

4H-5 統合OAシステムに発展するAS-Documents 文書データベース管理機能 DocRDB[†]

白井直裕^{††} 望月祐之^{††} 湯川敦司^{††} 今村泰介^{††} 角谷清夫^{††}
(株)東芝 府中工場^{††}

1. まえがき

ワードプロセッサやDTP*システムの普及に伴い、ドキュメントの電子化が進められている。電子化されたドキュメントは順次蓄積、管理され再利用されるようになった。しかしこれらの電子化されたドキュメントはその量が次第に膨大となり、目的とするドキュメントを探し出すのにも様々な工夫が必要となってきた。

文書処理システムAS-Documentsでは、大量に作成された文書を登録、検索する文書データベースシステム DocRDBを開発した。(今後は略して DocRDBと呼ぶ。)

文書データベースシステムはドキュメントを検索するためのキー情報をリレーショナルデータベースシステムであるORACLE**によって管理し、文書の実体はキー情報の1つであるファイル名とUNIXファイルシステムにより管理している。さらに機能の分散を行い、特定ワークステーションの負荷の軽減とヒューマンインタフェースの向上を図っている。

これによりデータ検索の効率が高く、文書実体をも同時に引き出すことができる文書データベースシステムを構築することができた。

2. 構成と機能

データベースアクセス用の言語SQL***は問い合わせ文、データ操作文(Data Manipulation Language statement)と、データ定義文(Data Definition Language statement)、データ制御文(Data Control Language statement)に分けられる。

このうち、問い合わせ/データ操作の機能をデータ操作ツール、データ定義/制御の機能をデータ定義ツール、そしてリレーショナルデータベースシステムORACLEと文書実体の管理を行う機能を文書データベース管理プロセスとして開発した。これら3種の機能を総合して文書データベース管理機能 DocRDBと呼んでいる。

以下でDocRDBの3種の機能について説明する。

2.1 データ定義ツール(DocDDT)

本機能は、キー情報を管理するための情報定義を行う。すなわちキー情報のテーブルの定義、編集、削除の指定、ビューの定義、削除、ユーザごとのアクセス権限の許可や取り消しなどのデータの保護に関する指定などを行う。

2.2 データ操作ツール(DocDMT)

本機能は、キー情報・文書実体の検索、登録、更新、削除を行う。すなわちORACLEに対するキー情報とそれに対応する文書実体の検索、登録、更新、削除、の条件の指定、実行、およびキー情報テーブルの検索結果の表示を行う。また検索した文書実体を文書編集機能DocMakerに通知して、文書の内容を直接表示することもできる。

2.3 文書データベース管理プロセス(DocDBM)

データ定義ツール、データ操作ツールにより入力された指定の条件でリレーショナルデータベースシステムORACLEと連係して、文書データベースを構築する。

3. DocRDBのシステムアーキテクチャ

DocRDBは図1に示すようなネットワークシステム上で動作する。各ワークステーションはイーサネット^{†††}で接続され、分散処理環境を実現している。これらのワークステーション上の一台にデータベースサーバーとして文書データベース管理プロセスがあり、その他のワークステーションはデータベースクライアントとなる。データベースクライアントはデータ定義ツール、データ操作ツールによって文書データベース管理プロセスとデータ通信を行い、データベースのアクセスを行う。

またこれらのワークステーションはNFS****(Network File System)によりネットワーク上で同一のディレクトリ構成を実現しているため、どのクライアントから見ても同一のファイル構造でアクセスできる。

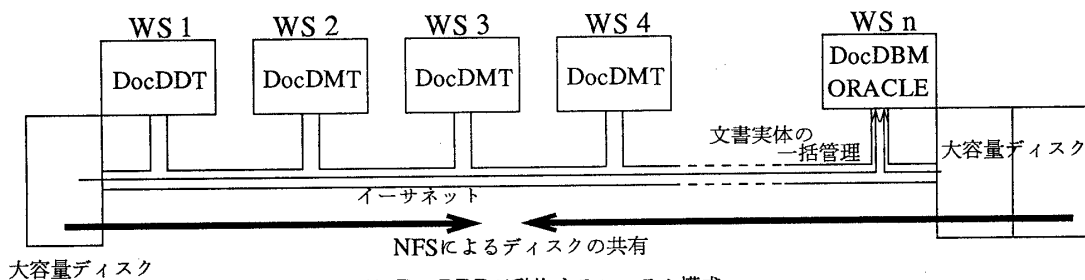


図1 DocRDBが動作するシステム構成

[†] AS-Documents growing for Integrated Office Automation System - Document Database Management System

^{††} TOSHIBA Co. Ltd FUCHU WORKS

4. データベースアクセスの分散処理

従来のデータベースアクセス方法は、データベースサーバとなっているワークステーションにリモート端末機能を用いてアクセスしていたため、データベースサーバが全ての処理の負荷を負う。従って端末を増やすと、仕事量に対してデータベースサーバの速度の低下が著しい。また入出力制御が全てキャラクタコードのため、ヒューマンインタフェースにポインティングデバイスを使用ができなかった。

