COMP9121 Assignment 2 2024 S2

Request for review:

Please read this instruction before you request a remark.

(1) Please firstly read the solutions and compare your answer against them.

(2) If you have no concern on each individual question, but you think there is a mistake in calculation the total mark, or late penalty, or special consideration, please send us an email titled, [COMP9121, Assignment 2, total mark calculation] and explain your situation. Send the email to liming.ge@sydney.edu.au.

(3) If you have concerns on individual questions, you can ask for remarking of those questions. Please give us reasons for each of them. You need to let us know which questions you want to remark and why you think your mark should be changed for those questions. Please note that we will remark the questions you ask for, but we do not guarantee that the final mark is the same or higher (it may be lower). Please send us an email titled, [COMP9121, Assignment 2, remark Qx, Qx …] (x is the question number) and explain your situation. Send the email to liming.ge@sydney.edu.au.

(4) If you do not follow the right procedure, your request may be ignored or there could be substantial delay in processing your request.

(5) We may ask you for an interview if we find your case is complicated.

In this assignment, some questions are student number dependent, you will get zero in that question if you use another student's number.

There are 10 questions, each question is equally weighted. Questions 6-10 will be released later. The due date is 27 Oct 2024 at 23:59 and we will open the submission site on Canvas one week before the due date.

1. SDN.

In the following network, the network administrator should achieve the load balancing function: If the traffic is from 11.11.11.0/24 to 11.11.44.0/24, the blue path should be used; If the traffic is from 11.11.22.0/24 to 11.11.44.0/24, the red path should be used; If the traffic is from 11.11.33.0/24 to 11.11.44.0/24, the green path should be used;

(1) Can this load balancing function be achieved by traditional non-SDN routers? Why or why not?

(2) If SDN is used in the network, please specify the SDN switch table at w. The four interfaces (ports) of the SDN switch are labelled as I_1—I_4 respectively.

（1） Traditional non-SDN routers cannot perform load balancing because they forward packets based solely on the destination, always sending traffic to the same next hop. To achieve load balancing, routers would need to distribute traffic to different hops, even when packets have the same destination.

Q1.1: 0 mark: Wrong

Q1.1: 2 marks: Miss Explanation

Q1.1: 3 marks: Insufficient Explanation

Q1.1: 4 marks: Correct

（2）

Ingress Port	Source MAC	Dest MAC	Eth Type	VLAN ID	Source IP	Dest IP	IP Protocol	Source Port	Dest Port	Action
I_1	*	*	*	*	11.11.11.0/24	11.11.44.0/24	*	*	*	Forward(I_3)
I_1	*	*	*	*	11.11.22.0/24	11.11.44.0/24	*	*	*	Forward(I_4)
I_2	*	*	*	*	11.11.33.0/24	11.11.44.0./24	*	*	*	Forward(I_3)

Q1.2: 6 marks (2 marks for each row)

2. TCP

Consider the figure below. Assume TCP Reno is the protocol experiencing the behaviour, and the TCP session has a large number of packets to send. Answer the following questions (1)—(3). There is only one single TCP session in this figure.

(1) The congestion window size is decreased to 1 (segment) at round 10. Is it caused by three duplicate ACKs or timeout? Why?

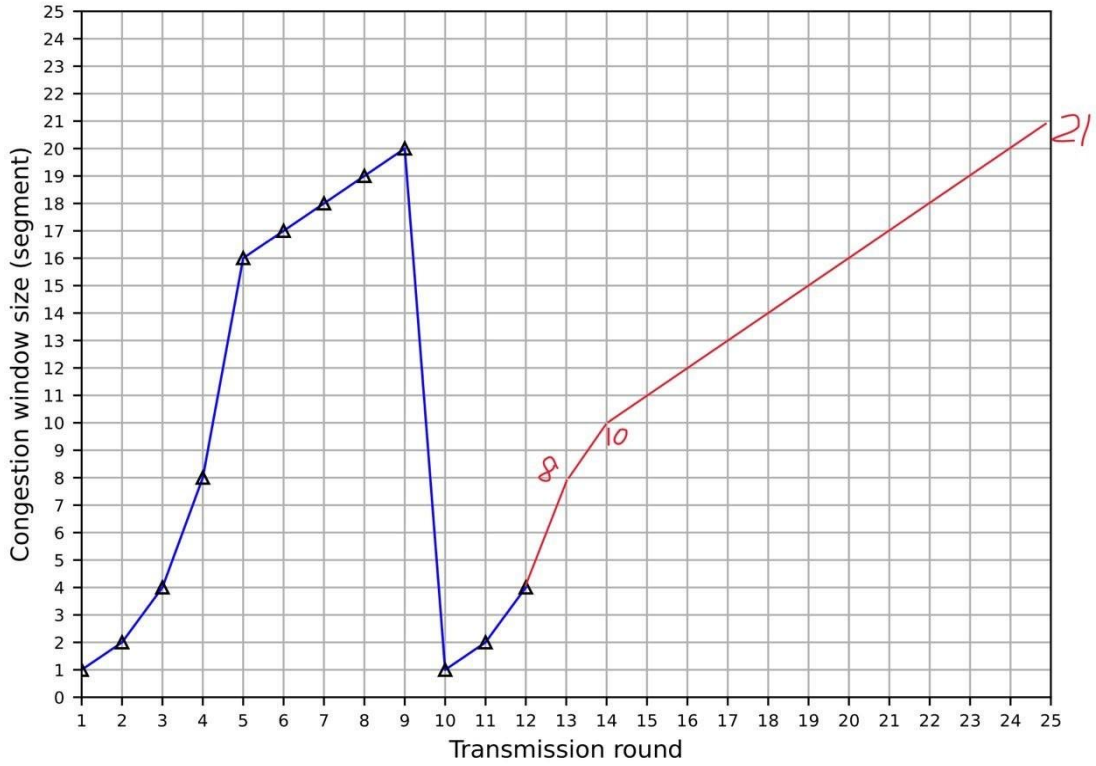
Timeout, timeout causes cwnd becoming 1. 3 duplicated ACK causes halved cwnd

(2) What is ssthresh at round 1? Why?

