COMP 431 Midterm Exam

Mike Ardiff

TOTAL POINTS

35.5 / 75

QUESTION 1

Persistent v. Pipelined Connections 4 pts

1.1 Persistent Connection Definition 1 / 1

✓ **+ 1 pts** **OK Definition**

+ 0.5 pts Partially correct defn, missing some important element

+ 0.5 pts Unclear, confusing definition

+ 0 pts Poor/Incorrect Definition

1.2 Pipelined connection Definition 1 / 1

✓ **+ 1 pts** **OK Definition**

+ 0.5 pts Partially correct def'n, missing some important element

+ 0 pts Poor/Incorrect Definition

1.3 Essential Differences 1 / 2

+ 2 pts OK response

✓ **+ 1 pts** **Missing some essential/important difference**

+ 0 pts Incorrect answer

QUESTION 2

Traffic Intensity v. Queueing Delay 4 pts

2.1 Traffic Intensity Definition 0.5 / 1

+ 1 pts OK Definition

✓ **+ 0.5 pts** **Partially correct definition/missing some important element**

+ 0.5 pts Confusing/unclear/incomplete def'n

+ 0 pts Incorrect

2.2 Queueing Delay Definition 0.5 / 1

+ 1 pts OK Definition

✓ **+ 0.5 pts** **Partially correct/Missing some important element**

+ 0.5 pts Confusing/unclear/incomplete

+ 0 pts Incorrect

2.3 Essential Differences 1 / 2

+ 2 pts OK Differences

✓ **+ 1 pts** **Missing some important element**

+ 0.5 pts Partially correct/unclear/incomplete

+ 0 pts Incorrect

QUESTION 3

Frequency v. Time Division Multiplexing 4 pts

3.1 Frequency Division Multiplexing 1 / 1

✓ **+ 1 pts** **OK Definition**

+ 0.5 pts Partial definition/Missing an important element

+ 0.5 pts Unclear/incomplete/partially correct

+ 0 pts Incorrect

3.2 Time Division Multiplexing 0.5 / 1

+ 1 pts OK Definition

+ 0.5 pts Partially correct/missing some important element

✓ **+ 0.5 pts** **Unclear/incomplete/partially correct**

+ 0 pts Incorrect

3.3 Essential Differences 1 / 2

+ 2 pts OK Differences

✓ **+ 1 pts** **Missing some important element**

+ 0.5 pts Partially correct/unclear/incomplete

+ 0 pts Incorrect

QUESTION 4

Bandwidth Delay Product v. Propagation Delay 4 pts

4.1 BDP Definition 1 / 1

✓ **+ 1 pts** **OK Definition**

+ 0.5 pts Partially correct/missing some important element

+ 0.5 pts Unclear/incomplete/partial answer

+ 0 pts Incorrect

4.2 Propagation Delay Definiton 0.5 / 1

+ **1 pts** OK Definition

+ **0.5 pts** Partially correct/missing some important element

✓ **+ 0.5 pts** **Unclear/incomplete/partial answer**

+ **0 pts** Incorrect

4.3 Essential Differences **1 / 2**

+ **2 pts** OK Differences

✓ **+ 1 pts** **Partial answer/missing some important element**

+ **0.5 pts** Unclear/incomplete/partial answer

+ **0 pts** Incorrect

QUESTION 5

HTTP Page Download Performance 20 pts

5.1 Performance w/ Non-Persistent Connections **4 / 5**

✓ **+ 5 pts** **Good formula/expression**

+ **2 pts** Good formula for getting the base page

+ **2 pts** Good formula for getting the embedded objects

✓ **- 1 pts** **You are not showing/explaining your work as required**

- 1 pts Your units are wrong -- you're mixing ms and secs

+ **0 pts** Incorrect

+ **0 pts** Click here to replace this description.

