Calypso's Voyage: Charting Traceroute Paths to Submarine Cables

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

Ying Zhang
Northwestern University
Evanston, IL, USA
yingzhang2024@u.northwestern.edu

Qianli Dong
Northwestern University
Evanston, IL, USA
qianlid@northwestern.edu

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

Ramakrishnan Durairajan
University of Oregon
Eugene, OR, USA
ram@cs.uoregon.edu

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

ABSTRACT

Submarine cables form the backbone of global Internet connectivity—yet they remain among the least understood and most fragile components of the infrastructure they sustain. Strengthening resilience demands knowing which cables and landing points are most critical and how severely their disruption would ripple through global connectivity. This depends on accurately mapping traffic onto the underlying infrastructure to identify vulnerabilities and assess their regional and global impact. Yet existing methods often lack the resolution and fidelity needed for such analysis, leaving researchers and policymakers without the insights to safeguard this vital system.

This paper introduces `Calypso`, a framework for mapping traceroute paths to the submarine cables they traverse. `Calypso` integrates ownership records, routing metadata, and geographic constraints to infer cable usage despite the opacity of the SCN and challenges such as route virtualization and inland infrastructure. It also defines *Route Stress*, a traceroute-derived metric for estimating the relative importance of submarine cables. Through expert validation, failure analysis, and regional case studies, we demonstrate `Calypso`'s utility in revealing SCN dependencies and informing resilience efforts.

1 INTRODUCTION

The submarine cable network (SCN) comprises over 600 cables and 7,500 segments, spanning 1.8 million kilometers across the ocean floor [19, 59, 62]. This underwater infrastructure carries the bulk of global Internet traffic and underpins global connectivity, supporting economies, communications, and international collaboration.

As reliance on the Internet and the underlying SCN grows, so do the risks. These range from accidental damage by fishing or anchoring [33] to natural disasters such as earthquakes and volcanic eruptions [22, 43, 60]. Geopolitical tensions add to these vulnerabilities, with submarine cables now seen as strategic assets in modern conflict. In February 2024, for instance, a Houthi attack in the Red Sea severed multiple cables, cutting four major systems, impacting 25% of regional traffic [46–48], and taking over four months to fully repair [18]. Other major failures such as the 2009 earthquake in Taiwan [49], Hurricane Sandy in 2012 [34], the 2022 volcanic eruption in Tonga [60], and 2023-2024 West African landslides [43] further underscore this fragility. Beyond their operational impact, cable failures are expensive. Repairs typically cost $1-3 million and take weeks to complete [67], yet despite the frequency, scale, and impact of these failures, we lack a scalable, measurement-based framework for understanding how end-to-end traffic traverses submarine cables or which cables and landing points are most critical to global connectivity.

This lack of visibility affects researchers, operators, and policymakers alike. While operators may know which cables they use, they typically cannot observe how other networks share or reroute across the same infrastructure, making it difficult to benchmark failover strategies, identify common bottlenecks, or assess resilience gaps. For regulators and international organizations (e.g., ITU), the absence of cable-level routing data impairs the ability to prioritize investments based on real-world dependencies.

Mapping traffic to submarine cables is central to addressing these challenges, yet the industry's opacity and the inherent limitations of end-to-end measurement make it stubbornly difficult. Several factors contribute: (1) cable usage is often governed by opaque leasing arrangements beyond physical ownership; (2) tunneling technologies like MPLS and SD-WAN obscure the physical path; and (3) ingress and egress points frequently lie far from coastal landing stations due to link-layer virtualization [21]. Prior work infers IP-level links based largely on geographic proximity to landing stations, but does not account for inland segments, multi-party leasing arrangements, or the decoupling between logical and physical paths.

Prior work [17, 40, 53] has taken steps toward automated cable mapping, but infers IP-level links based largely on geographic proximity to landing stations. These approaches do not account for inland segments, multi-party leasing arrangements, or the decoupling between logical and physical paths introduced by MPLS, SD-WAN, and other virtualization techniques — leaving a gap in our ability to map end-to-end traffic to the cables it traverses.

We introduce `Calypso`, the first framework for mapping traceroute paths to the submarine cables they traverse. `Calypso` combines data sources — cable topologies, AS rights of use, router geolocation, and inferred inland connectivity — to bridge the gap between logical L3 measurements and physical cable infrastructure. It consists of two components: `Chartbook`, a curated database encoding cable ownership and geography, and `Navigator`, a per-traceroute inference engine that maps traffic to cable segments using latency bounds, cable metadata, and path feasibility constraints.

`Calypso` is designed to serve researchers, operators, and policymakers seeking to understand and strengthen SCN resilience based on observed routing behavior rather than opaque or self-reported disclosures.

We validate `Calypso` through both expert feedback from SCN operators and empirical alignment with known cable failures. We illustrate its utility through two case studies: outbound routing from Australia and the 2024 disruptions affecting West and East Africa. These examples show how even a limited set of vantage points can surface rerouting patterns, identify fragility clusters, and highlight regional variation in resilience.

To quantify the operational importance of specific cables, we introduce Route Stress — a metric that captures how heavily network paths rely on individual submarine cables or landing points. Defined as the fraction of traceroutes that traverse a given cable, Route Stress provides a data-driven proxy for assessing infrastructure criticality and identifying concentrated points of failure. In §9.2, we show how a single metric—Route Stress—explains why two cable failures weeks apart devastated Kenya but barely touched South Africa, and vice versa.

Our paper makes the following contributions:

- We introduce `Calypso`, the first framework to map complete traceroute paths to their underlying submarine cables, jointly accounting for right-of-use agreements, inland routing, and virtualization artifacts.
- We validate our inferences using expert feedback from cable operators (56/66 confirmed mappings) and alignment with observed behavior during real-world cable disruptions.
- We demonstrate `Calypso`'s utility through two case studies: (1) a regional routing analysis of Australia's outbound traffic and (2) a comparative analysis of network responses to major 2024 failures in Africa.
- We introduce *Route Stress*, a metric that quantifies reliance on individual cables based on observed routing paths, enabling comparative analysis of cable-level importance and potential fragility.

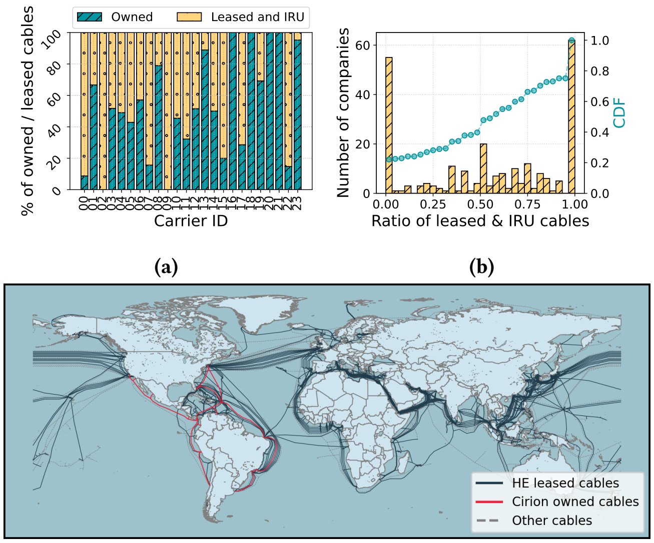
These contributions establish `Calypso` as a practical framework for revealing cable-level paths using large-scale, externally visible measurements. The remainder of the paper details `Calypso`'s design, implementation, and validation. *We will make Calypso available as an API for mapping traceroutes to submarine cables, with plans to open source the underlying code.*

2 BACKGROUND

The submarine cable network is the product of decades of investment, complex ownership arrangements, and evolving routing technologies. We summarize the key aspects of this ecosystem below.

2.1 Submarine Cable Network

The roots of the submarine cable network (SCN) date back to the 1850s with the deployment of telegraph cables. Since the late 1980s, their growth has been steady, with an average of one new cable activated per month. As of early 2025, there are ≈559 to ≈600 submarine cable systems that are active or under construction. These cables utilize advanced optical technologies capable of transmitting terabits of data per second, with the total capacity of the global submarine infrastructure growing by multiple orders of magnitude in recent decades [17, 63]. This growth is exemplified by recent projects such as Meta's 2Africa cable – designed to encircle much of the African coastline – and Google's Equiano and Grace Hopper systems, which connect key Internet hubs across Europe, Africa, and North America [32, 42].



(a) (b)

(c)

Figure 1: Leasing and ownership of submarine cable infrastructure. Fig. 1a shows ownership distribution among the 24 largest networks by customer cone. Fig. 1b: 25% of companies fully own their networks, while 22% rely entirely on leased cables. Fig. 1c contrasts two extremes: Cirion's fully owned network and Hurricane Electric's entirely leased one.

New cable projects are regularly announced to address challenges such as mitigating risks from natural disasters or geopolitical tensions, thereby improving the resilience of international connectivity [64]. Although cable landing points remain concentrated in major coastal hubs with existing infrastructure (e.g., colocation facilities), modern projects increasingly extend service to remote regions. These deployments leverage trunk-and-branch architectures [45, 68], primary cables with smaller branches delivering bandwidth inland, expanding global Internet access and fostering economic development in underserved areas (e.g., the Hawaiki Cable linking Australia, New Zealand, Hawaii and the continental US, has branches to New Caledonia, and a planned branch to Tonga [50]).

2.2 Cables' Owners and Tenants

There are two primary ways for an organization to gain rights to use submarine cables: ownership and leasing, each with distinct characteristics and trade-offs.

Ownership involves direct investment in submarine cable systems, either independently or through consortia. *Owners*, such as AT&T, Google, and Meta, hold stakes in these systems and often play a role in their governance. Given the substantial capital required to deploy submarine infrastructure, ownership is rarely undertaken by a single entity. For example, Vodafone independently funded the transatlantic Apollo cable, but this approach is an exception. More commonly, ownership is shared among multiple stakeholders through consortia. This model allows entities to pool resources and distribute risk, as seen in the case of Orange and China Telecom in the recently decommissioned SeaMeWe-3 [6]. Over time, the proportion of consortium-owned cables has grown, reflecting the industry's collaborative nature.

Leasing provides an alternative to ownership, allowing entities to access submarine cable capacity without high upfront costs. Tenants, such as ISPs and CDNs, acquire capacity through two main models: Capacity Leases and Indefeasible Rights of Use (IRUs). Capacity Leases offer short-term, renewable access to specific bandwidth amounts, enabling flexibility without long-term commitments. IRUs, in contrast, are long-term agreements (typically 20–30 years) granting exclusive, irrevocable rights to a portion of a cable's capacity, appealing to operators seeking stability while avoiding ownership complexities [7, 8]. Ownership and leasing strategies vary widely across the industry. Some entities, like Cirion Technologies (which acquired Lumen's Latin American assets), operate fully owned networks, while others, such as Hurricane Electric, rely entirely on leased capacity [29] (Fig.1c). However, the most common approach is a hybrid model, balancing flexibility and control by combining ownership and leasing. According to TeleGeography, 43% of entities follow this strategy, while 25% own their infrastructure outright (Figs.1a, 1b). This diversity underscores the resource-intensive and dynamic nature of the submarine cable industry.

2.3 Virtualization and Aggregation

Tenants have evolved to incorporate advanced routing technologies that enable fine-grained traffic control, moving beyond traditional IP's prefix-based, destination-driven forwarding, within Wide Area Networks (WANs).

Virtual Connections Through Logical Tunnels. The widespread adoption of technologies such as MPLS and SD-WAN has enabled the creation of virtual tunnels. These tunnels allow two network-layer devices (e.g., inland routers connected to either end of a submarine cable) to appear in traceroutes as if they were directly connected, even in the absence of a single, continuous physical link. By aggregating multiple physical segments into a single logical connection, these technologies enable: (i) seamless long-haul connections spanning multiple countries, and (ii) flexible, dynamic routing to optimize performance and resource utilization. For example, network paths can logically connect cities such as Seattle and Singapore while traversing numerous intermediate cable segments across diverse geographical regions [21].

