5J－9

インターネット運用のための トラヒック分析技術の一考察

池上　聡　浜田　雅樹　小松原　重之　三上　博英
NTTソフトウェア研究所

1．はじめに

現在、インターネットは急速に普及が進んでいる。これに伴い、ISP（Internet Service Provider）など、インターネットを運用する側では、トラヒックの増加に対して設備投資などの対応を行う必要が生じている。設備投資などを行うためには、トラヒックの増加が一過性のものか、増加傾向を示すものか等を判断する必要があり、そのために、トラヒックの分析が必要となっている。

今日、トラヒックに関連する技術や標準としてはプローブ（RMON、RMON2）やSNMP、SNMPv2等が存在する。これらは、トラヒックを収集したり、統計計算をするプリミティブな仕組みを提供するもので、ISPが運用する大規模なネットワークで、これらを使用するには問題がある。本稿では、それらの問題点を明らかにし、ISPが設備投資などを行うために必要なトラヒックの収集、分析方法について考察する。

2．トラヒック分析の必要性

ISPはトラヒックの増加に対して、設備投資やトポロジの再構成（回線帯域の増設やトポロジの変更等）を行う必要があるが、これらを適切に行う技術は確立されていない。

電話網においては、トラヒックの傾向を判断して設備投資を行う技術は確立されているが、この方法をインターネットにそのまま利用することは困難である。その大きな理由のひとつは、インターネットでは使用するアプリケーションによって発生するトラヒックが異なり、さらに、利用されるアプリケーションは多種多様で、現在も発展途上であることである。実際に、あるユーザが画像伝送を伴うアプリケーションを利用したために輻輳が発生した例などがある。この例では、あるユーザが一時的に画像伝送を伴うアプリケーションを利用しただけなので、すぐに設備投資を検討する必要はない。このように、ISPが設備投資や再構成を行うには、高負荷が一過性のものか、増加傾向を示すものか等を判断する必要がある。その判断を行うためには、統計的な情報だけでなく、高負荷時にどのようなアプリケーションが使用されているのか、また、特定のユーザによるものか不特定多数ユーザによるものかなどの分析が必要である。（以降、個別的分析と呼ぶ）

A Study of Traffic Analysis for Internet Operation
Satoru IKEGAMI, Masaki HAMADA,
Shigeyuki KOMATSUBARA, Hirohide MIKAMI
NTT Software Laboratories

3．現状のトラヒック収集技術とその問題点

3.1　現状のトラヒック収集技術

インターネットの管理には、SNMP(Simple Network Management Protocol)に代表されるネットワーク管理プロトコルを用いて、ネットワークノードよりデータを収集する方法や、プローブと呼ばれるパケット監視装置を用いて、ネットワークリンク上を流れるパケットを収集する方法が用いられている。前者では、転送パケット数、転送データ量、エラーパケット数等の情報が得られるが、個別的分析には不十分である。

プローブ（パケット監視装置）とは、ネットワーク上を流れるすべてのパケットを一旦取り込み、その全体または一部を保存、解析する装置のことである。プローブを用いてパケットのヘッダを解析すれば、２章で述べた個別的分析に必要な情報、つまり以下の情報を得ることができる。

①パケットを生成したアプリケーションの種類
②パケットの発信地
③パケットの着信地

以下、プローブにより得られるパケットのヘッダの集合をヘッダデータと呼ぶことにする。

3.2　プローブを大規模ネットワークで用いる際の問題点

①データ量が膨大

ISPのバックボーン等ではデータ量が多く、リアルタイムに解析することは困難である。例えば、省際研究情報ネットワーク(IMnet)[1]のFDDIバックボーンではその量が一日あたり数GBになる[2]。そのため、そのデータを保存しておき、後日解析したりする必要がある。さらに、保存する場合にも、データ量が膨大であるために多大なリソースが必要であり、長期保存も困難になる。

②プローブの設置数が多い

バックボーンでは収集できないトラヒックが存在する。例えば、図1のようなトポロジの場合には、接続組織間のトラヒック（図１の組織Aと組織Bのトラヒック）はバックボーンを通らないため、各接続組織とのリンク上（図1の＊）にプローブを設置する必要がある。そのため、大規模ネットワークでトラヒックを収集する際には多数のプローブが必要になる。

