DiffServ measurement

Lin Zhang
Columbia University
New York, NY 10027
USA
lz2110@columbia.edu

ABSTRACT

Differentiated services enhancements to the Internet protocol are intended to enable scalable service discrimination in the Internet without the need for per-flow state and signaling at every hop. A replacement header field, called the DS field, is defined, which is intended to supersede the existing definitions of the IPv4 TOS octet [RFC791]. Six bits of the DS field are used as a codepoint (DSCP) to select the PHB a packet experiences at each node. A two-bit currently unused (CU) field is reserved.

However , unfortunately Ipv4 packets containing DSCP are reputed to unsupported in some routers in the Internet, and the DSCP bits will get to lost (be zeroed) when the packets pass though those routers. To verity the view that some packets with DSCP do indeed lost those DSCP bits when they travel through the Internet, I have conducted a series of simple experiments, create the packets, set the DSCP bits , and send them to some destination across the Internet, and use probes distributed across the Internet to observe the header of the IP packets and measure whether ISPs allows DSCPs (DiffServ code points) in IP packets to cross unaltered , compute delay and compare with the packet without DSCP bits set.

1 . INTRODUCTION

To date, the Internet has mostly taken an egalitarian approach to packet scheduling in router queues. All packets receive equal service; no packets, including delay-sensitive audio and video packets, receive special priority in the router queues. No matter how much money you have or how important you are, you must join the end of the line and wait your turn! Due to the lack of any special effort to deliver packets in a timely manner, it is an extremely challenging problem to develop successful multimedia networking applications for the Internet.

Differentiated services is to introduce a small number of classes (possibly just two classes), assign each datagram to one of the classes, give datagrams different levels of service according to their class in the router queues, and charge users according to the class of packets that they are sending into the network.

The IP Precedence field is something of a forerunner of the DS field. IP Precedence, and the IP Precedence Field, were first defined in [RFC791].

Although early BBN IMPs implemented the Precedence feature, early commercial routers and UNIX IP forwarding code generally did not. As networks became more complex and customer requirements grew, commercial router vendors developed ways to implement various kinds of queueing services including priority queueing, which were generally based on policies encoded in filters in the routers, which examined IP addresses, IP protocol numbers, TCP or UDP ports, and other header fields. IP Precedence was and is among the options such filters can examine.

However , unfortunately Ipv4 packets containing DSCP are reputed to unsupported in some routers in the Internet, and the DSCP bits will get to lost or be zeroed when the packets pass though those routers. Despite this alleged lose of DSCP there does not appear to exist any recent measurements of how many percentage of packets lost their DSCP in their Internet trip .

The generally referred to reason, as to why packets containing DSCP lost it, is that some routers do not trust DSCP . A reason for this is not all QoS techniques are appropriate for all network routers. Because edge routers and backbone routers in a network do not necessarily perform the same operations, the QoS tasks they perform might differ as well.

To this end a series of simple experiments have been conducted using a tool name **Iperf** http://dast.nlanr.net/Projects/Iperf/ to create TCP and UDP packets with DSCP bits set to measure if routers actually zeroed the DSCP when the packets traveling on the Internet , and to examine if the packets with DSCP get more delay than otherwise. Even though our experiments are still at an initial stage, I believe that it is quite clear from out collected data that some packets carrying DSCP bits are in fact lose them during the end to end transmission on the Internet, briefly about one third from all DSCP set packets. However, the round trip time seems no difference between the DSCP set TCP connection and otherwise. Seems the router either zero the DSCP filed, which mean do not trust it, or the router just ignore it.

