

【出展】 Think IT

ウェブページタイトル：amazonがXMLDBを採用した理由！

URL：https://thinkit.co.jp/article/99/2

最終アクセス年月日：2008年7月11日

クラウド

技術解説

連載：あのサービスの裏側！

amazonがXMLDBを採用した理由！

□ 小野 雅史

□ 2008年7月11日(金)

XML技術とオブジェクト指向技術

今ITに求められているのは、既存の枠組みを超えた新しいビジネスモデルの実現です。特に刻々と変化するビジネス状況を現場に正確かつ瞬時に提供できるしくみは、Webを介した小売業には必須のものになっています。

こうしたITシステムには、大きく2つの要件が求められます。1つは「変化に対する柔軟性を持ったデータ管理の方法」、そしてもう1つは「その時々状況やイベントに応じた拡張性を実現するデータ基盤」です。

今回は、このようなシステムで使われているXML技術とオブジェクト指向技術についてお話しします。

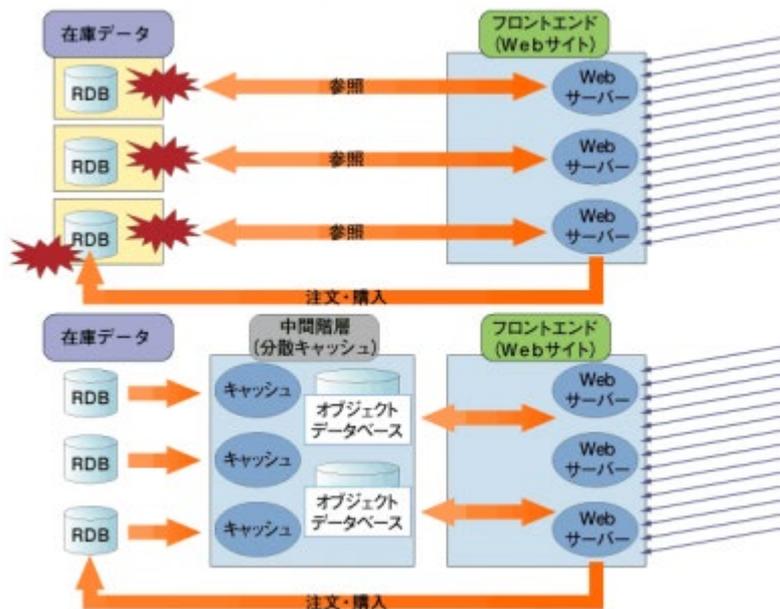


図1:XMLデータベース採用前(上)採用後(下)のアマゾンシステム

「今の在庫」を正しく把握せよ-米アマゾン・ドットコム

最初に紹介するのは、米アマゾン・ドットコム（以下アマゾン）の例です。

書籍の注文など、大規模なeコマースのシステムにはユーザーから大量のアクセスがあります。この時、ユーザーに対して「その時の在庫情報」を正しく提供することが重要で、これができないとせっかくのビジネス機会を失うことになってしまいます。特に大量アクセスが集中する時間帯などに、ユーザーからの注文に対して即座に在庫情報を提供できなければ、その機会損失は膨大になります。

アマゾンの従来のしくみでは、バックエンドのリレーショナルデータベース（以下RDB）にアクセス負荷が集中すると、パフォーマンスが劣化してしまい、最新の在庫情報を提供するのに非常に長い時間がかかっていました。そこで注目したのが、刻々と変化するデータをオブジェクトとして、それをメモリキャッシュ上で処理できる分散オブジェクト技術です。

さらに、キャッシュをWebサーバ側に複数配置し、メモリ上にキャッシュしたデータをメモリスピードで高速にアクセスする「キャッシュフォワード

（http://www.cybertech.co.jp/xml/xmlldb/luxeon/arch_c/cyber_luxeon_1.php）技術」を利用しています（図1）。

アマゾンでは、アクセス要求の増加に応じて、キャッシュの数を段階的に増やすことで、パフォーマンスを改善し、システムをスケールアウトしていきました。

ここで使われている「オブジェクトデータベース」や「キャッシュフォワード技術」は、XMLデータベース製品にも利用されており、現在日本市場では、XMLデータベース「Cyber Luxeon（サイバー・ラクセオン）

（<http://www.cybertech.co.jp/xml/xmlldb/luxeon/>）」やオブジェクト指向データベース「ObjectStore」に採用されています。

