CSci551 Spring 2006/Thursday Midterm Exam 9 March 2006

Name: ______________________________ E-mail: ______________________________

Student id: ___________________________ Location: ______________________________

(For location, list USC if you attend on campus. If you are a DEN remote student, list your DEN location.)

This exam will run 90 minutes (until 10:30am). At 10:40am we will start class in OHE studio F. This exam is closed book, closed note, closed neighbor. There are 3 enumeration questions (with 3 blanks), 11 short answer, 4 medium answer, and 1 long answer questions totaling 87 possible points. For each problem, make sure you have a distinct answer for *each* question mark or blank. For questions with multiple parts (indicated by letters), please label each part of your answer with the corresponding part of the question.

The last question should be number 19; please check to make sure you have all pages.

1 Enumeration

For each question, please fill in the blanks with the requested word or phrase. Each blank is worth 1 point.

1: What is the term for *throughput divided by delay*?

Answer: Power

2: Early Internet addressing was "classful". What is the term for the current Internet address allocation scheme?

Answer: CIDR or Classless Internet Domain Routing.

3: After a loss, what event triggers a fast retransmit in TCP?

Answer: A triple-duplicate ACK. OR the 3rd duplicate ACK.

2 Short Questions

Please use the space to write one or two sentences answering the question. Each question is worth 4 points. (Please answer in the large white space, not by the question itself. Please remember to *label each part* of your answer with a, b, c, etc.)

4: Suppose AS-101 purchases Internet service from two Internet service providers, AS-1 and AS-2. These providers peer with each other directly. AS-101 therefore has a router that has links to routers at both AS-1 and AS-2, and AS-1 and 2 have other routers joined by a direct link. (Assume these are all typical Internet ASes and routers.)

Suppose a backhoe cuts the direct link between AS-1 and AS-2. a) (1pt) Will the traffic go through the router in AS-101? b) (3pts) Why or why not?

Answer:

a) No, it will not. b) Because typical ASes establish policies that refuse to route through customers. OR because typical AS paths must be valley free.

5: a) (1pt) In Landmark routing, why do some nodes have multiple addresses? b) (1pt) What is one advantage Internet addresses have compared to Landmark routing addresses?

Both Internet addresses and Landmark addresses are hierarchical. c) (2pts) Why might Landmark addresses be more efficient to allocate than IP addresses?

Answer:

a) because they can see multiple landmarks at some level. b) IP addresses are fixed length; Landmark addresses are variable length.

c) Because they are dynamically allocated and reallocated, Landmark addresses can be more efficiently allocated than IP addresses (no wasted space for future hosts).

6: Labovitz et al, in their paper "Delayed Internet routing convergence" stated that the worst case convergence time for BGP is $O(n!)$ messages, because in a full mesh, BGP will try all possible combinations. a) What is one *protocol* reason BGP doesn't behave this badly in the real world, and why does that provide more rapid convergence? The network today has about 20,000 ASes, so $n!$ could be very bad. In addition to your answer in part (a), b) what is one *deployment* reason BGP doesn't behave this badly in the real world?

Answer:

a) BGP uses MinRouteAdver, which means routers learn more in each round and converge in $O(n-3)$ messages.

b) In the real world ASes are not connected as a full mesh.

7: Shaikh et al. looked at router behavior under load in their paper “Routing stability in congested networks: Experimentation and analysis”. To determine that the link is up, routing protocols send “hello” or “keep-alive” messages over the link, expecting that these messages would get through. Shaikh et al. found that routers would frequently declare links to be broken when there was high traffic on a link. They also found that BGP was much less susceptible to this problem than OSPF—the mean time between failures for BGP is about a factor of ten longer than OSPF. a) Why is BGP so much more resilient? Other than changing the routing protocol, b) What is a router level mechanism that could make this problem (of incorrectly declaring the link dead) much less likely? (Be brief in your answers—one sentence at most.

