Network Systems Report

IP Addresses

Upon investigating of the IP addresses, I noticed that a single website did often have multiple addresses. Upon research I gathered that the reason a website can have multiple IP addresses is because the sites can then be hosted at multiple locations. This is useful for load balancing and to compensate for a host that's down at that moment. It can also be to serve web pages based on the user location. Similarly, this is why sometimes the IP address for a site can change. When a website is requested the request in handled by a DNS server. The DNS server (probably using round-robin) then decides what IP address to give the user and therefore the application just uses the IP address it is given. This allows websites to balance load from multiple concurrent users. Approximately 36% of the IP address that my application retrieved were IPv6 while the rest were IPv4.

Router-level Topology Maps

The longest path was 25 nodes. Paths from different locations to the same destination are often disjoint and there are many destinations which can be accessed from multiple different routes. Looking at the IPv4 topology map clear organizational boundaries can be inferred. The bottom left is dominated by one ISP who owns the prefix 219. Similarly, in the IPv6 map, almost the entire left side of the map is dominated by an ISP that operates the prefix 2001.

IPv4 and IPv6

The IPv4 and IPv6 router level network topology maps do not match exactly; However, they have a very similar structure and in fact in my case, look like mirror images of each other. This is to be expected because I ran the traceroute on both IPv4 and IPv6 addresses from the same location and so the network maps turned out similar in structure.

The Trace Route Tool

Trace route is a network diagnostic tool that is primarily used to inspect the entire path that a packet travels through, the names and identity of routers and devices in your path and the time taken to send and receive data to each device on the path

To understand the Trace route tool, you must understand the concept of a packet's TTL. Each IP packet that you send on the internet has got a field called TTL. TTL stands for "Time To Live". It's the maximum number of hops that a packet can travel through across the internet, before its discarded. The TTL value is set by the sender. Each router that comes in between the source and destination will go on reducing the TTL value by 1 before sending to the next router. If a router receives a packet with TTL of 1 (which means no further traveling, and no forwarding), the packet is discarded. However, the router which discards the packet will inform the original sender that the TTL value has exceeded. The router that discarded the packet will send an ICMP "TTL exceeded message". ICMP (Internet Control Message Protocol) is an error-reporting protocol that network devices like routers use to generate error messages to the source IP address when network problems prevent delivery of IP packets. Hence when an ICMP TTL exceeded message is sent by a router, the original sender will know the address of the router.

Therefore, to get information about every device between the source and the destination, Trace Route works by initially setting the TTL for a packet to 1, sending it towards the requested destination host, and listening for the reply. When the initiating machine receives a "time exceeded" response, it examines the packet to determine where the packet came from - this identifies the machine one hop away. Then the tracing machine generates a new packet with TTL 2 and uses the response to determine the machine 2 hops away, and so on. This way can the traceroute program receives info about every intermediate device between the source and destination like the distance in hops and the time taken for a response.