

SPLASH

Trabajo Profesional

ALUMNOS

Aguerre, Nicolás Federico

naguerre@fi.uba.ar

Klein, Santiago

sklein@fi.uba.ar

TUTORES

Dr. Beiró, Mariano Gastón

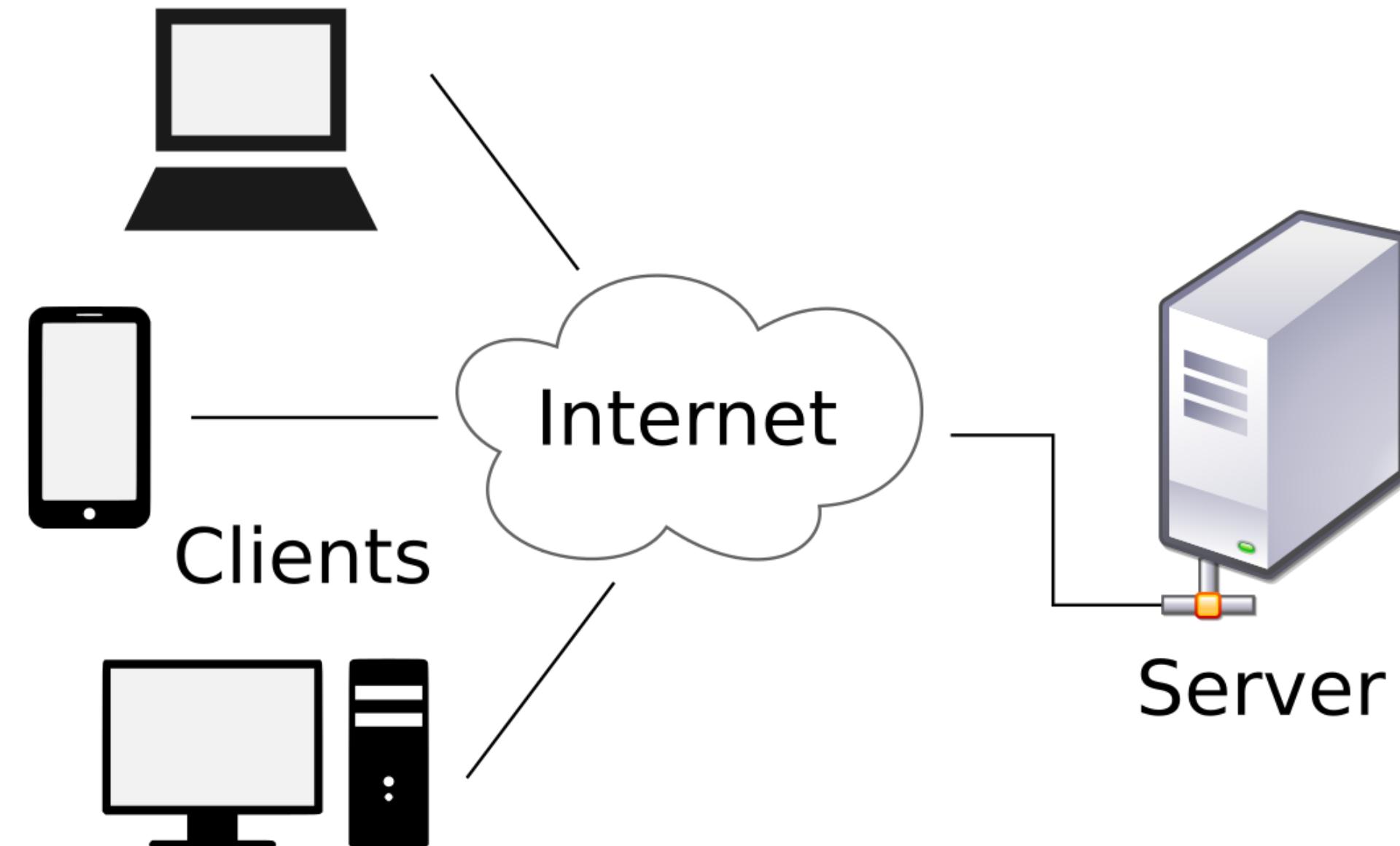
