Leeson Chen - Homework 4: Network Layer

CNT 4007C Computer Networks Fundamentals Instructor: Prof. A. Helmy

Assigned: Nov. 22hd, 2019. Due Date: Dec 3rd, 2019

(submission instructions similar to hwk3, on canvas then hard copy in class and ofc hrs)

Q1. What are the 2 most important network-layer functions in a datagram network, and what is the difference between them?

The two most important functions are forwarding and routing. Forwarding is analogous to getting through a single interchange on a trip, while routing is analogous to the planning process from source to destination on a trip. Forwarding means to move packets from a router's input to output; routing means to determine the route that packets take.

Q2. How many columns does a bare-bones forwarding table in a MPLS-capable router have? What is the meaning of the values in each of these columns? How many columns does a bare-bones forwarding table in a datagram network have? What is the meaning of the values in each of these columns?

A forwarding table in an MPLS router would have four columns: in label, out label, destination, and out interface. The labels are what forwarding is based on instead of IP addresses. In destination based forwarding, a forwarding table only has two columns: destination address range and link interface. The destination address range is a range of numerical addresses, and the link interface is the output.

Q3. An application generates chunks of 40 bytes of data every 20msec, and each chunk gets encapsulated in a TCP segment and then an IP datagram. What percentage of each datagram will be overhead, and what percentage will be application data?

Assuming that the TCP and IP additional data are both 20 bytes each (taken from the slides), that means the entire datagram is 20 bytes TCP + 20 bytes IP + 40 bytes original data = 80 bytes total. So 50% of the datagram is overhead.

Q4. What is the triangle routing problem in mobile IP? Suggest an improvement to it.

In mobile IP, the triangle routing problem is when there are two hosts, one mobile (e.g. a smartphone) and one fixed (e.g. a desktop computer). The mobile knows the fixed hosts' address, but the reverse is not true: the desktop does not know the phone's address, as it can change. The result is that different routing is needed for different directions. One possible improvement is if all datagrams to the mobile host must pass through the fixed host first, thus enabling the address to be shared.

Q5. Detail two examples in which tunneling helps in mobile IP and IPv6 deployment.

Mobile IP and IPv6 are the inevitable improvement that the future is moving towards, but it cannot be deployed all at once. This necessitates backwards compatibility and interaction with the old standard, IPv4. In an example scenario where a route of IPv6 routers are briefly interrupted by a sequence of older IPv4 routers, tunneling can be used to encapsulate the IPv6 datagram inside IPv4 compatible packets. In another scenario with mobile IP, imagine similarly that newer mobile routers are separated by older IPv4 routers. The IPv4 routers can encapsulate the original data and tunnel through to the other side of the mobile routers, ensuring nothing different is noticed by either side.

Q6. How does IPv6 help in supporting mobility, compared to IPv4?
With increased mobility comes an increased demand for more IP addresses. This demand is twofold; devices may require multiple IPs at different locations due to constant movement, and additionally more mobile devices are produced at a faster rate than fixed devices. As the demand for more IP addresses increases, the amount of IP addresses allowed by IPv4 reaches its limit (32 bit address space). IPv6 solves that problem by introducing 128 bit address space (2^128 is roughly equal to 340 billion billion billion). Additionally, the flags and header information in an IPv6 datagram are more streamlined compared to IPv4.

Q7. Why is MPLS sometimes called layer 2.5?
With layer 2 being the data link layer (traversing a single link) and layer 3 being the network layer (routing the path), MPLS exists in a strange intermediate state where it uses predetermined paths between routers (similar to data link), but doesn't do any "decision making" of the network layer. A compromise is saying it has attributes of both layers 2 and 3, making it layer 2.5.

Q8. Can label numbers be reused in MPLS? Why?
Yes, label numbers can be reused in MPLS because those label numbers are only specific to the input and output of a router and its directly connected neighbors. So two routers which share no neighbors in common can overlap their label numbers.

Q9. Mention two examples in which IP addresses can be reused, explaining the reasons for their reuse.
In DHCP, hosts dynamically obtain their IP addresses from the network server when joining. So any particular host only owns that IP for the duration of its connection, and when it leaves the IP can be used by a different host. Similarly, an ISP owns a range of IP addresses, and can allocate them to hosts that join the service. When someone leaves, that IP is usable by a different host.

