

ネットワークを利用した環境情報データ自動収集サーバシステムの開発

(環境計測) 西木陽平

1. はじめに

全国的な大気汚染物質のデータのオンライン提供が、2001年度から環境省の大気汚染物質広域監視システム、通称「そらまめ君」のページを通して行われるようになった。そのデータは、毎時更新され、地方別の図や表で表示される。従来、多くの人々にとって大気汚染物質のデータは、1年分が集計された紙のデータを入手するか、環境省からCD、磁気媒体などを入手するしか方法がなかったことから考えると大きな前進である。しかし、このシステムでも利用者側からの使い勝手は十分とはいえない。例えば、データの保存期間は1週間と短く、過去のデータをさかのぼって読み出せないし、提供されるデータが地図表示や表表示では、データの広域比較などの2次的利用には不便である。

そこで本研究では、長期間に渡り広域データを自動的にダウンロードし、格納し、2次的利用に対応できるデータサーバを、Linuxを用いて開発することにした。

このデータサーバを、2002年度卒業研究で丸橋[1]が作成したJavaによる環境情報グラフ化アプリケーションと結合させることで、環境情報提供の新たなスタイルを確立できると期待される。

2. ネットワークプログラムの開発の概要

インターネット上の通信は、socket(ソケット)を用いて行われる。socketをサポートする言語としては、C、Perl、Javaなどがあるが、今回はデータ取得後の処理を考慮し、保守や改良の容易なPerlを用いることにした。

socketとは、2つの通信し合うプロセス(サーバとクライアント)の両端を表す抽象的な概念をソフトウェアとして実現したもので、socketを介してプロセス間の通信ができるようになっている。ネットワーク上でデータがどこから来て、どこへ行くのかははっきりさせるため、IPアドレスとポートアドレスを使用する。

プログラムの概要は、クライアント(データサーバ)側がサーバ(そらまめ君のWWWサーバ)側に対してリクエストを送信し、次にサーバ側がクライアント側へヘッダとボディを送信する。送信が終了したらクライアント側がコネクションを切断する。これを利用し、そらまめ君のWWWサーバからHTMLファイルを取り出し、そこから必要な大気汚染物質のデータを取得する。その後は、サーバにデータを保存し、取得してきたままのデータ(そらまめ君のデータ形式)と京都府保健環境研究所のデータ形式(保環研のデータ形式)に変換したデータをweb上に公開する。(図1参照)

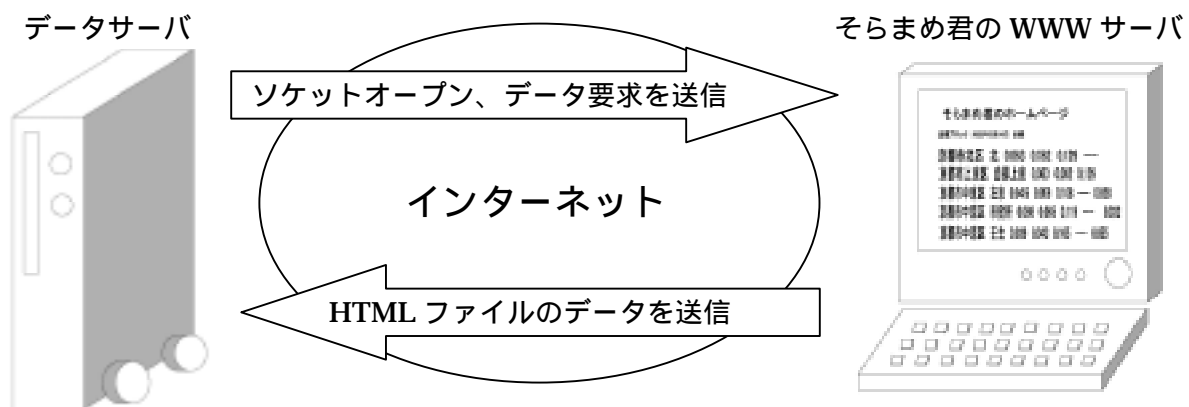


図1 ネットワークプログラムの概要

3. データサーバ構築の概要

長期間に渡り大気汚染物質のデータを保存し、誰もが容易に取得し、2次的利用ができるようにデータサーバの構築に取り組んだ。

開発環境として、プログラミング言語は、Perl(ver.5.8.0)を用い、サーバのOSにLinux(ver.2.4.18-14)、webサーバにはApache(ver.2.0.40-8)を利用した。