4．アプローチ

３章で明らかにした問題点に対する解決法として表１に示した方法が考えられる。

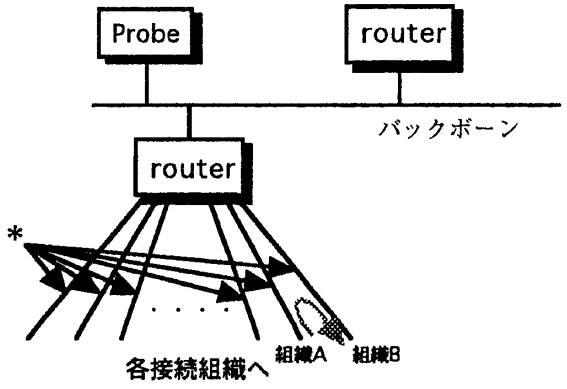

図1 プローブの設置場所

表1 問題点の解決法

問題点	手法
データ量が膨大	高負荷時にヘッダデータを保存する方法
設置数が多い	動的にプローブを配置する方法
	ノードでトラヒックデータを収集する方法

4.1 高負荷時にヘッダデータを保存する方法

通常のトラヒックの場合には、トラヒックの一定時間ごとの統計データを使った運用がほとんどである。例えば、IMnetではトラヒックデータをアプリケーション別、一日単位の統計データとして保存している[3]。統計をとるだけであれば、その数をカウントしてリアルタイムに統計データを作成すれば、ヘッダデータを保存する必要はない。一方、高負荷時には個別的分析が必要であるため、ヘッダデータを保存する必要がある。そこで、高負荷時にのみヘッダデータを保存する方法が確立できれば、データ量の問題は解決できる。実現方法のイメージを図2に示す。プローブでは、常に統計データを保存し（＊）、高負荷時には次の(1)~(3)のような制御を行う。

(1) ネットワーク上のある場所で高負荷が発生すると、NMS(Network Management System)がそれを検知する。
(2) 検知すると、NMSはプローブにヘッダデータを保存するように指示する。
(3) プローブがヘッダデータを保存する。

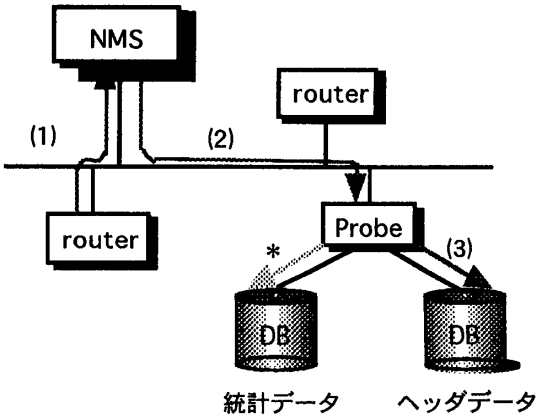

図2 高負荷時にヘッダデータを保存する方法

高負荷の検知はそれぞれの回線の回線使用率にしきい値を設けることで行う。回線使用率は以下のような方法で計算することができる[4]。ルータのMIBより、回線帯域、ルータが起動してからの時間（システムアップタイム）、インタフェースごとの伝送パケット量の情報を取得し、この値と前回の収集結果の差分により算出する。

この方法を実現するためには主に以下の事項について検討する必要がある。

・回線使用率のしきい値の決定法
・高負荷時におけるプローブ制御データの伝達方法
・高負荷時のヘッダデータの保存形式
・プローブの制御方式

4.2 動的にプローブを配置する方法

動的にプローブを配置できれば、トポロジの制約を受けずにトラヒックの収集が可能となる。ネットワーク負荷の度合に応じて、ファイヤーウォールなど、他の目的で使われているワークステーション上にヘッダデータ収集プロセスを送り込み、動作させる方法である。

この方法を実現するためには主に以下の事項について検討する必要がある。

・プロセスを起動させるしきい値の決定法
・データの回収方法
・動的配置のメカニズム
・CPUの負荷やデータの量
・セキュリティ技術

4.3 ノードでトラヒックデータを収集する方法

ルータでアプリケーションの種類、発信地、着信地のデータが取得できれば、プローブの設置数の問題は解決できる。しかし、ルータで、本来のルーティングに加え、データ収集を行うと大きなCPU負荷となる。よって、この方法は、前述したデータ量の問題やCPU負荷を考えると容易ではない。マルチプロセッサによる高速高機能化等の手法が求められる。大規模ISPを運営するためには将来的に有望な方法と考えられる。

5. おわりに

本稿では、大規模ネットワークにおいて、個別的分析を行うためのトラヒックを収集する際の問題点を明らかにし、その解決法のアイデアを紹介した。今後は更に検討を進め、実際のISPで有効性を検証する予定である。

参考文献

[1] 鈴木亮一、福田晴元、三上博英 "省際研究情報ネットワークの構築について" 情報処理学会 第52回全国大会論文詩集
[2] 串田高幸、佐藤卓由、山内長承 "インターネットにおけるトラヒック収集と解析" 情報処理学会研究報告 マルチメディア通信と分散処理 75-3, pp 13-18 1996
[3] 小松原重之、鈴木亮一、三上博英 "インターネットにおけるトラヒック解析法の研究" 情報処理学会 第52回全国大会論文詩集
[4] 福田晴元、小野諭、高橋直久 "インターネットにおけるQoSビジュアライザ" 情処第74回 DPS研究会 74-13, pp73-78 1996