mbeiro@fi.uba.ar

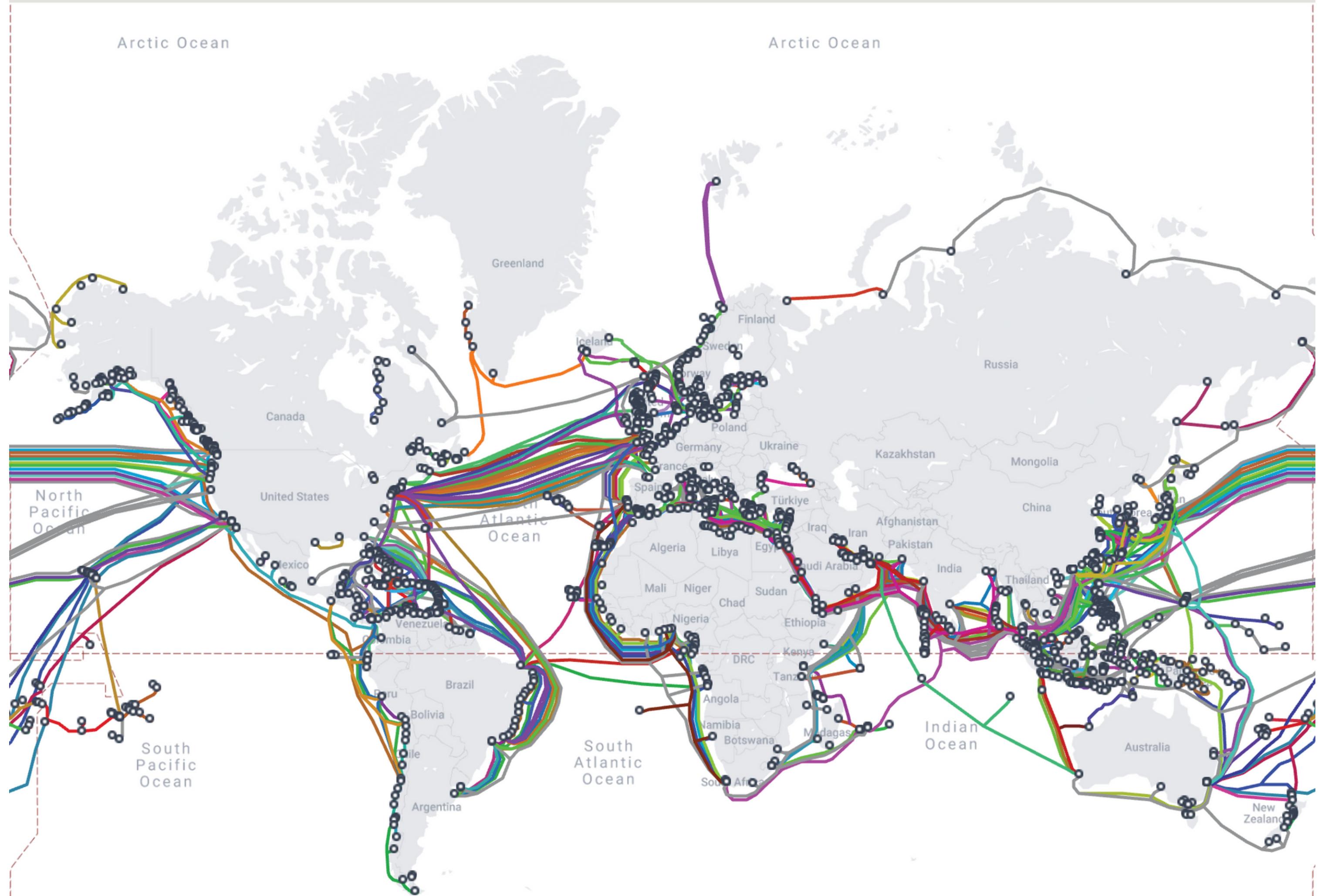
Dr. Carisimo, Esteban

ecarisimo@fi.uba.ar

1. Motivación

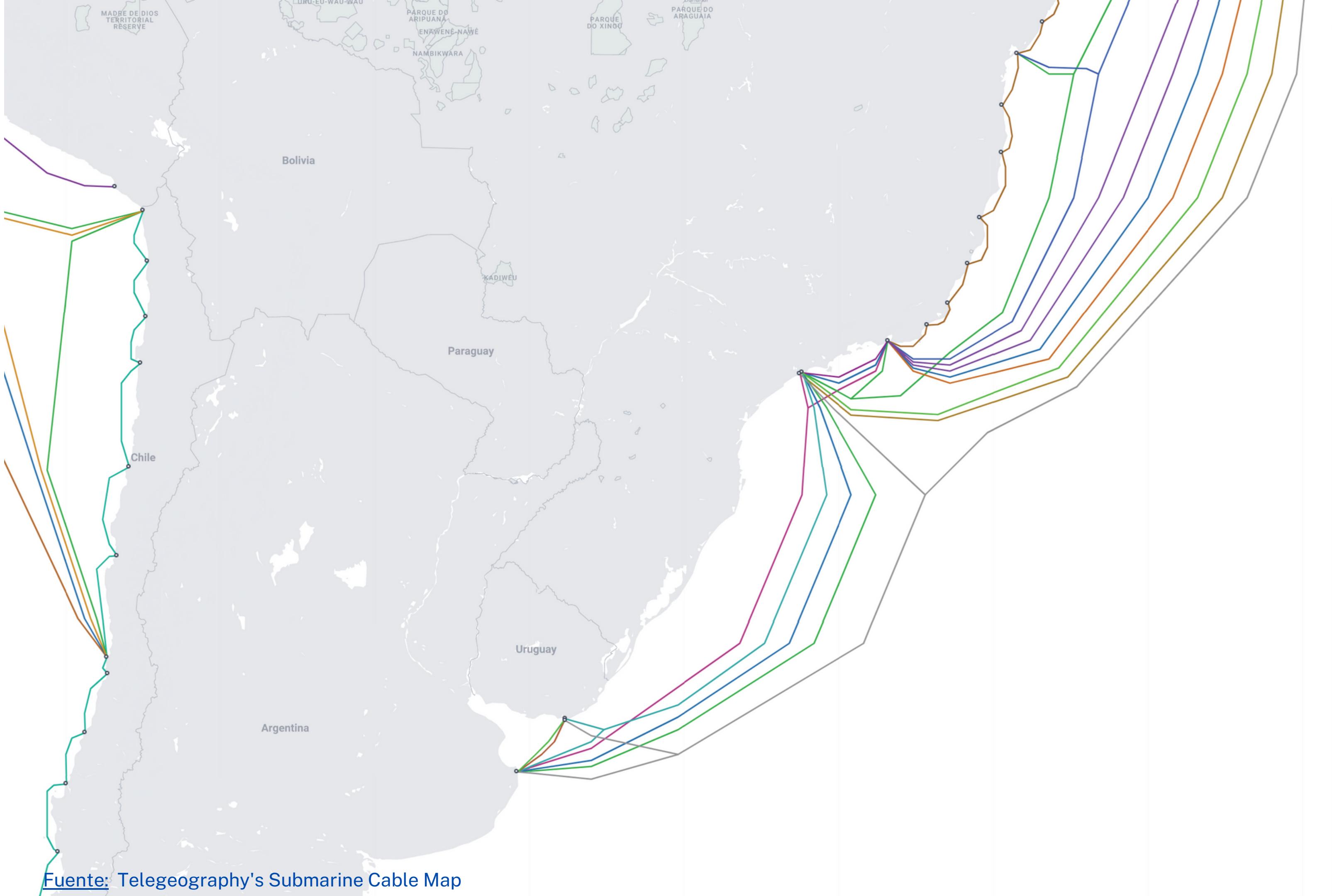
¿Qué es Internet?



