケージソリューション・マーケティング便覧」(2006年4月発行)

レベルで管理されているだけの企業も少なくない。従来はそれで問題なかったが、いまやこれらのデータは、情報漏洩に対するリスク管理やセキュリティ対策、あるいはコンプライアンス対応に伴う内部統制体制によって、適切に管理しなければならない。ネットワーク環境の進展によって電子メールによる取引や業務が増加し、電子メールの重要性が高まると同時に、法的な保管義務も課せられるようになり始めた。また、顧客の属性ごとに新しい商品やサービスを開発したり、マーケティング活動などで収集した顧客情報や企業のWebサイトに寄せられた顧客の声を、製品の開発部門やサービスの企画部門に反映させて、市場ニーズに対応した製品を迅速に開発しようとする状況を考えると、非構造化データの管理はますます重要になる。

このような要望に対して、リレーショナル・データベース管理システムをもとにして、データ変換ツールや連携ツールで

カスタム・アプリケーションを構築したり、基幹業務パッケージ製品にミドルウェアを組み合わせれば対応できるが、適用できる領域や業務、データベース構造、製品やシステム構築コスト、他システムとのデータ連携などを考慮すると、XMLデータベースを導入した方が容易に実現できるだろう。

また、非構造化データはコンテンツ・データベースやビジネス・インテリジェンス(BI)ツールでも扱えるが、これらはデータの利用目的が明確で、本来の目的を外れるような使い方を想定して設計されていない。

## 融合するシステム環境

変化への柔軟な対応が必要とされるのは、企業を取り巻く環境が多様化し、従来の枠組みでは対応できなくなっているためだ。企業の基幹系システムでは、在宅勤務や出向・派遣などの勤務形態の多様化に対応した人事給与管

理や、協業や提携を前提としたサプライ・チェーン管理、国際会計基準や監査基準に対応した財務会計などへの対応が求められ、情報系システムでも、顧客の要望や市場のニーズの収集・分析、顧客情報の事業への反映、事業に連動した情報発信、パートナー企業との情報連携などへの対応が必要になるなど、従来のような基幹系業務と情報系業務の明確に分けることが難しくなっている。また、本来の事業に経営資源を集中し、それ以外の業務を外部に委託するなど、企業の内外を分ける境界も曖昧になってきた。

こうした状況が、現在のXMLデータベースの需要をさらに加速している。

## 適用領域の拡大

富士キメラ総研が昨年4月に発表した国内XMLデータベース市場の調査結果によると、2005年度の市場規模は対前年度比50%増の9億6,000万円となった。また、2006年度は同87.5%増の18億円に達する見込みだ。2007年度はさらに成長し、同105.6%増の37億円と予測され、2005年度からの3年間で市場は4倍に拡大する。

XMLに対する認知度の向上を背景に、XMLデータが電子商取引システムでのデータ交換用途だけでなく、企業内の非構造化データの格納・管理に利用され始めたようだ。XMLデータに対する注目の高まりと需要拡大によって、XMLデータベース市場は今後も高い成長率で拡大し、2010年度には100億円を突破して120億円となり、2011年度には150億円に達すると予測されている。



## 鍵となるソリューション連携 適用領域を拡大するアプリケーションへの対応

XMLデータベースの商用製品は、XMLの最初の仕様がW3C (World Wide Web Consortium) から1998年2月に発表された翌年には市場に登場した。現在と同様にXMLの柔軟性がアピールされたが、格納できるデータの量や処理性能がリレーショナル・データベース(RDB)に劣り、XML技術を扱える技術者がほとんどいなかった。現在ではこうした状況が解消されつつある。現在市場には、2003年頃に市場投入された第2世代製品と、最近になって登場した第3世代製品が混在する。製品を提供するベンダーは国内と海外ベンダーを合わせても10社を超える程度で、市場の拡大を見込んで新規参入したベンダーもある。

### 実用に耐え得るレベルに

Web技術の標準化を推進するワールドワイドウェブ・コンソーシアム(W3C: World Wide Web Consortium)がXMLの最初の仕様「XML Version1.0」を発表したのは1998年2月。インターネットが一般の企業にも普及し始めた頃である。XMLは汎用のデータ記述言語「SGML (Standard Generalized Markup Language)」をもとに、W3CがWeb上で利用できるように開発した。

SGMLの特性を継承しているため汎用性が高いが、人間に理解しやすい形式でデータが記述されることから、コンピュータにとっては処理が重い。これがXMLデータベースの最初の克服すべき課題になった。また、HTMLのタグやリンク機能の考え方を取り入れているが、データと表示方法を分離してコンピュータとの連携性を高めている。

XMLデータを直接読み書きできるXMLデータベースは、XMLの仕様が発表されてから1~2年後に登場した製品を第1世代とし、性能と機能を向

上して2003年頃に登場した第2世代、現在登場している製品は第3世代に分類される。現在国内で提供されているほとんどの製品は第2世代製品だが、最近になって新版の投入が相次いでいる。

