

Gnutella 型 P2P プロトコルの可能性

The potential of the gnutella-like P2P protocols

上川 純一, 谷村 勇輔

Junichi Uekawa, Yusuke Tanimura

Abstract: An application known as gnutella is introduced briefly. Gnutella is an implementation of an idea known as P2P, which is spreading slowly, but surely.

1 はじめに

P2P とは, サーバとクライアントというような不均等な関係ではなく, 互いに対等な立場に立って通信することを基調とした一連のアイデアである. 固定 IP のサーバ間の通信だけでなく, 個人の IP 不定なシステムが対象である事を特徴とする. P2P という手法は, 近年大きく発展してきており, 注目をあびるようになってきている. 現在 PC 上で動く実例としては, ICQ 等のチャットをするシステムや, SETI@Home のような利用されていない計算資源を活用するシステム, そして napster¹ 等のように, ファイルの共有をするシステム等がある. 最近とみに流行している, パームや, 携帯電話も P2P の実現の一つの方向である.

ここでは, P2P の中でも gnutella について論ずる.

2 gnutella について

gnutella とは, P2P のファイルシェアリングを実現するプロトコルである. サーバを介さずに他のホストと通信し, ファイル名による検索とダウンロードを実現する.

gnutella を利用する人は, GnutellaNet に接続する事になる. GnutellaNet とは, インターネット上に構築された, 仮想的なネットワークである. それは, gnutella プロトコルで相互に接続された計算機によりつくられたネットワークである. それぞれのマシは, 各々4台程度のマシンに接続する. そして, その接続相手も, それぞれ他のマシンに接続して, その再帰的な構造により, 大きなネットワークを構成する. お互いに ping と pong という接続確認のメッセージを送りあい, GnutellaNet 上をバケツリレー形式で検索とその応答のメッセージを送る. 実際にファイルのダウンロードは, 直接的な HTTP によるセッションが開かれる (Fig. 1). gnutella の初期設定では, 一台のホストは, 4 台同時に接続するようになっている. そして, 5 つ先までの TTL²をもち, メッ

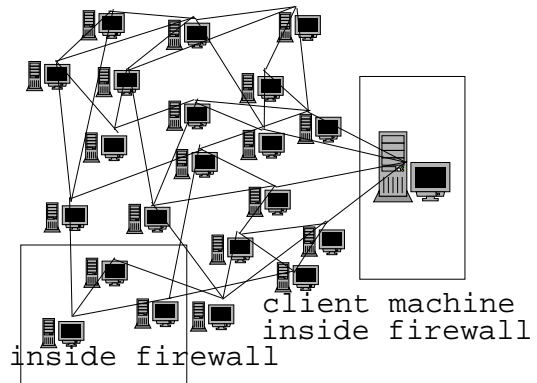


Fig. 1 Image of connection of the gnutellaNet

セージは, 5 台先まで届くようになっている³. 単純計算すると, $4^5 = 1024$ 台先まで, メッセージが届くという計算になっている.

gnutella は, ファイル名で GnutellaNet 上で検索をかける事ができる. そこで交換されているのは, 音楽ファイル, スピーチ, アニメや映画, コンピュータプログラム等のようなものである. gnutella は, ファイルの交換と検索サービスを提供している. しかし gnutella というものは基本的には相手ホストを検索するためのフレームワークであり, 拡張して他の用途に応用することは可能である.

3 gnutella の問題点

3.1 ネットワークの問題

まず問題として, gnutella では, 検索したあとの結果が返ってくるまでが遅いことがあげられる (図 Fig. 2). 5 台先までバケツリレーして検索している事になるので, 遅延が生じている. 10 分くらいたってからもまだ新しい検索結果が到着するということがしばしばある. リアルタイムで利用できる google 等の検索と比べると, 非常に遅い.

gnutella プロトコルはスケールしないと思われる.⁵

³gnutella プロトコルの仕様上, 5 以上の TTL を利用するには, 全員が同意する必要があるので, 難しい.

¹napster は, サーバを必要とするので真の P2P ではないと考えられる.