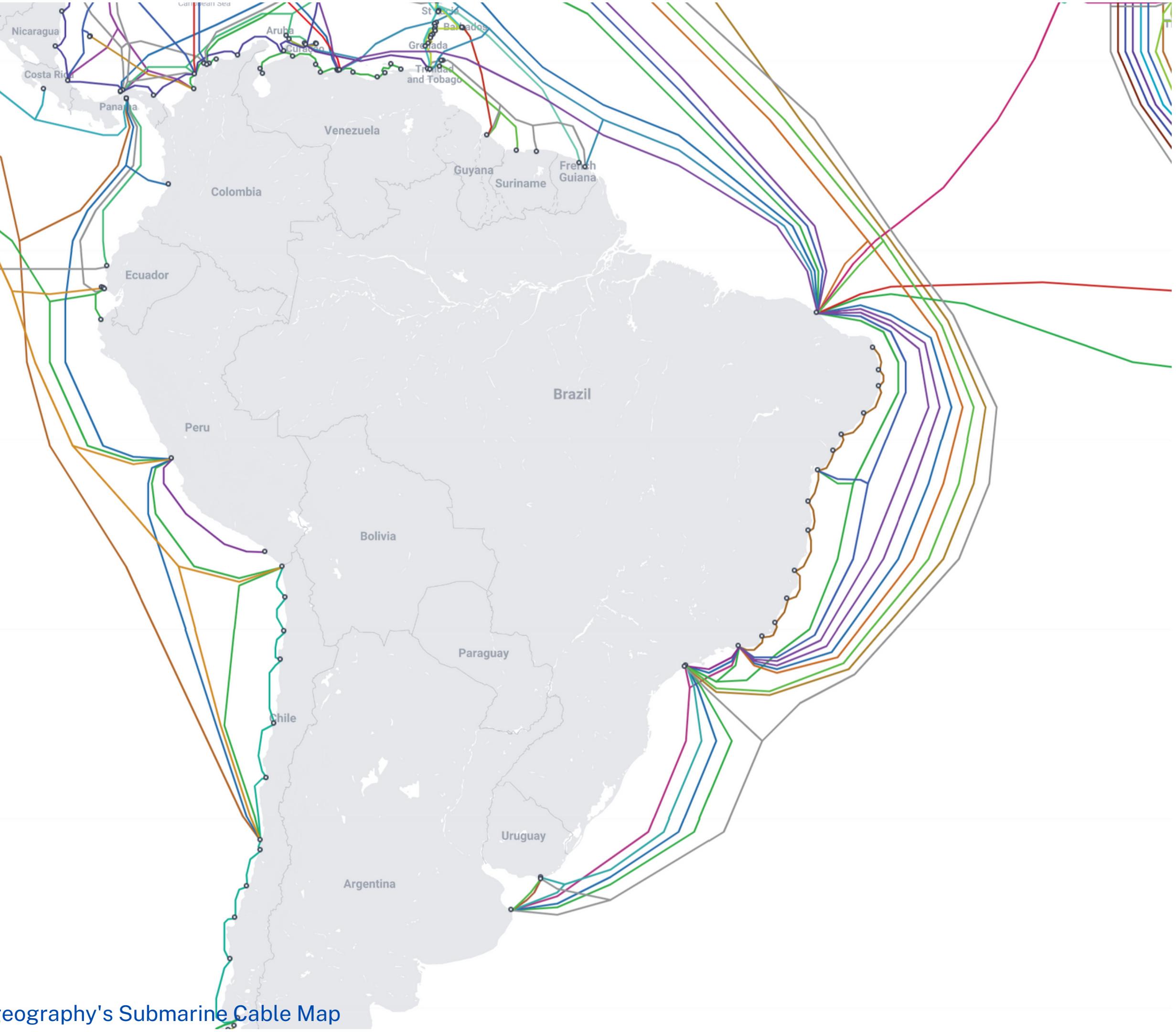


[Fuente: Telegeography's Submarine Cable Map](#)