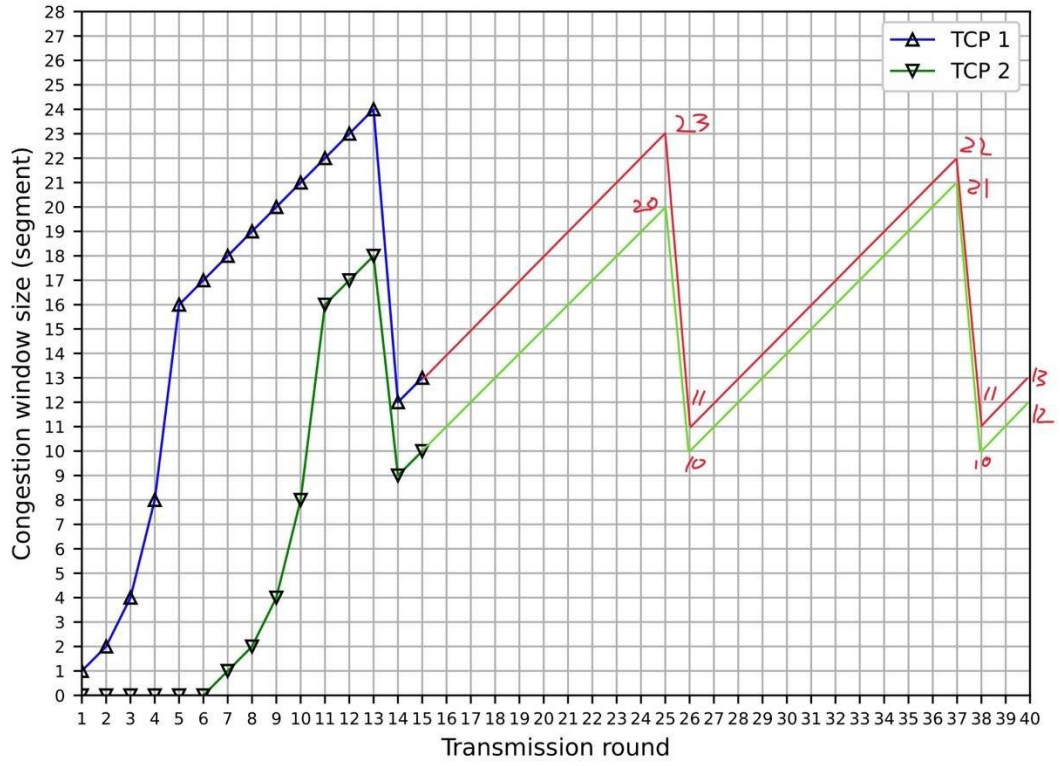
16 segments. As at cwnd=16 it turns from exponential to linear.

(3) Suppose there is no packet loss after round 10. In the figure, what are the congestion window sizes from round 12 to round 25?

Round	12	13	14	15	16	17	18	19	20	21	22	23	24	25
Windows Size	4	8	10	11	12	13	14	15	16	17	18	19	20	21



(4) In Figure below. Two TCP (Reno) sessions are sharing the same link and the link capacity is limited by 42 segments. Whenever the sum of the congestion window sizes reaches (or exceeds) 42, a packet is lost for each TCP session and the packet loss is detected by three duplicate ACKs. TCP session 2 starts later than TCP session 1. What are the congestion window sizes of the two sessions from round 15 to round 40? If the congestion window size is not an integer (in segment), you need to round it down.



3. Estimated RTT.

Consider the TCP procedure for estimating RTT. Let $EstimatedRTT_0$ = 100 ms be the estimated RTT when a TCP is initialised. Then, the TCP sender receives 5 ACKs and sample RTTs are measured as $SampleRTT_1$, $SampleRTT_2$, $SampleRTT_3$, $SampleRTT_4$, and $SampleRTT_5$. All of them are 110 ms. Let $EstimatedRTT_i$ denote the estimated RTT right after the ith ACK. We assume $\alpha = 0.125$ in this question.

(1) Calculate $EstimatedRTT_4$ and $EstimatedRTT_5$.

(2) Generalise your solution to n sample RTTs. The TCP sender receives n ACKs, with ith sample RTT $SampleRTT_i$. We assume all $SampleRTT_i$ are 110 ms. Express $EstimatedRTT_n$ as a function of n.

(3) For the formula in part (2), let n approach infinity. What is $EstimatedRTT_n$? Comment on why this averaging procedure is called an exponential moving average.

