



**HAL**  
open science

# Technical Perspective: Can We Uncover Private Backbone Infrastructures?

Sara Alouf

► **To cite this version:**

Sara Alouf. Technical Perspective: Can We Uncover Private Backbone Infrastructures?. Communications of the ACM, 2023, 66 (8), pp.94. 10.1145/3604621 . hal-04177294

**HAL Id: hal-04177294**

**<https://inria.hal.science/hal-04177294>**

Submitted on 4 Aug 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



# Technical Perspective

## Can We Uncover Private Backbone Infrastructures?

By Sara Alouf

THE INTERNET—THIS ever-expanding network of computers, information, and people—has gone through many steppingstones before being propelled to its current place as a worldwide social system. One of its major transformations resulted from the rise of distributed datacenters and the mobile revolution. Cloud computing has become a standard practice and many companies no longer invest in their own servers but rather use the services of cloud providers. People mostly connect to the Internet with mobile devices, including smartphones, and media consumption over the Internet has dramatically increased.

Large cloud and content providers like Google, Amazon, Microsoft, Meta (Facebook), Akamai, or Alibaba have become major players in today's Internet. They offer a plethora of services to the end users and are building and operating their own global-scale infrastructures. These *private* networks interconnect their numerous datacenters, located around the world, and the providers use them to privately transport their generated traffic, as far as possible, rather than having it flow over the public Internet.

There has been considerable effort to study the connectivity of the Internet in the past decades. The key point is to estimate the network distance between any two points of the network, and by “network distance” we mean “latency.” The benefits of knowing the network distance range from an informed choice in the selection of servers (for streaming and games, among others) to an optimized construction of overlay networks. There are also economic advantages for the stakeholders; for instance, informed decisions can be taken on the best location for infrastructure investments. To this day, two broad strategies have been developed to estimate the network distance: using direct measurements between nodes over either a pre-existing or a specifically deployed infrastructure


and using coordinate-based techniques together with a set of landmarks.

The increasing importance of private infrastructures within the Internet has motivated the authors of the following paper to look specifically at their network connectivity. Unlike the public Internet, private providers make measurements difficult. Rich measurements like those obtained with traceroute can no longer be taken for granted as traceroute packets can be modified or simply disabled by private cloud or Internet service providers. Salamatian et al. assemble techniques from the two broad strategies developed by the networking community and complement them with new ones using concepts from Riemannian geometry, with the additional specificity that all nodes are geo-located. While the methodology described in the accompanying paper has the potential to be applied to the Internet, the authors are particularly interested in intra-cloud provider networks. The central question addressed then is about the importance and the performance of the logical paths between the availability zones of the considered cloud provider.

I particularly enjoyed two things about this paper: The first thing is there is a combination of advanced mathematical notions and a practical solution. Advanced mathematics is needed to make sense of simple round-trip measurements. To identify which node pairs are most likely connected physically, the authors pinpoint those measurements that are approximately

proportional to the straight path between the endpoints. As Earth is not flat, Euclidean geometry goes out of the picture: straight paths are geodesics in Riemannian geometry. Once a graph of the network is constructed, the importance of an edge is assessed through its Ricci curvature. Edge performance is related to the latency beyond the minimal one over the straight path.

The second thing I liked is the new visualization tool the authors developed for this work. This tool produces a manifold view for a chosen latency threshold, and it fuses Riemannian geometry with graphical maps. This is how the importance of a logical path between two points can readily be observed. A path critical for the connectivity of the network tends to be heavily used; it would appear as a saddle in the manifold view (for those familiar with Riemannian geometry, the Ollivier-Ricci curvature would be negative). Considering a smaller latency threshold may make this path disappear in the corresponding manifold view. Comparing visually manifold views generated for different latency thresholds looks like comparing 3D views of the world map with different sea levels. In each manifold view, well-connected areas emerge as islands, whereas paths along which the latency is larger than the threshold appear flooded. Identifying where fibers should be deployed becomes a piece of cake! “Flooded” paths need larger bandwidth connections to enjoy a smaller latency and surface above sea level.

I hope I convinced you to check the paper and see for yourself how handy this tool can be in revealing features of the private provider infrastructure. As the authors have put their datasets and source code in a public repository, it can only facilitate its adoption by the networking community. Enjoy! 

Sara Alouf is a research scientist at Inria, Sophia Antipolis, France.

Copyright held by author.

**There has been considerable effort to study the connectivity of the Internet in the past decades.**