5.2 Performance w/ Persistent Connections **4.5 / 5**

✓ **+ 5 pts** **Good formula/expression**

+ **2 pts** Good formula for getting the base page

+ **2 pts** Good formula for getting the embedded objects

- 1 pts You are not showing/explaining your work as required

- 1 pts Your units are wrong -- you're mixing ms and secs

+ **0 pts** Incorrect

+ **0 pts** Click here to replace this description.

- 0.5 **Point adjustment**

Analysis for banana.com is wrong

5.3 Performance w/ Pipelined Connections **0.5 / 5**

+ **5 pts** Good formula/expression

+ **2 pts** Good formula for getting the base page

+ **2 pts** Good formula for getting the embedded objects

✓ **- 1 pts** **You are not showing/explaining your work as required**

- 1 pts Your units are wrong -- you're mixing ms and secs

✓ **+ 0 pts** **Incorrect**

+ **0 pts** Click here to replace this description.

+ 1.5 **Point adjustment**

Can't tell how pipelining is effecting your answer

5.4 Performance w/ Parallel Connections **0.5 / 5**

+ **5 pts** Good formula/expression

+ **2 pts** Good formula for getting the base page

+ **2 pts** Good formula for getting the embedded objects

✓ **- 1 pts** **You are not showing/explaining your work as required**

- 1 pts Your units are wrong -- you're mixing ms and secs

✓ **+ 0 pts** **Incorrect**

+ **0 pts** Need a full expression for the response time

+ 1.5 **Point adjustment**

Don't see how you are using the parallel connections

QUESTION 6

Content Distribution Networks 12 pts

6.1 Redirection to a Server in the CDN **1 / 6**

+ **6 pts** OK answer

+ **4 pts** Right idea but missing detail/errors in answer

+ **1 pts** Need a far more complete answer

✓ **+ 1 pts** **Not clear you understand what's going on**

here

+ **0 pts** Incorrect

You're not really answering the question.

6.2 Placing Requested Content in URL **2 / 6**

+ **6 pts** OK answer

✓ **+ 4 pts Right idea but missing detail/small errors in answer**

+ **1 pts** Need a far more complete answer

+ **1 pts** Don't understand your answer

+ **0 pts** Incorrect

- 2 Point adjustment

QUESTION 7

DNS Resource Record Types 12 pts

7.1 CNAME and NS DNS Records **3 / 4**

+ **2 pts** Good answer for CNAME records

✓ **+ 1 pts Partially correct answer for CNAME records**

+ **0 pts** Incorrect answer for CNAME records

✓ **+ 2 pts Good answer for NS records**

+ **1 pts** Partially correct answer for NS records

+ **0 pts** Incorrect answer for NS records

7.2 Purpose of Each Record Type **2 / 4**

+ **2 pts** OK answer for CNAME records

+ **1 pts** Partially correct answer for CNAME records

✓ **+ 0 pts Incorrect answer for CNAME records**

✓ **+ 2 pts OK answer for NS records**

+ **1 pts** Partially correct answer for NS records

+ **0 pts** Incorrect answer for NS records

7.3 When and Why Records Used **3 / 4**

+ **2 pts** OK answer for CNAME records

✓ **+ 1 pts Partially correct answer for CNAME records**

+ **0 pts** Incorrect answer for CNAME records

✓ **+ 2 pts OK answer for NS records**

+ **1 pts** Partially correct answer for NS records

+ **0 pts** Incorrect answer for NS records

QUESTION 8

Problems Browsing the Furby Website 15 pts

8.1 Cause of Performance Problems **4 / 9**

+ **9 pts** OK answer

+ **7 pts** Right idea, wrong justification

✓ **+ 4 pts Not correct but at least a coherent theory**

+ **2 pts** Your "most likely cause" was precluded in the problem description

+ **1 pts** Incorrect and not even a coherent theory

+ **0 pts** Incorrect

8.2 Determine if the Cause is the Correct Cause **1 / 6**

+ **6 pts** OK answer

+ **4 pts** Not the right answer but an OK answer given what you said in part (a)

+ **3 pts** Right idea but too many ideas/steps missing/wrong

+ **2 pts** This won't shed insight into the problem you listed in part (a)

✓ **+ 1 pts The problem description tells you what the result of your experiment will be**

+ **0 pts** Incorrect

QUESTION 9

9 Extra Credit: Finding the name of the requested video **0 / 0**

+ **5 pts** Correct!