Network operators frequently employ load-balancing techniques to aggregate and utilize multiple physical fiber strands within a submarine cable, which consists of multiple segments. This approach represents legitimate multi-fiber utilization rather than inaccuracies or measurement artifacts. For example, Australia's Academic and Research Network (AARNet)–which provides high-speed internet and data services to universities and research institutions–does not own physical infrastructure but instead secures long-term capacity agreements on various submarine cable systems. AARNet leases capacity from the Southern Cross Cable Network (SCCN) [2], which connects Sydney to Seattle via two fiber strands managed as an aggregated system. Private communications with cable operators confirm that traffic on the network's north branch (comprising segments G1, G2, and F) is actively load-balanced across both strands. Additional evidence supporting this can be seen in the network diagram published by AARNet [9], where the north branch is depicted with purple and green lines. This observation challenges the prevailing assumptions in SCN mapping efforts (e.g., [53]), which often rely on simplistic models that treat each traceroute as traversing a single cable. Such models fail to account for real-world operational practices, including load balancing.

Inland Submarine Path Ingress/Egress. The design and placement of submarine cable landing points are shaped by various geographic and topographic considerations [1, 35]. Gradually sloping seabeds, stable beaches, avoiding marine protected areas, and low-wave activity are critical factors. For instance, the South American Crossing (SAC) cable lands in Las Toninas, approximately 350 km south of Buenos Aires, due in part to the area's favorable seabed characteristics [36], and the ability to avoid ports and river deltas.

The aforementioned virtual connections also allow the ingress and egress points of submarine paths – the last router before and the first router after the landing point – to be located far inland. This decoupling of the network layer from the physical location of landing points facilitates efficient traffic routing within global-scale backbone networks, regardless of physical proximity to the landing sites.

3 CALYPSO'S GOALS AND CHALLENGES

Mapping network-layer paths to physical submarine cable infrastructure requires navigating the opacity of the SCN ecosystem and a set of interconnected technical challenges that existing approaches leave largely unresolved.

3.1 Goals

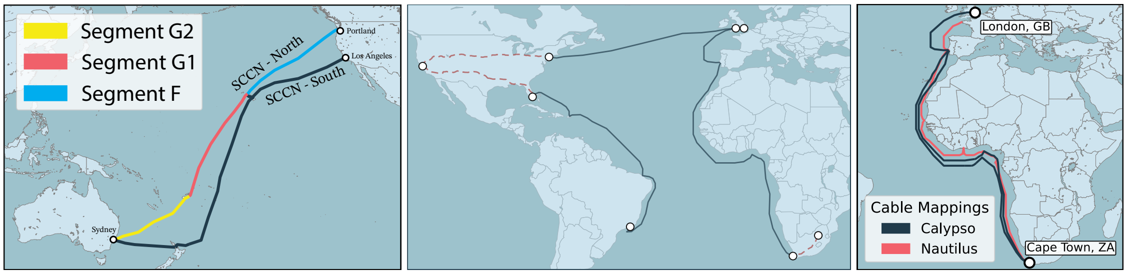
`Calypso` maps network-layer paths observed in traceroute data onto the physical submarine-cable segments that carry them, exposing how end-to-end Internet routes make use of the underlying cable infrastructure. To do so reliably, `Calypso` must map Layer-3 hops to the underlying physical infrastructure while resolving challenges such as multi-party cable ownership and leasing, cable aggregation, inland ingress points, and virtual links introduced by MPLS and SD-WAN.

`Calypso` addresses the needs of multiple stakeholders, including researchers, policymakers, and customers of network infrastructure. Researchers can gain valuable insights into the criticality and resilience of these networks, policymakers can be better equipped to assess and mitigate vulnerabilities in a vital component of national and economic security, while customers, ranging from network operators to enterprises, can better understand the infrastructure they depend on.

3.2 Challenges

The SCN ecosystem introduces four interconnected challenges that any mapping framework must confront.

Challenge 1: Determining who holds the right of use over a specific piece of infrastructure is complicated by the various roles entities play and their interrelationships within the SCN ecosystem



(a)

(b)

(c)

Figure 2: (a) Traffic is load-balanced across two fiber strands in segments G2, G1, and F over the North Branch, (b) Red dashed lines indicate distances between routers and landing points exceeding 1,000 km, preventing Nautilus from mapping the corresponding submarine cables, (c) Nautilus fails to infer multi-country cable sequences

(§ 2). More concretely, while cable ownership clearly provides right of use, with TeleGeography's public dataset [62] being the most widely referenced source for submarine cable ownership, capacity leasing and IRU are also options. For instance, Texlius, the global transit subsidiary of the Spanish-based company Telefonica, both owns and leases submarine cable capacity. Methods that consider only ownership relationships risk misattributing cable usage or overlooking legitimate traceroute–cable mappings.

Challenge 2: Determining which specific infrastructure segments (terrestrial and submarine) underlie each logical hop in a traceroute is complicated by the widespread use of route-virtualization techniques. As discussed in §2.3, technologies such as MPLS can make two consecutive routers appear directly connected in a traceroute even when the underlying path spans multiple submarine cables and inland fiber segments. Treating such virtual hops as single physical links leads to incorrect or failed cable mappings.

Challenge 3: A related challenge is distinguishing between mapping imprecision and genuine multi-cable aggregations. A single logical hop in a traceroute can correspond to one or several submarine cables. In some cases, multiple candidates arise from imprecision in the mapping process, where limited information leads to what we refer to as 'bundles', aggregations of possible cables associated with the hop. Additional information could reduce the size of these bundles, potentially isolating the mapping to a single cable. However, network operators also employ traffic engineering techniques that aggregate multiple cables, meaning that some hops truly rely on several cables, as seen in AARNet's use of SCCN described above.

Challenge 4: Determining where network ingress and egress points (e.g., routers in colocation facilities) connect to submarine infrastructure is also non-trivial. The difficulty stems from how landing stations are designed and how they physically connect to network-layer entry points. Even in otherwise *straightforward* cases (e.g., paths without MPLS), link-layer virtualization – such as long-haul optical circuits or pseudowires – can place these ingress or egress points far inland, in some cases over 1,300 km from the coast [21]. As a result, `Calypso` must explicitly account for inland ingress and egress when mapping traceroutes to physical cable infrastructure.

3.3 Limitations of Existing Work

Despite the critical role of the SCN in global connectivity [17], existing efforts to map traceroutes to submarine cables remain limited in scope and granularity.

Liu et al. [39] study the criticality of the SCN from the perspective of Internet end users, introducing a general methodology for analyzing a region's reliance on submarine connectivity. Their work intentionally abstracts the SCN as a single aggregate system rather than resolving individual cables, since the goal is to measure dependence, not topology. While this approach is well-suited for assessing regional exposure, it does not address the challenges involved in mapping traceroute segments to specific cables such as right-of-use constraints, link-layer virtualization, or cable aggregation that our work seeks to capture. Anderson et al. [11] introduced iGDB, a comprehensive global database for communication pathways across terrestrial infrastructures. Although the dataset includes TeleGeography's submarine cable information, its mapping efforts focus primarily on terrestrial infrastructure.[1]

More recent work has moved closer to cable-level inference. Ramanathan et al. [53] proposed Nautilus, a system for mapping IP links that likely traverse submarine cables, while Livadariu et al. [40] explored ownership- and naming-based heuristics to associate IP addresses with cables. Both represent important steps toward automated SCN mapping, yet share key limitations that restrict their applicability. We provide a detailed methodological comparison between `Calypso`, Nautilus [53], and Livadariu et al. [40]. We also quantify their respective mapping performance using the validation dataset introduced in §8.1, with results summarized in Table 4. We discuss traceroute geolocation details of the validation dataset in Appendix F.

Nautilus. Unlike `Calypso` and Livadariu et al. [40], Nautilus [53] maps only individual IP-level links to submarine cables by extracting pairs of consecutive hops with valid, non-private IPs. Any hop pair separated by non-responding routers ("*") is discarded, which means links containing missing responses are ignored. Using the authors' implementation [52], this filter eliminates 40.7% of links in our validation dataset (e.g., `210.57.30.93 -*- 203.50.6.239` in Table 4). To ensure a fairer comparison, we relax this constraint

---

[1] We confirmed this through communication with the authors.

so that Nautilus may attempt to map such links; this is the *only* modification we introduce.

Nautilus searches for landing points near each endpoint by expanding a geographic radius in 50 km increments up to 1,000 km. This heuristic works for coastal routers but does not hold for inland infrastructure. Routers in large countries like the United States, Brazil, Russia, Australia, and South Africa are frequently more than 1,000 km from the nearest relevant landing point. For example, row 3 of Table 4 shows an Ashburn–Hong Kong link; although Ashburn is on the U.S. East Coast, it is well beyond 1,000 km from the West Coast landing points used for trans-Pacific routes, causing Nautilus to return no match. Similar patterns appear for U.S. trans-Atlantic routes originating from the West Coast, and for Johannesburg–London and Sydney–Singapore paths. Figure 2b illustrates several such cases. Overall, one-quarter of the IP links in our validation dataset are unmappable by Nautilus due to this proximity heuristic. By contrast, `Calypso` does not assume that a router must lie within a fixed radius of its associated landing point.

Nautilus also cannot resolve links affected by *route virtualization*, where consecutive hops appear adjacent despite traversing physically disjoint infrastructure (*Challenge 2*, §3.2, Fig 2c). For example, these exists no direct cable connection between Cape Town and London[2]. Nautilus therefore is unable to provide any mapping between the two countries. These types of links account for approximately 24% of our validation dataset. On the other hand, `Calypso`'s `Navigator` module correctly identifies Portugal as the midpoint joining *WACS* and *Tata TGN-Western Europe* [44] between South Africa and London.

Finally, Nautilus does not incorporate leasing agreements or rights-of-use constraints, which leads to cable assignments that violate operator usage permissions.

Livadariu et al. Livadariu et al. [40] associate an IP address with a submarine cable if: (1) the IP's organization owns the cable; (2) its rDNS name contains a cable identifier; or (3) its prefix is announced by a cable owner. These criteria cover only a narrow subset of cables and do not account for *lessees* or *tenants*, leaving many legitimate operator-to-cable relationships unmapped. This limitation is reflected in the scale of their results with 717 IP links mapped to 11 cables. For example, although *Monet* is owned by a consortium (Algar, Angola Cables, Antel, Google), Telecom Italia Sparkle (AS 6762) operates it as part of its global backbone [57]. The IP `5.180.177.1` (AS 6762) has no rDNS entry and therefore cannot be mapped under the rules of [40]. `Calypso`'s `Chartbook`, by contrast, correctly associates it through encoded ownership and leasing relationships, together with reconciled organizational names from `Borges` [56].

As with Nautilus, the approach in [40] also relies on geographic proximity between link endpoints and cable landing points, making it unable to map links that span non-contiguous infrastructure.

4 CALYPSO: A FRAMEWORK OVERVIEW

`Calypso` consists of two components: `Chartbook`, a geographic and network-aware database, and `Navigator`, a per-traceroute mapping module. The foundation of `Calypso` is the `Chartbook` database, which integrates information on both submarine and terrestrial physical infrastructures, alongside the networks holding rights of use. `Calypso` leverages this database and network-layer path information to infer submarine cable segments. Using a hierarchical approach, it identifies physical connectivity across routers, networks, and countries, distinguishing between domestic and international links while accounting for both direct and indirect connections. The following two sections detail how `Chartbook` is constructed and how `Navigator` performs its inference. .

5 CALYPSO'S CHARTBOOK

The core of `Calypso` is `Chartbook`,[3] a database built by collecting, parsing, and organizing information on both terrestrial and underwater physical infrastructures, as well as the networks that hold rights of use. `Chartbook` contains (1) all submarine cables including various features such as landing point locations, (2) all networks with rights of use for each cable, and (3) terrestrial path information needed to interconnect submarine segments.

