Chapter 4 Problems

Problem 1

a) Data destined to host H3 is forwarded through interface 3

Destination Address	Link Interface
H3	3

b) No, because forwarding rule is only based on destination address.

Problem 2

a) No, you can only transmit one packet at a time over a shared bus.

b) No, as discussed in the text, only one memory read/write can be done at a time over the shared system bus.

c) No, in this case the two packets would have to be sent over the same output bus at the same time, which is not possible.

Problem 3

a) (n-1)D
b) (n-1)D
c) 0

Problem 4

The minimal number of time slots needed is 3. The scheduling is as follows.
Slot 1: send X in top input queue, send Y in middle input queue.
Slot 2: send X in middle input queue, send Y in bottom input queue
Slot 3: send Z in bottom input queue.

Largest number of slots is still 3. Actually, based on the assumption that a non-empty input queue is never idle, we see that the first time slot always consists of sending X in the top input queue and Y in either middle or bottom input queue, and in the second time slot, we can always send two more datagram, and the last datagram can be sent in third time slot.

NOTE: Actually, if the first datagram in the bottom input queue is X, then the worst case would require 4 time slots.

Problem 5

a) 112311231123…

b) 112112112…

Problem 6

a)

Packet Index	Leave queue time	Delay	Average Delay
2	2	2	3.091
3	3	2	
4	4	3	
5	6	3	
6	5	3	
7	7	4	
8	8	3	
9	9	4	
10	10	3	
11	11	3	
12	12	4	

b)

b)

Packet Index	Leave queue time	Delay	Average Delay
2	3	3	3.091
3	2	1	
4	7	6	
5	4	1	
6	8	6	
7	5	2	
8	10	5	

Packet Index	Leave queue time	Delay	Average Delay
9	6	1	
10	11	4	
11	9	1	
12	12	4	

c)

Packet Index	Leave queue time	Delay	Average Delay
2	3	3	3.091
3	5	4	
4	2	1	
5	4	1	
6	7	5	
7	6	3	
8	8	3	
9	10	5	
10	12	5	
11	9	1	
12	11	3	

d)

Packet Index	Leave queue time	Delay	Average Delay
2	3	3	3.091
3	2	1	
4	6	5	
5	4	1	

Packet Index	Leave queue time	Delay	Average Delay
6	8	6	
7	5	2	
8	10	5	
9	7	2	
10	11	4	
11	9	1	
12	12	4	

e) All average delay remains the same, no matter what algorithm is used.

Problem 7

a) They leaves at time 5, 6, 2, 4, 3, 7, 8, 10, 11, 9, 12 respectively.
b) They leaves at time 2, 4, 3, 7, 5, 6, 8, 10, 11, 9, 12 respectively.
c) They leaves at time 2, 3, 4, 10, 7, 5, 6, 8, 9, 12, 11 respectively.

Problem 8

a)

Prefix Match	Link Interface
11100000 00	0
11100000 01000000	1
1110000	2
11100001 1	3
otherwise	3

b) Prefix match for first address is 5th entry: link interface 3
Prefix match for second address is 3nd entry: link interface 2
Prefix match for third address is 4th entry: link interface 3

Problem 9

Destination Address Range	Link Interface
00000000 through 00111111	0
01000000 through	1

01011111

01100000 through 01111111	2
10000000 through 10111111	2
11000000 through 11111111	3

number of addresses for interface 0 = $2^6 = 64$

number of addresses for interface 1 = $2^5 = 32$

number of addresses for interface 2 = $2^6 + 2^5 = 64 + 32 = 96$

number of addresses for interface 3 = $2^6 = 64$

Problem 10

Destination Address Range	Link Interface
11000000 through (32 addresses) 11011111	0
10000000 through(64 addresses) 10111111	1
11100000 through (32 addresses) 11111111	2
00000000 through (128 addresses) 01111111	3

Problem 11

223.1.17.0/26
223.1.17.128/25
223.1.17.192/28

Problem 12

Destination Address	Link Interface
200.23.16/21	0
200.23.24/24	1
200.23.24/21	2
otherwise	3

Problem 13

Destination Address	Link Interface
11100000 00 (224.0/10)	0
11100000 01000000 (224.64/16)	1
1110000 (224/8)	2
11100001 1 (225.128/9)	3
otherwise	3

Problem 14

Any IP address in range 128.119.40.128 to 128.119.40.191

Four equal size subnets: 128.119.40.64/28, 128.119.40.80/28, 128.119.40.96/28, 128.119.40.112/28

