

**THE UNIVERSITY OF WESTERN ONTARIO
LONDON CANADA**

**COMPUTER SCIENCE 357b
MIDTERM EXAMINATION
FEBRUARY 12, 2005
2 HOURS**

NAME: _Marking Scheme_____

STUDENT NUMBER: _____

Question

1-25. _____

26. _____

27. _____

28. _____

29. _____

30. _____

31. _____

32. _____

33. _____

TOTAL _____

(Out of 150 marks)

There are no cheat sheets, books, or other reference materials allowed for this exam. No calculators or other electronic devices are permitted either.

Part I -- Multiple Choice, True/False -- Choose the best answer from the choices given. Circle your answer on the paper, and fill in the answer on the Scantron form. [50 marks total, 2 marks each]

1. The transport and application layers of the Internet protocol stack are implemented in end systems, but not in routers in the network core.
 a. True.
 b. False.
2. The SMTP protocol is used to:
 a. To define the format of multimedia messages and extensions.
 b. To transfer messages from one mail server to another.
 c. To transfer messages from a mail server to a user agent.
 d. All of the above.
 e. None of the above.
3. Suppose a client sends an HTTP request message with the If-modified-since: header. Suppose the object in the server has not changed since the last time that client retrieved the object. Then the server will send a response message with the status code:
 a. 304 Not Modified.
 b. 404 Not Found.
 c. 200 OK.
 d. 403 Permission Denied.
 e. None of the above.
4. An HTTP server typically sends multiple objects in a web page to clients within a multipart MIME message.
 a. True.
 b. False.
5. UDP has which of the following characteristics:
 a. A three-way handshake for connection establishment.
 b. Connection state at the server.
 c. A regulated send rate.
 d. All of the above.
 e. None of the above.
6. The transfer of an HTML file from one host to another is:
 a. Loss-intolerant and time insensitive.
 b. Loss-intolerant and time sensitive.
 c. Loss-tolerant and time insensitive.
 d. Loss-tolerant and time sensitive.
 e. None of the above.

7. Which of the following is an example of a peer-to-peer application?
- a. SMTP e-mail.
 - b. DNS.
 - c. FTP file transfer application.
 - d. HTTP web application.
 - e. None of the above.
8. Which of the following is used to contain an Internet standard?
- a. RFC.
 - b. IETF.
 - c. DNS.
 - d. PPP.
 - e. None of the above.
9. In theory, circuit switching supports more users on a network than packet switching.
- a. True.
 - b. False.
10. HFC is a technology typically deployed by telephone companies for high speed network access.
- a. True.
 - b. False.
11. The structure of the Internet is roughly:
- a. Hierarchical.
 - b. Linear.
 - c. Spherical.
 - d. Cubical.
 - e. None of the above.
12. Nodal processing typically provides the largest delays in a network.
- a. True.
 - b. False.
13. The Internet address for a host is sufficient to address data to an application running on that host.
- a. True.
 - b. False.
14. Which of the following does TCP not provide:
- a. Reliable, in order data delivery.
 - b. Flow control.
 - c. Congestion control.
 - d. Minimum bandwidth guarantees.
 - e. TCP provides all of the above.

15. All HTTP web servers use port 80 to listen for client requests.
- a. True.
 - b. False.
16. HTTP request and response messages are not humanly readable.
- a. True.
 - b. False.
17. The “Date” and “Last-Modified” response headers in HTTP convey exactly the same information.
- a. True.
 - b. False.
18. FTP does not encrypt passwords it sends to FTP servers to prevent eavesdropping and discovery of the passwords.
- a. True.
 - b. False.
19. All SMTP e-mail messages must be in 8-bit ASCII.
- a. True.
 - b. False.
20. A root DNS server contains all of the host-address mappings for every host on the Internet.
- a. True.
 - b. False.
21. A socket is a unidirectional communication mechanism.
- a. True.
 - b. False.
22. For every write() on a TCP socket, there must be a corresponding read() on the other end.
- a. True.
 - b. False.
23. For every sendto() on a UDP socket, there must be a corresponding recvfrom() on the other end.
- a. True.
 - b. False.

24. Which of the following are benefits of an ISP increasing the number of connections to other ISPs in the Internet?
- a. Improved reliability.
 - b. Increased bandwidth.
 - c. Reduced delays.
 - d. All of the above.
 - e. None of the above.
25. Which of the following approaches to peer-to-peer networking works without having a bootstrap node?
- a. The centralized directory approach.
 - b. The decentralized directory approach.
 - c. The query flooding approach.
 - d. All of the above.
 - e. None of the above.

