

平成15年度未踏ソフトウェア創造事業「未踏ユース」 (多種環境間で協調するファイル同期ツールの開発)

— 様々な場所、様々な形で情報を共有したい —

1. 背景

昨今、キーワードは「モバイル」からその先の「ユビキタス」へ向かっている。ノート PC 等を持ち歩き、出先や自宅でも作業を行う、それが当然のように行われている。そのような時代に求められるファイル同期とはどのようなであろうか？

オフィスや自宅、それぞれの環境によって周囲のコンピュータ環境も大きく異なっているだろう。オフィスにはファイルサーバが設置され、自身のコンピュータ環境を自動的にあるいは定期的にバックアップしているかもしれない。パワーユーザであれば、自宅にも作業 PC があり、自宅では作業 PC で作業を継続し、その結果をまたモバイル PC で持ち出すといった、ワーキングセット（作業領域）の同期が常に行われているであろう。一般のユーザであっても、出先や自宅で作業を行った結果を何処からでも取得したい、といったニーズは高まっており、インターネットディスク等が活用されているであろう。また、オフィス等で VPN (Virtual Private Network, 仮想プライベートネットワーク) が利用可能であれば、ファイルサーバへ直接バックアップすることも可能となるだろう。

2. 目的

本プロジェクトの目的は、多様な環境間で利用可能でかつユーザが容易に導入可能なファイル同期ツールの開発である。

当初の目的はネットワーク上の違いによって生じる通信環境の差異を吸収することとしていたが、途中よりプロジェクトの方向転換を検討することとなった。その背景として SoftEther (平成 15 年度未踏ユース「イーサネットのソフトウェア実装とトンネリングシステムの開発」、開発者：登 大遊) のように、通信環境の問題を解決する優れた VPN ソフトウェアが台頭してきたことがある。

結果、「多種環境間で協調する」というコンセプトを元に新たな目的を立てることとした。それは、同期相手とする計算機との間に環境上の隔りがあることを考慮したソフトウェアを実装することである。特にサーバーコンピュータとモバイル PC といったストレージ資源の隔りを考慮し、それらが混在する環境上での適切な動作・機能を念頭において設計・開発することとした。将来的には、PDA(Personal Digital Assistants, 携帯情報端末) までもサポート対象に加えることを考慮している。

3. 開発の内容

本プロジェクトで開発したファイル同期ツール（以降 psync と呼ぶ）の機能について説明する。本ソフトウェアはフリーソフトウェアとして公開する予定である。

3.1. 動作環境

主たる処理系には Ruby を利用し、GUI(Graphical User Interface)の構成には GTK+2(The GIMP Toolkit)と Ruby/GTK2 を利用した。オペレーティングシステムへの依存を避けるよう設計しているために移植性が高く、現在は Linux 環境をターゲットにしているが、今後は他プラットフォームのサポートも検討している。

Ruby には、UNIX 系や Windows 系など、簡単に多くのプラットフォームをカバーできる利点や、GUI 生成ライブラリである Ruby/GTK2 が利用できる利点がある。

一通りの実装を進めた後、性能面での見直しが必要となった場合には適宜 C 言語で書き換えることとしている。

3.2. ネットワーク環境によるポリシーの設定

オフィスや自宅、それぞれの環境によってコンピュータを取り巻く環境は大きく異なる。よって、ファイル同期を行う際にも環境に合わせたポリシーの設定・切り替えが必要となる。psync では「現在の環境」をネットワーク構成情報より識別し、予め設定されたポリシーを自動的に適用する機能を持つ。

3.3. 既存プロトコルをベースとしたファイル転送

ファイル転送には、転送処理を抽象化した上で既存のプロトコルを利用しており、独自プロトコルの実装を避けている。転送処理の抽象化には VFS(Virtual File System, 仮想ファイルシステム)の概念を踏襲している。

既存のプロトコルを利用することで、同期相手にソフトウェアをインストールすることが難しい場合でも psync を利用することができる。

3.4. ファイル内容の適切な変換

同期を取る双方の環境でハードディスク容量等のリソースに大きな差異がある場合、全ての情報を完全な形で同期することは難しい。

この場合、psync では必要最小限の情報（サムネイルやヘッダ情報）へ変換したファイルを転送することで、ファイルを同期した「つもり」で済ませることができる。これにより、これまでリソース上の問題で同期の対象外としていたファイルでも psync を利用することで大まかな情報のみを取り出して同期することができる。今後、完全な情報を必要に応じて復元する仕組みも検討している。

3.5. 更新ログによる効率的な同期

psync では、更新ログを利用して効率的な同期処理を行うことが可能である。

更新ログとは変更のあったファイル情報を時系列に記録したものであり、周期的なディレクトリ走査によって（もしくはオペレーティングシステムの機能を利用してリアルタイムに）生成される。この更新ログを利用することで、ディレクトリ全体のスキャンを行うことなく変更のあったファイルを知ることができる。

ファイル内容の変換機能と更新ログの機能を利用することで、モバイル環境向けのファイル同期を実現している。このイメージを「図 1：ファイル内容の変換機能と更新ログを利用したイメージ」に示す。

