REPORT
VISUALIZING NETWORK TOPOLOGY

Deliverables:

In this assignment, we have used 5 sources and 10+ destinations to visualize Internet Topology.
Sources are: -

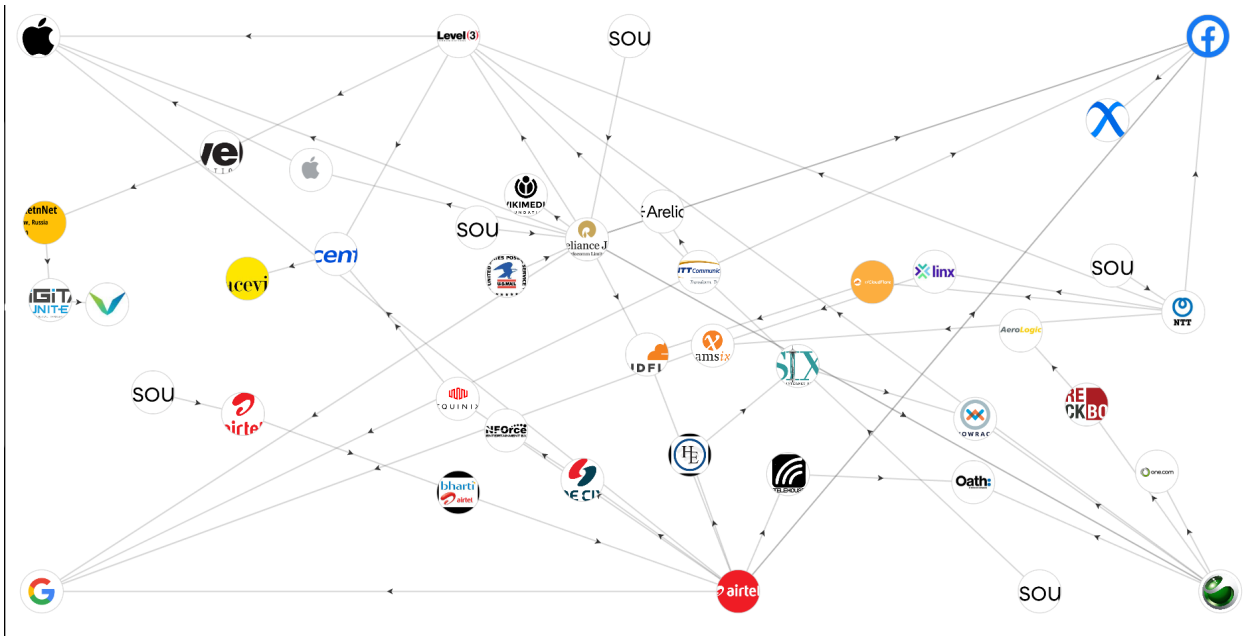
1. Jio Source 1
2. Jio Source 2
3. Airtel Source
4. Looking Glass Paris
5. Looking Glass Singapore

We are conducting traceroutes to various destinations from these sources. We collected the IPv6 addresses of different hops along the route from source to destination. We used WHOIS database and API to extract various information like Autonomous System Number (ASN), ISP name, Organization name, Geolocations etc.
(The raw data sheet is provided in the uploaded zip file.)

We used python code to extract the above mentioned information from the API. Various libraries like ***urlopen, numpy, pandas and os*** were used to extract tracerouting information directly from command line to the excel sheet just by giving the website name.
(The code file for this process is included in the uploaded zip file.)

We utilized flourish.studio, a visualization platform, to create data visualizations.
The visualization is presented as a directed graph, where each node represents an Internet Service Provider (ISP) and is connected to its peering ISPs. Each node has an image representing the particular organization and there are five nodes with an image showing SOU which represents the source. The data we used for visualization consists of two Excel sheets (one containing two columns representing node connections, and the other containing ISP names and their image addresses).

The link for visualization is :
https://flourish-user-preview.com/14822967/wwgMzUUL8718zhXaU4sHICQgsn9hYYvtqb-VB3l0udg0JnBFEde71o4ZITHRIY5P/



Note:- Please view the link, it is in interactive mode, image may seem naive but when you click on node you can check the respective peers and direction of packet flow.
We also used ASN to find the IP ranges along with the organization name. (It is in the uploaded zip file.)

Findings:

While making internet topology visualization we found that when we ping a website from our device, it doesn't go directly to the destination, it goes through different local and regional ISP's to reach the destination.

The path taken to reach the destination may not be the shortest path in terms of time and may not be the same every time we ping it.

When we used traceroute for IPv4 we were getting many private IP addresses(192.168.0.0-192.168.255.255) . We were not able to find data of them in whois database so we used IPv6 and were able to find the AS number and other related data using whois database. We also found that not every website has an IPv6 address such as store.steampowered.com, napster.com etc. Most of the IPv6 addresses were based in the USA and many of them were tech based companies. It was hard to identify which IP address is IXP, we had to look up the IXP database which was also limited.

Most of the early IP addresses were of network providers of particular source. There were some IP's with ASN 0, some of which were IXP, which shows that IXP rejects the packets from the traceroute.

While using different sources we found that IITH Ethernet is giving many requests timed out. It is because of security issues as it was blocking our request. Not only for IITH Ethernet but also for different local ISPs we faced the same problem. It might have happened due to problems on the return path or the destination's security device blocking your request.

Note:- All the files used for the assignment are given in the zip file.

Thank you:)

Deshmukh Divyansh Anil	ES21BTECH11012
Jagadeesh S	ES21BTECH11026
Saurabhkumar B Patel	ES21BTECH11028