Homework-1: Wireshark

Name: Shanqing Gu

What to Submit

Your final submission shall be a single pdf document that includes this document, screen captures of your exercises plus your answers to each of the written questions (if any). Note that you are expected to clearly label each section so as to make it clear to the instructor what files and data belong to which exercise/question.

Collaboration is expected and encouraged; however, each student must hand in their own homework assignment. To the greatest extent possible, answers should not be copied but, instead, should be written in your own words. Copying answers from anywhere is plagiarism, this includes copying text directly from the textbook. Do not copy answers. Always use your own words. For each question list all persons with whom you collaborated and list all resources used in arriving at your answer. Resources include but are not limited to the textbook used for this course, papers read on the topic, and Google search results. Note that 'Google' is not a resource. Don't forget to place your name on the document.

Exercise 1: Wireshark Protocol Layers

The goal of this exercise is to become familiar with Wireshark, a network packet sniffer and analysis tool, and to observe how protocols and layering are represented in packets. This exercise is adapted from the Protocol Wireshark Lab by David Wetherall.

This exercise uses the Wireshark software tool to capture and examine a packet trace. A packet trace is a record of traffic at a location on the network, as if a snapshot was taken of all the bits that passed across a particular wire. The packet trace records a timestamp for each packet, along with the bits that make up the packet, from the lower-layer headers to the higher-layer contents. Wireshark runs on most operating systems, including Windows, Mac and Linux. It provides a graphical user interface that shows the sequence of packets and the meaning of the bits when interpreted as protocol headers and data. It color-codes packets by their type and has various ways to filter and analyze packets to let you investigate the behavior of network protocols. Wireshark is widely used to troubleshoot networks. You can download Wireshark from www.wireshark.org if it is not already installed on your computer. We highly recommend that you watch the short, 5 minute video ?Introduction to Wireshark? that is on the site.

This exercise also introduces wget (Linux and Windows) and curl (Mac) to fetch web resources. wget and curl are command-line programs that let you fetch a URL. Unlike a web browser, which fetches and executes entire pages, wget and curl give you control over exactly which URLs you fetch and when you fetch them. Under Linux, wget can be installed via your package manager. Under Windows, wget is available as a binary; look for download information on http://www.gnu.org/software/wget/. Under Mac, curl comes installed with the OS. Both have many options (try wget --help or curl --help to see) but a URL can be fetched simply with wget URL or curl URL.

After installing Wireshark, perform each of the steps below.

Step 1: Capture a Trace

Proceed as follows to capture a trace of network traffic. We want this trace to look at the protocol structure of packets. A simple Web fetch of a URL from a server of your choice to your computer, which is the client, will serve as traffic.

1) Pick a URL and fetch it with wget or curl. For example, wget http://www.google.com or curl http://www.google.com. This will fetch the resource and either write it to a file (wget) or to the screen (curl). You are checking to see that the fetch works and retrieves some content. If the fetch does not work then try a different URL; if no URLs seem to work then debug your use of wget/curl or your Internet connectivity.

2) Close unnecessary browser tabs and windows. By minimizing browser activity, you will stop your computer from fetching unnecessary web content and avoid incidental traffic in the trace.

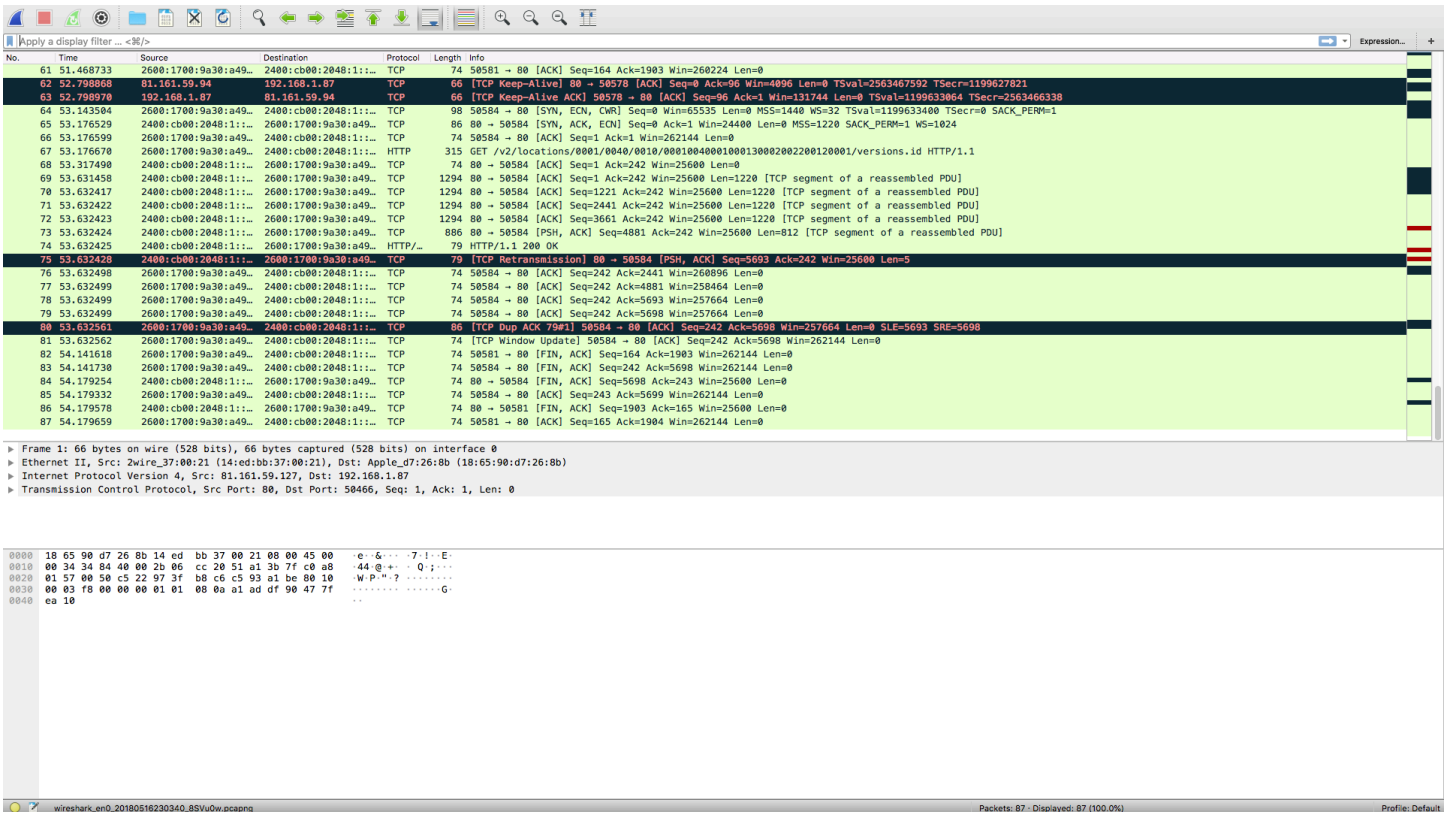
3) Launch Wireshark and start a capture with a filter of tcp port 80 and check *enable network name resolution.* This filer will record only standard web traffic and not other kinds of packets that your computer may send. The checking will translate the addresses of the computers sending and receiving packets into names, which should help you to recognize whether the packets are going to or from your computer. Select the interface from which to capture as the main wired or wireless interface used by your computer to connect to the Internet. If unsure, guess and revisit this step later if your capture is not successful. Uncheck *capture packets in promiscuous mode.* This mode is useful to overhear packets sent to/from other computers on broadcast networks. We only want to record packets sent to/from your computer. Leave other options at their default values. The capture filter, if present, is used to prevent the capture of other traffic your computer may send or receive. On Wireshark 1.8, the capture filter box is pre- sent directly on the options screen, but on Wireshark 1.9, you set a capture filter by double-clicking on the interface.

4) When the capture is started, repeat the web fetch using wget/curl above. This time, the packets will be recorded by Wireshark as the content is transferred.

5) After the fetch is successful, return to Wireshark and use the menus or buttons to stop the trace. If you have succeeded, the upper Wireshark window will show multiple packets, and most likely it will be full. How many packets are captured will depend on the size of the web page, but there should be at least 8 packets in the trace, and typically 20-100, and many of these packets will be colored green. Congratulations! You have captured your first trace.

Turn In: Screen capture your Wireshark trace and turn it in.

Step 1: Capture a Trace
Turn In: Screen capture your Wireshark trace and turn it in (8 points)



Step 2: Inspect the Trace

Wireshark will let us select a packet (from the top panel) and view its protocol layers, in terms of both header fields (in the middle panel) and the bytes that make up the packet (in the bottom panel). In the figure above, the first packet is selected (shown in blue). Note that we are using *packet* as a general term here. Strictly speaking, a unit of information at the Link Layer is called a frame. At the Network Layer it is called a packet, at the Transport Layer a segment, and at the Application Layer a message. Wireshark is gathering frames and presenting us with the higher-layer packet, segment, and message structures it can recognize that are carried within the frames. We will often use *packet* for convenience, as each frame contains one packet and it is often the packet or higher-layer details that are of interest.