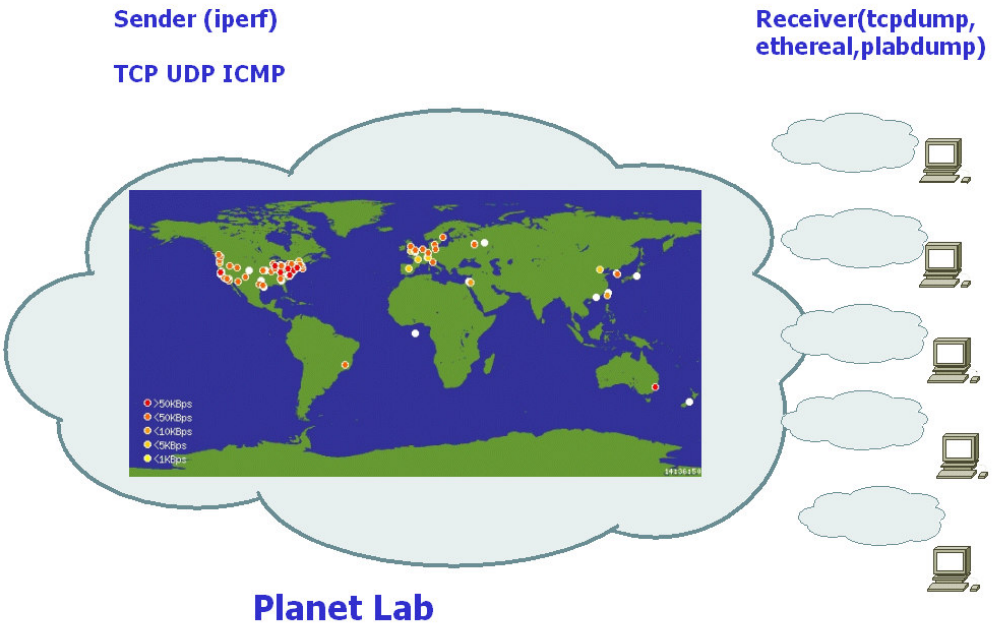
The remainder of this paper is structured as follows. In section 2, the setup and results of the experiments are presented. Section 3 some related work on IP packet measurement . Section 4 ,conclusions and future work.

2. Whether ISPs allows DSCPs (DiffServ code points) in IP packets to cross unaltered

Our hypothesis is that some packets with DSCPs bits will lost them , or be zeroed during travel across Internet , because some routers may not trust DSCP among the path . In this section I will present a set of experiments which support this hypothesis.

2.1. Measurement Methodology

To check whether ISPs allows DSCPs (DiffServ code points) in IP packets to cross unaltered, I use the network of **PlanetLab** http://www.planet-lab.org/php/top.php to did the measurement.



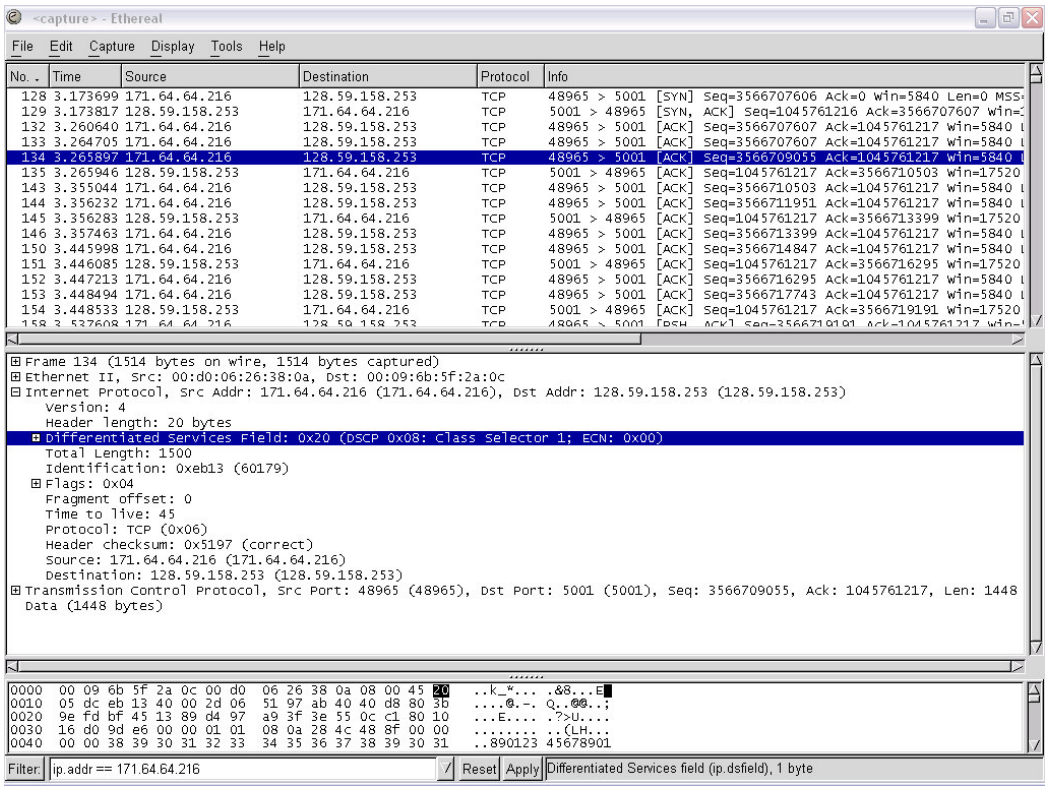
Do the measurement on the PlanetLab network