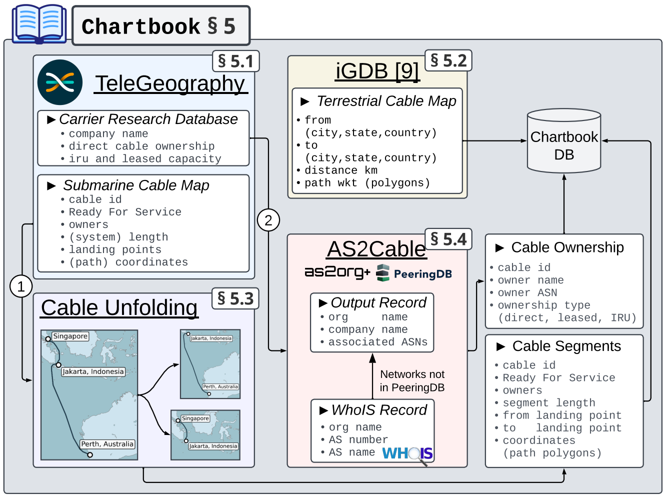


Figure 3: `Calypso`'s `Chartbook` Database.

Figure 3 provides an overview of the data sources (§5.1, §5.2) and processing steps (§5.3, §5.4) used to build `Chartbook`.

5.1 Submarine Cables and Rights of Use

`Calypso` relies on TeleGeography [63] as a primary data source for submarine cable information, which provides (1) a Submarine Cable Map and (2) a Transport Network Research Database detailing the rights of use for each cable. TeleGeography's Submarine Cable Map [62] serves as a key resource for understanding global submarine cable infrastructure, offering detailed information on cable systems, including cable polylines that describe their physical layout, landing points at both country and coordinate levels, cable lengths, and Ready-for-Service (RFS) dates. As of October 2024, the TeleGeography Submarine Cable Map contains information on 615 distinct cable systems.

At the core of `Calypso`'s `Chartbook` are *cable polylines* (GeoJSON RFC 7946 [20]), which represent physical paths of submarine

---

[2] *2Africa* was not yet operational at the time of our study.

[3] A chartbook refers to a bound collection of nautical charts used by sailors and navigators to plot courses across bodies of water.

cables as a series of geographic coordinates. While this data was previously available to the public, TeleGeography has since restricted access to it. The most recent publicly accessible dataset we identified dates back to 2022 [10, 11]. To obtain up-to-date cable layout information, we have secured a contractual agreement with TeleGeography.

The TeleGeography Transport Network Research Database is a proprietary resource that identifies carrier networks' rights of use, whether through ownership, leased capacity, or Indefeasible Rights of Use (IRUs), on each submarine cable network. As of October 2024, this database details the right of use via ownership, leased capacity, and IRUs of 450 distinct providers. We augment this dataset using automated web search and scraping techniques following Zhuang et al. [70], allowing us to incorporate publicly reported cable-capacity leasing agreements, including those used in our case study in Sec. 9.

5.2 Incorporating Terrestrial Connectivity

In addition to submarine cables, `Calypso` incorporates terrestrial data to capture connectivity between landing points and inland termination points. To do so, `Calypso` relies on iGDB [10, 11], a geographic database that integrates the Internet's logical and physical layers by identifying the locations of facilities housing network equipment (e.g., peering facilities, colocation centers, IXPs) and inferring the physical infrastructure that interconnects them. While public databases like PeeringDB [3] list network facilities, iGDB goes further by inferring physical connectivity between facilities, assuming that network cables follow rights-of-way associated with existing infrastructures such as roads, railways, or gas pipelines [27].

A key mismatch in integrating TeleGeography's submarine-cable data with iGDB's terrestrial infrastructure data lies in their differing geographic representations. iGDB partitions the world into 7,342 Thiessen polygons, each corresponding to the region closest to a specific urban area (e.g., a city, town, or other locality). Terrestrial paths are thus represented as pairs of polygons denoting source and destination regions. In contrast, `Calypso` represents submarine-cable segments as pairs of landing-station *coordinates* (§5.3).

To obtain a unified representation of both terrestrial and submarine infrastructure, we associate each cable landing point with its corresponding iGDB polygon by identifying the polygon that *spatially contains* [5] the landing-station coordinate (a standard point-in-polygon operation). This allows `Calypso` to seamlessly link submarine and terrestrial cable systems into a continuous topology. We use only terrestrial fiber data from iGDB to connect inland infrastructure with TeleGeography's submarine-cable data.

5.3 Submarine Cable Unfolding

While the TeleGeography and iGDB datasets provide extensive views of submarine and terrestrial infrastructure and their associated rights of use, their raw formats require parsing and restructuring before they can support cable-level inference. Submarine cable systems often include multiple landing points, which serve as ingress and egress locations for traffic entering the underwater infrastructure. TeleGeography reports the total system length, but this value reflects only the distance between the two most distant landing points. For precise analysis, we construct a database containing distances between all pairs of landing points through a process we term cable unfolding (① in Figure 3).

Cable unfolding generates a segment for each pairwise combination of landing points within a system, producing $\frac{N \cdot (N-1)}{2}$ segments for a cable with N landing points. The process leverages TeleGeography's cable polylines, which specify the cable layout as a sequence of coordinates. `Calypso` computes the great-circle distance between each consecutive pair of coordinates [54] and sums these values to obtain the length of each unfolded segment. Appendix C illustrates the procedure for the Indigo-West system connecting Australia, Indonesia, and Singapore. Starting with 615 submarine cable systems, the unfolding process yields 15,233 individual segments.

5.4 ASN to Submarine Cable Bindings

The final step in constructing `Calypso`'s database is reconciling TeleGeography's network identifiers with network-layer paths. Inferring submarine cable segments requires linking routers – and their corresponding operators and Autonomous Systems (ASes) – to specific submarine cables. While TeleGeography's Transport Network Database lists which networks are authorized to use each cable, it relies on commercial names rather than technical identifiers such as ASNs. To bridge this gap, `Calypso` builds a company-to-ASN mapping, integrating network-layer information with the commercial naming conventions used in TeleGeography's dataset (② in Fig. 3).

Mapping commercial names to ASNs is difficult because available data sources are incomplete, inconsistent, and often out of date. WHOIS records associate ASNs with organization names, but these names frequently reflect legal entities rather than operational brands. Many entries lag behind rebrandings, mergers, and acquisitions. For example, AS52320 (formerly GlobeNet) now operates as V.tal, yet WHOIS still lists it as GlobeNet.

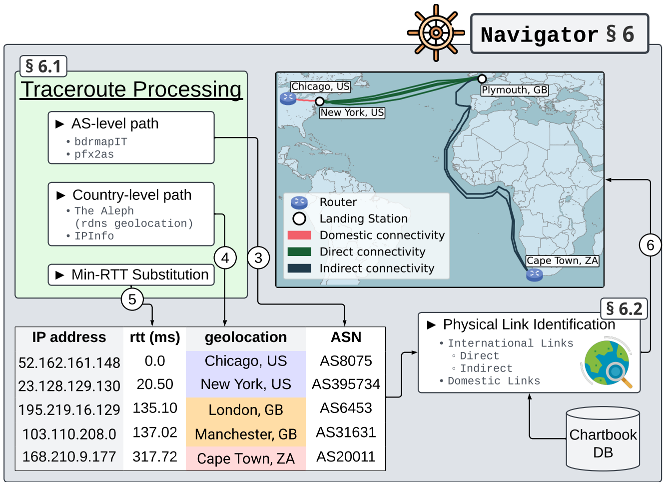
To resolve these inconsistencies, `Calypso` uses *Borges* [56] to group ASNs belonging to the same organization. It then searches PeeringDB for current organization names and consults WHOIS only for ASNs missing from PeeringDB, leveraging the latter's operator-maintained and therefore more current nature. `Calypso` applies fuzzy matching between these ASN-associated names and TeleGeography's company names; when a match is found, it links the corresponding ASNs to the relevant submarine cables. This process yields the ASN-to-cable bindings required to address *Challenge 1* in §3.2.

6 CALYPSO'S NAVIGATOR

In this section, we describe how `Navigator` processes network-layer path information to identify submarine cable segments using the data available in `Calypso`'s `Chartbook` database. Figure 4 presents an overview of this module.

6.1 Path Inference

The first step in identifying submarine cable segments along network-layer paths is converting traceroute data into a format compatible with `Calypso`'s database. Since `Calypso` maintains information at both the country level (submarine cable landing locations) and



IP address	rtt (ms)	geolocation	ASN
52.162.161.148	0.0	Chicago, US	AS8075
23.128.129.130	20.50	New York, US	AS395734
195.219.16.129	135.10	London, GB	AS6453
103.110.208.0	137.02	Manchester, GB	AS31631
168.210.9.177	317.72	Cape Town, ZA	AS20011

Figure 4: Overview of `Calypso`'s `Navigator` module

the network level (organizations with rights of use), we map each network-layer path into both an AS-level path and a country-level path to ensure alignment between traceroute results and `Calypso`'s dataset.

To determine AS-level paths ③, `Calypso` employs bdrmapIT [41] as its primary IP-to-ASN resolution method. bdrmapIT integrates RIR delegation files, AS2Org mappings, and AS-/prefix-level customer cone information to infer AS ownership and identify routers functioning as interdomain connection points. When bdrmapIT is unable to resolve an IP address, `Calypso` falls back on a prefix-to-AS resolution approach, using longest prefix matching against RouteViews route dumps to determine the originating network. For country-level paths ④, `Calypso` combines PTR-based geolocation inference with geolocation databases. If PTR records are available, `Calypso` uses The Aleph[37], a system that extracts location hints using an extensive set of parsing rules. When PTR records are unavailable, `Calypso` relies on IPInfo [25] and performs active latency measurements using RIPE Atlas probes to verify IPInfo's geolocation results.

In addition to path mapping, `Calypso` incorporates inter-hop latency measurements from traceroutes ⑤. Since traceroute results can be affected by congestion and routing anomalies, we mitigate these effects by using the minimum Round Trip Time (RTT) observed for each hop, providing a more stable latency estimate.

6.2 Mapping Traceroute Paths

`Calypso` integrates network and geographic information to systematically map the physical infrastructure underlying a traceroute path. It follows a hierarchical approach, representing connectivity as a graph, where vertices correspond to geographic locations (e.g., cities) and edges represent physical links such as terrestrial or submarine cables. Each edge in the connectivity graph is annotated with attributes, including cable type (`submarine` or `terrestrial`), cable length, and, for submarine cables, the specific cable name. Terrestrial connection data is sourced from iGDB, while submarine cable data comes from TeleGeography's Submarine Cable database. Edges that connect locations within a country are labeled `domestic`, while those spanning different countries are classified as `international`. Submarine cable landing points are explicitly marked, allowing `Calypso` to infer connections between submarine cables along a given route.

Mapping traceroutes to physical infrastructure requires identifying the most plausible path through the connectivity graph from source to destination. Exhaustive search is computationally impractical; `Calypso` therefore systematically prunes the route space before inference begins. It first detects international transitions by identifying country changes along the traceroute and grouping hops that remain within the same country. This preserves path structure while significantly reducing complexity. The following subsections describe how `Calypso` maps traceroutes onto physical infrastructure, with a full algorithm provided in Appendix D.

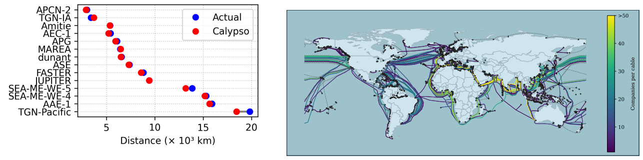
International Connectivity. When `Calypso` detects a country transition in a traceroute, it first checks whether the network has physical infrastructure that directly connects the two countries. If no such link exists, it searches for indirect connectivity by exploring intermediary countries through which the path could traverse while remaining within the network's right of use.