Answer:

a) Because BGP uses TCP and TCP tries more to retransmit the keep-alive for the routing protocol, while OSPF’s UDP doesn’t retry.

b) Several answers are possible. Easiest: the router could prioritize routing (or control) traffic over data traffic. OR the router could do Fair Queueing. (Assumes no huge number of flows.) OR just “separation of the traffic” (rather than prioritization)

8: Assume TCP is running with cwnd of 32 segments. Suppose the first packet in a round is lost. a) What algorithm allows TCP to recovery this packet without timing out? Suppose TCP wasn’t running this algorithm and so it ends up timing out. After the timeout, TCP will eventually return to steady state. b) What will be the cwnd at this time?

Answer: a) Fast retransmit. a) 16 segments.

9: IPv6 set out to address several related problems. It solved some of these well, while at least one problem was not really addressed. a) What is *one* of the problems that IPv6 successfully solved? b) What is one of the problems that IPv6 set out answer that it has *not yet* addressed?

Answer: a) Address space shortage OR inefficient address allocation OR routing table size.

b) routing table size.

10: Clark’s paper on the Internet Architecture listed “the Internet must permit host attachment with a low level of effort” as one of the original design goals. a) (2pts) What kinds of costs was Clark referring to in this statement? Today researchers are more concerned about other kinds of “effort to attach a host”. b) (1pt) What is one network protocol that reduces the level of effort required to attach a host? c) (1pt) What effort does it reduce?

Answer:

a) protocol implementation costs OR “implement the protocol stack”

b) DHCP, c) manual configuration of network addresses (and other parameters)

Other answers for b and c are possible.

b) CWA (Common Wrong Answer): IP or packet forwarding

11: Suppose you wished to become an ISP for five of your friends in your neighborhood. In addition to the means and right to forward packets, a) what will you need to obtain from your upstream Internet provider?

ISPs that are multi-homed can tolerate failures at one of their upstream providers by routing through the other. b) What is a technical reason related to your answer in part (a) that makes it unlikely that ISPs would like let you be multi-homed?

Answer:

a) A block of Internet addresses.

b) Your address space is small and multi-homed address spaces cannot be aggregated.

12: Improving Internet security has been a major goal. One huge hole in Internet security is routing: currently you typically trust what your peer tells you, so it's possible for a router to announce "I'm the shortest path to everywhere", then everyone sends their traffic there. (In fact, this even has actually happened multiple times!)

Consider the choice between a link-state routing protocol (where each router announces its state and that is flooded to everyone) and a distance vector protocol (where your neighbors tell you their best paths to everywhere). a) (1pt) Which protocol is more attractive from a security point of view? b) (3pts) Why?

Answer:

a) Link state.

b) Because with LS you trust only what is said about a node and its links, while with DV you must trust what you are told about *everything*.

CWA: a) distance vector, b) because you get info from only your neighbors and you trust your neighbors. (Wrong because your neighbors can be lied to).

13: Consider Ramakrishnan and Jain's DECbit system for router feedback to a sender.

Suppose the sender decides to react to congestion and reduce its sending window size from W to W'. a) Why does R&J's system *ignore* then next W congestion signals?

They compute average queue sizes at the router over the last regeneration period. b) Why not just use the most recent queue size as an indication of congestion?

Answer:

a) because it takes one RTT (one W's worth of packets) for a change in sending rate to propagate through the network.

b) Because queue size (1) varies frequently so the router would (2) change its feedback frequently, (3) reducing fairness. (Full credit requires at least 2 of these three things.)

14: An important question in network protocols is making sure the other end has gotten something. We talked about how BGP uses TCP to send routing table updates to the router on the other side of the link. Assume you have two computers, A and B, each peering with each other. A sends a route to B and wants to know if B got it and added it to B's routing table. a) (1pt) Is the ACK of the TCP connection sufficient to make this assertion? b) (3pts) Why or why not? (Briefly [1 or 2 sentences maximum] justify part (a) using reasoning from class.)

Answer:

a) No, because B may get the route but not execute it. b) The end-to-end argument says that you should wait for an application level acknowledgment to make sure the route is really added.

