Topology discovery by active probing

Bradley Huffaker, Daniel Plummer, David Moore, and k claffy

Abstract— **As the Internet has grown, so has the challenge of accurate measurement and modeling of its topology. Commonly used but coarse methods of measuring topology, e.g., BGP tables, suffer from several limitations. To pursue more accurate empirically-based topology modeling. CAIDA began its Macroscopic Topology Project in 1998, The project focus is actively measuring topology and round trip time (RTT) information across a wide cross-section of the commodity Internet. In this paper we describe CAIDA's topology measurement architecture and our analysis and visualization tools. We describe differences between IP and AS (BGP-based) granularities of topology modeling, including advantages and limitations of both, as well as how correlation between both types of data can yield more relevant insights. We introduce four new visualization metaphors for handling macroscopic topology data, as well as a tool for aggregating multiple IP addresses into the same physical router. We highlight results of our analyses, in particular relationships between RTT and topology data, and how source and destination selection and geopolitical boundaries affect those relationships.**

Keywords—**skitter AS IP topology RTT active measurement**

I. Introduction

Internet usage is increasing as access to the Net grows, and is critical for engineering, research, and many collaborative activities. It is difficult to imagine how the dynamically changing topological infrastructure of the Internet looks at any particular moment. We currently have limited understand concerning the impact that dynamic changes in traffic, topology, protocols, and business practices have on this new virtual frontier.

In its early years, monitoring Internet topology was a tractable problem. However, after experiencing exponential growth during the 1990's, inferring connectivity from traffic flow has become a daunting task. In 1998 CAIDA began its Macroscopic Topology Project to collect and analyze Internet-wide topology and latency (round trip time (RTT)) data at a representatively large scale.

In the course of this project CAIDA has created several innovative measurement, analysis, and visualization tools. The primary topology measurement tool we use is `skitter`, which collects forward IP path and round trip times (RTTs) from more than one-half million destinations. During our studies we found that we needed to create a router-level map, which requires aggregating IP addresses that belong to interfaces on the same router. As a result we developed the `iffinder` tool. For strategic development of probe destination lists relevant to the DNS system we created `dnsstat`.

The `skitter` tool requires a list of IP destination addresses to probe. We currently have five different IP address destination lists, each tailored to a specific problem. A given list has between a few hundred and more than one-half million destinations. The two primary lists we use focus on (1) covering as much of the global Internet routing system as possible and (2) analyzing performance to clients of the Domain Name Server (DNS) root servers. Currently 18 different source monitors, located in Asia, Europe, and North America, monitor these destination lists. CAIDA stores topology and latency (RTT) data daily for each server.[1] CAIDA has also created a set of tools for analyzing and visualizing the topology data, at various levels of aggregation granularity.

In this paper we also contrast an IP-level graph with an AS-level (Autonomous System) graph of the Internet. To clarify this distinction, it is important to understand that routing in the Internet occurs at two distinct levels, the IP address and AS.[2] Routing among ISPs occurs via the destination-based announcement of 'reachable' address space from one ISP to another. A typical 'core' Internet router has several IP interfaces that connect other routers, which belong to other ISPs (ASes). Within an AS, the IP hop count is typically relevant to intra-AS path selection, but such intra-AS IP hop count is neither known nor communicated across ASes. The Internet can thus be considered first as a collection of ASs and then as a collection of IP hops inside each AS.

While CAIDA's topology measurement tool (`skitter`) collects IP level topology information, we can abstract each IP address into its corresponding ('originating') AS. We will describe how AS graphs created from CAIDA's active probing methodology have several advantages for modeling and analysis of Internet topology, relative to commonly used techniques based on Border Gate Protocol (BGP) tables.

II. Background

One challenge of Internet research is to accurately model the topology and structure of hundreds of thousands of interconnected networks and machines. Such modeling can provide insight into how resources are used, how traffic flows, and where infrastructural vulnerabilities may lie. There are currently two primary methods for inferring Internet structure: using BGP inter-domain routing tables, and actively probing IP addresses to trace the actual paths that packets traverse from source to destination.

Many studies use the first method, e.g., Border Gateway Protocol (BGP) [1] tables from routers, to infer Internet structure (e.g. [2] [3] [4] [5]). BGP tables have the advantage that they are relatively easy to parse, process and comprehend. BGP data is useful for determining correspondence between IP addresses and network prefixes or ASes, and in analyzing different routing policies in the Internet [6].

However, despite widespread public availability, BGP data suffers from several limitations. BGP connectivity does not capture redundancy of different parts of the network or lateral connectivity among regional networks. It does not reveal public or private exchange points within the infrastructure or short-term

All authors are with: CAIDA, UC San Diego, San Diego Supercomputer Center MC 0505, 9500 Gilman Drive, La Jolla CA 92093-0505, USA. E-mail: {brad,djp,dmoore,kc}@caida.org.

