

## プライベートクラウドによる オンラインストレージサービスの統合

高橋 辰典<sup>†</sup>, 田中 秀直<sup>†</sup>, 木部 真一郎<sup>‡</sup>, 上原 稔<sup>†</sup>

クラウド時代では、データはクラウドに集中する。実際、クラウドを利用した数多くのオンラインストレージサービスが提供されている。しかし、クラウドとデバイスをつなぐネットワークには十分な帯域がないため、遅延が大きい。そこで、クラウド側で処理する仕組みが必要となる。我々はオンラインストレージサービス側で処理する方式をストレージサイド計算と名付け、これを推進している。本論文では、ストレージサイド計算の事例として電子ブック管理システムについて述べる。近年、電子ブックは急速に普及しつつある。電子ブック管理システムでは、個人が自身のPDF形式電子ブックをオンラインストレージサービスに保管している状況を前提とし、そのような電子ブックを便利に管理するシステムを提唱する。本システムは PC だけでなくスマートフォンのようなデバイスからもアクセスされ、移動中でもクラウド側で処理を進めることができる。

### Private Cloud based Integration of Online Storage Services

Tatsunori Takahashi<sup>†</sup>, Hidenao Tanaka<sup>†</sup>, Shinichiro Kibe<sup>†</sup>,  
Minoru Uehara<sup>†</sup>

In cloud era, data is centralized into cloud. Indeed, many online storage services are provided. However, the communication delay is long because the network connected devices to cloud does not have enough bandwidth. So, a mechanism of processing cloud-side is required. We call such a mechanism that processes data near online storage "storage-side computing." In this paper, we describe an electrical book management system as an example of storage-side computing. Recently, electrical book becomes popular. In our electrical book management system, we assume that a user stores his own personal electrical book formatted as PDF in several different online storage services. In such a situation, a user wants to manage them in an integrated way. Our system gives users such a service. Our system is allowed to be accessed from various kinds of devices such as Smart Phone and can process data during even off-line.

<sup>†</sup>東洋大学 工学部 情報工学科

Dept. of Information and Computer Sciences, Toyo University.

<sup>‡</sup>東洋大学大学院 工学研究科 情報システム専攻

Dept. of Open Information Systems, Toyo University.

### 1. はじめに

ストレージ技術の進歩とクラウドの普及により、オンラインストレージに対する要求はますます高まっている。クラウド時代では、データはクラウドに集中する。今日では、モバイル技術の進歩により一定水準のユビキタス環境が実現されている。データは分散しているより集中している方が容易に管理できる。よって、必然的にクラウドに集中する。しかし、今日のモバイル技術はユーザの要求を完全に満たすほどではない。そのため、クラウドとデバイス間に大きな遅延が存在する。この遅延は使い勝手を悪くする主因となる。そこで、遅延大きなネットワークを介さず、クラウド内で処理する方式が有効となる。我々は、このような方式をクラウド上のストレージサービスで実現されることからストレージサイド計算と名付ける。ストレージサイド計算は必ずしも物理的なストレージ層での処理を意味しない。むしろ、ストレージサービスを提供するクラウドに汎用な処理を包括する。デバイスにダウンロードすることなくクラウド上で処理することが重要である。

今日では、多くのオンラインストレージサービスが提供されている。しかし、それらは個別のサービスとして提供され、相互運用性は考慮されていない。また、それぞれに長所と短所があり、いずれのサービスも完全無欠ではない。むしろ、それぞれの特徴を生かしてうまく組み合わせることが重要である。

近年、オンラインストレージサービスを使用するデバイスも多様化している。PCだけでなく iPhone/iPad, Android Phone/Tablet などからもオンラインストレージサービスが利用されている。しかし、すべてのオンラインストレージサービスがこれらのデバイスから利用できるわけではない。今後は、PC 以外のデバイスでも便利にオンラインストレージサービスを利用できるようにする必要がある。

