COL334 Networks Assignment-1 Report

Kanishka Gajbhiye - 2021CS50131
Mrunal Kadhane - 2021CS10109

We performed several tasks to familiarise ourselves with handy tools such as traceroute, nmap, wireshark, ifconfig, etc. used in computer network diagnostics.

1 Network Analysis

a Some snapshots showing use of "traceroute(tracert)" command

Used cellular network (Airtel 4G)

```
C:\Users\Dell>tracert www.iitd.ac.in

Tracing route to www.iitd.ac.in [10.10.211.212]
over a maximum of 30 hops:

  1     4 ms     2 ms     3 ms  192.168.240.64
  2     *      249 ms    98 ms  10.184.32.13
  3   266 ms    93 ms    40 ms  10.254.175.5
  4   187 ms    20 ms    13 ms  10.254.236.26
  5    67 ms   304 ms   103 ms  www.iitd.ac.in [10.10.211.212]

Trace complete.
```

Figure 1: Trace to www.iitd.ac.in

```
C:\Users\Dell>tracert www.google.com

Tracing route to www.google.com [142.250.207.196]
over a maximum of 30 hops:

  1     2 ms     1 ms     2 ms  192.168.125.133
  2    10 ms     9 ms    10 ms  10.194.32.13
  3     9 ms     5 ms     5 ms  10.254.239.1
  4     6 ms     5 ms     7 ms  10.255.1.34
  5     9 ms     6 ms     4 ms  10.119.233.65
  6     *        *        *     Request timed out.
  7     *        *        *     Request timed out.
  8     9 ms     7 ms     6 ms  10.119.234.162
  9     *      228 ms    91 ms  72.14.194.160
 10    19 ms    11 ms    13 ms  108.170.251.97
 11    13 ms    15 ms    29 ms  142.251.76.169
 12    15 ms     7 ms     7 ms  del12s10-in-f4.1e100.net [142.250.207.196]

Trace complete.
```

Figure 2: Trace to www.google.com

b Some interesting observations

1. When running traceroute to www.iitd.ac.in and google.com sometimes the **path defaulted to IPv6**. Many modern networks are designed to support both IPv4 and IPv6 (dual-stack networks). Whether to use IPv6 or IPv4 depends on preference, network policies, network availability, etc. We can force IPv4/IPv6 using command options:

i. -4 for IPv4
ii. -6 for IPv6

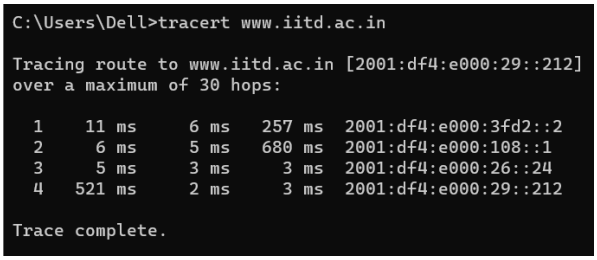


Figure 3: Trace to www.iitd.ac.in (IPv6)

Figure 4: Trace to www.google.com(IPv6)

2. In **MAC, maximum hops are 64 by default whereas in Windows it's 30** by default. This difference is due to the underlying implementation and default behavior of the operating systems.

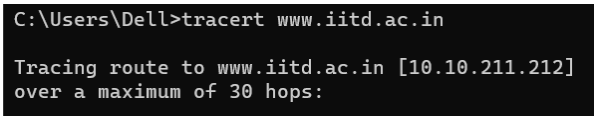


Figure 5: Default hops in Windows

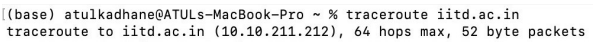


Figure 6: Default hops in MAC

3. We saw **missing routers along the path** that do not seem to reply to the traceroute. These routers are configured to deprioritise or automatically reject ICMP (Internet control message protocol) packets. If we see too many such rows in output, it signifies that network is congested.

Figure 7: Missing routers

4. **Private IP addresses not traceable**. Upon using traceroute on private IP, often found the command stuck in loops.