Support for this work is provided by DARPA NGI Contract N66001-98-2-8922, NSF grant NCR-9711092, and Caida members.

[1] The data is available to researchers upon request.

[2] An AS approximately maps to an ISP.

AS path variation and AS load balancing. Most importantly, BGP tables do not reflect how traffic *actually* travels toward a destination network. BGP tables provide only a single perspective from a router toward a destination, which, for several reasons, may not be directly reflected in traversed path data. As a result, we can make only limited inferences about Internet structure and function using BGP table data.

CAIDA built the `skitter` tool to overcome these limitations of existing data sources. In particular, `skitter` paths represent a finer grained and more precise view of topology than can be inferred from BGP tables.

III. Methodology

CAIDA's `topology measurement` tool consists of three main components.

- The `skitter` monitor implements Internet Control Message Protocol (ICMP) parsed traceroutes to collect the forward path from a monitor to a given destination. `skitter` assigns a value to the time to live (TTL) field of every packet. The initial packet to a given destination has a TTL value of 1, and subsequent packets increment the TTL by one for each hop. When routers receive a packet they decrement the TTL field by one and then forward the packet. If the TTL field equals zero at an intermediate router, that router responds to the probe with an ICMP TTL expired message. This process allows `skitter` to capture the addresses of intermediate routers in the path. `Skitter` continues to send packets with incremental TTL values; this process terminates when either the destination is reached or there is a timeout [7].
- The gathered topology data is stored in *arts++* [8] files on the local monitor for authenticated, encrypted daily transfer to a central repository.
- The centralized collector connects to the data server of each monitor box and stores the topology data in a central repository.

In addition to these primary components, several smaller functions enhance `skitter`'s functionality. Our central collector machine runs an *apache* web server that provides destination lists to each topology monitor. Each monitor has a command line client (with an SSL client certificate) that downloads its destination list twice daily. We also have a web server on each monitor that redirects queries to our topology project home page to forestall potential complaints from users of probed machines.

Destinations. We currently have 18 active topology monitors, segregated into four groups according to the destination list they probe. Seven monitors use a list of DNS clients, five monitors use an IPv4 space list, one monitor uses a 'small list' (approximately 1700 destinations), and the remaining five monitors use a list of web servers (approximately 15,000 destinations). To create the DNS clients list (58,000 destination IP addresses) we monitored DNS clients who requested information from DNS root servers and selected one address from each BGP-routable prefix in the RouteViews BGP table. To construct the 'IPv4 address space' list (approximately 661,000 addresses) we collected a large number of addresses from various sources and selected one responding IP within each routable /24 segment, breaking up prefixes larger then a /24. For the web list (approx. 15,000 destinations) we polled a group of webservers. The small list was culled from the web list but reduced in size to increase the sampling rate for each destination.

Cycles. A cycle represents the amount of time a monitor box requires to probe every destination in its list one time. Each topology monitor has a different cycle time, influenced primarily by the size of its destination list but also by the location of the monitor in the infrastructure, the maximum packets transmission rate of the source, and the amount of time the monitor spends probing each target server. The major contributor to variance in cycle duration is the length of the forward path and the number of non-responsive hops. Given the unique forward topology configurations across monitor sources, there may be significant variance among monitor cycle time even for monitors using identical lists.

Storage. IP topology traces from multiple sources generate a large volume of data. Organization of this data is critical for meaningful data analysis. We have developed long-term data storage techniques that allow us to correlate data from different days across multiple monitors. We classify and store individual files by server and day, where day is defined as the 24 hour period from midnight UTC.

Conversion from IP address paths to AS paths. For analyses that involve abstraction to ISP or network, we need to convert observed IP addresses to Autonomous System (AS) numbers. BGP tables contain AS paths that packets should traverse from a given router to their destination IP address (prefix). The AS at the end of an AS path in a core routing table should correspond to the AS administratively responsible for a destination IP address inside the announced prefix. To map IP addresses to ASes we use core BGP tables collected by the University of Oregon's RouteViews Project[**?**], which in conjunction with CAIDA's geographical IP address database allows us to depict several compelling aspects of inter-AS Internet structure.

IV. Advantages and Limitations

Any attempt to measure data from a dynamic system will have limitations and advantages. We outline several limitations of our topology mapping project and explain how we address them.

A. Active data collection

`skitter` uses active probing techniques to infer internal routing structure of the Internet. Because the current Internet is much larger than we can realistically probe, we recognize our inability to capture complete connectivity, particularly with regard to lateral connectivity [9] [10]. We address this limitation in three ways. First, we use many monitors strategically placed around the global Internet. The purpose is to probe from multiple locations and thus reduce dependence on downstream connectivity. Dispersed sources also improve estimates of lateral connectivity among nodes, a major limitation in any single-monitor probing system. Second, our topology project uses much larger destination lists than other studies, carefully screened and optimized to maximize reachability over time. Finally, we have collected over three years of skitter data, providing us with a large database for longitudinal studies.

