

WWWサービスにおける混雑状況の取得方式に関する研究

3P-9

鷲見卓哉* 坂本泰久**

NTTソフトウェア研究所* NTTソフトウェア(株)**

1. はじめに

現在のWWWサービスが内包する問題の一つに、利用者がファイルを取得するまでの待ち時間の変動が挙げられる。遅延の原因としてはネットワーク・サーバ等における種々の要因が考えられるが、この改善はインターネットが分散システムである性質上一元的な対処は難しく、遅延を一定に保つ事は難しい。

これらの現状をふまえて考えた場合、遅延時間、すなわちWWWサービスの混雑に関する情報は利用者及び管理者の双方にとって有意義な環境情報であると考えられる。また実際のファイル転送よりも短時間かつ低トラフィックでユーザの端末上における取得が可能であれば、本情報は利用者にとってWWWアクセス時に参照可能な有益な指標の一つとなりえ、不必要な接続による混雑の抑制と言うサーバ側にとっての利点ともなる。

本研究では、ユーザ端末上の測定プロセスによる取得データのみを用いてWWWサービス利用時のファイル転送時間の目安である混雑状況を取得する方式を提案し、実際に測定実験を行うことによって有効性の検討を行った。

2 混雑度情報の取得

インターネットの混雑という現象を考えた場合、観点により様々な解釈が考えられる。本研究では対象をWWWサービスに限定することにより「WWWサービスにおけるユーザの待ち時間=混雑度」と定義する。具体的にはユーザの待ち時間としてHTTPプロトコルのGET命令による実際のファイル転送に要した時間を基準とし、この値と相関性が高く、かつ短時間・低トラフィックで取得が可能なパラメータの探究を行う。

3.実験

3.1 測定項目

本実験ではユーザが実際のファイル転送に要する時間と比較する項目として、以下のラウンドトリップタイム(以下、RTTと省略)を測定した。

A. HTTPのConnection確立時間

B. HTTPのHEAD命令のRTT

C. HTTPのGET命令のRTT

D. ICMPエコー要求のRTT

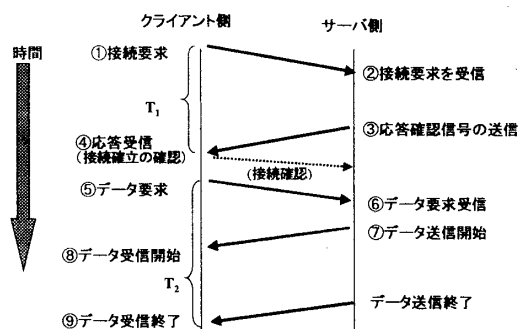


図1:接続(コネクション)確立の手順

WWWサービスに用いられているHTTPの低レイヤープロトコルであるTCPによる通信手順の概要について、図1に示す。

TCPにおいては実際のデータ転送前に、クライアント-サーバ間で接続の確立(①～④の処理)が行われ、データの伝達及びサーバでの接続処理に幾ばくかの時間が費やされる。実際のデータ転送処理は⑤～⑨となり、転送するデータの量によって所要時間は大幅に異なる。

実際のファイル転送に要する時間は①～⑨の総所要時間(T_1+T_2)であり、項目AのConnection確立時間は図1における T_1 、項目B,CのRTTは T_2 となる。項目BのHEAD命令によってサーバから返されるヘッダ情報は通常1パケット程度に収まるため、Bは⑤～⑧の所要時間と大まかにみなすことが出来る。CのGET命令のRTTは総所要時間の大部分を占め相関性は高い。しかし、測定に要する時間・パケット量も増大するため今回指標としては採用せず、総所要時間の算出のために測定する。

DのICMPエコー要求は一般にPingと呼ばれて

“A Research on the Method to Get WWW Services Traffic Information”

*Takuya Sumi **Yasuhisa Sakamoto

*NTT Software Laboratories

** NTT Software Corporation

いる、対象サーバに送出した ICMP パケットの応答が戻るまでの RTT を計測するものである。今回、項目 A~C との比較のために測定した。

3.2 測定及び結果

上記の A~D の 4 項目の RTT をインターネット上の複数のサイトに対して一定時間間隔で測定し、その結果から相関性の評価を行った。あるサイトに対する測定結果の時間変化を下図に示す。混雑時に最も値の増大が顕著な項目が項目 A であり、他は RTT 値の小さい順に D,B,C である。