PlanetLab currently consists of 160 machines, hosted by 65 sites, spanning 16 countries. Most of the machines are hosted by research institutions, although some are located in co-location and routing centers (e.g., on Internet2's Abilene backbone). All

of the machines are connected to the Internet, currently peering with 100 ISPs. The goal is for PlanetLab to grow to 1,000 widely distributed nodes that peer with the majority of the Internet's regional and long-haul backbones. Much of this infrastructure was supported with funding from Intel. PlanetLab nodes are 1.0GHz IA32, PIII class processor or greater , and currently run a Linux distribution that is largely based on RedHat 7.3. My data sources include the nodes of some research universities and some labs from research companies .

On the planetlab's network environment (Linux) I installed the application of **iperf** http://dast.nlanr.net/Projects/Iperf/ on some node machines random selected all over the world , for sending TCP and UDP packets with the DSCP filed set . And then use the application of **plabdump**, a wrapper for tcpdump that can be used to observe traffic based on TCP or UDP port, as probes to observe those received DSCPs set packet .

At Columbia campus, to observe the received packets, I set up a machine with the application of **ethereal** http://www.ethereal.com/ , a network protocol analyzer, and I another well known application **tcpdump** http://www.tcpdump.org/ , which to prints out the headers of the receive packets . I used those tools to verify that the iperf applications were setting the DSCP bits in the IP headers.



output of ethereal

The measurement consisted of:

1. Starting iperf to create a TCP connect with the DSCP field set to 0x20 , which means Class Selector 1: 001000 [RFC2474], and send the packet out to public Internet networks destinate to another node of planetlab node or the probe machine in Columbia campus with ethereal or tcpdump software installed ;
2. At the receiver use plabdump or ethereal as probes to observe the received packet, to print the header out, examine the DSCP filed to see has been changed or not , use tcpdump to verify the results;
3. Use ethereal to create the average RTT graph of 10 seconds TCP packet transmit with the DSCP set;
4. Do the test again without the DSCP set and get the RTT of normal packets;
5. For those packets which lost their DSCP bits (be zeroed), use the same sender send the packet to another receiver , or use another sender to send packets to this receiver ; if this time the DSCP does remain in the packet , use the "traceroute" application to track the routers on the path, to find out at which point the packets lost their DSCP bits ;
6. Compare the RTT (round trip time) of the packets with DSCP set which were successfully received, to the normal packets.

Below is a DSCP carrying packet received by a probe at the receiver , of the output of application ehtereal :

