

# XML と XMLDB を活用した分散データ管理アーキテクチャの提案

江川 純雄<sup>†</sup> 若原 俊彦<sup>‡</sup> 松本 充司<sup>†</sup>

早稲田大学大学院国際情報通信研究科<sup>†</sup>

早稲田大学国際情報通信研究センター<sup>‡</sup>

## 1. まえがき

近年インターネットにおいては様々なサービスが登場し多種の Web アプリケーションが普及している。しかしその一方で、ユーザーアクセスの集中やアクセス端末の多様化などによりハード機器やアプリケーション開発ソフトなどのサーバー構築コストは増加傾向にある。また、データ管理においてはリレーショナルデータベースにてデータ管理を行うというのが一般的であるが、組織内或いは組織間でデータの交換や登録を行う際にそれらの互換性の欠如が支障となることが多い。また、その他にも併用する Web アプリケーションサーバーの性能に依存してしまうことやデータベース内のデータテーブルが増大するにつれ、それらの管理が煩雑になりパフォーマンスを向上させるなどのチューニングにおいて専門的な知識が必要になってしまふといった欠点もある。そこでこれらを解決する方法として今回は XML ドキュメントおよび XML データベースを連携させてデータ管理を行うことを考え、そのデータ管理ではどのような利点があるか、どのような欠点があるかを洗い出すことにより優れたデータ管理手法を考察しようと考えた。また最近 SOAP, UDDI, WSDL といった最先端のテクノロジーを使う Web サービスが徐々に拡大し浸透しつつある。よって今回は Web サービスとも連携し分散型のデータ管理を柔軟に行うことができるアプリケーションアーキテクチャについても考察を行う。

## 2. XML と XMLDB の特徴

XML ドキュメントは任意のツリー構造を構築できることでユーザビリティに優れ、各ノードに任意の要素名と属性名を付与できることからタグの情報を元にした検索や更新が容易に行えるといった利点がある。しかしその一方で一つの XML ドキュメント内のデータ量が増大すると DOM ツリーの特性上オーバーヘッドが生じ検索パフォーマンスが低下すると言われている。また XML ドキュメント同士はノンリレーショナルである為データをリレーショナルに関連づける為のインデックスファイルが必要とする場合も出てくる。

これに対してネイティブな XMLDB は名前空間を含めた XML ツリー構造を保持したままドキュメントデータの保存とその取り出しが行なえる。そしてさらに XPATH による検索やツリー同士の関連づけなども行える為、

XML ドキュメントの弱点を補った特性を保持しているとも言える。

## 3. XML と XMLDB の連携

前章の XML ドキュメントと XMLDB の特性から、XML ドキュメントを主に使用しつつ、データ検索や XML ツリーのリレーショナル化、そしてデータバックアップなどで補助的に XMLDB を取り入れるという連携方法があり実際のシステム構築例も存在する。しかしこの連携だけでは XMLDB の利点を活かしてきていない。XMLDB の一番の利点は分散したデータ管理ネットワーク環境でもフレキシブルに XML ドキュメント単位のデータ通信や交換を行うことが可能となるという点であると考えられる。よって今回はその点を最大限に活かしたアーキテクチャモデルを考察することにした。図 1 は XML と XMLDB の連携アーキテクチャモデルを示している。