OR

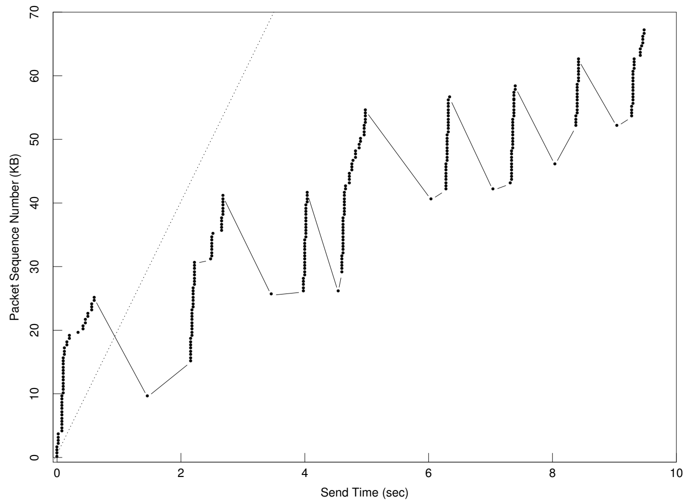
a) Yes b) Because fate sharing argues that if b gets it and crashes before executing it, there's nothing lost. (only get credit for yes if they give the fate-sharing answer.)

CWA: a) Yes, but b) just talking about how TCP works and that an ACK means data was received. (It does, but that doesn't say anything about the *application*.)

3 Medium Questions

Please use the large open space to write a few sentences answering the question. (But some parts to a question may require only a few words, not even a sentence.) Each question is worth 6 points. (Please remember to *label each part* of your answer with a, b, c, etc.)

15: This figure shows a time-sequence plot of a TCP connection taken over a 230kb/s link by Van Jacobson.



Other than general poor throughput, a) (1pt) What is one specific performance problem this graph indicates? b) (2pts) What about the graph demonstrates that that is a problem? (Your answer for part (a) should name the problem, while part (b) should describe what specifically in the graph shows the problem you mentioned in part (a).)

c) (1pt) Is this TCP connection doing slow start? d) (2pts) What about the graph supports your answer to part (c)?

Answer:

a) there are lots of timeouts, b) shown because there are large gaps in the send time when nothing is sent.

OR a) there are lots of retransmissions, b) shown because there are lots of packets that show up multiple times (on the same horizontal line)

c) No. d) The slows after transmit are nearly vertical, not 1-2-4-etc.

16: A simple equation about expected TCP throughput is $B \sim \frac{1}{RTT}\sqrt{\frac{3}{2p}}$ where p is the probability of packet loss on that path. Assume you have two flows, flow A between USC and UCLA, and flow B between USC and MIT. a) (2pts) What does this say about TCP performance of flows A and B? This simple equation ignores a number of details about TCP performance. b) (1pt) What is one aspect that is ignored? c) (3pts) In what circumstances would the aspect that you mention in (b) matter?

Answer:

a) USC-ISI get much higher throughput because it has lower RTT than USC-MIT.

Several answers to b/c:

b) timeouts, c) common when windows are small or loss is high

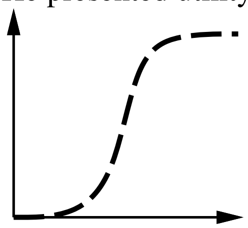
b) window limitation, c) common if you have small windows

b) delayed ACKs, c) commonly used

b) AQM, c) may altar routing

other answers possible

17: Shenker talked about the *utility* of an application operating at a given bandwidth.
He presented utility graphs that looked like this:



His graphs don't have units, but it is important to understand the qualities he measures on each graph. a) (1pt) What does the x-axis describe? b) (1pt) what does the y-axis describe?

Suppose you were drawing a similar curve for digital video. Assume the basic video stream takes up 19Mb/s. When the bandwidth falls below 19Mb/s, video packets are lost and so the viewing quality becomes blocky (you've probably seen it if you've ever used digital video over the Internet). Assume for the moment that at 15Mb/s it becomes unwatchable.

Considering this video system and a network that guarantees no packet loss, where the user has the option of purchasing a bitpipe of any speed from 0–32Mb/s. c) (4pts) What is the utility curve for this user? (Draw the graph, label both axes with their units, and label any inflection points in the graph.)

