ARIZONA STATE UNIVERSITY

CSE 434, SLN 70516 - Computer Networks — Fall 2022

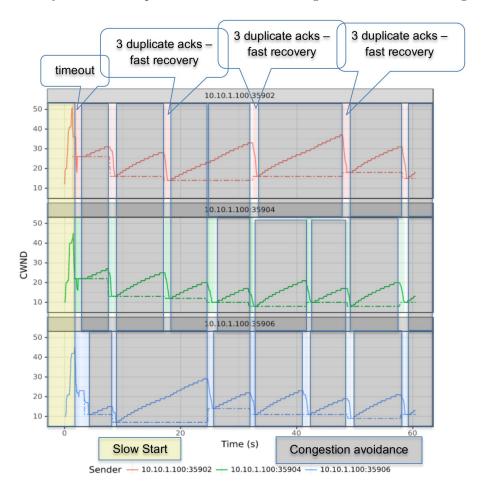
Lab #2 Solution

1 TCP Congestion Control

Exercise 1.1

Configure the topology as described in the "Run my experiment" section of the tutorial.

1. Include your annotated plot of the three flows sharing the bottleneck link using TCP Reno.



- 2. Using your plot and/or experiment data, explain the behaviour of TCP Reno in the slow start and congestion avoidance phases. In addition, explain what happens to both the congestion window and the slow start threshold when multiple duplicate ACKs are received.
 - Slow start periods, where the congestion window increases rapidly.
 - Congestion avoidance periods (when the congestion window increases linearly from than the slow start threshold)
 - Fast recovery, TCP Reno, will enter "fast recovery" in response to 3 duplicate ACKs. The slow start threshold is set to half of the CWND at the time of the loss event, the new CWND is set to the slow start threshold, and the flow enters "congestion avoidance" mode. In another words the receipt of triple duplicate ACKs indicates that some segments are getting through, and so

congestion may be less severe. Thus instead of wasting bandwidth, TCP tries a more aggressive strategy, by starting cwind at the current sthreshold value rather than at 1 MSS. This is $\frac{1}{2}$ the cwind value when the packet was dropped.

• slow start, Instances of ACK timeout, if any. This will cause the congestion window to go back to 1 MSS, and the flow enters "slow start" mode.

The behavior of TCP Reno during the slow start and congestion avoidance phases is depicted in the figure 1.

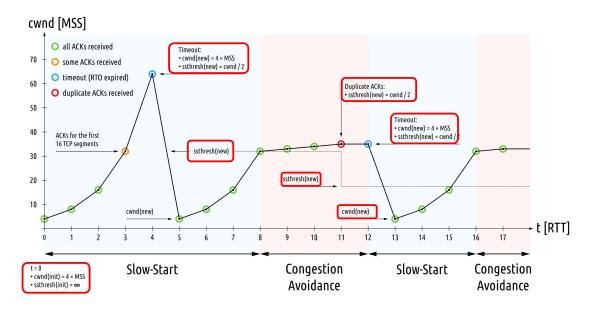


Figure 1: Image is GPLv3, via Wikimedia Commons.

Exercise 1.2

Under the "Optional: Other Congestion Control Algorithms" section of the tutorial, repeat the steps in the "Generating Data" section using the TCP Cubic variant. Download the updated sender-ss.svg image file to your computer. Study the TCP Cubic algorithm and use it to annotate your plot, as you did in Exercise 1.1 for TCP Reno.

- 1. Include your annotated plot of the three flows sharing the bottleneck link using TCP Cubic. Figure 2 shows the three flows sharing the bottleneck link. This figure could be different according to the network condition.
- 2. Using your plot and/or experiment data, explain the behaviour of TCP Cubic.
 - TCP CUBIC increases the size of cwind on the function of the cube between the current time and the future point of time when cwind will equal the value of cwind before the last loss event when operating in "congestion avoidance." So, while the current value of cwind is far from the old value of cwind, TCP CUBIC is rapidly increasing. TCP CUBIC gradually increases when the current value of cwind approaches the old value of cwind. After TCP cubic has successfully operated near (either slightly below or slightly above) the old value of cwind for a certain amount of time, it begins to rapidly increase to probe for a new operating point.
- 3. Why does TCP Cubic reach the available bandwidth faster than TCP Reno?