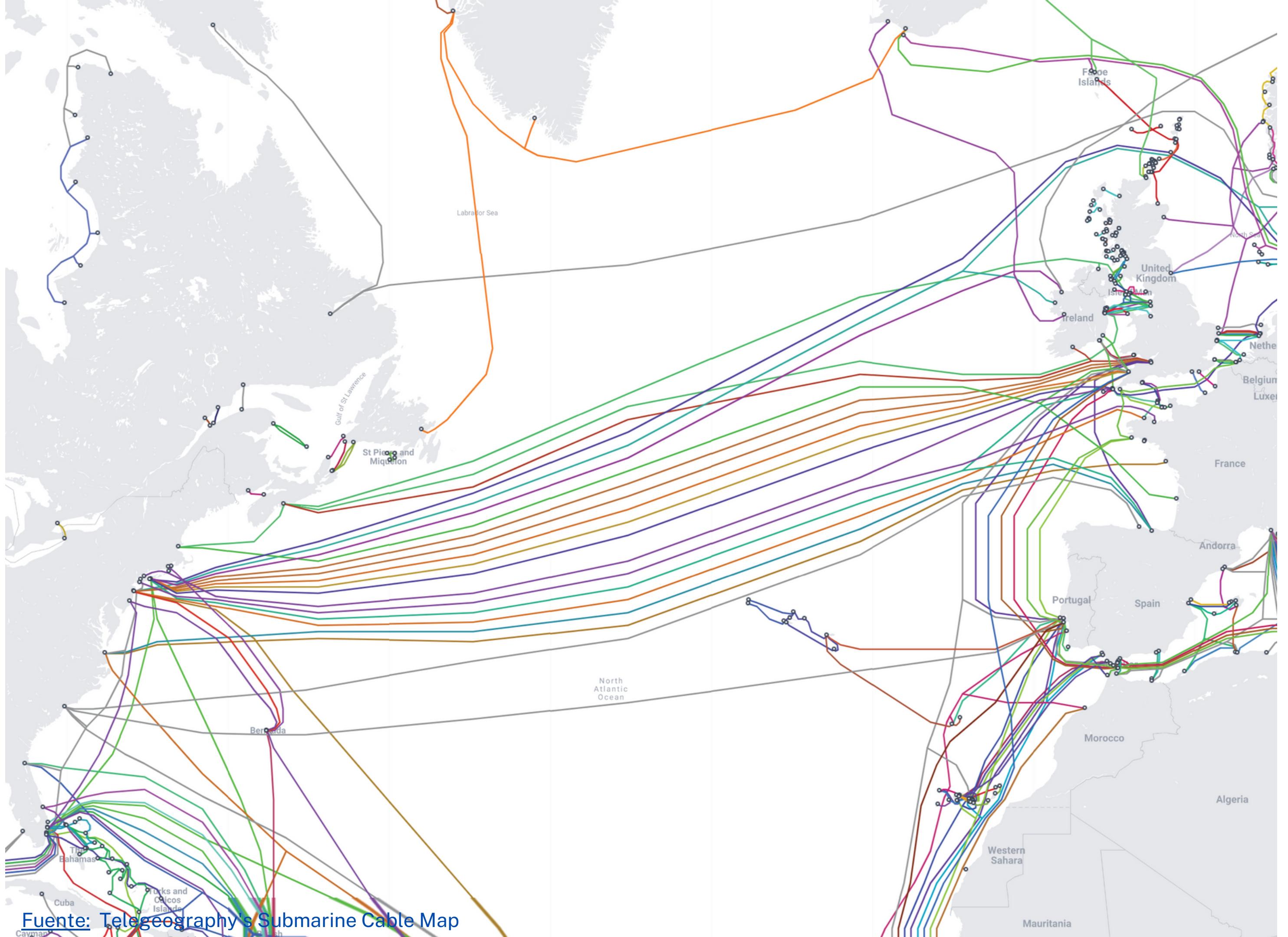
Souther



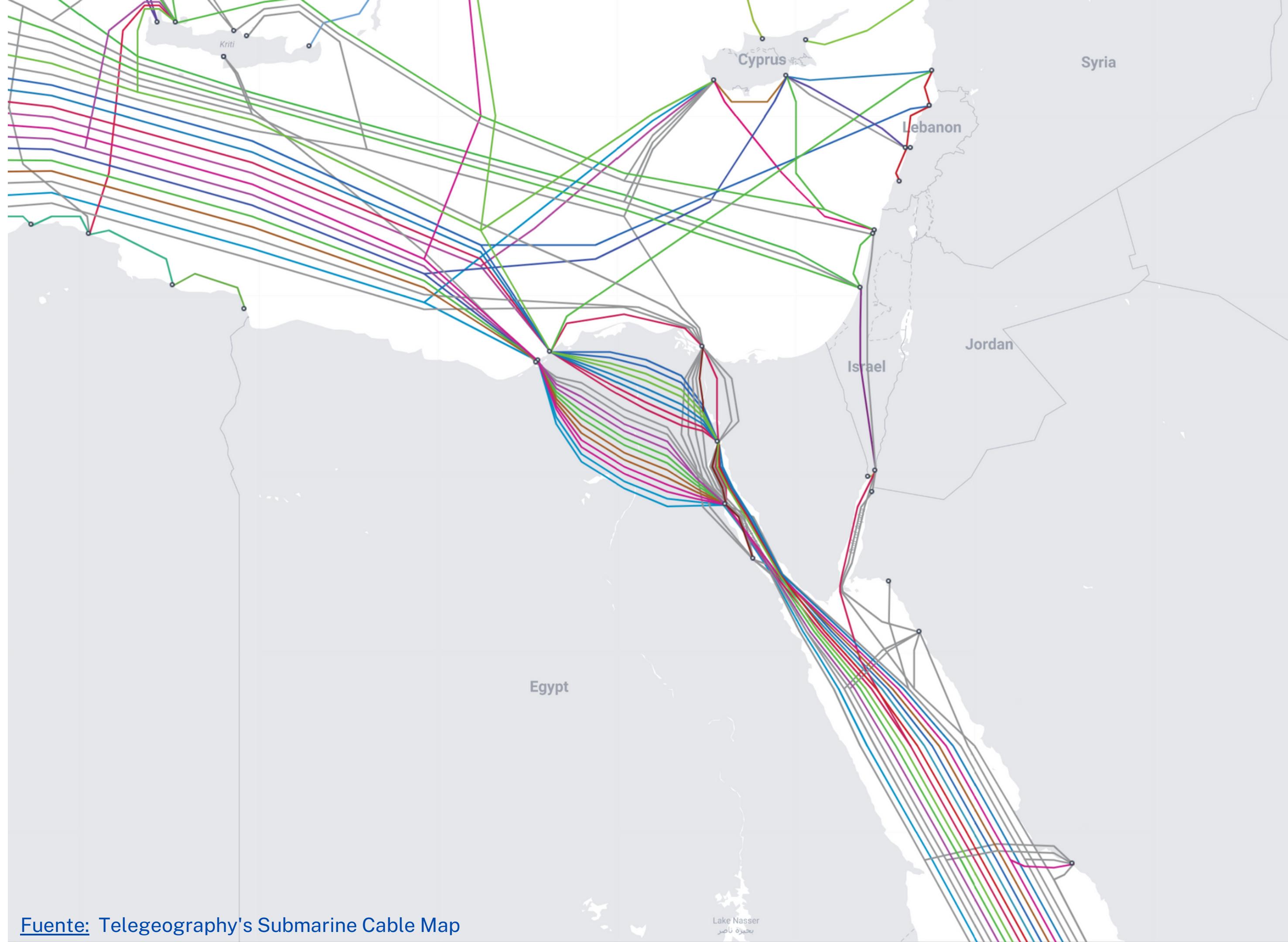
Fuente: Telegeography's