しかし、これらのデバイスでは PC に比べてストレージの容量が制限される。しかも、多くのオンラインストレージサービスはアップロード/ダウンロード方式を採用しており、ファイルを開くにはいったんダウンロードする必要がある。ファイルサイズが大きくなるとダウンロード時間が長くなるばかりか、3G 回線等を利用するスマートフォンでは通信料が発生する可能性もある。さらに、移動中に通信が切断されるとダウンロードに失敗することもある。このような場合、ストレージサイド計算が有効である。ストレージサイド計算では、通信が切断中でもクラウド側で処理が進行する。デバイスは最終結果だけに必要最小限のアクセスを行えばよい。

ストレージサービスの発展には iPod に始まる音楽データや iPad に始まる電子ブックも重要な背景の一つである。聴覚情報より視覚情報の方が重要である。電子ブックは、今後、教育や文化に多大な影響を与える重要な変革である。本研究では、電子ブックのライブラリ管理を題材に、ストレージサイド計算の有用性を示す事例を紹介する。そのために、複数の異なるオンラインストレージサービスを統合し、その統合さ

れたサービスにおいてストレージサイド計算を実現する。

本文の構成は以下の通りである。2 節で関連研究としてオンラインストレージサービスについて述べる。3 節では VLSD について述べる。4 節では、MeshRAID MP の概念と VLSD による実装法について述べる。5 節では、その評価を行う。最後に結論を述べる。

## 2. 関連研究

オンラインストレージサービスとは、インターネット上でファイル保管用のディスクスペースを貸し出すサービス。有料のものと無料のものがあり、無料の場合は使用時に広告が表示される。職場と自宅、あるいはノートパソコンなどとの間でのデータのやりとりや、複数人でのデータの共有などが行える。利用できる容量は数 MB～100MB 程度のサービスが多い。Web ブラウザや FTP クライアントから利用できるものが多いが、専用のソフトをパソコンに組み込んで、Windows のエクスプローラから他のディスクと同じような操作感で利用できるようにしたものもある。

また、オンラインストレージサービスは、リムーバブルメディアの代替、バックアップ、ファイル転送、共有ディスクの 4 種類に分類される。自宅で作成したファイルをこのオンラインストレージにアップロードし、リムーバブルメディアの代替は、会社でダウンロードして作業を再開するなどの利用方法である。有料のサーバーはメンテナンスが行き届いていることが期待でき、バックアップ先としての信頼性が高く、オンラインのバックアップに特化したサービスも多く提供されている。ファイル転送は、電子メール受配信時の容量制限を超えるファイルの受け渡しのために使う。企業コンプライアンスやセキュリティなどの問題から、無償サービスの利用を禁止する法人も多い。共有ディスク - グループで作業するための共有ディスクスペースとして使う。本論文で扱う Dropbox は共有ディスク以外の 3 つ全ての特性を持っている。

次に主要なオンラインストレージサービスの事例を紹介する

Amazon S3(Simple Storage Service)は多くのオンラインストレージサービスの基本となっている。Dropbox は米 Dropbox, Inc.が運営するオンラインストレージサービスである。無料で 2GB のディスクスペースを利用でき、専用クライアントソフトを利用した高機能なファイル同期・バックアップシステムが特徴である。SugarSync は無料で 5GB まで利用が可能である。SugarSync の最大の特徴は複数 PC の任意フォルダ同士を直接同期が可能であるという点あげられる。また、SugarSync は複数の PC を同期させた際にそのフォルダがどの PC のものなのかがわかりやすいように設計されている。OS が Linux に対応していないという欠点がある。 iCloud は Apple が iOS5 から提供するクラウドサービスである。5 GB が無料で利用可能で音楽、カメラロール、フォトストリーム、書類、アプリ、本、バックアップ、連絡先、カレンダー、メールなどを

