

電子図書館システム NetLibra の広域コンテンツ流通支援への適用検討

杉山主税, 高橋克則, 爰川知宏, 由比藤光宏

西日本電信電話株式会社 研究開発センター

〒180-8585 東京都武蔵野市緑町 3-9-11

Tel:0422-59-5156

Fax:0422-56-7498

E-mail: {c.sugiyama,takahashi,t.kokogawa,m.yuito}@rdc.west.ntt.co.jp

概要

我々は分散ネットワーク技術 CORBA を用いて、複数の電子図書館等を相互連携させ、情報流通の仕組みを提供するプロトタイプシステムであるネットワークライブラリシステム (NetLibra) を提案し、機能確認および評価を行ってきた。本稿では、その結果を踏まえ、(1)性能向上(2)運用管理(3)ユーザインタフェースの観点から NetLibra の機能拡張を行ったことについて記述する。

キーワード

横断検索, 地図検索, ファイアウォール, 電子図書館, マルチメディアアーカイブ, CORBA

Support for the Wide Area Contents Distribution of "Digital Library Systems NetLibra"

Chikara SUGIYAMA, Katsunori TAKAHASHI, Tomohiro KOKOGAWA, Mitsuhiro YUITOU

NTT West Research & Development Center 3-9-11, Midori, Msashino, Tokyo, 180-8585, JAPAN

Phone:+81 422-59-5156

Fax:+81 422-56-7498

E-mail: {c.sugiyama,takahashi,t.kokogawa,m.yuito}@rdc.west.ntt.co.jp

Abstract

We developed a prototype of an advanced digital library system called NetLibra, which was able to integrate some digital libraries using distributed objects technology and networks. From the results of evaluation of NetLibra, we got some requirements to use such kind of system effectively. To satisfy these requirements, from the viewpoint of (1) high performance, (2) operation management, and (3) user interface, we expanded the prototype.

Keywords

Cross Domain Search, Map Retrieval, Firewall, Digital Library, Multimedia Archives, CORBA

1. はじめに

近年、インターネットの普及により、ネットワークを介したデジタルコンテンツの流通が一般化してきた。図書館のように膨大かつ多様な情報を集め発信することが目的の施設等では、マルチメディアアーカイブ（マルチメディアコンテンツの格納庫）相互をネットワークを介して接続し、情報の発信／流通／共有を扱う仕組みが必要となる。我々のグループでは、分散処理技術である CORBA（Common Object Request Broker Architecture）[1] をベースとして、複数の電子図書館等を連携動作させるネットワーキングライブラリシステム NetLibra を開発し [2],[3],[4], プロトタイプシステムによる機能確認および評価を行ってきた。本稿では、NetLibra の機能概要とその評価について述べ、そこから得られた課題および実運用時に想定される要求条件を整理した。その結果を踏まえ、性能、運用、ユーザインタフェースの観点から NetLibra の機能拡張を行ったことについて記述する。

2. NetLibra のシステム概要

2.1 システムアーキテクチャ

NetLibra は、ネットワークで接続された複数電子図書館等を相互連携させるシステムであり、プラットフォームとして CORBA を採用することにより、複数の電子図書館に分散しているコンテンツ資源を一元的に検索／閲覧することができる横断検索機能を実現しており、また優れた拡張性を有している。