B. Load Balancing

Topological load balancing presents a unique obstacle to constructing Internet models via active probing. Consecutive probes

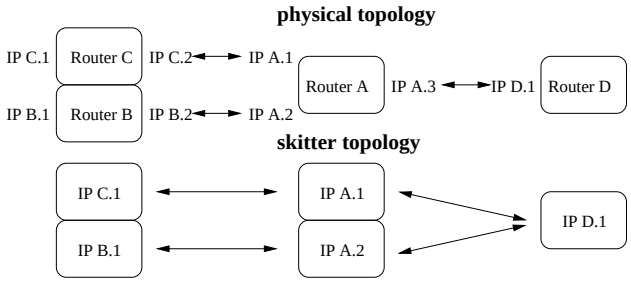

Fig. 1. The difference between the physical and skitter topology is caused by skitter recording only one interface on a link.

from a single monitor to a specific destination often produce paths that diverge along intermediate hops. This divergence may derive from routing changes or configured load balancing. Load balancing causes a periodic intentional 'path instability' in order to force a percentage of traffic onto an alternative link. Without a scheme to recognize load balancing, `skitter` will inaccurately classify load balancing instances as sets of independent paths.

C. *iffinder*

A single monitor will probe (and record) a single interface in each intermediate router along the path to a destination. When a second monitor probes the same destination, it is possible that its probe will be received on a different interface on the same intermediate router from that interface traversed by the first monitor's probes. This situation will introduce a spurious node because `skitter` will classify the different interfaces as separate routers. The resulting falsely created connectivity will affect subsequent analysis of IP graphs, most seriously the inflation of calculation of shortest paths (Figure 1).

To minimize the effect of this type of error we developed a tool called `iffinder` [11]. `iffinder` sends a probe UDP packet to an unused port on a router interface. Many routers will reply to such a packet with an ICMP PORT UNREACHABLE error, with the packet's source address set to that of the interface of the unicast route back to the probing source. Probing one interface and receiving such an error packet from a different interface allows us to infer that the two interfaces belong to the same physical router.

We have demonstrated the ability of `skitter` to generate and produce models of large sets of Internet data. In the next section we describe visualization techniques that allow researchers to analyze these complex data sets.

V. Overview of Visualization Techniques

A. Problem of large topologies

The dynamic nature of the Internet creates a challenge to visualize topological changes rather than static snapshots. We collected a macroscopic set of links during a time window of several days and assumed that all the links were valid during that window. Selection of window size involved tradeoffs. Larger windows allow collection of many more links because existing routes change and new paths increase the probability of observing new links. The disadvantage is that some of these links will

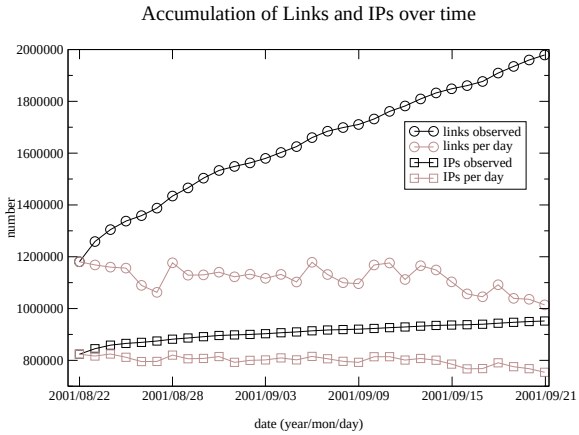

Fig. 2. Daily and cumulative counts of unique links and IPs addresses observed

be invalid by the end of the window interval if the route change was the result of transition to a new link. Figure 2 shows the different accumulation rates for links, routers, and all IP addresses observed in a month-long measurement window. While the cumulative number of observed links increases linearly over the observed time period, the number of links per day drops by about 20%. The total number of IP addresses observed also grows, albeit much more slowly than the link count, while the daily total of IP addresses seen remains fairly constant.

A future aim of CAIDA's topology project is to determine how much topology of a given AS (in particular large tier 1 ASes) we can capture on a given day. We will compare known physical topology of an AS with the `skitter`-observed structure, and vary the observation time window to determine the extent of false connectivity we mistakenly infer (i.e., connectivity loss that derives from natural routing dynamics). We hypothesize that the majority of false links occur outside of the domain of the major providers, whose networks appear to be more stable.

Given the size of the network and the large number of independent variables that can be assigned to each component, it is impossible to create a unified visualization that captures all pertinent topology information. Rather than creating a single tool, we have developed four different visualization techniques, which we describe in the next section, each of which emphasizes different aspects of Internet topology.

