COL334/672: Semester 2023-24-1

Minor exam: 120 minutes, closed-book.

Name: KUSHAGRA GUPTA

Entry number: 2021CS50592

Needless to say, please explain your answers. Zero marks will be awarded if you just state an answer without any explanation. Use the roughwork sheets to work out your answers and write them out neatly in the main answer sheets.

As a student of IIT Delhi, I will not give or receive aid in examinations. I will do my share and take an active part in seeing to it that others as well as myself uphold the spirit and letter of the Honour Code.

Signature: Kushagra

Q1 (out of 13)	
Q2 (out of 4)	
Q3 (out of 12)	
Q4 (out of 8)	
Q5 (out of 13)	
Total (out of 50)	

1. Answer the following link-layer related questions:

a. The transmission rate on a wireless link is 10Mbps. What is the bit-time to transmit one bit? 1Mbps = 10^6 bits per sec. Give your answer in ms.

$10^{-4} = 0.0001$ [1]

Explanation: Transmission delay = L/R, here $L = 1$ bit, $R = 10^7$ bps

$= 10^{-7}$ sec $= 10^{-4}$ ms

b. Continuing from the above, amplitude shift keying is used for modulation, at a signal frequency of 800MHz. How many wavelengths does 1 bit stretch over?

80 [1]

Explanation: BASK → 0 modulated as ⎯⎯
1 modulated as ∿∿ ($n\lambda$)

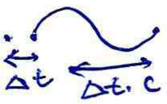
1 bit stretches over n wavelengths
then time to transmit 1 bit $= n\lambda/c$ = delay

$n = c/\lambda \times$ delay $= 800 \times 10^6 \times 10^{-7} = 80$ wavelengths

c. Continuing from the above, if the propagation speed is 3 x 10^8 m/s, what is the physical distance over which 1 bit is stretched? Give the answer in m (meters).

30 [1]

Explanation:



Since the bit time (to transmit one bit) $= 10^{-7}$ sec, each bit will get stretched over to $\Delta t \cdot c = 10^{-7} \times 3 \times 10^{8}$ meters $= 30$ m

d. Continuing from the above, 4B/5B is used as an encoding scheme with NRZI. What is the effective transmission rate? Give your answer in Mbps.

8 [1]

Explanation: Since with 4B/5B & NRZI, original 4 bits are transmitted as 5 bits, ∴ the effective transmission rate decreases by a factor of 5/4 ⇒ $10 \times 4/5 = 8$, (as time for n bits ⇒ $\frac{n}{10} \times \frac{5}{4}$ $= n/8$ ∴ effective rate)

e. Continuing from the above, a packet of size 1000 bytes travels over a network of 4 long-distance wireless links:

Source (S)------Router A------Router B-----Router C-----Destination (D)

If the bit error rate on a link is 0.000005, what is the packet error rate for the end to end transmission from the source to the destination? Assume no other source of packet or bit errors during the transmission.

0.16 or 0.1478.. $= 1-(1-0.000005)^{32000}$ [3]

Explanation: 1000 bytes $= 8000$ bits

For each link, probability no bit has an error for the packet $= (1-0.000005)^{8000}$

$\approx 1 - 8000 \times 0.000005$

For 4 links, probability no error $= (1-0.000005)^{32000}$

∴ packet error rate $= 1-(1-0.000005)^{32000} \approx 320000 \times 0.000005$ $= 0.16$

f. Continuing from the above, what is the one-way latency for the packet to travel from the source to the destination? Assume zero node processing delays, and no queueing delays either. The physical span for each link is 3km (1km = 1000m). The nodes follow a store and forward method. Give the answer in ms.

$10^{-3} \times 1.01$ sec $= 1.01$ ms or 0.81 ms (without 4B/5B) [2]

Explanation:

One way latency $= d_{proc}^{\,0} + d_{queue}^{\,0} + d_{trans} + d_{prop}$

$= L/R + d/s$ $\quad d = 3000$ m

$= \frac{8000 \text{ bits}}{8 \times 10^{6} \text{ bps}} + \frac{3000 \text{ m}}{3 \times 10^{8} \text{ m/s}}$

$= 10^{-3}$ sec $+ 10^{-5}$ sec

(with NRZI, R = 8 Mbps, 4B/5B → 1.01 ms)