(1) 4 marks

$$EstimatedRTT_4 = (1-\alpha)\,EstimatedRTT_3 + \alpha\,SampleRTT_4$$

$$= (1-\alpha)((1-\alpha)\,EstimatedRTT_2 + \alpha\,SampleRTT_3) + \alpha\,SampleRTT_4$$

$$= (1-\alpha)^2\,EstimatedRTT_2 + (1-\alpha)\alpha\,SampleRTT_3 + \alpha\,SampleRTT_4$$

$$= (1-\alpha)^2\big((1-\alpha)\,EstimatedRTT_1 + \alpha\,SampleRTT_2\big) + (1-\alpha)\alpha\,SampleRTT_3 + \alpha\,SampleRTT_4$$

$$= (1-\alpha)^3\,EstimatedRTT_1 + \alpha(1-\alpha)^2 SampleRTT_2 + (1-\alpha)\alpha\,SampleRTT_3 + \alpha\,SampleRTT_4$$

$$= (1-\alpha)^3\big((1-\alpha)\,EstimatedRTT_0 + \alpha\,SampleRTT_1\big) + \alpha(1-\alpha)^2 SampleRTT_2 + (1-\alpha)\alpha\,SampleRTT_3 + \alpha\,SampleRTT_4$$

$$= (1-\alpha)^4\,EstimatedRTT_0 + (1-\alpha)^3\alpha\,SampleRTT_1 + (1-\alpha)^2\alpha\,SampleRTT_2 + (1-\alpha)\alpha\,SampleRTT_3 + \alpha\,SampleRTT_4$$

$$= (1-\alpha)^4\,EstimatedRTT_0 + \sum_{i=1}^{4} \alpha(1-\alpha)^{4-i}\,SampleRTT_i$$

$$= (1-\alpha)^4 100 + \sum_{i=1}^{4} \alpha(1-\alpha)^{4-i} 110 \text{ (substituted value of } EstimatedRTT_0 \text{ and } SampleRTT)$$

$$= (1-\alpha)^4 100 + (1-(1-\alpha)^4)110 \text{ (due to geometric series)}$$

$$\approx 104.138 \text{ ms (2 marks)}$$

Similar calculation of $EstimatedRTT_5$

$$EstimatedRTT_5 = (1-\alpha)^5 100 + (1-(1-\alpha)^5)110 \approx 104.871\ ms \text{ (2 marks)}$$

(2) 3 marks

$EstimatedRTT_n = (1-\alpha)^n 100 + (1-(1-\alpha)^n)110$

(3) 3 marks

When all n approaches infinity, $(1-\alpha)^n$ is approaching to 0 and $EstimatedRTT_n$ is approaching to 110 (2 marks), which is SampleRTT. Therefore, it will approach to SampleRTT in a exponential function (1 mark).

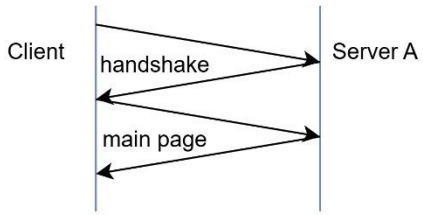
4. **HTTP and TCP.**

Client C requests a webpage from Server A. We assume that RTT between Client C and Server A is 10 ms. After obtaining the main page, Client C finds that there are 2 objects to be fetched. Both objects are also in Server A.

Client C starts to request for object 1 when the main page has been successfully downloaded. It starts to request for object 2 when object 1 has been successfully downloaded. The size of the main web page is small. It fits into 1 TCP segment. Each object fits into *n* TCP segments, where *n* is the last 3 digits of your student number plus 1000. For example, if you last three digits are 123, then *n* = 1123. We have ssthresh=64 segments at the beginning of a TCP session. *Non-persistent HTTP is used*. TCP termination delay is ignored. There is no packet loss.

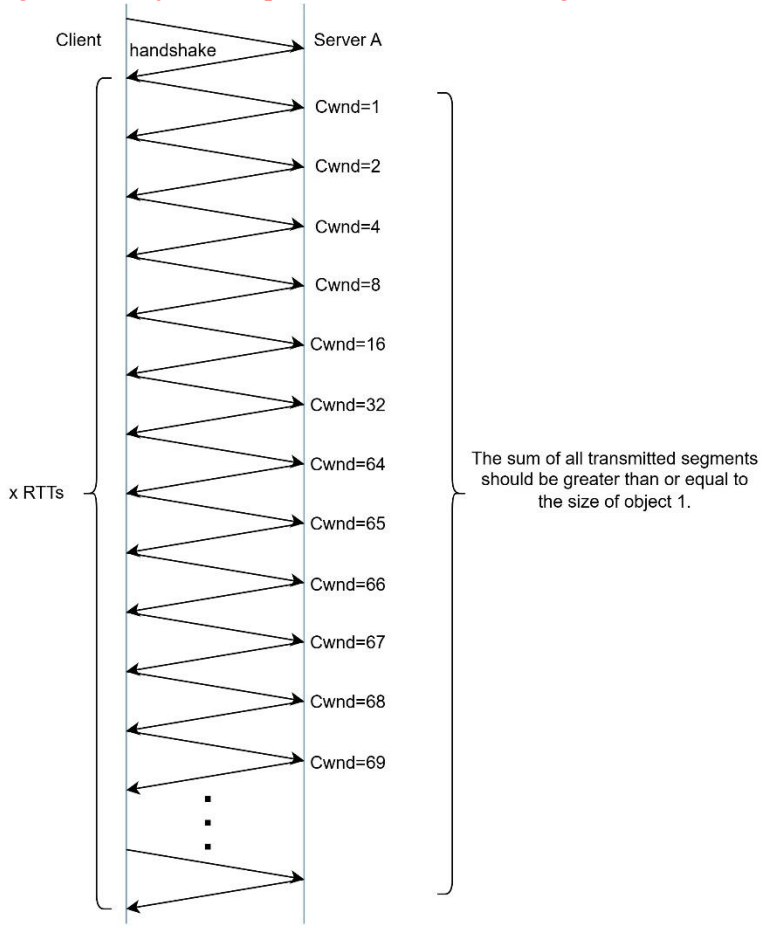
How long in total does it take for Client C to successfully obtain the webpage (including the main page and two objects).

