CSSE 477 –Review of the Web and HTTP

The Chapter 2 of the Kurose and Rose book (hand-out given to you during Tuesday's exam) is very important for the second half of the course. We will talk about web application servers, web services, and other interesting technologies and concepts that enable Service-Oriented Architecture (SOA). You must know how HTTP works in order to deeply understand these concepts.

Remember, the HTTP protocol is built on top of the transport layer. The web servers and clients (web browsers) all use the TCP Socket API for sending and receiving messages. These messages are plain text messages and sometimes accompany raw binary data such as pdf and image files. The HTTP protocol essentially defines the structure of these messages as well as their behavior. You have already done some socket programming in this course. Web browsers and servers essentially use those API to talk with each other using the HTTP protocol.

Please review Chapter 2 by answering the following questions:

1. What are the three essential components of the World Wide Web (WWW) or in short, the web?
2. What is the key advantage of having layered architecture for the Internet when it comes to building applications using HTTP?
3. HTTP is a stateless protocol. Explain why?
4. Assume you have entered the following URL on your browser: http://everest.csse.rose-hulman.edu/index.html. (Please check the contents of the web page before you answer this question.) Assume the browser uses non-persistent connection to the server. List down the sequence of activities that happens before the browser displays the content of the page.
5. What is the difference if the browser uses persistent connection for question #4?
6. What is the key advantage of having persistent connection?
7. Give a short description for each of the following field (or line) in an HTTP request message:
    a. GET /~test/index.html HTTP/1.1
    b. Host: everest.csse.rose-hulman.edu
8. Explain the purpose of the following request types: GET, POST, PUT, DELETE, and HEAD.
9. Give a short description of each of the following HTTP response codes: 200, 301, 400, 404, and 505.
10. What technique can one use to achieve a stateful behavior while using HTTP?
11. How does web caching works?
12. What is the significance of the conditional-GET request? How do you think the modern-day browsers take advantage of the conditional-GET requests?
13. Download the **HTTPTestClient.jar** software from Moodle (or use Postman or Advanced REST Client browser extensions for this exercise). Double-click to run it (or in terminal, type: **java –jar HTTPTestClient.jar**). Click **Connect** -> **Generate Persistent Request** -> **Send**. You should get a response back. (**Note**: The connection closed dialog is expected. You should see the response header in the log to understand why. Please feel free to try out other commands too.)
    a. Break-down the request according to Figure 2.8. (i.e., Identify request line, header lines, blank line, and entity body).
    b. Break-down the response according to Figure 2.9.
14. Based on your new found understanding of HTTP and Web Servers, if you were tasked with developing a Web Server from scratch in either Java or C#, explain what API you would use for communication between Web Browsers and Web Server? Draw an architecture diagram for the Web Server (not the client) identifying various modules required for the server. Detail the architecture using UML class diagrams where you identify various interfaces and classes in each module. Please briefly describe the purpose of each class and module if it is not very clear just from its name. The latest version of **UMLet** can be downloaded from **Moodle** (under the **Resources** section) that you can use for drawing. (**Note that I am looking for a meaningful attempt here, which does not have to be absolutely correct.**)