

40万件にもおよぶ 膨大な映像データの効率的な管理を HiRDBによって実現

HiRDBの提供する高度なデジタルコンテンツ管理機能を活用
読売テレビ放送株式会社

HiRDB 導入事例

読売テレビ放送株式会社（以下読売テレビ）は、メインフレームで運用していた番組ライブラリの管理システムを、HiRDBを軸にオープンシステムに全面的に置き換えた。このライブラリシステムは、従来からの番組などの映像テープ管理に加えて、個々の番組に関する複雑な権利・著作権の管理も支援する、より高度化されパワーアップした内容になっている。ここでは、HiRDB Universal Serverの提供するプラグインコンポーネント群が大きな威力を発揮している。

映像テープを メタデータベースで管理

ニュース、ドラマ、アニメからバラエティまで、日々放送されるテレビ番組の映像は、テープに記録され、膨大な映像アーカイブとしてテレビ局の倉庫に保管されている。ちょうど図書館の書庫と同じように、読売テレビの倉庫には、過去40年間に放送された番組やニュースやスポーツなどの取材映像を記録したテープがぎっしりと並んでいる。これらの膨大な映像のストックは、再放送に際して使用されるだけでなく、テレビ局内で資料として利用されたり、その一部が番組のなかで再利用されたりもする。したがって、膨大な数の映像テープを、番組、放送日時、内容や出演者などで検索し、目的の映像テープを迅速に探し出せるようにするには、ライブラリシステムが不可欠である。

映像テープを管理するライブラリシステムは、本を管理する図書館システムと似ているようでいて、独特の難しさがある。本ならば、1冊1冊がそれぞれで完結しているし、棚に並べられた本の背表紙をみれば、書名や著者名も目で探し出せる。しかし、テレビ番組などの映像テープになると、1本のテープに複数の番組や取材映像が記録されている場合があるため、本と同じようにはいかない。何よりも、映像に対する検索項目の種類やその組み合わせが、本の検索に比べて、格段に複雑なものになる。「むかしは倉庫に並んだテープの背ラベルを見たら、目的のテープを探せました。しかし、

記憶媒体がどんどん小さくなっているので、いまでは背ラベルにはその番号しか持たせられず、それを見ただけでは中に何が記録されているのかわかりません。記憶媒体が将来DVD-ROM/RAMになると、ますます薄くなってしまいます。そうすると、人間の目だけで倉庫から探すのは不可能で、媒体に記録されている内容を管理するシステムが不可欠です（読売テレビ 総務局情報システム部 木瀬浩平氏）

現場へのパソコンの普及が オープンシステム導入を後押し

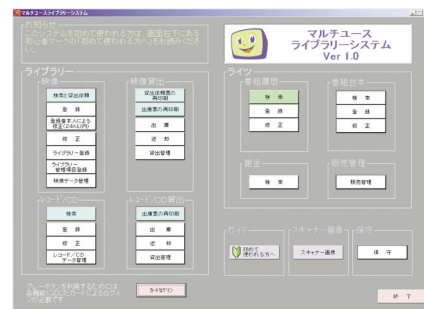
読売テレビにおける映像テープのライブラリシステムの歴史は長い。古くはライブラリ部の所管であり、ミニコンで運用していたが、レスポンスが悪い、トラブルが発生するなどの問題があり、1993年にシステム部に移管されることになった。

このとき、いったんはパソコンをベースとするオープンシステムの導入も検討された。しかし、現場の要望に応えられるだけの検索機能をもったソリューションは、まだ存在しなかったため、特許関係の検索システムなどで実績のあった、メインフレーム用のパッケージソフトが導入された。

その後、パソコンの性能と信頼性も向上し、またサーバで稼働する全文検索機能の品質も格段に向上したこともあって、資産の有効活用のために権利関係も含めて一元管理する必要からもメインフレームからオープンシステムへ

の移行が現実的に検討されるようになる。これと時期的に並行して、局内の現場にWindows®ベースのパソコンの導入が進展していったことも大きな要因であった。

「いつまでも（メインフレームの端末を使った）キャラクタベースの検索を使い続ける時代でもないだろう、一気にオープンシステムに作り直そう、という機運が盛り上がりました（読売テレビ 総務局情報システム部 林明彦氏）