26. The following question parts deal with packet switching. [12 marks total]

- a. Explain how pipelining in a packet switching network can improve performance and reduce delay. [4 marks]

Performance is improved because each link can now work in parallel transmitting part of the total data being transmitted, instead of transmitting all of the data in a single massive packet. This helps improve utilization and the throughput of the application.

Delay is improved for similar reasons. Since each link can process packets of data in parallel now, we are not forced into what was in essence a stop-and-wait approach to transmission as each node stores and then forwards the packet. With the links all working at once, delay can be reduced, as demonstrated on page 26 of the notes on Chapter 1.

- b. Give two advantages to breaking up messages into smaller packets in a packet switching network. [4 marks]

With messages broken into smaller packets, we get the performance and delay benefits as discussed above. Error handling is also now less painful, as a loss event loses less data, and the costs of retransmission are reduced as well.

- c. Why does breaking up messages into smaller packets in a packet switching network increase overhead? [4 marks]

Overhead is increased in many ways. First, with more packets, routers have more work and more decisions to make. Second, with smaller packets, the percentage of space in each packet used for header information as opposed to the real data itself increases. Since there are now more packets, the total amount of header information sent increases substantially, and the network is spending more time talking about the data than actually sending the data itself. Consequently, network overhead can increase substantially.

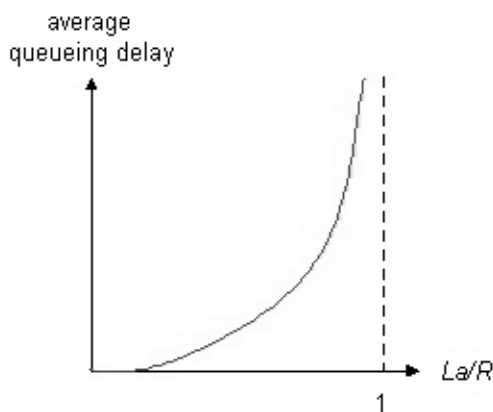
27. The following question parts deal with delays in a network. [12 marks total]
- a. If R is the outgoing link bandwidth from a router and L is the packet size, how would you calculate transmission delay at this router? [3 marks]

The transmission delay in this case is L/R .

- b. If the length of a physical link is D , and the speed at which data is propagated through the medium is S , how would you calculate the propagation delay? [3 marks]

The propagation delay in this case is D/S .

- c. Define traffic intensity mathematically as was done in class. What happens as traffic intensity approaches 0, approaches 1, and exceeds 1? Use a graph, if you feel it makes a better explanation. [6 marks]



Traffic intensity is $\lambda a/R$, where L is packet length, R is transmission rate (link bandwidth), and a is arrival rate of packets.

- $\lambda a/R \sim 0$: average queueing delay small
- $\lambda a/R \rightarrow 1$: delays become large
- $\lambda a/R > 1$: more “work” arriving than can be serviced, average delay infinite!

28. The following questions deal with authorization in the web. [18 marks total]

- a. Assuming no cookies, describe how authorization works in the HTTP protocol. [6 marks]

When a request is first made to a server, the client is notified (through a 401 authorization required message) that it needs to provide credentials to access the requested object. The client acquires the necessary credentials (usually in the form of a user name/password) and repeats the request, this time with an Authorization header line. If the credentials are acceptable, the server provides the requested object. Since HTTP is stateless, this Authorization header must be provided on each and every access to that object.

- b. If HTTP is stateless, why is a user prompted for authorization on the first access to a site by their web browser, and not on subsequent accesses? (Assume no cookies are used.) [6 marks]

The Authorization header is, in fact, being provided to the server on each subsequent access to the website. The web browser, however, is being smart, and remembering the authorization credentials entered by the user, and is automatically providing them to the server for you. So, the user is only prompted once, and the browser provides the authorization credentials every time on your behalf. It is a browser implementation decision, and not part of the protocol.