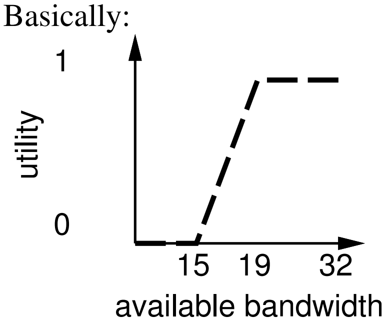
Answer:

a) (1pt) available bandwidth OR throughput

b) (1pt) utility OR goodness

c) (4pts) label axes available bandwidth and utility, inflection points at 15Mb/s and 19Mb/s, utility of 0 from 0–15Mb/s, linearly increasing from 15–19Mb/s, and 1 from 19Mb/s onwards.

Basically:



Scoring: 1pt for each inflection point at 15 and 19, 1pt for linear between, 1 for 0 before and 1 after.

18: A recent development is that the Chinese government has decided to field its own DNS root nameservers. These will support the usual generic top-level (gTLD) domains (.com, .net, .us, etc.), but also support Chinese-language alternate gTLDs (something like .shangye). This new gTLD will be available to anyone who uses the Chinese root nameserver, but not part of the usual ICANN root nameservers.

Suppose someone in China mails a URL to `ChinesePhotos.shangye` to a friend at USC. a) (1pt) What problems might arise when the friend here tries to click on that URL? b) (2pts) Why?

Although not directly in the Saltzer paper about naming, we talked in class about a mechanism that might help the problem mentioned in part (a). c) (3pts) What technical means might solve the problem mentioned in part (a)? (If you cannot remember what we discussed, you can *briefly* [1 or 2 sentences] describe a solution.)

Answer:

a) The name will not be resolvable b) because USC uses the ICANN root nameservers, not the Chinese root nameserver.

OR a) server not found

c) Passing a *context* with the name.

4 Long Questions

Please use the large open space to write a few sentences answering the question. (But some parts to a question may require only a few words, not even a sentence.) Each question is worth 12 points (unless otherwise indicated). (You should answer all parts of the questions in the large space provided; do not try and squeeze your answers in with the text of the question. Please remember to *label each part* of your answer with a, b, c, etc.)

19: XCP is designed explicitly to address problems that TCP encounters on high bandwidth-delay-product networks. One of its innovations is to provide separate fairness and utilization controllers. Its utilization controller is designed to ensure consistent high utilization of the network. a) (2pts) Why does this scenario (a on high bandwidth-delay product) require a different approach to utilization than TCP?

Consider TCP for a moment. b) (1pt) What congestion control algorithm does TCP run at steady state? c) (1pt) At steady state, how much can TCP increase a single flow's window each RTT?

XCP is *just* about congestion control. It does not address issues of loss repair. Current implementations of XCP use TCP algorithms to do loss recovery. d) (1pt) What TCP loss repair algorithm would be essential for XCP-style networks (i.e., high bandwidth-delay products)? e) (2pts) Why is this algorithm necessary, compared to just doing timeouts or even fast retransmit?

f) (2pt) In today's wired networks, what is the cause of most packet loss? Katabi et al (the XCP designers) argue that XCP can reduce this problem. g) (1pt) What mechanism in XCP reduces the problem mentioned in part (f)? h) (2pts) Why does the mechanism in part (g) reduce the problem in part (f)?

Answer:

a) Because a large changes in bandwidth (when flows enter or exit, or due to multiplicative decrease) leave large holes that need to be quickly filled. (TCP will not quickly fill unused capacity.)

b) congestion avoidance (or additive increase). CWA: AIMD.

c) 1 segment per RTT. CWA: 1/W or 1/cwnd (that's per ACK, not per RTT).

d) SACK. e) Because of the need to recover from multiple losses per RTT.

f) congestion.

g) utilization controller OR sends explicit feedback from the router. CWA: reacting *before* congestion (even XCP cannot predict the future!)

h) Because the utilization controller keeps queues small, so there are buffers free to handle bursts, so losses are rarer.