Chapter 3

Lab 2: Analyze the downloading of embedded objects in a web-page using Wireshark

3.1 Objective

To see how a webpage having embedded objects (image files) are downloaded. Understand the format of the HTTP messages exchanged between the client and the server. We will analyze the throughput and delay across a persistent connection.

3.2 Procedure

1. Start up your web browser, and make sure your browser's cache is cleared, as discussed above. Start up the Wireshark packet sniffer.

2. Enter the following URL into your browser http://gaia.cs.umass.edu/wireshark-labs/HTTP-wireshark-file4.html

3. Your browser should display a short HTML file with two images. These two images are referenced in the base HTML file. That is, the images themselves are not contained in the HTML; instead the URLs for the images are contained in the downloaded HTML file. Your browser will use the persistent TCP connection to retrieve the Publisher's logo from "gaia.cs.umas.edu". However, for the book logo, your browser makes a secure connection (https) to "kurose.cslash.net" and downloads it. Notice that for unsecure connections (http) the server port is 80 whereas for secure connections (https) the server port is 443.

4. Stop Wireshark packet capture, and enter "http" in the display-filter-specification window, so that only captured HTTP messages will be displayed.

5. You can filter "ip.src == 178.79.137.164" to see the packets (TCP in protocol column) which delivered the book logo.

3.3 Analyses

1. How many HTTP GET request messages were sent by your browser (excluding the object favicon.ico)? To which Internet addresses were these GET requests sent?

2. Fill the following table:

Object	**Source IP Address**	**Destination IP Address**	**Source Port**	**Destination Port**	**Transport Layer Protocol**
HTTP-wireshark-					

file4.html					
pearson.png					
8E_cover_small.jpg					

3. How many hosts were responding to your browser? What are the host names?
4. Can you tell whether your browser downloaded the two images serially, or whether they were downloaded from the two web sites in parallel? Explain.
5. What is the total delay incurred in downloading the webpage along with the embedded objects?
6. What was the throughput when downloading the publisher's logo?
7. Now inspect the HTTP GET message corresponding to the object "HTTP-wireshark-file4.html". What is the server name? How do we know it is persistent connection?
8. Now inspect the HTTP response message corresponding to the object "HTTP-wireshark-file4.html". What is the time at which the object was retrieved by the server? When did this version of the object become available on the server?
9. Now inspect the HTTP messages corresponding to the object "pearson.png". What is name of the server to which the GET message was sent? What is size of the object returned by the server? How could you tell?
10. What were the status and the code returned by the server when the HTTP GET message was sent by the browser to fetch the object "8E_cover_small.jpg"?