Q10. Compare and contrast link-state and distance-vector routing algorithms.
OSPF (open shortest path first) is an example of a link state algorithm, while RIP (routing information protocol) is an example of a distance vector algorithm. Both are subgroups of the broader category of routing protocols in the network layer. Distance vector algorithms assume every node approximates the distance to neighboring nodes, and optimizes this, but convergence time will vary. Link state algorithms can be bounded by node * link number of messages sent in n-squared time. So link state message complexity and speed of convergence can be bounded but distance vector algorithms cannot due to variance. In terms of robustness LS algorithms can advertise incorrect link costs but nodes are limited to computing only their own table. In distance vector algorithms the incorrect path cost can be advertised, and error propagates through then network making it more robust.

Q11. Compare and contrast the advertisements used by RIP and OSPF.
In OSPF, advertisements for link state flood the entire AS. This can be mitigated if there is a two level hierarchy to the AS, so that the advertisement is contained to a single area. A RIP advertisement is only sent to immediate neighbors, although the information in that advertisement is pertinent to all the routers in the topology.

Q12. Fill in the blank: "RIP advertisements announce the number of hops to various destinations. BGP updates, on the other hand, announce the _____ to the various destinations".
sequence of autonomous systems (AS) on the routes

Q13. Complete the following sentence in three different ways, each time using one of the following words: 'vector', 'domain', 'gateway'
"RIP is a _____ _____ protocol, while BGP is a _____ _____ protocol."
RIP is a distance vector protocol, while BGP is a border gateway protocol.
RIP is a intra domain protocol, while BGP is an inter domain protocol.
RIP is a intra AS routing protocol, while BGP is a between AS border gateway protocol.

Q14. Suppose you purchase a wireless router and connect it to your cable modem. Also suppose that your ISP dynamically assigns your connected device (that is, your wireless router) one IP address. Also suppose that you have five PCs at home that use 802.11 to wirelessly connect to your wireless router. How are IP addresses assigned to the five PCs? Does the wireless router use NAT? Why or why not?
In this scenario, the router would come built-in with a DHCP server. Dynamic Host Configuration Protocol allows for the router to assign 5 IP addresses to each of the 5 PC hosts. NAT (Network Address Translation) is also used in this scenario because this is how the router obtains the single IP from the ISP.

Q15. Consider the network setup in the figure below. Suppose that the ISP instead assigns the router the address 126.13.89.67 and that the network address of the home network is 192.168/16.