(without NRZI + 4B/5B, R = 10 Mbps, → 0.81 ms)

g. Continuing from the above, assume that only adjacent nodes are within one another's range, ie. the source node's transmission can only be heard by router A and does not interfere at any other node, similarly a transmission by router A can be heard by the source and router B only, and so on. Each of the routers have two NICs that can communicate on different non-interfering frequency channels at the same time. Therefore, while S is transmitting to A on frequency channel 1, A can be transmitting to B at frequency channel 2, B can be transmitting to C at frequency channel 3, and C and be transmitting to D at frequency channel 4. The spectral utilization in this case can be said to be 1/4, i.e. 25%. To improve spectral reuse, can B transmit to C at frequency channel 1, i.e. the same frequency at which S is sending to A?

Yes ______ $S \xrightarrow[1]{} A \xrightarrow[2]{} B \xrightarrow[3]{} C \xleftarrow[4]{} D$ [2]

Explanation:

Since only A can hear source node transmission, the transmission can never reach B or C. ∴ it is independent of S → A channel, so we can utilise that frequency channel to improve spectral reuse, provide A transmits to B & D → C or C → D at different frequencies.

h. Continuing from the above, what frequency channel assignment to the links can achieve the highest spectral utilization? Write your answer against the links below and also give the highest spectral utilization:

1 2 1 2

Source (S)------Router A------Router B-----Router C-----Destination (D) [1]

1/2 ______ (spectral utilization) [1]

Explanation:

In a similar manner as above, S--A & B--C can have same frequency channel as well as A--B and C--D can have the same channel. This way they can communicate on non interfering channels (each node A, B, C, D) and we can't do better than this as only 1 channel is not possible

2. Suppose a firewall is running on a proxy server to filter out packets that match a particular pattern. At what lowest layer of the network stack would the firewall be operating to be able to apply the following filters? [4]

a. Allow only a particular IP address range: Network Layer

b. Drop UDP data: Transport

c. Allow only HTTP data: Application

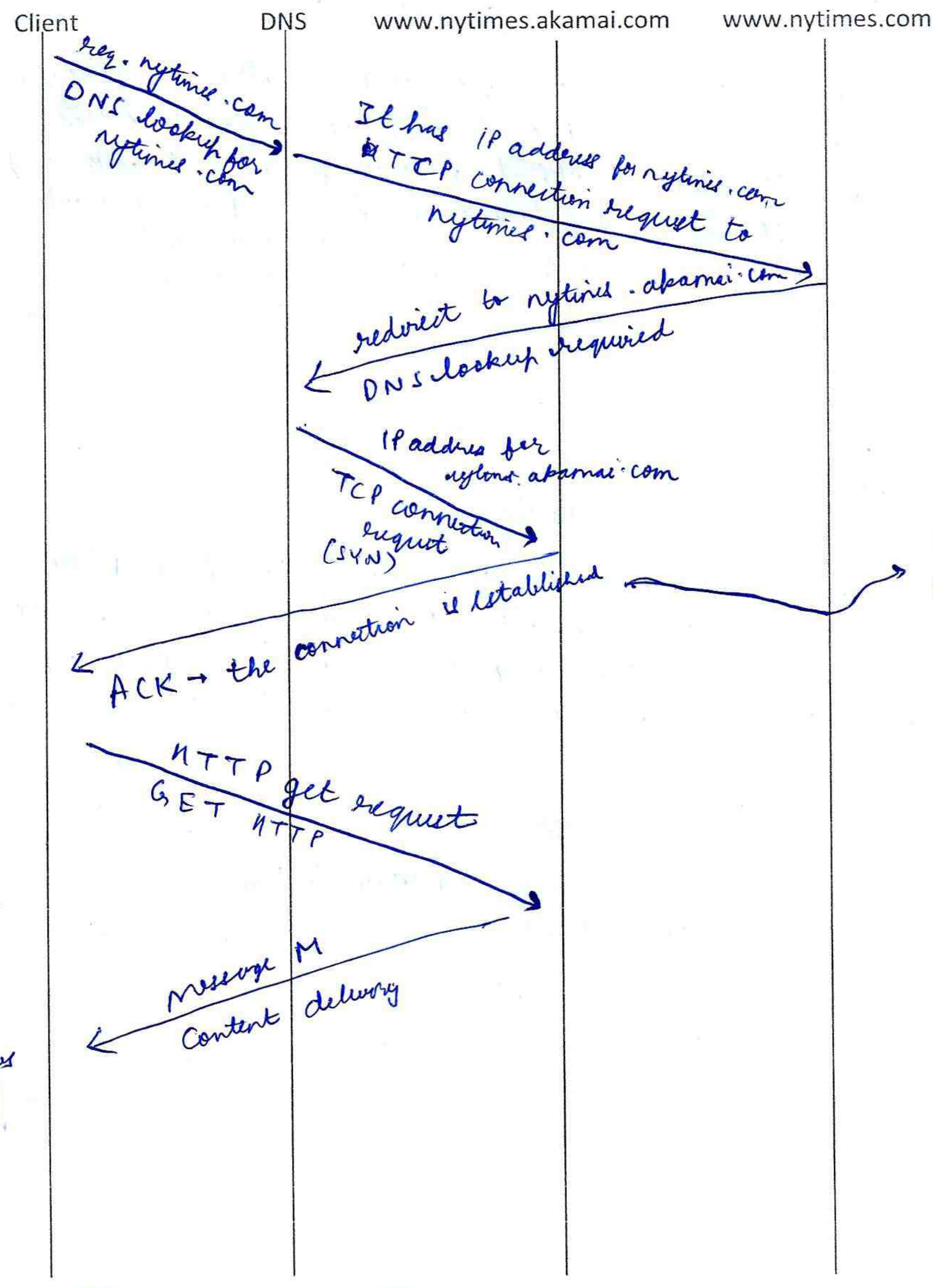
d. Filter out all .jpg images in web pages: Application