XMLデータベースは、市場に第2世代製品が登場してバージョンアップが何度か繰り返されたことで、ようやく実用レベルに達した。またその間に、XMLに対する市場の理解が進み、XML技術者の数も徐々に増えるなど、普及に向けた必要条件が揃ってきた。

### XMLデータベースの適用分野

富士キメラ総研の調査によると、国内XMLデータベース市場は2004年度に6億3,500万円となり、そこから対前年度比で50%を超える成長率で拡大し始めている。非構造化データの増加に伴うデータ管理需要の変化と、XML技術の利用拡大に伴う適用領域の拡大が大きな理由だ。

また、ミック経済研究所の「UNIX-

Windows対応ミドルウェアパッケージソフト市場展望2006年度版」によると、XMLデータベースは出荷金額で2005年度が対前年度対比96.5%増、2006年度が同25.1%の高成長が見込まれている。

国内でXMLデータベースを提供しているベンダー各社は、具体的な適用分野や事例の紹介を積極的に行なっている。これは、柔軟なデータ連携や統合を可能にするXMLデータの高い汎用性が適用分野を選ばないためで、そのことが、具体的なソリューションをイメージしにくい要因になっているからだ。

XMLデータベースで全社のデータを統合し、一元管理するという本来の活用方法以外にも、XMLデータベースを中心に基幹系システムを構築したり、データ連携ツールやXML開発・実行環境と連携することで、アプリケーション開発からデジタル・パブリッシングまで広範な用途に対応できる。

現在ベンダー各社が提示・推奨するXMLデータベース・ソリューション例としては、以下の分野がある。

### ●メタデータの管理

商品情報や素材管理、顧客管理、Webカタログなどの分野で、メタデータ(個々の情報やデータについての情報)をXMLデータベースで管理することによって、データの検索性と再利用性を高める。

### ●システムへの組み込み

XMLデータベースをコンテンツ管理システム(CMS)に組み込み、非構造化データを含むすべてのデータを一元管理する。ワークフロー管理システムにXMLデータベースを組み込み、ワークフローの変化に対応する。

### ●マニュアルの作成・管理

同じ構成が繰り返し使われるマニュアルの作成・管理にXMLデータベースを導入し、検索性と再利用性を高める。

### ●デスクトップ・パブリッシング

定型文書を大量に作成するデスクトップ・パブリッシング(DTP)にXMLデータベースを導入し、文書作成を自動化する。

### ●XML自動組版システム

XML対応のページレイアウト・ソフトとXMLデータベースを組み合わせ、XML自動組み版エンジンやXML自動組版システムを構築する。

### ●フォーム入力

フォームから入力したデータをXMLデータベースで蓄積し、再利用する

### ●システム間連携代理システム

XMLデータベースを中心として、Webアプリケーションと既存の業務システムを連携させる。

### ●財務レポート

XMLデータベースでXBRL(eXtensible Business Reporting Language)データを管理し、財務レポートを作成する。

### ●株価情報通知のパーソナライズ

株価売買情報を管理するバックエンドのデータベースとXMLデータベースを連携し、個人ごとの通知ルール・データをXMLで保持し管理する。

### ●G-XMLによる地図情報システム

G-XML(XMLで表現された地図情報)とXMLデータベースによって地図情報システムを構築する。

### ●未活用定型データのマイニング

社内に大量に蓄積された未活用の定型データにタグを付けてXML化し、XMLデータベースで管理する。

## 市場の主なプレーヤー

現在国内でXMLデータベース製品を提供しているのは、サイバーテック(Cyber Luxeon)、セック(Karearea)、ソニックソフトウェア(Sonic XML Server)、ソフトウェア・エー・ジー(Tamino)、東芝ソリューション(TX1)、日本アイ・ビー・エム(DB2 9)、日本オラクル(Oracle XML DB)、富士通(Interstage Shunsaku Data Manager)、三井物産・三井情報開発(NeoCoreXMS)、メディアフュージョン(EsTerra XML Storage Server)といったベンダーで、製品にそれぞれ特徴がある。

サイバーテックのCyber Luxeon、セックのKarearea、東芝ソリューションのTX1、富士通のInterstage Shunsaku Data Manage、メディアフュージョンのEsTerra XML Storage Serverは国産のXMLデータベースだ。

また、セックのKareareaは、XMLデータをメモリ上に展開して、独自の超高速データ処理エンジン(LFM)で高

速処理を行なうインメモリXMLデータベース。LFMはXMLデータをすべて主記憶上に展開して処理する。展開されたデータは独自のコンパイル処理によって非常にコンパクトに格納されるため、搭載するメモリ容量が比較的小さいPC(256MB以上)でも軽快に動作する。

また、ソニックソフトウェアのSonic XML Serverは、オブジェクト指向データベース「ObjectStore」をコアに持つネイティブXMLストレージで、エンタープライズ・サービス・バス製品「SonicESB」のアドイン・コンポーネントとして提供され、高速のXML処理、ストレージ、問合せサービスを提供する。