Direct Connectivity. If a network owns or leases physical infrastructure connecting two countries, `Calypso` assumes it will use the most direct route so long as the selected segment adheres to speed-of-light constraints. This applies to cases where countries are separated by a body of water (e.g., the US and Japan) as well as cases where a land border exists but submarine infrastructure is still available (e.g., Argentina and Brazil), where undersea cables may serve as high-capacity backbones. When multiple valid cable paths exist, `Calypso` considers all paths that meet speed-of-light constraints derived from inter-hop latency measurements, corresponding to the multimodal routing behaviors discussed in § 2.

In some cases, an interdomain hop suggests that the two networks may be connected via submarine infrastructure. When this occurs, `Calypso` considers all submarine cables that fall under the combined rights of use of both networks. If no such infrastructure exists, the hop may be virtualized, e.g., created by tunneling techniques, which implies that the physical path likely traverses additional intermediate countries, or the `Chartbook` may be missing information about one or both networks' rights of use.

Indirect Connectivity. When virtualization causes two routers to appear adjacent despite the absence of a direct link, `Calypso` searches for routes through intermediary countries. For each candidate sequence of countries, `Calypso` checks three conditions: (1) physical infrastructure must exist between each pair, (2) the network must have the rights to use that infrastructure, and (3) the resulting path must satisfy speed-of-light latency constraints. `Calypso` returns all intermediary-country paths that meet these criteria.

Domestic Connectivity. After identifying international connections, `Calypso` infers the domestic infrastructure linking the ingress and egress points within each country along the route. To achieve this, it repeats the pathway search process, focusing on infrastructure connecting landing points within the source, destination, and any intermediary countries.



(a) (b)

Figure 5: (a) comparison of cable lengths between TeleGeography and actual layouts; (b) global coverage of leasing & ownership information recorded in `Chartbook`

At this stage, `Calypso` considers both terrestrial infrastructure and domestic submarine cables, which connect domestic landing points. Notably, domestic submarine infrastructure is widely deployed, even in continental nations such as Brazil, Chile, and Venezuela, where it provides high-capacity connectivity between major population centers.

7 SENSITIVITY TO DATA COVERAGE AND ACCURACY

`Calypso`’s mapping fidelity depends on the coverage and accuracy of its input datasets. We assess in this section how discrepancies in TeleGeography’s cable layouts and cable ownership information, iGDBs’ terrestrial infrastructure, geolocation precision affect the mapping outcomes of `Calypso`.

TeleGeography. We discuss `Calypso`’s use of cable layout and ownership information data from TeleGeography.

Cable Layouts. TeleGeography provides submarine cable geometries as sequences of geographic coordinates in Well-Known Text (WKT) format. Rather than exactly representing the physical routes of deployed systems, these paths serve as approximations [65], which may introduce minor variations in speed-of-light (SoL) calculations. In practice, these differences are small, particularly with respect to cable *length*, and have negligible impact on SoL feasibility.

To quantify this, we compare lengths derived from TeleGeography’s layouts with those obtained from verified cable geometries. Because full submarine paths are typically confidential, broad validation is not possible. Instead, we use the subset of publicly documented routes from Arelion’s global network map [12]. We crawled, extracted, and parsed these coordinates, yielding real-world paths for 14 cable systems (Fig. 5a). The largest discrepancy arises for *TGN-Pacific*: TeleGeography’s layout differs from Arelion’s by 1,504 km over 20,000 km (7.5%), or roughly 15 ms of additional latency. Across all traceroutes traversing *TGN-Pacific*, we observe no cases where this difference affects SoL-based feasibility checks.

Cable Ownership. A key source of `Calypso`’s mapping accuracy is the *Transport Network Research Database* within `Chartbook` (§5.1), which encodes ownership and leasing arrangements for over 450 submarine cables. This information is key for distinguishing among parallel systems landing at the same or nearby stations and for determining which cables a given network is allowed to use.

Figure 5b summarizes global coverage of ownership data. Information is most complete for trans-Pacific systems and routes linking East Asia through South Asia, the Middle East, and Western Europe, enabling more confident inferences in these regions. Coverage is sparse for intra–Southeast Asian systems, where limited visibility leads to greater ambiguity and larger candidate sets.

Improving the completeness of ownership and leasing metadata remains important for further increasing `Calypso`’s precision. We are exploring the use of Large Language Models (LLMs) to automatically identify and extract such information from publicly available online sources.

iGDB. iGDB [10, 11] aggregates Internet measurement datasets into a unified, cross-layer repository. It organizes information across: (1) the *physical layer* (PoPs, IXPs, terrestrial and submarine fiber as GIS shapefiles); (2) the *logical layer* (IP–AS mappings, traceroutes, AS relationships, geolocations); and (3) *cross-layer links* connecting physical and logical entities. All datasets include timestamps to enable longitudinal analysis.

In our work, we rely exclusively on iGDB’s *terrestrial fiber* data. Submarine cable information is drawn from TeleGeography to ensure current coverage. Although iGDB’s terrestrial layer is older, such infrastructure changes slowly, reducing the risk of temporal mismatch. We omit iGDB’s logical and cross-layer components to avoid outdated dependencies.

Terrestrial fiber paths in iGDB originate from Internet Atlas [28]. While generally accurate, occasional omissions can lead to discrepancies. A notable example involved traffic between Tokyo and London: measured latencies were below the speed-of-light bound of any known submarine route. Further investigation revealed the Transit Europe–Asia (TEA) terrestrial network – a land route across Russia [61]. Traffic enters Russia via the Russia–Japan Submarine Cable Network (RJCN), producing round-trip latencies near 200 ms, well below the 260-300 ms of submarine-only paths. Private operator correspondence confirmed this routing.

Because iGDB lacked the TEA segment, `Calypso` initially considered only submarine alternatives and returned an incorrect mapping. Adding the TEA terrestrial path to `Chartbook` resolved the issue. This remains the only case in which iGDB’s terrestrial data materially affected `Calypso`’s results.

IP address	ASN	RTT	Geolocation
`168.209.1.166`	3741	`1.099`	Cape Town, South Africa
`168.209.100.214`	3741	`170.257`	Cape Town, South Africa
`83.231.235.221`	2914	`180.149`	London, United Kingdom

Table 1: The second hop is incorrectly geolocated to South Africa rather than its actual location in the UK. This leads to an RTT inconsistent with speed-of-light constraints.

Geolocation Data. Accurate hop-level geolocation is essential for `Calypso` and for related systems [40, 53]. Geolocation mistakes, especially those that violate speed-of-light (SoL) constraints, can invalidate `Navigator` mappings. Table 1 shows an example: the second hop is incorrectly placed in Cape Town instead of London, producing an RTT that is physically impossible and blocking any valid cable inference.

To mitigate such errors, `Calypso` uses IPInfo [25], which incorporates latency-based validation, and cross-checks rDNS-derived geohints using `The Aleph` [37] to ensure consistency. When rDNS records are not available, `Calypso` performs ping measurements from RIPE Atlas probes. In our ground-truth dataset, no traceroutes exhibit geolocation inaccuracies that affect `Calypso`'s mapping. Improvements in geolocation quality will increase the fraction of traceroutes that `Calypso` can map with high precision.

8 VALIDATION

Like much Internet infrastructure research, our methodology contends with a fundamental scarcity of ground truth. Despite this, we validate `Calypso` using two approaches: expert feedback and failure-based analysis. First, we consult submarine cable experts to verify our mappings. Second, we analyze traceroute changes during real-world cable failures to assess mapping accuracy, showing the robustness of our approach.

	c1	c2	c3	c4	c5	c6	c7	c8	c9	c10	c11	c12	c13	c14	c15	c16	c17	c18
● Q1	7/0	4/0	2/0	9/0	2/0	1/0	2/0	5/0	4/0	4/0	1/0	0/0	5/0	4/0	2/0	2/0	1/0	1/0
● Q2	1/0	0/0	1/0	1/0	1/0	0/0	1/0	0/0	0/0	1/0	0/0	1/0	1/0	0/0	0/0	0/0	2/0	0/0

Table 2: Validation result (carriers redacted for privacy reasons (*7/0: 7 yes* and *0 no* answers). All mappings for which we received response were confirmed by the experts, with no reported inaccuracies.

8.1 Direct Operator Feedback

We collaborated with submarine-cable experts[4] including operators, service providers, and independent consultants to validate `Calypso`'s outputs through interviews and case-specific assessments. Our validation centered on the traceroute inferences from our case studies (§9), where we asked operators whether the inferred cable paths matched their operational knowledge.

We obtained direct validation for 18 transport carriers (e.g., Telefónica, PCCW, Tata Communications) and 66 IP transport links, covering 80% of traceroutes used in §9.

For each case, we posed two key questions: ● (Q1) Is our identified mapping correct? ● (Q2) If full disclosure is restricted due to confidentiality concerns, do our inferences provide a plausible explanation?

Responses to Q1 directly assess `Calypso`'s accuracy, while Q2 helps account for privacy constraints and operator expertise on terrestrial routing, ensuring even partial confirmations contribute valuable insights. Table 2 summarizes the feedback from SCN experts on our findings from `Calypso`. We consulted experts on 66 mappings produced by `Calypso`, receiving responses for 56 of them (● Q1). Among these, experts confirmed that all inferences were correct, with no reported inaccuracies. For the remaining 10 mappings (● Q2), operators refrained from commenting due to the sensitivity of certain findings. Notably, eight of these involved only terrestrial infrastructure, while the other two included terrestrial and submarine infrastructures.

Date	Reported & Affected Cables	Affected Countries
2020.01.09	FALCON	Kuwait (and others)
2020.01.17	ACE, SAT-3, WACS	Cameroon
2020.08.09	SeaMeWe-4, SeaMeWe-5	Bangladesh
2022.06.07	AAE-1, PEACE, SeaMeWe-5	Djibouti (and others)
2023.01.28	AAE-1, APG, TGN-IA, AAG	Vietnam
2024.03.15	ACE, WACS, SAT-3/WASC	South Africa (and others)

Table 3: Reported cable failure events and our inferred cables. Green represents an inference that matches the reported affected cable, blue represents a cable that is not reported but appears in our inferences.

8.2 Failure-based Validation

Failure-based validation relies on detecting changes in traceroute data during submarine cable failures. A key assumption is that network routes remain stable during the validation window. If a route is stable, its inferred cable identification should also remain unchanged. Thus, a change in traceroute data during a reported cable failure supports the correctness of our mapping. The validation process follows these steps: (1) Identify cable failure events, (2) Select traceroutes from the affected geographic region, spanning before, during, and after the failure, (3) Map pre-failure traceroutes to submarine cables using `Calypso`, (4) Confirm the mapping if the failure coincides with route changes or latency spikes.

Traceroute Datasets. To capture failure events, we require high-frequency, stable traceroute measurements. RIPE Atlas' root server probes provide an ideal dataset, with traceroutes performed every 30 minutes. Prior work [69] has shown that anycast routes remain relatively stable, aligning with our validation assumptions.

Validation Process. We demonstrate failure-based validation using the June 7, 2022, outage affecting the AAE-1 and SeaMeWe-5 cables, which disrupted connectivity across Africa, the Middle East, Europe, and Asia. Analyzing RIPE Atlas traceroutes from Djibouti, we observed consistent paths before and after the failure but detected route changes and latency spikes during the event. Using `Calypso`, we inferred three cables (AAE-1, PEACE, and SeaMeWe-5) before the failure. Comparing our mappings against officially reported cable failures from 2019 to 2024 (Table 3) further strengthens our confidence in `Calypso`'s accuracy.

While full ground truth is unavailable, `Calypso` prioritizes precision by mapping only when multiple signals – including latency, geography, and ownership – are in agreement. As a result, some crossings may go undetected, which we conservatively treat as false negatives. In our South Africa dataset alone, `Calypso` maps 65.3% of over 700K traceroutes, with results aligning closely with known failures and operator feedback. This suggests that unmapped paths reflect conservative thresholds rather than systemic errors.