NetLibra は以下の三層のソフトウェア群で構成される。図 1 に NetLibra のソフトウェア構成を示す。

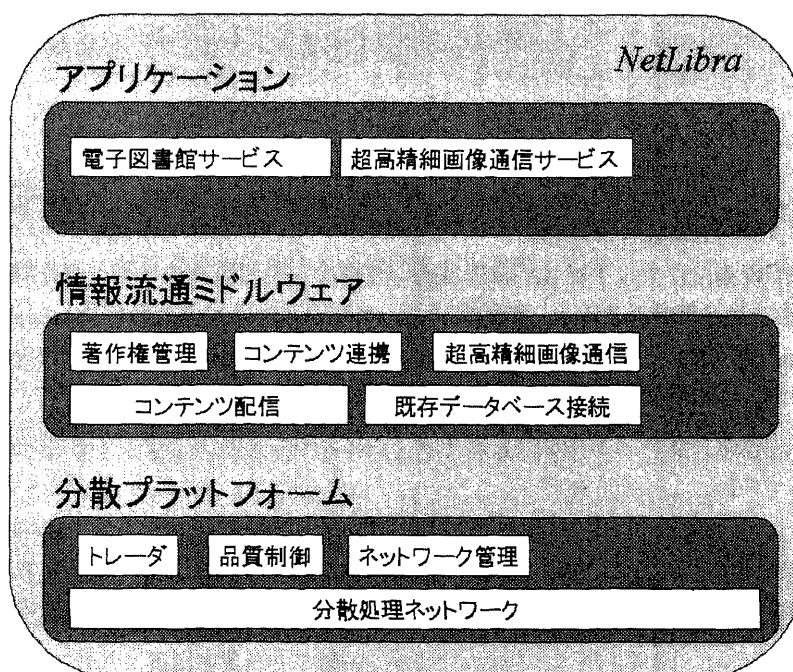


図1 NetLibraソフトウェア構成

(1) 分散プラットフォーム

CORBA による分散オブジェクト処理を用いたネットワーキング機能により、上位アプリケーションからネットワークを意識することなく、スケーラビリティを確保し、システムの拡張や相互連携が容易に行える。NetLibra では、CORBA 共通サービスで規定されているトレーディングサービスを用いることで、電子図書館の自律運用および相互連携機能を実装した。トレーディングサービスは、クライアントからの要求に対して、該当するサービスを実現するオブジェクトレファレンス（位置情報）を返すサービスであり、本サービスを実現する機能がトレーダである。

（2）情報流通ミドルウェア

検索などの基本処理や著作権管理等の情報流通機能群をミドルウェア化することで、上位アプリケーションに対する統一的なシステムインタフェースが提供できるとともに、アプリケーション開発効率を向上できる。

また、分散プラットフォームに対応しない従来のアプリケーションや別の電子図書館システムに対しても、個別のインタフェースミドルウェアを仲介することで、一体のサービスとして見せることができる。

（3）アプリケーション

分散プラットフォームや情報流通ミドルウェアの機能を共通に利用することで検索／閲覧といった基本的なサービスにとどまらず、司書や利用者支援のためのサービスといった図書館ごとに独自のサービスが容易に構築可能である。

2.2 システム構成

NetLibra のシステム構成を図 2 に示す。

利用者端末からは、ゲートウェイ上に実装した CGI（Common Gateway Interface）プログラムを介して、通常の WWW ブラウザによるアクセスで図書館サービスを利用することができる。NetLibra は、分散する電子図書館のもつ各サービスを 1 つのオブジェクトとして扱い、図書館サービスのオブジェクトレファレンスをトレーダにより管理している。

3. NetLibra の評価から得られた課題および要求条件

実験システムとして構築したプロトタイプは地理的に離れた 3 拠点に 4 つの電子図書館サーバを分散配置し、ATM ネットワーク（帯域 2 Mbps）にて接続し、地理的にも分散環境を構築した [5]。

本システム構成による書籍横断検索の性能評価を行った結果として、CORBA による処理のオーバーヘッドはそれほど問題ではなく、実用的な性能を確保できることが示された。その一方で、横断検索対象の図書館が増えたときの検索レスポンスの低下、および、検索処理によるヒット件数が増えた場合の検索レスポンスの低下が見られた。

前者は同時に検索を行うサーバが増えることによる、ネットワーク上のトラフィックの集中が要因の一つであり、評価結果からも、検索処理程度の負荷では深刻な影響は生じにくいと考えられるものの、サーバ数および同時利用者数がさらに増加することで、ネットワーク環境によってはさらなる性能低下が懸念される。

後者については、サーバ（データベース）側で処理される検索結果リストが膨大になることによる負荷増大が主な要因と考えられる。また、検索結果リストでなくコンテンツそのものの取得の場合は、検索時に比べて格段にデータ量が增大するため、サーバ負荷の増大は避けられない。