データサーバの備える機能は以下の通りである。

- cron (クローン) とシェルスクリプトを利用し、定時に自動でデータを取得する
- データ取得エラー時に管理者へメールを送信する
- データを取得してきた時に、そらまめ君のサーバ側で欠損が生じ、何ヶ所か測定局や測定時間が抜けている場合に欠損を表す 0 を代入する
- 取得したデータを保環研のデータ形式に変換する
- 自動で HTML を作成し、web 上にデータを公開する

Linux には、cron と呼ばれるデーモン (常駐プログラム) があり、crontab という設定ファイルに従って、決められた時間に決められた仕事をこなす。

今回、データサーバには、2つのシェルスクリプトを cron に設定した。シェルスクリプトとは、複数のコマンドが連続して動作するもので、シェルの文法と正規表現を組み合わせ記述されている。

初めに、データ取得プログラムとデータ取得エラー時にメールを送信するプログラムを記述したシェルスクリプトを、そらまめ君のサーバ側でデータの更新時間がずれることを考慮し、1時間に2回、30分おきに実行するように cron に設定する。

次に、測定局や測定時間が抜けた場合に 0 を代入するプログラム、保環研のデータ形式に変換するプログラム、自動で HTML を作成するプログラムを記述したシェルスクリプトを、1日のデータ取得が終了した時点で、1日に1回起動させるように設定する。

そらまめ君のサーバ側でエラーが発生し、データが更新されなかったり、更新されても何時間が抜けていたり、測定局が何ヶ所か抜けていた場合に、管理者にメールを送信するようにした。

最後に、そらまめ君のデータ形式と保環研のデータ形式を web 上で公開し、誰でも容易に利用できるようにした。(図2参照)

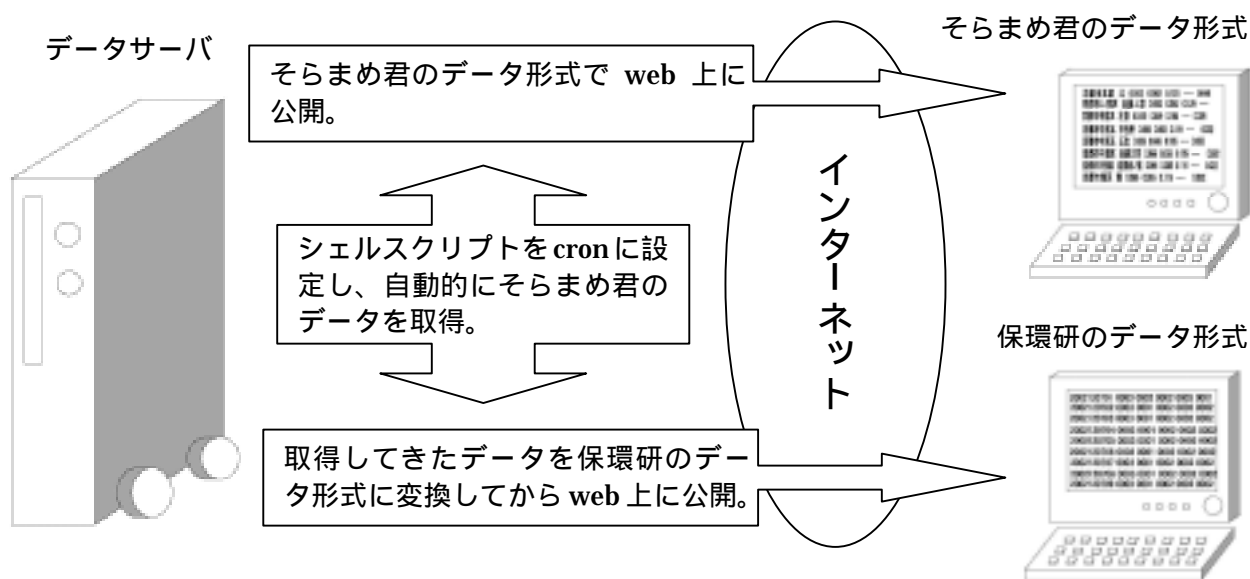


図2 データサーバの概要

4. 今後の展望と課題

本研究では、大気汚染物質のデータサーバの構築に限定したが、今後は、地震の即時データや気象データなど他にも様々な環境情報のデータを提供できるサーバに応用できる。

また、今回は京都府に限りデータを収集したが、全国 47 都道府県のデータを収集するデータサーバの構築にも使用できるように開発した。

しかし、課題として改善すべき点がある。今、そらまめ君のホームページのデータを取得できなかった時は、0 を代入している。また、ホームページのアドレスや構造が変わった時には、プログラムが動かなくなる。これらの課題に対しては、プログラムに汎用性をもたせ改善させる必要がある。

参考文献

- [1] 丸橋史 京都府立大学人間環境学部環境情報学科卒業論文 (2002)