The Flattening Internet Topology: Natural Evolution, Unsightly Barnacles or Contrived Collapse?

Phillipa Gill[1], Martin Arlitt[1,2], Zongpeng Li[1], and Anirban Mahanti[3]

[1] University of Calgary, Calgary, AB, Canada
[2] HP Labs, Palo Alto, CA, USA
[3] IIT Delhi, Delhi, India

Abstract. In this paper we collect and analyze `traceroute` measurements[1] to show that large content providers (e.g., Google, Microsoft, Yahoo!) are deploying their own wide-area networks, bringing their networks closer to users, and bypassing Tier-1 ISPs on many paths. This trend, should it continue and be adopted by more content providers, could flatten the Internet topology, and may result in numerous other consequences to users, Internet Service Providers (ISPs), content providers, and network researchers.

1 Introduction

Since its creation in 1969, the Internet has undergone several significant changes. From its beginnings as a research network, the Internet evolved into a commercial network by the mid-1990's [5]. The emergence of "killer applications" such as the World-Wide Web and Peer-to-Peer file sharing vastly expanded the Internet user base [11]. For a variety of reasons, including the commercialization and increased popularity of the Internet, it has become extremely difficult to make ubiquitous changes to the Internet infrastructure. This has led to the emergence of *architectural barnacles* [15], or ad hoc work-arounds for a variety of architectural problems. Architectural purists argue that barnacles may provide short-term relief to such problems, but over the long-term only exacerbate the underlying issues [15].

In this paper we examine a new trend that is emerging at the infrastructure-level of the Internet: large content providers are assembling their own wide-area networks. This trend, should it become common practice, could result in significant changes to the structure of the Internet as it exists today, and have numerous ramifications for users, ISPs, content providers, and network researchers.

We find that companies such as Google, Yahoo!, and Microsoft, are deploying large WANs. Google is leading the way, with a WAN infrastructure that covers much of the U.S., and extends to Europe, Asia, and South America. Yahoo! and Microsoft also have WANs covering the U.S., but do not (yet) extend to

[1] Our data is available at the Internet Traffic Archive - `http://ita.ee.lbl.gov/`.

other regions of the world. These efforts may force other Internet companies to follow suit, in order to remain competitive. For example, MySpace appears to be partnering with Limelight Networks, a Content Delivery Network, to build out a WAN for MySpace.

Our paper makes several contributions. First, we alert the network research community to this emerging trend, as it may affect the assumptions used in other studies. Second, we provide initial measurements on the number and size of the networks already in place for some large content providers. Third, we describe the potential implications of this trend, and discuss whether this is a natural evolution of the Internet architecture, an unsightly barnacle which will ultimately create additional problems, or a contrived attempt to disrupt the balance of power among the providers of the Internet architecture.

2 Background

2.1 Internet Architecture

The Internet architecture has evolved throughout its history. Initially, a single backbone network connected a small number of research networks, to enable researchers to remotely access computing resources at other institutions [5]. In the late 1980's, commercial ISPs began to form, and by 1995 the backbone network was completely transitioned to commercial operation [5]. This transformation resulted in the current three-tiered organization of the Internet infrastructure: backbone networks (Tier-1 ISPs), regional networks (Tier-2 ISPs), and access networks (Tier-3 ISPs) [5, 11]. Consumers and *content providers* access the Internet via Tier-3 ISPs. A Tier-2 ISP connects a number of Tier-3 providers to the Internet. The Tier-2 ISP peers with other Tier-2 ISPs to deliver their customer's traffic to the intended destinations. Tier-2 ISPs may also connect to some Tier-1 ISPs, to more directly reach a larger fraction of the Internet. There are only a few Tier-1 ISPs. Tier-1 ISPs *transit* traffic for their customers (Tier-2 ISPs), for a fee. Tier-1 ISPs peer with all other Tier-1 ISPs (and do not pay transit fees) to form the Internet backbone [11].

2.2 Motivations For Change

There are a number of reasons why content providers may be motivated to build their own wide-area networks, rather than utilize ISPs to deliver content to users. Three broad categories are *business reasons*, *technical challenges*, and *opportunity*. We discuss each in turn.