For main page: total 2 RTT, 1 RTT for handshake, 1 RTT for main page. 3 marks



For object 1: the transmission requires 1 RTT for handshake and an additional x RTTs to complete the transmission. The value of x, representing the minimum number of rounds needed to transmit the final

segment of object 1, depends on the last three digits of the SID.



SID	x (number of RTTs transmitting the object)
000-050	20
051-128	21
129-207	22
208-287	23
288-368	24
369-450	25
451-533	26
534-617	27
618-702	28
703-788	29
789-875	30
876-963	31
964-999	32

3 marks

For object 2: since it has the same number of segments as object 1, the required RTTs will match those for object 1—1 RTT for the handshake and x RTTs for transmission.

Total time required for the main page, object 1 and object 2:

SID	Total time
000-050	$(2 + (1 + 20) \times 2) \times 10 = 440\ ms$
051-128	$(2 + (1 + 21) \times 2) \times 10 = 460\ ms$
129-207	$(2 + (1 + 22) \times 2) \times 10 = 480\ ms$

208-287	$(2 + (1 + 23) \times 2) \times 10 = 500\ ms$
288-368	$(2 + (1 + 24) \times 2) \times 10 = 520\ ms$
369-450	$(2 + (1 + 25) \times 2) \times 10 = 540\ ms$
451-533	$(2 + (1 + 26) \times 2) \times 10 = 560\ ms$
534-617	$(2 + (1 + 27) \times 2) \times 10 = 580\ ms$
618-702	$(2 + (1 + 28) \times 2) \times 10 = 600\ ms$
703-788	$(2 + (1 + 29) \times 2) \times 10 = 620\ ms$
789-875	$(2 + (1 + 30) \times 2) \times 10 = 640\ ms$
876-963	$(2 + (1 + 31) \times 2) \times 10 = 660\ ms$
964-999	$(2 + (1 + 32) \times 2) \times 10 = 680\ ms$

4 marks

Q5. Wireshark (NAT).

(1) Use the trace files and investigate them carefully. Please write down the NAT translation table entries in the NAT router for the traffic you can find in Q5_home_side and Q5_ISP_side. Find as many entries as possible. For each entry, please give screenshots to justify your answer. (If your table has more than 5 entries, just give the screenshots for the first 5 entries in your table.) Please note that some other computers are also using the local network, and their traffic may also appear in Q5_ISP_side but not in Q5_home_side. You should ignore them when you figure out the NAT translation table.

Http:

Private (LAN)		Public (WAN)			Marks (5 screenshots)
192.168.137.20	55645	10.66.30.63	63677	A1.html	1 mark: row 0.5+ screenshot 0.5
192.168.137.20	55646	10.66.30.63	63678	object of A1.html	1 mark
192.168.137.20	55647	10.66.30.63	63679	object of A1.html	1 mark
192.168.137.20	55648	10.66.30.63	63686	A3.html	1 mark
192.168.137.20	55650	10.66.30.63	63688	object of A3.html	1 mark
192.168.137.20	55651	10.66.30.63	63689	object of A3.html	1 mark
192.168.137.20	55652	10.66.30.63	63694	A2.html	1 mark

TCP

Private (LAN)	Public (WAN)		Marks ()

192.168.137.20	55273	10.66.30.63	63111		
192.168.137.20	55624	10.66.30.63	63662		
192.168.137.20	55625	10.66.30.63	63623		

