

ECSE-6600: Internet Protocols

Spring 2007, Exam 1

SOLUTIONS

Time: 75 min (strictly enforced)

Points: 50

YOUR NAME (1 pt):

Be brief, but DO NOT omit necessary detail

{Note: Simply copying text directly from the slides or notes will not earn (partial) credit. Brief, clear and consistent explanation will.}

I. Short Questions: **Networking Ideas Review**

[8 pts] Briefly, explain the differences between:

- A) (2 pts) Go-Back-N vs Selective Repeat
 - B) (2 pts) IP (or L3) forwarding vs forwarding in L2 bridges
 - C) (2 pts) IP fragmentation/reassembly vs TCP segmentation of byte stream
 - D) (2 pts) ARP vs DNS
- A) Go-Back-N : Upon error (lack of acknowledgement at the sender), the complete window is retransmitted. Selective Repeat retransmits only the packets for which acknowledgement was not received. Go-Back-N leads to duplicate packets at receiver, while Selective Repeat leads to out of sequence packets at the receiver.
- B) IP forwarding includes a longest prefix match in the lookup table to determine the next hop, and then sending the packet to the next hop. L2 forwarding involves selective filtering and flooding on one of the ports of the switch to which the destination MAC address is directly connected. If lookup fails for L2 it results in a broadcast. Lookup never fails at L3 due to the presence of a default route.
- C) IP fragmentation/reassembly allows IP datagram to be transmitted over heterogeneous layers 2 networks which may have different MTU sizes. Fragment offset helps in out-of-sequence delivery of fragments. TCP segmentation of byte stream involves dividing application layer data to fit maximum segment size negotiated during the TCP connection establishment. Sequence numbers allow handling of out-of-sequence segment deliveries.
- D) ARP resolves the L3 address of destination to an L2 address. DNS resolves (human friendly) names to IP addresses.

2. [6 pts] **Internet Design:** Explain what mechanisms the Internet architecture uses to meet its goals of *scalability* and accommodation of *heterogeneity*? Compare/Contrast with alternative mechanisms that may be less effective for these goals.

Scalability: a) indirect connectivity using multi-hop forwarding, b) efficient filtering mechanisms to avoid broadcasts, c) hierarchical addressing (IP) d) routing, congestion control and other indirection mechanisms (DNS, ARP etc).

Heterogeneity: a) IP fragmentation to cope up with varying MTU sizes, b) Overlays c) Address + name resolution d) RTT estimation at TCP

Alternate mechanisms: 1) Translation instead of overlays 2) flat addressing instead of hierarchical addressing (does not scale) 3) direct connectivity instead of indirect connectivity (does not scale).

II. [10 pts] Statistical Multiplexing, Congestion Control

- (2 pts) Explain why you do not have the congestion control problem in circuit switching (and why it arises in packet switching with statistical multiplexing)?
- (4 pts) Explain why AIMD leads to fairness, and why AIAD or MIMD may not lead to fairness? (Hint: use phase plots like we did for AIMD in class)
- (4 pts) Consider a random traffic source that has an average rate (μ) 5 Mbps, standard deviation (σ) = 0.25 Mbps, and peak rate of 10 Mbps. You want to provision capacity C . Suppose you want to limit the probability of the short term rate (R) exceeding C (i.e. $P(R > C)$) to 4%, what value of C would you pick? (Hint: recall Chebyshev's theorem we saw in TCP RTO design)
- A) Circuit switching involves resource reservation during call/connection establishment, before data transfer. Since the resources are reserved a-priori, congestion does not occur. Packet switching allows resource sharing (data is chopped up into packets) and allows the total demand to exceed capacity leading to congestion.
- B) AIMD leads to fairness since it responds aggressively to congestion, and probes reluctantly for more resources when available. This allows the AIMD to converge to stability at the intersection of efficiency and fairness (in the diagram) thus guaranteeing fairness. AIAD and MIMD oscillate around efficiency and do not converge to the above intersection. (Students using diagrams to explain know better!)
- C) $P[|R - 5| > a] \leq 0.25^2/a^2 = 4/100$
 $\Rightarrow a = 1.25$
 $\Rightarrow P[R > a + 5] \leq 0.04$
 $\Rightarrow C = a + 5 = 6.25$ Mbps