画面1：新マルチユースライブラリシステム初期画面

User Profile

読売テレビ放送株式会社

1958年2月、新大阪テレビ放送株式会社設立総会東京支社として設立。同年8月には社名を現在の読売テレビ放送株式会社に变更し、準教育局として本放送を開始する。現在、その電波は近畿2府4県の全域と中部地区、四国地区の一部をカバーしており、山間部などの難視聴地域へは182か所にもおよぶサテライト局・ミニサテライト局を設置して、鮮明な画像を送り続ける。さらに地元へ密着した活動を行うために、次世代の放送システムの開発をはじめ、さまざまな観点から地域への貢献をはかっている。

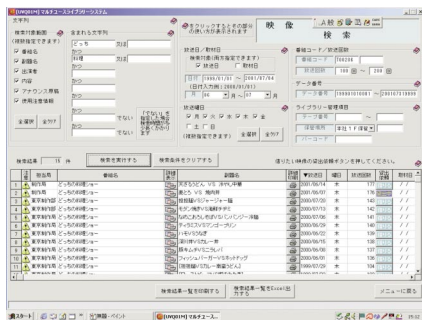
基幹システムとの連携から HiRDBを採用

2001年2月から本格稼働を開始したHiRDBベースの新ライブラリシステムには、過去40年にわたる映像約40万件分のデータが登録されている(画面1)。

基幹系から取り込むデータには、計数システムが管理する番組制作費関連データ、編成支援システムが管理する番組確定表データ、報道情報システムが管理するアナウンス原稿データの3種類がある(図1)とくに、最後のアナウンス原稿データがライブラリシステムに取り込まれる点が興味深い。つまり、ニュース番組の検索で、検索対象の1つとして、アナウンサーが読んだ原稿そのものも指定できるのである。

オープンシステムへの移行にあたって、HiRDB以外の他社製品も検討されたが、そのプラットフォームとしての機能・性能の優位性および、開発体制の信頼性などを総合的に検討して、日立のHiRDBを選んだという。

「将来的には映像自体をDBに取り込んで管理することを考えていますが、その際にも問題なく効率的に管理できることも大きな魅力でした。また(ライブラリシステムは)基幹系システムから一部データを取り込むのですが、この



画面2: 7つのフィールドをサポートする検索画面



読賣テレビ放送
株式会社
総務局
情報システム部
部次長職
木瀬浩平氏

基幹系は日立のメインフレームで稼働しています。メインフレームとHiRDBとのシームレスな連携を実現できる点も重視しました(木瀬氏)



読賣テレビ放送
株式会社
総務局
情報システム部
部次長職
林明彦氏

24時間を経過するとそのデータを参照できるのはライブラリ部だけになる。そして、ライブラリ部のチェックをパスした段階で、システムからそのデータが正規に検索できるようになる。

現在稼働しているライブラリシステムでは、利用者ごとの権限も細かく管理されている。実際に利用者がライブラリシステムを利用して、参照やデータ登録を行う際には、まず各人の社員カードをカードリーダーから読み取らせることで、社員の権限に応じて利用可能なメニュー項目だけが実行でき、それ以外のメニュー項目については、ロックがかかる仕組みになっている。

また、倉庫に保管されている映像テープの貸出と返却については、設置されているパソコン端末のバーコードリーダーを使って、映像テープに印刷されたバーコードを読み取らせることで、貸出返却システムへの操作が自動的に実行できるように作られている。

ユーザーの権限に応じた 細かなアクセス制限を実現

アナウンス原稿は、報道情報システムによって運用されているデータを、夜間のバッチ処理によってHiRDBに取り込むのに対して、番組名、出演者などの番組データのソースは、番組の制作現場で入力する。従来のシステムでは、現場のスタッフが「映像データ登録シート」という用紙に内容を記入したものを、ライブラリ部の専任オペレータがデータ入力していたのだが、新システムでは、現場のアシスタントディレクターなどが直接データ入力する運用になった。過去のデータの流用ができることもあり、データ入力に費やされてきた時間的コストが確実に削減された。

このようにして現場のパソコンから入力されたデータは、いったんライブラリ部が内容をチェックした上で、正式にデータベースに登録される。現場での入力から24時間以内であれば、データを登録した本人が内容を修正できるが、