さらに、三井物産・三井情報開発は2004年5月からNeoCoreXMSの無償版「Xpriori(エクスプリオリ)」を提供している。データベース・サイズの上限が約1GBで、データベースをインストールしたコンピュータと同じIPアドレスからのアクセスのみ利用可能、などの制限はあるが、NeoCoreXMSと同等の機能が利用できる。

また、メディアフュージョンは2005年12月に、日本で初めて開発されたXMLデータベース「Yggdrasil(イグドラシル)」のソースコードをGPLに基づいて無償で公開している。

このほか、XMLデータベースではないが、ジャストシステムがXMLデータベースと連携してXMLアプリケーションの開発・実行を可能とするXML製品の提供している。それが統合XMLアプリケーション開発・実行環境「xfy Enterprise Solution」で、IBM DB2 9対応製品とオラクルのOracle Database 10g対応製品を用意している。



## 日本市場での地歩を固める Tamino

ソフトウェア・イー・ジーは昨年12月に設立された独ソフトウェア・アーゲーの日本法人。日本市場には1974年に参入し、代理店経由で事業を展開してきたが、先進的な日本市場でのプレゼンス向上と直接的な顧客へのサポートを目的として日本法人を設立した。同社は1969年に設立された欧州最大手のシステム・ソフトウェア・ベンダーで、1999年に世界に先駆けてネイティブXMLデータベース「Tamino(タミーノ)」を発表した。

国内では2000年4月から販売開始しており、独ソフトウェア・アーゲーの最高経営責任者、カール・ハインツ・ストライビツヒ氏は、全世界で50%以上の市

場シェアを獲得している。日本にもすでに500社以上のユーザー企業に導入された。われわれにとって、今後も日本でのビジネスを拡大させるための重要な製品」と語っている。

Taminoは、標準のインターネット技術に基づくデータ管理を目的とした高性能XMLサーバとして提供され、構造化/非構造化データをサポートし、XMLドキュメントを処理、生成、交換するために必要なすべての機能を提供する。

Taminoの“XMLサーバ・コア”は、XMLサーバを稼働するため基盤を提供するもので、バッファ・プール管理、内部XMLパーサ(認証/非認証)、データマップ(スキーマ・コンテナ)、オブジェクトに関連するサーバ拡張/トリガ、リアルタイム結果変換、多様なインデックス・

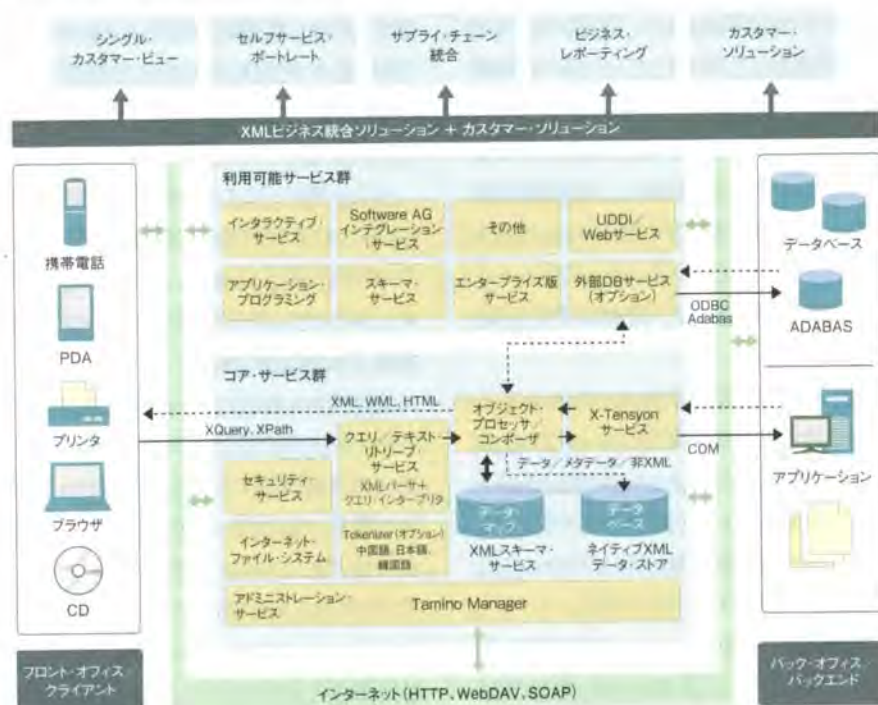
タイプ、高性能なディスク管理、2フェーズ・コミット、バックアップ、災害復旧リストアなどの機能が提供される。さらに、サーバ拡張や外部データベースへの接続が可能で、必要に応じたサービスを増強・利用できるように設計されている。