Problem 15

From 214.97.254/23, possible assignments are

a) Subnet A: 214.97.255/24 (256 addresses)
Subnet B: 214.97.254.0/25 - 214.97.254.0/29 (128-8 = 120 addresses)
Subnet C: 214.97.254.128/25 (128 addresses)

Subnet D: 214.97.254.0/31 (2 addresses)
Subnet E: 214.97.254.2/31 (2 addresses)
Subnet F: 214.97.254.4/30 (4 addresses)

b) To simplify the solution, assume that no datagrams have router interfaces as ultimate destinations. Also, label D, E, F for the upper-right, bottom, and upper-left interior subnets, respectively.

Router 1

Longest Prefix Match	Outgoing Interface
11010110 01100001 11111111	Subnet A
11010110 01100001 11111110 0000000	Subnet D
11010110 01100001 11111110 000001	Subnet F

Router 2

Longest Prefix Match	Outgoing Interface
11010110 01100001 11111111 0000000	Subnet D
11010110 01100001 11111110 0	Subnet B
11010110 01100001 11111110 0000001	Subnet E

Router 3

Longest Prefix Match	Outgoing Interface
11010110 01100001 11111111 000001	Subnet F
11010110 01100001 11111110 0000001	Subnet E
11010110 01100001 11111110 1	Subnet C

Problem 16

The IP address blocks of Polytechnic Institute of New York University are:

NetRange: 128.238.0.0 - 128.238.255.255
CIDR: 128.238.0.0/16

The IP address blocks Stanford University are:
NetRange: 171.64.0.0 - 171.67.255.255
CIDR: 171.64.0.0/14

The IP address blocks University of Washington are:
NetRange: 140.142.0.0 - 140.142.255.255
CIDR: 140.142.0.0/16

No, the whois services cannot be used to determine with certainty the geographical location of a specific IP address.

www.maxmind.com is used to determine the locations of the Web servers at Polytechnic Institute of New York University, Stanford University and University of Washington.

Locations of the Web server at Polytechnic Institute of New York University is

Hostname	Country Code	Country Name	Region	Region Name	City	Postal Code	Latitude	Longitude	ISP	Organization	Metro Code	Area Code
128.238.24.30	US	United States	NY	New York	Brooklyn	11201	40.6944	-73.9906	Polytechnic University	Polytechnic University	501	718

Locations of the Web server Stanford University is

Hostname	Country Code	Country Name	Region	Region Name	City	Postal Code	Latitude	Longitude	ISP	Organization	Metro Code	Area Code
171.64.13.26	US	United States	CA	California	Stanford	94305	37.4178	-122.1720	Stanford University	Stanford University	807	650

Locations of the Web server at University of Massachusetts is

Hostname	Country Code	Country Name	Region	Region Name	City	Postal Code	Latitude	Longitude	ISP	Organization	Metro Code	Area Code
128.119.103.148	US	United States	MA	Massachusetts	Amherst	01003	42.3896	-72.4534	University of Massachusetts	University of Massachusetts	543	413

Problem 17

MP3 file size = 5 million bytes. Assume the data is carried in TCP segments, with each TCP segment also having 20 bytes of header. Then each datagram can carry 1500-40=1460 bytes of the MP3 file

Number of datagrams required $= \left\lceil \frac{5 \times 10^6}{1460} \right\rceil = 3425$. All but the last datagram will be 1,500 bytes; the last datagram will be 960+40 = 1000 bytes. Note that here there is no fragmentation – the source host does not create datagrams larger than 1500 bytes, and these datagrams are smaller than the MTUs of the links.

Problem 18

a) Home addresses: 192.168.1.1, 192.168.1.2, 192.168.1.3 with the router interface being 192.168.1.4

b)

NAT Translation Table

WAN Side	LAN Side
24.34.112.235, 4000	192.168.1.1, 3345
24.34.112.235, 4001	192.168.1.1, 3346
24.34.112.235, 4002	192.168.1.2, 3445
24.34.112.235, 4003	192.168.1.2, 3446
24.34.112.235, 4004	192.168.1.3, 3545
24.34.112.235, 4005	192.168.1.3, 3546

Problem 19

a) Since all IP packets are sent outside, so we can use a packet sniffer to record all IP packets generated by the hosts behind a NAT. As each host generates a sequence of IP

packets with sequential numbers and a distinct (very likely, as they are randomly chosen from a large space) initial identification number (ID), we can group IP packets with consecutive IDs into a cluster. The number of clusters is the number of hosts behind the NAT.

