3Aa-03 インターネットにおけるトラッフィク解析

3Aa－3

串田高幸* 佐藤 卓由 山内長承
日本アイ・ビー・エム株式会社 東京基礎研究所
日本アイ・ビー・エム株式会社 コンテンツ事業開発部

はじめに

インターネットにおいて、トラフィックデータを収集解析することは、ネットワーク管理においてネットワークの状態を正確に把握するための重要な項目の一つである。インターネットでは、TCP/IPプロトコルを利用しているため、そのパケットのヘッダーを収集することによりネットワークのトラフィックを解析することができる。

一般にネットワークのバックボーンにおいては、大量のトラフィックが常に流れている。そのトラフィックデータのすべてを収集解析するには、測定装置も非常に高価になってしまう。さらに解析にも時間がかかる。もしネットワークトラフィックの測定に対して必要最低限度のパラメータを規定し、その規定に基づいてトラフィックの収集を部分的に行なうことで、ネットワークの状況がわかれば、その測定手段は効率がよいため有効な方式となる。

本稿では、インターネットにおけるトラフィックの収集解析を行なうためのシステムについて、その概要を述べる。そしてトラフィックの収集と解析についての方針について述べてゆく。また実際にインターネットにおいて収集したトラフィックのデータも示す。

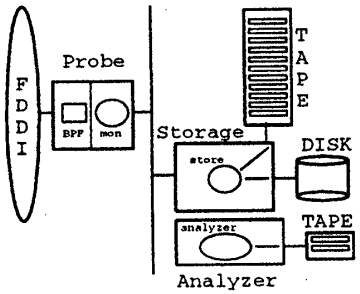

図 1: トラフィック収集解析システム

Analysis for the Internet traffic
Takayuki Kushida(kushida@trl.ibm.co.jp), Takuyoshi Satoh, Nagatsugu Yamanouchi
IBM Research, Tokyo Research Laboratory

トラフィックの収集、蓄積及び解析装置

本研究では、インターネットのトラフィックを収集解析するためにネットワークモニターとして汎用ワークステーションを利用したネットワーク収集解析システムを開発した。そして、そのシステムを実際に省際ネットワークで運用している[1][2]。図1が、本システムの概要である。このシステムでは、FDDIバックボーンに流れている全トラフィックを一旦、収集し、そのデータを連装型の自動切換の7GB磁気テープ装置に記録してゆく。

このシステムでは、全トラフィックのデータを収集しているため、パケットヘッダーに関して、すべての解析を行なうことができる。そのため事前に定義したデータを収集する場合とは異なり、何度も解析方法を変更して、トラフィックの解析を行なうことができる。

トラフィックの解析は、磁気テープに記録されたデータについて読み出して行なう。一旦、ハードディスクに記録されたデータをフィルター形式の構成ファイルによって、トラフィックの必要な情報だけの積算を行なってゆく。このプログラムを全トラフィックの積算を行なうことで、トラフィック情報を集めることができる。つまり、このフィルターをどのようにして構成し、そして結果をどのように使用するかにより、トラフィックの解析方法が決まってくる。

ネットワークとトラフィック

ネットワークの状況を知るためにネットワークのトラフィックを収集解析するという仕事は、以前から行なわれている。インターネットのトラフィック収集解析では、NSFNETバックボーンにおけるトラフィック収集とその解析が、よく知られている[3]。そこでは、専用の特別なルータを使用し、全米に散らばっているバックボーンノードからのトラフィックを収集し、その解析を行なっていた。

一般にネットワークにおいてトラフィックの状況を知ることによって、a) 現在のネットワークのトラフィック量、b) 過去からの変化により将来の予測、c) 日常のデータの解析から異常なデータの検出、d) 短期間の変化から長期間の変化を検知、e) 個別課金のために必要な情

報の入手、が可能になる。

またトラフィックは、1) 全体のトラフィック、2) プロトコルごとのトラフィック 3) アプリケーションごとのトラフィック及び4) 組織ごとのトラフィックというクラスに分けて収集することできる。また、これらのクラスには、それぞれパケット数とバイト数というの2つの単位がある。