Submarine Cable Map



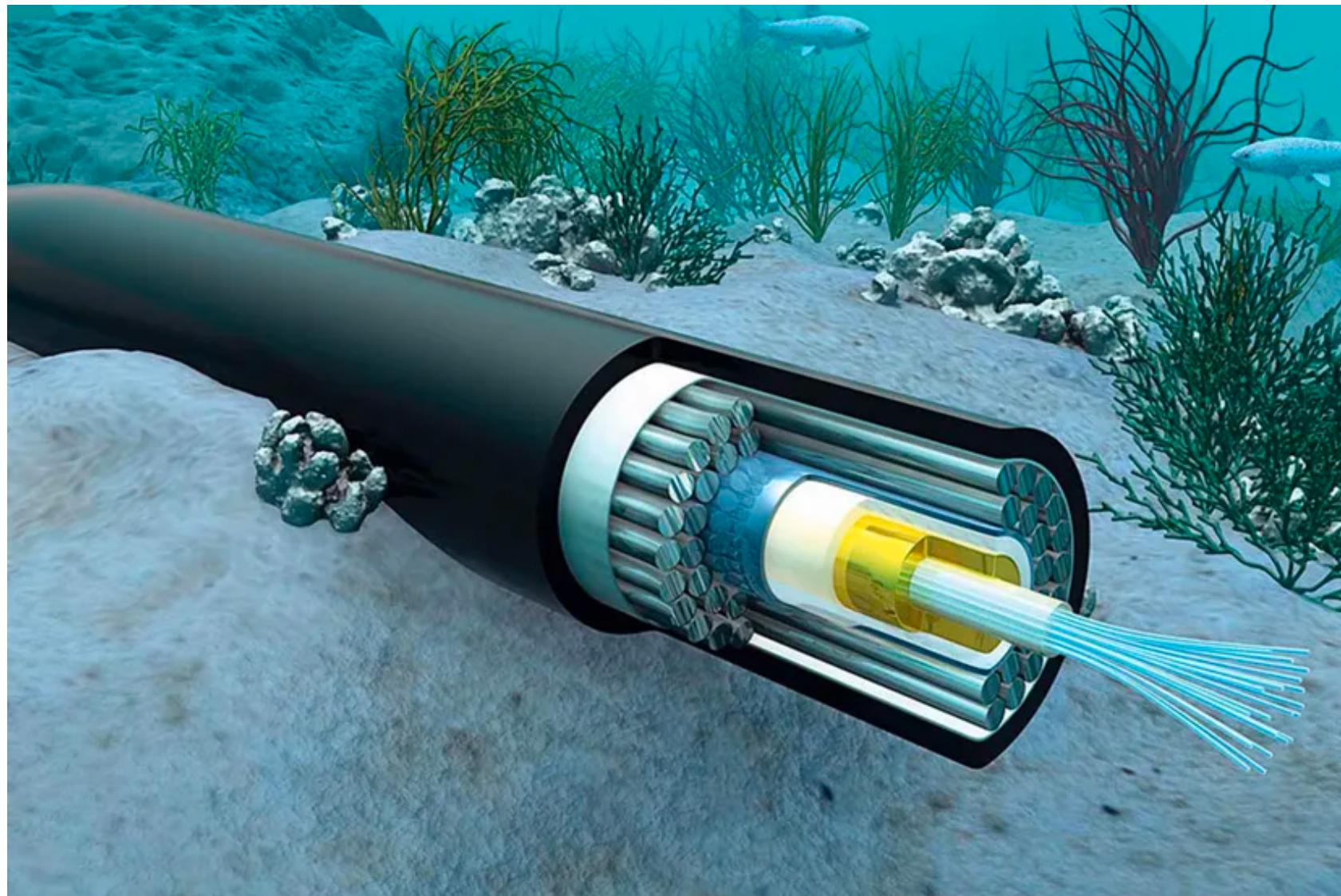
Fuente: Telegeography's Submarine Cable Map