Application
Transport
Network
Link
Physical

3. Akamai is a CDN (Content Delivery Network) provider. It places its content servers inside the networks of regional ISPs. An ISP like Airtel may have an Akamai server inside its network, and similarly Reliance may also have an Akamai server inside its network.

a. Suppose a content provider like NYTimes wants to use Akamai. It does this by responding to the very first HTTP GET request to www.nytimes.com, with a redirection to www.nytimes.akamai.com. The client will then use DNS to resolve www.nytimes.akamai.com, and since Akamai would have populated DNS servers to return the IP address for the Akamai CDN server that is closest to the client, therefore after the DNS resolution the client will be able to connect to the closest Akamai server to get the content. Trace out the TCP connection establishment, HTTP, and DNS requests, to the point where you actually start receiving content. [3]

HTTP protocol operated on TCP transport protocol



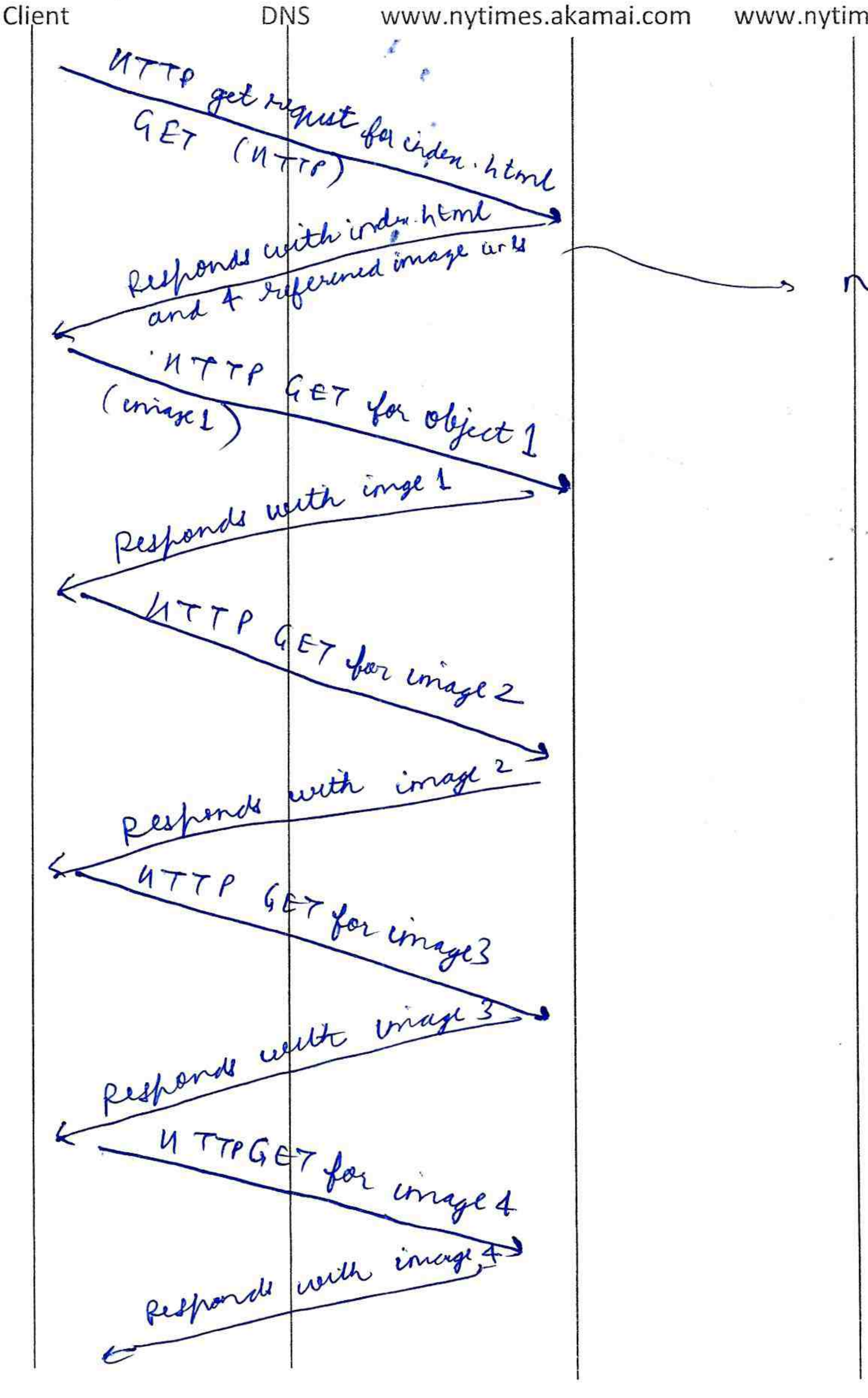
Note:
While sending acknowledgement message, DNS lookup is not required because the CDN server knows IP address of the client

Note ~~know~~ another DNS lookup is not required as the client knows IP address of the server

b. What is the key benefit of using content delivery networks? [1]

Since a large number of people use video streaming etc (80%), to avoid latency, congestion, queuing delay & provide end-to-end content delivery, organisations install CDNs routers / peer with local ISPs to create content delivery networks (CDNs)

c. Assume the web page structure is as follows: a base index.html file which refers to 4 images to be downloaded from www.nytimes.akamai.com. Complete the transaction diagram from where you left off in part (a), with connection keep-alive but without HTTP pipelining being used. [2]

