

## FTP サイトを検索するポータルレットの開発

渡邊 泰治<sup>†</sup> 井堀 佑美<sup>†</sup> 三浦 孝夫

法政大学工学部電気電子工学科<sup>‡</sup>

### 1. 前書き

本研究では、いつでも、どこでも、誰でも最適(効率、最新、確実性において)に FTP ファイルを入手するアプリケーションの実現開発について論じる。

FTP コマンドは現在に至るまでインターネットで広範囲に利用されているが、ファイルの入手に関して次のような問題がある。

1. GUI のための専用ソフトウェアが必要である
2. ファイル保管場所を調べるには個別にサイトディレクトリを探索する必要がある
3. 特定のスーパーFTP サーバに集中していない

この結果、FTP 利用者はホストやファイル構造、ファイル名を個別に特定しておく必要がある。

本論文ではこれらの問題点を解決するため、FTP サーバにあるファイルを半自動的に検索・取得し、ファイルリストを定期的に保持して、Web アプリケーションとして実現する方式と試作システムを論じる。

### 2. 機能統合のためのアイデア

FTP ファイルの半自動的な検索・取得のための制御を統合的に行うための概略を示す。利用者はファイル名を入力し、検索結果を選択することでファイルを取得する。そのために、あらかじめ FTP サーバに存在するファイルのリストを用意する。ファイル情報はデータベースとして管理する。図 1 に制御の流れを示す。

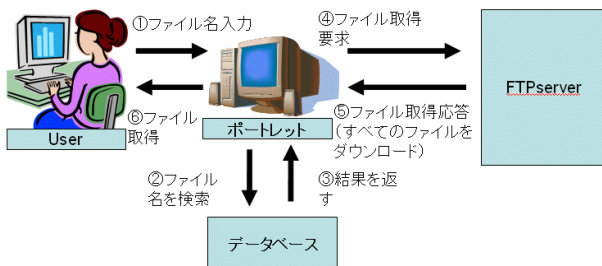


図 1. 制御の流れ

### 3. システム実現のための基本機能

#### 3.1 FTP

FTP(File Transfer Protocol)はファイルを転送するときに使われるシステムである。FTP サーバに接続するとき、最初にユーザ認証が行われる OS に

“Implementing Portlet for multiple FTP sites”

Yasuharu Watanabe, Yumi Ithori, Takao Miura  
Hosei University, Dept. of Elec. And Elec. Eng.  
Kajino-cho 3-7-2 Koganei, Tokyo, JAPAN

依存することなく使用可能な利点があるが、ファイルの検索やファイル情報の新旧の判定、ファイルの入手の手順を手動で行うという欠点がある。

#### 3.2 ポートレット

本試作システムはポートレットに基づいて開発する。ポートレットはポータルサイトを構成する部品であり、FTP ファイル取得用ポートレットならば Web 経由でファイルの取得が可能であること、複数あるポータルサイト間でファイルリストを更新することでファイルリストの保持を可能にすること等の利点を持つ。

### 4. ポートレットのデータ管理

本章ではポータルサイトで扱うデータベースを示す。これらは JDBC を介して利用する。

#### 4.1 ポートレットのデータ管理

FTP サーバと接続するために FTP サーバの URL、ユーザ名、パスワードが必要である。FTP サーバのファイル内容を最新に保つため、サーバを更新した日時を加えて FTP\_TABLE で管理する(表 1)。また、ファイルの検索・取得にはファイル名、ファイルの存在するディレクトリ、FTP サーバの URL が必要である。これらにファイルの情報(ファイルのサイズ、ファイルの更新日時)を加えて LIST\_TABLE で管理する(表 2)。

表 1. FTP\_TABLE

ftpUrl	ftpUser	ftpPwd	update
ftp.biglobe.ne.jp	anonymous	*****	2004 11 20
ftp.ring.gr.jp	anonymous	*****	2004 12 20

表 2. LIST\_TABLE

filename	directory	size	date	ftputl
md5	pub/48pub/security/bin	15884	1998 01 02	ftp.biglobe.ne.jp
Sum	pub/48pub/security/bin	5716	1998 01 02	ftp.biglobe.ne.jp

#### 4.2 ファイルリストのデータ管理

FTP サーバのファイル検索・取得のためにはデータが常に最新でなければならない。そのためファイルリストの更新をする。このため、(1) FTP サーバから直に情報を得てファイルリストを更新し、(2) 他のポータルサイトのファイルリストと比較

して更新するの 2 通り考えられる。(1)ならば常に最新の情報が得られるが、複数のポータルサイトが FTP サーバに接続するため FTP サーバにかかる負荷が大きくなる。このため本試作システムでは(2)にしたがってファイルリストを更新し、ポータルサイトからデータを取得し更新する。