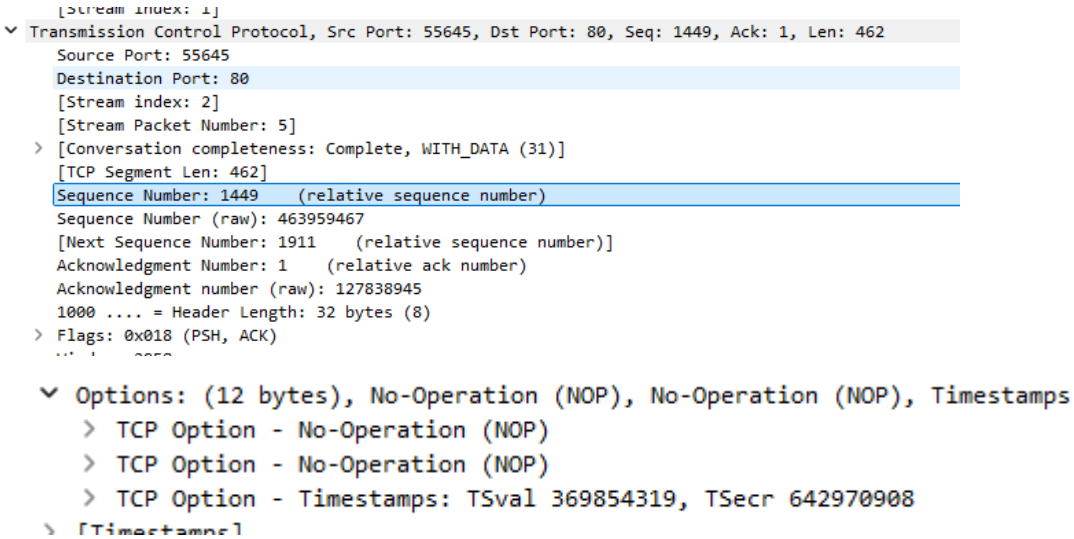
must 5 http + 2 any

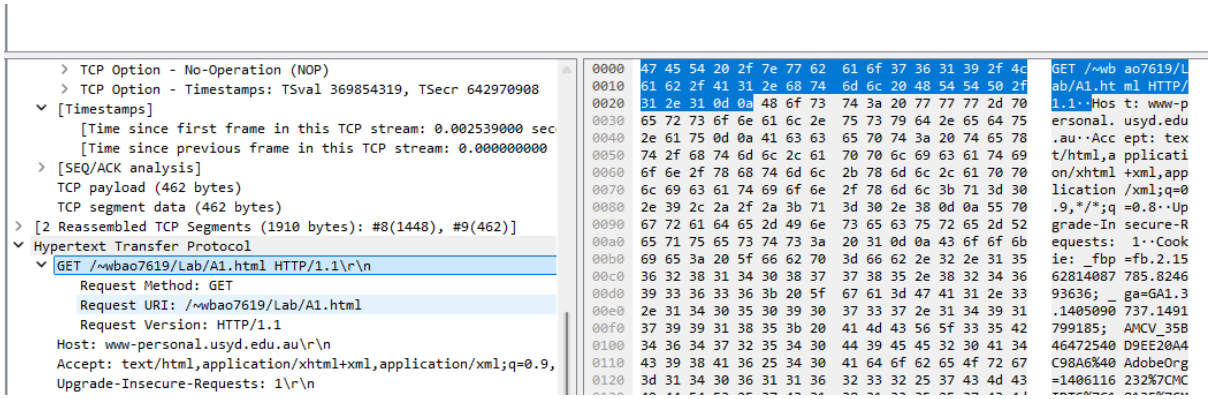
5 points for screenshots. In the screenshots, you need to give clear evidence showing that the payloads are identical. Examples include: absolute sequence number/ACK number in TCP header, timestamp in TCP header (optional field if possible), all bits in the HTTP payload are the same. You will lose mark if you show limited information which cannot guarantee that the payloads are identical. For example, only showing GET /~wbao7619/Lab/A1.html is not enough as multiple connections may have GET /~wbao7619/Lab/A1.html. Two examples are listed below.

Correct example: GET /~wbao7619/Lab/A1.html

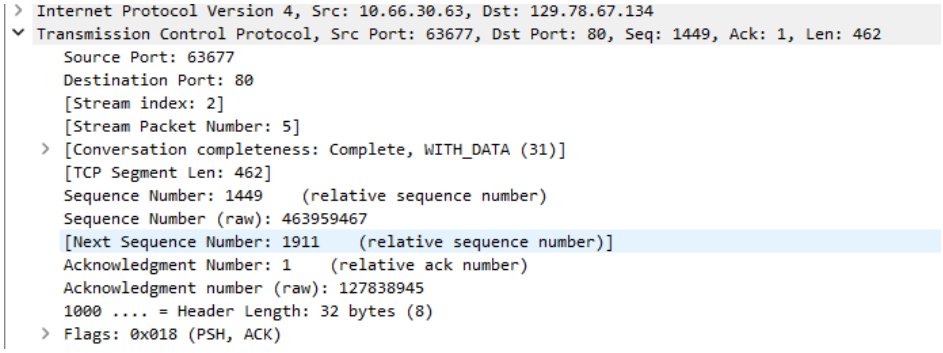
Display the sequence and acknowledgment numbers (absolute sequence numbers/ACK numbers), as well as the timestamp field in the TCP header (if the optional field is available). Additionally, ensure all bits in the HTTP payload are identical.

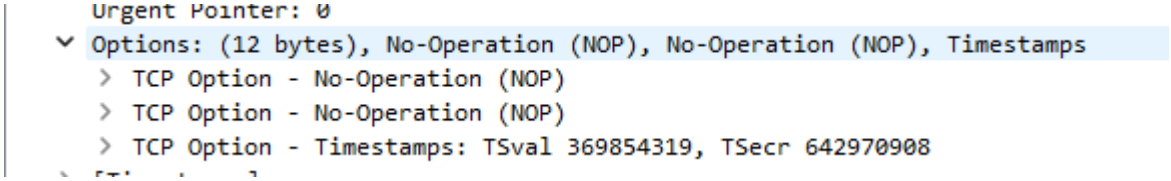
Private side:

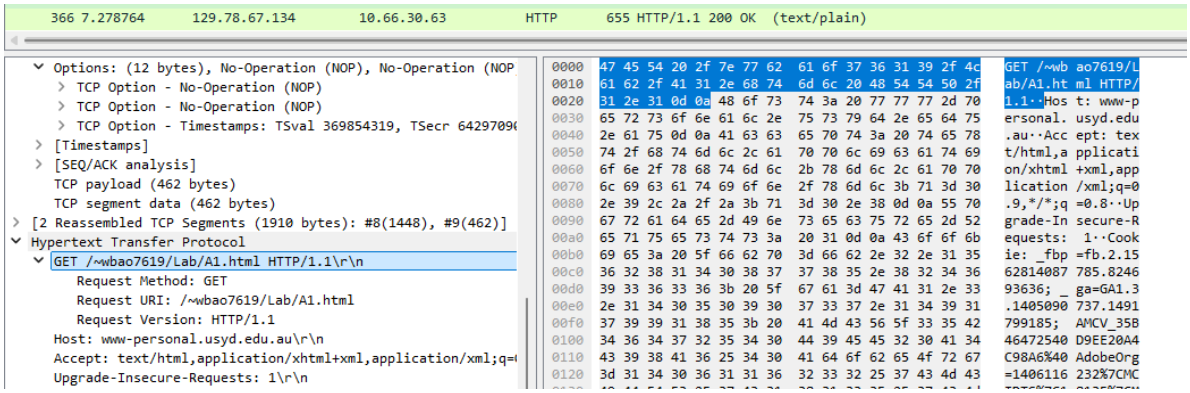




Public side:

