[4 marks] **IP addresses:**

1. **Looking at the addresses found for the different sites in Part 1, do some sites have more than one IP address, and what does it mean if they do?**

- There are several sites that have more than one IP address when executing dnslookup program for example "ebay.vn" at the time had 2 IPv4 addresses as 66.211.181.235 and 66.135.211.132. This implies that the site has multiple servers that registered with DNS server under one domain name. This also means that the site/domain is capable of implementing load balancing mechanism to balance the workload between servers. Even though it is not completely fault-proof by design, load balancing would really improve the site's reliability.

2. **If you run your dnslookup program several times, do you always get the same IP addresses for a site, and if not, why might this be?**

- When running dnslookup several times, for some sites, the IP addresses stay the same while for some others, the IP addresses changed in a rotation. The IP addresses stay the same for some sites is probably because either they don't have multiple IPs capacity or they have implemented CDN service, with which if is synchronised properly with DNS server would allow another layer of load balancing and user-oriented content distribution. For some other sites, the IPs either rotated around or changed randomly in a finite set. This change in the IP addresses is due to the nature of dnslookup which will return all of the IPs that is made available to the request. The dnslookup would then recursively sending requests to DNS server to query about the domain and trigger a load balancing mechanism for which is called "DNS round robin". This mechanism is implemented as a simple solution to balance the traffic load between the available servers. For each time the DNS server being hit with a DNS request, it returns one of the available IP of that domain and this would be shuffled around. Usually, if a browser sends the DNS request, the server would only reply with a single shuffled IP address. But the way dnslookup works would sometimes show up multiple addresses for some domains.

3. **Do you get the same IP addresses for a site if you run your dnslookup program from different locations, and if not, why?**

- When running dnslookup for a site at different locations, it is possible to retrieve multiple IPs.This is due to differences between ISPs that operating the network at the locations. Additionally, this could also be a result of implement CDN system as mentioned above, DNS server would return the IP of such system as an attempt to direct workload to a server that was prespecified to be dedicated for the dnslookup request.

4. **What proportion of sites has an IPv6 address?**

- For the sample sites that were used in this coursework, there were 6/22 sites that have implemented IPv6. This makes the proportion of IPv6 address is approximately 27% within this sample size of this report.

[4 marks] • **Router-level Topology Maps:**

- The longest path that was illustrated on the map for IPv4 is from "130.209.240.48" to "115.84.180.145" with 20 hops in between and for IPv6 is from "2001:630:40:f00:e22f:6dff:fe2c:ed80" to "2a03:2880:f029:11:face:b00c:0:2" with 12 hops in between.
- Paths from different locations are disjoint
- There were multiple routes to some destinations such as from "27.68.250.113" to "115.84.180.145" and"193.62.157.22" to "108.170.238.123".
- In terms of organisational boundaries, it could be seen on the IPv4 map, the first 3 nodes that start with "130.209.x.x" is operated by "University of Glasgow", "146.97.x.x" is operated by JANET, "62.115.x.x" is operated by TALIANET, "27.68.x.x" indicated the IPs is of Vietnamese origin and "203.133.x.x" is the military network for VIETTEL. From this investigation, it is reasonable to conclude that changes in IP prefixes would imply the changes in network operators.

[2 marks] • **IPv4 and IPv6:**

- In general, the topology map of IPv4 matches the map of IPv6 and this is as expected because most of the time, even if the infrastructure is capable of running on IPv6, it would probably still have the IPv4 running side by side with IPv6.
- This parallelism is due to the fact that it is practically impossible to switch all of the current systems from IPv4 to IPv6 overnight and the duration of this progress would in fact last about 10 more years. During this time, isolated IPv6 systems would need to use IPv4 to connect to each other and other IPv4 nodes at higher levels. This would result in the similar topology map structure as illustrated in the two maps.

[2 marks] • **The** traceroute **Tool:**

- The traceroute tool works by sending a dummy packet through network nodes with a specified TTL value to inform the nodes when to discard the packet and return a TTL Time Exceeded Message and eventually "ICMP Destination/PORT Unreachable" message which would indicate the destination has been reached.
- There are several types of traceroute commands such as UDP, ICMP and TCP. As default Linux machines would use UDP traceroute. When running this, the host machine would send out a UDP packet containing the source IP, which is the host's IP itself, the destination, which is the domain/IP in the query, and finally a random UDP port number from 33434 to 33534. This UDP packet would then be sent with TTL =1. When the packet reaches the host's internet gateway, TTL value would then be reduced by 1 which would be 0. The gateway would then reply to the host with a TTL Time Exceeded Message along with some information which contains its IP address. This is how the host comes to aware of the IP addresses of the nodes on the path to the destination. This process would then continue with another UDP packet from the host with TTL =2 and then 3 and so on. With every response, the host would increase the TTL value of the UDP packet until the packet reaches the destination and the destination would then reply with "ICMP Destination/PORT Unreachable" message and terminate the command. This exiting mechanism is bound to happen as the UDP port has always been in between 33434 to 33534.
- In another approach, Windows machines would use ICMP as the default protocol for traceroute command. ICMP traceroute is very similar to UDP but instead of UDP packet, the host machine would send out an "ICMP Echo Request", the hops in between would reply with an "ICMP Time Exceeded" and the final destination would reply with an "ICMP Echo". ICMP could really be helpful while using traceroute on networks that block UDP broadcasting.