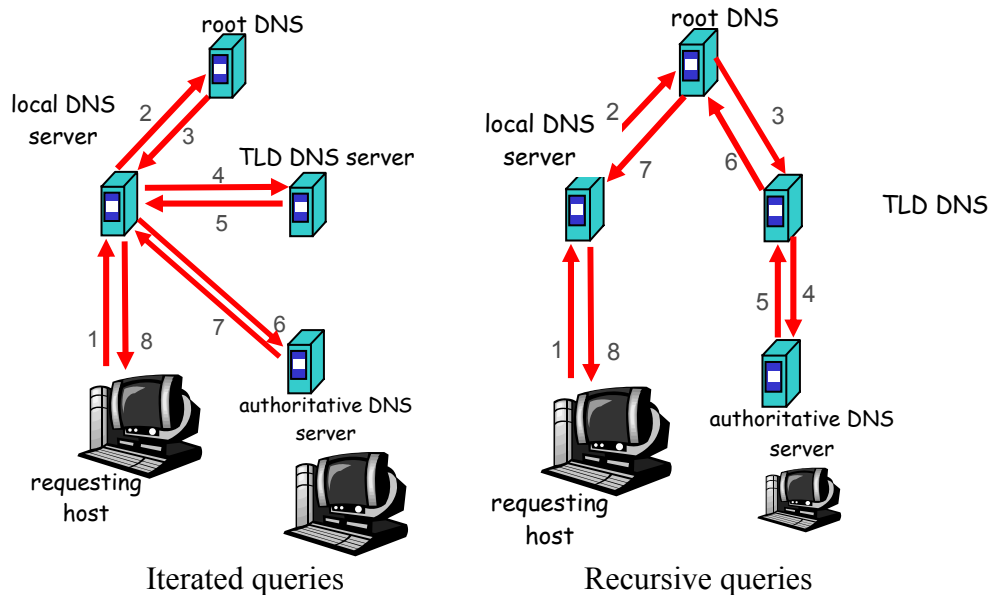
- c. How is HTTP authorization different when cookies are used? Explain your answer. [6 marks]

When cookies are available, a web site can issue you a cookie after your first authorization attempt on the website. Since you, and only you have this cookie, this is as good as requiring you to provide authorization credentials on each access.

So, after an initial authorization, the client receives a cookie with this response. On every subsequent access to the site, this cookie is provided with the request after the client retrieves this cookie from a local cookie repository. When the server sees the cookie, it consults some sort of back end cookie database, and authorizes the user this way. Since authorization is now contained in the cookie, and cookies are persistent across sessions, a user can launch a new browser and access the site without prompting for credentials.

29. The following questions deal with DNS. [18 marks total]

- a. Explain the differences between recursive and iterated DNS queries. Provide a diagram of each to help show these differences. [10 marks]



In a recursive DNS request, if the server receiving the request cannot respond directly itself, it forwards the request on to a server that can, which repeats the process over again. In an iterated DNS request, if the server cannot handle the request itself, it provides a redirection response, informing the requestor where it can find the answer itself; the server does not hunt for the information itself, but rather has the requestor do so itself.

- b. How can iterated DNS queries improve the overall performance of DNS? [4 marks]

Iterated request can improve overall performance by offloading the processing of requests from root and TLD servers to local servers. In recursive queries, root servers can be tied up ensuring the completion of numerous requests, which can result in a substantial decrease in performance. Iterated requests move that burden to local servers, and distributed the load more evenly throughout the Internet. With less work at the root servers, they can perform much faster.

- c. Explain how DNS caching can be used to increase the performance of DNS. What are the potential drawbacks involved in caching DNS records? [4 marks]

Caching improves performance by reducing the load on DNS servers on the Internet and by keeping copies of common resource records locally so that those requests do not take as long to satisfy. There is a risk, however, that cached records get stale and out of date, resulting in directions to wrong sites or bad addresses.

30. The following questions deal with socket programming. [12 marks total]
- a. Why does a call to `socket()` return an integer? [4 marks]

A call to `socket()` returns an integer because it is returning a file descriptor. A file descriptor is simply an integer index into the file descriptor table for the process. Everything that the process and operating system require for using that socket is linked from this file descriptor table entry. Consequently, nothing else is needed.