“ネイティブXMLデータ格納”では、HTTPベースのDOM、JDOMやSAX API経由でアクセス可能で、整形形式のXML(Well-formed XML)/妥当なXML(Valid XML)/非XML(Non XML)のドキュメント、メッセージ、メタデータのストレージ、自動バージョンングをサポートする。

このほか、“クエリ/テキスト検索サービス”“アドミニストレーション・サービス” DTDに基づくスキーマとXMLスキーマの作成を可能にする“スキーマ・サービス”や“セキュリティ”などの機能が提供される。

日本市場の事業方針についてストライビツヒ氏は「すでに200社以上のユーザー企業を日本に抱えている。ユーザーのIT環境をさらにレベルアップし、イノベーションを提供していくことがわれわれのゴールだ」としている。

Tamino XML Serverアーキテクチャ



出典:ソフトウェア・イー・ジー

## 協業による連携ソリューションを推進 NeoCoreXMS

三井物産が国内XMLデータベース市場に参入したのは2002年11月。米国ネオコア社と日本における独占販売契約を締結し、ネオコア社の開発したXMLデータベース「XML Information Management System (XMS)」を発表した。販売は関連会社の三井情報開発と提携し、共同で行なっている。

また、製品の販売・サポートについては、両社が共同で運営する「Mitsui NeoCore Center (ミツイ・ネオコア・センター)」で行なう。

NeoCoreXMSは、国内出荷数が現在までに500ライセンスを超え、国内XMLデータベースとしてはシェア第1位(富士キメラ総研調べ)を占めている。また、昨年11月に三井物産と三井情報開発が発表したところによると、NeoCoreXMSはミック経済研究所の調査でも出荷本数と出荷金額で第1位のシェアを獲得したという。NeoCoreXMSがシェア第1位になったのは2004年からで、今回で3年連続になる。両社では、NeoCoreXMSのサービスの質と連携ソリューションの数が市場に認められた結果としている。

NeoCoreXMSは、エクスプリオ社の独自検索技術“DPP (Digital Pattern Processing)”を実装することで、データ量やデータ構造に依存しない超高速検索を実現する点を特徴とし、1GBと100GBのデータ量での検索性能に差がないという。さらに、スキーマの定義が不要で、すべてのXML情報を柔軟に格納するほか、すべてのXMLタグを自動的にフル・インデックス化する。

両社によると、NeoCoreXMSの利用形態では、XMLデータベースとドキュメント・データベースの2つの活用方法に大別できるとしている。

XMLデータベースとしての活用では、企業内に蓄積されたメタ情報や属性情報をNeoCoreXMSに格納し、商品情報管理やコンテンツ管理などに活用したり、CMSなど組み込み、パッケージ・ソフトのデータベース・エンジンと

## XML管理システム



して活用されることが多い。また、ドキュメント・データベースでは、XML/SGMLで管理するマニュアル情報をNeoCoreXMSに格納し、各種マニュアルのデータ管理をアドビのInDesignやFrameMaker、マイクロソフトのMS-Officeなどと連携利用する。

連携ソリューションを推進する三井物産・三井情報開発は、昨年度(2006年4月～2007年3月)も他社との協業を積極展開した。昨年8月には日立システムアンドサービス(日立システム)とインフォテリアを含めた4社で、小売・卸業向けのソリューション展開で協業することに合意し、「商品情報エンタープライズDB」の提供を開始した。商品情報エンタープライズDBは、日立システムが提供するデータ連携ソリューション・スイート「ASTERIA (アステリア)」とNeoCoreXMSを組み合わせ、現場担当者に商品をキーにしたセールス・ナ

レッジを提供するための基盤を構築するもの。

基幹システムに存在する商品情報や在庫情報などの定量的な定型情報と、CRMやSFA、コールセンター・システムなどに蓄積された顧客の商品評価・販売上の注意点・クレーム情報・セールストーク・関連購買製品情報などの定性的な非定型情報をASTERIAで収集し、NeoCoreXMSに格納して商品情報統合データベースを構築する。

また9月に両社は、デジタルコンテンツ・メタデータ管理ソリューション分野で、住友電気情報システムとネクストコムと業務提携し、販売促進活動を共同で推進していくことで合意。これに基づき4社は、NeoCoreXMSと住友電気情報システムが提供する類似情報検索エンジン「QuickSolution」を基盤とし、ネクストコムがシステム・インテグレーションを担当して、デジタル・コンテンツ・メタ





データ管理ソリューションを提供する。

さらに、昨年12月にはプロフィールドと、XML技術によるCMSのパッケージ・ソリューション「XCMS」の販売・販促活動で提携した。NeoCoreXMSとプロフィールドが開発したXCMSを組み合わせて提供する。XMLデータベースをエンジンに採用したCMSソリューションは国内初となるもの。これにより、クライアント側での情報構造の変化にCMSが応答し、データベースを自動生成できるようになる。