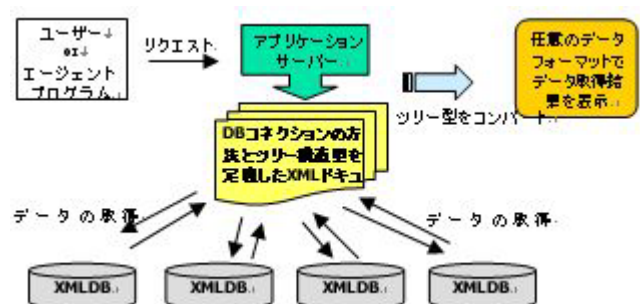


図 1. XML と XMLDB の連携アーキテクチャモデル

このようにアプリケーションサーバーが諸処の XMLDB と通信を行う際には DB コネクションの方法と各 XMLDB オリジナルなツリー構造型の定義情報が必要となるので、それらが記述された XML ドキュメントを予め作成し中核のアプリケーションサーバーに保持しておく。そうすればユーザー或いはエージェントプログラムが情報取得のリクエストを自律的に行い、アプリケーションサーバーが保持している XML ドキュメントの定義に従ってネットワーク上に分散しているそれぞれの XMLDB に情報を取得しに行き、その後アプリケーションサーバーのツリー構造にコンパイルしてデータ結果を表示できると考えられる。こうして分散したネットワーク環境においてフレキシブルなデータの取得や管理が実現できる。また、動的にデータ結果をコンパイルすることで Web ページの出力に限らず様々なデータフォーマット型でのデータ出力が可能となる。

## 4. アーキテクチャの構築

今回はこの XML と XMLDB の連携アーキテクチャを

<sup>†</sup>Proposal of the Distributed Data Management Architecture cooperated with XML and XMLDB

<sup>†</sup> Sumio EGAWA, Mitsuji MATSUMOTO · Graduate School of Global Information and Telecommunication Studies, Waseda University

<sup>‡</sup> Toshihiko Wakahara · Global Information and Telecommunication Institute, Waseda University

実践するにあたって、Java と XML を組み合わせたアプリケーションコンポーネントで以下のような Web アプリケーションシステムを構築した。

- (1) ユーザーはユーザーの個人情報を登録する。
- (2) 登録したユーザーは自分のスケジュール情報を公開し他のユーザーとスケジュールの共有ができる。

この Web アプリケーションシステムにおいて中核のアプリケーションサーバーに前章で述べた機能を実装した。そしてローカルエリアネットワーク内に XMLDB サーバーを三台設置し一台目では上記(1)のデータ管理、二台目では上記(2)のデータ管理、そして三台目ではログイン情報などのログデータの管理を行うこととした。図 2 はサーバと XMLDB とのデータの連携方法を示している。

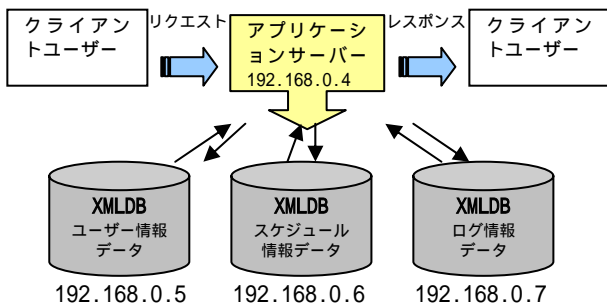


図 2. サーバと XMLDB とのデータ連携方法

このように外部インターネットのクライアントユーザーのリクエスト処理に対してアプリケーションサーバーは自身が保持している XML フォーマットの定義ファイルよりネットワーク内の使用可能な XMLDB を検索する。そして Java のソケット通信によって XMLDB サーバーに接続しユーザー情報の検索を行う。そしてその情報を得た後に次の XMLDB サーバーに接続しスケジュール情報やログ情報の処理を行うという分散型のデータ管理ネットワークアーキテクチャを構築した。

## 5. Web サービスへの拡張アーキテクチャ

しかし、前章のアプリケーションサーバーネットワークにおいては様々な問題点がある。例えばアプリケーションサーバーは XMLDB サーバーの IP アドレス、ポート番号、データベースディレクトリ、XML ツリー型などの各種情報を認知しておく必要があるため、アプリケーションサーバーの運用とアップデートが煩雑になり XMLDB サーバーに取り入れる条件も厳しくなってしまうといったことが挙げられる。よって XMLDB サーバーの情報、データへのアクセス方法などが共通化されたインターフェイス言語やファイルで記述しなければスタンダードになり得ないと考えられる。そこで今回はこのアーキテクチャを拡張させ、XML、SOAP、UDDI、WSDL などある程度標準化が進みながらも発展途上にあるプロトコルと技術要素を活用してインターネット上に分散する各種アプリケーションの連携を行うことができる Web サービスに対応させることにした。図 3 は Web サービスへの拡張アーキテクチャの例を示している。

