A First Look Into Distributed VPNs

Yunming Xiao[1], Matteo Varvello[2], Aleksandar Kuzmanovic[1]
[1]Northwestern University, [2]Nokia Bell Labs

Introduction

Ookla[1] reported that average Internet speeds in American homes grew 20x in the last 10 years, from 8 Mbps in 2010 to 180 Mbps in 2021, and a similar growing trend is observable worldwide. Recent studies[2, 3] suggest that the added cost for faster Internet speeds — e.g., $50 monthly to boost from 200 to 300 Mbps with Comcast Xfinity — is not worth to most residential users, which consume only 5% of their available bandwidth (**Figure 1**). Several systems were launched which allow users to monetize such *spare* bandwidth, e.g., distributed proxy or VPN services[4, 5, 6].

While such systems are attracting a considerable number of users, both as clients (buyers) and providers (sellers), little to nothing is known about such *bandwidth marketplaces*. We have identified *distributed VPN* (dVPNs) as the most concrete examples of such marketplaces because, in essence, a dVPN is a system which allows users to sell their spare bandwidth. The contribution of this paper is an investigation of the dVPN ecosystem. We run *active* and *passive* measurements to characterize their footprint and performance.

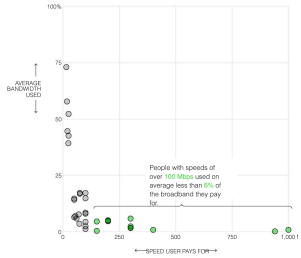

Figure 1. Added cost for faster Internet speed is not worth to most residential users[2].

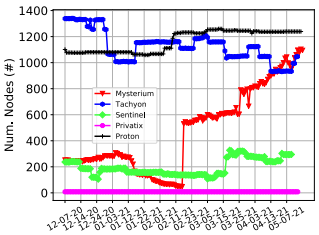

Figure 2. Evolution of the number of nodes.

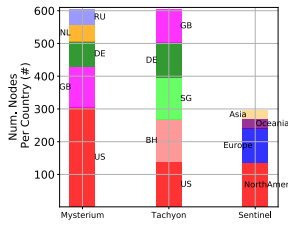

Figure 3. Number of nodes per country.

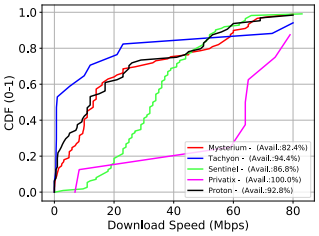

Figure 4. CDF of download speed.

Methods

DVPN Selection. Early dVPN attempts[7, 8] are not useful for our study since their users are not allowed to charge for their bandwidth. More recent dVPNs like Mysterium[4], Sentinel[5], Tachyon[6] are appropriate candidates.

Active experiments. We automate a dVPN client to discover and connect to available nodes. This allows to discover a dVPN footprint, pricing, and performance.

Passive experiments. Passive experiments consist of contributing bandwidth to such dVPNs by running several nodes. This is useful to characterize how much and which traffic a typical dVPN node carries. We have deployed 10 machines running nodes for mentioned dVPNs across 4 countries (US, UK, Italy, and China). Our passive measurements last for 3 months (February to April 2021) and account for ~16 TB of traffic. The IRB at our institution has determined that our work is not considered human research.

Active Measurement

Figure 2 shows that originally, only Tachyon had a footprint comparable with ProtonVPN (a commercial centralized VPN), *i.e.,* ~1000 nodes. However, Tachyon has lost 36% of its nodes over time while Mysterium's node count has been steadily increasing in the last three months, and it is now the largest dVPN with 1,100 nodes.

Figure 3 shows the US is currently the country where most nodes are located that irrespective of the dVPN. Germany (DE) and Great Britan (GB) are two other popular countries among dVPNs.

Figure 4 shows that only Tachyon is overall slower than ProtonVPN; Mysterium has comparable performance with ProtonVPN while both Sentinel and Privatix significantly improve bandwidth, by up to 3x and 6x.

Passive Measurement

Between February and April 2021, our nodes have served ~632 thousand dVPN sessions, ~623 million TCP/UDP flows, accounting for about 16 TB of data. Download traffic is the highest contributor, about 10x the amount of upload traffic.

Figure 5 aggregates data across the three dVPNs since no statistically meaningful difference was observed. The figure shows that the US has the most buyers, followed by Iran (IR) and United Arab Emirates (AE). The US is also the most popular destination regardless of which node (middle of the plot) is used, accounting for over half of the traffic. Russia (RU) is the second most popular destination, followed by China (CN), UK, and Netherlands (NL).

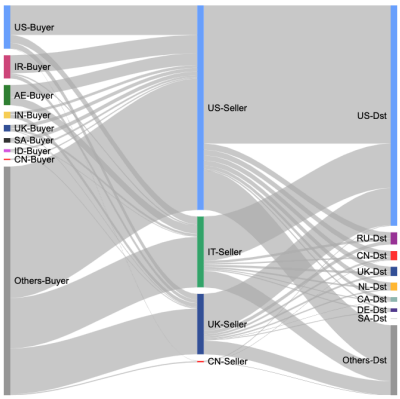

Figure 5. Visualization of dVPN traffic across Mysterium, Sentinel, and Tachyon. Buyer's locations are on the left, our machines where dVPN nodes are run in the center, and traffic destinations on the right.

Future Work

We plan to continue exploring the dVPN ecosystem with focus on the traffic *characteristics*, e.g., presence or lack of harmful traffic, to help asses the *risk* associated with running a DVPN node. We are further interested in understanding the economics between buyers and sellers. We will study the value of spare bandwidth in today's bandwidth marketplace and further explore the opportunities of optimizing the buyers' costs and sellers' income.

Contact

Emails:
yunming.xiao@u.northwestern.edu,
matteo.varvello@nokia.com,
akuzma@northwestern.edu
Website:
https://yunmingxiao.github.io/projectsfolder/ring/

References

1. Speedtest Global Index - United States. https://www.speedtest.net/global- index/united- states.
2. The Truth About Faster Internet: It's Not Worth It, The Wall Street Journal, 2019. https://www.wsj.com/graphics/faster- internet- not-worth-it/
3. F. Bronzino, et al. Inferring streaming video quality from encrypted traffic: Practical models and deployment experience. Proceedings of the ACM on Measurement and Analysis of Computing Systems, 3(3):1–25, 2019.
4. Mysterium network: Censorship free Internet for all. https://mysterium.network/
5. Sentinel: Secure yourself today with Sentinel. https://sentinel.co/
6. Tachyon VPN. https://tachyon.eco/
7. Hola: Get access to worldwide content. https://hola.org/
8. VPN Gate - Public VPN Relay Servers. https://www.vpngate.net/