### 第3世代の“ハイブリッドDB” DB2 9

日本アイ・ビー・エム (IBM) は昨年7月、XMLデータベースとリレーショナル・データベース (RDB) を融合した“ハイブリッド・データベース”として「DB2 9」を発表した。開発コード名で“Viper (バイパー)”と呼ばれていた製品で、XMLデータベースの技術をリレーショナル・データベース (RDB) のインフラ上で実装した。

DB2 9は、XMLデータとリレーショナル・データを1つのアプリケーションから活用できる第3世代の“ハイブリッド・データベース”として位置づけられている。XMLデータの処理に最適な格納方式を採用し、RDBの基本機能も大幅に拡張した。これにより日本IBMでは、XMLデータベースとRDBの二重投資が不要になるとしている。

DB2 9では、リレーショナル・データはリレーショナル・ストレージに格納され、XML文書は挿入時にパーシング (構文解析) されて、その階層構造を保持したままXMLストレージに格納される。

データを取り出す際には解析が不要なため、高いパフォーマンスを実現する。また、名前やIDなどの要素や属性に対して、必要なものについてのみ索引を作成できるため、検索のパフォーマンスが向上する。

さらに、DB2 9ではXML文書をRDBの列にマッピングすることなくそのまま格納するため、XML文書の形式の変更に合わせてテーブルの設計を変更する必要がない。データ領域を作成する際に、データ型をXML型に設定しておくだけで、ツリー形式のデータ構造になる。

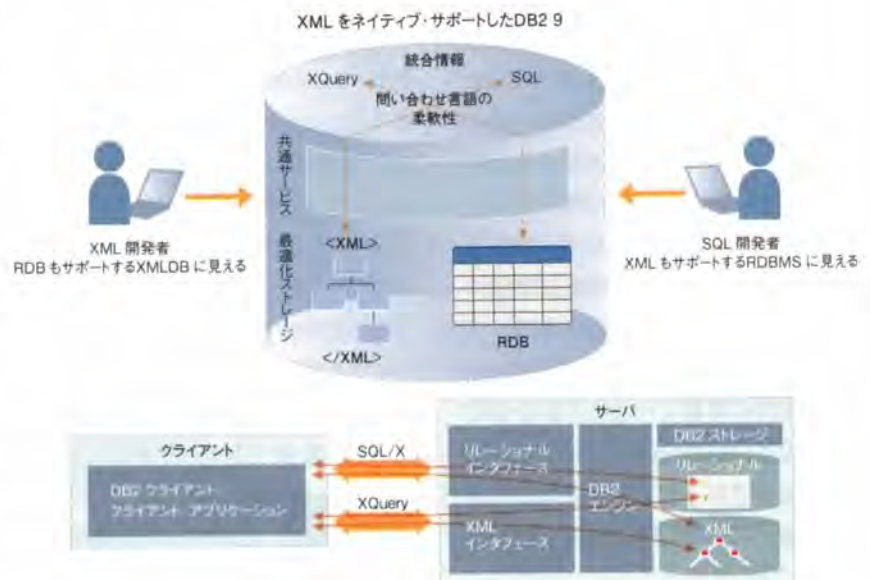
また、必要に応じて、XML文書ごとにXMLスキーマを使用したバリデーション (妥当性検査) を実行可能。リレーショナル・ストレージとXMLストレージはユーザーから透過的で、ユーザーは業界標準のXQuery、XPath、SQL/XMLのいずれを使用してもXML文書への照会、リレーショナル・データへの

照会、それらを結合したデータを取得できる。

IBMがDB2 XMLのサポート効果をDB2 9 Viperを使ったプロトタイプでテストしたところ、「検索および取り込みプロセスの開発」では、従来のXMLサポートがCLOBで8時間、shardで2時間を要していたが、DB2 Viperでは30分だった。また、「スキーマの変更」では1週間:5分、「I/Oコードの開発」では100:35、「検索処理」では24~36時間:10分以内といずれもDB2 Viperが速かった。さらに、検索の開発のために開発者が要する準備期間は、従来型XMLサポートの場合で1週間、DB2 Viperは5分しか、かからなかった。

DB2 9で拡張されたRDBの基本機能としては、スケーラビリティ (テーブル・パーティションのサポート、上限値の拡大)、管理コストの削減とパフォーマンスの向上 (オートノミック機能の強化、

DB2 9ハイブリッドXMLエンジン



オブティマイザの強化)、堅牢なインフラの提供(データ圧縮、セキュリティ機能(Label Based Access Control))、開発生産性の向上(最新のMicrosoft .NET環境対応、Eclipseのサポート)などがある。

また、パフォーマンス向上と大量データ処理で提供される機能には、オブティマイザ、クエリ並列処理、多次元クラスタリング、サマリ表(MQT: Materialized Query Table)、テーブル・パーティショニング、シェアード・ナッシング、ハッシュによるパーティショニング、再配置機能、パーティショニング機能の統合、対話型ワークロード管理(Governor)、予約型ワークロード管理(Query Patroller)