以上より、特にサーバおよびネットワークの負荷低減、負荷分散の観点で、横断検索、コンテンツ取得といった基本処理の効率化をはかる必要がある。

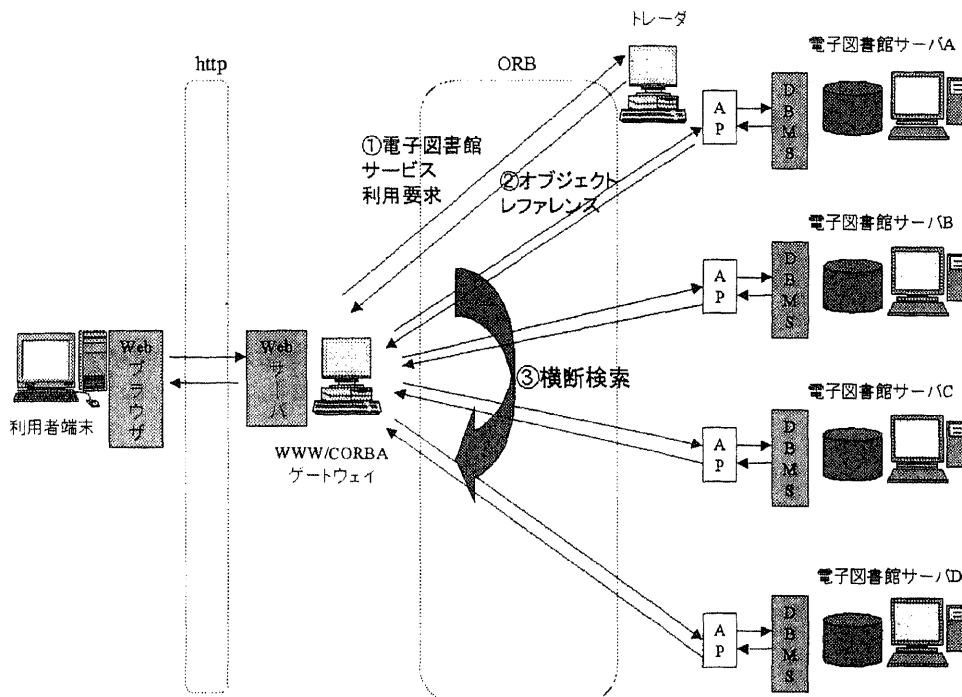


図2 NetLibraシステム構成

4. NetLibra の実運用を想定した場合の要求条件

4.1 ファイアーウォールを超えた通信の必要性

複数の電子図書館相互を連携する上では、電子図書館の本館や分館を同一 LAN 内で運用する場合もあるが、地理的に離れた他の図書館や分館とをネットワークを介して接続する場合、ネットワークの経済性や利便性の面から、インターネットを経由して接続することが考えられる。

この場合、インターネットを経由して別の LAN 内にある CORBA サーバと通信を行う場合には、CORBA 通信に使用される通信プロトコル IIOP (Internet Inter-ORB Protocol) が既存の Firewall を通過することができないという問題が生ずる。

IIOP は使用するポート番号が動的に割り当てられるため、ある特定のポート番号を IIOP 用に割り当てて、IIOP のフィルタをかけてセキュリティを確保するといったことが困難である。つまり、Firewall の外部から、Firewall 内の 1 つの CORBA サーバに対する通信を許可しようとするれば、Firewall 内の全てのサーバに対して、アクセスを許可することになるため、セキュリティ上問題が生ずる。

上記の理由により、インターネットを介した本システムの実運用を可能とするためにも、Firewall を越えた CORBA のサーバ間通信を実現する必要がある。

4.2 運用上の課題

4.2.1 分散データベースサーバにおける障害把握

実運用では、サーバのトラブルや休館によるサービス停止といった場合においても、利用者からのサービス利用要求に対しては、その影響を極力少なくする必要がある。ネットワーク上に分散配置された複数のデータベースサーバの稼動状況を定期的に監視し、稼動状況の把握や障害復旧を効率的に行える必要がある。

4.2.2 信頼性の向上

トレーダ上のサーバ管理情報や利用者情報といったシステム横断的に使用される機能、情報は、常時、正常動作していることが前提条件となるため、個別のサーバ以上に信頼性の向上と、トラブルの回避が必要である。

4.3 検索サービスの多様化

NetLibra の書誌情報（二次情報）による検索やインターネット等で使用されている検索エンジンでは、キーワード等を入力することによって、検索を行っている。

広く一般の利用者に受け入れられる検索方法の一つとして、キーボードからの入力を不要としたインタフェースが考えられる。例として、目的のコンテンツを、地図やリストをクリックすることで階層的に検索するサービス等が挙げられる。