When the "dot-com bubble" burst (around 2000), many Internet companies, including Tier-1 ISPs such as WorldCom, Genuity, and Global Crossing went bankrupt [13]. This *economic collapse* [13] motivated surviving (and new) Internet companies to increase their focus on "business essentials", such as *risk mitigation* and *cost control*. One risk mitigation strategy content providers may employ is to reduce their dependencies on partners. This could avoid disruptions in a content provider's core business, if, for example, a partner declared bankruptcy. Similarly, topics such as "network neutrality" may create uncertainty for content providers, and hence motivate them to build their own WAN infrastructures, to mitigate any possible or perceived risk. To control costs, a company may look for ways to reduce or eliminate existing costs. One strategy for content providers is to utilize settlement-free *peering* arrangements with ISPs, rather than traditional (pay-for-use) *transit* relationships [14]. For large content providers and small ISPs, peering can be a mutually beneficial arrangement.

Content providers may also be motivated to build their own WANs for technical reasons. For example, a content provider may wish to deploy a new "killer" application, such as video-on-demand. Although many scalable video on-demand delivery techniques exist, none have been widely deployed, owing to the lack of IP multicast on the Internet. This limitation is due to the "Internet Impasse" [15]; this predicament makes it nearly impossible to adopt ubiquitous architectural changes to the Internet that might improve security, enable quality-of-service or IP multicast [16]. A private WAN could avoid this impasse, and give content providers more control over their end-to-end application performance.

Some companies, such as Google, Yahoo!, and Microsoft, aim to provide "Software as a Service" (SaaS), which will deliver functionality via the Web that was previously only available through software installed on the user's computer. In response to the shift to SaaS, several companies are making multi-billion dollar investments in infrastructure such as large data centers [6, 12] and WANs. The motivations for these investments likely span both the business and technical categories described above.

Lastly, content providers may be motivated to build their own WANs because of opportunities that arise. For example, due to the bursting of the "dot-com bubble", a content provider may be able to inexpensively obtain WAN infrastructure (e.g., installed fiber optic network links) from bankrupt ISPs.

3 Methodology

3.1 Data Collection

Our measurement of the popular content provider networks utilizes the `traceroute` tool. `traceroute` is a tool that is commonly used to identify network topology.

To determine the extent of content provider networks, we decided on the following data collection methodology. First, identify a set of N popular content providers. For each of these content providers, select an end-point (i.e., a server). Next, select a set of M geographically-distributed nodes to issue `traceroute` queries, to gather topology information. Lastly, issue $N \times M$ `traceroute` queries. It is important to note that in this study we are only interested in identifying the end points of content provider networks; we are not trying to measure the end user experience, as this would require a different methodology (since end user requests are typically redirected to nearby servers).

For this study, we collected a single snapshot of the networks of the 20 top content providers, as ranked by Alexa [1], by querying from 50 different traceroute servers. The 20 top content providers we used are listed in Table 1. We believe this snapshot is sufficient for an initial view of these networks.

Table 1. Top 20 Content Providers, as Identified by Alexa.com

1	www.yahoo.com	6	www.myspace.com	11	www.hi5.com	16	www.friendster.com
2	www.msn.com	7	www.orkut.com	12	www.qq.com	17	www.yahoo.co.jp
3	www.google.com	8	www.baidu.com	13	www.rapidshare.com	18	www.microsoft.com
4	www.youtube.com	9	www.wikipedia.org	14	www.blogger.com	19	www.sina.com.cn
5	www.live.com	10	www.facebook.com	15	www.megaupload.com	20	www.fotolog.net

We resolve the hostnames of the popular sites only once, and only at a single location (the University of Calgary). We believe this approach will prevent our queries from being redirected to local instances of servers. Since our goal is to understand the size of content provider networks, and not to measure the end-user performance, we argue that our approach is reasonable.

Although we only selected 50 nodes to issue queries from, we selected the locations of these nodes such that they are (potentially) biased in two ways: towards the country in which the content provider is based; and towards areas with higher concentrations of Internet users. We argue this is reasonable as we expect content providers will expand their networks to areas with the largest numbers of (potential) users first. At the time of our study (September 2007), 15 out of 20 of the top global sites listed by Alexa were U.S. based. As a result, we selected 20 traceroute servers in the U.S. These servers were located in 20 different states, including the 10 most populous states. 18 of the U.S. based traceroute servers are at commercial sites, and the other two are at universities. The remaining 30 traceroute servers were selected from countries around the world. Although we intended to use the 30 countries with the most Internet users, some of these countries do not have public traceroute servers. Instead, we issued queries from two locations in Canada (a workstation at our university, and a public traceroute server at another) and from 28 additional locations from around the world, in countries which had working public traceroute servers listed on traceroute.org. Overall, the 30 countries (including the U.S.) we selected were among the top 40 countries in terms of most Internet users, according to Internet World Stats [10]. The 30 countries we used account for an estimated 82.7% of all Internet users.