次に「オブジェクトデータベース」「キャッシュフォワード技術」「ダイナミックモデリング技術」について、紹介しましょう。

XMLデータベースの技術

オブジェクトデータベースとは、データとそのデータの処理を1つのオブジェクトとしてとらえ、データを保持するデータベースです。現在幅広く利用されているリレーショナルでは、テーブルでデータが管理されており、オブジェクトデータベースとは構造が大きく異なります。

先ほどアマゾンの例で紹介した「ObjectStore」もオブジェクトデータベースです。今回は、このObjectStoreをコアデータベースに採用しているXMLデータベース「Cyber Luxeon」を例に、解説を進めていきましょう。

Cyber Luxeonは、W3C標準のXML処理エンジンを内蔵し、XMLドキュメントを階層構造のまま、DOMオブジェクトとして格納しています。

まずは、キャッシュフォワード技術について解説してきましょう。キャッシュフォワード技術は独自の分散型アーキテクチャーです。従来のデータベースは検索のパフォーマンス向上を目的として、データベースサーバ側にキャッシュを持っています。キャッシュフォワード技術とは、このキャッシュをアプリケーションサーバ側にも持たせる技術のことです。

キャッシュフォワード技術を用いて設計されたObjectStoreでは、データ検索を行う際に、直接データベースサーバのデータにアクセスするのではなく、アプリケーションサーバ側のキャッシュにアクセスすることで、単純な検索処理の向上だけでなく、サーバやネットワークの負荷を大幅に軽減することが可能です。

Cyber Luxeonでは、XMLデータをメモリに配置する「XMLキャッシュ」という技術にあたります。メモリ上にロードされたXMLを高速に検索することで、メモリスピードの処理速度で結果を返すことが可能です。さらに、このXMLキャッシュは別マシンへ分散することができ、クライアントのアクセス数が増大した場合や、サーバ側のXMLデータが増加した場合に、キャッシュサーバを増設するといった段階的な拡張（スケールアウト）ができるのが特徴です。

また、Cyber Luxeonでは「ダイナミックモデリング」と呼ばれる技術が採用されています。これは、XMLをノードレベルで処理（検索・更新）する技術です。この技術によって、システムを停止することなく、XMLのデータ構造を運用中にダイナミックに変更することが可能です。サービスを稼働し続けなければならないようなサービス（B2B、B2C向けの大規模ECサイトなど）で活用されています。

今回紹介した技術を図2にまとめました。

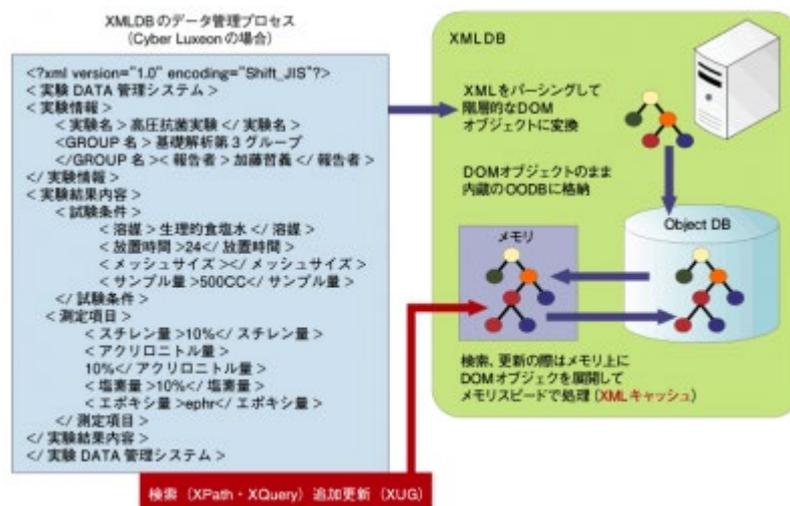


図2:XMLDBのデータ管理プロセス

データベースは適材適所

ここで、XMLデータベースを活用するにあたってのポイントとコツを整理してみましょう。

まずデータベースの選定のポイントは、適材適所です。変化が激しく予測がつきにくいデータを無理やりRDBで管理しても、データベースの設計変更による運用後のシステムメンテナンス回数が増大し、データの利活用を妨