9 USE CASES OF CALYPSO

`Calypso`'s ability to uncover critical SCN dependencies and quantify the impact of real-world disruptions is best illustrated through three case studies: an analysis of Australia's submarine cable infrastructure, emphasizing the role of its cables and landing points; an investigation into two recent cable failures in Africa, highlighting how Route Stress captures the asymmetric impact of disruptions on network performance; and a global applicability analysis that examines how input-data limitations manifest across regions at scale.

---

[4] Details elided for anonymity; will be added upon publication.

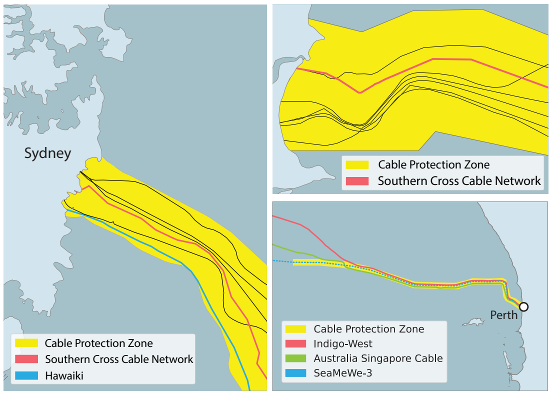


Figure 6: Original figures obtained and modified from [13–15]. Cable Protection Zones (CPZ) starting from the left and moving clockwise: South Sydney CPZ, North Sydney CPZ, and Perth CPZ.

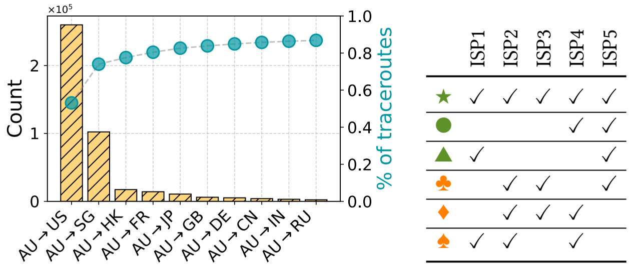
9.1 (Re)assessing Australia's CPZs

Escalating geopolitical tensions between the US and China are particularly pronounced in the Indo-Pacific, with Australia emerging as a focal point. This highlights the critical role of the submarine cable network in the region.

A recent report by the Australian Strategic Policy Institute (ASPI) underscores this concern, recommending regulatory adjustments, including reassessing submarine cable protection zones (CPZs). The Australian Communications and Media Authority established three CPZs (Perth, North Sydney, and Southern Sydney) in 2007 to protect critical infrastructure from external risks such as anchoring and fishing [23, 24]. These zones currently cover 18 landing points and 23 submarine cables, including AJC, SCCN, and the ≈39,000-km-long SeaMeWe-3. Given Australia's geographic position, maintaining a robust SCN infrastructure is essential for national and international connectivity. Figure 6 shows the three CPZs.

We focus on Australia for two key reasons. First, as ASPI notes, no new CPZs have been created since 2007. Second, Australia's telecommunications market is highly concentrated, with 80% of the population served by just five ISPs: Telstra, Optus, TPG, Vocus, and SuperLoop [38]. This raises concerns about overreliance on existing CPZs as the telecom sector's dependence on submarine infrastructure grows. The ASPI report emphasizes the need to "obtain and analyze data on cable disruptions within and outside of CPZs to assess the zones' effectiveness and inform decisions on whether more zones should be declared." `Calypso` offers a data-driven framework to support policy decisions such as those aligned with ASPI's recommendations.

To analyze Australia's international connectivity, we collected 480K traceroutes from 100 RIPE Atlas probes (measurement 5051 & 5151) during August 2024. `Calypso` is able to map 92.5% of traceroutes to submarine cables, while 7.5% either terminate within Australia or have its last visible hop domestically. Figure 7a shows over 50% of outbound routes are directed to the US and 20% to Singapore.



	ISP1	ISP2	ISP3	ISP4	ISP5
★	✓	✓	✓	✓	✓
●				✓	✓
▲	✓				✓
♣		✓	✓		✓
♦		✓	✓	✓	
♠	✓	✓		✓	

(a) (b)

Figure 7: (a) Outbound routes from Australia primarily to the US and Singapore, and (b) cables used by ISPs for US- and Singapore-bound routes (ISPs anonymized).

For US-bound routes, three submarine cables dominate: ★ Southern Cross Cable Network, ● Southern Cross NEXT, and ▲ Hawaiki. Singapore-bound routes depend on ♣ Indigo West, ♦ SeaMeWe-3, and ♠ Australia-Singapore Cable. Table 7b shows the submarine cables used by each ISP for US and Singapore-bound routes. Notably, `Calypso`'s `Chartbook` ensures high-fidelity mapping of traceroutes to physical infrastructure, and our validation in §8 confirms its accuracy across all five major ISPs.

`Calypso` provides a data-driven framework for mapping Australia's submarine-cable dependencies using only externally visible measurements. The systems identified by `Calypso` as central to Australia's connectivity correspond to those classified as protected infrastructure by the Australian government. As geopolitical tensions and reliance on submarine cables grow, empirical insights from `Calypso` can guide regulatory adjustments, such as expanding the CPZs to safeguard national and regional connectivity.

9.2 Criticality of SCN Failures in Africa

Mapping which cables a traceroute uses is essential, but it does not reveal how heavily a country depends on each system or why some failures have far greater impact than others. The February and March 2024 events in Africa make this clear: two disruptions weeks apart affected South Africa and Kenya in dramatically different ways, despite involving many of the same cable systems. Explaining this asymmetry requires characterizing traffic distribution across cables, not just identifying which cables appear in paths.

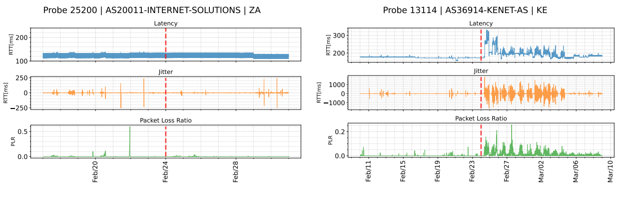
To demonstrate this, we examine two major incidents in early 2024. On February 24, a Houthi attack that sank the Rubymar severed three submarine systems in the Red Sea: AAE-1, Seacom/TGN, and EIG [46]. Figures 8a and 8b show the impact on two RIPE Atlas probes, one in South Africa (probe 25200, AS3741) and one in Kenya (probe 13114, AS36914), during this event. Kenya experienced sharp performance degradation, with latency spikes exceeding 300 ms, increased jitter, and packet loss above 0.2%, while South Africa remained almost entirely unaffected.[5]

Less than a month later, on March 14, while repairs to the February damage were still pending [16], a rockslide off Côte d'Ivoire disrupted four west-coast systems: WACS, ACE, SAT-3/WASC, and MainOne [43]. This time, the situation reversed: South Africa experienced marked performance degradation, whereas Kenya saw relatively little impact (Figs. 8c and 8d).

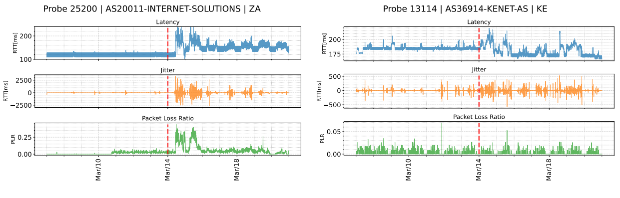
---

[5] While we present data from one probe per country for clarity, other probes within the same regions exhibited similar trends.

These contrasting outcomes highlight the need for a metric that captures how much a country depends on each cable, not just which cables appear in paths.



(a) 2024.02.24 ZA **(b) 2024.02.24 KE**



(c) 2024.03.14 ZA **(d) 2024.03.14 KE**

Figure 8: South Africa has limited impact by the 2024.02.24 event in the Red Sea (8a) while showing disruptions in the 2024.03.14 event in Cote d'Ivoire (8c). In contrast, Kenya shows disruptions in the first event (8b) and limited impact in the second (8d).

Quantifying the precise importance of a submarine cable is complex. We estimate relative importance by drawing from prior research [27, 55], which suggests that inter-AS traffic loads can be approximated by analyzing the number of traceroutes traversing a particular link. We define the *Route Stress* of a submarine cable as the proportion of routes in a traceroute dataset that rely on that cable. Formally, for a collection of traceroutes TR and a submarine cable A, the Route Stress S_A is:

$$S_A = \frac{\sum t_i}{|T_{SC}|} \forall (t_i \in T) \wedge (\mathcal{P}(t_i) \subset A)$$

where $\mathcal{P}(.)$ maps a Layer-3 traceroute to the physical infrastructure it traverses, and T_{SC} is the subset of traceroutes in TR that include at least one submarine cable.

We analyzed traceroute measurements from all RIPE Atlas probes in South Africa targeting the /24 address space during January 2024. This dataset comprised 115 probes across 56 distinct ASes, generating 700k traceroutes. `Calypso` mapped 65.3% of these traceroutes to submarine cables, with 23% relying on validated IP transit links (see §8). Figure 9 shows the resulting Route Stress distribution, highlighting ACE, WACS, and SAT-3/WASC on South Africa's west coast as among the most critical cables for international connectivity.

The disruption of these systems in the March 14, 2024 event had a substantial effect on South Africa's connectivity, reflected in the pronounced degradation across latency, jitter, and packet-loss metrics (Fig. 8c). In contrast, the February 24 Red Sea incident affected cables landing in both Kenya and South Africa but had a notably greater impact in Kenya, illustrating how cable importance varies significantly across regions, even within the same continent.

By quantifying Route Stress, we provide a metric for assessing the relative importance of submarine cables and anticipating the consequences of failures. The February 24 and March 14 incidents illustrate how differences in traffic distribution and available cable capacity shape the severity of disruptions. In South Africa, the high Route Stress of ACE, WACS, and SAT-3/WASC made their failure impactful, resulting in significant performance degradation.

These findings underscore the need for proactive infrastructure planning, including redundancy measures and alternative routing strategies, to mitigate risks and strengthen global connectivity.

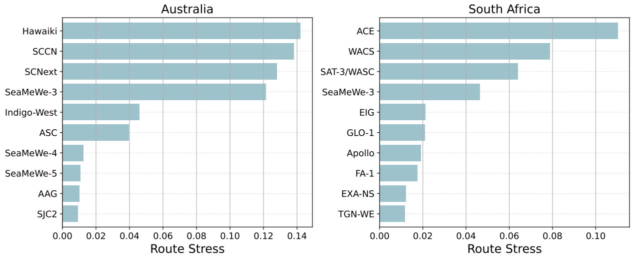


Figure 9: Route Stress values for case studies in Australia and South Africa.

9.3 Route Stress in Australia

Route Stress, introduced in §9.2, provides a quantitative view of how heavily a region depends on specific submarine cables. Applying the same analysis to Australia reveals a pronounced concentration of critical connectivity in the Southern Sydney region.

As shown in Figure 9, the cables carrying the highest Route Stress values are those linking Australia to the United States, consistent with the observations in §9.1. These high-stress systems all land within or adjacent to the Southern Sydney CPZ, which has long served as the country's principal international gateway. Importantly, recently deployed and in-deployment systems that `Calypso` identifies as key international paths also fall within the expanded CPZ boundaries announced by the Australian government in October 2025 [51]. This alignment between `Calypso`'s externally observable inferences and the government's protection priorities underscores the structural importance of this landing region and provides empirical support for the CPZ expansion.

Taken together, these results demonstrate how Route Stress enables a data-driven assessment of Australia's submarine-cable dependencies.

9.4 Global Applicability

Extending beyond our regional case studies, we assess `Calypso`'s applicability at global scale by applying it to a worldwide traceroute dataset. We use RIPE Atlas anchor mesh measurements [26], consisting of all-pairs traceroutes between 1,091 anchors over an 18-hour period on 2025-08-01. This yields approximately 120 million traceroutes and 15 billion hops.

