

P2P によるセキュア情報流通プラットフォーム “P2PWeb”

神谷俊之 加藤大志 仁野裕一 福岡秀幸 谷幹也

NEC インターネットシステム研究所

1.はじめに

近年、ブロードバンド環境の普及に伴い、P2P によるアプリケーションが注目されてきている。例えば、Winny, WinMX[1]などのファイル交換や、BitTorrent[2]といったコンテンツ配布、IP 電話など P2P の特徴を活かしたアプリケーションが実用的に使われ始めている。

我々は、P2P を用いて、情報発信の敷居を下げることによって、現在の Web 以上に多くの人が情報発信を行う大規模な情報共有・流通の場ができるのではないかと、いう仮説のもと、P2P ベースによる情報流通プラットフォーム “P2PWeb” を提案する。P2PWeb では、個人 (Person) が端末 (Peer) を介して “P2P” で情報発信を行うことで現在の “Web” 的な情報発信・共有の場をより簡単に実現することを目標としている。そのために P2PWeb では、P2P ファイル共有をベースとし、その上で各種の P2P ベースの情報発信アプリケーションを自由に開発できる API を提供する。

本稿ではまず P2PWeb のアーキテクチャについて述べた後、大規模な P2P ネットワークを実現するためのプロトコル GISP について述べ、アプリケーション例として現在試作を行っている P2PWeb ブラウザについて説明する。

2.個人からの情報発信と P2P プラットフォーム

Web の普及により、Web 以前の時代に比べて、個人からの情報発信は飛躍的に容易になったが、ホームページ (HP) 作成では内容よりもまずデザインするのに手間がかかる、整理した形での情報の提示が暗黙のうちに求められるなど、心理的な障壁もまだ存在する。2ちゃんねるや Blog は、どのような発言をしてもよいというような “場の空気” や思っていることをそのまま書けばデザインはシステムまかせというような枠組みで情報発信のための閾値を下げることで、従来の Web で扱われなかったような生の情報の発信に成功している。2ちゃんねるや Blog で発信される情報は粒度が小さく、信頼性も低いが、従来の Web で得られない膨大な情報を含んでいる。我々は P2P を基盤とする情報発信プラットフォームを提供することで情報発信のための閾値を下げ、さらに多くの情報をネット上に生み出すことができると考える。

P2P 型情報共有プラットフォームによる情報発信のメリットとしては、以下のような点が考えられる。

- ・大容量コンテンツの発信

今後、携帯電話などを利用したデジカメ画像やムービーなど大容量のコンテンツが増えることが予想されるが、ISP 側で管理されるホームページに対して、個人管理の PC では容量を増加することが容易である。また、整理されていない生の情報の場合、個々の情報に対してのアクセス頻度は高くないと考えられるが、その情報をサーバにアップロードする手間を省くことができる。

“P2PWeb” a P2P based secure information distribution platform

Toshiyuki KAMIYA, Dasihi KATO, Yuichi NINO, Hideyuki FUKUOKA and Mikiya TANI, Internet Systems Research Laboratories, NEC Corporation

- ・作成者自身による情報の所有・管理

サーバの故障等による停止や、事業者によるサービス変更、停止などによる影響を受けにくく、情報の発信者が情報の可用性を制御できる。また掲示板などのサービスについても、Web の管理者のような他者の管理を受けないアドホックなサービス・情報共有を実現可能である。

- ・大規模な情報共有プラットフォームの可能性

ネットワークへの参加者 (peer) の増加は、ネットワークへの情報の増加とネットワーク維持のための各種リソースの増加を同時にもたらすことから、単一サーバへ負荷集中によるボトルネックが発生しない。

P2PWeb では、従来の P2P 型システムよりも粒度が小さく、膨大となる情報を保持できるプロトコルと、その上で、情報共有のタイプによって、ファイル共有や掲示板、電子メール、IM、チャットなどの各種アプリケーションを構築可能な柔軟性を備える API を提供することを目標としてアーキテクチャを設計した。