Select a packet for which the Protocol column is 'HTTP' and the Info column says it is a GET. It is the packet that carries the web (HTTP) request sent from your computer to the server. (You can click the column headings to sort by that value, though it should not be difficult to find an HTTP packet by inspection.) Let's have a closer look to see how the packet structure reflects the protocols that are in use.

Since we are fetching a web page, we know that the protocol layers being used are from the HTTP stack (namely, HTTP at the Application Layer, TCP at the Transport Layer, IP at the Network Layer, and either Ethernet or WiFi at the Link and Physical layers). That is, HTTP is the application layer web protocol used to fetch URLs. Like many Internet applications, it runs on top of the TCP/IP Transport and Network layer protocols. The Link and Physical layer protocols depend on your network but are typically combined in the form of Ethernet if your computer is wired, or WiFi if your computer is wireless.

With the HTTP GET packet selected, look closely to see the similarities and differences between it and our protocol stack as described next. The protocol blocks are listed in the middle panel. You can expand each block (by clicking on the '+' expander or icon) to see its details.

- The first Wireshark block is 'Frame'. This is not a protocol, it is a record that describes overall information about the packet, including when it was captured and how many bits long it is.
- The second block is 'Ethernet'. Note that you may have taken a trace on a computer using 802.11 (WiFi) yet still see an Ethernet block instead of an 802.11 block. Why? It happens because we asked Wireshark to capture traffic in Ethernet format on the capture options, so it converted the real 802.11 header into a pseudo-Ethernet header.
- Then come IP, TCP, and HTTP, which are just as we wanted. Note that the order is from the bot- tom of the protocol stack upwards. This is because as packets are passed down the stack, the header information of the lower layer protocol is added to the front of the information from the higher layer protocol. That is, the lower layer protocols come first in the packet 'on the wire'.

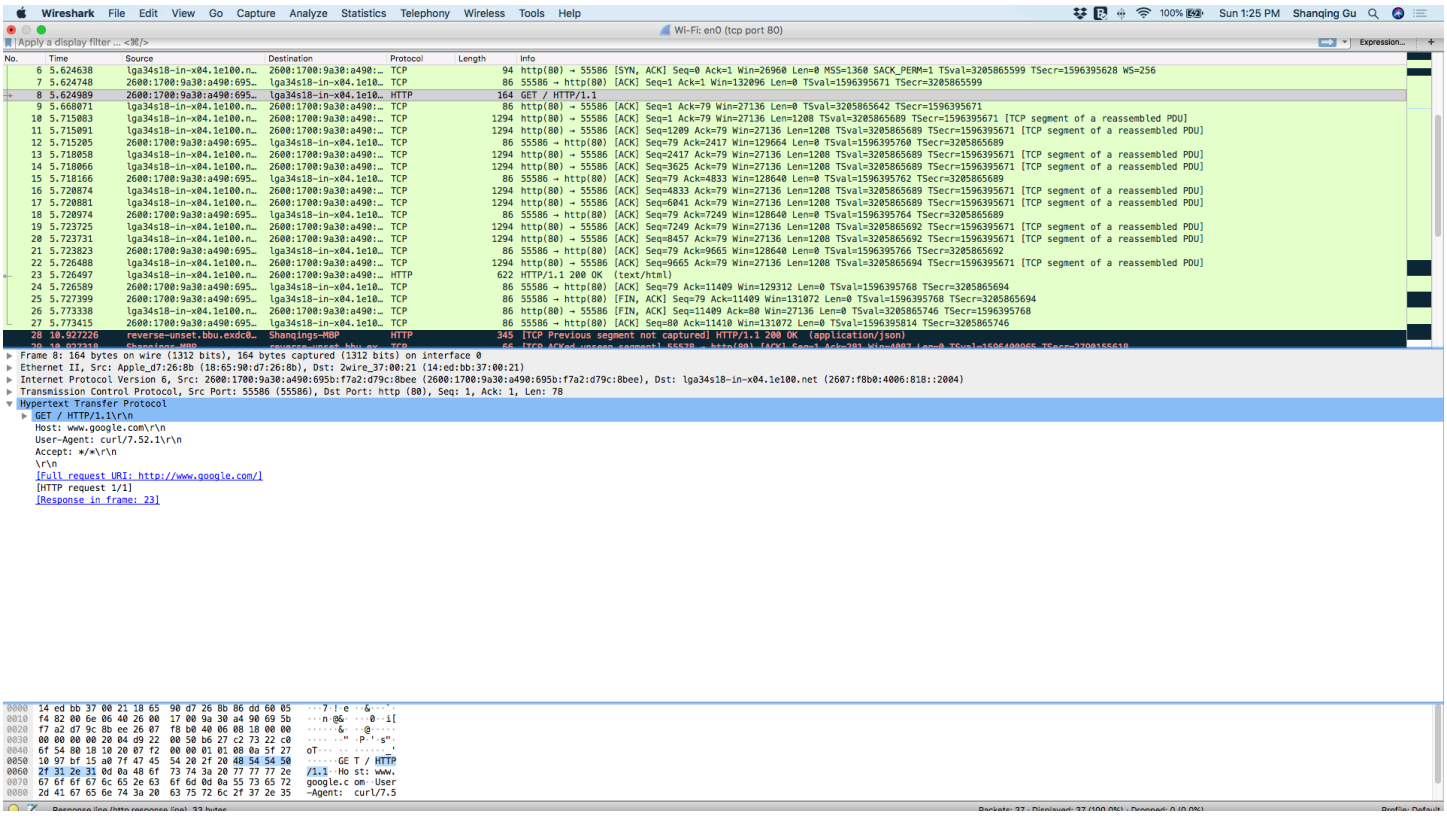
  Now find another HTTP packet, the response from the server to your computer, and look at the structure of this packet for the differences compared to the HTTP GET packet. This packet should have ?200 OK? in the Info field, denoting a successful fetch. In your trace, there should be two extra blocks in the detail panel.

• The first extra block says '[11 reassembled TCP segments ...]'. Details in your capture will vary, but this block is describing more than the packet itself. Most likely, the web response was sent across the network as a series of packets that were put together after they arrived at the computer. The packet labeled HTTP is the last packet in the web response, and the block lists packets that are joined together to obtain the complete web response. Each of these packets is shown as having protocol TCP even though the packets carry part of an HTTP response. Only the final packet is shown as having protocol HTTP when the complete HTTP message may be under- stood, and it lists the packets that are joined together to make the HTTP response.

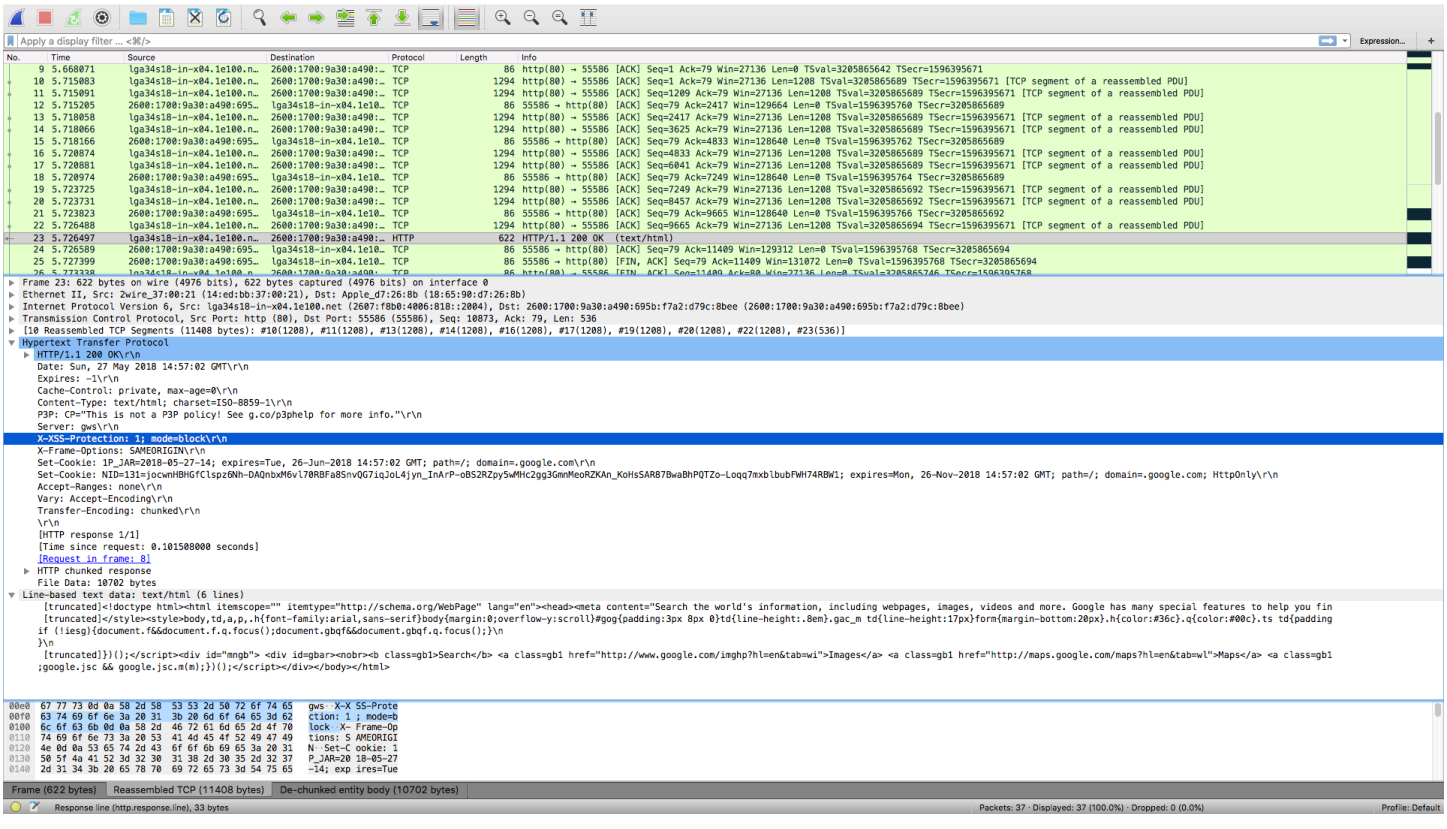
The second extra block says 'Line-based text data ...'. Details in your capture will vary, but this block is describing the contents of the web page that was fetched. In our case it is of type text/html, though it could easily have been text/xml, image/jpeg, or many other types. As with the Frame record, this is not a true protocol. Instead, it is a description of packet contents that Wireshark is producing to help us understand the network traffic.