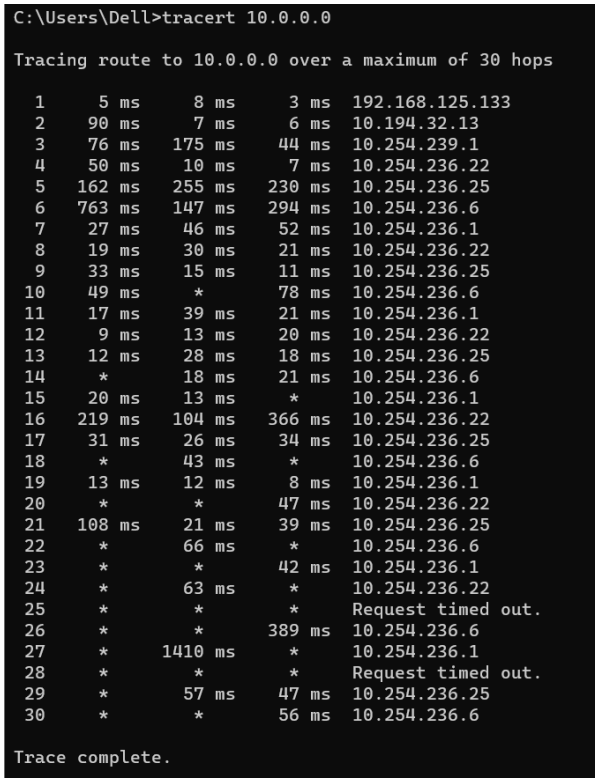


Figure 8: Traceroute stuck in loops

5. **Change in IP address of cellular network used**. This happens due to DHCP(Dynamic Host Configuration Protocol). Whereas IP addresses of popular websites and services, including Google, often remain relatively stable over time. This is done so that they are easily accessible.

6. **Reverse DNS lookup giving modified hostnames**. This is due to distributed nature of host's network. Some servers use load balancing and CDNs to optimize performance and distribute traffic.

Figure 9: Reverse DNS lookup of facebook

c Maximum size of ping packets

Upon using *ping* command, discovered that the maximum packet size that can be sent is **65500 bytes (On windows)**. Sending a packet of this size often results in "request timed out" almost everytime. Factors like network traffic and route can affect ping output.

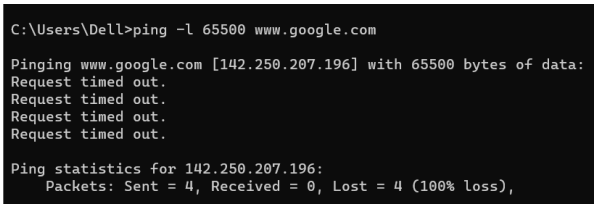


Figure 10: No ping successful

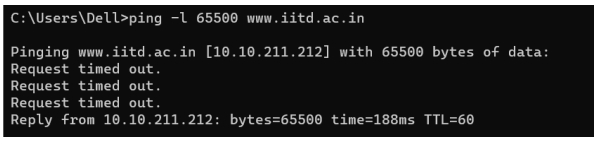


Figure 11: 1 out of 4 pings successful

2 Replicating traceroute functionality using ping

Since the code was made on Windows system, utilized nping to replicate tracert. The snapshots comparing in-built tracert and replicated tracert are shown.(Trace done to www.google.com)

On running in-built tracroute on some ip addresses it leads to failure (eg. 10.119.233.65), although such IPs can be seen if they are in path for some other IP. For such cases, the simulation outputs that the IP can't be traced.

