CSE 3214 Midterm Test
Winter 2014
February 25, 2014
Instructor: S. Datta

1. (10 points) (a) (1 point) Why does TCP not allow applications to control the transmission times of packets?

   Solution:TCP performs reliable delivery of packets and congestion and flow control, each of which requires it to schedule transmissions (including retransmissions) of packets at specific times. Hence it controls packet transmission times and does not let the sender application do this.

   (b) (2 points) The formula used by TCP to compute the estimated roundtrip time between 2 hosts is EstimatedRTT = $(1 - \alpha)$EstimatedRTT + αSampleRTT.
   What effect does α have on the estimated roundtrip time? What is the RTT value set to in TCP?

   Solution: The parameter α is a smoothing factor – it yields a smoother (lower variance) estimate of the RTT from measured RTT values. The higher the α value the less the smoothing.

   The RTT value is set to TimeoutInterval = EstimatedRTT + 4 DevRTT,
   where DevRTT is calculated using a smoothing formula very similar to the one given in the question.

   (c) (1 point) What are the primary disadvantages of proxy servers from the point of view of content providers?

   Solution: The main problems are that the server does not have real-time information about the clients. This hampers or prevents the delivery of customized advertisements or content (e.g. product suggestions at an e-commerce site).

   (d) (2 points) Why is reliable delivery implemented as a service in the transport layer instead of (say) the link layer in the internet?

   Solution: The main motivation is adherence to the simple-core-complex-edge philosophy.

   Note: There is nothing inherent about this choice. Architectures can and have been designed to provide congestion control at lower layers. However doing this would add to the computational burden on routers.

   (e) (2 points) Using the analogy of cars being packets and a tolled road network being the network, provide analogs of propagation delay, queuing delay and transmission delay.

   Solution: The analogy is more accurate for manual toll collection systems rather than electronic ones using transponders.

   Propagation delay is the time needed for a car to travel from one toll gate to another. Queuing delay is the time spent waiting behind other cars at toll gates before a car reaches the toll collector. Transmission delay is the time a car needs to get to full speed (say 100 km/h) after paying the toll.

   Note: Many students wrote that the transmission delay includes the time to pay the toll. This is not accurate – the processing delay is analogous to that.

   (f) (2 points) HTTP is called a pull protocol but SMTP is called a push protocol in the book. What do these terms mean?

   Solution: In HTTP, a file is transferred only when a client requests (or "pulls") it. In SMTP, the sender's email server sends an email to the receiver's email server without any action from the receiver. Hence SMTP is called a push protocol. SMTP does *not* push emails to the receiver's browser or email reading app, and it does not push email from the sender's email client to his email server.

   Note: Points were taken off if you said "server" instead of "sender's/receiver's email server" or equivalent.

2. (13 points) (a) (2+2 points) What are the pros and cons of having congestion control in the transport layer as opposed to the network layer? **Solution:**

   The main advantage is that the network core is simpler and complexity is pushed to the edge. The disadvantage is that the network does not have as much information and cannot react as quickly as it could if routers were involved in congestion control.

(b) (2 points) List some possible security/privacy problems created by the existence of cookies.

Solution:

Cookies set by a site can be read by others. Also cookies set by the same company ad banner on different pages can be used to track users. Finally the combination of user-suupiled inormation and cookies can be used to aggregate more information about users. All this information can be exchanged or cold without the consent or knowledge of the user.

(c) (3 points) Describe briefly the three main function of DNS *other than* providing IP address lookup service.

Solution:

Host aliasing, mail server aliasing, load balancing.

(d) (2 points) State Little's Law, defining all the variables involved.

Solution:

Little's Law is $L = \lambda W$, where L the average number of customers in the system, λ is the average arrival rate, and W is the average time that a job spends in system.

(e) (2 points) Briefly describe what optimistic unchoking means in the context of BitTorrent.

Solution: In BitTprrent a peer unchokes (or sends packets to) a few peers that have demonstrated their cooperativeness. In addition it selects a peer that it unchokes in the hope that that peer will cooperate (hence the adjective 'optimistic'). This is done to allow newly arrived nodes to start packet exchange.

3. (9 points) (a) (3 points) Suppose you are writing a video streaming application. Would you use a single TCP connection like HTTP or two TCP connections like FTP, and why? What are the benefits to making the control connection TCP-based and the data connection UDP-based?

Solution:

Having two connections would improve the responsiveness of the playback to commands like fast forward, pause, rewind, since these control commands could be sent using a separate connection to the data. An additional benefit is what the second part of the question asks. The data channel could use UDP for speed – multimedia data does not need the reliability provided by TCP. The control commands need reliability and should be sent over TCP.

(b) (2 points) What would happen to the usage of different servers if DNS queries are never cached?

Solution:

Without caching the hierarchical structure would be of little use and many more queries would travel up the hierarchy, completely overloading the servers ar higher levels.

(c) (4 points) Suppose a trunk router satisfies the requirements of a M/M/1 queue. That is, assume it has infinite buffer space and the incoming traffic is Poisson with a mean arrival rate λ and the packet processing times follow an exponential distribution with parameter μ. Assume $\mu = 10000$ pkts/sec. List the utilization and the expected number of packets in the buffer for $\lambda = 5000, 9000, 9900$ packets/second. What is the conclusion rom your numbers?

Solution: Utilization is given by $\rho = \lambda/\mu$ and the expected number of packets in the queue is $\rho/(1-\rho)$. By substituting the different λ values we see that the expected number of packets in the queue rises sharply as λ approaches μ.

4. (8 points) (a) (4 points) Imagine a wireless network where each node is mobile and is equipped with a GPS unit. The normal algorithms for routing would be problematic because routing tables would store stale information and further, a lot of communication would be spent updating routing tables. How would you do addressing and routing in such networks? Your system should try to minimize the overhead (number of bits transmitted for routing).

Solution:

This is a very open-ended question, meant to make you think of the underlying problems. Mobility is handled well in cellular systems, with the help of a complex non-mobile infrastructure. We would

like to think of systems that do not have such infrastructure. These are called ad hoc networks in the literature.

There is always a tradeoff between update complexity and the ease of location of nodes and routing of packets. We can prevent updates and use flooding to find nodes. This is very expensive since flooding uses lots of computation. On the other hand we can constantly update routing tables and use lots of communication. So how can we effectively use locations? If we want location speciic and not node specific information, we can use locations as addresses. This idea is similar to P2P systems indexing content in the same address space as nodes.

So if we query a specific location, the packet is sent to the vicinity using some routing algorithm and then any node in the vicinity can respond to the query. Of course the design of routing algorithms for such networks is a research problem. If you are interested, look up "geographical routing' on the Internet.

(b) (4 points) If the Internet designers were thinking of security issues, what would they do differently in DNS services and in the network layer?

Solution:

There are a few different issues here. First, it is very important that all DNS servers can be trusted. So all communication from DNS servers should carry authentication information. Note that encryption is not necessary as the information being sent need not be secret. Secondly, a router can bring down the Internet by advertising false information. So routers should authenticate themselves before sending routing table updates. Again there is usually no need for encryption of the routing tables. Third, in some systems it may be preferable to have encryption services in the network layer so that communication details are not viewable to intruders (we will see some instances of this in the Security chapter).