Turn In: Screen capture your Wireshark packet structure screen (8 points)

HTTP GET Packet: Request by an end-user's browser



HTTP packet with "200 OK": Response from the server



Step 3: Packet Structure

To show your understanding of packet structure, draw a figure of an HTTP GET packet that shows the position and size in bytes of the TCP, IP and Ethernet protocol headers. Your figure can simply show the overall packet as a long, thin rectangle. Leftmost elements are the first sent on the wire. On this drawing, show the range of the Ethernet header and the Ethernet payload that IP passed to Ethernet to send over the network. To show the nesting structure of protocol layers, note the range of the IP header and the IP payload. You may have questions about the fields in each protocol as you look at them. We will explore these protocols and fields in detail in future labs.

To work out sizes, observe that when you click on a protocol block in the middle panel (the block itself, not the '+' expander) then Wireshark will highlight the bytes it corresponds to in the packet in the lower panel and display the length at the bottom of the window. For instance, clicking on the IP version 4 header of a packet in our trace shows us that the length is 20 bytes. (Your trace will be different if it is IPv6, and may be different even with IPv4 depending on various options.) You may also use the overall packet size shown in the Length column or Frame detail block.

Turn In: Hand in your packet drawing (you can use paint program to draw) (8 points)

```
▼ Frame 15: 161 bytes on wire (1288 bits), 161 bytes captured (1288 bits) on interface 0
  ▶ Interface id: 0 (en0)
    Encapsulation type: Ethernet (1)
    Arrival Time: May 18, 2018 20:37:30.905144000 EDT
    [Time shift for this packet: 0.000000000 seconds]
    Epoch Time: 1526690250.905144000 seconds
    [Time delta from previous captured frame: 0.000103000 seconds]
    [Time delta from previous displayed frame: 0.000103000 seconds]
    [Time since reference or first frame: 31.575095000 seconds]
    Frame Number: 15
    Frame Length: 161 bytes (1288 bits)
    Capture Length: 161 bytes (1288 bits)
    [Frame is marked: False]
    [Frame is ignored: False]
    [Protocols in frame: eth:ethertype:ip:tcp:http]
    [Coloring Rule Name: HTTP]
    [Coloring Rule String: http || tcp.port == 80 || http2]
▼ Ethernet II, Src: Apple_d7:26:8b (18:65:90:d7:26:8b), Dst: 2wire_37:00:21 (14:ed:bb:37:00:21)
  ▶ Destination: 2wire_37:00:21 (14:ed:bb:37:00:21)
  ▶ Source: Apple_d7:26:8b (18:65:90:d7:26:8b)
    Type: IPv4 (0x0800)
▼ Internet Protocol Version 4, Src: 192.168.1.87, Dst: 81.161.59.127
    0100 .... = Version: 4
    .... 0101 = Header Length: 20 bytes (5)
  ▶ Differentiated Services Field: 0x00 (DSCP: CS0, ECN: Not-ECT)
    Total Length: 147
    Identification: 0x0000 (0)
  ▶ Flags: 0x4000, Don't fragment
    Time to live: 64
    Protocol: TCP (6)
    Header checksum: 0xeb45 [validation disabled]
    [Header checksum status: Unverified]
    Source: 192.168.1.87
    Destination: 81.161.59.127
▶ Transmission Control Protocol, Src Port: 58089, Dst Port: 80, Seq: 96, Ack: 559, Len: 95
▶ Hypertext Transfer Protocol
```

Packet Structure

Destination Address	Source Address	Type	Header	IP Data	Flags
6 Bytes	6 Bytes	4 Bytes	20 Bytes	127 Bytes	3 Bytes
Ethernet II Header				IPv4 Payload	

Step 4: Protocol Overhead

Estimate the download protocol overhead, or percentage of the download bytes taken up by protocol overhead. To do this, consider HTTP data (headers and message) to be useful data for the network to carry, and lower layer headers (TCP, IP, and Ethernet) to be the overhead. We would like this overhead to be small, so that most bits are used to carry content that applications care about. To work this out, first look at only the packets in the download direction for a single web fetch. You might sort on the Destination column to find them. The packets should start with a short TCP packet described as a SYN ACK, which is the beginning of a connection. They will be followed by mostly longer packets in the middle (of roughly 1 to 1.5KB), of which the last one is an HTTP packet. This is the main portion of the download. And they will likely end with a short TCP packet that is part of ending the connection. For each packet, you can inspect how much overhead it has in the form of Ethernet / IP / TCP headers, and how much useful HTTP data it carries in the TCP payload. You may also look at the HTTP packet in Wireshark to learn how much data is in the TCP payloads over all download packets.

Turn In: Your estimate of download protocol overhead as defined above (4 points)

TCP packet: SYN ACK

Destination Address	Source Address	Type	Payload	TCP
6 Bytes	6 Bytes	4 Bytes	32 Bytes	0 Bytes
Ethernet II Header			IPv6	

HTTP Packet

Destination Address	Source Address	Type	Payload	TCP
6 Bytes	6 Bytes	4 Bytes	861 Bytes	841 Bytes
Ethernet II Header			IPv6	TCP

Tell us whether you find this overhead to be significant (4 points)

As shown above, for TCP packet and HTTP packet, the Ethernet/IP/TCP header are the same, while the TCP payload is different. HTTP packet has more payload than TCP packet. The overhead difference is significant.

Step 5: Demultiplexing

When an Ethernet frame arrives at a computer, the Ethernet layer must hand the packet that it contains to the next higher layer to be processed. The act of finding the right higher layer to process received packets is called demultiplexing. We know that in our case the higher layer is IP. But how does the Ethernet protocol know this? After all, the higher-layer could have been another protocol entirely (such as ARP). We have the same issue at the IP layer ? IP must be able to determine that the contents of IP message is a TCP packet so that it can hand it to the TCP protocol to process. The answer is that protocols use information in their header known as a ?demultiplexing key? to determine the higher layer.

Look at the Ethernet and IP headers of a download packet in detail to answer the following questions:
1) Which Ethernet header field is the demultiplexing key that tells it the next higher layer is IP? What value is used in this field to indicate IP?
2) Which IP header field is the demultiplexing key that tells it the next higher layer is TCP? What value is used in this field to indicate TCP?

Turn In: Hand in your answers to the above questions (8 points)

1) Which Ethernet header field is the demultiplexing key that tells it the next higher layer is IP? What value is used in this field to indicate IP?

   Type; IPv4 (0x0800) indicates IP.

2) Which IP header field is the demultiplexing key that tells it the next higher layer is TCP? What value is used in this field to indicate TCP?

   Protocol; TCP (6) indicates TCP.

Exercise 2: Wireshark IPv4

The goal of this exercise is to become familiar with IPv4 (Internet Protocol version 4). This exercise is adapted from the IPv4 Wireshark Lab by David Wetherall.