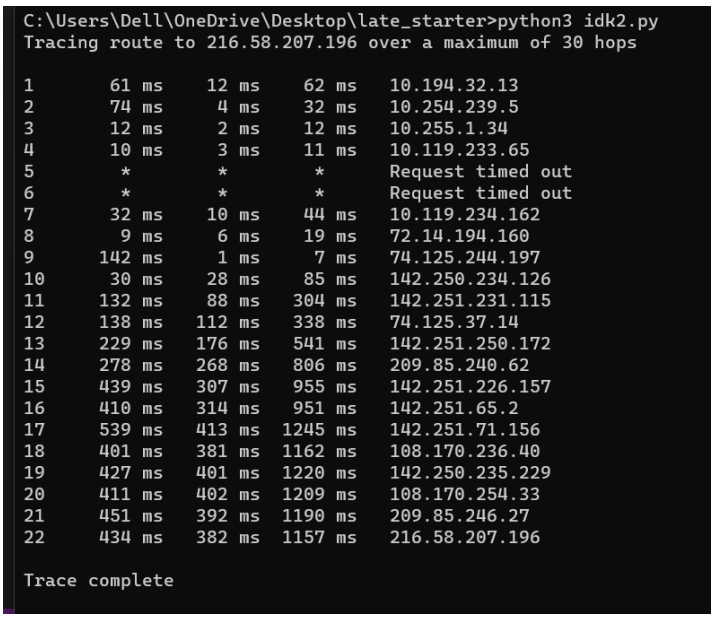


Figure 12: Code o/p replicating traceroute functionality

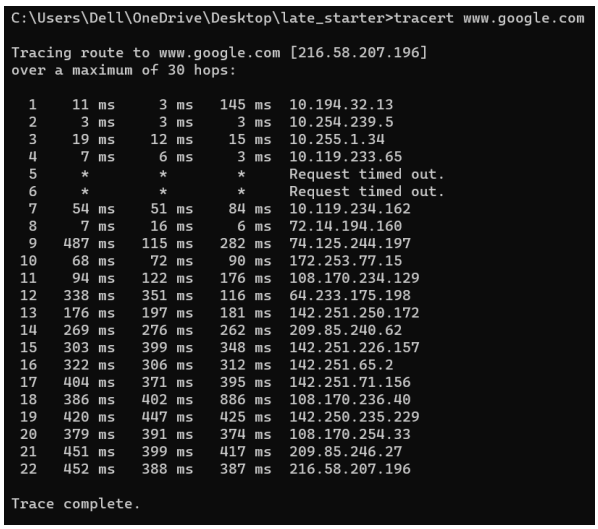


Figure 13: Traceroute output

Note: Python script submitted separately.

3 Internet Architecture

a Commands performed

We used the "Traceroute" command on mac or the "Tracert" command on windows to monitor the path traversed to reach the 5 DNS servers mentioned in the assignment pdf from three different locations

b Observations

A) Consult an AS-IP lookup service to figure out when traffic gets into the local ISP, transits to other intermediate ISPs, and finally into the destination domains.

Answer - AS-IP lookup service refers to a service that allows you to find the **Autonomous System (AS) associated with a specific IP address**. Autonomous Systems are essentially groups of IP networks and routers under the control of a single organization, often an Internet Service Provider (ISP) or a large company. If we closely examine the RTT time from a certain intermediate router to our own device has exceeded like 400ms, we can say that this router is handling heavy traffic and there are chances of packet drops and queing delay. We can use its ip address to find the local area of this router via the AS-IP lookup service.

B) In a neat tabular format, report the number of hops from the (3) traceroute sources to the above (5) destinations. If the pair of (traceroute source, destination) are geographically close to each other, does it roughly translate into fewer hops? Do Google and Facebook differ from the others in the number of hops required to reach them, irrespective of which traceroute source is used? Why would this be so?

Answer - Yes. If the traceroute source is closer to the destination, it results in fewer hops as observed in performing traceroutes to iitd.ac.in via Canada(more than 30) and via my own device in Delhi(4). Traceroutes to Google or Facebook are fewer in number of hops since these are accessed globally and hence it is necessary to reduce the latency for these websites.

Traceroute 1 - Own device

Destination	Hops
google.com	10(64 hops max)
University - utah	34(64 hops max)
Facebook.com	13(64 hops max)
University - IITD	4(64 hops max)
University - cape town	>64

Traceroute 2 - Canada

Destination	Hops
google.com	27(30 hops max)
University - utah	21(30 hops max)
Facebook.com	10(30 hops max)
University - IITD	>30(30 hops max)
University - cape town	>30

Traceroute 3 - USA

Destination	Hops
google.com	28(30 hops max)
University - utah	18(30 hops max)
Facebook.com	10(30 hops max)
University - IITD	>30(30 hops max)
University - cape town	>30

C) Report the latencies between the traceroute sources and the web-servers. Does the latency seem to be related to the number of hops, being higher when there are more hops? Why is this the case?