アナウンス原稿の高速検索を 実現するテキストサーチ

HiRDBの利用によって、ライブラリシステムの使い勝手だけでなく、検索そのものの性能も大幅に向上している。

メインフレームで運用していた旧システムでは、「複雑な検索条件を指定した場合、1分くらいレスポンスが返ってこないこともありました。新システムでは、このようなことがなくなり、瞬時に結果が返るようになりました。これは、メインフレームとUNIXサーバ(HITACHI 9000V L2000-1 HP-UX11.0)のマシンの性能差というよりも、HiRDBのデータベースとしての設計上の工夫や、全文検索エンジンの性能差が大きいように思います(木瀬氏)

検索画面(画面2)では、7つのフィールド(番組タイトル、サブタイトル、内容情報、映像情報など)について、まずどのフィールドを検索対象にするのかをチェックボックスで複数選択し、次

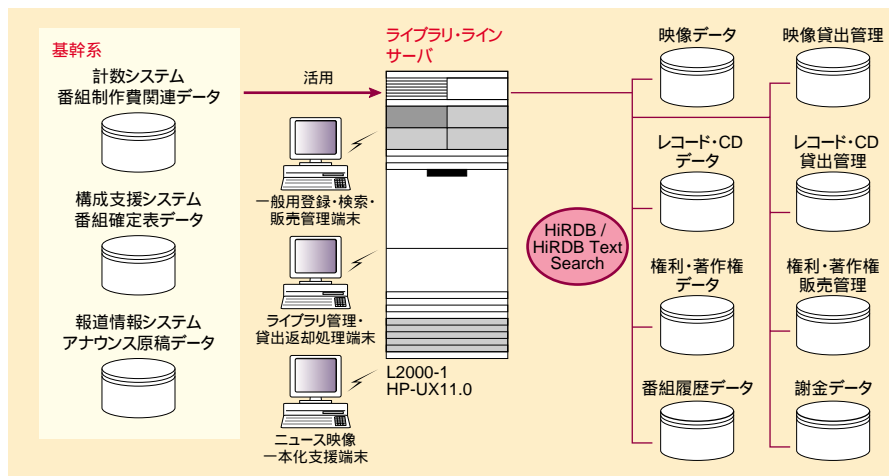


図1: 新マルチユースライブラリシステムシステム概要

にキーワード(最大9個)をAND、OR、NOTの組み合わせで設定できるユーザーインタフェースになっている。

ここで指定された検索対象フィールドは、単純にデータベースのテーブルの列名に対応するのではなく、フィールドの組み合わせパターンに応じて、特別なフィールドを検索対象にするという工夫が施されている。検索に際しては同義語(シソーラス)も参照しており、この同義語はユーザーが独自に追加・登録できる。

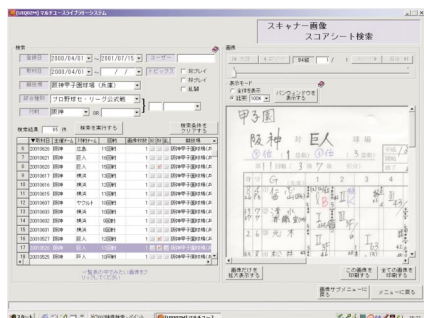
アナウンス原稿も検索対象になるため、実質的にはフルテキストサーチになる。このため、SQLの通常のLIKEを使うと、検索結果が返ってくるまでかなりの時間がかかってしまうことになる。これに対し、HIRDBのプラグインモジュールのひとつである日本語全文検索プラグイン(HIRDB Text Search Plug-in)を導入することで劇的な高速化を実現している。

この日本語全文検索プラグインは、検索対象となるテキストを、隣接するN個の文字の並びに分解したもの(n-gram)をもとにインデックスを作成し、このインデックスに対して検索を実行するため、通常のテキストサーチに比べ、はるかに高速な検索処理が実現できる。

こうした高い検索性能を実現しているのがHIRDBを中心にオプション群として用意されているHiRDB Universal Serverなのである。HiRDB Universal Serverは文書をはじめ、映像、画像、地図情報、空間情報といったデジタルコンテンツを効率的に管理・活用するための一連のプラグインを提供するものだ。