For more practical algorithms, see the following papers.

"A Technique for Counting NATted Hosts", by Steven M. Bellovin, appeared in IMW'02, Nov. 6-8, 2002, Marseille, France.
"Exploiting the IPID field to infer network path and end-system characteristics."
Weifeng Chen, Yong Huang, Bruno F. Ribeiro, Kyoungwon Suh, Honggang Zhang, Edmundo de Souza e Silva, Jim Kurose, and Don Towsley.
PAM'05 Workshop, March 31 - April 01, 2005. Boston, MA, USA.

b) However, if those identification numbers are not sequentially assigned but randomly assigned, the technique suggested in part (a) won't work, as there won't be clusters in sniffed data.

Problem 20

It is not possible to devise such a technique. In order to establish a direct TCP connection between Arnold and Bernard, either Arnold or Bob must initiate a connection to the other. But the NATs covering Arnold and Bob drop SYN packets arriving from the WAN side. Thus neither Arnold nor Bob can initiate a TCP connection to the other if they are both behind NATs.

Problem 21

S2 Flow Table	
Match	**Action**
Ingress Port = 1; IP Src = 10.3.*.*; IP Dst = 10.1.*.*	Forward (2)
Ingress Port = 2; IP Src = 10.1.*.*; IP Dst = 10.3.*.*	Forward (1)
Ingress Port = 1; IP Dst = 10.2.0.3 Ingress Port = 2; IP Dst = 10.2.0.3 Ingress Port = 1; IP Dst = 10.2.0.4 Ingress Port = 2; IP Dst = 10.2.0.4	Forward (3) Forward (3) Forward (4) Forward (4)
Ingress Port = 4 Ingress Port = 3	Forward (3) Forward (4)

Problem 22

S2 Flow Table	
Match	**Action**
Ingress Port = 3; IP Dst = 10.1.*.* Ingress Port = 3; IP Dst = 10.3.*.*	Forward (2) Forward (2)
Ingress Port = 4; IP Dst = 10.1.*.* Ingress Port = 4; IP Dst = 10.3.*.*	Forward (1) Forward (1)

Problem 23

S1 Flow Table	
Match	**Action**
IP Src = 10.2.*.*; IP Dst = 10.1.0.1	Forward (2)
IP Src = 10.2.*.*; IP Dst = 10.1.0.2	Forward (3)
IP Src = 10.2.*.*; IP Dst = 10.3.*.*	Forward (1)

S3 Flow Table	
Match	**Action**
IP Src = 10.2.*.*; IP Dst = 10.3.0.6	Forward (1)
IP Src = 10.2.*.*; IP Dst = 10.3.0.5	Forward (2)
IP Src = 10.2.*.*; IP Dst = 10.1.*.*	Forward (3)

Problem 24

S2 Flow Table	
Match	**Action**
IP Src = 10.1.0.1; IP Dst = 10.2.0.3	Forward (3)

IP Src = 10.1.0.1; IP Dst = 10.2.0.4	Forward (4)
IP Src = 10.3.0.6; IP Dst = 10.2.0.3	Forward (3)
IP Src = 10.3.0.6; IP Dst = 10.2.0.4	Forward (4)

S2 Flow Table	
Match	**Action**
IP Src =.*.*.*.*; IP Dst = 10.2.0.3; port = TCP	Forward (3)
IP Src =.*.*.*.*; IP Dst = 10.2.0.4; port = TCP	Forward (4)

S2 Flow Table	
Match	**Action**
IP Src =.*.*.*.*; IP Dst = 10.2.0.3	Forward (3)

S2 Flow Table	
Match	**Action**
IP Src = 10.1.0.1; IP Dst = 10.2.0.3; port = UDP	Forward (3)

Problem 25

We consider ICMP as a network-layer protocol, as ICMP packets are encapsulated by an IP packet.

Chapter 5. Review Questions.

1. Per-router control means that a routing algorithm runs in each and every router; both forwarding and routing function are constrained within each router. Each router has a

routing component that communicates with the routing components in other routers to compute the values for its forwarding table. In such cases, we say that the network control and data planes are implemented monolithically because each router works as an independent entity that implements its own control and data planes.