My device

Location	Latencies(ms)
Google.com	7.254
Facebook.com	36.594
Utah	313.844
Cape Town	420.25
IIT Delhi	3.823

Canada - Montreal

Location	Latencies(ms)
Google.com	46.741(not completely executed)
Facebook.com	8.475
Utah	57.124
Cape Town	215.34(not completely executed)
IIT Delhi	271.23(not completely executed)

USA - Chicago

Location	Latencies(ms)
Google.com	39.37
Facebook.com	18.84
Utah	36.821
Cape Town	196.887(not completely executed)
IIT Delhi	244.384(not completely executed)

Answer - The latencies are summarised in the table and these values are approximate values. It is possible that it may increase or decrease if the traceroute is performed again. A "hop" refers to a single point-to-point communication between two devices, often passing through intermediate routers or switches. Latency refers to the time it takes for data to travel from the source to the destination. Each hop introduces some delay due to the processing time of routers, switches, and other network equipment. More hops suggest that there is not a faster route to that server since such routers weren't placed to build a faster connection. This introduces **propagation delay, queuing delay, processing delay and routing delay**.

D) Which of the destination web-servers are resolved to the same IP address irrespective of from where you do a traceroute to them? Why do you think some web-servers are resolved to different IP addresses when queried from different parts of the world?

Answer - Web servers like that of Universities which are regional and not that heavily accessed on a global level have the same IP addresses for a traceroute from any part of the world. Web servers that are resolved to the same IP address regardless of the traceroute source are often the result of using **content delivery networks (CDNs)**. CDNs are distributed networks of servers that work together to deliver web content to users from the server closest to them geographically. This enhances performance and reduces latency by minimizing the distance data needs to travel.

E) If you do traceroutes from the same starting point to different IP addresses you found for the same web-server, do the paths appear different? Which ones are longer?

Answer - Web servers like that of Google, Facebook, Netflix have different IP addresses from the same starting point since these are heavily accessed and it is necessary to create a shorter route to reach them. That depends on the router closet to them leading it towards the destination router. The paths may or may not differ depending on the route chosen by the packet from the lookup table. Accordingly, the latencies will also change.

F) Try tracerouting to Google and Facebook from different countries of traceroute servers around the world. Are you able to find any countries that do not seem to have their local ISP's directly peered with Google and Facebook?

Answer - Yes, as visible from the above table, Canada seems to have a greater number of hops to Google and Facebook which might suggest that their local ISP's are not directly peered at. Although this may not be true since latency can also be the result from network congestion and other factors. Some servers which are directly peered with Google and Facebook show a direct jump from local ISP to routers in these website's network after some hops.

4 Packet Analysis

a Commands performed

Used wireshark to grab all packets while visiting the HTTP website such as http://act4d.iitd.ac.in from the browser. Performed a flushdns to clear the local DNS cache.

b Observations

A) Apply a "dns" filter on the packet trace, and see if you can find DNS queries and responses for www.iitd.ac.in. How long did it take for the DNS request-response to complete?

Answer - We received 4 dns query - response pairs for iitd.ac.in. Upon observing, we found that the average time between query and response time is around 10 ms.

