Problem Set 2

R1:

List five nonproprietary Internet applications and the application-layer protocols that they use.

1. Application: e-mail application-layer protocol: SMTP   
2. Application: WEB application-layer protocol: HTTP   
3. Application: streaming multimedia application-layer protocol: HTTP (eg Youtube), RTP  
4. Application: file transfer application-layer protocol: FTP  
5. Application: remote terminal access application-layer protocol: Tenet

R5:

What information is used by a process running on one host to identify a process running on another host?

The Information used is the source port number, the destination port number and ip addresses.

R11:

Why do HTTP, FTP, SMTP, and POP3 run on top of TCP rather than on UDP?

tcp is more reliable than udp,udp may have failures or data loss, so we can't afford to have losses in http,smtp,pop3 and so on.Accurate data is very important is all this protocols. Also none of the Protocols actually require TCP as its Transport protocol. They simply require the use of a reliable protocol at the Transport Layer where tcp is the best right now.

P1:

True or false?

a. A user requests a Web page that consists of some text and three images.

For this page, the client will send one request message and receive four

response messages. False

b. Two distinct Web pages (for example, www.mit.edu/research.html

and www.mit.edu/students.html) can be sent over the same persistent

connection. True

c. With nonpersistent connections between browser and origin server, it is possible

for a single TCP segment to carry two distinct HTTP request messages. False

d. The Date: header in the HTTP response message indicates when the

object in the response was last modified. False

e. HTTP response messages never have an empty message body. False

P4:

Consider the following string of ASCII characters that were captured by Wireshark when the browser sent an HTTP GET message (i.e., this is the actual content of an HTTP GET message). The characters <cr><lf> are carriage return and line-feed characters (that is, the italized character string <cr> in the text below represents the single carriage-return character that was contained at that point in the HTTP header). Answer the following questions, indicating where in the HTTP GET message below you find the answer.

1. What is the URL of the document requested by the browser?

gaia.cs.umass.edu /cs453/index.html

1. What version of HTTP is the browser running?

HTTP 1.1

1. Does the browser request a non-persistent or a persistent connection?

Persistent

1. What is the IP address of the host on which the browser is running?

This information is not contained in an HTTP message.

1. What type of browser initiates this message? Why is the browser type needed in an HTTP request message?

Mozilla/5.0. The browser type information is needed by the server to send different versions of the same object to different types of browsers.

P9:

Consider Figure 2.12, for which there is an institutional network connected to the Internet. Suppose that the average object size is 850,000 bits and that the average request rate from the institution’s browsers to the origin servers is 16 requests per second. Also suppose that the amount of time it takes from when the router on the Internet side of the access link forwards an HTTP request until it receives the response is three seconds on average (see Section 2.2.5). Model the total average response time as the sum of the average access delay (that is, the delay from Internet router to institution router) and the average Internet delay.

1. Find the total average response time.

Delta = L/R = (850,000bits) / (15000000 bits/sec) = .0567sec

BetaDelta = (16 requests/sec)(.0567 sec/request) = 0.907

Average access delay is (.0567 sec) / (1 - .907) = .6secs

Total average response time is .6sec + 3sec = 3.6sec

1. Now suppose a cache is installed in the institutional LAN. Suppose the miss rate is 0.4. Find the total response time.

Traffic intensity on the access link is reduced by 60% since 60% of the requests are satisfied within the institutional network.

Average delay is .0567 / (1 – (.4)(.907)) = .089sec

Average response time is .089 + 3secs = 3.089 secs for cache misses (40% )

Real Average response time is (.6)(0sec) + (.4)(3.089) = 1.24secs.

So the average response time is reduced from 3.6 secs to 1.24 secs.

P17:

Consider accessing your e-mail with POP3.

a. Suppose you have configured your POP mail client to operate in the

download-and-delete mode. Complete the following transaction:

C: List

S: 1 498

S: 2 912S: .

C: retr 1

S: blaah blah …

S: ………….blah

S: .

C: dele 1

C: retr 2

S: blaah blah …

S: ………….blah

S: .

C: dele 2

C: quit

S: +OK POP3 Server signing off

b. Suppose you have configured your POP mail client to operate in the

download-and-keep mode. Complete the following transaction:

C: list

S: 1 498

S: 2 912

S: .

C: retr 1

S: blah blah….

S: ………..blah

S: .

C: retr 2

S: blah blah …..

S: …………blah

S: .

C: quit

S: + OK POP3 Server signing off

c. Suppose you have configured your POP mail client to operate in the

download-and-keep mode. Using your transcript in part (b), suppose you

retrieve messages 1 and 2, exit POP, and then five minutes later you again

access POP to retrieve new e-mail. Suppose that in the five-minute interval

no new messages have been sent to you. Provide a transcript of this second POP session.

C: list

S: 1 498

S: 2 912

S: .

C: retr 1

S: blah blah….

S: ………..blah

S: .