2. Logically centralized control means that a logically central routing controller computes and distributes the forwarding tables to be used by each and every router, and each router does not compute its forwarding table, unlike the per-router control. In the case of logically centralized control, the data plane and control plane are implemented in separate devices; the control plane is implemented in a central server or multiple servers, and the data plane is implemented in each router.

3. A centralized routing algorithm computes the least-cost path between a source and destination by using complete, global knowledge about the network. The algorithm needs to have the complete knowledge of the connectivity between all nodes and all links' costs. The actual calculation can be run at one site or could be replicated in the routing component of each and every router. A distributed routing algorithm calculates the lease-cost path in an iterative, distributed manner by the routers. With a decentralized algorithm, no node has the complete information about the costs of all network links. Each node begins with only the knowledge of the costs of its own directly attached links, and then through an iterative process of calculation and information exchange with its neighboring nodes, a node gradually calculates the least-cost path to a destination or a set of destinations.

   OSPF protocol is an example of centralized routing algorithm, and BGP is an example of a distributed routing algorithm.

4. Link state algorithms: Computes the least-cost path between source and destination using complete, global knowledge about the network. Distance-vector routing: The calculation of the least-cost path is carried out in an iterative, distributed manner. A node only knows the neighbor to which it should forward a packet in order to reach given destination along the least-cost path, and the cost of that path from itself to the destination.

5. The count-to-infinity problem refers to a problem of distance vector routing. The problem means that it takes a long time for a distance vector routing algorithm to converge when there is a link cost increase. For example, consider a network of three nodes x, y, and z. Suppose initially the link costs are c(x,y)=4, c(x,z)=50, and c(y,z)=1. The result of distance-vector routing algorithm says that z's path to x is z→y→ x and the cost is 5(=4+1). When the cost of link (x,y) increases from 4 to 60, it will take 44 iterations of running the distance-vector routing algorithm for node z to realize that its new least-cost path to x is via its direct link to x, and hence y will also realize its least-cost path to x is via z.

6. No. Each AS has administrative autonomy for routing within an AS.

7. Policy: Among ASs, policy issues dominate. It may well be important that traffic originating in a given AS not be able to pass through another specific AS. Similarly, a

given AS may want to control what transit traffic it carries between other ASs. Within an AS, everything is nominally under the same administrative control and thus policy issues a much less important role in choosing routes with in AS.

Scale: The ability of a routing algorithm and its data structures to scale to handle routing to/among large numbers of networks is a critical issue in inter-AS routing. Within an AS, scalability is less of a concern. For one thing, if a single administrative domain becomes too large, it is always possible to divide it into two ASs and perform inter-AS routing between the two new ASs.

Performance: Because inter-AS routing is so policy oriented, the quality (for example, performance) of the routes used is often of secondary concern (that is, a longer or more costly route that satisfies certain policy criteria may well be taken over a route that is shorter but does not meet that criteria). Indeed, we saw that among ASs, there is not even the notion of cost (other than AS hop count) associated with routes. Within a single AS, however, such policy concerns are of less importance, allowing routing to focus more on the level of performance realized on a route.

8. False.

With OSPF, a router broadcasts its link-state information to all other routers in the autonomous system to which it belongs, not just to its neighboring routers. This is because with OSPF, each router needs to construct a complete topological map of the entire AS and then locally runs Dijkstra's shortest-path algorithm to determine its least-cost paths to all other nodes in the same AS.

9. An area in an OSPF autonomous system is refers to a set of routers, in which each router broadcasts its link state to all other routers in the same set. An OSPF AS can be configured hierarchically into multiple areas, with each area running its own OSPF link-state routing algorithm. Within each area, one or more area border routers are responsible for routing packets outside the area. The concept of area is introduced for scalability reason, i.e., we would like to build a hierarchical routing for a large scale OSPF AS, and an area is an important building block in hierarchical routing.

10. A subnet is a portion of a larger network; a subnet does not contain a router; its boundaries are defined by the router and host interfaces. A prefix is the network portion of a CDIRized address; it is written in the form a.b.c.d/x ; A prefix covers one or more subnets. When a router advertises a prefix across a BGP session, it includes with the prefix a number of BGP attributes. In BGP jargon, a prefix along with its attributes is a BGP route (or simply a route).

11. Routers use the AS-PATH attribute to detect and prevent looping advertisements; they also use it in choosing among multiple paths to the same prefix. The NEXT-HOP attribute indicates the IP address of the first router along an advertised path (outside of the AS receiving the advertisement) to a given prefix. When configuring its forwarding table, a router uses the NEXT-HOP attribute.