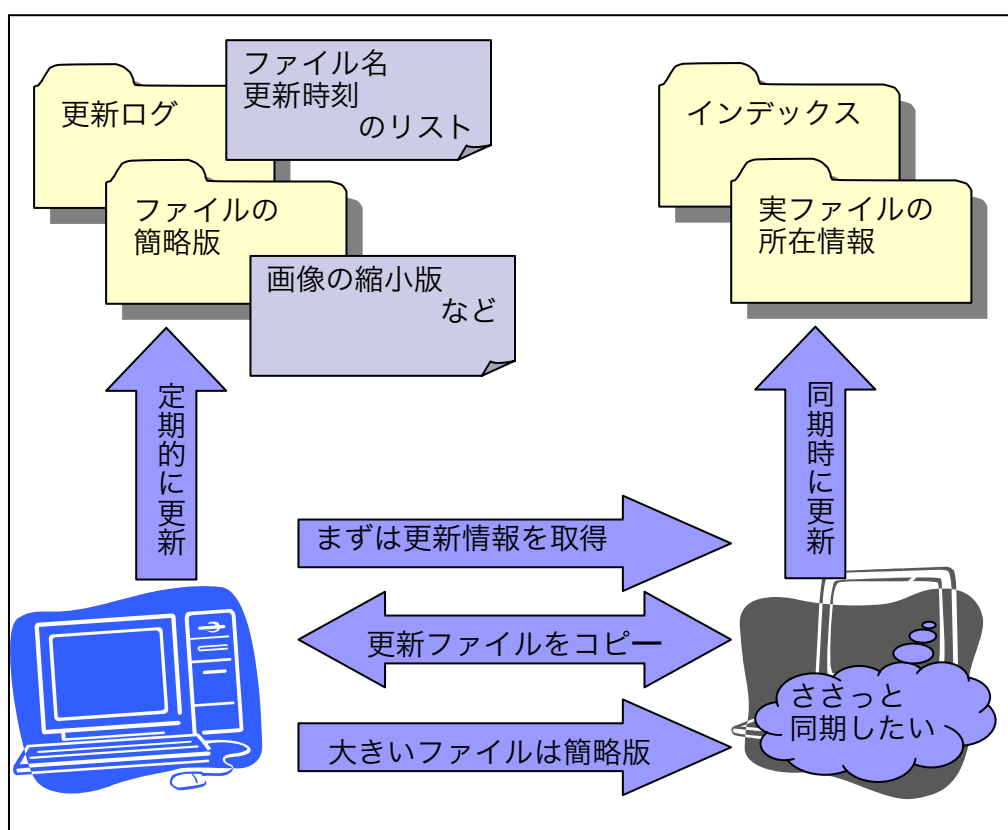


図 1：ファイル内容の変換機能と更新ログを利用したイメージ

4. 従来技術との相違

ファイル同期をテーマとした研究はこれまでも数多く行われており、本テーマと類似した技術を備えたソフトウェアも多い。ここでは psync が備えていると考えられるアドバンテージについて述べる。

- 極力、オペレーティングシステムの機能から独立していること
psync は高い移植性を目標としており、オペレーティングシステムへの極端な

依存を避けるような設計が行われている。しかし、オペレーティングシステムの機能を用いることでファイル同期を効率よく行えることも事実であり、これが如実に表れる例として変更ファイルの検出作業がある。

psync ではこのような作業を更新ログの生成として予め行うことにより、高速な更新チェックを実現するよう努力している。

- ファイル内容の変換機能

多くのファイル同期ツールでは、その動作においてファイルの内容を考慮することが無い。またモバイル環境における切断状態（ネットワーク上のファイルを操作できない状態）を考慮し、そのような状態でもファイルへのアクセスを実現している分散ファイルシステムも存在するが、その場合でも必要とするファイルのプリフェッチ（先読み）を必要とする。切断状態でプリフェッチされていないファイルは、そのファイルの内容はおろか、それがどのようなファイルであったかも知ることが出来ない。

これに対して psync では、ユーザの指定に応じてファイルの概要版（プレビュー）のみを常に保持しておくことができる。この機能は、デジタルカメラが生成する画像ファイルのような、ファイル名だけで内容を推し量ることが困難なケースではアドバンテージになると考えられる。

5. 期待される効果

psync が対象としている分野としてインターネットストレージが挙げられる。ファイル同期では、同期相手でファイルの転送を可能にしておくこと、すなわち常時稼働しているサービスを必要とする。

インターネットストレージは、サービスを常に稼働させることがないエンドユーザを対象とした「同期の仲介役」として機能すると考えられる。psync では、この「仲介役」を利用したファイル同期の実現も視野に入れている。

6. 普及の見通し

本ソフトウェアは未だ開発中であり、現在のところ成果物の公開は実施していない。そのため普及の見通しを具体的に示すことは難しい。

7. 開発者名

村崎 大輔（電気通信大学大学院 情報工学専攻 violet@takopen.cs.uec.ac.jp)

(参考) 開発者 URL

<http://ne.cs.uec.ac.jp/~violet/>