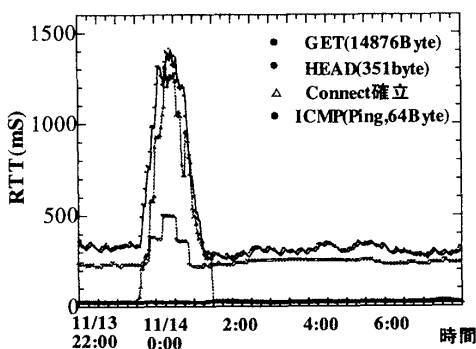


図2:各項目のRTT時間変化

結果として最も総所要時間と相関性が高い項目は A の Connection 確立時間である。項目 B の HEAD 命令の RTT はある程度の相関が見られるものの、項目 A ほどではなかった。

尚、項目 D の A~C との相関に関しては、高い相関が見られるサイト、RTT が一定で相関が見られないサイトなど、サイトによる相違が大きい結果となった。

4. 考察

項目 B の所要時間との相関がさほど高くなかった理由として、パケット数の相違が推測できる。サーバからクライアントに送り出されるパケットは、サーバが複数の要求を処理しているような場合などは連続しているとは限らない。そのため、1 パケット分のデータの RTT ではサーバ側の混雑によるパケット間の時間間隔を十分には把握できず、相関性が低下したのではないかと考えられる。

項目 A の Connection 確立時間が項目 C、ひいては総所要時間と相関性が高くなった理由として、サーバが Connect 要求を迅速に受理できないため生じた接続待ちの時間が、サーバの負荷である接続受付数に比例しているからではないかと考えられる。

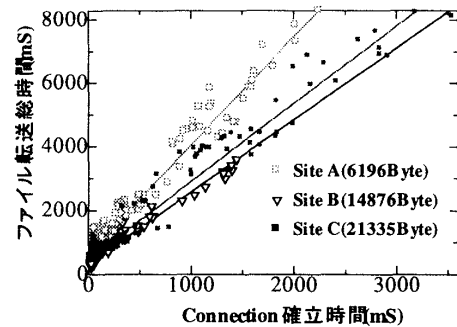


図3: Connect時間と総所要時間の相関

図 3 に、図 2 のサイトを含む 3 つのサイトについての項目 A と総所要時間との散布図を示す。一次関数プロットで近似が可能な正の相関が観測されており、この相関が該当するサイトでは混雑度の指標を項目 A より算出することが可能となる。

また項目 D と ICMP の RTT と A~C の値との相関については、ICMP がネットワークの要因による遅延により強く影響されると考えると、両者の相関度より各サイトの混雑の要因が途中のネットワーク経路によるのか、サーバの接続要求過多によるものかの区分が可能になると考えられる。

5. おわりに

本研究では、ユーザ端末から少ないトラフィック量で、WWW サービスにおける待ち時間を混雑情報という形式で取得するための方式を提案し、実験による検証を行った。その結果より、HTTP における Connection の確立に要する時間が実際のコンテンツ転送時間と相関があり、混雑度のための有効な指標となりえる事が確認された。Connection 確立時間は HTTP のセッションにおいて容易、かつ実転送時間より短時間で取得できるため、混雑度取得の方式として有効である。また、混雑情報はユーザ作業の効率化以外にもネットワーク及びサーバ状況把握などにおいて有益となると考えられるため、本方式の多方面への応用が期待できる。

参考文献

- [1]岡田,住友,谷口:「インターネットの測定」Bit Vol30, No7,8,9 1998
- [2]Graham,I. Donnelley,S. Martin,S. Mertens,J. Cleary,G.J「Nonintrusive and Accurate Measurement of Unidirectional Delay and Delay Variation on the Internet」INET'98 Conference
http://www.isoc.org/inet98/proceedings/6g/6g_2.htm