動画の取り込みは運用上困難

HiRDB Universal Serverには、テキストサーチのほか、静止画や動画に対する柔軟な検索・管理機能などもあるが、読売テレビのライ



画面3 : スキャン画像の検索画面

ブラリシステムでは、番組に関連する新聞の切り抜き、スポーツの試合の手書きスコアなどをスキャンした画像もサポートされている(画面3)。これらの静止画をHIRDB自体のデータとして取り込むことも可能なのだが、Windows® 2000のファイルサーバ機能を使って、別マシン(HA8000/170)を専用の画像サーバにして、HIRDBには画像のファイル名とパスや登録者、画像コメントなどの文字情報だけが登録されている。画像データをHIRDBで管理するとディスク容量が膨大になってしまうことから、画像サーバとHIRDBを分離した構成になっている。

HIRDBに動画ファイルを取り込める機能を利用して、たとえば番組の頭出し部分の動画をライブラリシステムに格納することも検討されたが、これについては技術的な制約よりもむしろ、運用上の制約が大きいいため、今後の対応課題として今回のシステムでは実現を見送られた。

「過去の映像はすべてアナログのテープに録画されているので、かりにビデオキャプチャボードを使って頭出しを次々にデジタル化していったとすると、1時間の番組だと1時間かかってしまいます。文字データを登録するだけでもオペレーションが手一杯な状況なので、さすがに動画の取り込みは難しいと判断しました。ただし、将来的な技術革新によって、アナログテープではなくディスクに媒体が変わり、1時間の番組が15分くらいで頭出しのカットが撮れるようになれば、HIRDBの動画ファイルの取り込み機能を利用して、番組の動画をライブラリシステムでサポートすることも、検討していきたい(木瀬氏)

多チャンネル時代には番組の権利関係の整備が必須

読売テレビのライブラリシステムは、メインフレームで運用していた旧システムをオープンシステムに置き換えただけではない。個々の番組の権利や著作権などの管理(ライセンス機能)も、もうひとつの大きな柱であり、2001年9月の本稼働をめざして、現在開発が進められている。

読売テレビのライブラリシステムは、映像テープの管理、テープに記録されている映像に関するデータの管理、そして、映像に付随するいろいろな権利関係の管理、この3つを一元管理するためのシステムとして構想されているのだ。「地上波デジタルの時代になると、チャンネルが増えるので、過去のコンテンツの再利用が課題

になります。しかし現状では、映像の権利関係の整理は、まだまだ遅れています。従来は、そうした権利関係に関するデータが、番組担当者個人の頭の中にしかなくて、担当者からそうした事実関係を引き出して、権利関係をクリアしていく、というやり方でした。それではいけないだろう。地上波デジタルの始まる2003年までに、早急に権利関係を整備しておかなければならない。そのため、権利関係の管理も含む、マルチユースライブラリシステムを実現しようというのが、今回の移行の基本目的です(木瀬氏)

番組の収支管理にむけて

ひとつの番組には、その出演者、制作会社、脚本家、構成作家など、多くの人が関わり、その各々に対してすべて権利関係が発生する。通常の契約では、地上波での一次放送と3年以内の再放送に対する使用権を、テレビ局側から関係者に支払うかたちになる。そのため、過去の番組をビデオで販売する、あるいは海外のテレビ局に販売するといった二次使用が発生する場合には、再度契約を結び直すことになる。

このため、過去の番組を二次使用するにあたっては、その番組に誰が関係していて、その関係者とのような契約を結んでいたかを確認する必要がある。当然、再契約には契約金の調整も必要になるし、二次使用によって発生した収入を、著作権者に対してどのように配分するかも調整しなければならない。こうした複雑な権利関係の調整を支援するのが、マルチユースライブラリシステムのライセンス機能なのである。

このライセンス機能は、権利関係や二次使用における売上げの配分を扱うだけでなく、番組収支という考え方を推進する上でも、重要な土台になるものとして位置づけられている。

「番組収支とは、ある番組について、費用がいくらで、収入がいくらを押さえるということです。これは計数システムによって実現されているわけですが、海外での販売で発生した収入については、現状では、別扱いになっています。本来は番組ごとに二次使用分も含めたかたちで一元管理すべきなのですが、なかなかそこまで実現できていない。そのため、ライセンス機能を実現した次の段階では、こうした二次使用における収入と計数システムのデータがリンクするシステムに発展させる計画です(林氏)

