

# 効果的なデータ拡散に向けたオーバレイトポロジの比較

齋藤 義文<sup>†1</sup> 阿部 洋文<sup>†1</sup> 梅村 恭司<sup>†1</sup>

## 1. はじめに

インターネット上でなんらかの障害が発生すると、ネットワーク上のデータへのアクセスが不可能になることがある。障害発生時でもデータにアクセスするためにデータの複製を複数個用意しておき、ネットワーク内の多くの場所に配置しておく、データ拡散システムがある。データ拡散システムにおいて複製データに変更が生じた場合、更新データをネットワーク内のノードに配布する必要がある。この更新データの配布には時間がかかる。更新データをすべてのノードに配布し終えるまでの時間が長くなると、更新中にまた別の障害が発生してしまう恐れや、すでに新しい更新データがある場合でも、古い更新データを配布してしまうといった恐れがある。

これらの問題を解決するため、いかに速くすべてのノードにデータを配布し終えるかや、いかに早い時期に多くのノードがデータを持っているか、といったことが重要になる。データの配布時間は更新データをネットワーク内のどこを経由してどこに配置するのかという、データ拡散方式によって左右される。データ拡散方式はそれを構成するオーバレイネットワークの構造で性能が異なるため、データ配布に適したオーバレイネットワークの構造を求めたい。こういったオーバレイトポロジがデータ配布に適しているかを調べるため、様々なトポロジの特徴を知る必要がある。

今回は、ネットワーク実験のテストベッドとして PlanetLab を用い、PlanetLab 上で代表的なトポロジのオーバレイを構築してデータ転送実験を行い、各トポロジでのデータ転送時間の比較を行った。

## 2. 評価方法

### 2.1 PlanetLab とは

ネットワーク実験の評価方法として、実際の環境を用いる場合やシミュレーションを行う場合などがある。

実際の環境を用いる場合、機器の用意や実験規模といったコスト面の問題、シミュレーションの場合は結果の精度といった問題がある。そこでテストベッドを用いる。テストベッドとは大規模なシステム開発で用いられる実際の運用環境に近づけた試験用プラットフォームの総称であり、PlanetLab<sup>1)</sup> もその一つである。

PlanetLab とは、一定の条件を満たすことで企業、大学を問わずそこに所属する研究者は誰でも参加可能な研究ネットワークである。実際のインターネットに接続された何百台の PC を使用してオーバレイネットワークを構築し、従来、一人では不可能だった分散アプリケーションなどの研究を行うことが可能になっている。PlanetLab は 2008 年 11 月 4 日現在、473 サイト 919 ノードから成っている。

### 2.2 評価方法

一般的なデータ転送の評価ではネットワークの帯域幅を比較することが多い。我々は、データ拡散において重要となる、ノードにデータを配布し終える速さと、データを配布し終えたノードの割合という 2 つの観点を合わせ持った指標として、

$$\int_0^{\infty} (1 - f(x)) dx$$

で表される面積を比較することを提案する。 $f(x)$  は時間に対するデータ配布が完了したノードの割合の関数である。

例として、図 1 のような横軸にデータ配布時間、縦軸にデータ配布が完了したノードの割合をとったグラフを考える。このグラフでは、配布の過程において、「データ配布開始直後はあまり多くのノードがデータを所持していないが配布完了直前に急激に割合が増える」破線で示した曲線、「最初から多くのノードがデータを所持しているが全体に行き渡るまで時間がかかる」実線で示した曲線、といった様々な傾向が考えられる。このような各曲線において、斜線で示した部分はデータを持っていないノードの期待値であるため、この面積を比較し、より小さな面積である方が「速い時間で多くのノードがデータを所持している」といえる。こ

<sup>†1</sup> 豊橋技術科学大学

Toyohashi University of Technology

これはデータ配布中に障害が発生した場合でも、既に多くのノードにデータを配布し終えた状態であることを示す。このような状態の方が配布時間が多少遅い場合でもデータ拡散に優れていると考えられる。