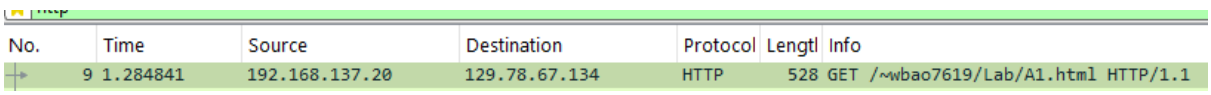




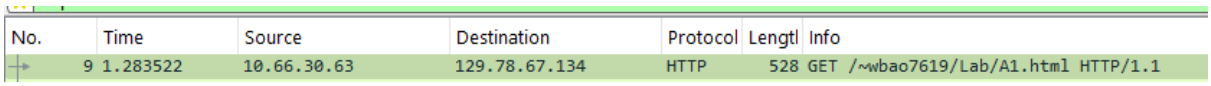


Wrong Example: GET /~wbao7619/Lab/A1.html

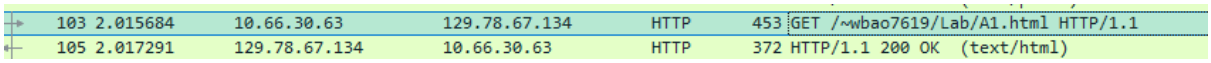
Private side:



Public side:



However, this is insufficient, as the Q5_ISP_side.pcapng file contains an additional GET /~wbao7619/Lab/A1.html request. Displaying only the GET /~wbao7619/Lab/A1.html request alone is not enough to determine whether they are identical.



(2) What is the MAC address of the client PC? What are the MAC addresses at the two interfaces of the NAT router? Please give you screenshots.

MAC addresses of client PC: 3c:15:c2:db:0c:30 1 mark

Interface of MAC address towards WAN: 54:bf:64:a1:ab:8c 1 mark

Interface of MAC address towards LAN: 3a:00:25:f9:68:5c 1 mark

6. P2P. A server distributes a file with the size of 10^9 bytes to n hosts. The upload rate of the server is 3.5 Mbps. The i-th host (i = 1,2,…,n) has a download rate of 0.5i Mbps. For all hosts, the upload rate is 1 Mbps. Calculate the minimum distribution time as a function of n for P2P distribution and client-server distribution, and plot the two curves. (n = 1,2,…,100)

P2P distribution: (3%)

$$D_{p2p} \geq max\ (F/u_s, F/d_{min}, nF/(u_s + \sum_{i=1}^{n} u_i))$$

$$D_{p2p} \geq max\ ((10^9 * 8)/(3.5 * 10^6), (10^9 * 8)/(0.5 * 10^6), (10^9 * 8 * n)/(\left(3.5 + \sum_{i=1}^{n} u_i\right) * 10^6))$$

$$D_{p2p} \geq max\ (2285.714, 16000, (10^3 * 8 * n)/(\left(3.5 + \sum_{i=1}^{n} u_i\right)))$$

Client-server distribution: (3%)

$$D_{c-s} \geq max\ (nF/u_s, F/d_{min})$$

$$D_{c-s} \geq max\ ((10^9 * 8 * n)/(3.5 * 10^6), (10^9 * 8)/(0.5 * 10^6))$$

$$D_{c-s} \geq max\ ((10^9 * 8 * n)/(3.5 * 10^6), 16000)$$

We can see the results as follows: x (number of users) y (number of 10^3 seconds), blue: P2P, orange: C-S. When the number of users >7, P2P shows performance gain compared with C-S

