**An Experimental Visualization of the Internet Topology**

|  |  |
| --- | --- |
| Thomas Kennedy  Southern Methodist University tkennedy@smu.edu | Alec Siems  Southern Methodist University asiems@smu.edu |
| Jacob Adkins  Southern Methodist University jpadkins@smu.edu | Vipul Kohli  Southern Methodist University vkohli@smu.edu |

# ABSTRACT

The outcome of this experiment is a visualization of the topology, media, and delays related to sending packets across all the major routing nodes and links of the Internet. This visualization defines peering between local Internet Service Providers (ISPs) and backbone ISPs as the infrastructure which makes worldwide communication possible. As the nature of networking is based on more on economics rather than geography, the design of this visualization is completely original to better reflect the topology and connectivity of the Internet rather than simply the geography of its hosts. This paper experimentally analyzes the data aggregated from the tools and registries of IP2Location, BGPlay, Whois, and Graphviz in order to visualize the Internet connections between all the countries and continents of the world.

# INTRODUCTION

# RELATED WORK

Cheswick and Burch [1] collected Traceroute data of more than a hundred thousand networks every day since 1998. While their project included a visualization, their main purpose was to portray the growth of the Internet over time. Our experiment, instead, was to focus on mapping the Internet topology in its current state by specifically characterizing the major ISP nodes and their peering links, a concept beyond the scope of Cheswick and Burch’s research. However, we compared our findings to their work on a general level. #need to expand here after experiment#

In February 2012, H.D. Moore [2] attempted to contact every IP address on the Internet. Out of the 3.7 billion IP addresses contacted, over 3 million responses were received, amounting to over two terabytes worth of data. His results give a good representation of the number of IP addresses accessible to everyone in the world. However, our timeframe and computing resources did not allow for such an extensive data collection and analysis as Moore’s, so chose to limit our experimental scheme to the most meaningful IP addresses and URLs to draw the Internet to each country in the world.

Two possible resolution directions have appeared to improving the current Internet architecture’s scalability problem: separation from the main transit core and elimination of the edge networks. We may see which is empirically implemented through our research.[[1]](#footnote-1)

There is one popular tool that will help up locate the longitude and latitude of a particular IP address. The developer created a database of CSV files to help users pinpoint any IP address that is within the database. This will come in handy after we execute all traceroute commands and need to fill in our map.[[2]](#footnote-2)

3. EXPERIMENTAL METHODOLOGY

This sections describes the tools and procedure used to collect and analyze ISP and routing data in order to produce a visualization of the Internet to every country and continent which shows the topology, material, and routing involved.

#expand here

3.1 Data Collection and Analysis Software   
 In order to map the Internet to all of the countries in the world, a list of IP addresses of servers in each country in the world was required. Therefore, we chose a list of Uniform Resource Locators (URLs) of the Web sites of universities in every country provided by Univ.cc [3]. We targeted university Web sites over commercial Web sites since universities are more likely to have their Web servers in-house whereas commercial Web sites mostly outsource their server needs. Each country’s list from Univ.cc is proportional to the country’s area. For example, larger countries like Argentina has 83 URLs while a smaller country like Andorra has only 1 university.

After the URLs are collected, the next task is to determine the Internet routing to each of the URLs. The IP2Location Traceroute [4] software converts these URLs to their respective IP addresses and determines the route to each of these IP addresses from the IP address of the chosen IP2Location host server. The IP2Location Traceroute application allows users to choose a server either in the United States (Phoenix, Dallas, or Los Angeles), Canada (Montreal), France (Paris), Germany (Dusseldorf), Netherlands (Amsterdam), England (London), Singapore, Japan (Tokyo), or Malaysia (Kuala Lumpur), a total of twelve servers in ten countries on three continents. To originate in many other countries servers were chosen from Traceroute.org [5] but first preference was given to IP2Location. Traceroute.org provided access to servers which returned the route trace only with each node’s IP address, URL, and communication delay whereas IP2Location provided the same information along with the associated city, state, and country of the node which was invaluable with our intent to target every country.

After the Traceroute data was collected, the next task was to plot the nodes and links to determine the topology to most effectively construct the final visualization. Graphviz [6] best suited this purpose. Graphviz is an interpreter of the Dot graphical programming language. Dot computer code is used to assign relationships between different parent and children text elements. Graphviz interprets the Dot code to develop a diagram with bubbles and connecting arrows. This graphical web representation of the network in Graphviz extracted the network topology since the backbone nodes would have more connectors than the local destination nodes.

1. Meisel, D, et al. ”Towards A New Internet Routing Architecture: Arguments for Separating Edges from Transit Core.”

   http://conferences.sigcomm.org/hotnets/2008/papers/18.pdf [↑](#footnote-ref-1)
2. http://dev.maxmind.com/geoip/legacy/geolite/ [↑](#footnote-ref-2)