 The TCP Cubic modifies the linear window growth function of existing TCP standards to be a cubic function. It also achieves more equitable bandwidth allocations among flows with different RTTs (round trip times) by making the window growth to be independent of RTT

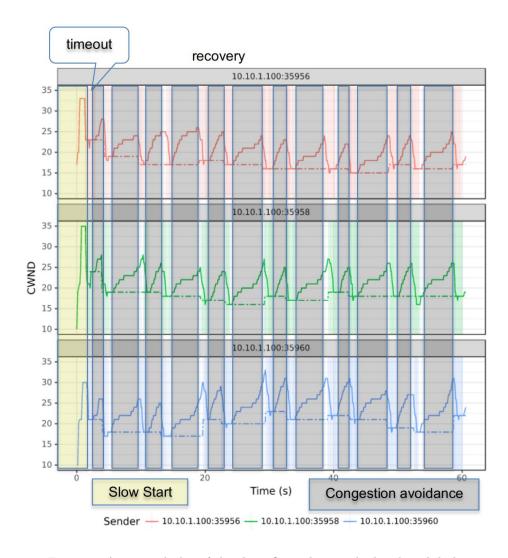


Figure 2: Annotated plot of the three flows sharing the bottleneck link

Exercise 1.3

Under "Optional: Low Delay Congestion Control" section of the tutorial, repeat the steps in the "Generating Data" section using the TCP Vegas variant. Make a note of the iperf3 throughput and the RTT estimated by ping during the TCP Vegas flow. TCP Vegas does not work well when it shares a bottleneck link with a flow using a loss-based algorithm, e.g., TCP Reno. To help understand the problem, you will send one TCP Reno flow and one TCP Vegas flow through the bottleneck router. For your lab report:

- 1. Make a note of the throughput reported by iperf3 for the TCP Reno flow and the TCP Vegas flow. Note: the numbers could be different depending on network condition. The bottleneck link is 1 Mbps. The purpose of this question is to demonstrate that TCP Vegas's performance is extremely poor.
 - TCP Reno, Throughput=909 Kbps
 - TCP Vegas, Throughput=60 kbps
- 2. Using your throughput measurements, and knowledge of the congestion control mechanisms used by each algorithm, can you explain why TCP Vegas does not work well in this scenario?
 - TCP Reno's congestion window grows rapidly without considering the delays in the network. TCP Vegas, on the other hand, determines congestion from delays and reduces its send rate accordingly. TCP Vegas is more concerned with avoiding packet losses.

Exercise 1.4

Under "Optional: Explicit Congestion Notification (ECN)" section of the tutorial, use active queue management to configure a queue in both directions on the router, and enable ECN in TCP on each host. Prepare to capture the TCP flow on each host, and then start the experiment. When the experiment finishes, transfer the packet captures (i.e., the .pcap files) to your computer using scp.