Fuente: Telegeography's Submarine Cable Map



Fuente: Telegeography's Submarine Cable Map



Hard Failure

El servicio sufre una **interrupción**, dejando a los usuarios **desconectados** de internet.

Soft Failure

El servicio se **degrada**, ofreciendo una conectividad parcial o de **peor calidad**, pero no se interrumpe completamente.

¿Qué pasa si se rompe uno de estos cables?

The screenshot shows the homepage of Clarín. At the top, there is a navigation bar with the Clarín logo, a "Suscribite por \$30" button, and an "Ingresar" button. Below the navigation bar, there are links for "En vivo" (live), "Debate presidencial", "Todo sobre el Superclásico", and "Elecciones 2023". A sidebar on the left is labeled "Sociedad". The main headline reads: "Se rompió un cable submarino en Las Toninas y hay problemas con Internet en el país".

The screenshot shows the homepage of Infobae. At the top, there is a navigation bar with the Infobae logo and links for "Últimas Noticias", "Política", "Economía", "Dólar hoy", "Deportes", "Sociedad", "Policiales", and "Newsletters". Below the navigation bar, there is a category link "TECNO >". The main headline reads: "Una falla masiva de Internet afectó a millones de usuarios en Argentina". A descriptive text below the headline states: "El desperfecto generado por proveedores internacionales alcanzó a empresas de cable, Internet y telefonía celular. El servicio se normalizó pasadas las 20 horas".

¿Qué pasa si se rompe uno de estos cables?

The screenshot shows the Clarín website's header with navigation links for 'En vivo', 'Debate presidencial', 'Todo sobre el Superclásico', and 'Elecciones 2023'. Below the header, a section titled 'Sociedad' contains the main headline: 'Se rompió un cable submarino en Las Toninas y hay problemas con Internet en el país'.

The screenshot shows the Infobae website's header with categories like 'Últimas Noticias', 'Política', 'Economía', etc. Below the header, under the 'TECNO' category, is a prominent headline: 'Una falla masiva de Internet afectó a millones de usuarios en Argentina'. A descriptive paragraph follows: 'El desperfecto generado por proveedores internacionales alcanzó a empresas de cable, Internet y telefonía celular. El servicio se normalizó pasadas las 20 horas'.

The screenshot shows the Cloudflare Blog's header with sections for 'Product News', 'Speed & Reliability', 'Security', 'Serverless', 'Zero Trust', 'Developers', 'Deep Dive', and 'Life @Cloudflare'. It features a search bar and a 'Subscribe' button. The main article title is 'AAE-1 & SMW5 cable cuts impact millions of users across multiple countries', dated 06/10/2022, by David Belson.

¿Qué pasa si se rompe uno de estos cables?

The screenshot shows the Clarín website's header with navigation links for 'En vivo', 'Debate presidencial', 'Todo sobre el Superclásico', and 'Elecciones 2023'. Below the header, a section titled 'Sociedad' contains the headline: 'Se rompió un cable submarino en Las Toninas y hay problemas con Internet en el país'.

The screenshot shows the Infobae website's header with categories like 'Últimas Noticias', 'Política', 'Economía', etc. Below the header, under the 'TECNO' category, is the headline: 'Una falla masiva de Internet afectó a millones de usuarios en Argentina'. A descriptive text below the headline states: 'El desperfecto generado por proveedores internacionales alcanzó a empresas de cable, Internet y telefonía celular. El servicio se normalizó pasadas las 20 horas'.

The screenshot shows a Cloudflare blog post by David Belson. The title is 'AAE-1 & SMW5 cable cuts impact millions of users across multiple countries'. The post was published on 06/10/2022. The author's profile picture and name, 'David Belson', are visible at the bottom.

The screenshot shows a Cloudflare blog post by João Tomé. The title is 'Internet is back in Tonga after 38 days of outage'. The post was published on 02/22/2022. The author's profile picture and name, 'João Tomé', are visible at the bottom.

¿Qué pasa si se rompe uno de estos cables?

The screenshot shows the Clarín website's header with navigation links for 'En vivo', 'Debate presidencial', 'Todo sobre el Superclásico', and 'Elecciones 2023'. Below the header, a section titled 'Sociedad' contains the main headline: 'Se rompió un cable submarino en Las Toninas y hay problemas con Internet en el país'.