This exercise uses the Wireshark software tool to capture and examine a packet trace.
This exercise uses wget (Linux and Windows) and curl (Mac) to fetch web resources.
This exercise uses traceroute to find the router level path from your computer to a remote Internet host. traceroute

is a standard command-line utility for discovering the Internet paths that your computer uses. It is widely used for network troubleshooting. It comes pre-installed on Window and Mac, and can be installed using your package manager on Linux. On Windows, it is called tracert. It has various options, but simply issuing the command traceroute www.uwa.edu.au will cause your computer to find and print the path to the remote computer (here www.uwa.edu.au).

Perform each of the steps below.

Step 1: Capture a Trace

Proceed as follows to capture a trace of network traffic. We want this trace to look at the protocol structure of packets. A simple Web fetch of a URL from a server of your choice to your computer, which is the client, will serve as traffic.

1) PickaURLataremoteserverandfetchitwithwgetorcurl.Forexample,wget http://www.google.com or curl http://www.google.com. This will fetch the resource and either write it to a file (wget) or to the screen (curl). You are checking to see that the fetch works and retrieves some content. With wget you want a single response with status code *200 OK.* If the fetch does not work then try a different URL; if no URLs seem to work then debug your use of wget/curl or your Internet connectivity.

2) Close unnecessary browser tabs and windows. By minimizing browser activity you will stop your computer from fetching unnecessary web content and avoid incidental traffic in the trace.

3) Performatraceroutetothesameremoteservertocheckthatyoucandiscoverinformationaboutthenetwork path. On Windows, type, e.g., tracert www.uwa.edu.au. On Linux / Mac, type, e.g., traceroute www.uwa.edu.au. If you are on Linux / Mac and behind a NAT (as most home users or virtual machine users) then use the ?I option (that was a capital i) to traceroute, e.g., traceroute ?I www.uwa.edu.au. This will cause traceroute to send ICMP probes like tracert instead of its usual UDP probes; ICMP probes are better able to pass through NAT boxes. Save the output as you will need it for later steps. Note that traceroute may take up to a minute to run. Each line shows information about the next IP hop from the computer running traceroute towards the target destination. The lines with ∗'s indicate that there was no response from the network to identity that segment of the Internet path. Some unidentified segments are to be expected. However, if traceroute is not working correctly then nearly all the path will be ∗'s. In this case, try a different remote server, experiment with traceroute.

4) LaunchWiresharkandstartacapturewithafilteroftcp port 80andcheck*enablenetworknameresolution.* This filer will record only standard web traffic and not other kinds of packets that your computer may send. he checking will translate the addresses of the comput- ers sending and receiving packets into names, which should help you to recognize whether the packets are going to or from your computer. Select the interface from which to capture as the main wired or wireless interface used by your computer to connect to the Internet. If unsure, guess and revisit this step later if your capture is not successful. Uncheck *capture packets in promiscuous mode.* This mode is useful to overhear packets sent to/from other computers on broadcast networks. We only want to record packets sent to/from your computer. Leave other options at their default values. The capture filter, if present, is used to prevent the capture of other traffic your computer may send or receive. On Wireshark 1.8, the capture filter box is pre- sent directly on the options screen, but on Wireshark 1.9, you set a capture filter by double-clicking on the interface.

5) When the capture is started, repeat the web fetch using wget/curl above. This time, the packets will be recorded by Wireshark as the content is transferred.

6) After the fetch is successful, return to Wireshark and use the menus or buttons to stop the trace. If you have succeeded, the upper Wireshark window will show multiple packets, and most likely it will be full. How many packets are captured will

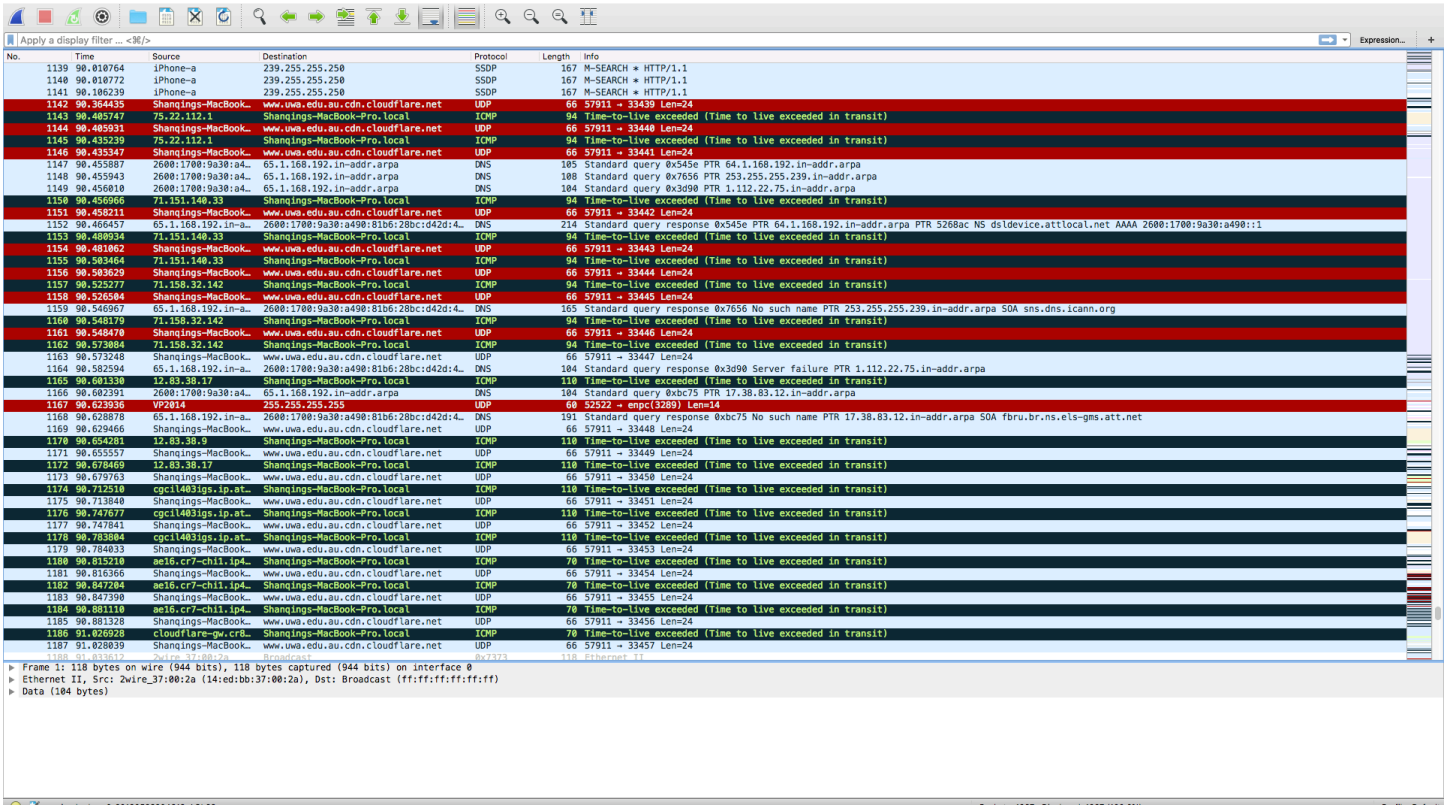
depend on the size of the web page, but there should be at least 8 packets in the trace, and typically 20-100, and many of these packets will be colored green.

Turn In: Screen capture your Wireshark trace and your traceroute, and turn it in. (8 points)

Traceroute

```
[Shanqings-MBP:~ shanqinggu$ traceroute www.uwa.edu.au
traceroute: Warning: www.uwa.edu.au has multiple addresses; using 104.20.10.164
traceroute to www.uwa.edu.au.cdn.cloudflare.net (104.20.10.164), 64 hops max, 52 byte packets
 1  homeportal (192.168.1.254)  4.183 ms  3.566 ms  2.720 ms
 2  75.22.112.1 (75.22.112.1)  27.625 ms  44.308 ms  27.495 ms
 3  71.151.140.33 (71.151.140.33)  24.195 ms  42.305 ms  28.401 ms
 4  71.158.32.142 (71.158.32.142)  23.824 ms  23.149 ms  21.382 ms
 5  12.83.38.17 (12.83.38.17)  25.334 ms
    12.83.38.9 (12.83.38.9)  25.817 ms
    12.83.38.17 (12.83.38.17)  24.882 ms
 6  cgcil403igs.ip.att.net (12.122.133.33)  35.831 ms  33.317 ms  32.041 ms
 7  ae16.cr7-chi1.ip4.gtt.net (173.241.128.29)  31.942 ms  32.741 ms  43.791 ms
 8  cloudflare-gw.cr8-chi1.ip4.gtt.net (69.174.23.14)  31.244 ms  31.315 ms  31.963 ms
 9  104.20.10.164 (104.20.10.164)  30.371 ms  33.447 ms  29.808 ms
Shanqings-MBP:~ shanqinggu$
```

Wireshark trace



Step 2: Inspect the Trace

Select any packet in the trace and expand the IP header fields (using the '+' expander or icon) to see the details. You can simply click on a packet to select it (in the top panel). You will see details of its structure (in the middle panel) and the bytes that