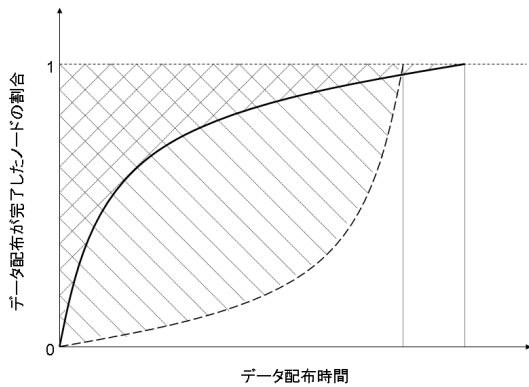


図 1 データ配布率の例

### 3. 評価システム

評価の際のデータを収集するため、ノード間の転送速度を計測するためのプログラムを作成した。この計測プログラムは BitTorrent を基にしており、Peer-to-Peer によるデータ転送を行い各ノードにデータがすべて配布し終えるまでの時間を計測するものである。BitTorrent との違いは、ノードの通信先を制御できることである。

BitTorrent の特徴のようにデータを断片化し、断片化したデータ単位で転送を行うことにより通信を高速化している。どのノードがどの断片を所持しているかという情報をまとめるトラッカーが存在し、ノード間の転送を制御する。

### 4. 実験方法

作成した計測プログラムを用いて PlanetLab 上でデータ転送実験を行った。PlanetLab 上で表 1 に示す国内 7 サイトのノードを確保し、ソースノード (pl2.planetlab.ics.tut.ac.jp) に置かれた 10MB のデータをすべてのノードに配布し終えるまでの時間を計測した。

今回は例として、オーバーレイトポロジは、メッシュ型とリング型の 2 つのトポロジを用いた。各トポロジでノードの配置を固定して 200 回の転送実験を行い、その平均を求めて比較に用いた。

参考に、実験で得られた値を元に、横軸を転送時間、縦軸をデータ配布が完了したノードの割合としたグラ

表 1 実験で使用了 PlanetLab 上のノードリスト

pl2.planetlab.ics.tut.ac.jp
pl2-higashi.ics.es.osaka-u.ac.jp
planet1.jaist.ac.jp
planetlab-04.naist.jp
planetlab1.iit.u-tokyo.ac.jp
planetlab2.sfc.wide.ad.jp
pub2-s.ane.cmc.osaka-u.ac.jp

フを図 2 に示す。

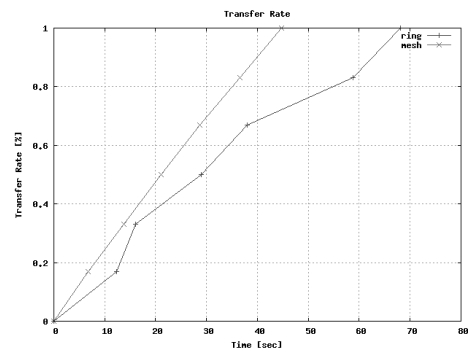


図 2 各トポロジのデータ配布率

メッシュ型のトポロジはほぼ時間に比例して割合が増加しているのがわかる。リング型のトポロジは、遅いノードがボトルネックになり、全体の完了時間が遅くなっているように思われる。

得られたグラフに我々の提案する評価指標を用いると、ノード数が 7 サイトの場合はメッシュ型の方が面積が小さく、データ拡散に適していると結論付けられる。

また、実験全体を通して遅いノードがボトルネックとなり、結果に大きく影響している。実際の環境では速いノードから遅いノードという様々な速度のノードが存在しているため、どのような状況下でもうまく動作することを念頭に置かなければならない。しかし実験時は極端に遅いノードの扱い方を考慮する必要がある。

### 5. おわりに

効果的なデータ拡散に向けた評価方法を提案し、そのための評価システムを作成した。今後の課題として、実験で得られた結果から評価指標を元に評価を行い、効果的なデータ拡散に向けたオーバーレイトポロジの比較を行う。

### 参考文献

- 1) PlanetLab. <http://www.planet-lab.org/>.