To keep the load on the 20 selected servers low, we issued only a single traceroute query from each server to each destination, and only one query at a time. Furthermore, we throttled the rate at which the queries were issued (this is in addition to throttling done by some of the traceroute servers). Our data collection occurred between September 27 and October 1, 2007. In future work, we plan to collect data periodically, to understand rate of expansion of content provider networks.

3.2 Data Analysis

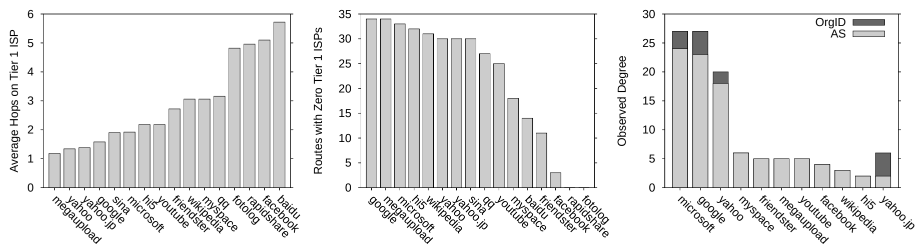
In order to analyze the `traceroute` data, several challenges had to be overcome. First, automating the parsing of the data was problematic. Among the 50 different `traceroute` servers there were 10 different output formats. Thus, a parser was needed that could handle all of these. Second, the `traceroute` output only contained a portion of the data of interest. This meant it was necessary to find additional sources of data (e.g., IP address to organization mappings, organization to Autonomous System (AS) number mappings, etc.) Lastly, there were no obvious metrics for quantifying the size of the WAN of each content provider; this meant a lot of manual inspection of the data was needed in order to determine what the (automated) analysis should evaluate.

We overcame the first two challenges by developing a program to parse the outputs of the various traceroute servers. This program extracts the sequence of IP addresses for each of the traceroute queries. Once the sequence of IPs for a traceroute query is extracted, additional data about each of the IPs is gathered. First, the identity of the organization that registered the IP address is queried from the regional Internet registries. Second, the AS number for the IP address is resolved using an AS number lookup tool [21]. Gathering this extra information increased the potential analyses that we could perform on the data. Specifically, we were able to identify which of the hops in the traceroute path belonged to Tier-1 ISPs using a list of the nine Tier-1 ISPs and their AS numbers [23].

We selected four metrics to facilitate the comparison of the content provider networks, and to examine whether the Internet topology is flattening. We use the *average number of hops on Tier-1 networks* as a measure of how involved such ISPs are in the path. A related metric is the *number of paths that involve no Tier-1 ISPs*. Our third metric, which we call *degree*, provides a conservative estimate of the number of different ISPs a content provider is connected to. This examines the AS number for the router that immediately precedes the first router belonging to a content provider, on each distinct path. Lastly, we consider the *number of geographic locations* in which a content provider's routers appear to be located. We acknowledge that all of these metrics have their shortcomings. For example, it may not be meaningful to compare hop counts when examining differences in the paths. Hu and Steenkiste [9] describe similar issues for identifying metrics for comparing the similarity of end-to-end Internet routes. However, we believe our metrics nevertheless provide some interesting insights. For example, with the traditional Internet model we might expect popular content providers to peer exclusively with a number of Tier-1 ISPs, to ensure global coverage with a minimal number of exchanges on each end-to-end path. If, however, the Internet is flattening, we might expect to see more extensive peering with lower tier ISPs.