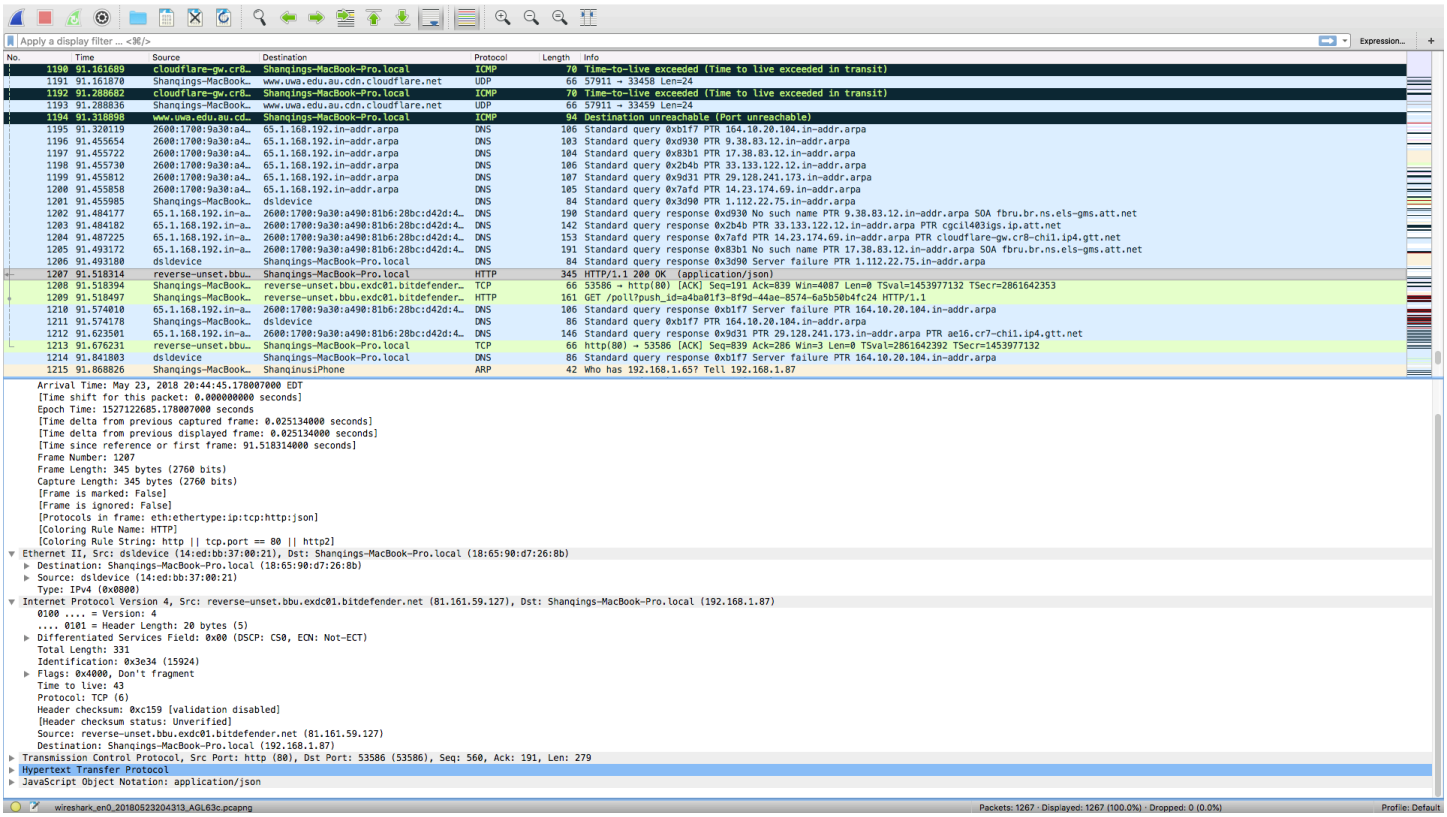
make up the packet (in the bottom panel). Our interest is the IP header, and you may ignore the other higher and lower layer protocols. When you click on parts of the IP header, you will see the bytes that correspond to the part highlighted in the bottom panel.

Let us go over the fields in turn:

The version field is set to 4. This is *IPv4* after all.

- Then there is the header length field. Observe by looking at the bytes selected in the packet data that version and header length are both packed into a single byte.
- The Differentiated Services field contains bit flags to indicate whether the packet should be handled with quality of service and congestion indications at routers.
- Then there is the Total Length field.
- Next is the Identification field, which is used for grouping fragments, when a large IP packet is sent as multiple smaller pieces called fragments. It is followed by the Flags and the Fragment offset fields, which also relate to fragmentation. Observe they share bytes.
- Then there is the Time to live or TTL field, followed by the Protocol field.
- Next comes the header checksum. Is your header checksum carrying 0 and flagged as incorrect for IP packets sent from your computer to the remote server? On some computers, the operating system software leaves the header checksum blank (zero) for the NIC to compute and fill in as the packet is sent. This is called protocol offloading. It happens after Wireshark sees the packet, which causes Wireshark to believe that the checksum is wrong and flag it with a different color to signal a problem. A similar issue may happen for the TCP checksum. You can remove these false errors if they are occurring by telling Wireshark not to validate the checksums. Select ?Preferences? from the Wireshark menus and expand the ?Protocols? area. Look under the list until you come to IPv4. Uncheck ?Validate checksum if possible?. Similarly, you may uncheck checksum validation for TCP if applicable to your case.
- The last fields in the header are the normally the source and destination address. It is possible for there to be IP options, but these are unlikely in standard web traffic.
- The IP header is followed by the IP payload. This makes up the rest of the packet, starting with the next higher layer header, TCP in our case, but not including any link layer trailer (e.g., Ethernet padding).

Turn In: Screen capture your Wireshark packet structure screen. Modify your screen capture to indicate each of the components of the IPv4 packet (8 points)



Step 3: IP Packet Structure

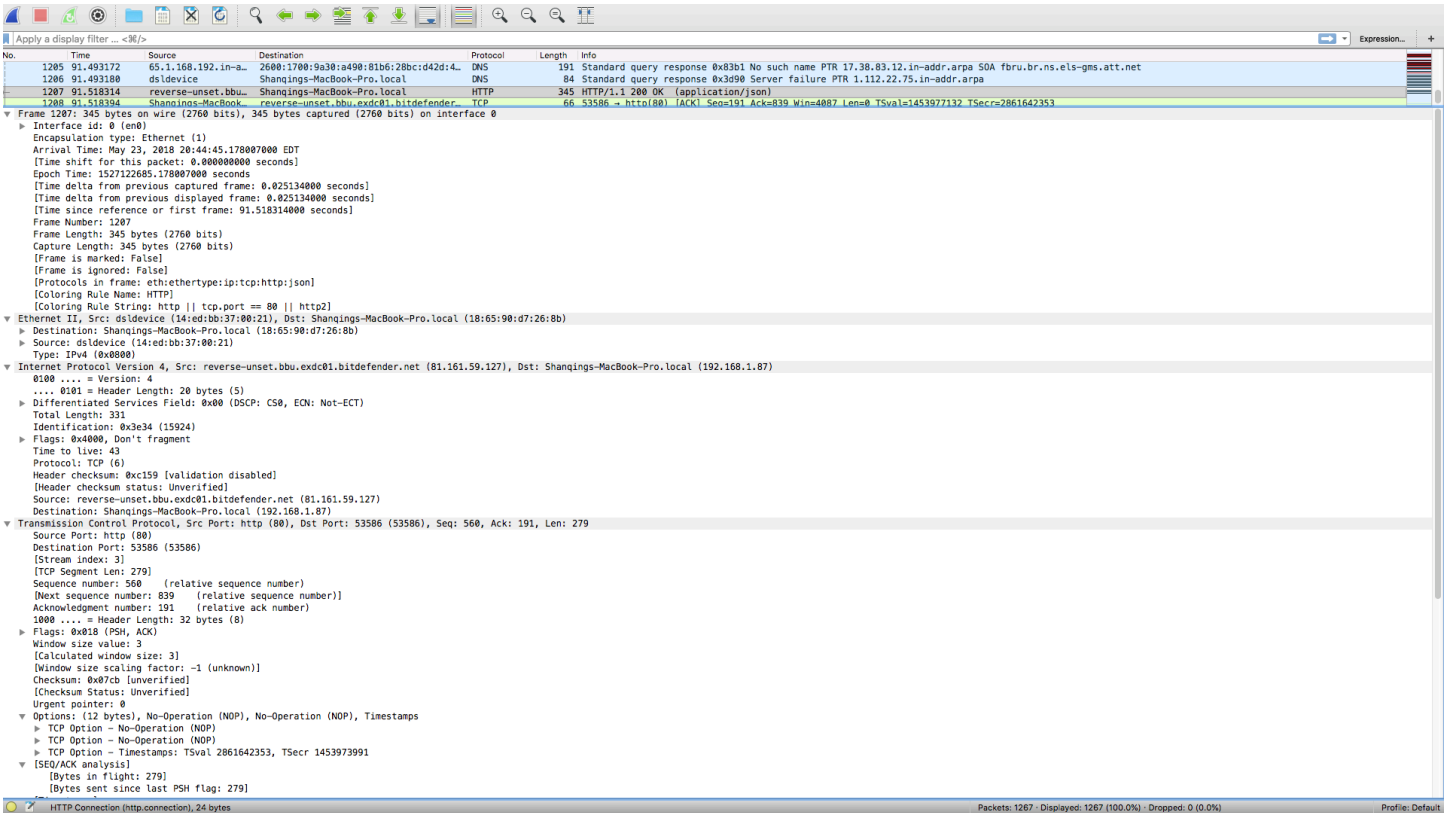
To show your understanding of IP, sketch a figure of an IP packet you studied. It should show the position and size in bytes of the IP header fields as you can observe using Wireshark. Since you cannot easily determine sub-byte sizes, group any IP fields that are packed into the same bytes. Your figure can simply show the frame as a long, thin rectangle. Try not to look at the figure of an IPv4 packet in your text; check it afterwards to note and investigate any differences.