C: retr 2

S: blah blah …..

S: …………blah

S: .

C: quit

S: + OK POP3 Server signing off

The second transcript is the same as the first because the POP client has been closed

P18:

1. What is a whois database?

A whois database stores the users of certain applications on the web. It does this by finding information like the IP address, and domain name a user is on.

1. Use various whois databases on the Internet to obtain the names of two DNS servers. Indicate which whois databases you used.

<http://www.dawhois.com/>

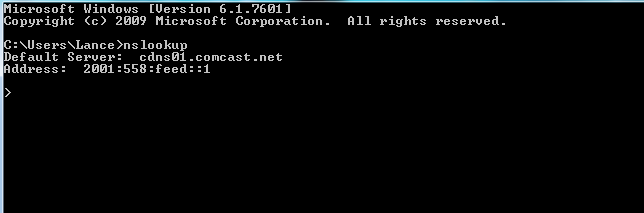
For [walmart.com](https://walmart.com/) a DNS server is:

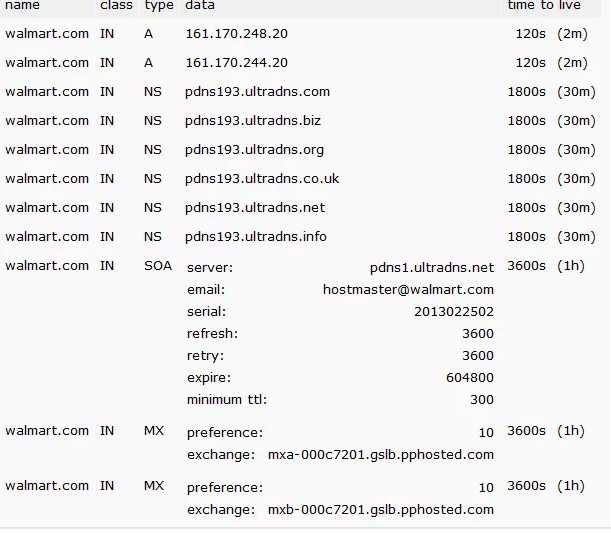
NS1-137.AKAM.NET

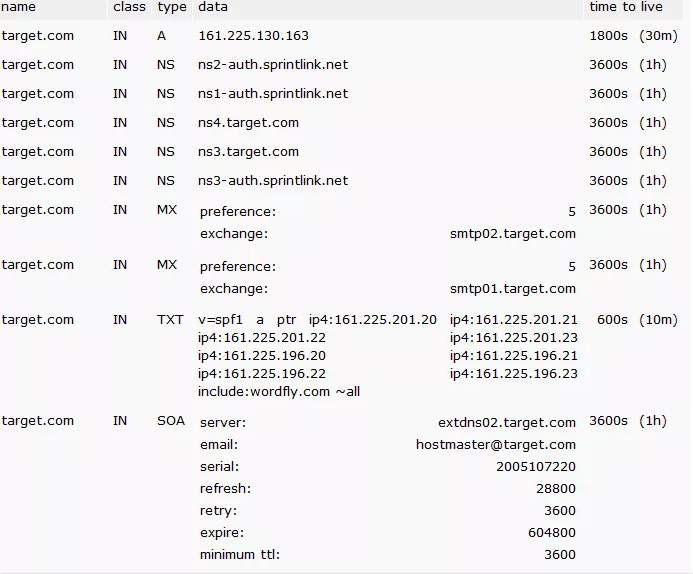
For [Target.com](https://target.com/) a DNS server is:

NS1-AUTH.SPRINTLINK.NET

1. Use nslookup on your local host to send DNS queries to three DNS servers: your local DNS server and the two DNS servers you found in part (b). Try querying for Type A, NS, and MX reports. Summarize your findings.





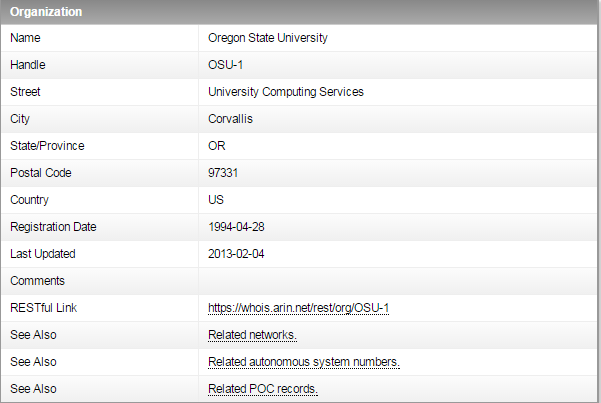


1. Use nslookup to find a Web server that has multiple IP addresses. Does the Web server of your institution (school or company) have multiple IP addresses?

[Walmart.com](https://walmart.com/) has multiple IP addresses, ECSU only has one.