さらに、「可用性・セキュリティの強化」で提供される機能には、パフォーマンス監視・分析、データ行圧縮、バックアップ圧縮、高可用(オンラインでの管理による高可用の実現)、高可用性災害時リカバリ(HADR: High Availability Disaster Recovery)、自動クライアント・リクルート機能、セキュリティの強化(暗号化、ラベル・ベースのアクセス制御)などがある。

#### データベース上で

#### XMLとSQLを相互運用

### Oracle Database 10g (Oracle XML DB)

オラクルのXMLデータベース「Oracle XML DB」は「Oracle 9i Database Release 2」上に実装され、2002年8月に国内出荷されている。標準仕様に準拠し、XML Typeデータ型(XML文書格納用のデータ型)に

#### Oracle XML DB



よって、「1つのXML文書」を「1つのデータベース・オブジェクト」として実現している。また、XMLスキーマに基づく構造化マッピング、高レベルなXML用のAPIの実装、リポジトリ機能の実装(Oracle XML DB Repository)などを特徴とする。

「Oracle Database 10g」で組み込まれた新機能および強化点として、XMLスキーマ進化、NLS対応、データ型への対応強化、XPathリライトの強化、SQL/XMLの最適化がある。

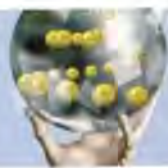
XMLスキーマの進化によって開発者は、アプリケーション要件の変更に応じてXMLスキーマを変更できるようになった。スキーマ展開プロセスにより、XMLスキーマに加えられた変更の影響を受けることなく、既存のドキュメントと参照を維持できる。

またNLS対応としてOracle Database 10gでは、Oracle XML DBリポジトリに対するプロトコル・ベースの接続に関連して、異機種間クライアント・キャラクタ・セットの完全サポートが提供された。プロトコル・サーバにより、クライアント・キャラクタ・セットとデータベース・

キャラクタ・セット間の双方向変換が可能になった。これは、リポジトリを使用して、他のキャラクタ・セットを使用するクライアント間で情報を共有できることを意味する。

さらに、データ型への対応強化では、TIMESTAMP WITH TIME ZONE、BINARY\_FLOAT、BINARY\_DOUBLEなどのXMLスキーマと、SQLスカラー・データ型間におけるマッピングの新しいオプションがサポートされた。

また、XPathリライトの強化では、Oracle Database 10gのXPathリライト機能が改善され、これまで以上に多くの数のXPath式をネイティブSQL文にリライトできるようになった。XPathリライトは、継承された単純型、列挙、型などのXMLスキーマ構造でも使用できる。コレクションを参照するXPath式のリライトも大幅に拡張されおり、XPathリライトは主要なXPath関数に対応している。Oracle Database 10gでは、XPath式のバインド変数に対するサポートも導入された。これらの拡張により、クエリ操作と更新操作のパフォーマンス



スが改善されている。

このほか、SQL/XMLの最適化では、XML生成機能の最適化により、Oracle Database 10gのパフォーマンスが大幅に向上する。これによりオラクルでは、リレーショナル・データからXMLを生成するメソッドとして、SQL/XML演算子の使用を推奨している。XPathリライトは、現在はSQL/XML演算子でのクエリにも対応。これは、SQL/XML演算子ベースのクエリやビューに対するXPath操作により、基盤となるリレーショナル表のSQL問合せのパフォーマンスと同じレベルのパフォーマンスが提供されることを意味する。

Oracle XML DBによってオラクルでは、アプリケーション作成コストの削減をはじめ、シンプルな開発プラットフォーム、複雑さの排除、コンポーネントの削減、統合されたセキュリティの実現、信頼性/可用性/スケーラビリティなどが実現できるとしている。

## 大規模・高速検索向けXMLDB TX1

東芝ソリューションが2005年4月に市場投入した「TX1」は、検索性能で世界最高水準を目指して開発されたXMLデータベース。大容量データの高速検索と複雑化する検索ニーズに対応するため、XMLデータの階層構造の自動抽出（構造の自動抽出と問い合わせ最適化）による構造索引と、全文検索のための語彙索引を実装し、抽出した構造情報と語彙情報を統計的に分析して最適な問合せプランを生成する。これらの技術で検索対象への参照を最小限に抑え、検索速度の高

速化を実現している。

また、XMLデータの高い柔軟性を維持しながら、新規構築から部分的な修正・変更まで、開発生産性を大きく向上させることができる。XMLの問合せ言語にXPath、XQueryを採用し、標準に準拠したインタフェースで効率的なアプリケーション開発を支援する。

さらに、データベースに求められる信頼性を確保するためのデータベース管理機能をサポートしており、更新と参照の並列化、トランザクション一貫性、リカバリ機能などを実装している。このほか、オンライン・バックアップ、データのインポート/エクスポートなどの運用機能もサポートしている。