3.P2PWeb のアーキテクチャ

P2PWeb のアーキテクチャは、基本的にはファイル共有をベースとし、各 Peer では OS の提供するディレクトリツリーをそのまま利用する。P2PWeb ではアプリケーション開発のために、仮想共有ディレクトリ上に通常の OS のファイルシステムと同様の API を提供する予定である。P2PWeb のアーキテクチャ階層図を図 1 に示す。

プロトコル: P2PWeb のネットワークは、各 peer がお互いに通信しあうことによって維持される。Peer 間のファイル転送には HTTP を使い、インデックスの検索には GISP (Global Information Sharing Protocol) と呼ぶ独自プロトコルを用いている。プロトコルの特徴は次節で述べる。

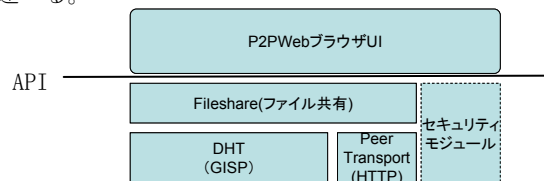


図 1 P2PWeb のアーキテクチャ

ファイル共有: FileShare モジュールでは、各 Peer が持つディレクトリツリーをマージして仮想的に単一のディレクトリ構造として利用者・アプリケーションに提供する。例えば、利用者 A が/abc/def.txt というファイルを公開していて、利用者 B が/abc/ghi.txt というファイルを公開しているとすると、両方の利用者からはこれらを仮想的に重ねあわせて、/abc の下に def.txt と ghi.txt が両方存在しているように見える。

仮想ディレクトリ中のファイルの閲覧時に、ローカルに実体がない場合には GISP で実体の所在を発見し、HTTP でローカルにコピーする処理を行う。これにより、利用者はファイルを参照する際に、その実態がどこにあるのかを気にせず閲覧することが可能である。また、一旦、コピーが行われた後は、そのファイルがオリジナルか複

製かは区別されない。ファイルのダウンロードは、他の peer がファイルの閲覧を要求した場合のみ行われ、要求した Peer に複製が作成される。

セキュリティ：仮想ディレクトリにおけるアクセス権限管理やファイルのトレース機能を提供するセキュリティモジュールがプラグインとして組込可能である。

4. GISP

GISP は P2PWeb においてインデックスを管理するプロトコルである。GISP は DHT (Distributed Hash Table, 分散ハッシュテーブル) 技術をベースとしており、ID (key) をオブジェクト (value) と結び付けて登録・取得する機能が、サーバレスで効率よく構成できることを特徴とする。DHT では、各 Peer で ID(key) とオブジェクト (value) の関係を分散して保持する。また、どの peer がその関係を保持しているかという情報自体も接続情報という形で分散管理されている。典型的な DHT の実装である Chord[3]、Kademlia[4] などでは、ルーティングテーブルの大きさと、peer を特定する処理にかかる時間がネットワークの大きさの log のオーダーで処理できることが知られており、ネットワークが非常に大きなものとなった場合にもメモリ使用量、遅延を抑えることができ、大規模な分散情報管理を実現することができる。

GISP は、Chord や Kademlia と同様の Hypercube タイプの DHT の一実装であり、大規模な分散ファイル共有を実現するために以下のような特徴を持っている。

- 複数のオブジェクトを同じ key に対し put することを許している。これによりディレクトリ構造を素直に表現できる。
- 各 peer が管理する他の peer node 数が固定されず、多数の peer 情報を持つことができる。大きなネットワークに対して高速なアクセスが可能。
- パケットレベル・ネットワークプロトコルレベルでのネットワーク利用の効率性について配慮して、UDP の上の best effort 型プロトコルとして実装されており、多数の peer からの同時接続が可能。

5. P2PWeb ブラウザの試作

P2PWeb のアプリケーションとして、エクスプローラ風 UI を持つ P2PWeb ブラウザを開発している (図 3)。以下、P2PWeb ブラウザの各機能について説明する。

