Improved Bandwidth Estimations using the IPv6 Hop-by-Hop Extension Header

M. Crocker, G. Lazarou, J. Picone and J. Baca
Intelligent Electronic Systems
Center for Advanced Vehicular Systems
Mississippi State University
{crocker, glaz, picone, baca}@cavs.msstate.edu

***Abstract*—Accurate measurement of bottleneck and available bandwidths in network paths is a problem that has intrigued researchers for years. Numerous tools and techniques have been developed to determine bottleneck and available bandwidths but none of these methods provide a simple solution that is accurate, efficient, and flexible. Such a solution is simply not practical for IPv4 networks. Fortunately, the next generation Internet Protocol, IP version 6, has the functionality necessary to implement feedback mechanisms to assist in accurate bandwidth estimations. In this paper we present a method to accurately and efficiently determine the network bandwidth in IPv6 networks through the use of a proposed timestamp option for the IPv6 hop-by-hop extension header.**

***Index Terms*—Internet, Protocols, IPv6, Bandwidth, Timestamp**

I. Introduction

Accurately estimating the bottleneck and/or available bandwidths in a network path is essential for the correct and efficient operation of many Internet applications and protocols as well as network management applications. End-to-end flow control, server selection for downloads and streaming media, peer-to-peer host selection and content delivery, and multicast configuration protocols are ~~a~~ just a few examples where accurate bandwidth estimations are valuable. Accurate bandwidth estimations are also valuable to network designers and administrators to aid in network troubleshooting, capacity provisioning, and traffic engineering.

Knowing the capacity of a network path is valuable in a number of situations, but the best methods for measuring bandwidth are essentially limited due to the nature of the network. The basic problem with all the current measurement techniques is that they are trying to infer characteristics about a network that is not designed to reveal any information to data that traverses it; this is especially true of the Internet. The network simply transports data to its destination as best it can, which is why the Internet is considered a best effort network. There are no guarantees for data traveling through the Internet and there is no way to accurately predict how data will be handled. The only way to determine link capacities and utilization of the links is by examining how the network delivers a single packet or sequence of packets to the destination.

Despite these limitations, researchers have been able to devise methods that can measure bandwidths in a network, even the Internet, with some degree of accuracy. Each method has its own limitations and tradeoffs as we

Summary of Comments on Microsoft Word - paper_v3.doc

Page: 1

Author: bisdik
Subject: Cross-Out
Date: 11/6/2005 5:35:42 PM

unpredictable network behaviors. The simple fact is that none of the methods currently in use can consistently and accurately measure bandwidth without help from the network itself.

III. IPv6

A. IPv6 Overview

Internet Protocol version 6 (IPv6) [7] will soon replace the current Internet Protocol version 4 (IPv4) (although IPv4 may never completely go away [8]). Full deployment of IPv6 can be expected to be completed within the next 10 years or sooner depending on the push from other countries and government agencies, not to mention the demand due to the increase in Internet capable devices and the constant threat of router meltdown due to unmanageable routing tables [8]. Initial deployment of IPv6 has already begun with the establishment of 6bone [9], the IPv6 Internet backbone.

When IPv6 is fully deployed, the current tools and techniques for measuring bandwidths should still be applicable but some of the techniques will have to be reevaluated. IPv6 is different in many respects from IPv4, so much so that some of the assumptions may not hold true and the approaches may need to be adjusted. IPv6 primarily differs with IPv4 in that it offers expanded addressing, simplified header format, and improved extension and option support [8].

These differences should not present any immediate problems to the present IPv4 measurement techniques. The differences that will demand the need for reevaluation is the increased header size, MTU, and fragmentation properties of IPv6. The header size in IPv6 is fixed at 40 octets whereas the IPv4 header is variable between 20 and 40 octets. The minimum MTU for IPv6 is set to 1280 octets, more than double the 576 octets of IPv4. The tailgating technique described in [2] uses a smaller packet followed by a larger packet in order to give a higher probability that the two packets will be queued adjacently at each router. The increase in the IPv6 header size will reduce the minimum packet size and therefore reduce the large packet to small packet ratio. This ratio is critical in maintaining back-to-back queuing in network paths where adjacent link capacities increase by a factor greater than the lead/tailgate packet ratio. One would expect the larger MTU of IPv6 to increase the ratio but most researchers assume that packets of 1500 octets can be sent without fragmentation and so the ratio commonly used is 1500 octets to 40 octets.

One benefit of IPv6 is that it eliminates fragmentation at the router level. Hosts must not send packets larger than the MTU of the receiver or an ICMP error message will be returned to the host. The elimination of fragmentation at the router level was decided in order to reduce the complexity and improve efficiency for all routers in the Internet. This decision should prove to benefit bandwidth estimations as the overall network traffic will have a more consistent composition and may result in more accurate assumptions about the network.

B. Timestamps

A timestamp option was initially defined for IPv4 in RFC 760 [10]. This original definition was a very limited

Page: 4

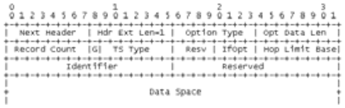
Author: bisdik
Subject: Highlight
Date: 11/6/2005 5:39:37 PM

Reference [7] dates back to 1998 and still IPv6 deployment is relatively sporadic at best. What facts support your statement for IPv6 "soon" replacing IPv4?