Given this scale, we omit latency-based geolocation validation and rDNS lookups via `The Aleph`, and instead remove traceroutes that violate speed-of-light (SoL) constraints (e.g., Table 1). We also

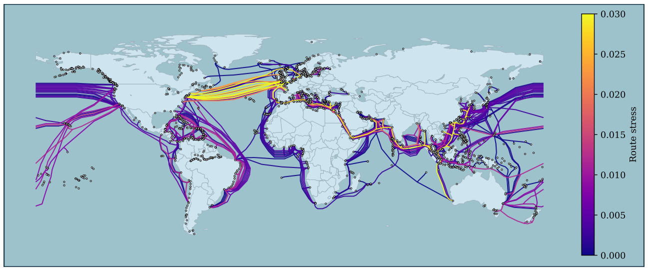


Figure 10: Route Stress for global RIPE Atlas anchor measurements

exclude routes that do not traverse submarine cables such as paths between anchors within the US or within Europe where only terrestrial connectivity is observed.

Figure 10 shows the resulting global Route Stress distribution. Trans-Atlantic cables exhibit the highest stress values, primarily reflecting the geographic distribution of RIPE Atlas anchors rather than limitations in input datasets. Most anchors are located in the US (164), Germany (123), the Netherlands (57), France (50), and the United Kingdom (44), producing a disproportionate number of Europe–North America traceroutes. Regions with sparse anchor coverage (e.g., Africa, South America, parts of Southeast Asia) generate fewer observable routes, which correspondingly reduces the statistical confidence of inferred stress values rather than the correctness of individual mappings.

Across all regions, we observe that limitations in iGDB, geolocation precision, or TeleGeography coverage manifest as localized mapping gaps rather than systematic regional biases. In other words, where these datasets are complete, `Calypso` behaves consistently; where information is missing, `Calypso` returns candidate bundles or marks links as indeterminate. The global results therefore reflect the availability of measurement vantage points, not region-specific inaccuracies in the underlying datasets.

10 RELATED WORK

Our work builds on foundational research into the physical infrastructure of network communications, with particular emphasis on the submarine cable network. Prior efforts have examined both terrestrial [11, 27] and SCN infrastructures [40, 58]. Durairajan et al. [27] conducted an exhaustive investigation into the physical backbone of the U.S. Internet. More recently, Anderson et al. [11] introduced iGDB, a global-scale database providing information on both terrestrial and submarine communication pathways. The SCN has garnered increasing attention in research and, more recently, in popular media. Bischof et al. [17] proposed a research agenda to better understand this complex global infrastructure. Fanou et al. [31], Liu et al. [39], Ramanathan et al. [53], and Livadariu et al. [40] leveraged these insights to explore the impact of specific deployments, assess SCN criticality from an end-user perspective, map IP links onto their underlying submarine cables, and translate public SCN announcements into information to facilitate research into SCNs. As discussed in §3.3, these studies have largely sidestepped the challenges of mapping entire traceroutes to submarine cables. In a related study, Carisimo et al. [21] analyzed the *long-haul link network*, highlighting the adoption of virtualization techniques for interconnecting remote locations in a single hop, while Bhosale et al. [30] assessed using LEO networks as backups for submarine cable failure events. Additionally, Starosielski developed a gamified mapping tool, Surfacing, to enhance SCN literacy [58].

11 DISCUSSION AND LIMITATIONS

Beyond the case studies, Calypso's results carry broader implications for how the submarine cable network is deployed, operated, and governed — and expose limitations that contextualize the framework's current scope.

11.1 Broader Implications

Much of the existing research on SCN resilience focuses on physical robustness while underemphasizing the regulatory forces that shape how new infrastructure is approved and deployed. Delays and restrictions imposed by bodies such as Team Telecom introduce systemic vulnerabilities: they slow critical upgrades, constrain transpacific capacity expansion, and create chokepoints that magnify economic and operational risks. Network operators have detailed visibility into their own infrastructure but limited insight into how other networks route traffic or expose shared dependencies across jurisdictional boundaries. Policymakers and organizations such as ISOC and ITU must assess whether today's infrastructure can meet future demands for security and resilience. `Calypso` offers what has so far been missing from this dialogue: a data-driven, externally observable characterization of cable-level dependencies that neither operators nor policymakers need to self-report.

11.2 Limitations

Securing ground-truth validation remains challenging: many operators are reluctant to share detailed information, and comprehensive validation of cable-level paths remains an open problem. `Calypso` also inherits limitations from TeleGeography and iGDB, which contain gaps, inconsistencies, and temporal staleness that place natural limits on coverage and precision. Finally, `Calypso` does not model operational factors such as on-demand capacity reallocation or time-varying traffic engineering policies — incorporating such dynamics is an important direction for future work.

12 SUMMARY

We introduced `Calypso`, a framework that maps complete traceroutes to the submarine cables they traverse, and Route Stress, a traceroute-derived metric for estimating the relative importance of individual cables. `Calypso` addresses the challenges posed by the opacity of the submarine cable industry and the complexity of mapping network-layer paths to physical infrastructure. Through rigorous validation and case studies – including an analysis of Australia's submarine cable network and the impact of recent events in Africa – we demonstrated the efficacy of `Calypso` and its potential value in assessing SCN vulnerabilities, and provide actionable insights for researchers and policymakers aiming to enhance the resilience of the global submarine cable infrastructure.

REFERENCES

[1] [n. d.]. Currents, Tides, Wrecks, and More – The Impact of the Environment on Undersea Connectivity. https://sub.co/articles/currents-tides-wrecks-and-more-the-impact-of-the-environment-on-undersea-connectivity/. ([n. d.]). Accessed: 2024-12-12.

[2] 2016. AARNet and Southern Cross Cable Network boost trans-Pacific research network to 100G. (2016). https://www.aarnet.edu.au/aarnet-and-southern-cross-cable-network-boost-trans-pacific-research-network-to-100g Accessed: 2025-01-28.

[3] 2023. PeeringDB. https://catalog.caida.org/dataset/peeringdb. (2023). Dates used: 2023-12-08. Accessed: 2023-12-08.

[4] 2023. RouteViews Prefix to AS mappings. https://catalog.caida.org/dataset/routeviews_prefix2as. (2023). Dates used: 2023-12-08. Accessed: 2023-12-08.

[5] 2024. GeoSeries.contains: Test if geometries contain another geometry. GeoPandas Documentation. (2024). https://geopandas.org/en/stable/docs/reference/api/geopandas.GeoSeries.contains.html Accessed: 2025-01-30.

[6] 2024. SEA-ME-WE 3 cable retired after 25 years in service. https://www.datacenterdynamics.com/en/news/seamewe-3-cable-retired-after-25-years-in-service/. (December 4 2024). Accessed: 2024-12-12.

[7] 123NET. 2024. Understanding Indefeasible Rights of Use (IRU). (May 2024). https://www.123.net/blog/understanding-indefeasible-rights-of-use-iru/ Accessed: 2024-12-15.

[8] 123NET. 2024. Understanding IRU Fiber: A Comprehensive Guide. (March 2024). https://www.123.net/blog/understanding-iru-fiber-a-comprehensive-guide/ Accessed: 2024-12-15.

[9] AARNet. 2024. AARNet and Southern Cross Cable Network Boost Trans-Pacific Research Network to 100G. (2024). https://aarnet.edu.au/aarnet-and-southern-cross-cable-network-boost-trans-pacific-research-network-to-100g Accessed: 2025-01-30.

[10] Steven Anderson and contributors. 2022. iGDB: The Internet Geographic Database. https://github.com/standerson4/iGDB. (2022). Accessed: 2025-11-15.

[11] Scott Anderson, Loqman Salamatian, Zachary S. Bischof, Alberto Dainotti, and Paul Barford. 2022. iGDB: connecting the physical and logical layers of the Internet. In *Proc. of IMC*.

[12] Arelion. 2025. Arelion Global Network. https://www.arelion.com/our-network. (2025). Accessed: 2025-11-06.

[13] Australian Communications and Media Authority (ACMA). 2019. Northern Sydney Protection Zone. Online. (2019). https://www.acma.gov.au/publications/2019-10/mapping/northern-sydney-protection-zone Accessed: 2025-01-30.

[14] Australian Communications and Media Authority (ACMA). 2019. Southern Sydney Protection Zone. Online. (2019). https://www.acma.gov.au/publications/2019-10/mapping/southern-sydney-protection-zone Accessed: 2025-01-30.

[15] Australian Communications and Media Authority (ACMA). 2019. Submarine Telecommunications Cables Perth Protection Zone Map. Online. (2019). https://www.acma.gov.au/publications/2019-10/mapping/submarine-telecommunications-cables-perth-protection-zone-map Accessed: 2025-01-30.

[16] Harry Baldock. 2024. Red Sea Cable Repairs Delayed As Yemeni Govt Probes AAE-1 Consortium. (May 2024). https://subtelforum.com/yemen-government-probes-delay-aae-1-cable-repairs/

[17] Zachary S. Bischof, Romain Fontugne, and Fabián E. Bustamante. 2018. Untangling the world-wide mesh of undersea cables. In *Proc. of HotNets*.

[18] Bloomberg News. 2024. Damaged Internet Cables Repaired in Red Sea as Houthis Attack Ships. (July 2024). https://www.bloomberg.com/news/articles/2024-07-17/damaged-internet-cables-repaired-in-red-sea-as-houthis-attack-ships Accessed: 2025-01-22.

[19] Lucy Bricheno, Isobel Yeo, Michael Clare, James Hunt, Allan Griffiths, Lionel Carter, Peter J. Talling, Megan Baker, Stuart Wilson, Matthew West, Semisi Panuve, and Samuiela Fonua. 2024. The diversity, frequency and severity of natural hazard impacts on subsea telecommunications networks. *Earth-Science Reviews* (2024). https://www.sciencedirect.com/science/article/pii/S0012825224003003

[20] Howard Butler, Martin Daly, Allan Doyle, Sean Gillies, and Tim Schaub. 2016. The GeoJSON Format. Internet Requests for Comments. (2016). https://datatracker.ietf.org/doc/html/rfc7946

[21] Esteban Carisimo, Caleb J. Wang, Mia Weaver, Fabián E. Bustamante, and Paul Barford. 2023. A Hop Away from Everywhere: A View of the Intercontinental Long-haul Infrastructure. *sigmetrics* (2023).

[22] Kenjiro Cho, Cristel Pelsser, Randy Bush, and Youngjoon Won. 2011. The Japan Earthquake: the impact on traffic and routing observed by a local ISP. In *conext special workshop on the internet and disasters*.

[23] Australian Communications and Media Authority. 2024. Zone to protect Perth submarine cables. (Feb 2024). https://www.acma.gov.au/zone-protect-perth-submarine-cables

[24] Australian Communications and Media Authority. 2024. Zone to protect Sydney submarine cables. (2024). https://www.acma.gov.au/zone-protect-sydney-submarine-cables#sydney-protection-zones

[25] Omar Darwich, Hugo Rimlinger, Milo Dreyfus, Matthieu Gouel, and Kevin Vermeulen. 2023. Replication: Towards a Publicly Available Internet Scale IP Geolocation Dataset. In *Proc. of IMC*.

[26] Alun Davies. 2017. Anchoring Measurements: Bringing Back the Balance. https://labs.ripe.net/author/alun_davies/anchoring-measurements-bringing-back-the-balance/. (June 2017). RIPE Labs article.

[27] 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*.

[28] Ramakrishnan Durairajan, Subhadip Ghosh, Xin Tang, Paul Barford, and Brian Eriksson. 2013. Internet atlas: a geographic database of the internet. In *Proceedings of the 5th ACM Workshop on HotPlanet*. https://doi.org/10.1145/2491159.2491170

[29] Hurricane Electric. 2024. Hurricane Electric Internet Services - Peering Report. (2024). https://bgp.he.net/report/peers#_adjacencies Accessed: 2024-12-07.