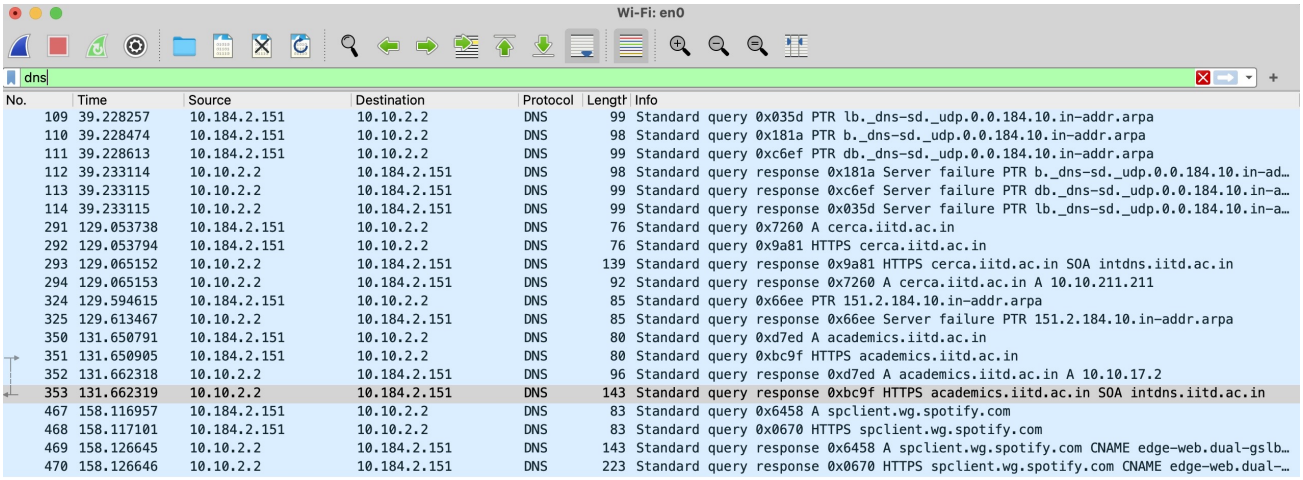


Figure 14: DNS Query

B) Apply an "http" filter on the packet trace and report the approximate number of HTTP requests that were generated. What can you tell from this observation about how web- pages are structured, and how browsers render complex pages with multiple images and files?

Answer - We are capturing packets for the site act4d.iitd.ac.in. As this is not a secure connection, we are able to see http request via which data is getting sent to the receiver. There are a minimum of 10 HTTP requests on wireshark. The site has complex pages with multiple images and javascripts which requires us to send it through different HTTP requests. As the browser parses the HTML, it encounters references to **external resources like images, CSS files, and JavaScript files**. It sends separate requests to the server for each of these resources. These resources are fetched in parallel, which helps improve the loading speed of the page.

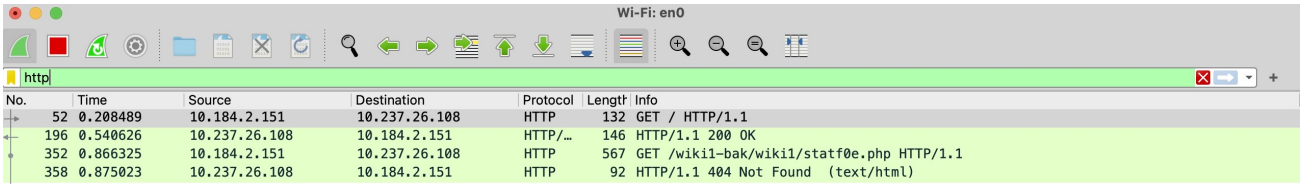
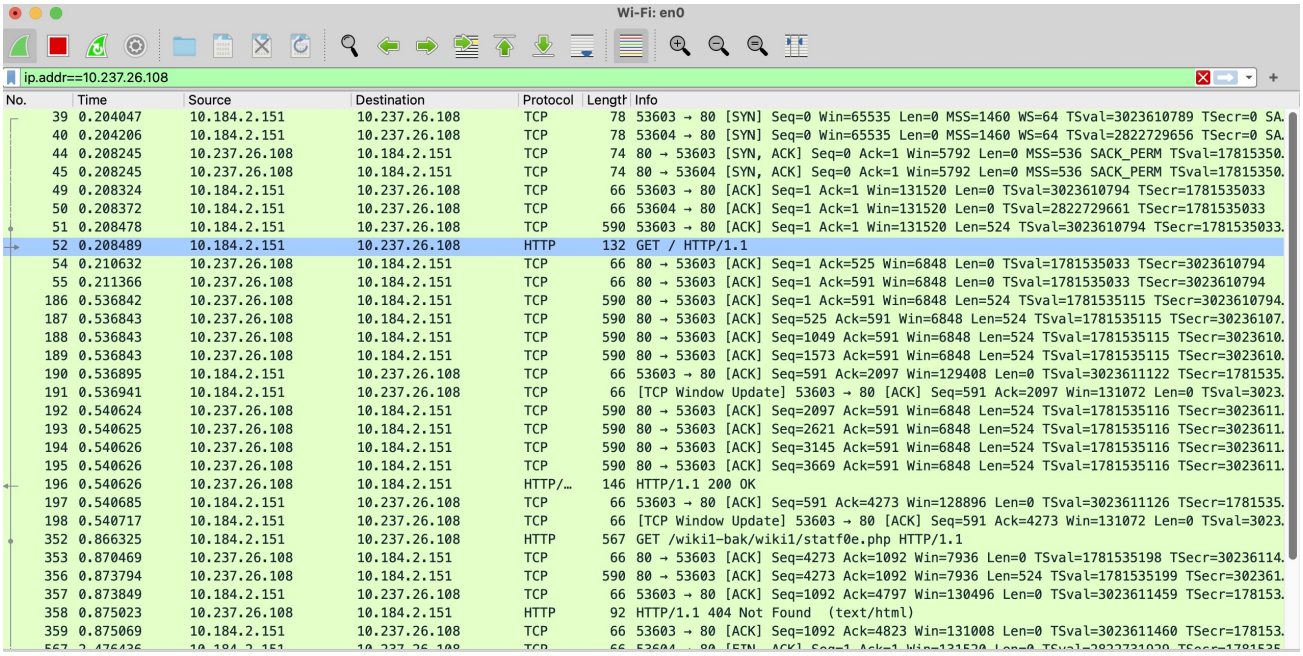