+ **1 pts** Not great but maybe you know what's going on!

✓ **- 0 pts Incorrect**

- **0 pts** No attempt

COMP 431 INTERNET SERVICES & PROTOCOLS

Fall 2018
Kevin Jeffay

Midterm Examination

October 8, 2018

There are thirteen (13) pages to this exam (including this page). Please *print* your last name at the top of *each* page (except this cover page). You have until 4:50 PM (1 hour and 15 minutes) to complete this exam.

This is a closed book/closed notes exam. You may *not* use any reference material or electronic aid (laptop, PDA, calculator, cell phone, *etc.*) during this exam.

Notes:

- Each question is labeled with a numeric value representing *both* the amount of time you should expect to spend on the question (in minutes) *and* the number of points the question is worth. The total value of all questions is 75 points. Allocate your time accordingly. (*E.g*, if a question is worth 10 points you probably should not spend more than 10 minutes on the question.)
- Answer all questions on this exam. DO NOT WRITE ON THE BACK OF ANY PAGE! If you need extra space to answer a question extra pages will be provided to you (just ask!).
- You should read all parts of a question before answering any part of the question.

Please sign the following pledge:

I have neither given nor received unauthorized aid on this exam.

Signed: Michael Ardiff

Printed Name: Michael Ardiff PID: 730088498

1. [16 pts.] Define the following terms so as to highlight the essential difference(s) between them.

a) Persistent versus pipelined connections...

What is a persistent connection?

A persistent connection is a connection in which the client and server maintain their connection so they can have multiple request/responses occur with only 1 handshake.

What is a pipelined connection?

A pipelined connection is a persisted connection where the client can send requests one after another before even receiving a response.

What is the difference between the two?

A pipelined connection is a type of persistent connection that does not have to wait for a response. At worst, a pipelined connection performs at the level of a persistent connection (when there is only 1 request between client and server).

b) Traffic intensity versus queuing delay...

What is traffic intensity?

The average utilization of a network device that transmits data.

What is queuing delay?

The average amount of requests waiting to get transmitted at a router or any network device that transmits data.

What is the difference between the two?

Queing delay is a function of intensity. $q = \frac{I}{1-I}$.

c) Frequency versus time division multiplexing...

What is frequency division multiplexing?

Frequency division multiplexing is sending different data across a link at the same time by assigning them different frequencies.

What is time division multiplexing?

Time division multiplexing is sending short piece of data one after another so fast that even the computer cannot tell that full datas were broken into segments

What is the difference between the two?

FDM is relied upon by cable, which will send video 1, video 2, and audio all at the same time next to each other on a link. TDM is used by things like telephones, and different data is sent in small parts one after another to ensure fairness.

d) Bandwidth delay product versus propagation delay...

What is the bandwidth delay product and why is it an important concept?

Bandwidth delay product is the amount of data that can occupy an access link. It's important because it lets us know how much data can actually be in the process of being sent at 1 time.

What is propagation delay and why is it an important concept?

Propagation delay is the time it takes for data to travel across an access link. It is important because it helps us point the overall RTT for data transmission, and possibly address bottlenecks to improve RTT.

What is the difference between the two?