ひとつの Apple ID で同期することが可能である。

### SkyDrive

SkyDrive は、マイクロソフトが提供する Windows Live のサービスの 1 つである無料オンラインストレージである。SkyDrive は 1 アカウントあたり、25 ギガバイトを利用できる。しかし、100 メガバイト以上のファイルは利用できない。また、あらゆる種類のファイルをアップロードできるという特徴もある。Yahoo Box は、パソコン用アプリではドラッグ&ドロップによる操作で複数のファイルを一括でアップロードすることができる。ファイル形式はどんなものでも保存できる。そして、アップロード途中で止まってしまっても続きから再開して行うことができる。Google Documents はドキュメントやコンテンツはいつでも共有することができる。そして、すべての更新履歴データにアクセスすることができるので保存のミスがなくなる。また、複数のユーザがチャットで連携しながら同時に 1 つのドキュメントの編集作業を行うことができる。Gmail Drive とは、ローカルドライブのように扱うことができるシステムである。ファイルを添付ファイルとして送信している形である。それを WEB 上から Gmail にアクセスしてファイルを保存することも可能である。Pogoplug は、パーソナルクラウド環境を構築することができるようにしたファイル共有デバイスで、Dropbox などのクラウドサービスとは異なり、自分のパソコンに全データを残したままで自分専用のクラウドを実現できて、Pogoplug 経由で外部からもデータにアクセスができる。また、容量は自分で用意したハードディスクなどにデータを保存しているのほぼ制限がない。セキュリティに関しては、自分次第である。Ajaxplorer とはブラウザ上から利用出来るファイルマネージャーであり、ユーザごとのファイル管理、ファイルの共有、保存ファイルの閲覧、編集などが可能である。Ajaxplorer はサーバとして利用でき LAN 内だけでなく外部からも閲覧可能である。

赤間ら<sup>1)</sup>はフォルダ・プログラミングを提唱した。フォルダ・プログラミングでは、フォルダを処理に関連付け、ファイルをフォルダに移動するだけで処理を適用可能とした。この方式では汎用性の高い操作が特別なプログラミングの知識なしに実行できる。フォルダ・プログラミングをネットワーク・フォルダに拡張したネットワーク・フォルダ・プログラミングはストレージサイド計算の一種といえる。

メディア処理向けクラウド基盤「虹雲」は、前述のフォルダ・プログラミングを発展させたシステム POLDER と追記・参照型データ管理システム DMS を組み合わせたものである。ここでは、DMS は分散データストリーム管理システム (DSMS) と DBMS を統合した情報統合基盤である。ビッグデータにおける CEP(Complex Event Processing)として考えることもできる。



#### 4. 電子ブック管理システム

本論文では、ストレージサイド計算の事例として電子図書管理システムについて述べる。近年、タブレット PC やスマートフォンの進歩と普及によって、書籍の利用形態がアナログからデジタルへと移行してきている。デジタルでの書籍は一般に電子書籍と呼ばれ、その購入方法、利用媒体、データ形式などは様々であり複雑化している。

また、電子書籍の普及に伴い個人レベルで電子書籍を作成し利用しようとする、電子書籍の自炊と呼ばれる行為も多く見受けられるようになってきた。既存のサービスでは電子媒体で提供されていない書籍を電子書籍で読みたい、膨大な書籍の管理をデータ上で行いたいというニーズの高まりがこうした自炊という方法を生み出した。自炊した書籍の管理には、出先での閲覧を考慮し、オンラインストレージサービスに保存することが最適であるといえる。また、オンラインストレージサービスの高い安全性と信頼性により、ローカルで管理するよりもデータを損失してしまう可能性が少なく、定期的なバックアップの必要もない。しかし、こうしたオンラインストレージサービスはローカルで利用することを前提とした現在の電子ブックリーダーとはオンラインストレージ間の連携が不十分で利便性が低い。

そこで本論文では、複数の異なるオンラインストレージサービスを統合し、その統合されたサービスにおいてストレージサイド計算を実現するために、次のような要件を定義するものとする。

まず、電子ブックは PDF 形式とする。これは電子書籍のデータ形式として最も広く普及したデータ形式であり、多くのデバイスで利用することが可能なためである。また、電子書籍の貸出を考慮した際に貸出期間を制限するスクリプトをデータ内に埋め込むことが容易であることも理由の一つである。

デバイスからのアクセスにはインターネットブラウザを用いる。使用者は AjaXplorer にアクセスし一つのインターフェースから複数のオンラインストレージサービスにアクセスし電子書籍を管理することができる。AjaXplorer 上から手持ちのデバイスのローカルに保存するにはサーバ上で自動的に貸出期間の処理が行われ、期限を過ぎると閲覧できなくなる。