データ定義ツール、データ操作ツール、文書データベース管理プロセスは、互いに完全に独立したプロセスとして機能分散されている。そのうちデータ定義ツール、データ操作ツールの各プロセスはポインティングデバイスとしてマウスを採用した入出力制御を行っており、ヒューマンインタフェースを向上させた。

ユーザはデータ操作ツールでデータベースアクセス条件を設定し、文書データベース管理プロセスに対してアクセス要求を出す。文書データベース管理プロセスは要求を受け取ると、ORACLEに対してデータベースのアクセス要求を実行しその結果を受け取り要求元のデータ操作ツールに通知する。要求を発行したデータ操作ツールでは結果を受け取りユーザに表示する。

データ定義ツール、データ操作ツールと文書データベース管理プロセス間のデータ通信に遠隔手続き呼び出しRPC(Remote Procedure Call)を用いることで、文書データベース管理プロセスは必要な時のみ動作するため、単位時間あたりの負荷が軽減され、データベースアクセスの高速な分散処理を実現している。

5. データベースアクセスの負荷分散

文書データベース管理プロセスは内部で負荷の分散を行っている。通常のデータベースシステムでは1つのデータベースサーバに対して複数のデータベースクライアントがアクセス要求を出す。複数のデータベースアクセス要求に対し、データベースサーバ内には複数の要求を同時に受け付けられるように同じ機能のプロセスを複数起動する。

これらのプロセスは通常データベースへのログオン要求時に起動され、ログオフ時に終了するが、一定時間内にこのプロセスに対してアクセス要求がない場合自動的に終了する。これによってデータベースサーバの負荷を少なくして、データベースクライアントに対する効率的な応答を確保した。

6. 文書実体の管理

文書に対応するキー情報のアクセスはデータ操作ツール、文書データベース管理プロセス、ORACLEの手順で行われるが、文書実体の管理、アクセスは文書データベース管理プロセスのみが行っている。

*DTP = Desk Top Publishing 卓上出版システムのこと

**ORACLEはORACLE社の登録商標で、リレーショナルデータベースである。

***SQL データベースアクセス用の言語。世界的に標準化が図られている。

****NFSはSun Microsystems社の登録商標である。

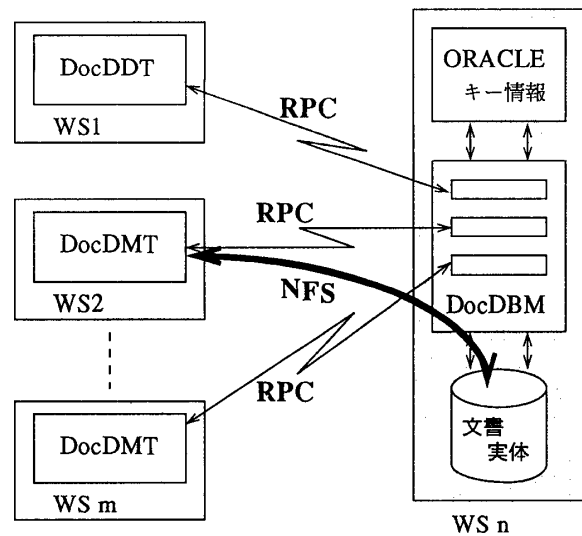


図2 DocRDBの構成

ユーザはデータベースに登録したい文書のファイル名をデータ操作ツールを用いてキー情報の1つとして設定する。他のキー情報設定が終了したらキー情報と格納する文書実体の名称を文書データベース管理プロセスに通知する。文書データベース管理プロセスは、キー情報をORACLEに対して登録し、文書実体はNFSを通じて文書データベース管理プロセスによって読み出され、システム構成時にあらかじめ定義された領域へ書き込まれる。文書データベース管理プロセスによって管理される文書実体は文書データベース管理プロセスに従ったファイルのアクセス権定義が行われるため、一端登録されたドキュメントを勝手に書き替えることができなくなる。

7. あとがき

DocRDBはユーザが複雑なSQL言語を操作しなくても、データベースをアクセスする機能とヒューマンインタフェースを実現し、電子化されたドキュメントをキー情報付きで管理することを可能とした。また、文書の索引情報のみをアクセスするだけでなく、文書実体をも同時にアクセスすることを可能とした。

その結果、DocRDBによって汎用的な文書データベース、及び文書アクセス環境を構築することが可能となった。

また、DocRDBは複数のデータベースに対してもRPCによるアクセス要求が発行できるシステムアーキテクチャを持っており、ネットワーク分散データベースを志向している。

参考文献

- [1]芝野 耕司:データベース言語SQL、情報処理、Vol.28、No.3、p.208-214
- [2]データベース言語標準SQL、日経コンピュータ、1987.12.21、(no.163)、p.50-74