A Hop Away from Everywhere: A View of the Intercontinental Long-haul Infrastructure

Esteban Carisimo
esteban.carisimo@northwestern.edu
Northwestern University
Evanston, Illinois, USA

Caleb J. Wang
caleb.wang@northwestern.edu
Northwestern University
Evanston, Illinois, USA

Mia Weaver
mweaver6@wisc.edu
University of Wisconsin-Madison
Madison, Wisconsin, USA

Fabián E. Bustamante
fabianb@cs.northwestern.edu
Northwestern University
Evanston, Illinois, USA

Paul Barford
pb@cs.wisc.edu
University of Wisconsin-Madison
Madison, Wisconsin, USA

ABSTRACT

We present a longitudinal study of intercontinental long-haul links (LHL) – links with latencies significantly higher than that of all other links in a traceroute path. Our study is motivated by the recognition of these LHLs as a network-layer manifestation of transoceanic undersea cables. We present a methodology and associated processing system for identifying long-haul links in traceroute measurements, and report on our findings from. We apply this system to a large corpus of traceroute data and report on multiple aspects of long haul connectivity including country-level prevalence, routers as international gateways, preferred long-haul destinations, and the evolution of these characteristics over a 7 year period.

CCS CONCEPTS

• **Networks → Topology analysis and generation**.

KEYWORDS

Long-Haul Links (LHL), submarine cables

ACM Reference Format:
Esteban Carisimo, Caleb J. Wang, Mia Weaver, Fabián E. Bustamante, and Paul Barford. 2024. A Hop Away from Everywhere: A View of the Intercontinental Long-haul Infrastructure. In *Abstracts of the 2024 ACM SIGMETRICS/IFIP PERFORMANCE Joint International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS/PERFORMANCE Abstracts '24), June 10–14, 2024, Venice, Italy.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3652963.3655041

1 INTRODUCTION

As part of ongoing work focused on the criticality of the submarine cable network [10], we set ourselves to map traceroute measurements to submarine cables as a first step towards understanding its potential vulnerabilities. We meant to follow an approach [3] building on the assumption that given a traceroute, one could identify the link with the largest latency and, if the associated routers were found near submarine landings, that this link could be mapped to one or a handful of cables. Alas, a preliminary analysis of traceroute datasets disabused us of this assumption.

While we found traceroutes matching these expectations, we also found many others in which the routers associated with a submarine traversing link were far inland from the closest landing points, some as far as 700 km. Further analysis revealed many of these long-haul links, covering distances larger than 10,000 km and connecting every country worldwide. Figure 1 illustrates an example long-haul link connecting Seattle, US, with Singapore by concatenating several submarine cable segments.

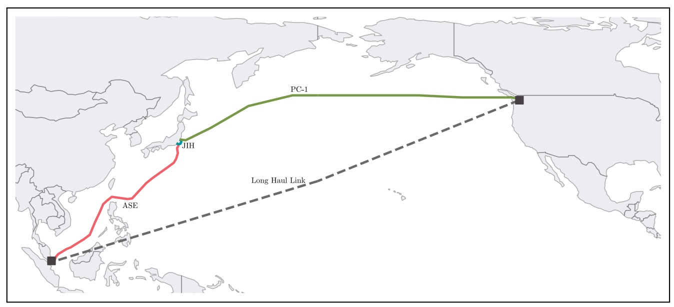

Figure 1: A long-haul link connecting Seattle, US with Singapore, via the a concatenation of several submarine cable segments: PC-1, JIH and ASE.

This paper presents the first in-depth, longitudinal study of intercontinental long-haul links and their preferred destinations as the network-layer manifestation of critical transoceanic undersea cables.

We make the following key contributions:

- We present a methodology and associated processing system for identifying intercontinental LHLs in large traceroute datasets.
- We apply this methodology to a large corpus of traceroute data collected by the CAIDA Archipelago platform [5] and report on multiple aspects of this LHL network, including the number, length, and popular destinations of LHLs.
- We report on the evolution of LHL properties over a multi-year period starting in 2016.

SIGMETRICS/PERFORMANCE Abstracts '24, June 10–14, 2024, Venice, Italy

ACM ISBN 979-8-4007-0624-0/24/06
https://doi.org/10.1145/3652963.3655041

- We release source code and artifacts to facilitate the reproducibility of our study[1].

This study contributes to the community effort to create consistent maps across layers of the Internet, critical to a range of important analyses from performance and robustness to security [1, 6, 7].