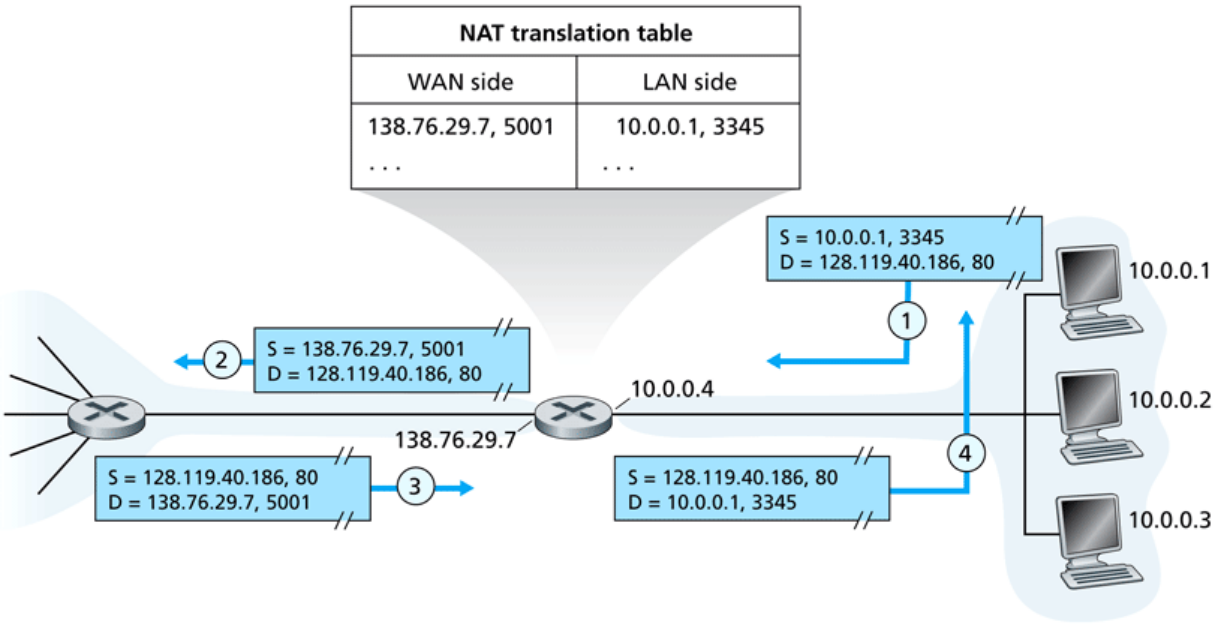


Figure 4.22 ♦ Network address translation

a. Assign addresses to all interfaces in the home network.
192.168.0.1

192.168.0.2
192.168.0.3
192.168.0.4 (router)

b. Suppose each host has two ongoing TCP connections, all to port 80 at host 128.119.40.86. Provide the six corresponding entries in the NAT translation table.

WAN:	*LAN:*
126.13.89.67, 4000	192.168.0.1, 3345
126.13.89.67, 4001	192.168.0.1, 3346
126.13.89.67, 4002	192.168.0.2, 3445
126.13.89.67, 4003	192.168.0.2, 3446
126.13.89.67, 4004	192.168.0.3, 3545
126.13.89.67, 4005	192.168.0.3, 3546

Q16. What is the 'rendezvous problem' in multicast? How can it be solved? (mention three main approaches/algorithms to the solution along with the protocols that use them)
The rendezvous problem in multicast is where, when switching from a shared tree to a shortest path tree (or vice versa), you will need to select a particular point where shared trees meet (i.e. rendezvous). However, optimizing the choice of this meeting point is NP complete, and therefore hard to optimize for realistic scenarios with scaled data. This problem can be avoided with PIM-SM which avoids optimizing the RP; PIM-DM which broadcasts and prunes for dense groups; or just switching to the source tree after a certain threshold of data rate.

Q17. How does SDN extend the notion of matching to define firewalls and NAT boxes, while traditional longest-matching is limited?
SDN is a logically centralized controller for switches so that routers coordinate their forwarding tables. SDN is implemented in three areas: the SDN controlled switches in the data plane, and the SDN controller (network OS) and network control applications in the control plane. The middle SDN controller acts as an intermediary for the switches and applications with northbound and southbound APIs. Firewalls and NAT boxes take place above, in the northbound API. Matching with respect to firewalls and NAT boxes occurs there, while traditional longest matching takes place below in the southbound API.

Q18. Do you see a layer violation in BGP peering? Explain.
BGP is the method through which different autonomous systems (AS) connect, referred to as the glue holding the internet together. BGP peering, where two routers of separate AS exchange messages over TCP, is supposedly a part of the network layer, because it is a type of routing protocol. However, because it is built upon TCP, that would imply it is above TCP transport layer, making it an application layer protocol.

Q19. Provide example of route oscillation in link state routing (e.g., Dijkstra's algorithm).
In link state routing, every node will propagate a map of the available paths. but compute their own table. So any given node will compute the least cost path from itself to all other nodes in a forwarding table. However, due to traffic dynamics the cost of paths can fluctuate. If path A is better and more traffic travels there, suddenly path A becomes more congested, while path B

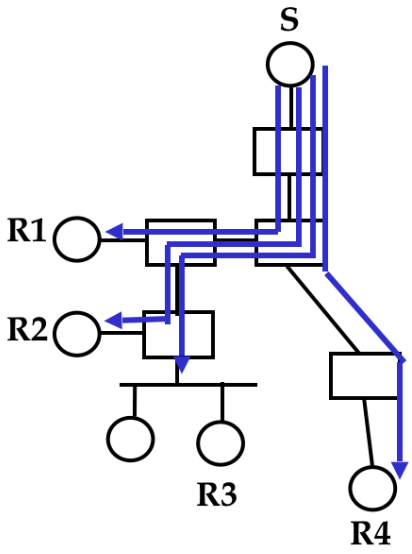
becomes less congested. This gets calculated and then traffic is sent to path B. The inverse happens and traffic oscillates between paths A and B.