B. AS Core Graph

Our first visualization implements the technique discussed in Section III to convert IP addresses to AS numbers and display peering relationships among these ASes. Because the AS graph exhibits highly meshed connectivity within its core, all central nodes are largely interconnected. In contrast, the majority of *leaf ASes* connect to relatively few ASes and in many cases, only to one. As a result, depicting leaf ASes may clutter a visualization, obscuring the connectivity density of central nodes rather than elucidating relationships.

To reduce the visual complexity imposed by leaf ASes, we position nodes with lower connectivity at the edge of the image,

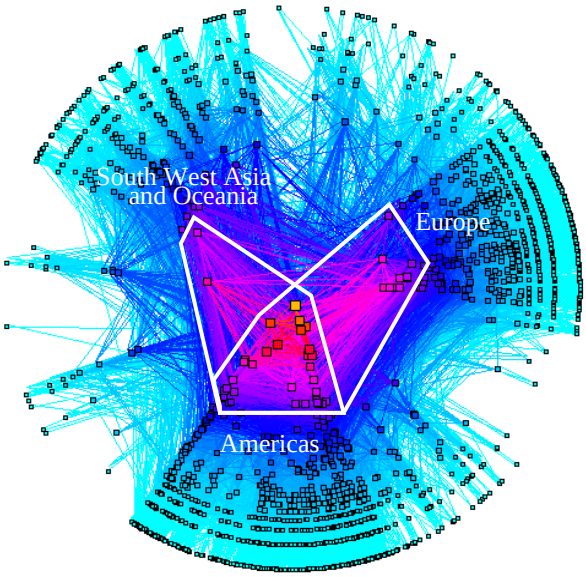

Fig. 3. AS Core Graph: Areas in white represent the derived core Internet graph. Interestingly, most larger provider links in Asia and Europe are within their own continent or within the Americas; few direct links go between Asia and Europe.

by setting the radius of the node equal to the outdegree (Equation 1).

$$r = 1 - log\Big(\frac{outdegree(AS)+1}{maximumOutdegree+1}\Big) \quad (1)$$

This method of node differentiation also reduces server-based bias because we use few servers to collect paths to many destinations. Paths from a single or small number of sources to a larger number of destinations resemble a tree. The largest fanout occurs at points that provide transit to many different locations and are typically located toward the center of the infrastructure. Outdegree values capture this phenomenon better than indegree values.

We also rank a link's importance with a *weight* metric, which is set to the smaller of the two weights of the nodes at each end of the link. Those links of lower importance are plotted first, so that more important links will overwrite them and increase in prominence. To depict geographical relationships among ASes we use the geographic longitude value of the AS headquarters calculated by our *NetGeo*[12] tool to compute the node's angular position (Equation 2).

$$\theta = (longitudeofheadquarters) \quad (2)$$

Figure 3 shows the AS graph seen from 15 CAIDA topology monitors during the first week of August 2001.

C. Dispersion Graph

Although the AS Core Graph provides a useful macroscopic Internet visualization, it obscures connectivity of any individual server. A different technique, the *dispersion graph*, allows visualization of the *AS dispersion* of paths observed from a `skitter` source. Each path contains IP addresses of intermediate nodes between the source and the destination.

Figure 4 is an example of an *AS dispersion* graph from our San Diego `skitter` monitor. This graph reflects complete traces to 21,574 different destinations observed during a 24-hour period on April 27th, 2000. The x-axis represents the IP hop number along the path. The gray scale [3] and numeric label in the vertical bars at each hop identify the AS responsible for the IP address at this hop. The height of the bar represents the proportion of paths that passed through a particular AS at a given hop. Areas are gray when the set of paths disperse into too many distinct ASes to delineate clearly in the plot. We sort the data from the bottom by proportion of paths traveling through each AS. Black bars indicate paths that have terminated in fewer than 24 IP hops.

D. Hyperbolic Space

`Walrus` is a visualization tool that can display large graphs (relevant to characterize a large IP) in 3D hyperbolic space, based upon techniques developed by Tamara Munzner [13]. Similar to the AS Core graph technique, `walrus` is designed to capture IP topology. It can be used to visualize tree-like graphs that have a meaningful spanning tree with a relatively small number of non-tree links. The hyperbolic layout technique overcomes traditional computational difficulties of visualizing large graphs in two ways. First, the computational cost of layout includes only the spanning tree in the calculations since tree layout techniques scale better than those for general graph layout. Second, the problem of displaying a large graph on a small screen scales well with hyperbolic geometry, which provides a focus-and-context view that resembles a continuous fish-eye distortion in three dimensions. This approach allows the user to examine fine details of a small area while maintaining a view of the whole graph as a frame of reference. The user can examine arbitrary areas of the entire graph by interactively moving the focus. `Walrus` does have the limitation that it requires the graph to have a spanning tree, which in turn requires artificial imposition of a tree for most large scale Internet graphs. This artificial spanning tree can distort intuitive expectation of node placement.