[30] Bhosale et al. 2025. Assessing LEO Satellite Networks for National Emergency Failover *(IMC 2025)*. https://saeed.github.io/files/cosmosim-imc25.pdf

[31] Rodérick Fanou, Bradley Huffaker, Ricky Mok, and KC Claffy. 2020. Unintended Consequences: Effects of Submarine Cable Deployment on Internet Routing. In *Proc. of PAM*.

[32] Google. 2023. Equiano Cable Project. (2023). https://cloud.google.com/blog/products/infrastructure/announcing-the-equiano-subsea-cable.

[33] Mick Green and Keith Brooks. 2011. *The Threat of Damage to Submarine Cables by the Anchors of Cables Underway*. Technical Report. Centre for International Law, National University of Singapore. https://cil.nus.edu.sg/wp-content/uploads/2011/04/Mick-Green-and-Keith-Brooks-The-Threat-of-Damage-to-Submarine-Cables-by-the-Anchors-of-Cables-Underway.pdf

[34] Geoff Huston and George Michaelson. 2012. Superstorm Sandy and the Global Internet. The ISP Column. (November 2012).

[35] International Cable Protection Committee Ltd (ICPC). 2009. Submarine cables and the oceans: connecting the world. https://www.iscpc.org/publications/icpc-unep_report.pdf. *UNEP-WCMC Biodiversity Series* 31 (2009). Accessed: 2024-12-12.

[36] Infobae. [n. d.]. El cable que, desde Las Toninas, conecta a Internet toda la Argentina. https://www.infobae.com/2014/10/13/1601116-el-cable-que-las-toninas-conecta-internet-toda-la-argentina/. ([n. d.]). Accessed: 2024-12-12.

[37] Fabian E. Bustamante Kedar Thiagarajan, Esteban Carisimo. 2025. The Aleph. Conext 2025. (2025). https://estcarisimo.github.io/assets/pdf/papers/2025-the-aleph.pdf Accessed: 2025-01-30.

[38] APNIC Labs. 2025. AS Population Statistics for Australia - 28 January 2025. (2025). https://stats.labs.apnic.net/cgi-bin/aspop?c=AU&d=28/01/2025 Accessed: 2025-01-29.

[39] 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 Conference. In *Proc. of IMC*.

[40] Ioana Livadariu, Ahmed Elmokashfi, and Georgios Smaragdakis. 2024. Tracking submarine cables in the wild. *Computer Networks* (2024).

[41] Alexander Marder, Matthew Luckie, Amogh Dhamdhere, Bradley Huffaker, kc claffy, and Jonathan M. Smith. 2018. Pushing the Boundaries with bdrmapIT: Mapping Router Ownership at Internet Scale. In *Proc. of IMC*.

[42] Meta. 2023. 2Africa Project Overview. (2023). https://about.fb.com/news/2020/05/2africa/.

[43] Admire Moyo. 2024. Multiple Subsea Cable Breaks Knock SA's Internet. (March 2024). https://subtelforum.com/south-africa-internet-hit-by-multiple-subsea-cable-breaks/

[44] Submarine Networks. 2025. West Africa Cable System (WACS). (2025). https://www.submarinenetworks.com/en/systems/euro-africa/wacs Accessed: 2025-01-26.

[45] Md. Nooruzzaman and Toshio Morioka. 2018. Multicore fibers for high-capacity submarine transmission systems. *Journal of Optical Communications and Networking* (2018).

[46] Mohammed Hatem Olivia Solon. 2024. Houthi-Sunk Ship's Anchor Likely Severed Sea Internet Cables. (March 2024). https://www.bloomberg.com/news/articles/2024-03-06/anchor-from-houthi-sunk-ship-likely-damaged-undersea-cables

[47] Lipika Pelham. 2024. Cargo ship attacked by Houthis sinks off Yemen coast. (Mar 2024). https://www.bbc.com/news/world-middle-east-68457445

[48] The Associated Press. 2024. Cargo ship sinks in Red Sea nearly 2 weeks after being hit by Yemen's Houthi rebels. (Mar 2024). https://www.npr.org/2024/03/02/1235521730/red-sea-cargo-ship-sinks-houthis-rubymar

[49] Winston Qiu. 2011. Submarine Cables Cut by Taiwan Earthquake and Typhoon Morakot. Submarine Cable Networks. (March 2011).

[50] Winston Qiu. 2024. Australian and New Zealand Governments to Fund Tonga Hawaiki Branch System. *Submarine Cable Networks* (June 2024).

[51] Winston Qiu. 2025. Australian ACMC Enlarges Sydney Submarine Cable Protection Zone upon Google's Request. *Submarine Networks* (2025). https://www.submarinenetworks.com/en/nv/insights/acmc-enlarges-sydney-submarine-cable-protection-zone

[52] Alagappan Ramanathan. 2025. Nautilus-Code Repository. https://gitlab.com/netsail-uci/nautilus/-/tree/main/code. (2025).

[53] Alagappan Ramanathan and Sangeetha Abdu Jyothi. 2024. Nautilus: A Framework for Cross-Layer Cartography of Submarine Cables and IP Links *(SIGMETRICS '24)*. https://doi.org/10.1145/3652963.3655068

[54] P.N. Ruane. 2014. Heavenly mathematics: the forgotten art of spherical trigonometry by Glen Van Brummelen, pp 192, £24.95, ISBN 978-0-691-14892-2, Princeton University Press (2013). *The Mathematical Gazette* (2014).

[55] Mario A. Sanchez, Fabián E. Bustamante, Balachander Krishnamurthy, Walter Willinger, Georgios Smaragdakis, and Jeffrey Erman. 2014. Inter-Domain Traffic Estimation for the Outsider. In *Proc. of IMC*.

[56] Carlos Selmo, Esteban Carisimo, Fabián E. Bustamante, and J. Ignacio Alvarez-Hamelin. 2025. Learning AS-to-Organization Mappings with Borges *(IMC 2025)*. https://doi.org/10.1145/3730567.3732918

[57] Telecom Italia Sparkle S.p.A. 2017. Sparkle Global Backbone. https://www.globalbackbone.tisparkle.com/. (2017). Accessed: 2025-11-08.

[58] Nicole Starosielski. 2013. Surfacing: A Digital Mapping of Submarine Systems. SubOptic 2013. (2013). https://www.academia.edu/3492742/Surfacing_A_Digital_Mapping_of_Submarine_Systems Accessed: 2025-01-30.

[59] Nicole Starosielski, Iago Bojczuk, John Hooft-Toomey, Ramakrishnan Durairajan, Hannah Ellis, Erick Contag, and John Tibbles. 2025. Enhancing Strategic Resilience of Subsea Cables in the Caribbean. In *Proceedings of SubOptic '25*. Lisbon, Portugal.

[60] Chris Stokel-Walker. 2022. Tonga's volcano blast cut it off from the world. Here's what it will take to get it reconnected. *MIT Technilogy Review* (January 2022).

[61] Submarine Networks. 2011. Transit Europe-Asia (TEA) Terrestrial Cable Network. https://www.submarinenetworks.com/en/systems/asia-europe-africa/tea/tea-cable-network. (2011). Accessed: 2025-11-09.

[62] TeleGeography. 2023. Submarine Cable Map. (2023). https://www.submarinecablemap.com/.

[63] TeleGeography. 2024. TeleGeography Official Website. (2024). https://www2.telegeography.com/

[64] TeleGeography. 2025. Ready for Service 2025. (2025). https://www.submarinecablemap.com/ready-for-service/2025.

[65] TeleGeography. 2025. Submarine Cable Frequently Asked Questions. https://www2.telegeography.com/submarine-cable-faqs-frequently-asked-questions. (2025). Accessed: 2025-11-08.

[66] Telstra International. n.d.. Global Network. https://www.telstrainternational.com/en/global-network. (n.d.). Accessed: 2025-01-22.

[67] Dean Veverka. 2011. Under the sea. International Cable Protection Committee. (September 2011). www.shippingandmarine.co.uk.

[68] Zengfu Wang, Qing Wang, Bill Moran, and Moshe Zukerman. 2020. Optimal Submarine Cable Path Planning and Trunk-and-Branch Tree Network Topology Design. 28, 4 (August 2020), 1562–1572.

[69] Lan Wei and John Heidemann. 2017. Does anycast hang up on you?. In *Proc. of TMA*.

[70] Shuying Zhuang, Jessie Hui Wang, Jilong Wang, Zujiang Pan, Tianhao Wu, Fenghua Li, and Zhiyong Zhang. 2021. Discovering obscure looking glass sites on the web to facilitate internet measurement research. In *Proc. of CoNEXT*.

A ETHICS

This paper introduces a framework that leverages existing traceroute datasets and information on the SCN. As such, it raises no ethical issues related to user privacy and places no new burden on network resources.

B SAMPLE OF TELEGEOGRAPHY'S DATASET

As illustrated in Figure 11a, TeleGeography's Submarine Cable Map provides structured data on the Grace Hopper submarine cable, owned by Google, not only showing its layout but also describing its 7,000 km span, connecting Spain and England in Europe to the northeastern US across the North Atlantic. Figure 11b illustrates a sample entry from this database. In this example, Telstra has direct ownership of the Telstra Endeavour & Asia-America Gateway cables, and holds IRUs and leased capacity on the SCCN & Hawaiki cables.[6]

```
{
    "cable_name": "Grace Hopper",
    "RFS": "2022 September",
    "owners": ["Google"],
    "length": 7,191 km,
    "landing_points": [{"name":"Bude, United Kingdom"}, ...],
    "coordinates": [(-16.19, 46.27) ,(-9.89, 46.89), ...]
}
```

(a) JSON data for cable profile information

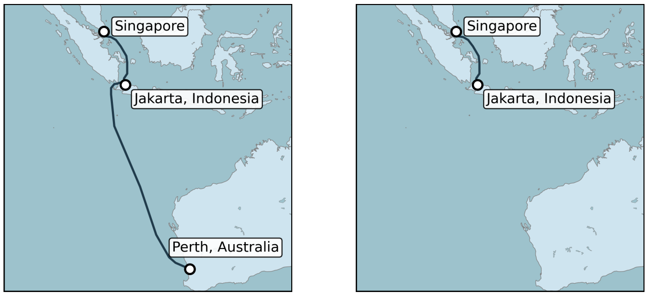
```
{
    "company_name": "Telstra",
    "direct_cable_ownership": [ "Telstra Endeavour",
        "Asia-America Gateway (AAG)",  ... ],
    "iru_and_leased_capacity": ["Southern Cross, SCCN", "Hawaiki", ...]
}
```

(b) JSON data for carrier cable ownership information

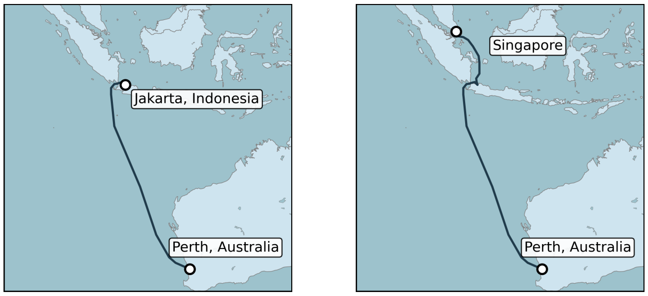
Figure 11: Examples of data provided by Telegeography.

C `CALYPSO`'S CABLE UNFOLDING

The example below illustrates the cable-unfolding step, in which `Calypso` expands a multi-landing-point system into explicit segments between each landing-point pair. As shown in Figure 12a, the Indigo-West cable system connects Australia to both Indonesia and Singapore. Although its total reported length is 4,600 km, unfolding reveals three distinct segments: `AU` → `ID` (3,400 km), `AU` → `SG` (4,600 km), and `ID` → `SG` (1,040 km). A packet traveling from Australia to Indonesia would therefore traverse only the `AU` → `ID` segment rather than the entire system length.



(a) Full cable system **(b) SG to ID segment**



(c) ID to AU segment **(d) SG to AU segment**

Figure 12: (a) Indigo-West cable system with total length of 4,600 km, and (b), (c), (d) each representing a distinct segment

---

[6] Our contractual agreement with Telegeography does not let us redistribute disaggregated information appearing in this database. A full map of Telstra's network infrastructure can be found at [66].