Figure 15: HTTP Request

C) Find the number of TCP connections that were opened between your browser and the web-server. Is this the same as the number of HTTP requests for content objects that you found in the previous part? Do you find that some content objects are fetched over the same TCP connection? Note that TCP connections are distinguished from one another based on the source port and destination port.

Answer - There are a 30 of TCP connection opened between my browser and the web server which is greater than the HTTP requests. Each packet being received by the receiver needs to send an acknowledgemnt message which creates a TCP Connection. The message is chopped at the transport layer and until all the packets are not received for a particular message, the message is not sent out. All the chopped packets for a particular message are assigned the same destination port number and as the packets are received , the transport layer of the receiver checks the sequence numbers along with the port numbers and accordingly waits until it receives that particular packet. The sender is alerted by the receiver through the [ACK] signals for each packet.



No.	Time	Source	Destination	Protocol	Length	Info
39	0.204047	10.184.2.151	10.237.26.108	TCP	78	53603 → 80 [SYN] Seq=0 Win=65535 Len=0 MSS=1460 WS=64 TSval=3023610789 TSecr=0 SA.
40	0.204206	10.184.2.151	10.237.26.108	TCP	78	53604 → 80 [SYN] Seq=0 Win=65535 Len=0 MSS=1460 WS=64 TSval=2822729656 TSecr=0 SA.
44	0.208245	10.237.26.108	10.184.2.151	TCP	74	80 → 53603 [SYN, ACK] Seq=0 Ack=1 Win=5792 Len=0 MSS=536 SACK_PERM TSval=17815350.
45	0.208245	10.237.26.108	10.184.2.151	TCP	74	80 → 53604 [SYN, ACK] Seq=0 Ack=1 Win=5792 Len=0 MSS=536 SACK_PERM TSval=17815350.
49	0.208324	10.184.2.151	10.237.26.108	TCP	66	53603 → 80 [ACK] Seq=1 Ack=1 Win=131520 Len=0 TSval=3023610794 TSecr=1781535033
50	0.208372	10.184.2.151	10.237.26.108	TCP	66	53604 → 80 [ACK] Seq=1 Ack=1 Win=131520 Len=0 TSval=2822729661 TSecr=1781535033
51	0.208478	10.184.2.151	10.237.26.108	TCP	590	53603 → 80 [ACK] Seq=1 Ack=1 Win=131520 Len=524 TSval=3023610794 TSecr=1781535033.
52	0.208489	10.184.2.151	10.237.26.108	HTTP	132	GET / HTTP/1.1
54	0.210632	10.237.26.108	10.184.2.151	TCP	66	80 → 53603 [ACK] Seq=1 Ack=525 Win=6848 Len=0 TSval=1781535033 TSecr=3023610794
55	0.211366	10.237.26.108	10.184.2.151	TCP	66	80 → 53603 [ACK] Seq=1 Ack=591 Win=6848 Len=0 TSval=1781535033 TSecr=3023610794
186	0.536842	10.237.26.108	10.184.2.151	TCP	590	80 → 53603 [ACK] Seq=1 Ack=591 Win=6848 Len=524 TSval=1781535115 TSecr=3023610794.
