Networked Systems - Assessed Exercise 2

Peter Dodd
2308057D

IP addresses

Looking at the IP addresses found from the execution of dnslookup, there is a varied range of responses. Some sites have multiple addresses for a single host name. This means that the site is hosted on multiple servers, which is used for load balancing so one server doesn't get overloaded.

Some bigger sites, such as Google or Facebook, that are also hosted on multiple servers will only return a single address, or one address for each of IPv6 and IPv4. This address might also change each time dnslookup is executed. This is because these websites run over a Content Distributed Network. The goal of CDNs is to reduce latency time and to balance loads so a central server doesn't become overloaded. A single data centre would not be able to handle the amount of requests a website like this would receive. The CDN is made up of caches in different servers across the world that cache the website's information. The CDN will return an IP address to a local cache instead of that to the main server. This IP address may change so one cache doesn't become overloaded.

If the website is hosted on a single server, then no matter what location you're in, the IP address will remain the same. If the website is hosted on multiple servers or uses a CDN, then the IP address use may change with different locations depending on what address the DNS thinks is closest to you based on your IP address.

Out of the 17 hostnames used in my dnslookup, 12 of them, around 71%, had IPv6 addresses as well as IPv4. This is substantially higher than the proportion of users using IPv6 in the UK, which is currently around 35%.

Router-level Topology Maps

The longest path in both the maps is 18 nodes long, and is present in the IPv6 map. It connects the IPs of 2001:630:40:f00:e22f:6dff:fe2c:ed80 and 2001:41d0:0:50::a:5a0f. It is expected that IPv6 would have the longest route, as there are less routers that support it.

The majority of different paths from varying sources to a single destination are disjoint, although there are some examples in the maps that are not. Some destinations have multiple routes, although the majority again do not.

The IP addresses for certain areas of the maps share prefixes. Clear examples of these on the IPv4 map are 146.97, 130.209 and 108.170 amongst others. Each of these prefixes belong to a different Autonomous System (AS), who are typically internet service providers (ISPs). Each AS will have a range of IPs that systems within it can use, hence why they share a prefix. From this, it is possible to infer which parts of the network are operated by different ISPs.

IPv4 and IPv6

The IPv4 and IPv6 router-level topology maps are not identical, but have similarities. Close to the source of each, the graphs share the same structure. However, the further away the nodes are from the source, the more changes there are between the two maps.

The only reason the maps would be identical is if every router in the network used both IPv6 and IPv4. However, this is not the case, meaning that paths that are possible through IPv4 have no guarantee to be possible through IPv6. This is what creates the differences between both the maps.

The Traceroute Tool

Traceroute works by utilising the Time To Live (TTL) property in IP headers. This is used to let a router know when it should discard a packet, and is useful for if a packet gets caught in a router loop. The TTL is set initially at some value. With each router that receives the packet, this value is decreased by one. When TTL is equal to one, the packet is discarded and the router sends back a 'time exceeded' response to the source of the packet.

This is used in the traceroute call to be able to determine the IP addresses of each of the routers in the path. Initially, traceroute sends out a packet to the destination IP with TTL equal to one. This means that it gets to the first router in the path, then is discarded with the error response sent back to the source. Once this is received, the TTL is increased to two, and another packet is sent. This then gets the IP address of the next router in the path. This is repeated until either the destination is reached, or the something in the network blocks the functionality for traceroute.

The 'time exceeded' response messages sent by the routers is done so through the Internet Control Message Protocol (ICMP). It relates to IP, UDP, and TCP as it is a protocol. Unlike UDP and TCP, it is not used for the transfer of data between users in network applications, but rather by routers to return meaningful responses for different errors or operations that run on the network. One of the only uses of ICMP in network applications is through the use of the traceroute call. The Windows version of the traceroute command, tracert, sends packets out in ICMP, whilst Mac OS and Linux send them over UDP. All versions of the command receive response from routers in ICMP.