- 1. Compare the delay performance of Reno with ECN to that from the experiment (Exercise 1.3) showing the delay performance without ECN.
 - When ECN is enabled, the RTT and throughput both decrease. This reduction should be reflected in the answers.
- 2. Look for the ECN-related fields in the IP header and TCP header, during connection establishment and during data transfer. Annotate your packet capture files by drawing a circle or a box around the fields to show:
 - The ECN handshake, i.e., the ECN-setup SYN packet and corresponding ECN-setup SYN-ACK. It can be seen in figure 3.

```
10.10.2.100
       14 0.046932 10.10.2.100
                                       10.10.1.100
                                                                74 5201 → 37294
                                                                                   [SYN, ACK,
Flags: 0x0c2 (SYN, ECN, CWR)
                                                       Flags: 0x052 (SYN, ACK, ECN)
      .... = Reserved: Not set
                                                          000. .... = Reserved: Not set
   ...0 .... = Nonce: Not set
                                                          ...0 .... = Nonce: Not set
   .... 1... = Congestion Window Reduced (CWR): Set
                                                          .... 0... = Congestion Window Reduced (CWR): Not set
  .... .1.. .... = ECN-Echo: Set
                                                         .... .1.. .... = ECN-Echo: Set
   .... ..0. .... = Urgent: Not set
                                                          .... ..0. .... = Urgent: Not set
   .... ...0 .... = Acknowledgment: Not set
                                                         .... = Acknowledgment: Set
   .... .... 0... = Push: Not set
                                                          .... 0... = Push: Not set
                                                                  .0.. = Reset: Not set
   .... .... .0.. = Reset: Not set
 ..1. = Syn: Set
            ...0 = Fin: Not set
                                                                  ...0 = Fin: Not set
                    ECN Syn
                                                                           ECN Syn Ack
```

Figure 3: SYN and SYN-ACK packets

• Select a data packet and show the two ECN bits in the IP header set to either either 01 or 10. It can be seen in the figure 4.

```
> Frame 10: 168 bytes on wire (1344 bits), 80 bytes captured (640 bits)
> Ethernet II, Src: 02:57:5c:75:f7:26, Dst: 02:47:07:8a:a9:4f

V Internet Protocol Version 4, Src: 10.10.1.100 (10.10.1.100), Dst: 10.10.2.100 (10.10.2.100)
0100 ... = Version: 4
... 0101 = Header Length: 20 bytes (5)

V Differentiated Services Field: 0x02 (DSCP: CS0, ECN: ECT(0))
0000 00.. = Differentiated Services Codepoint: Default (0)
... ..10 = Explicit Congestion Notification: ECN-Capable Transport codepoint '10' (2)
```

Figure 4: Data packet with the two ECN bits in the IP header set to 10

• Find an instance of a "congestion experienced" signal, i.e., the two ECN bits in the IP header set to 11, and subsequent "congestion window reduced" (CWR) flag in the TCP header of the next packet.

First, the receiver gets the congestion-experienced signal from the router, Figure 5 (ECN = 11). Second, the receiver replies with an ACK where the ECN-Echo bit is set in the TCP Header. Third, the sender receives the echo bit and reduces their CWND, figure 6.

```
> Frame 762: 1514 bytes on wire (12112 bits), 80 bytes captured (640 bits)
> Ethernet II, Src: 02:57:5c:75:f7:26, Dst: 02:47:07:8a:a9:4f

Variable Protocol Version 4, Src: 10.10.1.100 (10.10.1.100), Dst: 10.10.2.100 (10.10.2.100)

0100 .... = Version: 4

.... 0101 = Header Length: 20 bytes (5)

V Differentiated Services Field: 0x03 (DSCP: CS0, ECN: CE)

0000 00... = Differentiated Services Codepoint: Default (0)

.... .11 = Explicit Congestion Notification: Congestion Experienced (3)
```

Figure 5: Congestion experienced

```
Flags: 0x090 (ACK, CWR)

000.... = Reserved: Not set

...0 .... = Nonce: Not set

....1 .... = Congestion Window Reduced (CWR): Set

....0 .... = ECN-Echo: Not set
```

Figure 6: When the sender receives the echo bit, reduce CWND.

2 Basic Home Gateway Services: DHCP, DNS, NAT

Exercise 2.1

In the section "Observe a DHCP request and response:"

1. Take a screen shot showing the DHCP discover message sent by the client to try and find DHCP servers on the LAN. What are the source and destination IP addresses in this request? Why are these addresses used?

```
| October | Octo
```

Source IP address: 0.0.0.0

Source MAC address: 02:18:ef:27:07:7b Destination IP address: 255.255.255.255.255 Destination MAC address: ff:ff:ff:ff:ff 2. Take a screen shot showing the DHCP offer sent by the server. What IP address does the server offer? What is the range of addresses that the server in our experiment may offer? (You can refer to the dnsmasq configuration file.)