187	0.536843	10.237.26.108	10.184.2.151	TCP	590	80 → 53603 [ACK] Seq=525 Ack=591 Win=6848 Len=524 TSval=1781535115 TSecr=30236107.
188	0.536843	10.237.26.108	10.184.2.151	TCP	590	80 → 53603 [ACK] Seq=1049 Ack=591 Win=6848 Len=524 TSval=1781535115 TSecr=3023610.
189	0.536843	10.237.26.108	10.184.2.151	TCP	590	80 → 53603 [ACK] Seq=1573 Ack=591 Win=6848 Len=524 TSval=1781535115 TSecr=3023610.
190	0.536895	10.184.2.151	10.237.26.108	TCP	66	53603 → 80 [ACK] Seq=591 Ack=2097 Win=129408 Len=0 TSval=3023611122 TSecr=1781535.
191	0.536941	10.184.2.151	10.237.26.108	TCP	66	[TCP Window Update] 53603 → 80 [ACK] Seq=591 Ack=2097 Win=131072 Len=0 TSval=3023.
192	0.540624	10.237.26.108	10.184.2.151	TCP	590	80 → 53603 [ACK] Seq=2097 Ack=591 Win=6848 Len=524 TSval=1781535116 TSecr=3023611.
193	0.540625	10.237.26.108	10.184.2.151	TCP	590	80 → 53603 [ACK] Seq=2621 Ack=591 Win=6848 Len=524 TSval=1781535116 TSecr=3023611.
194	0.540626	10.237.26.108	10.184.2.151	TCP	590	80 → 53603 [ACK] Seq=3145 Ack=591 Win=6848 Len=524 TSval=1781535116 TSecr=3023611.
195	0.540626	10.237.26.108	10.184.2.151	TCP	590	80 → 53603 [ACK] Seq=3669 Ack=591 Win=6848 Len=524 TSval=1781535116 TSecr=3023611.
196	0.540626	10.237.26.108	10.184.2.151	HTTP/…	146	HTTP/1.1 200 OK
197	0.540685	10.184.2.151	10.237.26.108	TCP	66	53603 → 80 [ACK] Seq=591 Ack=4273 Win=128896 Len=0 TSval=3023611126 TSecr=1781535.
198	0.540717	10.184.2.151	10.237.26.108	TCP	66	[TCP Window Update] 53603 → 80 [ACK] Seq=591 Ack=4273 Win=131072 Len=0 TSval=3023.
352	0.866325	10.184.2.151	10.237.26.108	HTTP	567	GET /wiki1-bak/wiki1/statf0e.php HTTP/1.1
353	0.870469	10.237.26.108	10.184.2.151	TCP	66	80 → 53603 [ACK] Seq=4273 Ack=1092 Win=7936 Len=0 TSval=1781535198 TSecr=30236114.
356	0.873794	10.237.26.108	10.184.2.151	TCP	590	80 → 53603 [ACK] Seq=4273 Ack=1092 Win=7936 Len=524 TSval=1781535199 TSecr=302361.
357	0.873849	10.184.2.151	10.237.26.108	TCP	66	53603 → 80 [ACK] Seq=1092 Ack=4797 Win=130496 Len=0 TSval=3023611459 TSecr=178153.
358	0.875023	10.237.26.108	10.184.2.151	HTTP	92	HTTP/1.1 404 Not Found (text/html)
359	0.875069	10.184.2.151	10.237.26.108	TCP	66	53603 → 80 [ACK] Seq=1092 Ack=4823 Win=131008 Len=0 TSval=3023611460 TSecr=178153.

Figure 16: TCP Request

D) Now try doing a trace for http://www.indianexpress.com and filter for "http". What do you find, is there any HTTP traffic? Browse through the entire trace without any filters, are you able to see the contents of any HTML and Javascript files being transferred? What just happened?