Q20. Explain the 'count-to-infinity' problem and why it happens in distance-vector routing algorithms (like Bellman-Ford in OSPF), then suggest an improvement. Mention a limitation of your suggestion.

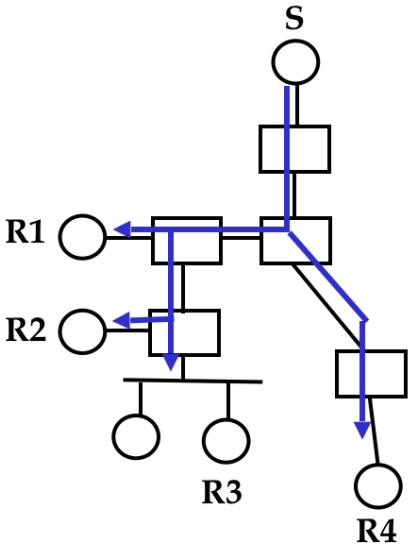
The count to infinity problem appears in distance vector routing, when loops appear in the routes. These routing loops can occur when two routers send each other updates in a pattern, referred to as how 'bad news travels slow'. A bad route will only update after the packet is being sent through, so the packet bounces back around and tries the other way but the new way is also slow. One possible solution is to implement routers so that any single junction never has only two choices of output to choose from; but this comes with the cost of more overhead and may not solve more complicated routing loops. The poisoned reverse solution / route poisoning solution suggests spreading news of route failure by poisoning the route (give it a value of infinity). However, this proposal increases the size of routing announcements.

Q21. What is the advantage of using multicast over unicast? Give example and provide a quantitative analysis (i.e., numbers) to solidify your argument (you can use a drawing of the network topology).

The fundamental principle of multicast is multipoint to multipoint transmission, which is what the majority of internet applications today are reliant on. Multicast conserves bandwidth by replicating packets only when necessary, and requires less traversals than unicast (in the given example, only half as many). Multicast has applications in news, weather, web content, stock prices, sensors, etc.



Multiple unicasts
(18 link traversals)



Multicast
(9 link traversals)

Q22. [Extra points] Consider the network topology in the figure below with nodes x, y, z. Assume that distance-vector routing is used with Bellman-Ford algorithm. The link cost between x and y changes from a to a' , where a' >> a, and a' > c. Focus on nodes y and z, their routing table updates, and messages exchanged between them.

I. Derive an equation for the number of iterations needed for the routing to converge as a function of the link costs a, a', b and c.

The problem is an example of the count to infinity problem stated earlier. The number of iterations is c - (a + b + b). In the example where a, b, and c are 4, 1, and 50 respectively, the answer comes out to 50 - (4+1+1) = 44.

II. Suggest an improvement that would speed up the convergence and prevent looping in the above case. Explain your solution and how it works.

A suggested improvement is making the bad path unavailable by calculating its cost as impossibly high (infinity). This prevents it from being incorporated in the next round of calculations, and so the transmission just takes the single route that's available.

Q23. [Extra points] Mention the two main reasons why broadcast-and-prune multicast does not scale.

Broadcast and prune multicast is one of two types of multicast (the other being explicit join). A subtype of broadcast and prune is MOSPF (multicast OSPF) and DVMRP (distance vector multicast routing protocol). However, this cannot be scaled everywhere because MOSPF is limited only to AS where OSPF is used, and by flooding they limit their usability. MOSPF and DVMRP flood membership info and packets respectively, which may not be ideal for certain AS, and is extremely congesting and cumbersome for larger networks, limiting this approach to smaller networks.

Q24. [Extra points] Mention two reasons why an optimal RP placement is not used (or needed) in PIM-SM.

As mentioned earlier, the rendezvous problem is an NP complete problem when shifting between the shared and shortest path trees. Optimally choosing the RP is hard, but PIM_SM sidesteps this problem. Instead, the RP is just the tree center per group. So the tree center is where members and senders meet, and where explicit joins happen, and where packets are sent. This is a sometimes suboptimal, but overall more efficient compromise of combining the trees but avoiding the RP issue.