げる結果となってしまいます。

定型的なデータはRDBで管理し、そのデータを時系列でまとめて、ほかのデータとの関連付けをする必要が出た場合には、柔軟で拡張性のあるXMLデータベースに格納するのが効率的ではないかと考えます。

ではXMLデータベースの設計のコツはなんでしょうか。ここまで述べた通り、XMLデータベースは、データ項目の変化に対しての柔軟性と拡張性を持っています。ですから、データ設計も「ナナサン（7：3）の法則」が良いわけです。

つまり、サービス当初に分かっている基本項目が全体の70%だとすると、予測不可能な30%の項目については、運用後、必要に応じて拡張していけば良い、といった割り切りで設計すると良いのです。

ところで、皆さんのまわりにRDBに格納できなかった、つまり捨てられていたデータはありませんか？RDBに格納しなかった理由は、「格納するための項目がなかった」「1カラムの文字制限にあふれた」「文章なのでそもそもデータベースに入れることすら考えなかった」というものが大半を占めているはずですよ。

しかしよく考えてみてください。こういったデータこそ、まさに顧客や現場の声であり、捨ててはいけぬ貴重なナレッジではないでしょうか。ですから、そのナレッジはXMLデータベースで管理できるよう工夫することが大事です。

XMLデータベース、現在そして未来

ここまでは、皆さんがあまり知ることのない、事例を元にXMLデータベースを紹介しました。最後は、今後のXMLデータベースの可能性について書きます。

今後、XMLデータベースが活躍する分野は「クロスメディア」です。クロスメディアとは、紙、Web、携帯など多様なメディアでコンテンツを配信したものを、ユーザーが横断的に利用できるしくみです。

例えば、主婦が夕食のレシピを雑誌で見ながら、その出版社の会員制Webサイト上に掲載されているレシピ情報と必要な献立情報を、自動的に携帯メールに配信する、というように雑誌・PC・携帯という複数のメディアを横断して情報を利用するようなくみです。

クロスメディアとXMLが注目されているのには、理由があります。

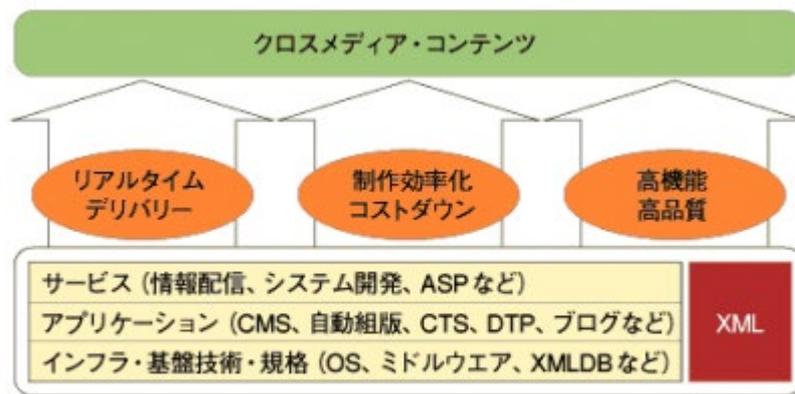


図3:クロスメディア

XMLDBがオンデマンドでデータの生成

クロスメディアでは、フリーペーパーのような紙の印刷物を製作する際、データをXML化することで、Webサイトのコンテンツを簡単に（オンデマンドで）生成することも可能になります。

ブログシステムの裏側にXMLデータベースを追加し、再利用のためのコンテンツを蓄積しておけば、Webサイトからブログコンテンツを選択して紙の出版物をオンデマンドで制作するような「オンデマンドブログ出版」も容易に実現できます。

業界を超えた新しいメディアの枠組みの中で、Webサイトの利用者とコンテンツ提供者、サイト管理者が、自由自在につながるためのしくみが今後続々と登場してくるでしょう。

その中でもXMLデータベースは、サービスの裏側のデータ管理部分を支える重要なキーテクノロジーになることは間違いないのです。

著者 小野 雅史

この著者の
記事一覧



株式会社サイバーテック

株式会社サイバーテック 取締役。XMLDBの普及に命をかけるマーケティング屋。XMLDBの発展はエンジニアの感性と遊び心にかかっていると信じて疑わない。既成概念にとらわれないオープンでキャッチなXMLアプリケーションの登場を待ち望んでいる。1967年1月生まれ。出生地は名古屋市。現在の生息地は川崎市。水瓶座のO型。