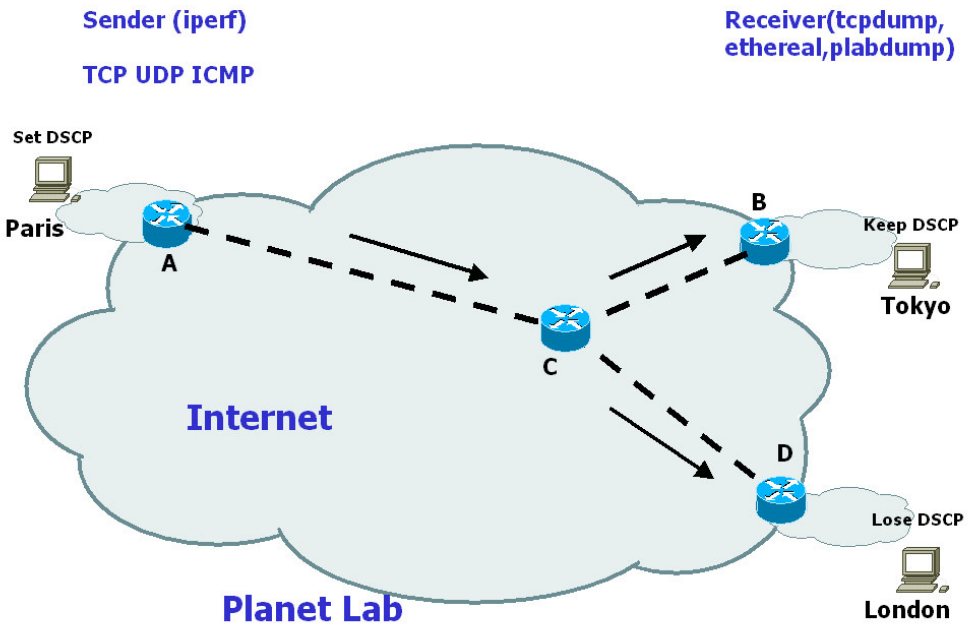
```
⊟ Internet Protocol, Src Addr: 171.64.64.216 (171.64.64.216), Dst Addr: 128.59.150.9 (128.59.150.9)
    Version: 4
    Header length: 20 bytes
  ⊟ Differentiated Services Field: 0x20 (DSCP 0x08: Class Selector 1; ECN: 0x00)
      0010 00.. = Differentiated Services Codepoint: Class Selector 1 (0x08)
      .... ..0. = ECN-Capable Transport (ECT): 0
      .... ...0 = ECN-CE: 0
    Total Length: 60
    Identification: 0x6f1c (28444)
```

In this special case , the packet was sent from a node of IP address 171.64.64.216 (planetlab-1.stanford.edu) , to IP address 128.59.150.9 , a node of Columbia University, and I seized this packet use one of the probes , ethereal ,(I could use other tools like Tcpdump, plabdump too get this packet too), and print out the whole packet . At the sender (Stanford node), I create the packet using iperf, and set the DSCP field as 0x20 (Class Selector 1), and from the analyst of the packet at the probe, we see the DSCP bits are kept in the header . For I set the probe at the receiver side, in Columbia campus in this case, means there is no router in the path strips DSCP bits or zero them.

I gathered a list of web-sites to a number of sites including Universities and some companies around the world and performed the above procedure . **In my more than one hundred runs , about 20% to 30% packets observed lost their DSCP at receiver traveling on the Internet.**

2.2. Discusson

Among my tests, in some cases , the DSCP bits did disappeared .Then I tried to find out at which point of the path , those bits changed to zero. I created another DSCP packet from the same sender which we already know lost its DSCP in the previous test , sent the packet to another destination (another probe). It may also lost the DSCP at receiver too . Or, the result could be one receiver get DSCP and one lost it, as in graph 1 :

