Characterizing and measuring path diversity of internet topologies

Renata Teixeira, Keith Marzullo, Stefan Savage, Geoffrey M. Voelker

HAL Id: hal-01097527

https://inria.hal.science/hal-01097527

Submitted on 24 Dec 2014

Characterizing and Measuring Path Diversity of Internet Topologies

Renata Teixeira,[*] Keith Marzullo, Stefan Savage, Geoffrey M. Voelker
Computer Science and Engineering, University of California, San Diego
{teixeira,marzullo,savage,voelker}@cs.ucsd.edu

Categories and Subject Descriptors

C.2.1 [**Computer Communications and Networks**]: Network Archtecture and Design—*network topology*

General Terms

Measurement

1. INTRODUCTION

The use of multiple network paths between a pair of hosts has been proposed for a wide variety of network technologies in order to achieve higher bandwidth in data transfers, to select paths with low latency, to balance load in the network, and to protect against failures.

Path diversity describes the number of disjoint paths between hosts and characteristics of those paths, such as path length, latency, and bottleneck bandwidth. Despite wide interest in such techniques, little work has been done that actually characterizes and measures the path diversity that exists in large-scale deployed packet networks like the Internet. Our work characterizes layer-3 path diversity [1] in the Internet.

2. PATH DIVERSITY IN THE INTERNET

The Internet is an interconnection of autonomous systems. An *autonomous system* (AS) is a network under a single administrative entity, such as an Internet Service Provider (ISP), an enterprise, or a campus. Each ISP network is a collection of physical locations called Points of Presence, or PoPs. Customers connect to the ISP network in one of the ISP's PoPs. PoPs are interconnected by inter-PoP links. ASes interconnect at peering points. Routers that interconnect ASes exchange routing announcement updates using the Border Gateway Protocol (BGP), a policy-based protocol.

As a start, we characterize the path diversity between a pair of nodes in a network topology by the number of node- and link-disjoint paths between them. The number of disjoint paths between two hosts in the Internet depends on where and how they connect to the network. Paths between two hosts in the same AS are determined by the AS's internal network infrastructure and by the *intra*-domain routing protocol. Paths between two hosts connected to different ASes are determined not only by the network infrastructure and routing policies inside each network on the path, but also by peering connections and *inter*-domain routing policies used by BGP. Thus, it is important to study path diversity both inside an ISP network and across multiple networks.

In our work, we analyze a set of network topologies (both individual AS and multiple-AS topologies). In this paper we present the results for two topologies: the topology of the Sprint network, which illustrates the path diversity inside an AS; and a topology derived from CAIDA, which illustrates the path diversity available across multiple ASes in the Internet.

2.1 Path Diversity inside an ISP Network

As an example, we study path diversity in the continental US Sprintlink topology. The exact topology was provided by Sprint under a non-disclosure agreement[2]. We study a PoP-level topology that is composed of 17 PoPs. Each node in the topology represents the set of fully-meshed routers in a PoP. Links represent both active and backup IP links between PoPs, which allow us to assess the maximum potential for path diversity.

We measure path diversity for every pair of Sprintlink PoPs. Figure 1 shows the cumulative distribution function of the number of disjoint paths for all pairs of PoPs. The solid line represents the maximum possible number of paths for each pair of PoPs as determined by the minimum between the out-degree of the source PoP and the in-degree of the destination PoP.

Figure 1 shows that the exact Sprintlink topology has a high level of path diversity. Almost 90% of pairs of PoPs have at least four link-disjoint paths between them and 40% of pairs of PoPs are connected through eight or more link-disjoint paths. For protection reasons, when there is more than one link between two PoPs, each of these links connects physically different routers. Hence, link-disjoint paths in this topology are also router disjoint.
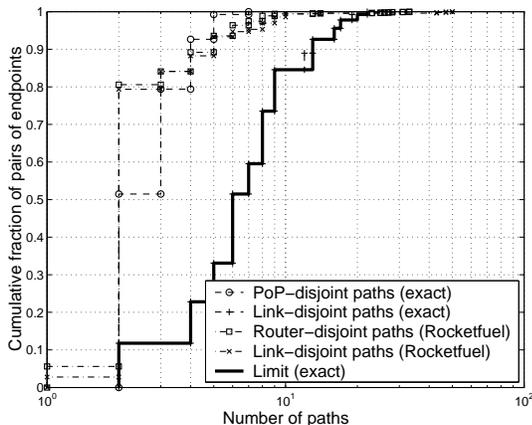
---

[*]Supported by Capes/Brazil

[1]Note that layer-3 path diversity is distinct from physical-level path diversity: several virtual links may share a physical link.

[2]An approximate topology can be found at http://www.sprint.net

Figure 1: Path diversity in the Sprintlink topology.