FESS で検索させたい電子書籍ファイルがあるフォルダ内をクロール設定によって指定して、電子書籍ファイルの情報を取得した後検索すると欲しい電子書籍の情報を検索することができて、それにアクセスすることでファイルを保存することもできる。

次に、システム概要を説明する。本論文でサポートするオンラインストレージサービスは Dropbox, Pogoplug の 2 種類とする。Pogoplug とは Cloud Engines が提供する、本体に接続した USB ストレージを、LAN の中だけでなく外からもアクセスできるようにするデバイスである。サービスの形態としては NAS に近いものがあるが、Pogoplug は NAS と違い外からアクセスする際に独自のドメインを必要とせず、Cloud

Engines が持つ中継サーバを介してブラウザ上からアクセスできるためより手軽にプライベートサーバを利用することができるものである。

AjaXplorer によるストレージサイド計算の定義は、オンラインストレージサービス間でのコピー、移動をバックグラウンドで、容易に行えることとする。使用者は電子書籍のコピー、移動を複雑な手順を踏まなくとも、一つのインターフェイス上で行うことができ、すべてのファイルの操作はサーバによりバックグラウンドで処理される。

また、書籍データである PDF ファイルは使用者からの貸出要求を受けると、ダウンロード前に Javascript による貸出期限スクリプトを埋め込まれ、自動的に閲覧できる期限が設定される。PDF は前述したように Javascript を埋め込むことができ、デバイスから日付を取得し PDF ファイルを標示させないようにすることが可能である。Javascript の埋め込みには Adobe Acrobat や専用の有効期限設定用のアプリを用いる。

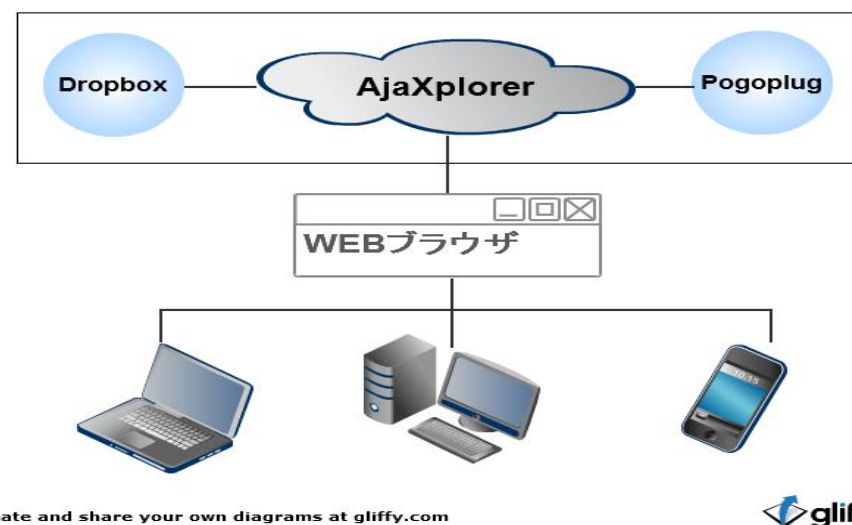


図 2. システム概要  
Fig.2 System Overview

次に電子ブックの検索のしくみについて述べる。今回の電子書籍を検索するのに用いる検索エンジンは FESS である。FESS とは、全文検索サーバである。Java の実行環境があればどの OS でも実行することができる。Seasar2 をベースとして構築されてい

て、検索エンジンの部分に Solr を利用している。ドキュメントのクロールには S2Robot を利用することで、Web ページやファイルシステムに対するクロールができるようになり、MSOffice 系のドキュメントや PDF ファイル、zip などの圧縮ファイル、データベースも検索対象とすることができる。

次に検索する方法を説明する。AjaXplorer で立ち上げたサーバにアクセスするように、Fess で設定を行う。フォルダの所有者権限とグループが違っていると、クロールをさせることができないので、フォルダの所有者権限とグループを Apache のユーザに合わせ、ファイルの置いてあるフォルダをクロールさせるように設定する。フォルダ内をクロールさせ、データ取得後、検索をすることで、保存してあるファイルを検索することができる。