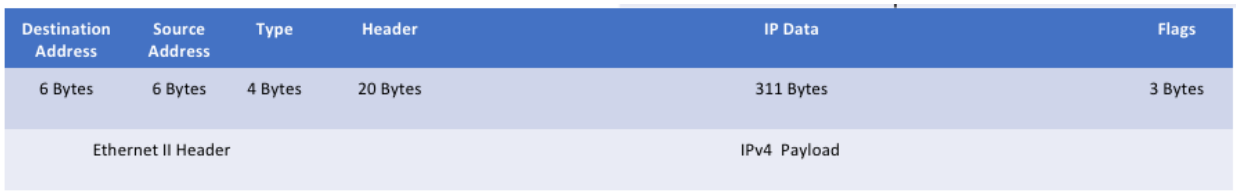
To work out sizes, observe that when you click on a protocol block in the middle panel (the block itself, not the '+' expander) Wireshark will highlight the corresponding bytes in the packet in the lower panel, and display the length at the bottom of the window. You may also use the overall packet size shown in the Length column or Frame detail block. Note that this method will not tell you sub-byte positions.

By looking at the IP packets in your trace, answer these questions:

1) What are the IP addresses of your computer and the remote server?
2) Does the Total Length field include the IP header plus IP payload, or just the IP payload?
3) How does the value of the Identification field change or stay the same for different packets? For instance, does it hold the same value for all packets in a TCP connection or does it differ for each packet? Is it the same in both directions? Can you see any pattern if the value does change?
4) What is the initial value of the TTL field for packets sent from your computer? Is it the maximum possible value, or some lower value?
5) How can you tell from looking at a packet that it has not been fragmented? Most often IP packets in normal operation are not fragmented. But the receiver must have a way to be sure. Hint: you may need to read your text to confirm a guess.
6) What is the length of the IP Header and how is this encoded in the header length field? Hint: notice that only 4 bits are used for this field, as the version takes up the other 4 bits of the byte. You may guess and check your text.

Turn In: Hand in your drawing of an IP packet and the answers to the questions above (8 points)





1) What are the IP addresses of your computer and the remote server?
IP address for my computer: 192.168.1.87 and the remote server is: 81.161.59.127

2) Does the Total Length field include the IP header plus IP payload, or just the IP payload?
The total length field is the total length of the IPv4 datagram in bytes (331 bytes), so it only includes IP payload (331bytes), while the IP header is 20 bytes.

3) How does the value of the Identification field change or stay the same for different packets? For instance, does it hold the same value for all packets in a TCP connection or does it differ for each packet? Is it the same in both directions? Can you see any pattern if the value does change?
a. The value of the Identification field will change for different packets.
b. It holds different values between TCP packets because the data is different.
c. It is different in both directions.
d. The value of identification filed will change depending on the IP payload.

4) What is the initial value of the TTL field for packets sent from your computer? Is it the maximum possible value, or some lower value?
The initial value of the TTL field for packets sent from my computer is 331 bytes. It is some lower value.

5) How can you tell from looking at a packet that it has not been fragmented? Most often IP packets in normal operation are not fragmented. But the receiver must have a way to be sure. Hint: you may need to read your text to confirm a guess.
It is not fragmented. From Flag field, showing "Don't fragment".

6) What is the length of the IP Header and how is this encoded in the header length field? Hint: notice that only 4 bits are used for this field, as the version takes up the other 4 bits of the byte. You may guess and check your text.
IP header length is 20 bytes, and 331 bytes total length, this gives 311 bytes in the payload of the IP datagram.

Step 4: Internet Paths

The source and destination IP addresses in an IP packet denote the endpoints of an Internet path, not the IP routers on the network path the packet travels from the source to the destination. traceroute is a utility for discovering this path. It works by eliciting responses (ICMP TTL Exceeded messages) from the router 1 hop away from the source towards the destination, then 2 hops away from the source, then 3 hops, and so forth until the destination is reached. The responses will identify the IP address of the router. The output from traceroute normally prints the information for one hop per line, including the measured round trip times and IP address and DNS names of the router. The DNS name is handy for working out the organization to which the router belongs. Since traceroute takes advantage of common router implementations, there is no guarantee that it will work for all routers along the path, and it is usual to see ?*? responses when it fails for some portions of the path.

Using the traceroute output, sketch a drawing of the network path. If you are using the supplied trace, note that we have provided the corresponding traceroute output as a separate file. Show your computer (lefthand side) and the remote server (righthand side), both with IP addresses, as well as the routers along the path between them numbered by their distance on hops from the start of the path. You can find the IP address of your computer and the remote server on the packets in the trace that you captured. The output of traceroute will tell you the hop number for each router.

To finish your drawing, label the routers along the path with the name of the real-world organization to which they belong. To do this, you will need to interpret the domain names of the routers given by traceroute. If you are unsure, label the routers with the domain name of what you take to be the organization. Ignore or leave blank any routers for which there is no domain name (or no IP address).

This is not an exact science, so we will give some examples. Suppose that traceroute identifies a router along the path by the domain name arouter.cac.washington.edu. Normally, we can ignore at least the first part of the name, since it identifies different computers in the same organization and not different organizations. Thus we can ignore at least ?arouter? in the domain name. For generic top-level domains, like ?.com? and ?.edu?, the last two domains give the domain name of the organization. So for our example, it is ?washington.edu?. To translate this domain name into the real-world name of an organization, we might search for it on the web. You will quickly find that washington.edu is the University of Washington. This means that ?cac? portion is an internal structure in the University of Washington, and not important for the organization name. You would write*?University of Washington* on your figure for any routers with domain names of the form ∗.washington.edu.

Alternatively, consider a router with a domain name like arouter.syd.aarnet.net.au. Again, we ignore at least the arouter part as indicating a computer within a specific organization. For country-code top-level domains like '.au' (for Australia) the last three domains in the name will normally give the organization. In this case the organization?s domain name is aarnet.net.au. Using a web search, we find this domain represents AARNET, Australia?s research and education network. The 'syd' portion is internal structure, and a good guess is that it means the router is located in the Sydney part of AARNET. So for all routers with domain names of the form *.aarnet.net.au, you would write 'AARNET' on your figure. While there are no guarantees, you should be able to reason similarly and at least give the domain name of the organizations near the ends of the path.

Turn In: Hand in your drawing, and traceroute output (8 points)

Path from computer to www.osu.edu found by traceroute

```
[Shanqings-MBP:~ shanqinggu$ traceroute www.osu.edu
traceroute to wh-prdosuedu130-vip.it.ohio-state.edu (140.254.112.130), 64 hops max, 52 byte packets
 1  homeportal (192.168.1.254)  3.652 ms  2.923 ms  3.000 ms
 2  75.22.112.1 (75.22.112.1)  21.578 ms  22.627 ms  27.587 ms
 3  71.151.140.33 (71.151.140.33)  22.404 ms  21.445 ms  25.810 ms
 4  71.158.32.142 (71.158.32.142)  20.701 ms  26.956 ms  21.656 ms
 5  71.158.32.141 (71.158.32.141)  23.024 ms  29.531 ms  28.253 ms
 6  12.83.38.21 (12.83.38.21)  22.185 ms
    12.83.38.13 (12.83.38.13)  24.816 ms  22.183 ms
 7  12.123.241.201 (12.123.241.201)  30.387 ms  29.457 ms  29.470 ms
 8  12.244.6.6 (12.244.6.6)  29.580 ms  31.996 ms  29.406 ms
 9  clmbs-r5-et-4-0-0s100.core.oar.net (199.218.20.29)  30.038 ms  30.268 ms  31.047 ms
10  199.18.169.10 (199.18.169.10)  29.199 ms  34.673 ms  30.033 ms
11  socc4-forg6-4.net.ohio-state.edu (164.107.1.129)  29.554 ms  29.864 ms  91.510 ms
12  * * *
```

Drawing based on traceroute

My computer										Destination www.osu.edu
IP address: 192.168.1.254										IP address: 140.254.112.130
Hops 1	2	3	4	5	6	7	8	9	10	11
	sbcglobal.net	Philadelphia.hfc.comcastbusiness.net	?	?	?	?	?	Oar.net	?	ohio-state.edu

Exercise 3: Wireshark DNS

The goal of this exercise is to become familiar with the Domain Name System (DNS).
This exercise uses the Wireshark software tool to capture and examine a packet trace.
This exercise uses a web browser to find or fetch pages as a workload. Any web browser will do.
This exercise uses dig to issue DNS requests and observe DNS responses. dig is a flexible, command-line tool for querying

remote DNS servers that replaces the older nslookup program. It comes installed on Mac OS. On Window, you can download dig from ISC?s BIND web site as part of the bind download. (Note that there may be some dependencies. Check for online instructions to set up *dig* on Windows.) On Linux, install dig with your package manager. It is normally part of a dnsutils or bindutils package.