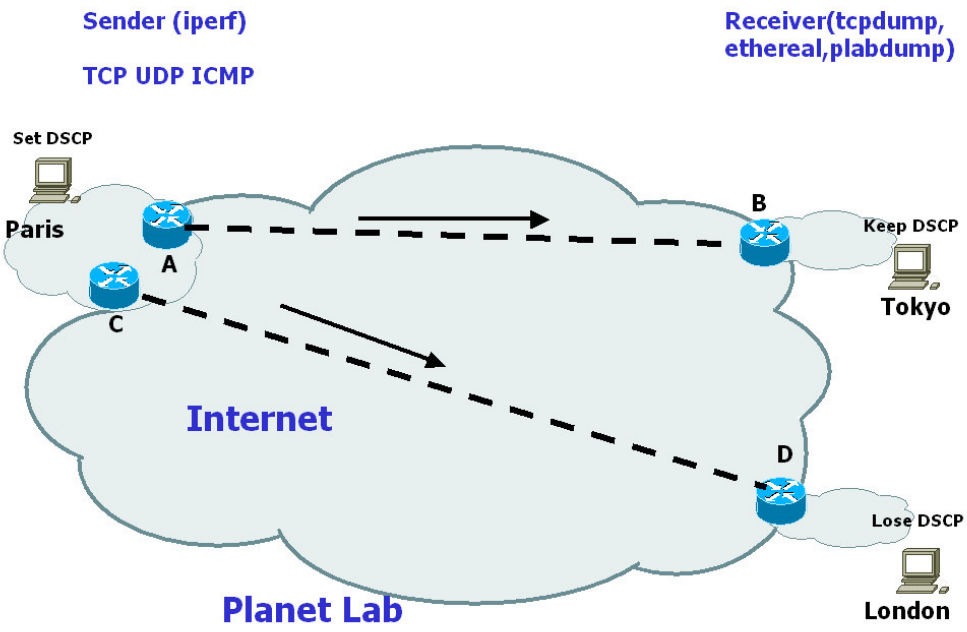


Graph 1

I sent packets from Paris to Tokyo first and to London for the second try. Tokyo received the packet with its DSCP kept and London got all zero in packet's DSCP field . For find out the point which at the packet lost its DSCP , I used the "traceroute" application to track the routers on the path, from those three nodes to each other , the DSCP must lost between the path of router C and D . Later, I use London, which lost the DSCP when acted as receiver , to sent packets out , to other nodes . If all the packets received at the other side of the Internet lost DSCP bits, then it is most likely the DSCP lost at router D, which is the London's access router.

From my test, about 20% packets lost DSCP at the local access router , in the other word, the packet carrying DSCP sent out from that kind of network can not be received with DSCP successfully in any receiver in the Internet.

Another case is the host in the **multihome network** , means the local network has two Internet connections , as in graph 2:



Graph 2

DSCP packets were sent from Paris to Tokyo first and to London for the second try. Tokyo received the packet with its DSCP kept and London got all zero in packet's DSCP field . From the traceroute , the results tells all packets come out from route A keep the DSCP and all packets go through router C lost it . And if let Tokyo or London act as sender , send DSCP packet to probes, will keep DSCP bits . So, it is most likely, router A support DSCP, while router C does not trust it at all .

It is the case of Columbia campus network . By discussing with out **AcIS (Academic Information Systems)** staff , I get to know Columbia campus has two gateways to Internet, one is Columbia commodity Internet gateway (nyser-gw) which is Cisco 6509 , the other is the Internet2 gateway (nn2k-gw) using Cisco 7500/RSP4 . The observation is all packets go through (nyser-gw) lost their DSCP , while nn2k-gw keep the DSCP . This clearing of the DSCP bits since packets on 100Mbps links are not trusted so the resetting occurs at the 100Mbps interface on the Cisco 6509 Sup2/MSFC2 when the packet in or out to Columbia campus.

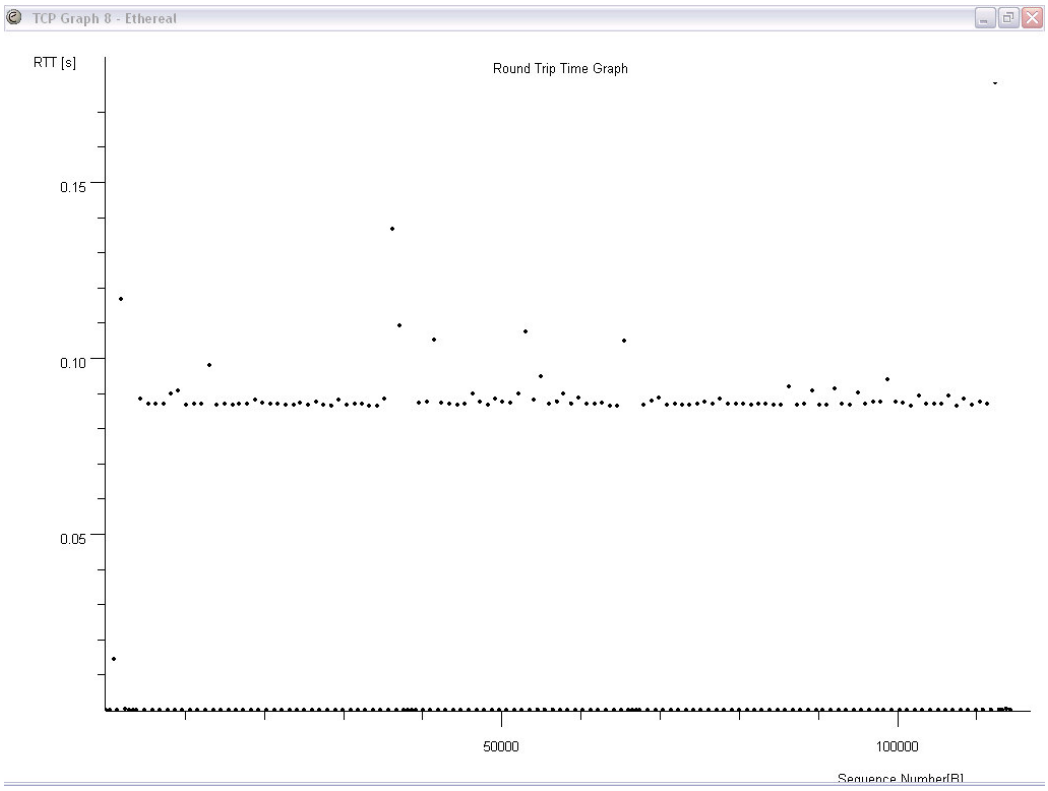
That give the rise of all packets go through Internet2 which from most university keep their DSCP , while all packet transmit from companies on the commercial Internet lost DSCP, change to all zero. This multihome environment makes the packets go out or come in to Columbia campus have about 50 percentage possibility to lost their DSCPs , base on what path the packet select to use .