PLAY BACK! HiRDB Case Study その特長と注目ポイント

デジタルコンテンツ検索の 劇的な高速化を実現

読売テレビのマルチユースライブラリシステムでは、アナウンス原稿を含む膨大なテキストデータに対する高速な検索が実現されている。これを支えているのがHiRDBのプラグインオプションとして提供されるHiRDB Universal Serverの「日本語全文検索プラグイン(HiRDB Text Search Plug-in)」である。日本独自のインクリメンタルn-gramインデックス方式を採用することで、大量の文書情報が正確かつ高速に検索可能になっている。

HiRDBの日本語全文検索プラグインには、このほかに、SGMLで記述した文書に対して構造を意識した検索を実現する「構造検索」、検索条件に指定された単語だけでなく、その同義語や異表記も同時に検索条件として扱う「同義語・異表記検索」、検索条件を満足する度合いによって得点(スコア)を付加し、さらにスコア情報によって検索結果のランキングによるソートを可能にする「スコアリング機能・

マルチメディアユースをはじめとする 次世代データベースの構築をサポート

今回の記事では、読売テレビにおいて実現されているHiRDBでのデジタルコンテンツ管理・活用を中心に紹介した。最後にそのポイントを確認しておきたい。

ランキング機能」などが用意されている。

さらに高度な拡張を実現する イメージ検索機能

読売テレビのライブラリシステムでは、動画の管理は運用上の問題から見送られたかたちだが、今後番組の制作過程におけるフルデジタル化が進展すれば、動画データの管理も要請されることになるだろう。

HiRDB Universal Serverには、日本語全文検索だけでなく、画像・映像に対して色や形などによるイメージ検索を実現するためのコンポーネントも提供されている。キーワード検索ではなく、画像そのものをキーにすることで、キー画像と類似している画像を検索する「画像検索用コンポーネント(HiRDB Image Search Plug-in)」や、画像データをファイルシステムのような外部記憶装置に格納し、大容量コンテンツの入出力性能を従来のBLOB方式に比べて、飛躍的に向上させる「データ管理用コンポーネント(HiRDB File

Link)」や、画像を一覧するためのサムネイル画像の作成や画像の合成など、複雑な画像操作部品を提供し、さらには独自の画像合成技術により、画像上に透かしマークを重ねて表示することで著作権情報を画像に埋め込める「システム構築用コンポーネント(Image Database Access)」などがある。

またHiRDB Universal Serverの「映像管理システム構築コンポーネント」は、動画のどこにどんな映像が含まれているのか探し出すのを支援する機能を提供する。これを使うことで、たとえば、映像テロップやシーンの変化点を自動的に検出した上で、映像の構成内容をインデックス化して表示するという高度な機能も実現できる。

また、記述内容をXML準拠の形式で保存することもできる。具体的には、映像データのデジタイズ(MPEG-4へのエンコード)と検索用インデクス付け機能を提供する「Video Indexing System」、映像データの簡単なデータベース登録・検索用インタフェースを提供する「Video Database Access」、そしてビデオサーバと連携したスケーラブルな映像データ管理機構を提供する「HiRDB File Link Stream Option」などが用意されている。

お問い合わせ 日立オープンミドルウェアお問い合わせセンター
TEL***** 0120-55-0504
URL***** <http://www.hitachi.co.jp/soft/hirdb/>
e-mail****hirdb@itg.hitachi.co.jp



HiRDBの優位性と自社製品ならではの 迅速なサポートが評価いただけたと思います

株式会社 日立システムアンドサービス プロダクトソリューション事業部
主任技師 笹尾和正氏(写真右)
株式会社 日立システムアンドサービス 産業・流通システムサービス事業部
山口 賢氏(写真左)

読売テレビ様のシステムのポイントは何と言っても、同社が過去40年にわたって蓄積してきた約40万件にもおよび映像データについての属性情報やアナウンス原稿などの膨大なテキストデータをいかに効率的に管理し、高速な検索機能を提供するかということにほかなりません。こうした厳しい要求に応え得たのは、HiRDB Universal Serverの提供するハイパフォーマンスな高速全文検索をはじめとするコ

ンポーネント群があればこそと自負しています。また、基幹系システムからのデータ取り込みにおけるメインフレームとのシームレスな連携を実現している点でも、HiRDBのメリットが存分に活かされています。今後は、HiRDB Universal Serverの提供する優れた機能群を活用して、映像データの管理・検索を含めた、さらに高度なマルチメディアコンテンツ管理機能を提供していく予定です。