5. NetLibra の機能拡張

3章,4章での課題および要求条件をもとに、以下の機能拡張を行った。

5.1 評価結果を考慮した性能面の機能拡張

5.1.1 負荷分散機能

サーバの負荷状態を監視する機能を追加することにより、各図書館のサービスが稼動するサーバの負荷情報と、ゲートウェイとデータベースサーバとの間のネットワーク負荷情報を常に収集しておき、利用者からのサービス利用要求に対して、これらの情報をもとに最も負荷の小さいサーバに優先的にアクセスをさせることを実現した。

●負荷情報の取得方法

トレーダでは、サーバのオブジェクトリファレンスから、各サーバの CPU 使用率と CPU クロックを取得する。また、全てのサーバに対して、同時にアクセスを行い、各サーバのレスポンス時間を測定する。CPU 使用率とレスポンス時間からサーバの優先順位を決定し、定期的に監視を行う。

5.1.2 属性情報（分野情報）の管理

トレーダで管理する電子図書館サーバの属性として、そのサーバで扱われるコンテンツの分野等を持たせることにより、利用者側で分野を選択した場合、その分野に該当するコンテンツの存在するサーバのみを対象として検索を行うことを実現した。これにより横断検索処理の効率化を図る。

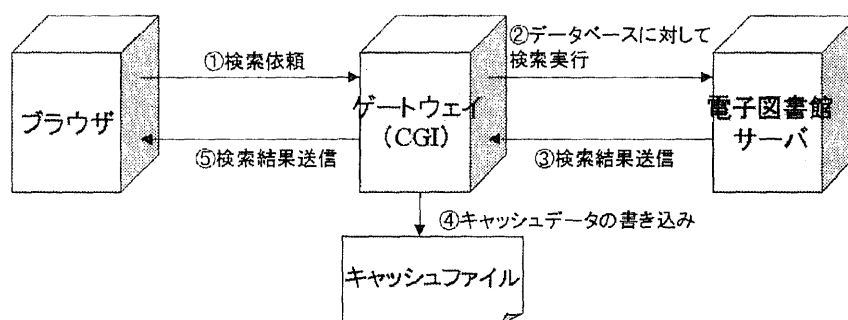
5.1.3 検索結果取得単位および上限数の設定

データベースサーバからゲートウェイに対して、一度に転送される検索結果の件数を抑えるため、データベースにて一度に取得する件数に変更可能な上限値を設けることとし、検索結果の件数に対しても、同様に上限数を設けた。これにより、データベースの必要以上のパフォーマンス低下を抑えた。

5.1.4 キャッシュ機能

ゲートウェイにおいて、利用者からの検索要求に対する結果、あるいはコンテンツの表示に対して保持期間を設けたキャッシュを行うこととした。同一の検索条件でアクセス対象のサーバから検索結果がキャッシュされていれば、キャッシュの内容を読み込み、キャッシュされていない場合は検索要求を発行して、得られた結果をキャッシュする機能を実現した(図3)。これにより、横断検索時のデータベース、ネットワーク負荷の低減を図った。

※キャッシュファイル未存在時の処理フロー図



※キャッシュファイル存在時の処理フロー図

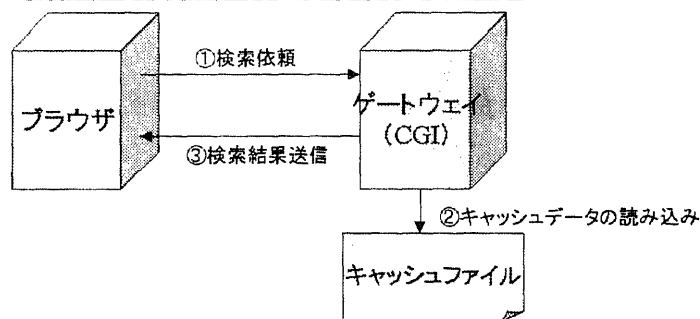


図3 キャッシュ機能

5.2 実運用を想定した要求機能の実現

5.2.1 ファイアウォールトンネリング機能

インターネットを介して Firewall を越えた別の LAN 内にある CORBA サーバへ通信を行う時に、CORBA の通信プロトコル IIOP を WWW で主として使用される通信プロトコル HTTP に変換をすることにより、固定ポートによる CORBA のサーバ間通信を実現した。