During my tests, I met another case of this kind of multihome network at lanetlab1.ashburn.equinix.planet-lab.org 65.77.223.82 Equinix – Ashburn , from the observation of “traceroute”, I found all traffic out from one access gateway lost the DSCP and from another gateway keep it .

2.3. Effects on DSCP transmit time

From some recently researches and papers published on the Internet, I get to know that packets carrying IPv4 options are both delayed to a greater extent and dropped more often than packets without options. But will packets with DSCPs get more delay or not?

To find out if packet with DSCP set get more delay or not , I used etheareal at testing nodes to capture TCP session for 10 seconds , for TCP session with DSCP set and normal TCP session, and compare the RTT of those two sessions .



From the statistics of RTT for ten seconds TCP transmit , there is seems no difference between the packet with DSCP set to the normal TCP packets . So in most case , the back bone routers probably just ignore DSCP bits , for normally , any packet requiring extra work ends up in the slow path .

2.4. Results

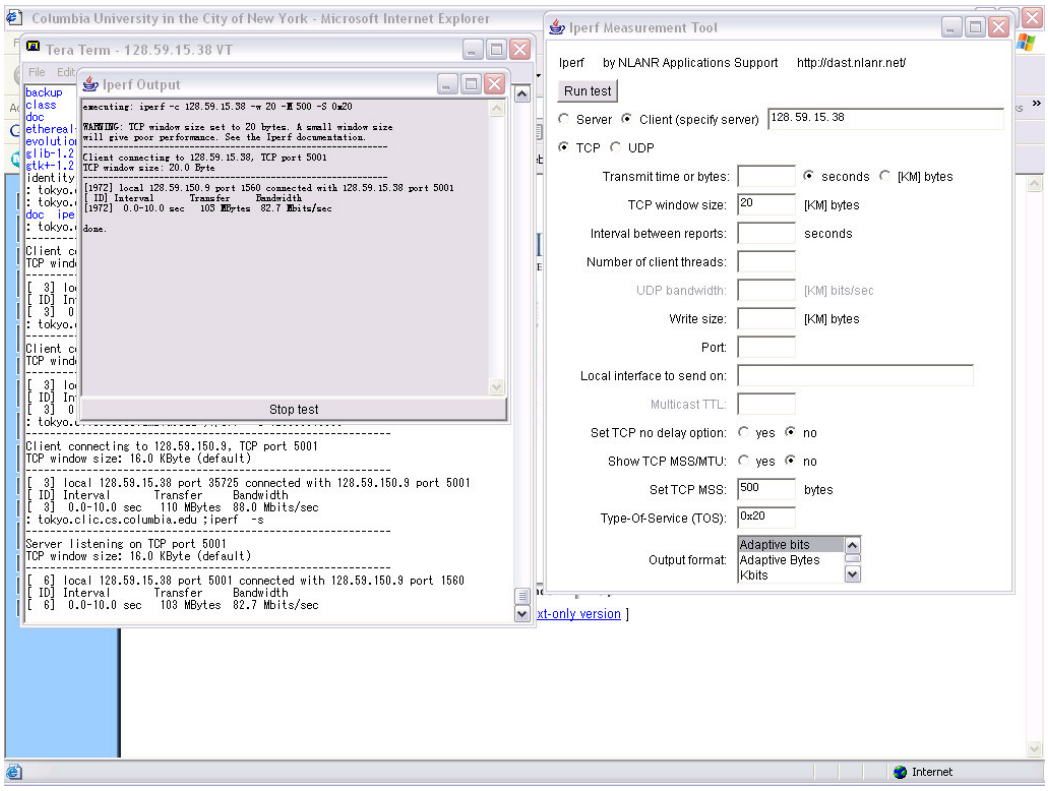
By sending packet from 30 senders to 5 probes respectively , the following results were achieved:

