COMP 431 Internet Protocols & Services

Fall 2018  
Kevin Jeffay

Worksheet 9, October 17

1) Redo problem 2 from the midterm (the problem of downloading homepage for *www.banana.com*. And embedded object). In each case make sure you give 1 coherent expression and *pay attention to your units!*

*a*) How long would it take to download the *banana*.*com* homepage and all the embedded objects if the browser uses only non-persistent connections? Hint: Your expression should evaluate to 480.8 *ms*.

4\*2\*50 + 2\*2\*10 + 2\*20 ms + (20,000+15,000+4\*10,000+5,000) bits / 100,000,000 bps \* 1000 ms / 1 s

*b*) Repeat your analysis for (*a*) now assuming the browser uses only persistent connections. Hint: Your expression should evaluate to 370.8 *ms*.

4\*50 + 3\*10 + 2\*50 + 2\*20 ms + (20,000+15,000+4\*10,000+5,000) bits / 100,000,000 bps \* 1000 ms / 1 s

*c*) Repeat your analysis for (*a*) now assuming the browser uses only pipelined connections. Hint: Your expression should evaluate to 310.8 *ms*.

3\*50 + 2\*10 + 2\*50 + 2\*20 + (20,000+15,000+4\*10,000+5,000) bits / 100,000,000 bps \* 1000 ms / 1 s

*d*) Repeat your analysis for (*a*) now assuming the browser can use up to four parallel connections. Hint: There are several possible answers here but two obvious values your expression might evaluate to are 300.34 *ms* or 200.35 *ms*.

There’s 4 connections, so we can only grab 4 resources at a time.

If I arrange my connections in this fashion:

Homepage and first banner ad (1): 2\*50ms (for handshake) + 2\*50 ms (for propagation) + (20000 + 15000) bits / 100,000,000 bps \* 1000 ms = 200.35 ms

Juicy apples, buy apples, and orange growers (2): Apples resources total to 2\*10ms + 2\*10 ms + (10,000 + 10,000) bits / 100,000,000 bps \* 1000 ms = 40.2 ms, and orange resource is 2\*20ms + 2\*20 ms + (10,000) bits / 100,000,000 bps \* 1000 ms = 80.1 ms, so they add to 120.3 ms

Team banana (3): 2\*50ms (for handshake) + 2\*50 ms (for propagation) + (10,000) bits / 100,000,000 bps \* 1000 ms = 200.1 ms

Pie Recipe (4): 2\*50ms (for handshake) + 2\*50 ms (for propagation) + (5,000) bits / 100,000,000 bps \* 1000 ms = 200.05 ms

So since the slowest connection is 200.35 ms, that one determines how fast we can retrieve all of the resources.

2) Redo problem 3 from the midterm (the News and Observer website using the *cdn-name* CDN for delivering video content).

*a*) The URL the user clicks on:

|  |  |
| --- | --- |
| *http://hs7vtga89ql.sktn3.z23.cdn-name.net/* |  |

consists of just a hostname and does not contain a file path to any object. Explain in detail 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 this link? Make sure you answer what the question is asking. That is, if you explain how the user is directed to a certain server, you need to also explain how it is that that server will have a copy of the cat burglar video.

Points you should pay attention to:

* This URL was created and embedded in the News and Observer home page by the N&O server and this server had *no knowledge* of where content is stored in the *cdn-name* CDN network.
* This question is *not* asking about whether or not the server the user is directed to is close to the user.
* When the server *hs7vtga89ql.sktn3.z23.cdn-name.net* receives the user’s HTTP GET request, the request line of the request will simply be GET / HTTP/1.1. This means that no content/object name or path appears in this request line. This further means that at least on its face, this is a request for a home page (and specifically is not a request for the cat burglar video).

When the news and observer decides to store their cat burglar video on a third-party CDN, they will upload it to some cloud based architecture that then hands the storage over to the third party. From there, the third party CDN will distribute copies of this cat burglar video to different servers across the CDN in a way that it feels will minimize retrieval time for users at a justifiable cost.

Now that the video is inside different servers across the CDN, when the user clicks on the resources on the N&O website, they will issue a get request that will return a manifest for the file that returns a list of CDNs that contain the file. From there, using DASH streaming, the client can pick and choose (and switch during streaming) which CDN to retrieve the file from.

*b*) An alternate way the N&O web server could have changed 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* |  |

What’s wrong with this approach?

The thing that’s wrong with this approach is that these files are stored in chunks throughout the CDN, and that the first response is just going to be a manifest file. That means that a .mp4 doesn’t directly exist in the cdn-name.net server, because that server is going to return a manifest with lists of chunks of the mp4s across the CDN.

3) Redo problem 5 from the midterm (the awful Furby download problem).

*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? Hint: Read the problem carefully. The writeup is designed to eliminate many possible performance problems.

My new guess is that the furby website has P2P file distribution, and everyone who uses this furby website is at NC State. Thus, when the state student begins retrieving the files, they quickly bolt up the leaderboard in terms of uploading and downloading when they become optimistically unchoked, whereas the Duke student will fail to join the top 4 of others because they are not faster than other NC State students.

*b*) Describe an experiment or set of experiments you could perform (at either Dook, State, or UNC) to confirm your hypothesis. Make sure you describe how data gathered from your experiment would prove or disprove your hypothesis.

An experiment I would do is to go to Duke with a lot of laptops and try to raise the level of the P2P file distribution there because they would choose each other as their top 4 by going to the furby website all at once. If I have enough laptops there, then the file distribution time will quickly increase at Duke since it obviously was not operating at anywhere near it’s full efficiency beforehand.