4 Results

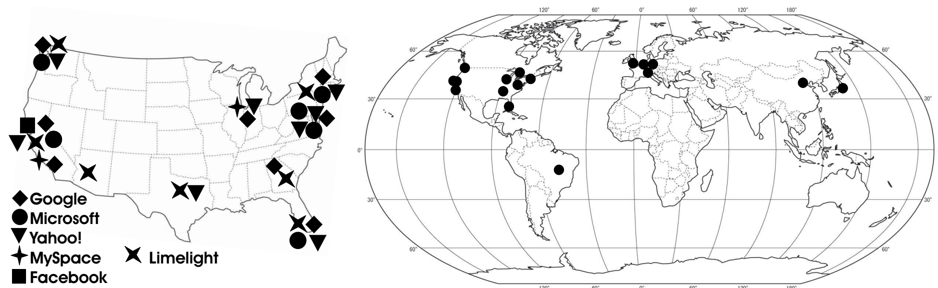
In our analysis we observe that some companies own multiple top 20 sites. Specifically, we observe that Orkut and Blogger are both owned by Google, and traffic for these sites is carried on Google's network. We observe a similar trend for the sites owned by Microsoft, namely MSN and Live. Paths for all four of these subsidiary sites is carried on the same network as their parent companies, and thus the results are very similar. As a result, we only consider one site for each company when the traffic is carried on the same network. Therefore, for our results we omit Orkut, Blogger, MSN and Live, and only show the results for Google and Microsoft, the parent companies. Although Google has recently acquired YouTube, traffic for YouTube has not yet (completely) migrated to Google's network. Thus for our study, we consider YouTube separately from Google. Also, Yahoo! Japan has a unique AS number, so we consider it separately from Yahoo!.



(a) (b) (c)

Fig. 1. Comparison of Network Paths to Top Sites: (a) Average Tier 1 Hops per Path; (b) Number of Paths with No Tier 1 Hops; (c) Connectedness of Each Site

Figure 1 shows the results for three of our metrics. Figure 1(a) shows the average number of hops on a Tier-1 network, for each of the sites. The most notable observation is that our `traceroute` probes traversed significantly more Tier-1 hops on average for some sites than for others. The more established "big three" content providers (Microsoft, Yahoo!, Google) were among those with the lowest averages. Figure 1(b) shows the number of (`traceroute`) paths to each site that contained no Tier-1 hops. For some content providers, including the "big three", 60% (30 paths out of 50) or more contained no Tier-1 hops. Figure 1(c) examines the degree of *connectedness* for each of the content providers that have their own AS number. This graph reveals a clear distinction between the "big three" and the other content providers. Our `traceroute` results show that Microsoft connect to at least 24 different ASes, Google to at least 23, and Yahoo! to at least 18. The next highest is MySpace, at only six. Some paths included IP addresses that we were unable to map to an AS number. For these IP addresses only, we used the organization identifier (OrgID) as retrieved from the corresponding Internet registry. This method enabled us to identify an additional three connection points for Microsoft (27 in total), four for Google (27), and two for Yahoo! (20). The only other content provider affected by this issue was Yahoo! Japan.



(a) (b)

Fig. 2. (a) Location of network end-points in the United States for selected content providers. (b) Our measurement of Google's current WAN.

Figure 2 shows the geographic distribution of entry points into the WANs of selected content providers. Figure 2(a) shows the location of entry points across the U.S. The figure reveals that Microsoft, Google, and Yahoo! all have networks that span the country. The entry points appear to be located (as one would expect) in large centers where carrier hotels or Internet Exchanges exist. Google has the most extensive (live) WAN of any of the content providers we examined. Entry points into Google's WAN are shown in Figure 2(b). Our probes entered the Google network in 10 different North American cities, as well as four European, two Asian, and one South American location.

Other than the "big three", we did not detect any other content providers with large network infrastructures. For example, we only saw Facebook connect to ISPs in the San Francisco Bay area. We did, however, observe several things that suggest others are rolling out WAN infrastructures, in different ways. First, MySpace is peered with ISPs in two separate locations (Los Angeles and Chicago), and appears to partner with Limelight Networks, a Content Delivery Network, to reach other locations. Of 14 probes we sent from European `traceroute` servers to MySpace, eight entered the Limelight network in Europe (in Amsterdam, Frankfurt, or London), which entered Limelight's U.S. network in New York. Six other probes from different locations traversed the Limelight network in the U.S., before reaching MySpace. Second, YouTube (recently acquired by Google) appears to peer with numerous ISPs around the U.S. (We also noticed signs that YouTube's traffic is migrating to Google's infrastructure.)

5 Discussion

In this section we consider the potential ramifications of the identified trends. We discuss these from the perspectives of content providers, users, and ISPs.

If content providers build extensive network infrastructures, they could reap a number of benefits. In particular, they could gain greater control over network-related issues that affect their business. They could deploy applications that have been stymied by the “Internet Impasse”. For example, there are reports that Google will deploy (or has deployed) computation and storage resources at the edge of their network [4]. This could enable Google to provide true video on-demand, and participate in the cable television market. Similarly, they would reduce their reliance on external providers, who might wish to compete against them. There are also many disadvantages. Perhaps the most significant is the cost of deploying, operating and maintaining the infrastructure. Although a few large content providers may have the funds to attempt this, it will be difficult for a large number to follow. In addition, as large content providers move their traffic off the (shared) Internet, small content providers may be faced with larger bills, if ISPs need to recover lost revenues. These issues may lead other content providers to re-examine cost/control tradeoffs; e.g., using VPNs rather than deploying physical networks.