E. Bidirectional paths

In some cases it is convenient to display only a subset of paths. This technique allows one to focus on a specific set of paths between a given source and destination (e.g., load-balanced, or flapping). CAIDA developed a technique to depict a set of bidirectional paths from one source to a small set of destinations (30 or fewer). The technique uses horizontal space to depict the AS responsible for routing to a given IP address, and vertical space to depict the number of IP hops from the source. Figure 6 shows the set of paths from our San Diego topology monitor to all other nodes in the destination list XXX for 1-3 January 2001.

VI. Overview results from `skitter` analyses

We present results from three individual CAIDA studies that use `skitter` and associated tools. Each study demonstrates

[3] Color avialable on the CAIDA site

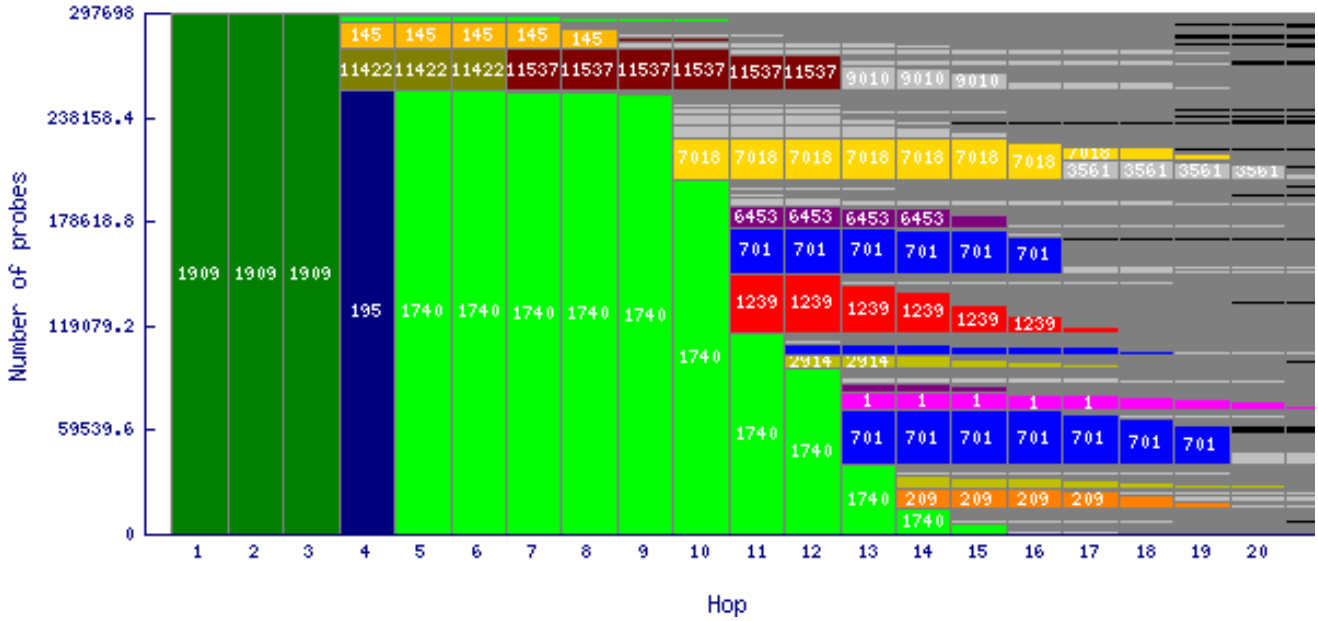

Fig. 4. AS dispersion plot for CAIDA's San Diego, US *skitter* topology monitor, 24 hours, 27 April 2000.