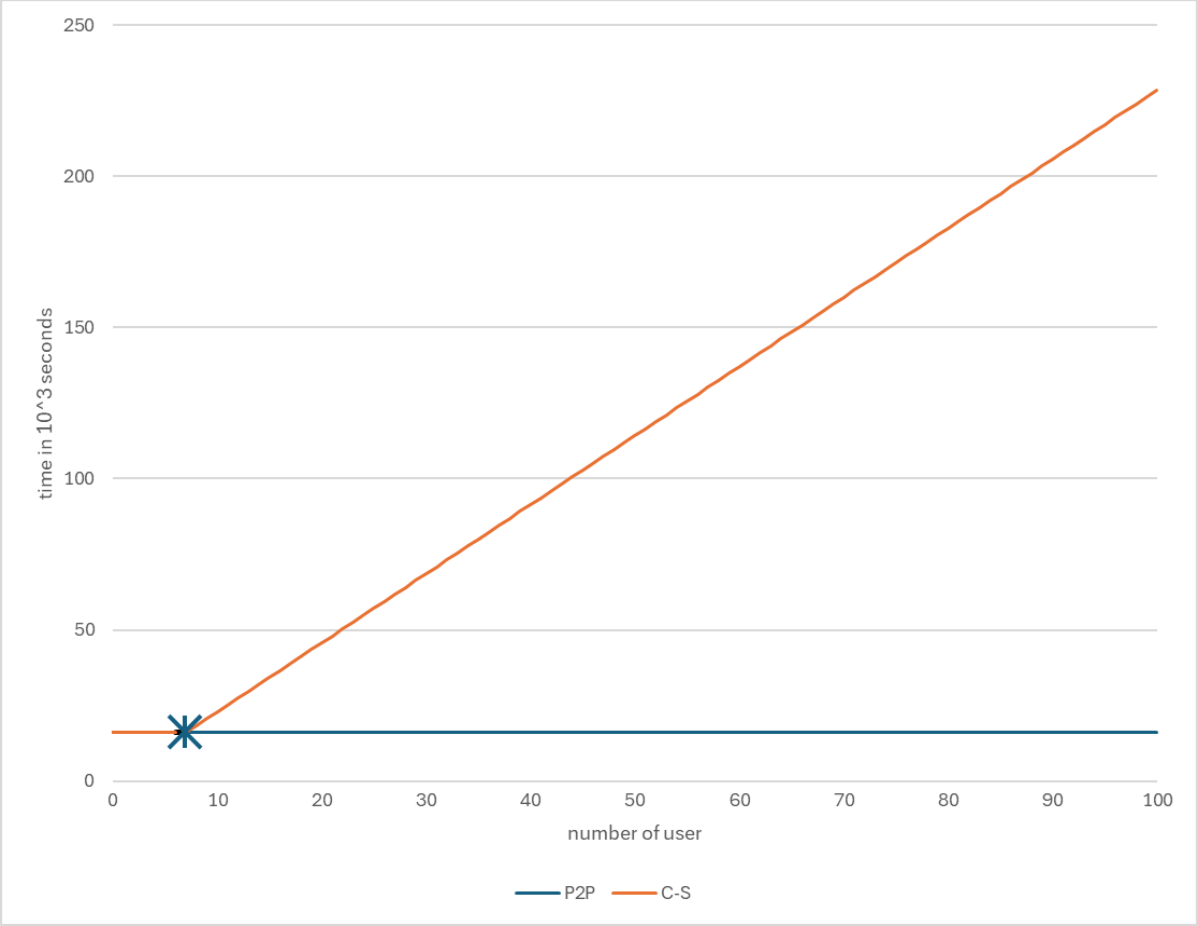


Figure: 4%

7. Cross layer: Routing, UDP/TCP, DNS/HTTP

Client-LC one way delay:

Client-A-B-LC = 4 ms

LC-Client one way delay:

4 ms, reverse of Client-LC

LC-Root one way delay:

LC- B -E-AS4-AS5-Root = 47 ms

LC-B-E-AS2-AS3-AS5-Root is not correct, as it contains more inter-AS hops.

Root-LC one way delay:

47 ms, reverse of Root-LC

LC-TLD one way delay:

LC-B-E-AS4-TLD = 42 ms

(LC-B-D-AS6-TLD) is not correct as both AS4 and AS6 has one AS hop and we should use hot potato to exit at E

TLD-LC one way delay:

42 ms, reverse of TLD-LC

LC-AA one way delay:

LC-B-D-C-AS10-AA = 36 ms

AA-LC one way delay:

36 ms, reverse of LC-AA

Client-Web Server A one way delay:

Client-A-C-Web Server A = 33 ms

Web Server A-Client one way delay:

33 ms, reverse of Client-Web Server A

Client-Web Server B one way delay:

Client-A-C-Web Server B= 33 ms

B Web Server-Client one way delay:

33 ms, reverse of Client-Web Server B

Overall Delay DNS for Web A = Client-LC+LC-Root+Root-LC+LC-TLD+TLD-LC+LC-AA+AA-LC+LC-Client

= (4+47+42+36)*2 = 258

Overall delay for getting Main page: One RTT for handshake (Client-WebA+ WebA-Client) + One RTT for mainpage (Client-WebA+ WebA-Client)

= 2*2*33=132

Overall delay for getting object 1 in WebA: One RTT for object 1 (Client-WebA+ WebA-Client)

=2*33=66

Second DNS for Web B = Client-LC+LC-Root+Root-LC+LC-TLD+TLD-LC+LC-AA+AA-LC+LC-Client

= (4+47+42+36)*2 = 258

Overall delay for getting object 2 in WebB: One RTT for handshake(Client-WebB+ WebB-Client) + One RTT for object 2 (Client-WebB+ WebB-Client)

= 2*2*33=132

Total delay = 258+132+66+258+132=846

If the DNS time delay for Web Server A is correct, 2 points.

If the DNS time delay for Web Server B is correct, 2 points.

If the time delay to obtain the main page is correct, 2 points.

If the time delay to obtain Object A is correct, 1 point.

If the time delay to obtain Object B is correct, 2 points.

If the final time delay calculation is correct, 1 point.

Q8. RSA. Use RSA algorithm to answer the following questions. You should use decimal numbers instead of binary numbers in this question. Correct solution is not unique.

(1) Select proper values of p and q, larger than 10 but less than 20. Then, calculate n and z.

(2) Select proper values of e and d. Then, encrypt the last two digits of your student number.

(3) Decrypt your answer for Part (2), and verify if the last two digits of your student number can be recovered.

(4) Trudy (the intruder), eavesdropped on the encrypted number and the public key (n, e), how can she decrypt the number in this example? In reality, if p and q value are very large numbers, and Trudy eavesdropped on the encrypted number and the public key, can she still decrypt the number? Why or why not?