Users could benefit from this trend in several ways. First, these “private” networks may provide better quality of service than the existing Internet, since the content providers could optimize their networks for the applications and services they provide. Second, users may get access to new applications and services much sooner than if they need to wait for a large number of ISPs to agree on a common set of supporting technologies to deploy. Over the long term, however, users could suffer if small content providers are unable to survive, as creativity may be stifled and the variety of content may decrease as a result.

Tier-1 ISPs may notice the greatest changes from this trend. In particular, if this trend becomes widely adopted, Tier-1 ISPs may need to adapt (e.g., vertically integrate, offer content services of their own, or implement new network functionalities that content providers desire, such as IP multicast), or face bankruptcy as revenue dries up. However, since large content providers are unlikely to carry transit traffic, the need for Tier-1 ISPs may not disappear. In fact, a possible (and counter-intuitive) side-effect of large content providers moving their traffic to private networks is lower costs for Tier-1 ISPs, as they may not need to increase the capacity of their networks as often (assuming large content providers are responsible for a significant fraction of the volume of Internet traffic). At the “bottom” of the hierarchy, competing with the “last-mile” ISPs (Tier-3) is unlikely to be attractive to content providers, as the last-mile is expensive to install, and the Return-On-Investment (ROI) relatively low. However, nothing should be assumed; Google recently qualified to bid on wireless spectrum in the United States, which could be interpreted as an initial step in providing last-mile wireless Internet service.

Our data suggests that the Internet topology is becoming flatter, as large content providers are relying less on Tier-1 ISPs, and peering with larger numbers of lower tier ISPs. Content providers are clearly exploring an alternative; only time will determine if this “mutation” becomes the new “norm”, or an “abomination” which will eventually die off. However, this remains a hypothesis, as our results provide only two certain answers: (1) several large content providers are indeed deploying their own networks, and (2) it will be necessary to perform a more rigorous and longitudinal study to determine whether this trend is a short term *barnacle* (e.g., as inexpensive dark fiber disappears, will the trend end?), a slow, but certain evolution of the Internet (e.g., if greater peering between content providers and small ISPs occurs, the Internet topology could flatten), or a contrived collapse (e.g., content providers cunningly defending their territory against ISPs who wish to move into seemingly more profitable content services).

6 Related Work

Our interest in this topic was piqued by an article on a telecom news site [17]. This article stated that Google is building a massive WAN, and speculated that other large "Internet players" are likely doing the same. Thus, we wanted to determine if companies like Google have operational WANs, and if so, how large they are.

We are not aware of any work that has examined this specific trend. However, there are numerous prior works on tools, methodologies, or Internet topology measurements that we leveraged, or could leverage in future work, to answer the questions of interest to us. We describe some of the most relevant works below.

In this study we utilized `traceroute`; however, it has a number of known weaknesses [5]. Tools such as `tcptraceroute` [22] or Paris Traceroute [2] could be used in conjunction with PlanetLab to address these known limitations of `traceroute`. Sherwood and Spring propose additional methods for addressing these weaknesses [18].

The closest work to our own is Rocketfuel, which created router-level ISP topology maps [19]. A key difference is their paper focused on mapping the network topologies for specific ISPs, while we are interested in the network topologies for specific content providers. Spring *et al.* also proposed *scriptroute*, a system to conduct network measurements from remote vantage points [20]. Given the similarity in objectives, we will likely revisit Rocketfuel and *scriptroute* in the future. Additionally, scalability and efficiency of collection will be important for larger and repeated data collection efforts. Donnet *et al.* [8] and Dimitropoulos *et al.* [7] have investigated these issues for topology discovery.

A number of papers have discussed the need to evolve the Internet architecture, and proposed ways in which change could be enabled within the current (static) architecture [3, 15, 16]. In this paper, we examine a change that is occurring in the Internet architecture. Depending on how this change is viewed (e.g., is it a fundamental shift, or just an unsightly barnacle), it may be necessary to revisit the predictions of what the future Internet will look like.