²time to live

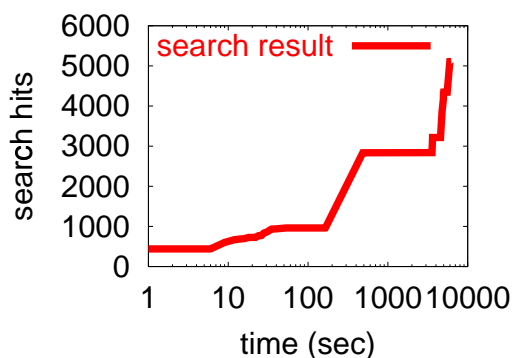


Fig. 2 Number of replies from GnutellaNet searching for "lain"

Table 1 Data flow in gnutella when idle

時間	流量
19:00-20:00	8.0MB/h
20:00-21:00	4.4MB/h
21:00-22:00	7MB/h
22:00-23:00	4.5MB/h

台先までメッセージが届くようになっている場合に、それぞれのピアに直接4台づつつながっているなら1000台程度がつながっている事になる。これより多くの人と交流しようとしてTTLを増やすと、指数関数的に流量が増えることになる¹⁾。

実際に自分で現実の状況を計測した所、一秒あたり6回くらいのペースで検索が流れてきた⁴⁾。これは、検索結果とあわせると大量のデータ流を生み出している(表Fig. 2)。

3.2 検索効率の問題

現在の状況としては、gnutellaでは基本的にファイル名でキーワード検索がかかるようになっている。一部のgnutellaソフトウェアはmp3等に埋め込まれているID3tagを読んでくれるようになっているようである。このことを応用して、もっと知的な検索方式を採用する事もできるはずである。ただ、それぞれのピアのソフトウェアが独立して返事のかえすことになっているので、形式の統一ということは難しい。

通常のクライアントサーバの世界ではダウンロードだけするという人の存在は、特に問題は無い。しかし、gnutellaのようなシステムでは、全ての人がクライアントでありサーバであるゆえに、その中で、サーバである事を放棄するピアがいると問題がある。TTLが5で到達できるマシンの数は、有限である。その中のシステム

⁴⁾4月23日の午前10時において1分間計測した。

のうち、現在ダウンロードを許しているのは、10%程度らしい。これですますダウンロードしかなしない人が参加すると、TTLが5のままでは十分多くのダウンロードができるホストまで届かず、有用な検索結果が得られないという事となる。

3.3 著作権関連の問題

GnutellaNet上で配布されている音楽のmp3形式のファイルや、アニメや映画のmpegやaviファイルは、これらは、個人的な興味で交換しているようである。この動画と音声の交換から企業利益が導出できるのだろうか、という疑問がわく。

現在の資本主義経済の考えとgnutellaの考えの間の乖離は、問題である。著作権周辺の話題に関して、その効力を強制できるだけのものが紳士協定以上の物が無い事が問題になる。法律によってのみ守られる著作物は、容易に複製できる、販売する権利を寡占することにより利益を得るというビジネスモデルは、gnutellaなどのファイル共有システムと真向から対立する。

4 展望

今現在考えられている次世代ファイル共有プロトコルとしてfreenet⁵⁾や、gnutella²⁾6)があがっている。現実問題として、よりスケール可能なシステムが切望されている。プロキシの問題もあるので、IRCのように、複数のサーバがP2Pで接続し、それらにクライアントが接続するというようなシステムも要求されるようになると思われる。

現状では、普及しすぎるとネットワークが輻輳し、十分な検索能力も無いということになるので困る。より効率のよい通信方法というものが重要である。

ネットワークのバンド幅が相当大きくなると、gnutellaはこのままでは利用できないと思われる。検索エンジンとしての利用をする時も、より速く検索結果が帰って来ることが要求される。これは、gnutellaのみの問題ではなく、P2Pのかかえる本質的な問題である。この問題点を解決したプロトコルが登場した後でないとして、gnutella型P2Pプロトコルは普及しないだろう。

参考文献

- 1) <http://www.darkridge.com/~jpr5/doc/gnutella.html> Why Gnutella Can't Scale. No, Really.

⁵⁾暗号化されたP2Pのプロトコル。

⁶⁾gnutellaを後継するプロトコル。作業された形跡はみあたらない