2 METHOLODOGY

Defining the length of Long-Haul Links. The analysis of intra- and inter-continental distances between peering locations reported on PeeringDB [9] reveals that these distributions have minimal overlap beyond 5,700 km. We use this observation to establish a 57 ms RTT (at $\frac{2}{3} \cdot c$ speed) threshold for identifying intercontinental LHLs.

Identyfing Long-Haul Links in traceroutes. Our methodology utilizes ATDK's LevelShift [2] to detect significant latency shifts as a hint of LHLs and then applies the derived latency threshold to discard hops that do not meet our criteria. We also augment the data using CAIDA's ITDK [4] for alias resolution and router geolocation and rely on Speed-of-Light constraints to remove any geolocation inaccuracy.

Applying our methodology. Our analysis uses traceroute datasets from CAIDA's Archipelago (Ark) platform [5], spanning seven years and involving 231.45 million traceroutes across all vantage points available. We utilize three consecutive measurement cycles in each annual snapshot to enhance link detection and minimize transient latency effects.

3 RESULTS

Our study focuses on the lengths and termination points of LHLs using 2022 CAIDA Ark data. We discover that LHLs primarily follow East-West paths, often exceeding the direct submarine cable routes in latency, due to the growing adoption of path virtualization technologies in the lower layers. Our analysis reveals the presence of *super routers*, central nodes in the network connecting multiple countries, and the prominence of global connectivity in the United States, a primary termination point for these links.

Our findings show a notable decoupling between LHLs and physical submarine cable locations, enabling routers to be located far inland, challenging prior assumptions for submarine cable mapping efforts. Many identified LHLs result from the wide adoption of virtualization technologies (e.g., MLPS), which hide physical links in virtual network-layer links connecting pairs of nodes, as far as Sao Paulo and Tokyo, to each other in a single hop. We investigated the adoption of identifiable MPLS tunnels and found a wide range of adoption levels across networks, with an average of 2.54% but with networks with adoptions higher than 90%. Beyond complicating infrastructure mapping efforts, adopting MPLS tunnels challenges routing optimization and debugging in the presence of path inflation, in some cases resulting from auto-bandwidth algorithms [8].

Our longitudinal study shows that the number of edges grew 2.9x, from 18,026 to 52,066, while the number of vertices grew 2.4x (from 9,560 to 23,267). Despite this growth, some properties of the LHnet have remained stable, including the inter-hop latency distribution as well as the prevalence of the intra-AS LHLs (79% to 89%).

[1] Code available at https://github.com/NU-AquaLab/intercont-LHL-2024

4 ACKNOWLEDGEMENTS

This work is supported by National Science Foundation grants CNS-2107392, CNS-1703592, CNS-2039146, and CNS-2106517. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. The authors also thank Erick W. Contag of SubOptic for valuable conversations related to this work.

REFERENCES

[1] Scott Anderson, Loqman Slamanatian, Zachary S. Bischof, Alberto Dainotti, and Paul Barford. 2022. iGDB: Connecting the Physical and Logical Layers of the Internet. In *Proc. of IMC.*
[2] ARUNO. 2022. ADTK Detectors. (2022).
[3] Zachary S. Bischof, Romain Fontugne, and Fabián E. Bustamante. 2018. Untangling the world-wide mesh of undersea cables. In *HotNets.*
[4] CAIDA. 2019. Macroscopic Internet Topology Data Kit (ITDK). http://www.caida.org/data/internet-topology-data-kit/.
[5] CAIDA. 2023. Archipelago Measurement Infrastructure Updates. https://catalog.caida.org/details/media/2011_archipelago. Accessed: 2021-9-30.
[6] Ramakrishnan Durairajan, Paul Barford, Joel Sommers, and Walter Willinger. 2015. InterTubes: A Study of the US Long-haul Fiber-optic Infrastructure. In *Proc. of ACM SIGCOMM.*
[7] Shucheng Liu, Zachary S Bischof, Ishaan Madan, Peter K Chan, and Fabián E Bustamante. 2020. Out of Sight, Not Out of Mind: A User-View on the Criticality of the Submarine Cable Network. In *IMC.*
[8] Abhinav Pathak, Mingn Zhang, Y Charlie Hu, Ratul Mahajan, and Dave Maltz. 2011. Latency inflation with MPLS-based traffic engineering. In *Proc. of IMC.*
[9] PeeringDB. 2019. https://www.peeringdb.com.
[10] Colin Wall and Pierre Morcos. 2021. Invisible and Vital: Undersea cables and transatlantic security. *CSIS* (2021).