7 Conclusions

In this paper, we utilized an active measurement (`traceroute`-based) approach to demonstrate that large content providers are deploying their own WANs. We show that established companies such as Google, Microsoft, and Yahoo! already have sizable WAN infrastructures, and find that some smaller (but very popular) content providers appear to be following their lead. While there are many possible motivations for this trend, we believe it is more important to consider the potential ramifications. Specifically, it could alter the way in which the Internet operates, either (eventually) eliminating the need for Tier-1 ISPs, or forcing such ISPs to evolve their businesses. Network researchers also need to understand whether this is a long or short term trend, as it will affect the importance of research topics.

Significant work remains to be done on this topic. Increasing the breadth of the study, conducting a longitudinal study, and considering alternative metrics are some of the dimensions of our future work.

Acknowledgements The authors greatly appreciate the providers of the public traceroute servers as well as the feedback of Bala Krishnamurthy, Dejan Milojicic, Jeff Mogul, Carey Williamson and the anonymous reviewers.

References

1. Alexa's Top 500 Sites, http://www.alexa.com/site/ds/top_500.
2. Augustin, B., Cuvellier, X., Orgogozo, B., Viger, F., Latapy, M., Teixeira, R.: "Avoiding Traceroute Anomalies with Paris Traceroute", *Internet Measurement Conference*, Rio de Janeiro, Brazil, 2006.
3. Clark, D., Wroclawski, J., Sollins, K., Braden, R.: "Tussle in Cyberspace: Defining Tomorrow's Internet" *ACM SIGCOMM*, Pittsburgh, PA, 2002.
4. Cringely, R.: "Google-Mart: Sam Walton Taught Google More About How to Dominate the Internet than Microsoft Ever Did", Nov. 17, 2005. http://www.pbs.org/cringely/pulpit/2005/pulpit_20051117000873.html
5. Crovella, M., Krishnamurthy, B.: *Internet Measurement: infrastructure, traffic & applications.* J. Wiley & Sons, Ltd., West Sussex, England, 2006.
6. Data Center Knowledge Web site. http://www.datacenterknowledge.com/
7. Dimitropoulos, X., Krioukov, D., Riley, G.: "Revisiting Internet AS-level Topology Discovery", *Passive and Active Measurement*, Boston, MA, 2005.
8. Donnet, B., Friedman, T., Crovella, M.: "Improved Algorithms for Network Topology Discovery", *Passive and Active Measurement*, Boston, MA, 2005.
9. Hu, N., Steenkiste, P.: "Quantifying Internet End-to-End Route Similarity", *Passive and Active Measurement*, Adelaide, Australia, 2006.
10. Internet World Statistics, http://www.internetworldstats.com/; statistics from June 30, 2007.
11. Kurose, J., Ross, K.: *Computer Networking: A Top Down Approach.* Addison Wesley, Boston MA, 2008.
12. Mehta, S.: "Behold the server farm", *Fortune* magazine, July 26, 2006. http://money.cnn.com/2006/07/26/magazines/fortune/futureoftech_serverfarm_fortune/
13. Norton, W.: "The Evolution of the U.S. Internet Peering Ecosystem", Equinox White Paper, 2003.
14. Norton, W: "A Business Case for ISP Peering", Equinox White Paper, 2001.
15. Peterson, L., Shenker, S., Turner, J.: "Overcoming the Internet Impasse through Virtualization", *IEEE Computer*, April 2005.
16. Ratnasamy, S., Shenker, S., McCanne, S.: "Towards an Evolvable Internet Architecture", *ACM SIGCOMM*, Philadelphia, PA, August 2005.
17. Raynovich, R.: "Google's Own Private Internet". http://www.lightreading.com/document.asp?doc_id=80968
18. Sherwood, R., Spring, N.: "Touring the Internet in a TCP Sidecar", *Internet Measurement Conference*, Rio de Janeiro, Brazil, 2006.
19. Spring, N., Mahajan, R., Wetherall, D.: "Measuring ISP Topologies with Rocketfuel", *ACM SIGCOMM*, Pittsburgh, PA, August 2002.
20. Spring, N., Wetherall, D., Anderson, T.: "Scriptroute: A Public Internet Measurement Facility", *USENIX Symposium on Internet Technologies and Systems (USITS)*, Seattle, WA, March 2003.
21. Team Cymru IP to ASN Lookup page, http://www.cymru.com/BGP/asnlookup.html.
22. Toren, M.: `tcptraceroute`, http://michael.toren.net/code/tcptraceroute/.
23. Wikipedia article, "Tier 1 network", http://en.wikipedia.org/wiki/Tier_1_network.