BDP is how much data can be on a link at once. Propagation delay is how long it will take for that data to cross the link

2. [20 pts.] Consider the problem of downloading homepage for the website the *www.banana.com*. The HTML file for this homepage is 20,000 bits and contains the following embedded URLs:

http://www. banana.com/banner-ad.jpg	(15,000 bits)
http://www.apple.com/juicy-apple.jpg	(10,000 bits)
http://www.orange.com/orange-growers-ad.jpg	(10,000 bits)
http://www.apple.com/buy-apples-ad.jpg	(10,000 bits)
http://www. banana.com/team-banana-ad.jpg	(10,000 bits)
http://www2. banana.com/pie-recipe.jpg	(5,000 bits)

For this problem you are going to determine how long would it take to download the *banana.com* homepage and all the embedded objects assuming the browser always used HTTP connections of a certain type. For each question below you are to give an expression (*i.e.*, a formula) that gives the required download time. You only need give the expression, you need not evaluate the expression (*i.e.*, compute a number) however, you *must* show/explain your work. To analyze this problem you may assume:

- All network links connecting the browser to each server operate at 100 Mbps.
- The size of the HTTP request is negligible as are the headers in the HTTP response.
- The browser has an internal (empty) browser cache and that there is no proxy cache on the network.
- The round-trip time from the browser to all servers in the *banana.com* domain is 50 *ms*, the round-trip time from the browser to all servers in the *www.apple.com* domain is 10 *ms*, and the round-trip time from the browser to all servers in the *www.orange.com* domain is 20 *ms*.

There are four parts to this question spread out over this page and the next three pages.

NO CACHING AT ALL?

a) How long would it take to download the *banana.com* homepage and all the embedded objects if the browser used only non-persistent connections?

Non-persistent connections means we need to handshake before every request. That means it's going to be 2 RTT for every content, plus the transmission time of the response.

Transmission Time = : $\frac{20000 + 15000 + 10000 + 10000 + 10000 + 10000 + 5000 \text{ bits}}{100{,}000{,}000 \text{ bps}} \times \frac{1000 \text{ ms}}{s}$

Overall RTT Time = $3 \times 2 \times 50 \text{ ms} + 2 \times 2 \times 10 \text{ ms} + 2 \times 20 \text{ ms}$

Total Time = Transmission Time + Overall RTT Time, but transmission is negligible

The HTML file for the *banana.com* homepage is 20,000 bits and contains the following embedded URLs:

http://www.banana.com/banner-ad.jpg	(15,000 bits)
http://www.apple.com/juicy-apple.jpg	(10,000 bits)
http://www.orange.com/orange-growers-ad.jpg	(10,000 bits)
http://www.apple.com/buy-apples-ad.jpg	(10,000 bits)
http://www.banana.com/team-banana-ad.jpg	(10,000 bits)
http://www2.banana.com/pie-recipe.jpg	(5,000 bits)

All network links connecting the browser to each server operate at 100 Mbps. The round-trip time from the browser to all servers in the *banana.com* domain is 50 *ms*, the round-trip time from the browser to all servers in the *www.apple.com* domain is 10 *ms*, and the round-trip time from the browser to all servers in the *www.orange.com* domain is 20 *ms*.

b) How long would it take to download the *banana.com* homepage and all the embedded objects if the browser used only persistent (non-pipelined) connections?

Persistent means 1 - handshake per server.

$$\text{Transmission is still} = \frac{20k + 15k + 4 \times 10k + 5k \text{ bits}}{100{,}000{,}000 \text{ bps}} \times \frac{1000 \text{ ms}}{5}$$

$$\text{All RTTs} = \underbrace{3 \times 50 \text{ ms}}_{\text{www.banana}} + \underbrace{3 \times 10 \text{ ms}}_{\text{www.apple}} + \underbrace{2 \times 50 \text{ ms}}_{\text{www2.banana}} + \underbrace{2 \times 20 \text{ ms}}_{\text{www.orange}}$$

Time = Transmission + ALL RTTs, but transmission is negligible

The HTML file for the *banana.com* homepage is 20,000 bits and contains the following embedded URLs:

http://www.banana.com/banner-ad.jpg	(15,000 bits)
http://www.apple.com/juicy-apple.jpg	(10,000 bits)
http://www.orange.com/orange-growers-ad.jpg	(10,000 bits)
http://www.apple.com/buy-apples-ad.jpg	(10,000 bits)
http://www.banana.com/team-banana-ad.jpg	(10,000 bits)
http://www2. banana.com/pie-recipe.jpg	(5,000 bits)

All network links connecting the browser to each server operate at 100 Mbps. The round-trip time from the browser to all servers in the *banana.com* domain is 50 *ms*, the round-trip time from the browser to all servers in the *www.apple.com* domain is 10 *ms*, and the round-trip time from the browser to all servers in the *www.orange.com* domain is 20 *ms*.

c) How long would it take to download the *banana.com* homepage and all the embedded objects if the browser used only pipelined connections?