The Infobae website's header includes categories like 'Últimas Noticias', 'Política', 'Economía', 'Dólar hoy', 'Deportes', 'Sociedad', 'Policiales', and 'Newsletter'. The main article under the 'TECNO' category is titled 'Una falla masiva de Internet afectó a millones de usuarios en Argentina', with a subtitle explaining it affected cable, Internet, and cellular phone services.

The Cloudflare blog post is titled 'AAE-1 & SMW5 cable cuts impact millions of users across multiple countries'. It was published on 06/10/2022 by David Belson. The post discusses how two major submarine cables were cut, leading to widespread connectivity issues.

The Cloudflare blog post is titled 'Internet is back in Tonga after 38 days of outage'. It was published on 02/22/2022 by João Tomé. The post reports that Internet service has been restored to Tonga after nearly a month of being down due to a volcanic eruption.

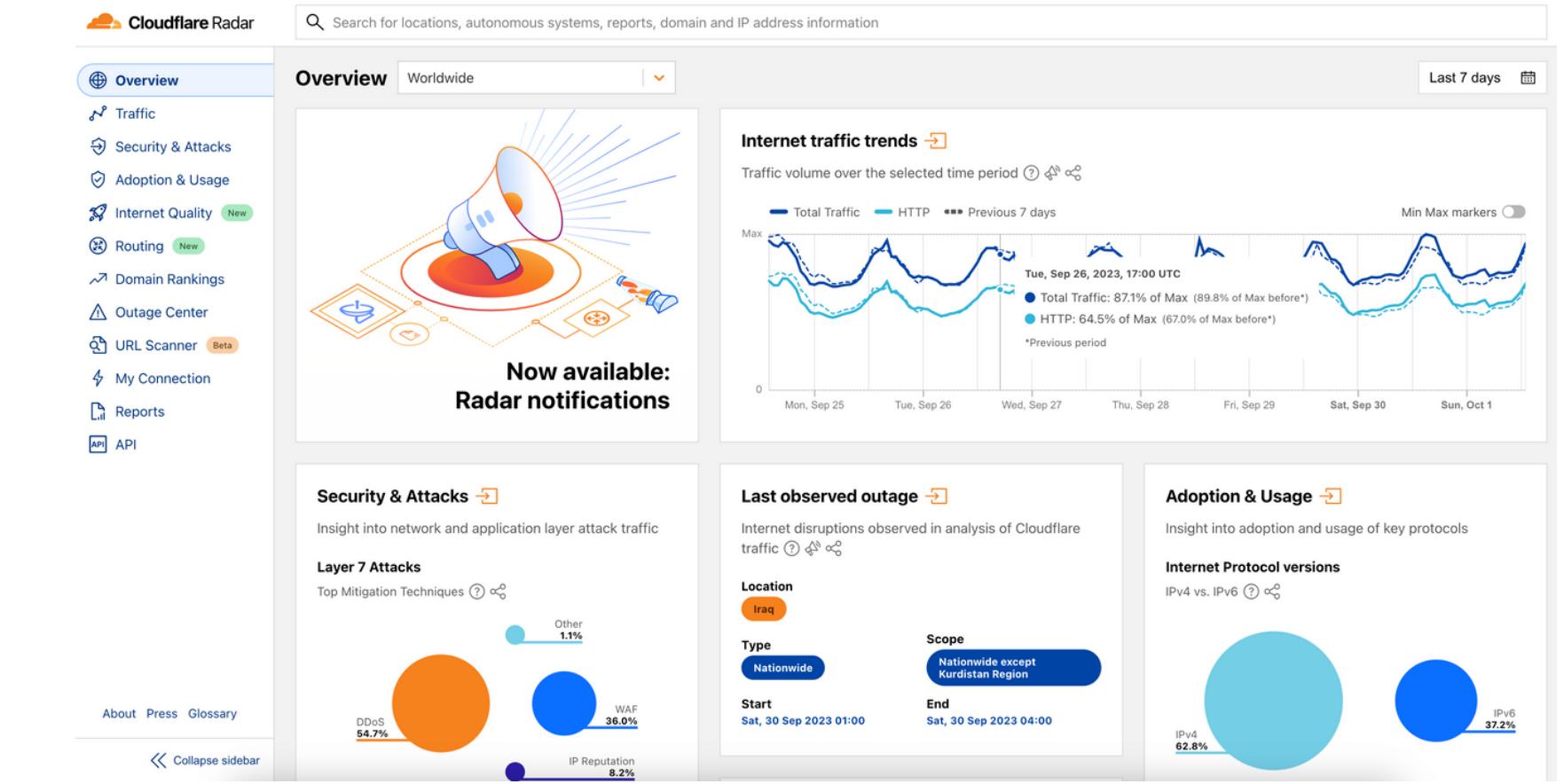
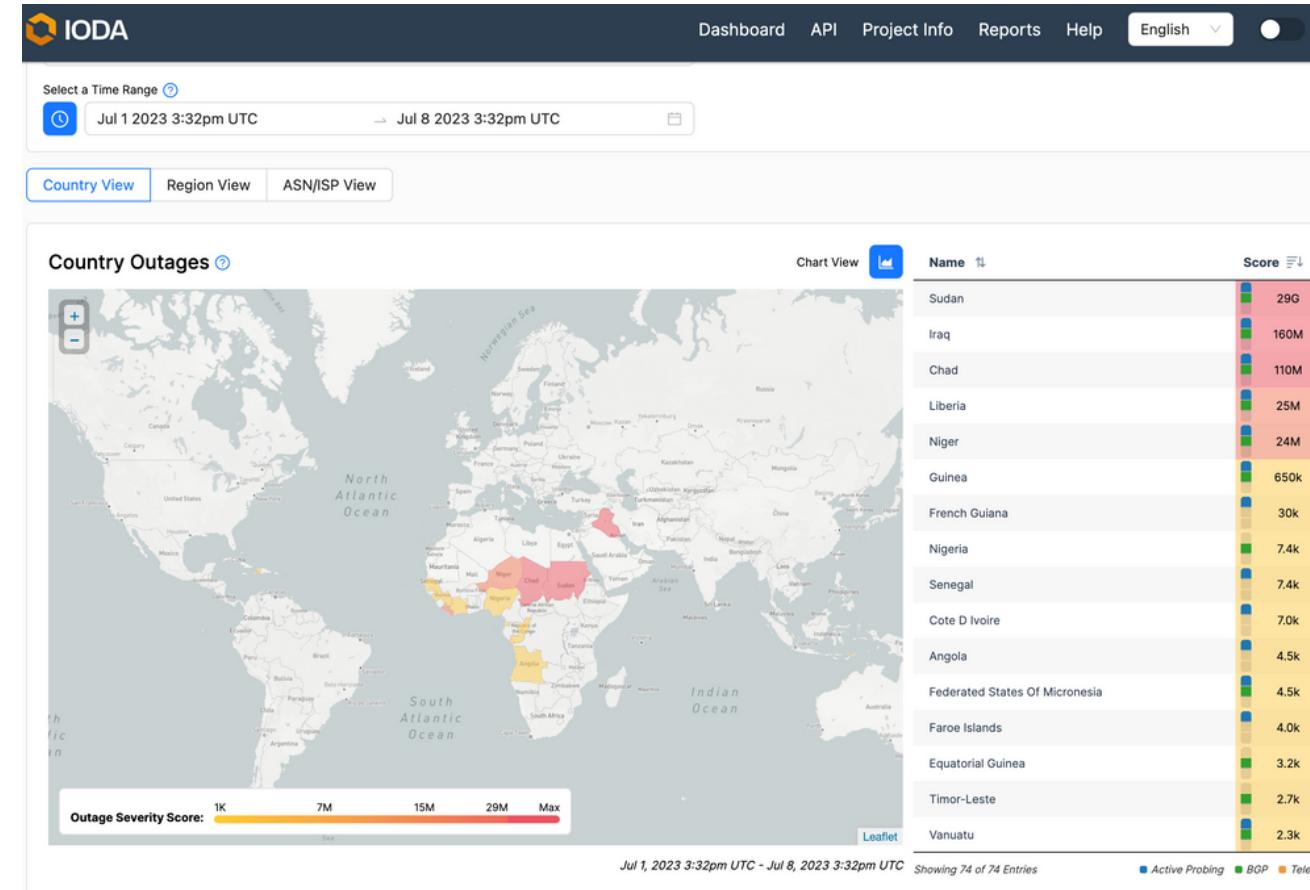
The Vietnam+ news article is titled 'Internet slows as four out of five submarine cables suffer disruptions in unprecedented incident'. It was published on Sunday, October 1, 2023. The article discusses how four out of the five submarine cables connecting Southeast Asia to the rest of the world were disrupted, causing significant slowdowns in internet speeds.