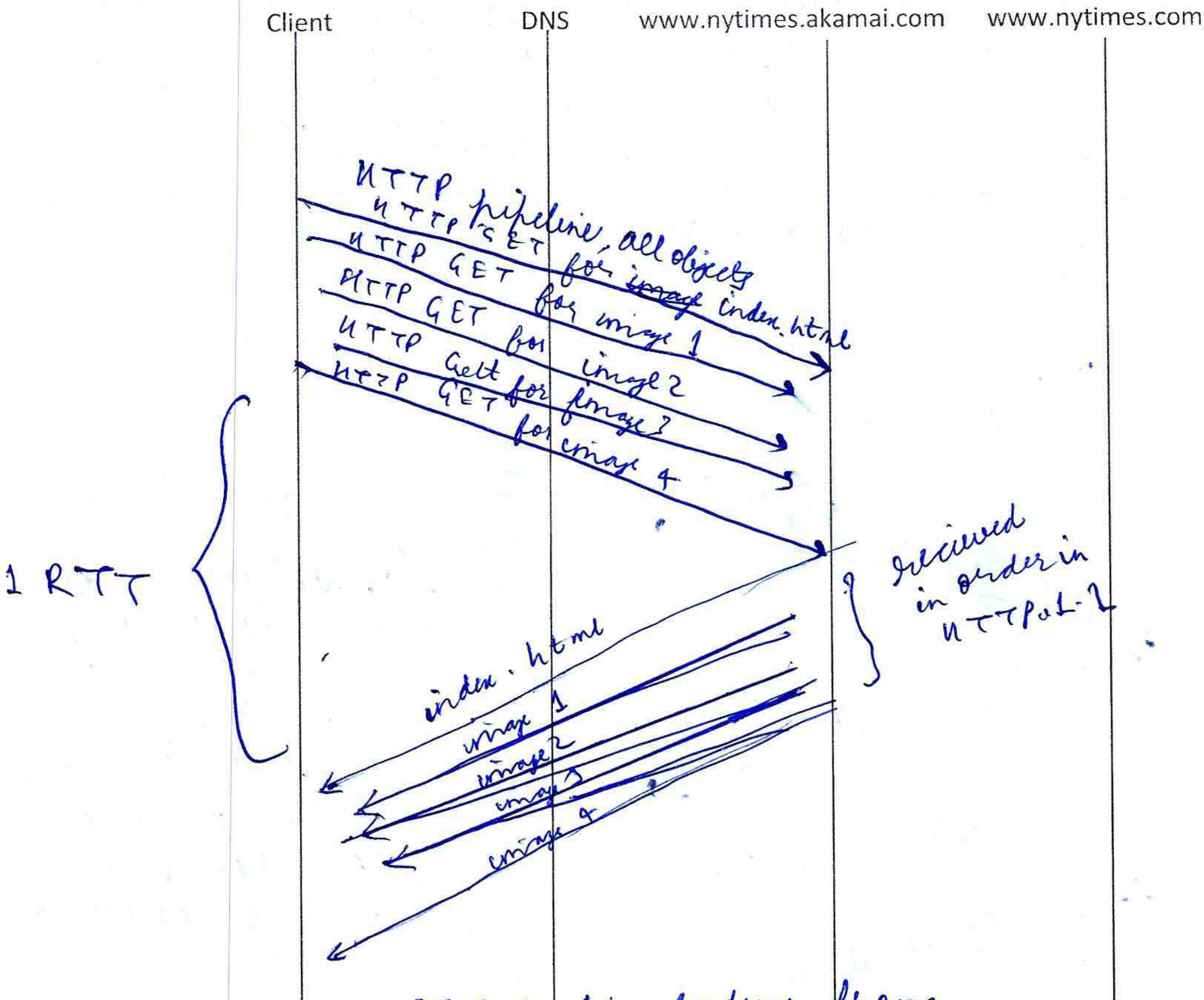


note: No need for DNS lookup here as the images are stored in the same server & client already has the IP address for that server.

Connection keep alive:- Didn't need new TCP connections to be formed.

Without HTTP pipelining

d. Now trace the same web page download request with connection keep-alive and HTTP pipelining being used. [1]



Total time reduces from
5 RTT ~~to~~ + transmission delay
1 RTT + transmission delay
where
transmission delay = time to transmit index.html & 4 referenced images.

e. Suppose a modification is made in the HTML standard which allows multiple sources to be specified in the base index.html from where subsequent objects could be downloaded. For example, this is how the web page could indicate that 1.jpg can be downloaded from either of www.nytimes.com or www.nytimes.akamai.com:

In this case, the client could open multiple TCP connections, to www.nytimes.com and www.nytimes.akamai.com and fetch different objects from different connections. Give at least two advantages of such a setup and when would these advantages be realized. [3]

* In this case, the multiple TCP connections opened in parallel, allow the objects to be fetched simultaneously, reducing the transmission time from $t_{index.html} + \sum t_{all\ 4\ images}$ to $\max(t_{index.html}, t_{image1}, \dots, t_{image4})$.

