

ユビキタス計算機環境のためのオンラインストレージによる分散仮想ディスクシステムの開発

野尻祐亮[†] 並木美太郎[‡]

[†] 東京農工大学工学部

[‡] 東京農工大学共生科学技術研究院システム情報科学部門

1 はじめに

個人データやOSのデスクトップ環境を、自宅・外出先など場所を問わず利用したいという要求がある。本研究ではそれに応えるため、高い機能性や利便性を持ち、ユビキタスにアクセス可能な仮想ディスクシステムを開発する。

概念図を図1に示す。本研究の基本概念は、インターネット上のHTTP, FTPサーバやオンラインストレージサービスなどを、仮想的なディスクデバイスとしてアクセスできるようにすることである。ユーザは、インターネット上に仮想的に存在するディスクに、ドキュメントファイルやメールなどの個人データを保存し、取り出すことができる。仮想ディスクは、ブロックデバイスのインタフェースを提供することにより、通常のディスクデバイスと同じ使い勝手に利用できる。

対象とするプラットフォームはx86版Linuxカーネル2.6.17とする。

2 関連研究

インターネット経由でブート可能なLinuxディストリビューションであるHTTP-FUSE-KNOPPIX [1]では、システムファイル取得のため仮想的なブロックデバイスをHTTP-FUSE-loopドライバを用いて実現する。ただし、この仕組みはデータの読みみに特化しているため、ユーザデータの保存には応用できない。分散ストレージシステムPetal [2]は、冗長性確保、動的な再構成やバックアップなどをサポートする高機能なシステムであるが、インターネット経由での利用が想定されていない。

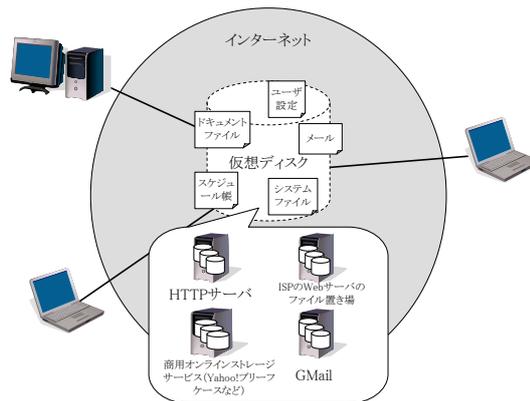


図1: 概念図

3 設計

本開発では、ファイルシステムのドライバではなく、ブロックデバイスのドライバを作成する。つまり、カーネル上では通常の物理的なディスクデバイスと同列の扱いとなり、ext3のようなファイルシステムや、ソフトウェアRAIDといった、ストレージデバイスのための既存の枠組みをそのまま適用することができ、応用が利く。

クライアントマシン上で動作させるドライバの構成を図2に示す。ドライバは、カーネルモード部分とユーザモード部分の2つの部分からなる。実際のネットワーク入出力処理などをユーザモードで行うことで、プログラミングが容易となる・各種プロトコルごとにモジュール化できる・拡張が容易となるという利点がある。

カーネルモード・ユーザモードの各部分について述べる。

カーネルモジュール `vdisk.ko` カーネルモジュール `vdisk.ko` は、カーネル側に向けたインタフェースとしてブロックデバイス `/dev/vdiskn` (n はマイナー番号) ユーザモード側に向けたインタフェースとしてキャラクタデバイス `/dev/vdiskio` を持つ。カーネルモジュールの役割は、ブロックデバイスを通してカーネルから受け取ったブロックI/Oリクエストを、独自定義のI/Oパケットに変換し、

Development of Distributed Virtual Disk System by Online Storage for Ubiquitous Computing Environment

Yusuke NOJIRI[†], Mitaro NAMIKI[‡]

[†] Faculty of Engineering, Tokyo University of Agriculture and Technology

[‡] Division of Systems and Information Technology, Institute of Symbiotic Science and Technology, Tokyo University of Agriculture and Technology