Firewallを越えてNetLibraの各電子図書館サーバ連携を実現するシステムの構成を図4に示す。図4ではサーバA～Cは同一LAN上にあるが、サーバDはFirewallを介してインターネット経由で接続されている。相互のFirewall内にはIIOPをHTTPに変換するプロキシサーバを設置する。プロキシサーバによりIIOPがHTTPに変換され、相手側のFirewallを通過して、目的のCORBAサーバに到達する。CORBAサーバ側では、HTTPをIIOPに復元するAPIを実装しており、このAPIにより取り出されたIIOPリクエストにより、一連の通信を実現する。

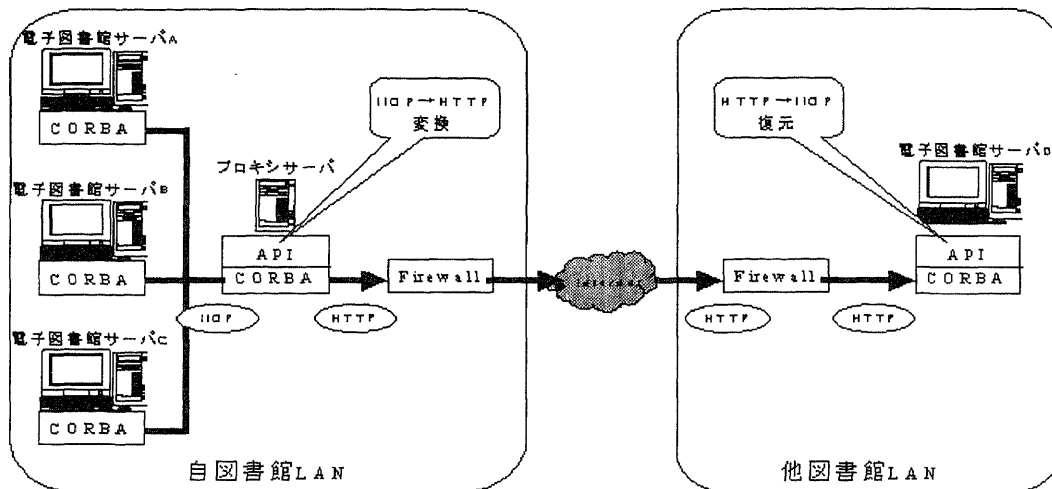


図4 Firewallトンネリング構成

5.2.2 運用面の機能拡張

5.2.2.1 サービスの稼動状態監視

トレーダの情報から、定期的に稼動状況を監視し、サーバからの応答が無い場合は自動的にサービスの再起動を行うことを実現した。なお、サーバからの応答が無い場合は応答がないサーバが提供するサービスの情報をトレーダから削除することで、利用者からの検索要求に対して、常にサービスが稼動しているサーバのみへ横断検索を行うことを可能とした。

5.2.2.2 トレーダ, 利用者サービスの2重化

トレーダおよび利用者情報を管理する機能を2重化し、1つ目のサーバがアクセスエラーとなった場合に、2つ目のサーバへの切替えを行うことで、システムの信頼性向上を図った(図5)。

5.3 階層型地図検索サービスの実現

5.3.1 階層型地図検索サービスの概要

新機能として、キーボードレスでコンテンツを選択/表示できる階層型地図検索サービスを設けた。具体的には、一般利用者にコンテンツを提供するサービスの一例として、日本地図や世界地図をクリックする