The client is offered the IP address: 192.168.100.159. The client also learns the server's IP address is 192.168.100.1 and that the server should be the default gateway.

3. Take a screen shot showing the DHCP request sent by the client. What is the destination address in this request? Why?

In this example, the client was given only one offer. Therefore, it requested the IP address 192.168.100.159 from the server at 192.168.100.1.

4. Take a screen shot showing the DHCP acknowledgement sent by the server to complete the configuration.

```
③ Selectnicolean@gateway → □ ×

15:42:32.194989 02:70:69:96:22:8f > 02:18:ef:27:07:7b, ethertype IPv4 (0x0800), length 342: (tos 0xc0, ttl 64, id 25463, offset 0, flags [none], proto UDP (1.7), length 328)

192.168.100.1.67 > 192.168.100.159

Server-IP 192.168.100.15

Client-Ethernet-Address 02:18:ef:27:07:7b

Vendor-rfc1408 Extensions

Magic Cookie 0x63825363

DHCP-Message Option 53, length 1: ACK

Server-ID option 54, length 4: 192.168.100.1

Lease-Time Option 51, length 4: 14400

RN Option 58, length 4: 72600

RB Option 59, length 4: 12600

Subnet-Mask Option 1, length 4: 255.255.255.0

BR Option 28, length 4: 12600

Subnet-Mask Option 1, length 4: 192.168.100.1

Default-Gateway Option 3, length 4: 192.168.100.1
```

5. Step 5: Take a screen shot showing the configuration of the client's eth1 interface after receipt of the DHCP acknowledgement. How can you tell what IP address the client is using? Does this correspond to the IP address in the request and in the acknowledgement?

```
incolean@client-1:-

nicolean@client-1:-

nicolean@client-1:-

sudo dhclient -d eth1

Internet Systems Consortium DHCP Client 4.2.4

Copyright 2004-2012 Internet Systems Consortium.

All rights reserved.

for info, please visit https://www.isc.org/software/dhcp/

Listening on LPF/eth1/02:18:ef:27:07:7b

Sending on Socket/fallback

DHCPDISCOVER on eth1 to 255.255.255.255 port 67 interval 3 (xid=0x43e45e0e)

DHCPREQUEST of 192.168.100.159 from 192.168.100.1

DHCPOFER of 192.168.100.159 from 192.168.100.1

DHCPACK of 192.168.100.159 from 192.168.100.1

bound to 192.168.100.159 -- renewal in 6121 seconds.

**Cnicolean@client-1:-

tink encap: thernet. HWaddr 02:18:ef:27:07:7b

inter addr: 102.168.100.159 looses of the condition of
```

The ifconfig eth1 command lets us see IP address associated with eth1. In this case, the client's eth1 is shown to be 192.168.100.159. This matches the IP address that was offered, requested, and acknowledged in the previous steps. In addition to receiving a new IP, the client received the default gateway, DNS nameserver, and subnet mask.

Exercise 2.2

In the section "Observe a DNS request and response:"