## 5. 評価

ここでは本論文のプライベートクラウドによるオンラインストレージサービスの統合の評価を行う。

まず、オンラインストレージサービスが統合されたことを示す。統合のベースとなる AjaXplorer はサーバ上のファイルパスを指定してリポジトリを追加することが可能である。このリポジトリの設定はログインしたユーザごとに保存され、AjaXplorer のユーザ管理のみで、LAN 上に存在する複数のクラウドサービスを管理することが出来るようになる。1 つ目に、Pogoplug との統合について述べる。Pogoplug は Linux 版では公式サイトで配布されている FUSE を利用した pogoplugfs というファイルにより mount させることが出来る。これにより、マウントした場所のパスを指定することで AjaXplorer からファイルの閲覧が可能となる。2 つ目に、Dropbox との統合について述べる。Dropbox は GUI 上であればインストーラーを起動して、ローカルエリアに同期用のフォルダを作成することが可能だが、CUI ではセキュリティ上の理由から全ての導入作業をコマンドだけで完結することは出来ない。そのため、ユーザはコマンドラインに標示された URL に Dropbox へログイン済みの Web ブラウザからアクセスすることで機器の登録し、導入を完了すること出来る。また、Dropbox は一つのアプリケーションに一つのアカウントしか設定することができないので、この問題は Dropbox のプロセスを複数起動させることで、マルチユーザに対応した。

次に、AjaXplorer で Dropbox, Pogoplug の PDF を閲覧、編集できることを示す。AjaXplorer 自体には PDF を閲覧するための機能はないため、使用しているデバイスのブラウザが PDF に対応していれば問題なく閲覧することは可能である。さらに、統合した Dropbox と Pogoplug も設定したリポジトリに移動することで、同様に閲覧が可能である。次に編集については、PDF ファイルの移動、コピー、削除、リネーム、権限の変更が可能である。移動、コピーはリポジトリを跨いでの編集も可能なので Dropbox、

Pogoplug 間でも同様の操作が可能である。

さらに、ストレージサイド計算が可能であることを示す。このシステムはサーバにアクセスしファイルの編集を行うため、デバイスを切断してもサーバ側で処理が実行され続ける必要がある。そのための検証に次のような操作を行った。処理の経過が監視できるように、70MB ほどの比較的大きなファイルを各リポジトリ間を移動させる命令を外部のデバイスから行った後、デバイスのブラウザを終了しサーバ側でファイル移動の経過を監視した。結果、ファイルは徐々にファイルサイズを増加させ移動され続け、命令を完了することが確認できた。

また、このシステムの運用を考えた際には電子書籍の閲覧だけでなく貸し出しを考慮した PDF ファイルの有効期限設定も必要である。PDF に期限を設ける方法についてはシステム概要で説明したとおりである。期限を設定された PDF ファイルは埋めこまれたスクリプトにより日付を取得し、指定した期限と照らし合わせ閲覧の可否を判定する。期限切れの PDF ファイルは指定した領域に、白紙などのレイヤを上から標示させられ閲覧できなくなる。しかし、PDF にスクリプトを組み込めるアプリケーションは限られており、現状では有効期限の設定は別アプリケーションから行うことしかできない。また、期限設定用のアプリケーションは AjaXplorer から呼び出しが出来ず、ユーザの貸し出し要求に対する自動的な有効期限設定は行えないことをここで断っておく。



図 3. 検索結果  
Fig.3 Search results

FESS でクロールの設定をして、クロールさせ、ファイルの情報を登録させた後、検索する。図3に実行結果を示す。検索すると、情報を取得したファイルを検索することができる。また、どこにそのファイルがあるのかも表示されている。現状では、クロール設定を行い、電子書籍ファイルの情報を取得して検索できることしかできていない。アクセスをして、電子書籍ファイルの保存を行うことができていないことをここでは断わっておく。