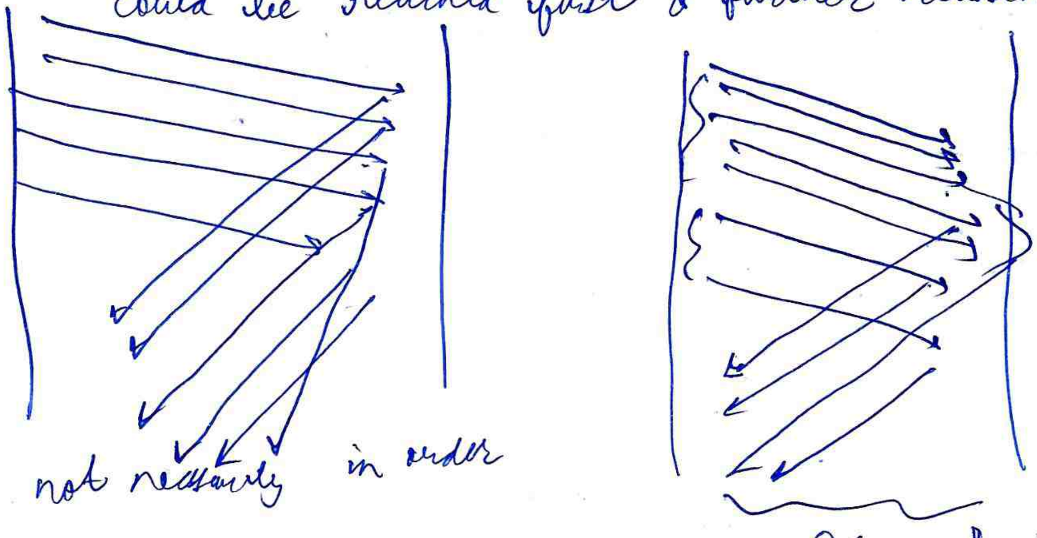
* The site could get loaded faster, as soon as the objects are fetched, they could be loaded onto the website. But it requires multiple TCP connections.

This setup prevents head of line (HOL) blocking and is realised when there is a large request / object being sent in a pipeline, which consumes bandwidth & prevents further objects from being loaded.

f. In part (e) above, give one disadvantage of such a setup. Does HTTP/2 address this issue? [2]

Opening parallel TCP connections consumes bandwidth

HTTP/2 addresses the issue of HOL Blocking by interleaving frames of objects (divided into frames) & reassembling them in order to prevent the issue of maintaining order in pipeline. The frames sent later on could be reached first & further reassembled at the client



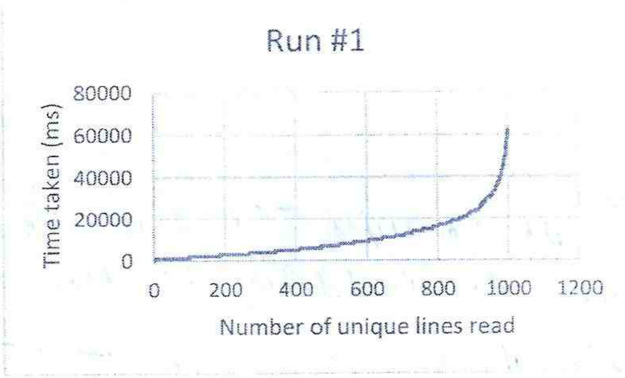
not necessarily in order

this is only possible in HTTP v2

in HTTP v1.1 interleaving is not possible

Opening frames

4. In Assignment-2, our experiments brought up graphs such as the following for the time taken for a single client to receive a given number of unique lines, against the maximum number of unique lines to be received (1000 in this case). We will now try to derive this theoretically, step by step.



a. Without loss of generality, it can be assumed that the server has pre-decided the order in which it would send unique lines. Therefore, for any SENDLINE request, the random choice the server would have to make is whether to send a line it has already sent, or advance to the next line that it should send. At time *i*, assume the server has sent x_i unique lines, out of *N* maximum unique lines it has to send. This can be represented graphically as follows. What is the probability p_i that the server will send a new line at this point in response to a SENDLINE request? The server makes a draw from a uniformly random distribution that produces a number between 0..N-1. [1]

$\underbrace{\quad}_{x_i}$ $\quad p_i = \dfrac{N - x_i}{N}$ (uniformly random distribution)

Since there are $N - x_i$ unique lines & it chooses one line from N lines, $p_i = \dfrac{N - x_i}{N}$

b. Now calculate the expected value of the waiting time at this point, to receive a new line, i.e. after how many SENDLINE requests is the server likely to send a new line. Hint: The chance of waiting for n units of time would be $(1 - p_i)^{n-1}p_i$. Explain why, and then calculate the expected value w_i as the waiting time at point *i*. [4]

$$E[\#\text{ of SENDLINE requests}] = \sum_{k=1}^{\infty} k \cdot P[\text{wait for } k \text{ units of time or } k \text{ requests need}]$$