- 1. Sample answer to this part is the following:
 - (a) Hostname = website.Lab2Grier.ch-geni-net.instageni.colorado.edu
 - (b) Type = A
 - (c) Address = 192.12.245.168
 - (d) ns.emulab.net (155.98.32.70); ns.instageni.colorado.edu (192.12.245.132)
 - (e) 192.168.100.1 (gateway acts as DNS server)

```
client-1:- dig website.Lab2Grier.ch-geni-net.instageni.colorado.edu

; ; <<>> DiG 9.11.3-lubuntul.15-Ubuntu <<>> website.Lab2Grier.ch-geni-net.instageni.colorado.edu
;; global options: +cmd
;; got answer:
;; ->>>HEADER</- opcode: QUERY, status: NOERROR, id: 44555
;; flags: qr aa rd ra; QUERY: 1, ANSWER: 2, AUTHORITY: 2, ADDITIONAL: 2
;; OPT PSEUDOSECTION:
; EDMS: version: 0, flags:; udp: 4096
; COOKIE: aa81e3d79eb6a9ff6ofc157b617edlc63344a61f2624bd75 (good)
;; QUESTION SECTION:
;website.Lab2Grier.ch-geni-net.instageni.colorado.edu. IN A
;; ANSWER SECTION:
website.Lab2Grier.ch-geni-net.instageni.colorado.edu. 1 IN CNAME pcvm2-12.instageni.colorado.edu.
pcvm2-12.instageni.colorado.edu. 30 IN A 192.12.245.168
;; AUTHORITY SECTION:
instageni.colorado.edu. 30 IN NS ns.emulab.net.
instageni.colorado.edu. 30 IN NS ns.instageni.colorado.edu.
;; ADDITIONAL SECTION:
ns.instageni.colorado.edu. 30 IN A 192.12.245.132
;; Query time: 3 msec
;; SRVER: 192.168.100.1853(192.168.100.1)
;; WHEN: Sun Oct 31 11:26:33 MDT 2021
;; MSG SIZE rcvd: 208</pre>
```

```
tcpdump: listening on ethl, link-type EN10MB (Ethernet), capture size 262144 bytes

11:24:04.139420 IP (tos 0x0, ttl 64, id 34139, offset 0, flags [none], proto UDP (17), length 121)

192.168.100.109.37160 > 192.168.100.1.531 40984+ [lau] A? website.Lab2Grier.ch-geni-net.instageni.colorado.edu. (93)
```

Figure 7: Dig and TCP dump output

2. Sample answer to this part is the following:

For the hierarchical DNS resolution with +trace, take a screen shot of the dig command and its output. Draw a diagram showing how the hostname was resolved recursively, starting from the implied " at the end and moving toward the beginning.

```
| Second | S
```

Figure 8: Screen shot of the dig command and its output

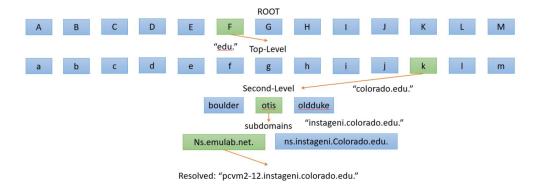


Figure 9: Sample diagram showing how the hostname was resolved recursively

Exercise 2.3

Follow the instructions in the section "Use NAT" set up the gateway to use NAT. For your lab report:

- 1. Take screen shots to show the three-way TCP handshake for a connection between client and website as seen by tcpdump at the website, and as seen by tcpdump at the gateway (on the LAN). Make sure you can see the IP addresses and port numbers used in the connection. sample answer would be similar to figure 10.
- 2. Draw a diagram showing how NAT is used between client and website, similar to this diagram but with the IP addresses, hostnames, and ports from your experiment.



Figure 10: 3-Way Handshake

The diagram should be similar to figure 11.

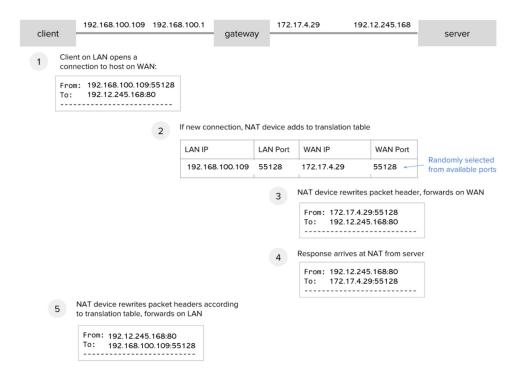


Figure 11: Nat diagram