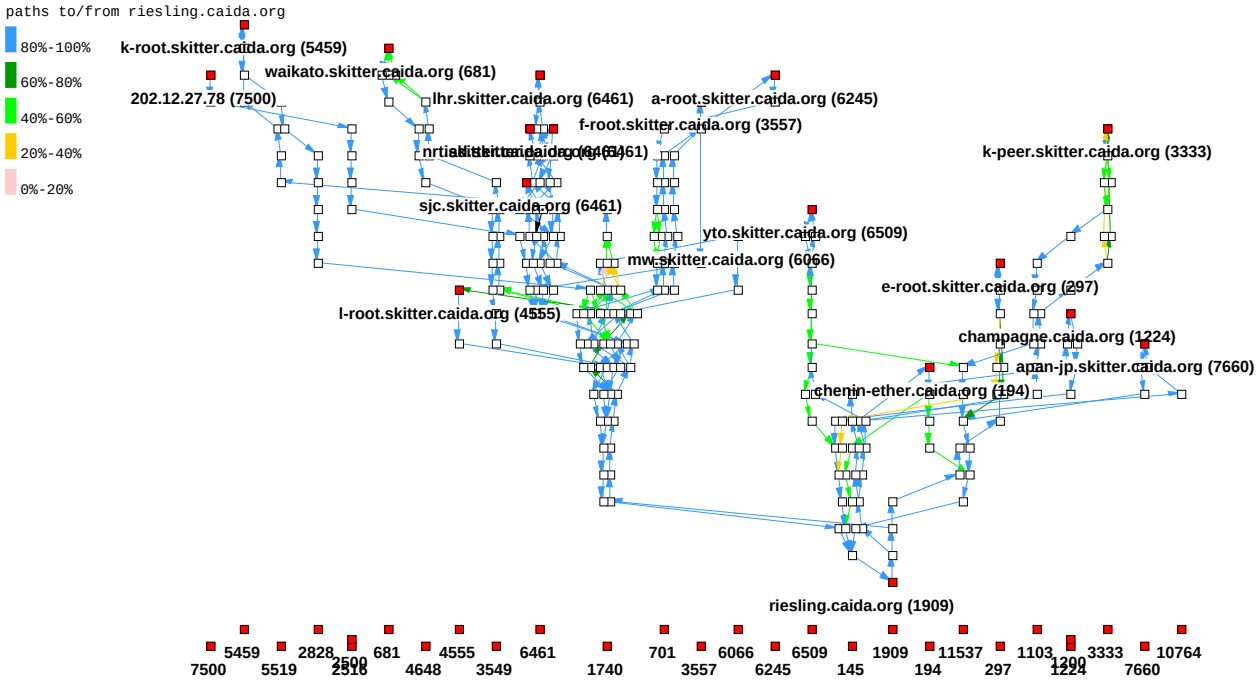

Fig. 6. Multipath visualization technique: bidirectional IP paths from CAIDA's San Diego topology monitor to all other CAIDA monitors for 1-3 January 2001.

the ability of active probing to reveal Internet topology for a unique, specific problem.

A. Transit providers for international links in the Pacific Rim

Goal: To study which companies and countries provide international intra-Pacific Internet transit.

Data: We used the members of the Asia Pacific Economic Cooperation forum (APEC) as our list of target countries. The list includes 14 countries in Asia, three in North America, two in South America, and two in Oceania. For those countries large enough to span across a continent (US, Canada, Russia), we limited the destinations to locations that we could determine were on or near the Pacific coast. While there are multiple Pacific coast destinations in the USA and Canada, Russia had none outside of Europe and was removed from this study.

Method: The first two weeks in November 1999 we measured the percentage of times a server was observed to provide transit for a given path. We classified an AS (or country) as providing transit if it was neither the source nor the destination AS (or country) but appeared at least once in the path.

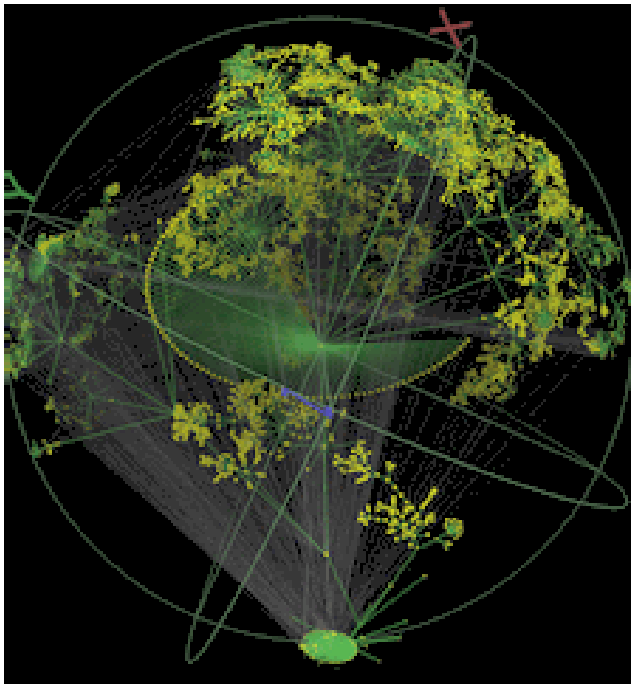

Fig. 5. Hyperbolic layout visualization: CAIDA's San Diego, CA topology monitor; link tree

Results: Four major ISPs appear in 52% of all traces. Of these four, only one is not registered in the US (TELEGLOBE, ASN 6453, registered in Canada). Table I shows that US providers were used for transit for over 71.5% of all measured international paths. This data does not imply that no backup paths existed, but that preferred paths passed almost exclusively through the US.[14]

B. Geopolitical classification of large RTT DNS clients

Goal: To correlate DNS latency performance with geographical location of DNS clients, for the DNS root system.

Method: For each probe cycle we classify the latency (RTT) to a destination as large if it is greater than the 90th percentile of the overall RTT distribution for this cycle. Typically, large RTTs have values greater than 500ms and occasionally as great as 1000ms. We consider RTT distributions independently across cycles, because of significant diurnal variations in the data (networks are more congested during business hours, less so at night).

