Lab Worksheet: What Does it Take to Get a Web Page?

Name:				
Email:				
Discussion Leader:	Bestavros	Hall	Kolaczyk	Reyzin

Task 1: Looking up the IP address of the remote host

Follow the steps described for this task to lookup the IP address of www.cnn.com or www.facebook.com and answer the following questions:

- How many answers did you get?
- Did you get an authoritative answer?
- Write down the IP address of a www.cnn.com or a www.facebook.com host:

Task 2: Getting an authoritative answer for host name lookup

Follow the steps described for this task to get an authoritative answer for the IP address of www.cnn.com or www.facebook.com and answer the following questions:

- What name server did you use to obtain an authoritative answer?
- Did you get answers different from those you got before?
- Write down the IP address of www.cnn.com or www.facebook.com obtained from an authoritative name server:

Task 3: Contacting the web server program on the remote host

Using telnet on port #80 for one of the cnn servers you got in the previous task, get the contents of the web page at http://www.cnn.com/robots.txt and answer the following question:

- How many directories are not available for crawlers on www.cnn.com?

- Would a crawler be allowed to fetch the contents of the web page at http://www.cnn.com/aol/story/2004/12/14/cnn_ALLPOLITICS_dnc.chair.html? Why?

- Are you allowed to fetch the above page using a regular browser? Why?

(Optional) Task 4: What exactly is "robots.txt"?

In your own words explain why it is not fair to accuse crawlers (such as those used by Google and Microsoft for indexing purposes) as making content that is not for public distribution available for searching?

Task 5: What does it take for a packet to go from East Coast to West Coast?

Using traceroute, identify the path that packets would take when sent from BU to the following destinations.

- cs.stanford.edu (West Coast, US)
- cs.berkeley.edu (West Coast, US)
- www.usc.edu (West Coast, US)

Now, answer the following questions:

1. How many hops (routers) are common on the path from BU to West-Coast servers

2. Approximately, how many times could a packet go from the East Coast to the West Coast and back in one second? Show your work.

Using traceroute, identify the path that packets would take when sent from BU to at least three of the following destinations in Europe.

- latency.hosteurope.de (Germany)
- www.unige.ch (Switzerland)
- www.supporttechnique.net (France)
- support.lbn.fr (France)
- www.rhnet.is (Iceland)

Now, answer the following questions:

3. Identify the RTT (round-trip time) of the transatlantic hop on the path from BU to the European servers you used.

4. Approximately, what is the one way delay for a transatlantic hop?

5. Assuming that packets travel at 120,000 mph (66% of the speed of light), what is the distance of the transatlantic hop. Show your work. Does this result make sense?

(Optional) Task 6: Looking at the Reverse Path

Using traceroute, identify the path that packets would take when sent from one of the following destinations to BU. The destinations below are some of those you used in the previous task.

- www.usc.edu -- use http://www.usc.edu/cgi-bin/traceroute
- latency.hosteurope.de -- use http://latency.hosteurope.de/lg/
- www.unige.ch -- use http://www.unige.ch/cgi-bin/people/hugen/trace
- www.supporttechnique.net -- use http://www.supporttechnique.net/traceroute.ihtml
- support.lbn.fr -- use http://support.linkbynet.com/ping_tracert.ashx
- www.rhnet.is -- use http://www.rhnet.is/cgi-bin/rh-traceroute

Answer the following questions:

1. Can you identify any paths that are not symmetric (i.e., the path from BU to the remote server is different from the path from the remote server to BU)?

Now, from one of the sites in Europe (e.g., latency.hosteurope.de) do a traceroute to the sites in the West Coast.

2. How long is the one-way delay? How does it compare to the one-way delay between Europe and BU + BU to WestCoast?

3. Do the paths from Europe to Eastern US and to Western US have the same transatlantic hop?