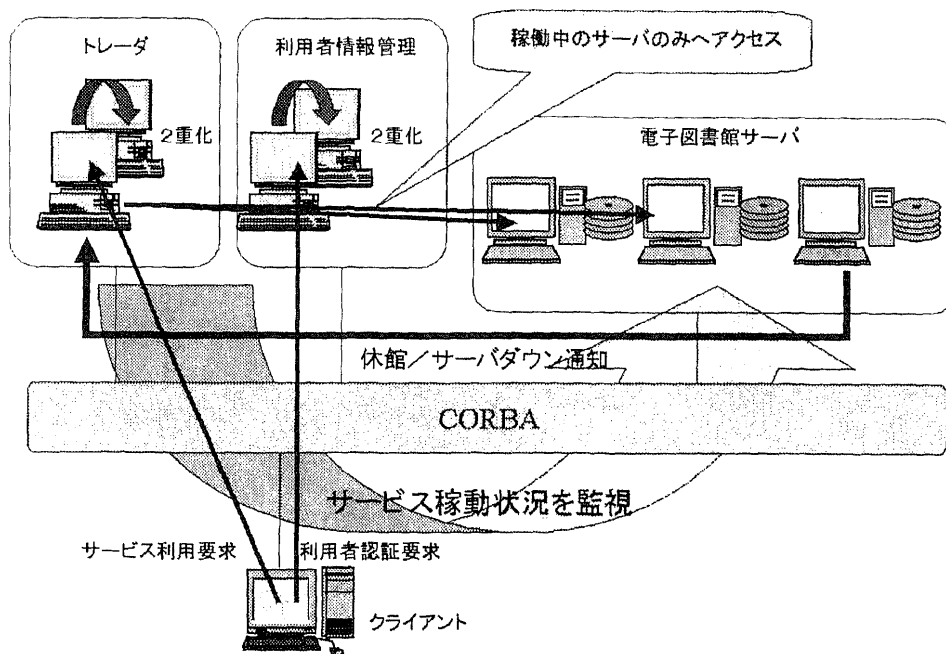


図5 NetLibra運用管理

ことにより地域を特定し、さらに時代と分野により絞り込むことによって、コンテンツを導き出すサービスである。本地図検索サービスによって導かれた検索結果はサムネイルによって表示され、サムネイルをクリックすることにより、コンテンツを参照することができ、また、そのコンテンツに関連する情報を参照することができる（図6）。

階層型地図検索サービスで使用するクリックابلマップは、利用者側で自由に作成することができる。検索結果の一覧表示でのサムネイル画像ファイルは、データベースではなくファイルとして持っているため、その表示速度は、データベースから表示するよりも高速化が図られている。しかしサムネイル画像の表示数が増えることによって、その表示速度や表示イメージに影響が及ぼされるため、設定ファイルをゲートウェイ上に設けることによって、サムネイル画像ファイルのTサイズや表示数等を常時変更が行えるようにした。階層型地図検索サービスでの検索結果の一覧表示画面イメージを図7に示す。

5.3.2 階層型地図検索サービスの応用例

本サービスでは、クリックابلマップの階層を自由に変更することができるため、地図の地域指定を店舗単位まで細かく指定することによって、電子図書館における地域情報の閲覧だけでなく、特定の映画館の上映内容や上映時間等の情報、スーパーやデパートの値引き商品等の情報を動画や音声にて簡単に紹介するといった広範囲のマルチメディアアーカイブサービスへの応用が挙げられる。