同社は昨年11月、TX1の新バージョン「TX1 V2」を販売開始した。TX1V2では、テラバイト級のデータ容量でも高速検索を実現するTX1の特徴を継承し、単語の意味に着目して検索性能（精度と速度）を向上する“自然言語処理機能”と、RDBやNotesなどの多様な形式のデータをXML形式に変換し登録する「データ連携機能」を新たに搭載した。コンテンツ管理システム

(CMS)やデータ統合システム分野向けに提供される。

TX1 V2に搭載されたのは、形態素解析方式（意味のある単語で索引となる文字列を切り出す方式）の検索が行なえる自然言語処理機能。これは東芝ソリューションのナレッジ・マネジメント支援システム「Knowledge Meister」にも搭載されるなど、同社が長年培ってきた高度な日本語処理技術をもとに開発された。これにより、従来からのNグラム方式（隣接する文字列で索引となる文字列を切り出す方式）の検索に加えて、単語の意味に着目した検索が可能になった。Nグラム方式と形態素解析方式は、型番や本文などXMLのタグ単位で使い分けられるため、データに適合した検索設定で目的のデータを素早く抽出できる。

また、データ連携機能では、元データからXML形式へのデータ変換方法を指定して実行するだけで、TX1へ容易にXMLデータとして登録できる。これらの新機能により、既存システムはそのまま運用しながら、複数システムに分散したデータを一元管理するデータ統

データ連携機能を活用したTX1によるデータ統合システムの例



合システムを最適に構築できるようになる。たとえば、販売情報や問合せ情報、保守情報が個別のシステムで運用されている場合、あるユーザーに関する情報をすべて集めるには、それぞれのシステムのデータを個別に検索する必要があったが、TX1でデータ統合システムを構築すると、システムを横断した迅速な検索や分析が可能になる。

## 超高速インメモリXMLデータベース Karearea

セックは2002年4月に超高速インメモリXMLデータベース「Karearea (カレアレア)」を公開、12月から Karearea 1.1 のリリースを開始した。

Kareareaは、データのリアルタイム処理システムで求められるゼロ・レイテンシー（システムが限りなく瞬時に反応すること）と大量のXMLデータの高速処理に対応するネイティブXMLの超高速インメモリXMLデータベースとして開発された。

KareareaはXMLをメモリ上に展開し、超高速データ処理エンジンとして搭載したターボデータラボラトリー社製の「LFM」によって高速処理を行なう。

また、Kareareaでは、LFMで処理するために主記憶上に展開したデータが、独自の“コンパイル”処理によってコンパクトに格納されるため、重複したデータが多くなりがちなXMLに対して、メモリ搭載容量が比較的小さいPC (256MB)でも軽快に動作する。さらに、KareareaがXMLのスキーマ情報を解釈して自動的に最適な内部構造に変換するため、XML文書をそのまま格納して使用することが可能で、ユーザーが面倒なスキーマ設計やインデックス作成を行なう必要がない。これによりユーザーは、性能面を気にすることなくXML文書をそのまま格納でき、性能をチューニングしなくても高速な検索・多次元集計・ソートなどが行なえる。

Kareareaデータベースに格納されたXMLデータに対しては、専用のJavaAPIを使用してJavaアプリケーション・プログラムからアクセスすることが可能。Kareareaは、インメモリデータベースの利点を最大限に活かして高速な検索を実現しているだけでなく、一般のネイティブXMLデータベースにはほとんど装備されていない高速ソート機能や多次元集計機能を装備していることが大きな特徴になっている。

### Kareareaの構成



- LFM Engine  
超高速データ処理エンジン
- LXP (LFM XML Processor)  
LFMとKareareaのインタフェース部
- Karearea Java API  
Kareareaデータベース・アクセス用のJava API  
Karearea Java APIを使用してアプリケーション・プログラムを開発可能

## 次世代XMLソリューション xfy Enterprise Solution

ジャストシステムが昨年9月に販売開始した「xfy Enterprise Solution for DB2 9」は、XMLアプリケーションの開発基盤とクライアント側の実行環境を統合的に提供するプラットフォーム。「xfy Enterprise Solution 1.0」と「xfy Database Adapter for DB2 9」で構成される。

xfy Enterprise Solution 1.0は、XMLを利用して情報の再利用と一元管理を実現し、柔軟性の高いアプリケーション開発基盤とクライアントの実行環境を統合的に提供する基盤環境。また、xfy Database Adapter for DB2 9は、IBMのXMLDB / RDBハイブリッド・データベース「DB2 9」との連携機能を提供する。xfy Enterprise Solution for DB2 9は、クエリ対象のデータベースを選択するだけで、そのデータベースに適したクエリ文を自動生成する。また、RDBの参照結果を条件にして、XMLDBをクエリできるXQueryを自動生成するなど、XQueryとSQLの違いを意識することなく利用できる。

さらに、利用するWebサービスを選択して必要な情報を設定するだけで、自動的にリクエストを発行し、取得したXMLデータをもとに自動的にビューを作成する。このビューは自由に組み合わせることが可能で、これにより使い勝手に優れたアプリケーションをコーディングなしで作成できる。