D CALYPSO'S MAPPING TRACEROUTE TO PHYSICAL INFRASTRUCTURE

Algorithm 1 detail `Calypso`'s approach to mapping traceroutes to physical infrastructure.

Algorithm 1 TRACEROUTE2CABLES(**tr**)

```
$C \leftarrow$ getConnectivityGraph()
$P \leftarrow$ getCountryPairs(tr)
$Q \leftarrow \emptyset$   // international links
$R \leftarrow \emptyset$   // country landing points
$S \leftarrow \emptyset$   // domestic links
for all $p$ in $P$ do
    // identify int'l connectivity for each landing
    // point pair (p) in P
    if $C.hasEdge(p)$ then $L \leftarrow C.getDirectLinks(p)$
    else $L \leftarrow C.getIndirectLinks(p)$
    end if
    $Q.insert(L)$
    // add ingress landing point to src country
    // add egress  landing point to dst country
    $R.ccsrc.insert(L.src)$; $R.ccdst.insert(L.dst)$
end for
for all $r$ in $R.ccsrc \times R.ccdst$ do
    // identify domestic connectivity for landing
    // point pairwise combinations (r) in R
    $S.insert(C.getDomesticLinks(r))$
end for
return $(Q, S)$
```

E IMPACT OF CABLE RIGHTS-OF-USE ON BUNDLE SIZE IN AUSTRALIA

In this section, we extend the analysis from § 9.1 by examining how bundle sizes change when *rights of use* constraints are applied. We normalize bundle-size reductions by the link frequencies observed in our traceroute dataset and present the results in Fig. 13. The figure shows how bundle sizes shift across country–country and ASN–ASN links. Enforcing rights of use consistently reduces the number of plausible cables for many links, demonstrating how operator-specific constraints sharpen `Calypso`'s inferences.

F COMPARING TRACEROUTE- AND LINK-LEVEL MAPPING APPROACHES

We compare `Calypso` against two recent systems that tackle related but narrower problems [40, 53]. Whereas these systems map individual IP links to submarine cables, `Calypso` operates at the *traceroute* level, enabling end-to-end mapping across entire submarine segments. For a fair comparison, we extract from our ground-truth dataset the IP links relevant to both methodologies and evaluate all systems under identical conditions.

All systems depend on accurate hop-level geolocation. We use a shared ground-truth dataset, geolocating hops via IPInfo [25] (with latency-based validation) and verifying results through RIPE Atlas measurements. For hops with rDNS records, we cross-check locations using `The Aleph` [37]. All submarine-related hops were successfully validated. We further annotate hops with AS and organizational context using CAIDA's prefix-to-ASN [4] and `Borges` [56] datasets to ensure consistency across systems. Table 4 summarizes the comparison.

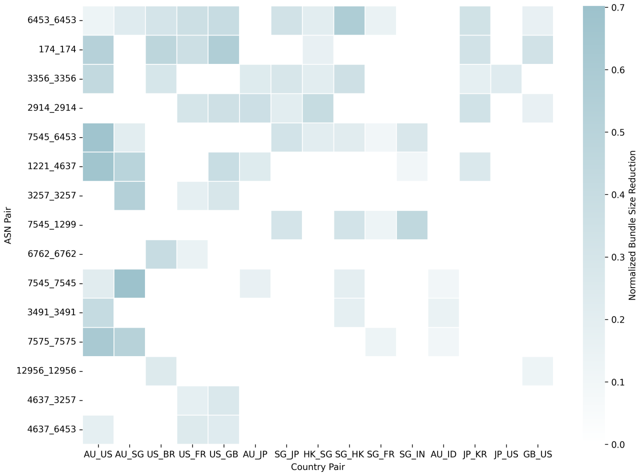


Figure 13: Normalized Bundle Size Reductions by Country and ASN Pairs under Cable Rights-of-Use

Cable Links to Map	Calypso	Nautilus [53]	Livadariu et al. [40]
5.53.1.37 -*- 102.130.70.202	SACS	landing point > 1000km	SACS
177.124.128.129 -*- 84.16.15.69	SAM-1, GlobeNet, BRUSA	©Seabras-1	ownership-missing
84.16.15.69 -*- 63.220.194.190	AAG	landing point > 1000km	ownership-missing
177.124.128.129 -*- 94.142.99.163	SAM-1, GlobeNet, BRUSA	BRUSA,©Seabras-1	ownership-missing
94.142.99.163 - 94.142.99.170	MAREA	MAREA	MAREA
177.124.128.129 -*- 94.142.99.170	BRUSA → MAREA	LHL-unmapped	LHL-unmapped
176.52.248.24 -*- 84.16.15.69	SAM-1, GlobeNet, BRUSA	©Seabras-1	ownership-missing
84.16.15.69 -*- 213.140.38.94	AC-1 → terrestrial	LHL-unmapped	LHL-unmapped
94.142.99.163 -*- 94.142.98.241	SAM-1, GlobeNet, BRUSA	©Dunant,©Apollo,©FA-1©MAREA,©AC-1,©FA-1	ownership-missing
94.142.98.241 - 5.53.6.192	terrestrial	terrestrial	no mapping
213.140.53.51 - 129.250.2.12	SAM-1, GlobeNet, BRUSA	©Seabras-1	ownership-missing
129.250.6.177 - 129.250.4.143	PC-1, TGN-Pacific, JUPITER, TPE	PC-1, TGN-Pacific, TPE	APG, APCN-2, ASE ownership-missing
129.250.4.143 -*- 129.250.5.242	APG, APCN-2, EAC-C2C, ASE	APG, APCN-2, EAC-C2C, ASE, ©SJC	JUPITER, TPE, ownership-missing
203.50.13.86 - 202.84.222.134	SCNext,Hawaiki,SCCN	SCCN	ownership-missing
203.50.11.176 - 210.57.30.93	Indigo-West	landing point > 1000km	Indigo-West
210.57.30.93 -*- 203.50.6.239	EAC-C2C, APCN-2	EAC-C2C, APCN-2, ◷SeaMeWe-3, ©APG	APCN-2, ◷SeaMeWe-3
114.31.206.51 - 193.251.248.35	SeaMeWe-3	landing point > 1000km	SeaMeWe-3
4.68.68.125 - 200.189.213.114	Monet, GlobeNet	landing point > 1000km	ownership-missing
203.50.11.176 - 210.57.30.93	Indigo-West, SeaMeWe-3	landing point > 1000km	ownership-missing
141.136.105.130 -*- 199.168.63.126	FA-1 → SeaMeWe-3	LHL-unmapped	LHL-unmapped
203.50.9.11 - 202.84.140.109	SeaMeWe-4 → EIG	LHL-unmapped	LHL-unmapped
154.54.0.222 - 154.54.82.37	Apollo,Yellow,EXA-NS, FA-1	EXA-NS	ownership-missing
64.86.160.0 - 64.86.143.83	Apollo,Yellow, FA-1	landing point > 1000km	ownership-missing
80.231.20.82 -*- 52.85.22.1	SeaMeWe-3, FEA	SeaMeWe-3, ©EIG (wrong cable)	ownership-missing
198.142.249.250 - 203.208.143.185	SCNext,SCCN	SCCN	ownership-missing
110.175.58.250 -*- 114.31.192.70	Hawaiki,SCCN	SCCN	ownership-missing
216.6.99.21 - 216.6.99.70	Americas-II,Seabras-1	©GlobeNet,Seabras-1	ownership-missing
198.142.249.242 - 203.208.147.113	SeaMeWe-3	landing point > 1000km	SeaMeWe-3
63.217.237.125 - 63.223.19.186	EAC-C2C, SeaMeWe-3 , FNAL	EAC-C2C,SeaMeWe-3,FNAL,©APCN-2,©APG	SeaMeWe-3 , FNAL, ownership-missing
206.148.24.164 -*- 213.202.6.222	OAC	©SeaMeWe-3	ownership-missing
206.82.129.19 - 66.110.59.114	AAG → TGN-Indicom	LHL-unmapped	LHL-unmapped
123.176.118.33 -*- 63.222.112.50	SCCN	SCCN	ownership-missing
116.51.31.166 - 103.87.124.69	SeaMeWe-3	SeaMeWe-3,©SeaMeWe-4,-5	SeaMeWe-3
110.175.58.250 -*- 203.219.107.114	SeaMeWe-3, Indigo-West	landing point > 1000km	ownership-missing
203.219.107.114 - 27.111.228.83	PEACE Cable, SeaMeWe-4,-5	SeaMeWe-4,-5	ownership-missing
203.29.134.122 - 64.86.197.96	SCCN	SCCN	ownership-missing
64.86.197.56 -*- 5.180.177.1	Americas-II, Seabras-1, Monet	landing point > 1000km	ownership-missing
203.221.3.35 - 213.248.95.232	SCCN	SCCN	ownership-missing
58.178.249.248 -*- 203.134.4.5	ASC	landing point > 1000km	ASC
129.250.3.214 -*- 129.250.2.111	FA-1, Apollo,Yellow	Apollo,©TGN-SA	ownership-missing
202.171.171.1 -*- 63.217.25.225	ASC,Indigo-West,SeaMeWe-3	landing point > 1000km	Indigo-West
4.68.75.214 -*- 67.14.18.70	AJC	AJC	ownership-missing
193.251.248.35 - 81.52.166.55	SeaMeWe-3,-4,-5	SeaMeWe-3	SeaMeWe-3,-4,-5
114.31.201.29 - 4.69.218.150	AJC → AAG	LHL-unmapped	LHL-unmapped
63.220.195.67 - 193.251.240.71	AAG	landing point > 1000km	ownership-missing
102.130.69.5 -*- 213.248.100.20	WACS → TGN-west	LHL-unmapped	LHL-unmapped
168.209.1.166 - 168.209.100.214	WACS → TGN-west	LHL-unmapped	LHL-unmapped
41.84.12.109 -*- 63.218.230.70	WACS → terrestrial	LHL-unmapped	LHL-unmapped
81.52.179.79 - 193.251.133.35	SAT-3/WASC, ACE	landing point > 1000km	SAT-3/WASC, ACE
63.218.151.2 - 154.54.5.34	WACS → TGN-west	LHL-unmapped	LHL-unmapped
154.54.5.34 - 154.54.30.185	Apollo,Yellow,EXA-NS,FA-1	Apollo,Yellow,FA-1,©AC-1,©TGN-AS	ownership-missing
63.222.250.1 -*- 102.130.68.142	SAT-3/WASC	no mapping (bg_te)	SAT-3/WASC
63.222.250.1 -*- 63.220.65.45	WACS → terrestrial	LHL-unmapped	LHL-unmapped
196.37.155.169 - 168.209.129.144	WACS,SAT-3/WASC,ACE	SAT-3/WASC	ownership-missing

Unless stated otherwise, all Nautilus cable inferences listed in the table are categorized as `S,B` (definitely submarine & both geolocations good; see original paper [53] for definitions). These correspond to the label bg_oc (both-good, ocean) in their codebase [52] and represent the highest-confidence cable classification.

Table 4: Comparison of submarine cable mappings produced by `Calypso`, Nautilus [53], and Livadariu et al. [40] for representative IP links extracted from our traceroute ground-truth dataset. Each row lists an IP link and the corresponding cable(s) identified by each system. `landing point > 1000km` indicates cases where the nearest cable landing point is more than 1,000 km from the router's location, making Nautilus unable to identify a cable. Livadariu et al. rely on ownership cues derived from hostnames and prefixes, for which coverage is limited; we mark instances as `ownership-missing` where no cable can be inferred due to insufficient information. `©cable-name` denotes cable inferences that violate ownership or leasing rights as recorded in `Chartbook` from TeleGeography, while `◷cable-name` highlights cable inferences that violate SoL constraints of cable segment lengths. Finally, long-haul virtualized links requiring two or more cables to connect routers are marked as `LHL-unmapped`.