12. A tier-1 ISP B may not to carry transit traffic between two other tier-1 ISPs, say A and C, with which B has peering agreements. To implement this policy, ISP B would not advertise to A routes that pass through C; and would not advertise to C routes that pass through A.

13. False.
    A BGP router can choose not to add its own identity to the received path and then send that new path on to all of its neighbors, as BGP is a policy-based routing protocol. This can happen in the following scenario. The destination of the received path is some other AS, instead of the BGP router's AS, and the BGP router does not want to work as a transit router.

14. The communication layer is responsible for the communication between the SDN controller and those controlled network devices, via a protocol such as OpenFlow. Through this layer, an SDN controller controls the operation of a remote SDN-enabled switch, host, or other devices, and a device communicates locally-observed events (e.g., a message indicating a link failure) to the controller.

    The network-wide state-management layer provides up-to-date information about state a network's hosts, links, switches, and other SDN-controlled devices. A controller also maintains a copy of the flow tables of the various controlled devices.

    The network-control application layer represents the brain of SDN control plane. The applications at this layer use the APIs provided by a SDN controller to specify and control the data plane in the network devices. For example, a routing network-control application might determine the end-end paths between sources and destinations. Another network application might perform access control.

15. I would implement a new routing protocol at the SDN's network-control application layer, as this is the layer where a routing protocol determines the end-to-end paths between sources and destinations.

16. The following is a list of types of messages flow across a SDN controller's southbound from the controller to the controlled devices. The recipient of these messages is a controlled packet switch.
    - Configuration. This message allows the controller to query and set a switch's configuration parameters.
    - Modify-state. This message is used by a controller to add/delete or modify entries in the switch's flow table, and to set switch port properties.
    - Read-state. This message is used by a controller to collect statistics and counter values from the switch's flow table and ports.
    - Send-packet. This message is used by the controller to send a specific packet out of a specified port at the controlled switch.

    There are also messages that network-control applications (as senders) send to the controller across the northbound interfaces, for example, messages to read/write network state and flow tables within the state-management layer of the controller.

17. Two types of messages from a controlled device to a controller:
    - Flow-removed message. Its purpose is to inform the controller that a flow table entry has been removed, for example, by a timeout or as the result of a received modify-state message.
    - Port-status message. Its purpose is to inform the controller of a change in port status.

    Two types of messages from a controller to a controlled device:
    - Modify-state. The purpose is to add/delete or modify entries in the switch's flow table, and to set switch port properties.
    - Read-state. The purpose is to collect statistics and counter values rom the switch's flow table and ports.

18. The service abstraction layer allows internal network service applications to communicate with each other. It allows controller components and applications to invoke each other's services and to subscribe to events they generate. This layer also provides a uniform abstract interface to the specific underlying communications protocols in the communication layer, including OpenFlow and SNMP.

19. Echo reply (to ping), type 0, code 0
    Destination network unreachable, type 3 code 0
    Destination host unreachable, type 3, code 1.
    Source quench (congestion control), type 4 code 0.

20.
ICMP warning message (type 11 code 0) and a destination port unreachable ICMP message (type 3 code 3).

21.
A *managing server* is an application, typically with a human in the loop, running in a centralized network management station in a network operation center. It controls the collection, processing, analysis, and/or display of network management information. Actions are initiated in a managing server to control network behavior and a network administrator uses a managing server to interact with the network's devices.

A *managed device* is a piece of network equipment (including its software) that resides on a managed network. A managed device might be a host, router, switch, middlebox, modem, thermometer, or other network-connected device.

A *network management agent* is a process running in a managed device that communicates with a managing server, taking local actions at the managed device under the command and control of the managing server.

Management Information Base (MIB) collects the information associated with those managed objects in a managed network. A MIB object might be a counter, such as the number of IP datagrams discarded at a router due to errors in an IP datagram header, or the number of UDP segments received at a host, or the status information such as whether a particular device is functioning correctly.

22.
GetRequest is a message sent from a managing server to an agent to request the value of one or more MIB objects at the agent's managed device.

SetRequest is a message used by a managing server to set the value of one or more MIB objects in a managed device.

23.
A SNMP trap message is generated as a response to an event happened on a managed device for which the device's managing server requires notification. It is used for notifying a managing server of an exceptional situation (e.g., a link interface going up or down) that has resulted in changes to MIB object values.