また時間に対する変動は、どちらかというと周期的になっている。そのため時間は、以下のようなクラスに分けられる。

- 昼夜のトラフィック変動
- 曜日 (平日と週末) による変動
- 月の間での変動
- 月ごと (季節) による変動

部分的な収集と解析

主要なバックボーンネットワークにおいて流れているデータは、パケットヘッダーだけでも1日約数GBと膨大なデータである。そのため、その膨大なデータを解析し、その結果を出すには、処理時間が非常にかかってしまう。もし部分的な収集解析によって、ネットワーク全体の状態がわかるならば、全体のトラフィックデータを取得せずにすむので記憶容量を少なくでき、さらに処理の時間を短くすることができる。そのため、効率のよい収集解析のために部分的な収集方法の規定を考える必要がある。

トラフィックデータの部分的な収集方法は、時間に対して周期的な部分を選択する方法と、時間に対してランダムに選択する方法がある。また収集データ量に対して常に一定量を選択する方法も適用することができる。さらに特定のトラフィックのイベントによって、トラフィックを収集することもできる。以下に部分的な収集方式についてまとめておく。

- 時間に対して周期的な収集
- 時間に対してランダムな収集
- 常に一定データ量の収集
- イベントによって収集

部分的に収集したトラフィックから全体のトラフィックを算出するためには、部分的に収集したデータに対して、特定の関数を適用する。最も簡単な方法としては、例えば、一定の時間、部分的な収集を行なった場合に、その関数として単純な正比例を使うことである。その結果、全トラフィックは、部分的な収集の積算結果を単純に時間でかければよいことになる。

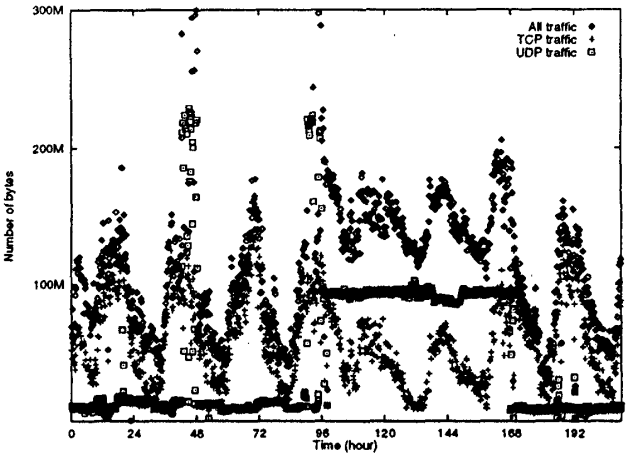

図 2: トラフィック収集結果

実際の収集データ

図2では、実際に省際ネットワークバックボーンにおいて、数日間、収集したデータのうちIP/ARPの全データ量、TCPデータ量及びUDPのデータ量を10分おきに加算した結果をグラフに示している。

このデータでは、全パケット数が、TCPのパケット数とほぼ同じ傾向で変化している。また昼夜の時間変化や、平日と週末の変化も顕著でている。

一方、UDPデータは、全データ量とはまったく関係なく、時間に対して比較的緩やかな変化をしている。しかしUDPには、特定の時間に極端なトラフィックの変化がみられる。これは、UDPを利用したアプリケーションが、その時間に動作したと予測される.

参考文献

1. 串田高幸,「ネットワーク解析ツールの設計」. 平成7年情報処理学会第51回全国大会. 1995年9月.

2. 串田高幸,「インターネットのトラフィックを測定及び解析するためのツールの設計及び開発」. 情報処理学会マルチメディアと分散処理ワークショップ. 1995年10月.

3. K. C. Claffy, H. Braun, and G. C. Polyzos. *Tracking Long-Term Growth of the NSFNET.* CACM Vol.37 No.8 1994.