III. [10 pts] Virtualization, Indirection, Multiplexing, Reliability:

- (4 pts) Explain the concepts of indirection, virtualization and multiplexing.
- (4 pts) Explain how you could create a virtual router out of a set of physical routers to provide reliability. Walk through what happens (how the indirection & virtualization is done) when any important router in your set of physical routers fails.
- (2 pts) How do you ensure that remote nodes will not see any changes (in terms of L3 and L2 addresses they have in their forwarding tables and ARP tables) ?
- A) Indirection is a mapping which allows for dynamic binding and unbinding at any instance of time to allow for better flexibility. Multiplexing refers to sharing if resource across time or frequency or a mix of these. In combination of a multiplexed resource, indirection allows for virtualization – a notion of having an unshared virtual resource. E.g. forwarding maps a destination address to an output port using table lookup thus providing support for indirect connectivity through a virtual link abstraction. Address resolution dynamically resolves an L3-address to L2 address thus allowing an interface to be configured with an arbitrary IP address.
- B) The virtual router has a virtual IP address as well as a virtual MAC address. Mapping between virtual IP address and physical IP address (similarly for MAC address) can change in a manner transparent to remote nodes. The network treats the set of physical routers as one virtual router. All the physical routers share the same forwarding information (next hop for destination). They all have one port each connected to each of the subnets the virtual router belongs to. If one of the routers fail, one of the other routers starts forwarding the packets destined to the failed router. This transparency provides the rest of the network with a virtual router resource abstraction.

- ❑ (2 pts) How do you ensure that remote nodes will not see any changes (in terms of L3 and L2 addresses they have in their forwarding tables and ARP tables) ?

- ❑ The set of routers advertise the routes over a virtual IP address. The remote nodes set their next hop as this virtual IP address. A packet destined to the virtual IP address can be intercepted (processed/forwarded) by any one of the set of physical routers. Virtual IP address and virtual MAC address are the key. Upon ARP for the virtual IP address, a *representative* router responds with the virtual MAC address. However, any of the routers in the set can intercept the layer 2 frame with destination MAC address set to the virtual MAC address. If the representative router fails, a leader election algorithm can be run internally to elect a new representative router to respond to ARP queries. The remote nodes have next hop set as the virtual IP address and its layer 2 mapping set as the virtual MAC address.

IV. [15 pts] TCP & Congestion Control:

- [5 pts] Explain why TCP self-clocking could lead to burstiness? How does it constrain TCP's throughput in asymmetric links? How could you rectify the situation & restore performance?
- [5 pts] Consider a large bandwidth-delay product (BDP) path (eg: between two supercomputers in NYC and SFO), and small buffers relative to the BDP. What performance issues would TCP face in such paths? What would happen to TCP performance if the delay portion of this bandwidth-delay product became smaller (while the product remained high)? Explain crisply.
- [5 pts] If you had multi-bit explicit feedback, could you make TCP performance better? Why/how? How could multi-bit feedback help in cases where you had huge volatility in the BDP?
- A) Self clocking – send one packet upon receiving one ack. Thus if the acks get batched together, it leads to burstiness at the sender. E.g. cumulative acks. In asymmetric links, the acks traversing the slow reverse path might arrive late thus causing sender to wait longer than necessary before sending more traffic. This leads to under-utilization of the fast forward link constraining TCP throughput. Header compression, link level support, cumulative acks can be used to rectify.
- B) TCP does not scale well for large BDP and small buffers, and quickly reaches congestion. Small buffers lead to confusion between burstiness and congestion. In response, TCP reduces window size using multiplicative decrease, and is followed only by additive increase. However, this additive increase is too slow to completely utilize the available bandwidth. Thus TCP throughput remains low. If the delay portion grew smaller leading to smaller RTT, feedback (acks) would be quicker leading to faster increase in utilized bandwidth and resulting in better performance.

[5 pts] If you had multi-bit explicit feedback, could you make TCP performance better? Why/how? How could multi-bit feedback help in cases where you had huge volatility in the BDP?

- A) Yes. Multi bit explicit feedback can increase TCP performance by decoupling efficiency and fairness control. Multi bit explicit feedback can help distinguish packet loss due to error from those due to congestion. It would provide TCP with more information about what is happening in the network, allowing it to better control its window/sending rate. This could help in cases where there is huge volatility in the BDP by telling TCP to slow down its sending rate much before reaching congestion avoidance (and thus preventing TCP from using multiplicative decrease). Also multi bit feedback can help identify underutilized regions where multiplicative increase can be applied to quickly utilize the available bandwidth, and later followed by additive increase. Explicit rate feedback can specify the exact rate at which the sender should send.