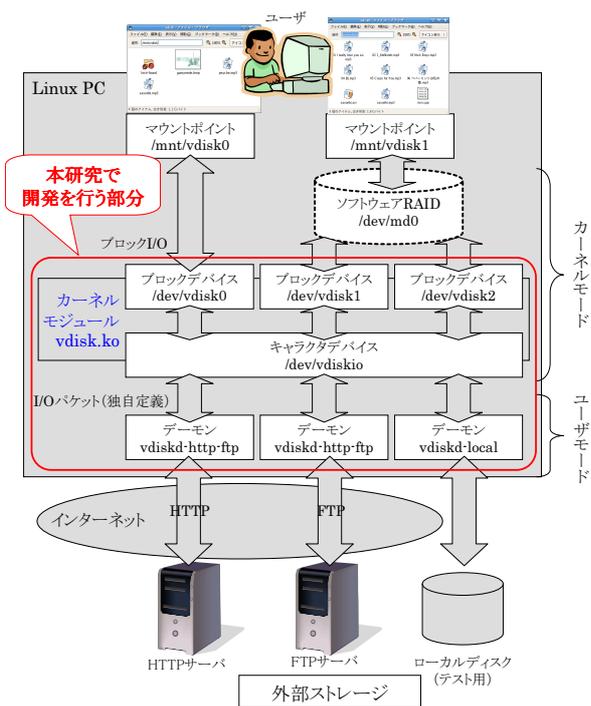


図 2: 仮想ディスクシステム的设计

キャラクタデバイスを通してユーザモード側に転送することである。

デーモン `vdiskd-http-ftp`, `vdiskd-local` ユーザモードで実行するデーモンプログラムは、キャラクタデバイス `/dev/vdiskio` を通じて受け取った I/O パケットを解釈し、実際のネットワーク入出力処理を行う。デーモンは、各種外部ストレージごとに特化したプログラムとする。デーモン `vdiskd-http-ftp` は、HTTP または FTP によりブロックファイルの入出力を行う。デーモン `vdiskd-local` は、ローカルディスクに対してブロックファイルの入出力を行うテスト用デーモンである。

4 評価

本システムの性能を評価するため、5 通りの条件で、スループットを計測するベンチマークテストを行った。クライアントマシンの CPU は Intel Pentium 4 2.80GHz、メモリは 512MB、OS は KNOPPIX 5.0.1 である。クライアント・サーバ間のネットワークは 1000BASE-T である。

ベンチマークの結果を図 3 に示す。ローカルディスクに対する入出力で、本ドライバを介する場合と介さない場合は、スループットが読み込み時で約 36%、書き出し時で約 59% 低下している。この値は本ドライバのオーバーヘッドであるといえる。RAID-0 (ストライピング) のディスクアレイ化した場合では、むしろソフト

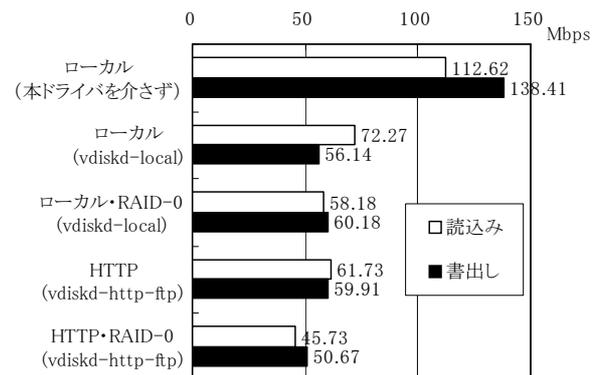


図 3: 仮想ディスクデバイスのスループット

ウェア RAID のオーバーヘッドが災いしたためか、多くの場合でスループットが低下している。ネットワークを介した場合のローカル環境に対するスループットの低下は小さく、最悪でも約 21% (「ローカル・RAID-0」/「HTTP・RAID-0」の読み込み時)にとどまっております、妥当な結果である。

5 おわりに

ディスクインタフェースをブロックデバイスドライバとしたことで、高い透過性を提供できた。これを応用し、ホームディレクトリや OS そのものを本仮想ディスクシステムに載せることも可能である。

ベンチマークテストの結果から、開発したドライバのオーバーヘッドは実用に堪える程度であるといえる。ネットワークを介した場合の仮想ディスクのスループットは、一般家庭に提供されている光ファイバ程度であることを考えると、満足のいく値である。インターネットを介した場合、レイテンシの影響によりいいパフォーマンス得られないことも考えられるが、先読みの実装やブロックサイズの調整である程度改善できると考えている。

セキュリティに関しては、プロトコルの認証機構や暗号化に頼っている。本システムが独自にデータを暗号化する仕組みを持つなどして、さらにセキュリティを高めることは今後の課題である。

参考文献

- [1] 産業技術総合研究所: HTTP-FUSE-KNOPPIX. <http://unit.aist.go.jp/itri/knoppix/http-fuse/>
- [2] Edward K. Lee and Chandramohan A. Thekkath. Petal: Distributed Virtual Disks. In *Proceedings of the Seventh International Conference on Architectural Support for Programming Languages and Operating Systems*, pp. 84-92, 1996.