これまでXMLで記述された文書をブラウザで見るとは、XMLデータをHTMLに変化する必要がある、



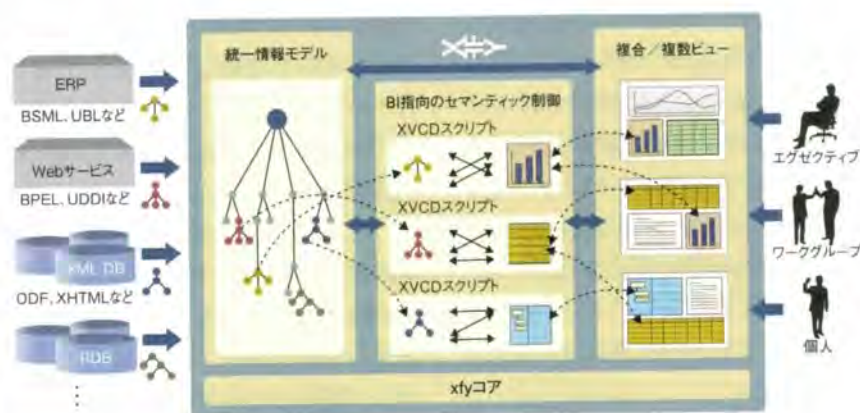
付加情報を活用できなかった。xfy Enterprise Solution for DB2 9の環境では、DB2 9に格納されたXMLデータを、そのままクライアント上で直接編集・加工できるという。また、従来のWebアプリケーション開発はJavaとSQLの実装で、柔軟性に欠けるため、環境やニーズの変化に即応するには限界があったが、xfy Enterprise Solution for DB2 9では、xfyの柔軟な拡張性を活用したGUIでの開発が可能になるため、全社レベルの高度なXMLアプリケーションを短時間で構築でき、エンドユーザーも業務の目的に最適なXMLアプリケーションを作成できる。

なお、ジャストシステムは昨年10月に、日本のソフトウェア・メーカーとして初めて米国IBMとの間で包括的な協業契約を締結し、グローバルな規模でxfyビジネスを推進していくことを発表している。これにより両社は、世界市場を視野に入れて共同で市場展開計画を策定し、具体的なセールス活動をグローバルに展開する。

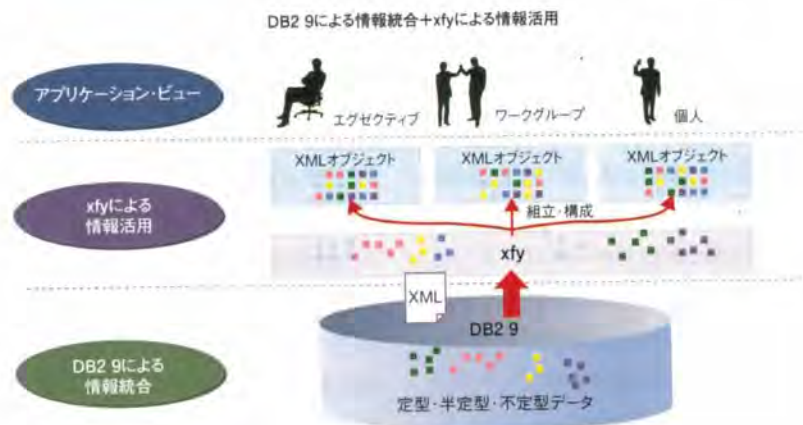
また、ユーザー事例や適用業務の研究を踏まえた新規事業開拓も行なう。さらに、相互に技術を評価し、ユーザー企業から必要とされるソフトウェアを共同で開発する。

さらに昨年11月には、日本オラクルとの間で、XMLの新規市場開拓で戦略的な協業に合意したことを発表した。これにより両社は、共同で「xfy Enterprise Solution for Oracle Database 10g」のプロモーションや企業ユーザーへの提案を行なうなど、販売施策やパートナー施策で協業する。この合意に基づく協業活動として、早期導入企業へのシステム提案をはじめ

## xfyアーキテクチャ



## xfy Enterprise Solution for DB2 9



め、共同セミナーな各種イベントへの出席、共同販促活動、業界別ソリューションの開発、パートナー企業への販売支援活動、などを進めていく。

この協業に合わせてジャストシステムは、xfy Enterprise Solution for Oracle Database 10gを発売した。Oracle Database 10gとxfy Enterprise Solution for Oracle Database 10gの連携ソリューションを共同で提案し、新規市場を共同で開拓

していく。

xfy Enterprise Solution for Oracle Database 10gでは、標準でサービス指向アーキテクチャ (SOA) のクライアント環境が提供されるため、対象とするWebサービスを指定するだけでXMLデータを取得でき、自動的にビューに視覚化できる。さらに、生成したビューを登録して貼り付けるだけで、SOAのクライアント・アプリケーションを開発できる。