Results: Figure 7 compares geographical distribution of IP addresses in the target list with geographical distribution of those IP addresses with large RTTs seen in December and March. This data shows that IPs from Asia, South America and Africa appear disproportionately relative to their representation in the target list.

C. Comparison of multiple distance metrics

Goal: To compare different measurements of distance between source and destination in terms of their utility in server selection.

Method: We used *skitter* to gather forward IP hops and RTT to selected destinations. *skitter* is similar to `ping` and

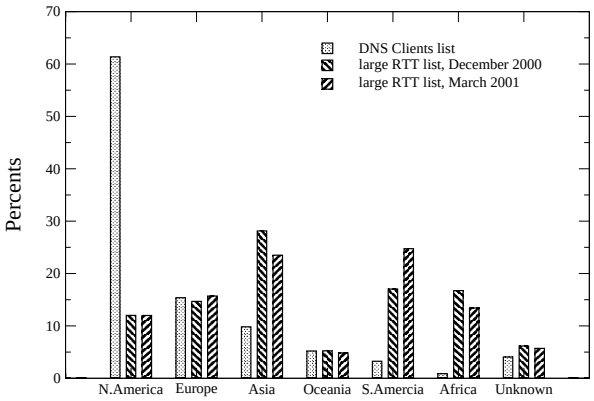

Fig. 7. Geographical distribution of IP addresses in the entire DNS client list versus those with statistically large RTTs in December and March.

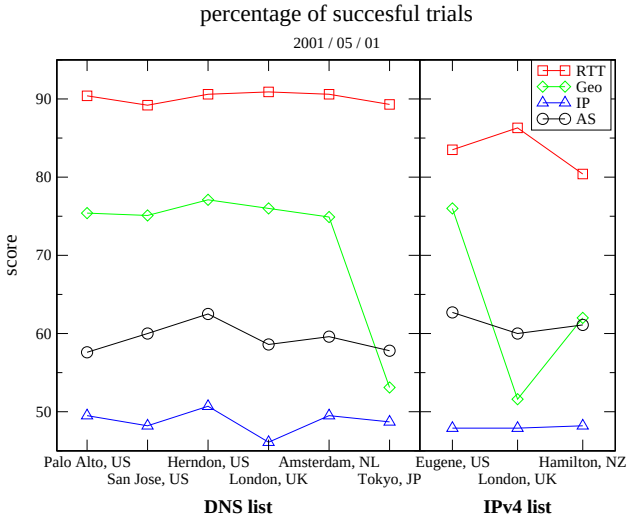

Fig. 8. Success rates for the different distance/performance metrics across monitors. Data collected April 5th, 2001

`traceroute` but uses more accurate kernel timestamps. Success is the percentage of trails in which that metric successfully selected the destination with the lower RTT.

Results: Figure 8 presents results stratified by metric, plotted as a function of (i.e., x-axis) monitor location and destination list. *RTT* refers to the median RTT for a destination for the previous day, *Geo* uses the great circle (circumference of globe) distance between the monitor and the destination, *IP* is the IP path length and *AS* is the AS path length.

The percentage of successful trials varied widely among metrics. For each metric (with the exception of the geographic distance) variance among monitors was small. The RTT metric resulted in a lower score for the IPv4 list than the DNS list, likely because the IPv4 list is polled only once daily. As a result, we under-sample IPv4 destinations, rendering it less likely that the median will accurately reflect typical behavior.

TABLE I
TRANSIT COUNTRY MATRIX

	all	AU	CA	C_H	JP	KR	MX	NZ	SEA	SWA	TW	US
US	71.5	77.8	82.0	90.3	49.5	61.6	100.0	79.6	63.0	97.8	83.5	
CA	13.3	8.3		4.9	37.5	2.1			27.5	22.3	1.3	0.2
AU	2.8			18.4				46.1	1.6			0.4
JP	1.2		1.4	7.4		10.5			12.0			0.3
NZ	0.9	3.7										
EUR	0.7			2.1		1.7			4.2	27.0		
UK	0.7	0.0	0.0		0.1			0.0	5.8	21.1		0.2
SEA	0.3	0.7		5.6								
AR	0.1									5.2		
AE	0.1								1.9			
CH	0.1									2.8		
MM	0.1								1.6			

- EUR : European countries, except for the United Kingdom
- SEA (South-East Asia): Brunei, Indonesia, Papua New Guinea, Philippines, Singapore, Thailand and Viet Nam
- C_H: China and Hong Kong
- SWA (South-West-America): Chile and Peru

An empty space means we had no traces of that category; 0.0 means that the value is less than 0.1%.

VII. CONCLUSIONS

Since 1998, CAIDA has used its IP topology probing tool (`skitter`) to infer properties of macroscopic Internet topology and performance. These active probing techniques hold several advantages over other topology inference techniques. By deploying 18 source monitors worldwide, many probing greater than one-half million destination addresses, CAIDA has been able to gather data that allows for modeling Internet infrastructural characteristics that have thus far been only examined tangentially if at all.