In addition to a timestamp option, ICMP messages may need to be defined to request and report timestamps. The ICMP ping message could be modified to copy the timestamp header from the request message into the reply message. In this manner, any host that is reachable with a ping message has the capability to return timestamp information about the incoming and return path to the sender. Integration into an ICMP message will eliminate the need for special software running on the destination host as well. The CSI mechanism defines ICMP messages to be used for link investigation but for the sake of brevity we will assume that a modified ICMP echo message exists that can carry the timestamp hop-by-hop extension to the destination host and back.

B. Timestamp Hop-by-Hop Format

The proposed timestamp option format is shown in Figure 2. This figure represents fields in the hop-by-hop extension header to specify desired handling of the timestamp option and it also includes fields for holding information about this packet.



```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Next Header  | Hdr Ext Len=1 |  Option Type  | Opt Data Len  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Record Count  |G|  TS Type    |  Resv | IfOpt | Hop Limit Base|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Identifier           |           Reserved            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
+                          Data Space                           +
|                                                               |
```

Figure 2 – IPv6 Timestamp Header

Each of the fields is described as follows:

Option Type: 8-bit integer identifying this option as the timestamp option. The IPv4 timestamp option type value is 68 but this cannot be used for IPv6. The IPv6 option must start with 001 indicating that routers should skip this option if they do not support it and that the data in this option may change en route to the destination.

Opt Data Len: 8-bit length of the timestamp option in number of octets. Option length starts from record count to the end of data space.

Record Count: 8-bit unsigned integer indicating the number of records contained in data space. Each router must increment this field after inserting a record in the data space. The position for the next record can be calculated using [Record Count] * [Record Length] + 12.

TS Type: The TS Type field indicates the desired timestamp behavior. This is an 8-bit field with the upper bit R specifying a high resolution timestamp (finer than 1 millisecond) or a normal resolution timestamp (1 millisecond or less). The TS Type currently has 3 values specified below that follow the flags field in RFC 791. The unused values for this field are reserved for future use.

0 – insert timestamp record only

1 – insert Internet address of the registering entity before the timestamp record

3 – the Internet address fields are pre-specified. An IP module only registers a timestamp if it matches its own

Page: 7

Author: bisdik
Subject: Highlight
Date: 11/6/2005 6:05:36 PM

This and the next subsection IV.C read more than a IETF draft rather than a tutorial paper. Your readership are networking professionals and researchers in general and not (necessarily) just standards committee participants.

perform load balancing, traffic engineering, and network optimization. Current techniques for measuring available and bottleneck bandwidths are overall acceptable in certain applications but they suffer in accuracy, simplicity, and efficiency.

The measurement techniques in IPv4, although inaccurate, should be applicable in IPv6, but IPv6's improved options provide the resources necessary to implement a timestamp option that will make these techniques simpler, accurate, and more efficient. Our simulations show that timestamps used in conjunction with IPv6 provide an accurate solution and are up to 70% more accurate than the best estimation method for IPv4.

The greatest obstacle to making our technique a reality is getting the IETF to accept a definition for a timestamp option and making it a standard. The IETF may have reservations for such an option due to hardware limitations and the ineffectiveness of the IPv4 timestamp. The IPv6 specification is constantly undergoing revisions and improvements so a timestamp option would be a low priority next to completely defining IPv6. Even so, we feel that a timestamp option would benefit not only bandwidth measurements, but other areas as well. Until there is feedback from the network itself, consistent accurate bandwidth measurements will be a difficult challenge to overcome.

VIII. References

[1] V. Jacobson. Pathchar: A Tool to Infer Characteristics of Internet Paths. ftp://ftp.ee.lbl.gov/pathchar.

[2] K. Lai and M. Baker. "Measuring Link Bandwidths Using a Deterministic Model of Packet Delay," ACM SIGCOMM Computer Communication Review, vol.30, pp.283-294, Oct. 2000.

[3] K. Lai and M. Baker. "Nettimer: A tool for Measuring Bottleneck Link Bandwidth," in *Proc. of USITS*, pp. 123-134, Mar. 2001.

[4] D. Sisalem and H. Schulzrinne, "The Loss-delay Based Adjustment Algorithm: A TCP-friendly Adaptation Scheme", Workshop on Network and Operating System Support for Digital Audio and Video, 1998.

[5] M. Jain and C. Dovrolis. "End-to-end Available Bandwidth: Measurement Methodology, Dynamics, and Relation with TCP Throughput," in *Proc. of ACM SIGCOMM,* pp. 295–308, Aug. 2002.

[6] J. C. Bolot. "End-to-end Packet Delay and Loss Behavior in the Internet," in *Proc. ACM SIGCOMM*, pp. 289-298, Sept.1993.

[7] Internet Engineering Taskforce and Internet Engineering Steering Group: IPv6 Working Group. RFC 2460: Internet Protocol, Version 6 (IPv6) Specification. [Online] Available: http://www.rfc-editor.org/rfc/rfc2460.

[8] P. Loshin, "IPv6 Theory, Protocol, and Practice", 2nd ed., San Francisco, CA: Morgan Kaufmann, 2004, pp. 22-23.

[9] IPv6 Internet Backbone. [Online]. Available: http://www.6bone.net

[10] Internet Engineering Taskforce and Internet Engineering Steering Group: Network Working Group. RFC 760: DoD standard Internet Protocol. [Online]. Available: http://www.rfc-editor.org/rfc/rfc760.

Page: 14

Author: bisdik
Subject: Highlight
Date: 11/6/2005 6:08:41 PM
These statements are not appropriate for the paper. They read like a criticism of actions by IETF. Once again, your readership of this paper is not standards committee member participants to whom you're making a presentation.