(1) p and q are the prime number between 10 and 20. We need to check whether the student's p and q are prime and ensure they are between 10 and 20.
For example:
let p=11 q=13
n = p*q = 143
z=(p-1)(q-1)=(11-1)(13-1)= 120
Correct p 1 mark, correct q 1 mark, correct n 1 mark, correct z 1 mark

(2) e should be less than n, and e and z must be relatively prime. We need to check whether the student's e and z are relatively prime and ensure that e is less than n.
Select d such that ed-1 is divisible by z.
Use the last two digits of your SID as m
Encrypt the message using c = m^e mod n

For example:
e = 7
z = 120
d= 103
m = 60
c = 60^7 mod 143 = 135
Correct e 1 mark, correct d 1 mark

Correct encrypt message 1 mark

(3) To decrypt. m = c^d mod n. The decrypted m should be the same as the original m. We need to decrypt the student's message and check if it matches the last two digits of the student's SID.

For example:

m = 135^103 mod 143 = 60

Give 1 mark if the decrypted message matches the original message.

(4) 1. In RSA encryption, n is the product of two large prime numbers p and q. If Trudy can factor n (i.e., find p and q), she can compute z=(p−1)(q−1), which is necessary for calculating the private key d. Once she has the private key d, she can decrypt the message using it.

1 mark

2. When p and q are very large, factoring n becomes extremely difficult using current technology. As a result, it is not feasible for Trudy to decrypt the message by factoring n within a reasonable time frame.

1 mark

Q9. Hash.

(1) Write down your magic sentence and submit Python 3 code as supplementary material.

(2) How long does it take to find the nonce?

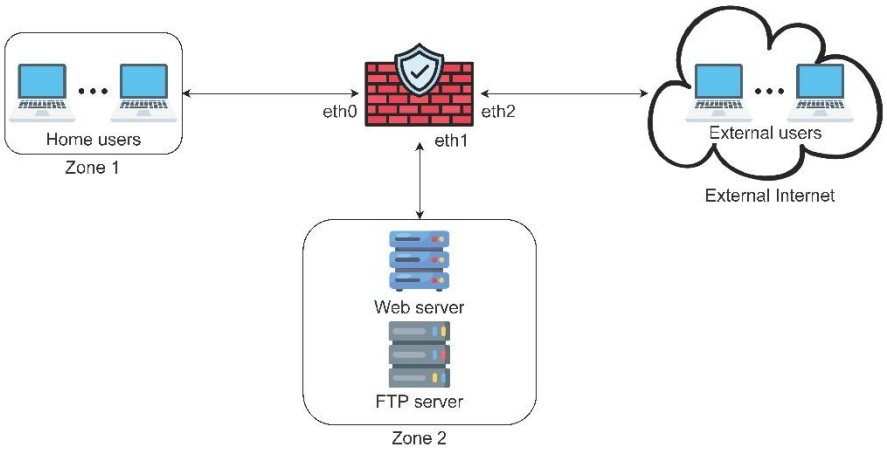
(3) In this case we find the nonce generating for 6 leading zeros. Comment on what will happen if you are required to find a nonce that generates a hash with 20 leading zeros. (This is what miners of bitcoins should do!)

(1) & (2) For questions 1 and 2, the answers will differ due to varying inputs. However, it can be solved using an exhaustive search approach. We will first use the nonce provided by the student along with the student's SID to form a sentence. Then, we will generate a SHA-384 hash for that sentence and check whether the hash has 6 leading zeros. We also need to run the student's code, first checking whether there is a timer included, and then verifying whether it can find the corresponding nonce within the time specified by the student.

(3) Requiring 20 leading zeros (80 bits of zero) demands significantly more computational effort than 6 leading zeros, exponentially increasing the difficulty and time needed to find the correct nonce.

(1) Magic sentence 4%, python code 2%. (2) 2%. (3) 2%.

10. Firewall. We will design firewall configuration. The figure below shows the configuration we want to achieve. The home users reside in Zone 1. The external users reside in the External Internet. A Web server (111.111.11.1) and an FTP server (111.111.11.2) are located in Zone 2. The Web server supports both http and https requests.



The security policy is as follows:

1) The home users can freely access any Web service, anywhere, but only if they initiate the connection themselves.

2) The home users and the external users can establish the telnet session with each other.

3) Everyone (home users and external users) can access the web server in Zone 2.

4) Only the home users can access the FTP server in Zone 2.

5) Block all other incoming and outgoing traffics.

Complete the table to define a stateless firewall configuration for the given scenario. You can refer to the IP ranges (source and destination) by their "zone name". Use "Ext" to indicate external Internet. Use "*" to indicate "all".

Interface	Source IP	Destination IP	Source Port	Destination Port	ACK	Action
Eth0	Zone 1	*	*	80,443	Any	Allow
Eth1	Zone 2	Zone 1	80,443	*	Set	Allow
Eth2	Ext	Zone 1	80,443	*	Set	Allow
Eth0	Zone 1	Ext	*	23	Any	Allow
Eth2	Ext	Zone 1	23	*	Set	Allow
Eth2	Ext	Zone 1	*	23	Any	Allow
Eth0	Zone 1	Ext	23	*	Set	Allow
Eth2	Ext	111.111.11.1	*	80,443	Any	Allow
Eth1	111.111.11.1	Ext	80,443	*	Set	Allow
Eth0	Zone 1	111.111.11.2	*	20,21	Any	Allow
Eth1	111.111.11.2	Zone 1	20,21	*	Set	Allow
*	*	*	*	*	*	Deny