最新のファイルリストを持つポータルサイトの情報を管理するデータベース(更新用データベース)が必要となる。更新用データベースが持つ情報は更新した FTP サーバの情報(URL, ユーザ名, パスワード)とその FTP サーバの最新の情報を持つポータルサイトのデータベースの情報(URL+データベース名, ユーザ名, パスワード)と更新を実行した日付である。図 2 にファイルリスト更新の流れを示す。

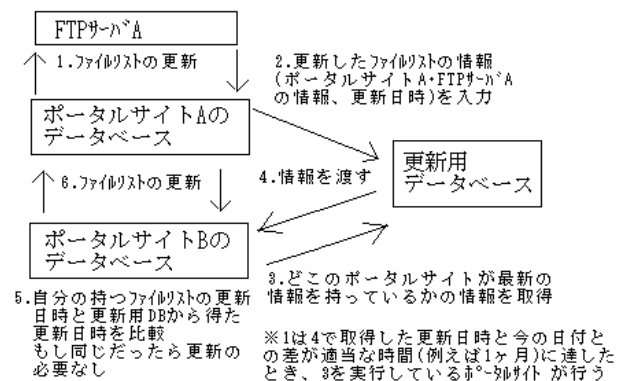


図 2. ファイルリスト更新の流れ

## 5. 統合システム

本システムは FreeBSD4.6.2Release, Postgresql, Apache1.3.29, Tomcat4.1.30, Jetspeed(ポータルフレームワーク), J2sdk1.3.1 を用いて実現する。

本システムはポートレット, データベースの検索, ファイルの取得, ファイルリストの管理から構成される。ポートレットの作成手順は次に従う。

- (1) ポートレットを Jetspeed に認識させるための xml 形式のレジストリファイルを作成
  - (2) 画面表示するための JSP ページとコントローラとなる Java クラス, そして画面に表示する内容を格納する JavaBean を作成
  - (3) 1, 2 で作ったファイルを所定の場所に配置
- 作成されたポートレットは利用者ページに容易に追加・配置することができる。

データベースの検索には JDBC を使い, 入力されたファイル名に該当するファイル情報を応答として返す。これによって FTP サーバのアドレス, ユーザ名, パスワード, ファイル名, ファイルが存在するディレクトリ, ファイルのサイズを得る。

ファイル自体を取得するために FTP サーバに対して(ソケット通信で)FTP コマンドを送信する。送信するコマンドはデータベースの検索で得られた

情報より LOGIN FTP サーバの URL, USER ユーザ名, PASS パスワード, CWD ファイルが存在するディレクトリ, PORT IP アドレス+ポート番号(プログラムから取得), RETR ファイル名, QUIT となる。各コマンドを送信する毎に FTP サーバからの応答が正しいかを判定している。取得したファイルは任意のフォルダに保存し, ポートレットにはハイパーリンクを表示する。

ファイルリストにない FTP サーバの情報を新たに入力する場合には LIST と CWD コマンドを用い FTP サーバのディレクトリを調べデータベースに入力する。更新には ftpurl をキーとして更新用データベースの更新日時と FTP\_TABLE の更新日時を比較。もし古かった場合, 最新のファイルリストを持つポータルサイトのからファイルリストを取得し, 自身のファイルリストと比較し差分だけを更新する。これらの機能をひとつのポートレットにまとめる。ファイル取得の実行例を図 3 に示す。

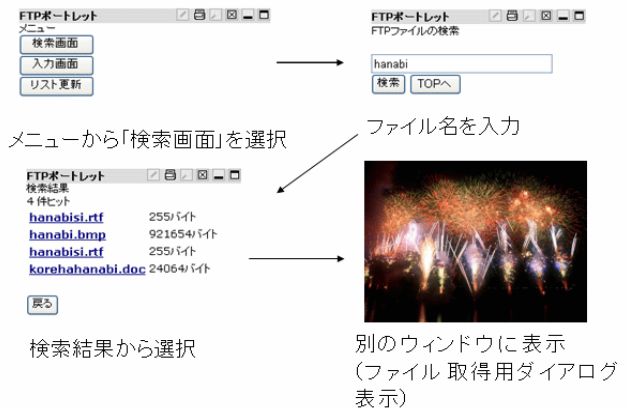


図 3. 実行例

## 6. 結論

本試作システムにより, ポートレットを使用するために特別なソフトは不要となった。データベースを用いることで複数の FTP サイトとあらゆるディレクトリを検索できるためファイルの検索の効率は上がる。複数のポータルサイトでファイルリストを更新するため FTP サーバにかかる負荷が小さくなる。

今回開発したポートレットは先に述べた FTP の問題点を解決している。また, ポートレットを用いることにより Web 上での FTP ファイルの容易な検索を可能にし, FTP サーバの情報をデータベースにすることで検索と更新の効率を康成させることができるものとなった。

## 7. 参考文献

- [1] 下村隆夫: Java によるインターネットプログラミング, 近代科学社, 2002