Perform each of the steps below.

Step 1: Manual Name Resolution

Before we look at how your computer uses the DNS, we will see how a local nameserver resolves a DNS name, i.e., we will interact with remote nameservers. To do this exercise, you will pretend to be the local nameserver and issue requests to remote nameservers using the dig tool.
Pick a domain name to resolve, such as that of your web server. We will use www.smu.edu. Find the IP address of one of the root nameservers by searching the web. For example, the Wikipedia article on root name servers includes the IP address of the root nameservers a through m. Any one of these should do, as they hold replicated information. You need this information to begin the name resolution process, and nameservers are provided with it as part of their configuration.

Use dig to issue a request to a root nameserver to perform the first step of the resolution. You are assuming that you have no cached information that will let you begin a resolution below the root. The format of a dig command is dig @aa.bb.cc.dd domainname. It instructs dig to send a request to a nameserver at a given IP address (or name) for the given domain name. The reply from the root does not provide the full name resolution, but it does tell us about nameservers closer to having the information for you to contact.

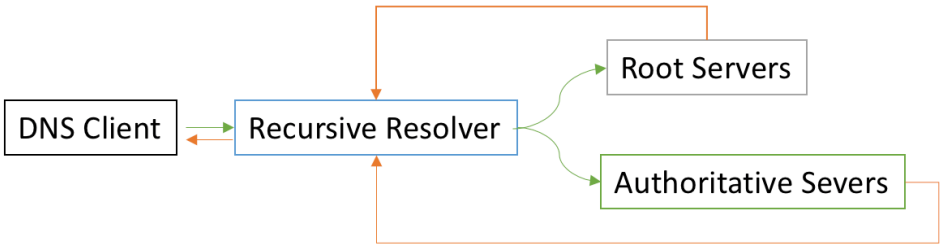
Continue the resolution process with dig until you complete the resolution. When you have alternatives to choose, prefer IPv4 nameservers and select the first one in alphabetical order. If this nameserver has multiple IP addresses then select the numerically smallest IP address. You can complete the resolution without these tie-breaking rules and will likely obtain the same result since the DNS information is replicated. Keep these dig commands handy, as you will repeat them in the next step when you capture a trace.

Draw a figure that shows the sequence of remote nameservers that you contacted and the domain for which they are responsible. Note that future name resolutions are likely to be a much shorter sequence because they can use cached information. For example, if you looked up a domain name in ?.edu? then when you look up a different domain name in ?.edu? you already know the name of the ?.edu? nameserver. Thus you can start there, or even closer to the final nameserver depending on what you have cached; you do not need to start again at the root nameserver.

Turn In: Hand in your drawing (8 points)

```
[Shanqings-MBP:~ shanqinggu$ dig @192.41.162.30 www.smu.edu

; <<>> DiG 9.10.6 <<>> @192.41.162.30 www.smu.edu
; (1 server found)
;; global options: +cmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 15820
;; flags: qr rd; QUERY: 1, ANSWER: 0, AUTHORITY: 4, ADDITIONAL: 5
;; WARNING: recursion requested but not available

;; OPT PSEUDOSECTION:
; EDNS: version: 0, flags:; udp: 4096
;; QUESTION SECTION:
;www.smu.edu.                   IN      A

;; AUTHORITY SECTION:
smu.edu.                172800  IN      NS      pony.cis.smu.edu.
smu.edu.                172800  IN      NS      seas.smu.edu.
smu.edu.                172800  IN      NS      xpony.smu.edu.
smu.edu.                172800  IN      NS      epony.smu.edu.

;; ADDITIONAL SECTION:
pony.cis.smu.edu.       172800  IN      A       129.119.64.10
seas.smu.edu.           172800  IN      A       129.119.3.2
xpony.smu.edu.          172800  IN      A       129.119.64.8
epony.smu.edu.          172800  IN      A       128.42.182.100

;; Query time: 63 msec
;; SERVER: 192.41.162.30#53(192.41.162.30)
;; WHEN: Sat May 26 18:05:56 EDT 2018
;; MSG SIZE  rcvd: 186
```

Drawing



Briefly review how a query recursively receives a response in a typical DNS resolution scenario:

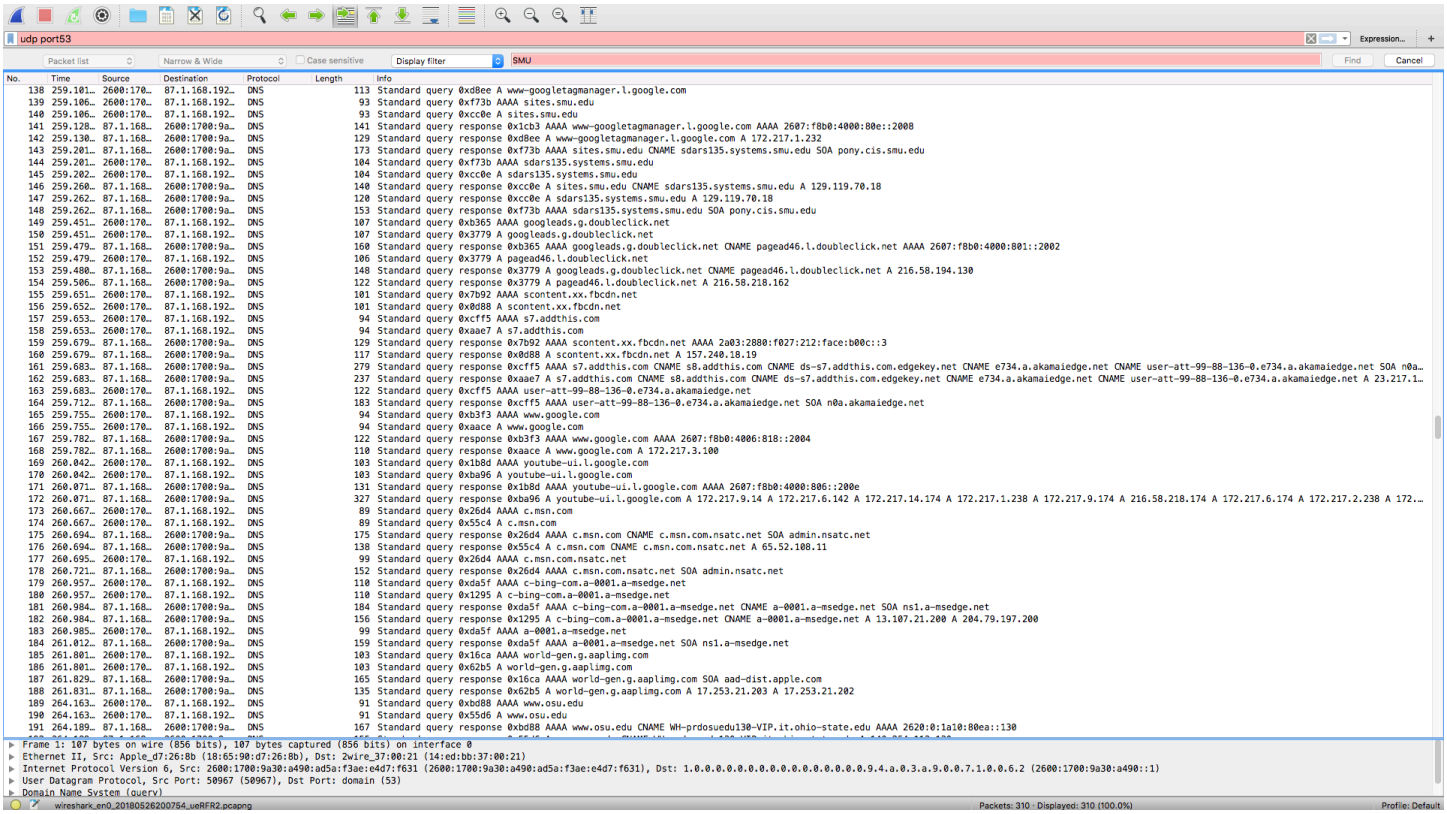
1. Local server (DNS client or stub resolver) queries recursive resolver for www.smu.edu
2. Recursive resolver queries the root name server for www.smu.edu
3. The root name serve refers my recursive resolver to the .com Top-Level Domain (TLD) authoritative serve.
4. My recursive resolver queries the .com TLD authoritative server for www.smu.edu
5. The .com TLD authoritative server refers my recursive server to the authoritative serves for www.smu.edu
6. My recursive resolver queries the authoritative servers for www.smu.edu and receives 192.41.162.30 as the answer.
7. My recursive resolver caches the answer for the duration of the time-to-live (TTL) specified on the record and returns to me.