Mostly , at the probe of the receiver, I could get the DSCP unchanged , and the same RTT of non-DSCP packet , means the ISP backbone routers just forward the DSCP set packet without any additional process.

In some cases , less than one third of all tests, the access routers (or layer 3 switches like Cisco 6509) of the network do not trust DSCP, could cause the DSCP lost at the edge of the Internet. This kind of situation include only one access router and multihome cases .

In the case of compare the transmit time of packet carry DSCP with normal data packet, I used iperf to send TCP packets with different DSCP , the average RTT was no difference from the output of RTT graph of ethereal , regardless of whether the DSCP bits were not set or set . So the SP routers normally just ignore the DSCP bits.

Below is the graph of using iperf to create special Ipv4 packet and send to the probe on the Internet.



Use Iperf to create a IP packet carrying DSCP bits

3. RELATED WORK

A ping delay with Ipv4 header options measurement has been done by Prof. Michael Welzl [5] , he used RTT to measure the delay . Another measurement of the Ipv4 header options delay has been done by Pierre Fransson [6], using ICMP header , and the solution is Ipv4 options do cause more delay in the Internet. There does not seem to exist any substantial amount of measurements on the support of DSCP field however , so I did this series of experiments base on the measurement focus on the field of Differentiated Services.

4. CONCLUSIONS AND FUTURE WORK

I have gathered data from 150 different runs, by sending packet from 30 random selected nodes to 5 probes . Out of these, I observed some packets carrying DSCP bits are in fact lose them during the end to end transmission on the Internet, about 20% to 30% from all DSCP set packets, depends on the node selection. However, the round trip time seems no difference between the DSCP set TCP connection and otherwise. Seems the router either zero the DSCP filed, which mean do not trust it, or the router just ignore it.

Finally it is worth noting that when the host using private IP address , which means it need to pass SP's NAT translate router to get on the Internet, the DSCP bits get lost.
The next step , I plan to use this test environment to observe other kind of packet,for example, I could set the IP or TCP options, and send the packet though the Internet , to find out whether ISP support other slightly-unusual IP and TCP features such as IP header options.

5. ACKNOWLEDGMENTS

Thanks to professor Henning Schulzrinne of Columbia University , this project is under his instruction . I greatly benefited from discussions with Weibin Zhao , who is my mentor of this measurement , he gave me many very good suggestions . This test was supported by the equipment of planetlab network services .

6. REFERENCE

[1] K. Nichols, S. Blake, F. Baker and D. Black . Definition of the Differentiated Services Field (DS Field). RFC 2474 , Internet Engineering Task Force , December 1998.
[2] S. Blake , D. Black , M. Carlson , E. Davies , Z. Wang and W. Weiss . An Architecture for Differentiated Services . RFC 2475 , Internet Engineering Task

Force , December 1998.
[3] IEPM. TOS: Type of Serveice . August 1999
[4] IEPM. SLAC QBSS Measurements . August 9, 2001.
[5] Michael Welzl . The impact of "Slow Path" processing . September 2002.
[6] Pierre Fransson and Andreas Jonsson . The Need for an Alternative to Ipv4-Options .
[7] Andre Broido , Ryan King, Evi Nemeth , and kc claffy . Radon spectroscopy of packet delay .