CAIDA's probing infrastructure provides a richer model of Internet topology than one based on BGP tables. In particular, active probing techniques using multiple monitors capture much more lateral connectivity than BGP tables. CAIDA has used this data to correlate Internet structure to geographical location as well as to compare different metrics that measure performance [15].

We have integrated several tools with `skitter`, some of which integrate a wide variety of information into our existing data, i.e. using geographical data to map connectivity. Other tools assist with resolving ambiguities in the data, e.g, use of `iffinder` to aggregate IP addresses into router nodes.

Many difficulties in understanding changing Internet topology rely on the integrity of data collection and large data sets without mechanisms to filter and aggregate. Network data inherently lends itself to graph-based visualization, and CAIDA has developed a suite of tools in pursuit of greater insight from these data sets. Each tool has been designed to allow researchers to focus on specific aspects of the data set, such as components that constitute the network core in an AS graph.

We have presented synopses of results from several CAIDA macroscopic topology studies. Our results demonstrate that our measurement approach is sufficiently flexible to support a wide variety of analyses. This flexibility is the result of engineered software integration of other data modules (e.g. conversion of IP addresses to ASes through BGP tables) with which we can generate geographical mappings, compare metrics, and perform other analyses beyond the scope of traditional (i.e., BGP-based) topology analysis techniques.

VIII. FUTURE WORK

We have four immediate research goals that rely primarily on IP-level topology data.

Coverage comparison. As mentioned in Section V, we are trying to use `skitter` data to compare a known backbone topology of a given AS with topology inferred by CAIDA's active probes. In particular, we hope to quantify the accumulation of false links over time, and in general the ability to capture precise topology of a given AS via remote but strategically designed active probing.

Load balancing. CAIDA will use `skitter` data to identify and study load balancing in the Internet. First, for links that are observed as unstable over time, we will determine which instabilities are due to load balancing versus true instabilities (i.e. routers removed from the network). Second, we will identify and taxonomize different types of load balancing. Finally, we will attempt to build techniques to derive physical network topology from that observed by IP topology probes. In particular, we will develop graph-theoretically-based algorithms to remove false links.

Path length. We will quantify measures of path lengths. In particular, when `skitter` returns a complete path to a given destination, is that path the shortest possible in the IP address graph? If not, what is the distance of the shortest path?

Monitors We plan to continue to increase our global Internet topology coverage by placing additional topology monitors worldwide. We recognize that it is critical to continue the careful, strategic selection of both source and destination addresses to maximize the marginal utility of adding either (source or destination) to CAIDA's measurement infrastructure. The selection of monitor location has received little research attention, and we recognize its importance. Our goal for the next 12 months is

to target increased source and destination in underrepresented regions (e.g. Asia, Africa and South America).

REFERENCES

[1] K. Lougheed and Y. Rekhter., "RFC 1106: Border Gateway Protocol (BGP)," June 1990.

[2] D. Meyer, "University of oregon route views project," http://www.antc.uoregon.edu/route-views/.

[3] "NLANR routing tables," http://moat.nlanr.net/Routing/rawdata.

[4] Sean McCreary and Bill Woodcock, "PCH RouteViews archive," http://www.pch.net/documents/data/routing-tables.

[5] "Routing Information Service," http://www.ripe.net/ris/ris-index.html.

[6] A. Broido and kc claffy, "Analysis of route views bgp data: policy atoms," in *Proceedings of Network-Related Data Management workshop*, May 2001, p. 18.

[7] R. Braden, "RFC 1122: Requirements for Internet Hosts – Communication Layers," Oct. 1989.

[8] "Arts++," http://www.caida.org/tools/utilities/arts/.

[9] H. Burch and B. Cheswick, "Mapping the internet," in *Proceedings of IEEE Computer*, 1999, pp. 32–36.

[10] H.Tangmunarunkit. Heuristics for Internet map discovery R.Govindan, "In proceedings of ieee infocom," Tel Aviv, Israel, March 2000.

[11] Ken Keys, "iffinder," http://www.caida.org/tools/measurement/iffinder/.

[12] D. Moore, R. Periakaruppan, J. Donohoe, and k. Claffy, "Where in the world is netgeo.caida.org," http://www.caida.org/outreach/papers/inet_netgeo/.

[13] Tamara Munzner, "H3: Laying out large directed graphs in 3D hyperbolic space," in *Proceeding of the 1997 IEEE Symposiu on Information Visualization*, 1997, pp. 2–10.

[14] Bradley Huffaker, Marina Fomenkov, David Moore, Evi Nemeth, and k claffy, "Measurements of the Internet topology in the Asia-Pacific Region," 2000, http://www.caida.org/outreach/papers/asia_paper/.

[15] A. Broido and kc claffy, "Internet topology: connectivity of ip graphs," in *Proceedings of SPIE International symposium on Convergence of IT and Communication*, 2001.