Transmissin is negligible

RTT totals across www.ban... 2 × 50ms + 2 × 10ms + 2 × 50ms + 2 × 20ms

The HTML file for the *banana.com* homepage is 20,000 bits and contains the following embedded URLs:

http://www.banana.com/banner-ad.jpg	(15,000 bits)
http://www.apple.com/juicy-apple.jpg	(10,000 bits)
http://www.orange.com/orange-growers-ad.jpg	(10,000 bits)
http://www.apple.com/buy-apples-ad.jpg	(10,000 bits)
http://www.banana.com/team-banana-ad.jpg	(10,000 bits)
http://www2.banana.com/pie-recipe.jpg	(5,000 bits)

All network links connecting the browser to each server operate at 100 Mbps. The round-trip time from the browser to all servers in the *banana.com* domain is 50 *ms*, the round-trip time from the browser to all servers in the *www.apple.com* domain is 10 *ms*, and the round-trip time from the browser to all servers in the *www.orange.com* domain is 20 *ms*.

d) How long would it take to download the *banana.com* homepage and all the embedded objects if the browser used up to 4 parallel, non-persistent connections?

Since "key .. and there's 6 resources, you would have two go across, meany the two apple resorces have finish, so the maxmm time now be 10 + 50 ms.

3. [12 pts.] The Raleigh newspaper, the News and Observer (the "N&O"), maintains a website *www.newsobserver.com*. On the newspaper's home page there are a number of links to videos of current news stores. Originally, the newspaper hosted all of the content found on its website on its own servers but now the video content is delivered by a third-party CDN. As such, whereas before, the HTML for the N&O's home page contained URLs for links to videos of the form

http://www. newsobserver.com/assets/videos/breakingnews/cat-burglar.mp4 (1)

now, the URL for the same video appears in the HTML for the N&O home page as

http://hs7vtga89ql.sktn3.z23.cdn-name.net/ A NS /MX CNAME (2)

As a result of this change, today, when the user clicks on a link for some video content embedded in the rendered *www. newsobserver.com* web page, the media will be delivered to the user from one of the CDN's thousands of servers that is close to the user and has the requested content. For example, if the user clicks on link (2) above, a video of story on a local cat burglar will be streamed from one of the CDN's servers that is close to the user. Importantly, when the N&O converted all of their video URLs to reference the third-party CDN, it did so without any knowledge of how the CDN stored content in its network.

There are <u>two parts</u> to this question spread out over this page and the next page.

a) Technically, URL (2) above consists of just a hostname and does not contain a file path to any object. As such, it appears to be a URL for a home page for some site in the *cdn-name.net* domain. How is it that the user is directed to a CDN server that has a copy of the cat burglar video when they click on link (2)? Explain all the steps in this process.

When the user clicks on link (2), they will be directed to their local DNS. Because the local DNS won't know where to go since it's a new domain, it will be pointed to an authoritative DNS that may suggest to try an ANS such as dns1.cdn-name.net in an NS record. Since this name server is now hierarchically above the full domain name, it will either keep providing new NS records to look at, or just provide an A record with the address of the full domain. At that point the address will be given to the host which will then establish a connection with the full domain's CDN server.

b) An alternate way the N&O web server could have changes its URLs for video assets to be served by the third-party CDN would have been to change URL (1) above to something like

http://www.cdn-name.net/<some path name>/cat-burglar.mp4 (2)

What's wrong with this approach?

First off, the N&O does not own that domain, so it cannot directly place that there. Also, the nice thing about CDNs is they take advantage of the CNAME resource record, so there's no need to try to put the URL as such when the DNS can figure out the alias very quickly.