Figure 2: Cumulative distribution of the number of partially-disjoint paths in the CAIDA topology.

Path diversity inside an ISP network depends on how ISPs engineer their network and thus it varies from one ISP to another. In an attempt to characterize path diversity for a larger set of ISPs, we also studied ISP topologies made publicly available by the Rocketfuel project [2]. These are router-level topologies extracted using end-to-end measurements. Since the topologies are generated by sampling, Rocketfuel infers the set of IP addresses belonging to a single router and only captures active links.

Overall, our results for the Rocketfuel topologies showed that there is significant variance of path diversity between different ISPs, but in general, tier-1 ISPs have at least two paths between an entry and an exit point in their network. Figure 1 also shows the results for the Rocketfuel-derived Sprint topology. Comparing the number of link-disjoint paths in the exact Sprint topology with the number of router-disjoint paths measured by Rocketfuel for the same topology, we see that the path diversity in topologies measured by Rocketfuel is underestimated[3]. Rocketfuel gives us the paths that are simultaneously active between two PoPs, instead of the maximum set of available paths in the complete Sprint topology.

2.2 Path Diversity across the Internet

To analyze path diversity across multiple ASes, we use a topology from CAIDA[1]. CAIDA has been recording active measurements from various monitors scattered around the world since 1998. Paths in the CAIDA topology are composed of active links. In our study, we use a snapshot of a core of the CAIDA Internet topology collected from 16 monitors with over 57,000 nodes and nearly 200,000 edges.

We compute path diversity among the 16 monitors. Since most of the monitors have an out-degree of one, 90% of the pairs of monitors are connected by only one path. The lack of path diversity is not a property unique to CAIDA monitors: most hosts connected to the Internet are single-homed. This lack of diversity led us to investigate alternative definitions of path diversity. We define *partially-disjoint* paths as those for which there are some nodes that appear in more than one path. We identify the routers on the border of

the ASes of the source and the destination, and compute path diversity for those nodes. This value represents the maximum path diversity available between the monitors' ASes.

Figure 2 presents the cumulative distribution of the number of partially-disjoint paths. It also shows the cumulative distribution of fully-disjoint paths and the limit imposed by the degree of the border nodes for comparison. The percentage of pairs of monitors connected by only one path drops to 24% when allowing nodes inside the ASes of the source and destination CAIDA monitors to be shared along the paths.

The CAIDA topology is constrained by routing policies and can only capture links that are active, therefore the path diversity presented in Figure 2 represents a lower bound of the number of disjoint paths for the CAIDA monitors. Path diversity in such a network is a useful approximation to the path diversity that an overlay network, for example, could exploit in practice if deployed on the Internet.

3. CONCLUSION

Our results show that there is potential for high level of path diversity inside an ISP, as illustrated by the Sprint network. The utilization of this path diversity, however, is limited by the routing protocols. Current intra-domain routing protocols use multiple paths in the event of a failure or for balancing load among equal-cost paths. Results from the CAIDA topology indicate that even though there is usually only one path between two end-hosts, there is significantly more path diversity across multiple ASes in the core of the Internet. This implies that in order to achieve higher end-to-end path diversity, clients should be multi-homed.

This work is a first step toward a model of path diversity in the Internet. We are generalizing these results for a larger number of ISP topologies and inter-domain topologies collected from a varied set of monitors.

4. REFERENCES

[1] B. Huffaker, D. Plummer, D. Moore, and k claffy. Topology Discovery by Active Probing. In *Symposium on Applications and the Internet (SAINT)*, Jan. 2002.
[2] N. Spring, R. Mahajan, and D. Wetherall. Measuring ISP Topologies with Rocketfuel. In *Proceedings of ACM SIGCOMM*, Pittsburgh, PA, Aug. 2002.

---

[3]The fact that the number of disjoint paths in the Rocketfuel topology is larger than the maximum for some pair of nodes is a consequence of IP add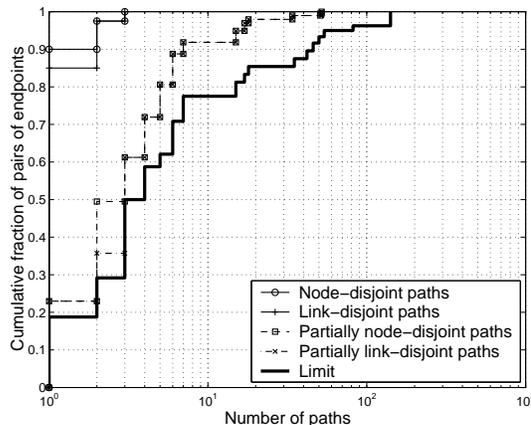resses that were not accurately mapped to the same router.