1. (PD) One of the central problems faced by a protocol such as MIME is the vast number of data formats available. Consult the MIME RFC to find out how MIME deals with new or system-specific image and text formats.

2. (PD) Consult the MIME RFC to find out how `base64` encoding handles binary data of a length not evenly divisible by three bytes.

3. (PD) Suppose a very large website wants a mechanism by which clients access whichever of multiple HTTP servers is "closest" by some suitable measure.

   (a) Discuss developing a mechanism within HTTP to do this.

   (b) Discuss developing a mechanism within DNS to do this.

   Can either approach be made to work without upgrading the browser?

4. (PD) Get the WSDL for some SOAP-style Web Service and choose an operation. In the messages that implement that operation, identify the fields.

5. (PD) DNS servers also allow reverse lookup; given an IP address 128.112.169.4, it is reversed into a text string 4.169.112.128.inaddr.arpa and looked up using DNS PTR records (which form a hierarchy of domains analogous to that for the address domain hierarchy). Suppose you want to authenticate the sender of a packet based on its host name and are confident that the IP *address* is genuine. Explain the insecurity in converting the source address to a name as above and then comparing this name to a list of trusted hosts. (Hint: Whose DNS servers would you be trusting?)

6. (PD) What is the relationship between a domain name (e.g., cs.princton.edu) and an IP subnet number such as 192.12.69.0? Do all hosts on the subnet have to be identified by the same name server? What about reverse lookup?

7. (PD) What DNS cache issues are involved in changing the IP address of, say, a web server host name? How might these be minimized?

8. (KR) Consider distributing a file $F = 15$ Gbits to N peers. The server has an upload rate of $\mu_s = 30$ Mbps, and each peer has a download rate of $d_i = 2$ Mbps and an upload rate of μ. For $N = 10$, 100, and 1000 and $\mu = 500$ Kbps, 800 Kbps, and 2 Mbps, prepare a table giving the minimum distribution time for each of the combinations of N and μ for both client-server distribution and P2P distribution.

9. (KR) Consider an overlay network with N active peers, with each pair of peers having an active TCP connection. Additionally, suppose that the TCP connections pass through a total of M routers.

   (a) How many nodes and edges are in the corresponding overlay network?

10. (KR)Suppose Bob joins a BitTorrent, but he does not want to upload data to any other peers (so called free-riding).

    (a) Bob claims that he can receive a complete copy of the file that is shared by the swarm. Is Bob's claim possible? Why or why not?

    (b) Bob further claims that he can further make his "free-riding" more efficient by using a collection of multiple computers (with distinct IP addresses) in the computer lab in his department. How can he do that?

11. (KR) In this problem, we are interested in finding out the efficiency of a BitTorrent-like P2P file sharing system. Consider two peers Bob and Alice. They join a torrent with M peers in total (including Bob and Alice) that are sharing a file consisting of N chunks. Assume that at a particular time t, the chunks that a peer has are uniformly at random chosen from all N chunks, and no peer has all N chunks. Answer the following questions.

(a) What is the probability that Bob has all the chunks that Alice has, given that the numbers of chunks that Bob and Alice have are denoted by n_b and n_a?

(b) Remove part of the conditioning in part (a) to find out the probability that Bob has all the chunks that Alice has, given that Alice has n_a chunks.

(c) Suppose that each peer in BitTorrent has 5 neighbors. What is the probability that Bob has data that is of interest to at least one of his five neighbors?

12. (KR)A circular DHT consists of peers 1, 4, 5, 6, 9, 10, 12, and 15. Suppose that a new peer 7 wants to join the DHT and peer 7 initially only knows peer 15's address. What steps are taken?

13. (KR)In the circular DHT example above, suppose that peer 4 learns that peer 6 has left.

    (a) How does peer update its successor state information?
    (b) Which peer is now its first successor?
    (c) Which peer is now its second successor?

14. (KR) As DHTs are overlay networks, they may not necessarily match the underlay physical network well in the sense that two neighboring peers might be physically very far away; for example, one peer could be in Asia and its neighbor could be in North America.

    (a) If we randomly and uniformly assign identifiers to newly joined peers, would this assignment scheme cause such a mismatch?

    (b) How would such a mismatch affect the DHT's performance?