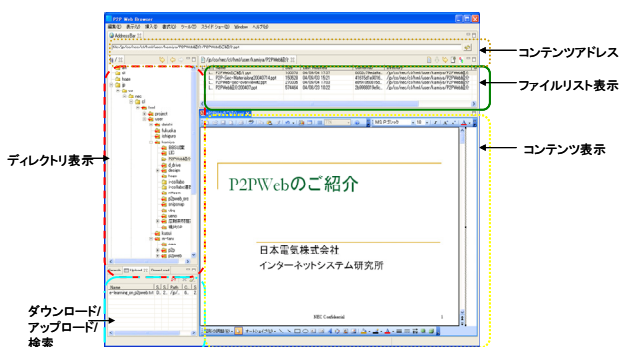


図 2 P2PWeb ブラウザ

• 仮想ディレクトリ共有

各 Peer がローカルに持っているディレクトリ構造を重畳した形で仮想的に単一のディレクトリ構造を表示。

• リンクと Reputation

リンクとはファイルとファイルとを結びつける機能である。例えば、P2PWeb 内で関係した内容のファイルを見つ

けたときに、利用者が自由にリンクを張ることができる。Reputation はファイルに対して評価値と自由文をつける機能である。利用者はファイルを取得する前に評価値の集計を知ることができ、ファイルを取得するかどうかの判断ができる。

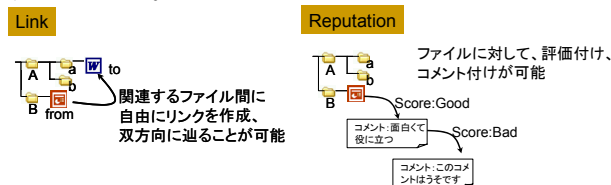


図 3 リンクと Reputation

• 多重ダウンロード

P2PWeb では、一つのコンテンツのコピーが複数の peer に存在している場合がある。このことを利用して効率よくファイルの取得を実現する複数の peer からの分割多重ダウンロードの機能を実装している。

• BBS 機能

ファイル共有をベースとしたアプリケーションの一例として XML ファイルベースでの BBS 機能を実装している。本機能では BBS はスレッド単位で XML ファイルとして管理されており、スレッドへの新たな発言は元ファイルに発言を追記した新しいファイルとして書き込まれる。スレッドを読む際には同じディレクトリ内にある全ての同一スレッドのファイルをマージする処理を行う。

• ディレクトリアクセス制御

ファイル共有で一般に行われるようなディレクトリごとにグループの情報を共有する用途の場合、ディレクトリ単位でアクセスの制御が行えることが望ましい。これを実現するために、グループ毎の暗号鍵によって DHT のデータ、およびファイル自身を暗号化することでディレクトリアクセス制御を実現している。

• トレーサビリティ

P2PWeb 上でどれだけの人がダウンロードを行ったかを把握するため著作権管理技術 (DRM) とダウンロード時の暗号化を組み合わせ、各 peer 間でのコンテンツのダウンロードをトレースし、専用のトレーサビリティサーバ上で確認できる仕組みを実現している。本方式では P2PWeb 外に流通したコンテンツについては再生を防止する特徴がある。[5]

6. おわりに

本稿では、P2PWeb のコンセプトを提案し、DHT をベースにした P2P ファイル共有プロトコル GISP、P2PWeb ブラウザの試作について説明した。現在 GISP のスケーラビリティの検証を進めており、今後、セキュリティ強化、アプリケーション開発のための API 検討など実用化に向けて開発を行っていく予定である。

参考文献

[1] WinMX, <http://www.winmx.com/>
 [2] BitTorrent, <http://www.bittorrent.com/>
 [3] I. Stoica et al., “Chord: A scalable peer-to-peer lookup service for internet applications”, In Proc. of ACM SIGCOMM, 2001
 [4] P. Maymounkov et al., “Kademlia: A Peer-to-Peer Information System Based on the XOR metric”, In Proc of IPTPS02, 2002
 [5] 仁野他, “P2P ネットワークにおける情報の流通監視方式の提案”, 第 67 回情処全大, 2004