今回においては Web サービスアプリケーションサーバーとして Sun Microsystems がオープンツールとしてリリースしている JWSDP (Java Web Services Developer Pack) を使用し、これに付随されている XMLDB ベースの UDDI レジ

ストリにローカルエリアネットワーク内の XMLDB にアクセスするインターフェイスである WSDL ファイルの情報を登録した。そしてその後 WSDL ファイルの記述に従って XMLDB サーバーに接続を行い、XMLDB からデータの取り出しを行うことが可能となった。今後はアプリケーションサーバーに Web サービス利用ユーザー別のアクセス制限機能などを実装させる事でセキュリティ対策を実現しようと考えている。また、さらに Web サービス利用のクライアントユーザーの判別によってサーバー側で様々な付加情報を付与するロジック処理を行い、それに従ってオリジナルなデータ取得結果を伝えるというセマンティック Web 的な機能も実装できると考えている。

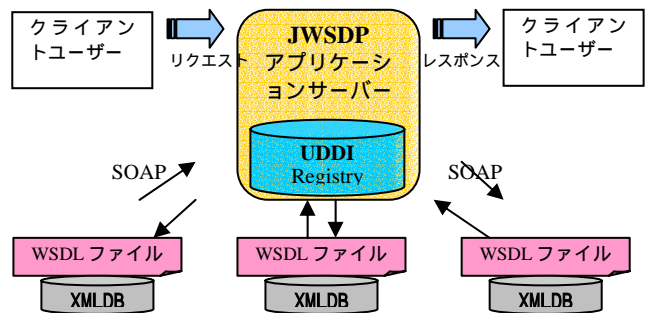


図 3. Web サービスへの拡張アーキテクチャの例

## 6. まとめ

XML ドキュメントと XML データベースを組み合わせたデータ管理には様々な連携方法があり、扱うデータの種類や目的によってフレキシブルな連携を行うことができることが判った。また、この連携により従来には存在し得なかった斬新な分散型のデータ管理ネットワーク環境のアーキテクチャを構築できることが判明した。

しかし、この様なアーキテクチャを実現するためには独特のオリジナルなシステム環境条件を定めてもスタンダードとして広がるとは考えられない。RFC などである程度標準化が進んでいるプロトコルやインターフェイスを取り入れる事が重要であると考えた。そこで最近仕様や各種ツールの開発が活発に進んでいる Web サービスの各コンポーネントを導入することとした。そしてプライベートな UDDI レジストリを構築し XML フォーマットのドキュメントを活用し複数の XMLDB を連携させることで、今までにない分散型のデータ管理ネットワーク環境のアーキテクチャを構築できることが判明した。

今後は XMLDB のツリー構造型の定義を行うことを可能とする様々なスキーマを探索しデータ取得後のスムーズなツリーコンバートを実現させる予定である。また同時に XMLDB の連携方法やデータの検索を行う際の最適な検索アルゴリズムの開発と実装を図っていく予定である。

## 文 献

- [1] 垂水浩幸, “グループウェアとその応用”, 共立出版, 2000.
- [2] 米持幸寿, “Web サービス完全解説”, 翔泳社, 2002.
- [3] 江川 純雄, 松川 哲平, 若原 俊彦, 松本 充司, “グループウェアにおけるデータ管理手法の提案”, 2002 第一回情報科学技術フォーラム, 一般講演論文集第四分冊, no.O-4, pp.333-334, Sept.2002.