4. [12 pts.]

a) Explain what CNAME DNS records are and what NS DNS records are.

CNAME DNS records are DNS resource records that take a domain and give back another domain that is an alias. NS records are resource records that take a full domain name and give the Authoritative Name Server appropriate for that domain.

b) What purpose does each record type serve?

CNAMEs are useful, for example in agreements between different domains where one is using the other as an alias, such as a redirect.

NS records are useful when a DNS does not have an A record for a domain, but it can point to the ANS in order to get into the correct family of name servers.

c) When and why might a DNS server return one of these types of records versus the other or versus an A record?

An A record is only returned when a DNS server knows the IP address of a server. If it does not, it will look for a CNAME or NS because it is trying to get closer to its destination. With a CNAME, the domain can be a very different record (like www.ncstaterejects.com having a CNAME entry that points www.unc.edu), whereas a NS record is likely returning a high level server in a hierarchy (like dns1.unc.edu for gb.cs.unc.edu).

5. [15 pts.] A Dook and an NCSU student share a love of Furbies and constantly visit a Furby website to read Furby postings, learn Furby news, *etc*. On a Furby social media site the Dook student learns that his time to download the main Furby web page, a modestly decorated web page with a few dozen embedded objects for adds, display graphics, images, *etc*., is 10 times longer than the download time for the State student. This is strange because Dook and State (and UNC) connect to the Internet at the same access point in RTP and both students have similar computers running the same OS and browser software configured in a similar manner (similar cache settings *etc*.). Moreover, neither campus operates a proxy cache and the performance disparity occurs every time these students access the page and even occurs when they use other computers on their campus to access the page. The students also determine that they are getting the same exact content from the website (*i.e*., there's no site specific generation of content occurring).

The Dook and State students know a little bit about networking and they each perform a series of traceroute experiments to the main Furby web server and a few of the ancillary servers that deliver content for the Furby web page. The results of these experiments confirm that other than the "first mile" connection to the Internet (the connection from each campus to the RTP Internet access point), the paths to all servers on the Internet are exactly the same. And the first mile (or the campus network) cannot be the problem because, the network path from the Dook student's laptop to the Furby website turns out to give the same performance in all the traceroute experiments as the path from the State student's laptop on the State campus to the website.

Confused (as usual), the Dook and State student approach **YOU** — a friend at UNC who they know is in the famous COMP 431 course (and hopefully has better things to do in life than visit Furby websites). Premium Furby Access? Net Neutrality?

There are <u>two parts</u> to this question spread out over this page and the next page.

a) What would you suspect is the most likely cause of the 10:1 performance difference in downloading the content necessary to display the Furby web page at Dook and State? Justify your answer.

Because we know that both students' machines take the same time to reach the Furby website itself, the problem cannot be in the main website retrieval. Instead, the issue must lie in the retrieval of other content on the home page. Thus there is likely a CDN server very close to the NCSU campus that is making it so much faster.

b) Describe an experiment or set of experiments you could perform (at either Dook, State, or UNC) to confirm your hypothesis.

I would run traceroute to see if I could find the CDN's address of the actual content. From there, I would ping the CDN's from their servers, and check to see how long it took to access these CDNs.

Extra Credit Do NOT attempt this question unless you are completely done with the rest of the exam and are satisfied with your answers. Little to no partial credit will be given for this question so don't attempt it unless you are confident in your answer.

Returning to question (3), as previously stated, technically the URL (2)

http://hs7vtga89ql.sktn3.z23.cdn-name.net/ (2)

consists of just a hostname and does not contain a file path to any object. As such, it appears to be a URL for a home page for some site in the *cdn-name.net* domain. That is, the GET request a browser will emit after a user clicks on a link such as (2) will — ignoring headers — simply look like

GET / HTTP/1.1

How is it that the third-party CDN knows which video to send to the user when the user clicks on link (2)?

Becau there's only one video on this server.