1. Use the ARIN whois database to determine the IP address range used by your university.

IP range is: 209.206.253.184 - 209.206.253.191



1. Describe how an attacker can use whois databases and the nslookup tool to perform reconnaissance on an institution before launching an attack.

Whois can be used to look up a huge IP range, which would give the attacker a huge amount of people to target. After this, by using nslookup for those IP addresses, this attacker could obtain even better info about any of those people, and better target the attack.

1. Discuss why whois databases should be publicly available.

Whois databases should be publicly available becasue they make finding information about domains much easier. Without whois databases, it would be very difficult to find any domain information without contacting people.

P19:

In this problem, we use the useful dig tool available on Unix and Linux hosts to explore the hierarchy of DNS servers. Recall that in Figure 2.21, a DNS server higher in the DNS hierarchy delegates a DNS query to a DNS server lower in the hierarchy, by sending back to the DNS client the name of that lower-level DNS server. First read the man page for dig, and then answer the following questions.

1. Starting with a root DNS server (from one of the root servers [a-m].rootservers.net), initiate a sequence of queries for the IP address for your department’s Web server by using dig. Show the list of the names of DNS servers in the delegation chain in answering your query.

The following delegation chain is used for gaia.cs.umass.edu a.root-servers.net E.GTLD-SERVERS.NET ns1.umass.edu(authoritative)

First command: dig +norecurse @a.root-servers.net any gaia.cs.umass.edu

;; AUTHORITY SECTION:

edu. 172800 IN NS E.GTLD-SERVERS.NET.

edu. 172800 IN NS A.GTLD-SERVERS.NET.

edu. 172800 IN NS G3.NSTLD.COM.

edu. 172800 IN NS D.GTLD-SERVERS.NET.

edu. 172800 IN NS H3.NSTLD.COM.

edu. 172800 IN NS L3.NSTLD.COM.

edu. 172800 IN NS M3.NSTLD.COM.

edu. 172800 IN NS C.GTLD-SERVERS.NET.

Among all returned edu DNS servers, we send a query to the first one. dig +norecurse @E.GTLD-SERVERS.NET any gaia.cs.umass.edu

umass.edu. 172800 IN NS ns1.umass.edu.

umass.edu. 172800 IN NS ns2.umass.edu.

umass.edu. 172800 IN NS ns3.umass.edu.

Among all three returned authoritative DNS servers, we send a query to the first one. dig +norecurse @ns1.umass.edu any gaia.cs.umass.edu

gaia.cs.umass.edu. 21600 IN A 128.119.245.12

1. Repeat part a) for several popular Web sites, such as google.com, yahoo.com, or amazon.com.

For google.com:

a.root-servers.net

E.GTLD-SERVERS.NET

ns1.google.com(authoritative)

P22:

Consider distributing a file of F = 15 Gbits to N peers. The server has an upload rate of us = 30 Mbps, and each peer has a download rate of di = 2 Mbps and an upload rate of u. For N = 10, 100, and 1,000 and u = 300 Kbps, 700 Kbps, and 2 Mbps, prepare a chart giving the minimum distribution time for each of the combinations of N and u for both client-server distribution and P2P distribution.

using the formulas:

Dcs = max{NF/Us , F/Dmin}

DP2P =max { F/Us ,  NF/(US+U1+…UN)}

F = 15 Gbits = 15 \* 1024 Mbits

us = 30 Mbps

dmin = di = 2 Mbps

Client Server N

|  |  |  |  |
| --- | --- | --- | --- |
|  | 10 | 100 | 1000 |
| 300Kbps | 7680 | 51200 | 512000 |
| 700Kbps | 7680 | 51200 | 512000 |
| 2Mbps | 7680 | 51200 | 512000 |

Peer to Peer N

|  |  |  |  |
| --- | --- | --- | --- |
|  | 10 | 100 | 1000 |
| 300Kbps | 7680 | 25904 | 47559 |
| 700Kbps | 7680 | 15616 | 21525 |
| 2Mbps | 7680 | 7680 | 7680 |

P26:

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

1. 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?

Yes. His first claim is possible, as long as there are enough peers staying in the swarm for a long enough time. Bob can always receive data through optimistic unchoking by other peers.

1. 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?

His second claim is also true. He can run a client on each machine, and let each client do “free-riding”, and combine those collected chunks from different machines into a single file. He can even write a small scheduling program to let different machines only asking for different chunks of the file. This is actually a kind of Sybil attack in P2P networks.

P33:

Can you configure your browser to open multiple simultaneous connections to a Web site? What are the advantages and disadvantages of having a large number of simultaneous TCP connections?

1. Yes, you can configure your browser to open multiple simultaneous connections to a website.
2. The advantages of having a large number of TCP connections is the fact that the data will be transferred reliably, the data packets will be in order, and lost packets will be re transmitted.
3. A major disadvantage is that TCP has a congestion control mechanism, which will slow down the transfer of data when the connection is congested. In this instance, the transfer rate would be very slow since there are many simultaneous connections open at once.