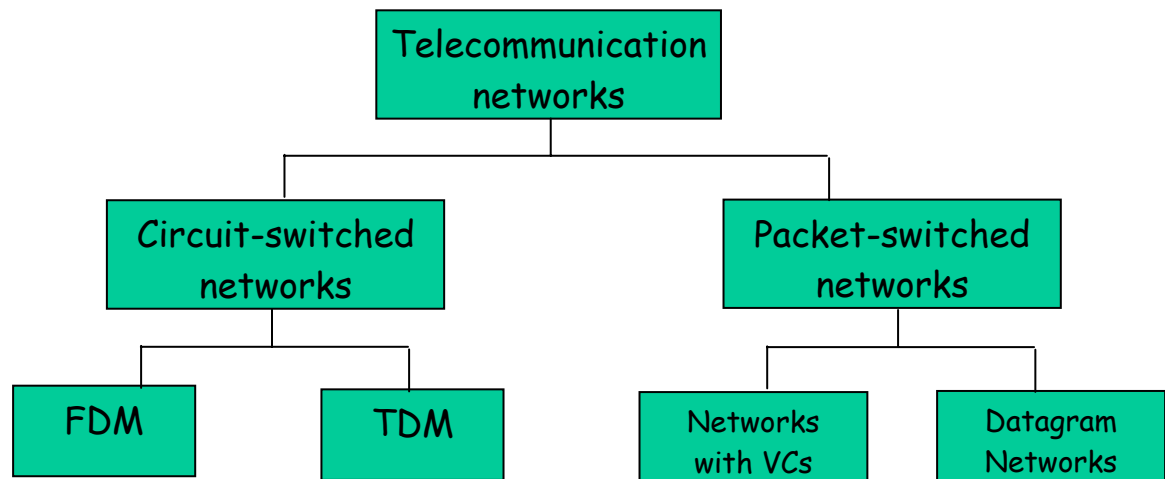
- b. When a UDP server receives a UDP datagram from a client, how does it know where to send a response to? [4 marks]

When a server receives a UDP datagram using `recvfrom()`, it also obtains a socket address structure that contains the IP address and port number of the sender of the data. Using this information, it knows exactly where the message originated from, and how to return a response back to the sender.

- c. When a TCP server receives a TCP message from a client, how does it know where to send a response to? [4 marks]

When a server receives a TCP message from a client, it must be receiving this message over a TCP connection. The transport layer maintains information in this connection for the remote end (the remote IP address and port number). This information was accessible when `accept()` accepted the connection. Consequently, when the server wishes to reply, it simply has to send the data back down the connection, and it will automatically be sent to the originator of the message as the connected socket has the required remote addressing information.

When examining the network core in our early overview of concepts in networking, we examined several networking approaches and produced a taxonomy for classifying networks. Draw this taxonomy in diagram form. [6 marks]



31. With all of the features provided by TCP, why do we even bother to have a UDP protocol? [4 marks]

There are several reasons:

- UDP is a bare-bones protocol, and allows the programmer to essentially build their own protocol to meet their own needs
- UDP does not have connection overhead, and can send data quicker without the RTT required to establish a connection
- UDP takes less resources (no connections to be maintained, etc.)
- The lack of flow control, congestion control, retransmissions make it better suited for multimedia data.
- UDP has a smaller header, making it more efficient.
- UDP supports broadcasting.

32. The following questions deal with content distribution. [18 marks]
- a. Suppose you work for a large company, and you want to increase performance to accessing web sites for as little cost as possible. Which would you recommend: increasing the amount of bandwidth on the corporate access link, or installing a proxy server web cache? Why? Would the web site access patterns of corporate users affect your decision? Why? [9 marks]

Installing a proxy server web cache would cost less in the long run, and reduce traffic over the corporate access link resulting in better performance. This is based on the assumption of course that the web cache results in a performance improvement, which requires company users to be accessing the same content at least a decent percentage of the time. If there are no common web objects being accessed by the users, the cache hit rate will be low, and all of the requests will have to go over the corporate access link anyways, causing a performance bottleneck. So, the web site access patterns could affect the decision. If access patterns result in what would be a low cache hit rate, the proxy server will not provide a performance improvement, and the best approach would be to increase the bandwidth on the access link, even if it is more costly.

- b. Suppose you work for a company that needs to distribute content to a very large number of users over the Internet. Describe in detail how a Content Distribution Network (CDN) would operate in connecting users with your content. [9 marks]

A CDN would populate its servers around the world with your content. To connect your users with the content maintained by the CDN, references to the content in your web pages would be modified to refer to the CDN content instead.

When the user downloads your web pages, they will get the links to the CDN site. When they need to access the data, DNS will need to resolve the address of the CDN server given. A smart DNS server in the CDN will automatically load balance client requests and direct them to a node in the CDN network that is closest to the user. The user then downloads the object from this node, reducing delay and spreading load evenly over the entire CDN.