Answer - Doing a trace for Indian Express website, we see **no HTTP requests** on wireshark. Indian Express has a secure connection and hence all the data being transferred is encrypted. It uses the TLSv1(Transport Layer Security version 1) protocol. TLSv1 is a **cryptographic protocol that provides secure communication over a network. It ensures data confidentiality, integrity, and authenticity between two parties (client and server) by encrypting the data being transmitted**. It uses encryption algorithms and cryptographic techniques to protect the data from eavesdropping and tampering. Hence we are not able to see the contents of the HTML and Javascript files being transferred.

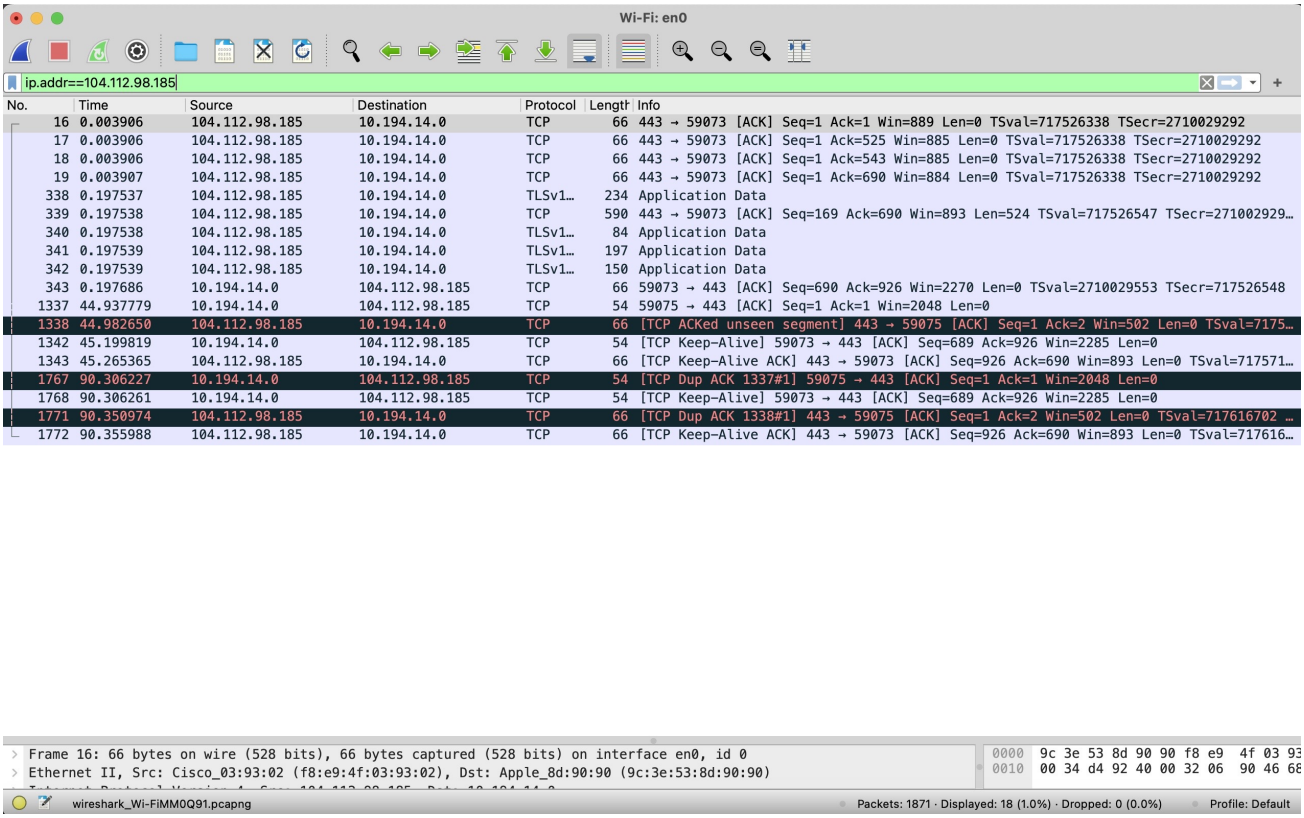


Figure 17: Analysing trace to Indian Express Website