Step 2: Capture a Trace

Capture a trace of your browser making DNS requests as follows. Now that we are familiar with the process of name resolution, we will inspect the details of DNS traffic. To generate DNS traffic you will both repeat the dig commands and browse web sites.

1. 1) Close all unnecessary browser tabs and windows. Browsing web sites will generate DNS traffic as your browser resolves domain names to connect to remote servers. We want to minimize browser activity initially so that we capture only the intended DNS traffic.
2. 2) Launch Wireshark and start a capture with a filter of *udp port 53.* We use this filter because there is no shorthand for DNS, but DNS is normally carried on UDP port 53. Select the interface from which to capture as the main wired or wireless interface used by your computer to connect to the Internet. If unsure, guess and revisit this step later if your capture is not successful. Uncheck 'capture packets in promiscuous mode'. This mode is useful to overhear packets sent to/from other computers on broadcast networks. We only want to record packets sent to/from your computer. Leave other options at their default values. The capture filter, if present, is used to prevent the capture of other traffic your computer may send or receive. On Wireshark 1.8, the capture filter box is present directly on the options screen, but on Wireshark 1.9, you set a capture filter by double-clicking on the interface.
3. 3) Repeat the dig commands from the previous step. This time, you should see the DNS request and reply packets that correspond to your commands captured in the trace window. Note that there may be some background DNS traffic originating from your computer if any process needs to resolve names to make a network connection. We are assuming that there will be little of this traffic so that you can
4. 4) Wait 10 seconds, then open your browser and browse a variety of sites. Using your browser will generate DNS traffic as you visit new domains, and also as your browser runs its background tasks such as auto-completion. Unlike the dig traffic, this will be DNS traffic between your computer and the local nameserver.
5. 5) Stop the capture when you have a good sample of DNS traffic. We would like enough traffic to see a variety of behavior. DNS traffic is generated fairly quickly as you browse so it should only take a short while to collect this DNS traffic.

Turn In: Hand in your drawing (8 points)



Step 3: Inspect the Trace

To explore the details of DNS packets, select a DNS query expand its Domain Name System block (by using the '+' expander or icon). The first packets should correspond to your dig commands, followed by DNS traffic produced by your browser. Select the first DNS query that corresponds to your dig commands and expand its DNS block. Likely this query is the first packet in your trace, with the first several packets corresponding to your dig commands, followed by other DNS traffic produced by your browser. To check, see if there are several queries that list the domain you chose in the Info column, each followed by a response. We will use these DNS messages to study the details of the DNS protocol. Sometimes there may be other DNS traffic interspersed with these queries due to background activity; you should ignore these extraneous packets.

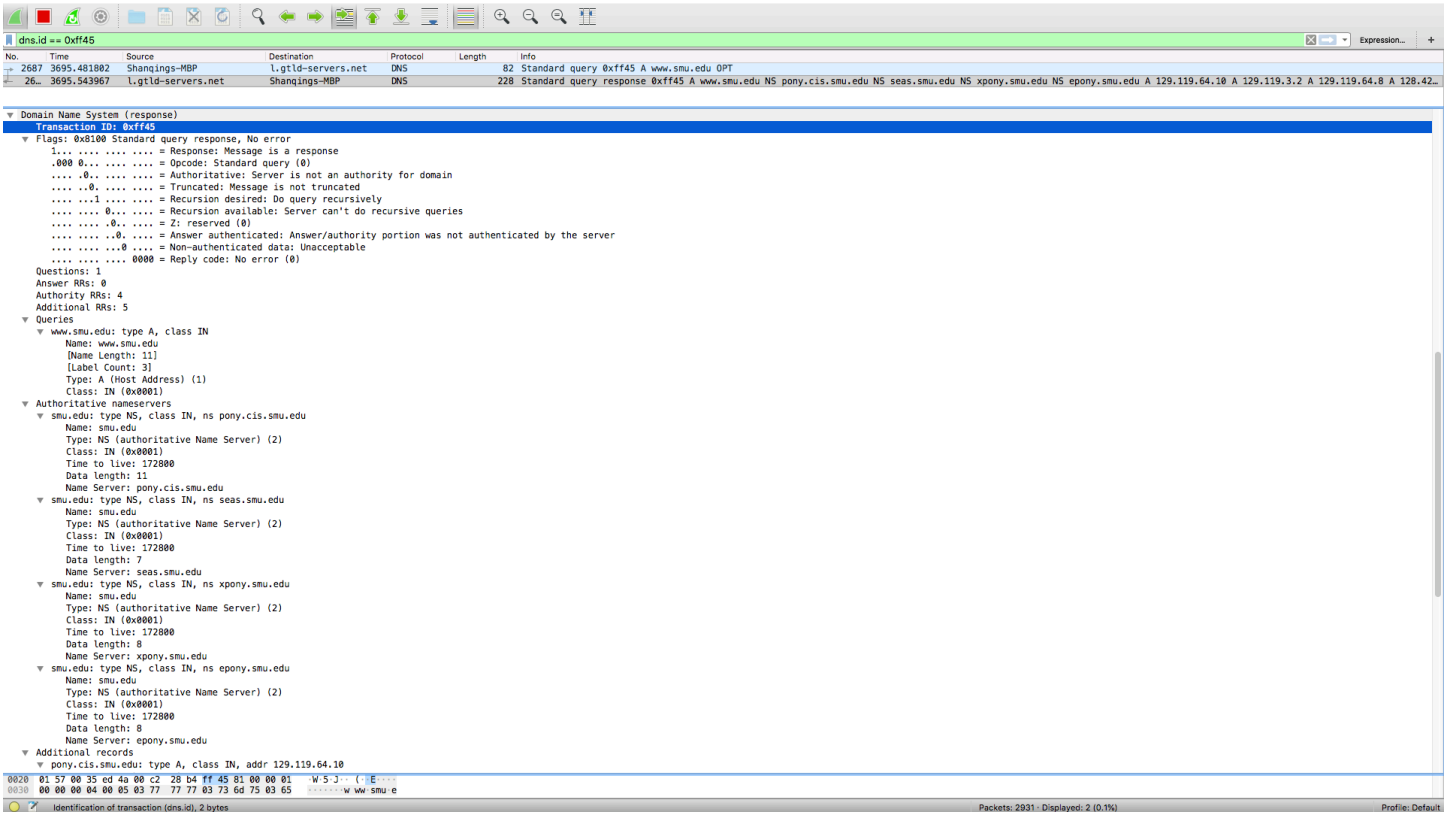
Look at the DNS header, and answer the following questions:

1) How many bits long is the Transaction ID? Based on this length, take your best guess as to how likely it is that concurrent transactions will use the same transaction ID.
2) Which flag bit and what values signifies whether the DNS message is a query or response?
3) How many bytes long is the entire DNS header? Use information in the bottom status line when you select parts of the packet and the bottom panel to help you work this out.

Now examine the responses to the dig DNS queries you made. The initial response should have provided another nameserver one step closer to the nameserver, but not the final answer. You should find that it includes the original query in its Query section. It will also include records with both the name of the nameservers to contact next, and the IP addresses of those nameservers. The final response in this series will include the IP address of the domain name ? this is the answer to the query. Look at the body of the DNS response messages, and answer the following questions:

4) For the initial response, in what section are the names of the nameservers carried? What is the Type of the records that carry nameserver names?
5) Similarly, in what section are the IP addresses of the nameservers carried, and what is the Type of the records that carry the IP addresses?
6) For the final response, in what section is the IP address of the domain name carried?

Turn In: Hand in your answers to the above questions (12 points)



1. The Transaction ID is 16 bits long, which is unlikely there are concurrent transactions. The host computer needs 2^{16} query or response pairs at the same time to cause a collision. This value is an extremely high DNS load for normal computer.

2. The first flag bit signifies query or response. A “0” indicates a query and a “1” indicates a response.

3. The length for DNS header is 12 byes.

4. The names of nameservers are carried in the Authority section in an NS (Authoritative Name Server) record.

5. The IP addresses of the nameservers are carried in the Additional section. The Type of record is A for an IPv4 address.

6. The IP address of the queried domain name is carried in the Answer section.