## 6. まとめ

本論文では、電子ブックのライブラリ管理を題材に、ストレージサイド計算の有用性を示す事例を紹介することを目的に、複数の異なるオンラインストレージサービスを統合し、その統合されたサービスにおいてストレージサイド計算を実現することを目標として、設計と実装及び評価を行った。

その結果、まず AjaXplrer に Dropbox と Pogoplug を統合し、Dropbox を複数のプロセスで動かしマルチユーザに対応させた。さらに、ブラウザ上から AjaXplorer を介してデバイスを選ばず、クラウド間の違いを意識することなく電子書籍データである PDF を閲覧、編集できるように改良した。また、PDF ファイルの編集作業である、移動やコピーはデバイスでブラウザを終了してもサーバー側のバックグラウンドで実行され続けることを検証した。本システムにおけるブラウザでの電子書籍の閲覧という機能の延長線上にある、電子書籍の貸出についても調査を行った。電子書籍の貸出しに必要な PDF ファイルの有効期限の設定は別アプリケーションを用いることで行えることが分かり、実際に有効期限を設定し期限が切れると閲覧が不可なることを確認した。

次に FESS で設定を行い、クロールさせ情報を取得した後、検索してみると取得した情報のファイルが検索できている。また、そのファイルが保存されている場所も表示されていてどこに保存されているかはわかる。

今後の予定は、AjaXplorer のユーザ登録に際した、Dropbox アカウントの同時設定を行う機能と、AjaXplorer の貸出操作に対応した自動的な有効期限の設定機能の開発と実装が挙げられる。現状、AjaXplorer と Dropbox のユーザ登録は別々に行う必要があり、さらに管理者によるディレクトリ情報の設定など、すべて手動で行わなければならない。そのため、これらを一括で行えるようなシステムの開発を行う。また、電子書籍の有効期限の設定も AjaXplorer とは別でアプリを使用する必要があり、そのための操作を自動で行うことができないので、この点についても考察を行う。

検索については、検索したあと、そのファイルにアクセスしようとする则表示すること、また、ファイルを保存することができなかつたので、今後この問題を解決することにより、扱いやすくすることを目指す。

## 参考文献

- 1) 赤間 浩樹, 内藤 一兵衛, 内山 寛之, 山中 真和, 谷口 展郎, 長谷川 知洋, 三井 一能, 山室 雅司: "フォルダ・プログラミングとネットワーク・フォルダ・サービス", DICOMO2008, pp. 1435-1442, (2008)
- 2) Shinichiro Kibe, Minoru Uehara: "Proposal for a Cloud-based Educational Environment", In Proc. of 3rd International Workshop on Information Technology for Innovative Services(ITIS2011) in conjunction with the 14th International Conference on Network-Based Information Systems(NBiS2011), pp.523-528, (2011.9.7-9,Tirana,Albania)
- 3) Shinichiro Kibe, Minoru Uehara, Motoi Yamagiwa: "Evaluation of Bottlenecks in an Educational Cloud Environment", In Proc. of the 13th International Symposium on Multimedia Network Systems and Applications (MNSA2011) in conjunction with the 3rd IEEE International Conference on Intelligent Networking and Collaborative Systems (INCoS2011), pp.520-525, (2011.11.30-12.2, Fukuoka, Japan)
- 4) Shinichiro Kibe, Motoi Yamagiwa, Minoru Uehara: "Proposal for Improving Throughput in Supersaturated Cloud", ITIS-2012-03, (TBA)
- 5) 木部 真一郎, 上原 稔: "クラウド環境に基づく情報教育環境の構築", 信学技報 RIS No.3, Vol.3, pp.1-6, (2011.4.16)
- 6) 部 真一郎, 上原 稔: "教育環境のための過飽和クラウドにおけるボトルネック評価", DPSWS2011 マルチメディア通信と分散処理ワークショップ論文集(DPSWS2011), pp.188-193, (2011.10.5-7,十和田)
- 7) 木部 真一郎, 上原 稔: "プライベートクラウドによる実習環境の構築", 第73回情報処理学会全国大会, 5X-6, pp.409-410, (2011.3.2-4, 東工大)