$$P[\text{it waits for } k \text{ units of time}] = \underbrace{(1-p_i)^{k-1}}_{\text{does not receive unique in } k-1 \text{ time}} \; \underbrace{p_i}_{\text{receives unique finally}}$$

$$= \sum_{k=1}^{\infty} k(1-p_i)^{k-1} p_i = S \qquad -(1)$$

$$(1-p_i)S = \sum_{k=1}^{\infty} k(1-p_i)^k p_i \qquad -(2)$$

(1) − (2)

$$S p_i = \sum_{k=2}^{\infty} k(1-p_i)^{k-1} p_i + p_i - \sum_{k=1}^{\infty} k(1-p_i)^k p_i$$

$$SP_i = \sum_{t=1}^{\infty} (t+1)(1-P_i)^{t} P_i + P_i - \sum_{k=1}^{\infty} k(1-P_i)^{k} P_i$$

$(k-1=t)$

$$SP_i = \sum_{t=1}^{\infty} \underbrace{(1-P_i)^{t}}_{GP} P_i + P_i - \cancel{N(1-P_i)^N P_i}$$

$$= \cancel{P_i\left[\frac{1-(1-P_i)^N}{1-(1-P_i)}\right] + P_i - N(1-P_i)^N P_i}$$

$$SP_i = P_i\left[\frac{1-P_i}{1-(1-P_i)}\right] + P_i = P_i + 1 - P_i \Rightarrow S = \frac{1}{P_i}.$$

$$\therefore \text{the } w_i = 1/P_i = \frac{N}{N-x_i}$$ (can also say using binomial distribution)

c. The total expected waiting time to receive all N unique lines would then be $w_0+w_1+...+w_{N-1}$. Simplify this using the expressions obtained above. You may need to use the harmonic series sum: $1+1/2+1/3...+1/N$ = approximately $\ln N + c$, where $c \sim 0.577$ as $N\to\infty$. [3]

Now total waiting time to receive N unique lines

$$= \sum_{i=0}^{N-1} w_i \quad \& \quad x_i = i$$ (no. of unique lines read at point i where i is the point when i unique read)

$$w_i = \frac{N}{N-x_i} = \frac{N}{N-i}$$

$$= \frac{N}{N-x_0} + \frac{N}{N-x_1} -- \frac{N}{N-x_{n-1}}$$

$$= \frac{N}{N} + \frac{N}{N-1} --- \frac{N}{N-(N-1)} = \sum_{i=0}^{N-1} \frac{N}{N-i}$$

$$= N\left[\frac{1}{1} + \frac{1}{2} -- \frac{1}{N-1} + 1/N\right]$$

$$\approx N\left[\ln N + c\right]$$

here $c = 0.577$ as $N \to \infty$.

$$\approx N \ln N + 0.577N.$$

5. Short-answer questions:

a. Ethernet requires a minimum frame size to ensure collision avoidance — [1]

b. The IP address 127.0.0.1 refers to host [1]

c. Each router has one unique IP address. True/false? False [1]

(as many IP addresses as number of interfaces (may have more interfaces, IP address, also public & private IP)

d. Even when using transport layer security for SMTP connections, so that your data transmission is encrypted, SMTP servers can still read your email. True/false?

True for TLS (but False for end-to-end encryption) [1]

e. When you switch to a new network, you typically contact a DNS server to get an IP address in this network. [1]

f. ARP operates on UDP. True/false? ~~True~~ False (TCP as reliable) [1]

g. The 700MHz spectrum is considered prime property because it has a considerable wavelength to avoid deflection/loss and less bit time to transmit (due to high Freq) so it is a sweet spot between higher freq & λ, ∴ it is expensive & prime property. [1]

(as f ↑, bit time reduces ↓, but λ ↓ causing more loss) (Wifi operates at 2.4GHz)

h. Ping uses the ICMP protocol. [1]

i. The latency for data transmission over the Starlink network is expected to be similar to that over a terrestrial fibre optic network. True/false? True [1]

j. Ad networks like Doubleclick use Cookies to track users across different websites. [1]

k. TCP transmits data at a constant rate. True/false? True [1]

l. A device with two NICs will have two MAC addresses. True/false? True [1]

m. Which error detection scheme is useful to detect bursty bit errors? Cyclic Redundancy Check (CRC) [1]

(as Parity bits don't work for bursty errors).

f) DNS or Traceroute uses the UDP Protocol