Test Preparation #2

Covers: Udacity Lessons 6 through 8 and associated papers.

Instructions: These questions are ungraded. They are provided as a tool to help you prepare for the closed book tests in this course. It is recommended that you work through these questions as you complete sections of material and papers, rather than leaving them for the week before the Test. You are encouraged to discuss these questions on Piazza with your fellow students!

We will post an answer key the week before the test with model answers.

Congestion Control and Streaming

1. What causes congestion collapse to occur?

2. What is the difference between *fairness* and *efficiency* in a congestion control scheme?

3. Assuming traditional TCP Reno with AIMD behavior (i.e., the version presented in the lecture videos), suppose a TCP flow's bottleneck link has 1 Gbps capacity, and that link is not being shared with any other flows. What will the *average* throughput of that flow be, in megabits per second (Mbps)?

4. What circumstances lead to the incast problem? (In other words, what factors must be present for incast to occur?)

5. Suppose you are working on some live video call software (think Skype or Google Hangouts) and decide to build your application protocol on top of UDP (rather than TCP). Give as many different points as you can (minimum two) to help justify that decision.

Reading – CUBIC TCP

6. Why does the linear growth rate of TCP-RENO (1/RTT) perform poorly for short lived flows in networks with large bandwidth and delay products?

7. Describe the operation of BIC-TCP's binary search algorithm for setting the congestion window. Once stable, how does BIC-TCP react to changes in available bandwidth, i.e. what happens when there is a sudden increase or decrease in available bandwidth?

8. How does the replacement of this congestion control algorithm with a cubic growth function in CUBIC-TCP improve on BIC-TCP? Discuss.

9. What is the purpose of the following regions of the CUBIC growth function:

    a. Concave

    b. Plateau

c. Convex

10. How does CUBIC's fast convergence mechanism detect a reduction in available bandwidth (i.e. a new flow competing for bandwidth)?

Reading – TCP Fast Open

11. What kinds of web traffic stand to benefit most from utilizing the TFO option? How does TFO improve the performance of these flows?

12. Describe how a trivial implementation of TCP Fast Open (in which the server replies to a all HTTP GET requests with a TCP SYN-ACK packet with data attached) can be exploited to mount a source address spoof attack. How does TFO prevent this?

Reading - Multi Path TCP (MPTCP)

13. What threat do network middleboxes pose to negotiating MPTCP connections? How does the design of MPTCP mitigate this?

14. Why are receive buffer sizes required to be larger for MPTCP enabled connections? What controls does MPTCP put in place to maximize memory usage?

Reading - Dynamic Adaptive Streaming over HTTP

15. Explain two of the benefits of using HTTP as a streaming media delivery protocol. Make sure you explain the "why" as well as the "what".

Rate Limiting and Traffic Shaping

16. Would you use a leaky bucket or a token bucket to traffic shape a constant bit rate (CBR) audio stream? Explain why.

17. If you want to traffic shape a variable bit rate (VBR) video stream with average bit rate 6 Mbps and maximum bit rate 10 Mbps, should you use a leaky bucket or a token bucket? What values should you use for rho and beta if you want to allow bursts of up to 500 ms?

18. Suppose you're running an ISP and get the crazy idea that implementing Power Boost for your own network would be a good idea. For the 6 Mbps service plan (i.e., customers can have a sustained rate of 6 Mbps), you'd like to allow them to burst up to 10 Mbps for up to 10 seconds. In megabytes (MB), what should you set the beta parameter of your token bucket to? (Round to the nearest tenth of a MB, i.e., one decimal place.)

19. Read about the following two Active Queue Management (AQM) techniques: Random Early Detection (RED) and CoDel. Although they vary in specifics, these two algorithms share a common basic approach to solving the buffer bloat problem. Explain what that approach is and why it works.

20. If you want to find out if a remote host (i.e., not *your* server) is currently under a DoS attack, would you use active or passive measurement? Explain why.

21. If you want to compute the traffic intensity, I=La/R, on a router interface (i.e., the ratio between arrival rate and forwarding rate), would you use Counters, Flow Monitoring, or Packet Monitoring? Explain why.

Reading – Sizing Router Buffers

22. Discuss the drawbacks to over-buffering routers. If memory is widely available at low cost, why is it a bad idea to use massive buffers to ensure high link utilization?

23. Under what conditions was the "rule-of-thumb" for buffer size ($B = RTT \times C$) originally conceived? How does this fundamentally differ from current, real world conditions?

24. Statistical modeling of desynchronized long lived flows indicates that smaller buffer sizes are sufficient to maintain link utilization as the number of these long lived flows increases. However, not all flows can be expected to be long lived. Discuss why short lived flows (less than 100 packets) do not significantly detract from these findings.

Reading – Controlling Queue Delay

25. Explain how standing queues develop in network buffers at bottleneck links. Why is a standing queue NOT correctly identified as congestion?

26. Consider the CoDel active queue management algorithm. How does the algorithm decide whether or not drop a flow's packets? What effect does dropping the packet have on the TCP sender?

Content Distribution

27. If your browser has a page in the cache and wants to know if the page is still fresh and can be used, or is too old and should be replaced with fresh content from the server, which HTTP method should it use to find out?

(If you are familiar with the If-Modified-Since header field, which we have not discussed in this class, please assume that we are not using If-Modified-Since.)

28. Consider the HTTP protocol. What will cause a server to send a response message with the status code...

a. 404 Not Found ?

b. 302 Moved Temporarily (also sometimes called 302 Found) ?

c. 200 OK ?

29. Consider the HTTP protocol. What would the following header fields be used for?

a. Last-Modified

b. Host

c. Cookie

30. Of the various methods to redirect web requests to CDN servers, DNS redirection is the most common. Why would this method be more popular than the alternatives?

31. How does BitTorrent implement the tit-for-tat algorithm? Be sure to explain in detail, including the roles of both choking and unchoking.

32. In a distributed hash table, how many steps (hops) are required to lookup an item if the finger table is a constant size (i.e., its size does not depend on the number of DHT nodes in the system)? Explain why that is the right answer.

33. For a more typical DHT setup where the finger table has O(log N) entries, for N DHT nodes in the system, explain why the number of steps to access an item is O(log N).