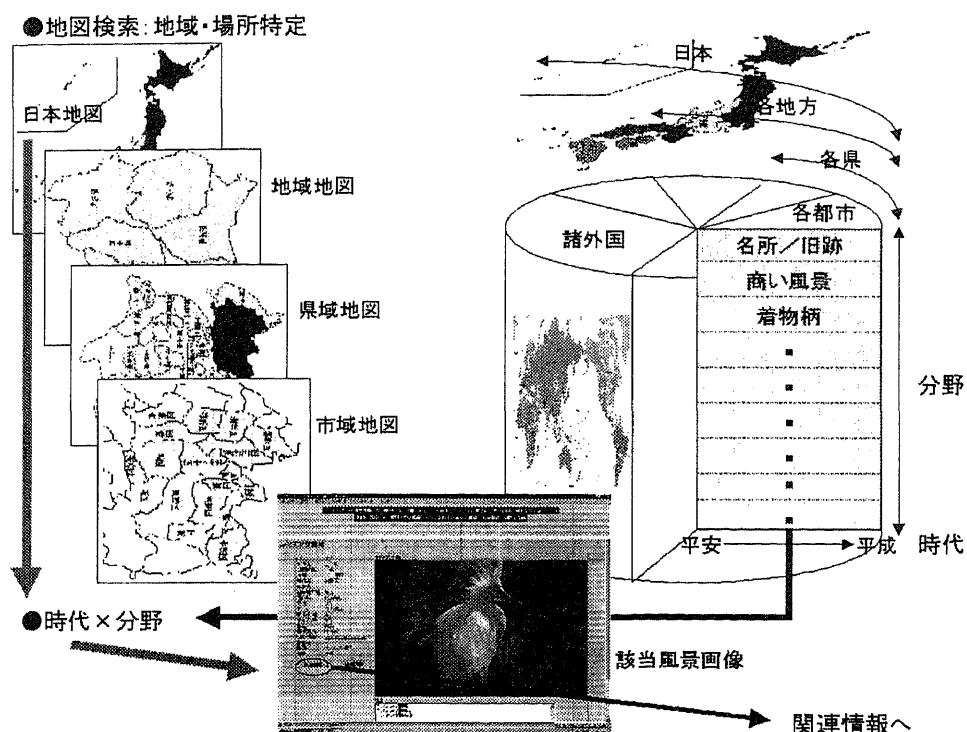


図6 階層型地図検索サービス概要

6. おわりに

性能、運用、ユーザインタフェースの観点から、電子図書館システムとしての課題および要求条件を整理し、機能拡張を行った。

本システムはインターネットでの分散配置されたサーバ間を相互に連携動作させることができる分散アーカイブシステムであり、負荷分散機能、運用管理機能の高度化、利用者にキーボードレスでコンテンツを提供する階層型地図検索サービスを実現した。

現在その評価中であるが、同時アクセス等の高負荷時や大容量のデータベースを接続したときの性能検証、システムの信頼性といった部分を中心に行っていく予定である。ユーザインタフェースの部分は、内部での検証が難しいため、電子図書館システムの利用実験等を通して、利用評価を行う予定である。

本評価試験から得られた結果に基づき、次世代電子図書館システムとしての要求条件を精査し、コンテンツ流通の促進に寄与できるよう努めていきたい。

7. 参考文献

- [1]Object Management Group: "The Common Object Request Broker Architecture and Specification", CORBA Version 2.0, 1995
- [2] 吉開, 西野, 萱野, 若原: "分散オブジェクト環境におけるネットワークライブラリの検討," 情報処理学会, 情処研報 Vol.97, No.104, pp.219-224, 1997.
- [3] 萱野, 西野, 関, 爰川: "分散ネットワーク電子図書館 NetLibra の提案," 第57回 情報処理学会全国大会, 3-1, 1998.10

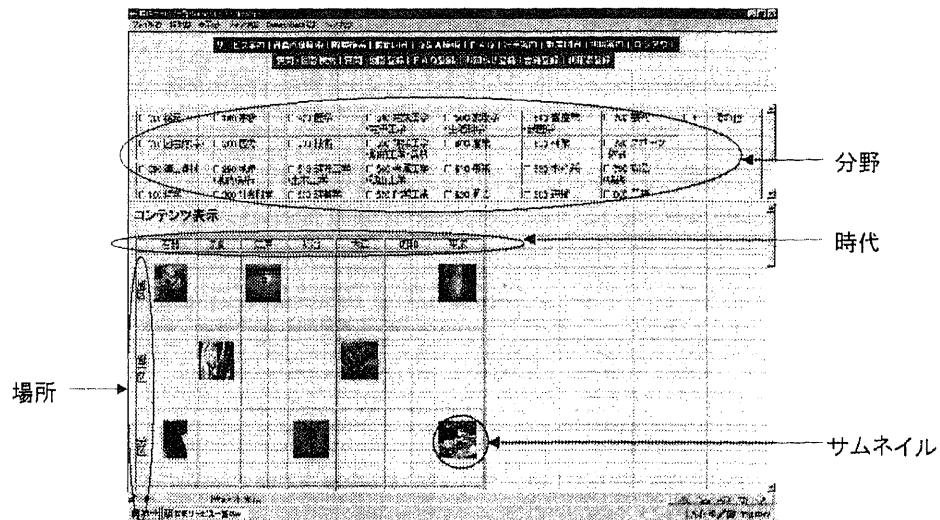


図7 階層型地図検索サービスの検索結果一覧表示画面

[4] 由比藤, 西野, 爰川, 杉山, 萱野: "分散ネットワーク図書館 NetLibra の機能構成の検討," 1999 年信学会総合大会, 1999.

[5] Kokogawa, T., Yuitou, M., Nishino, M., and Seki, Y. : NetLibra: Toward Networking Library Systems, International Symposium on Digital Libraries 1999 (ISDL'99), Tsukuba, Japan, September, 1999.