Preguntas de investigación que motivaron este trabajo

1. ¿Cómo se ve impactado el tráfico de red cuando ocurre una soft failure?
2. ¿Es posible distinguir fallas en cables submarinos de otros tipos de fallas? ¿Tienen una impronta propia?
3. ¿Qué impacto tiene una falla en un cable submarino en los países que lo usan?
¿Cómo varía?
4. ¿Cuán frecuentemente ocurren?
5. ¿Cuáles son los países más afectados por ellas?

Estado del arte



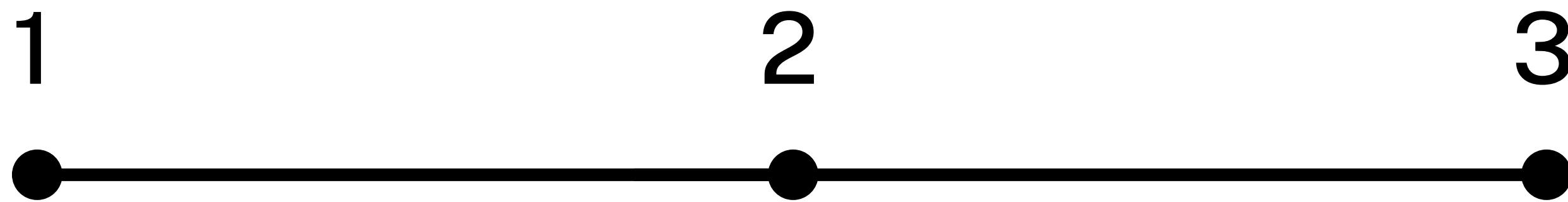
Internet Outage Detection and Analysis, una plataforma de detección de hard failures desarrollada por Georgia Institute of Technology



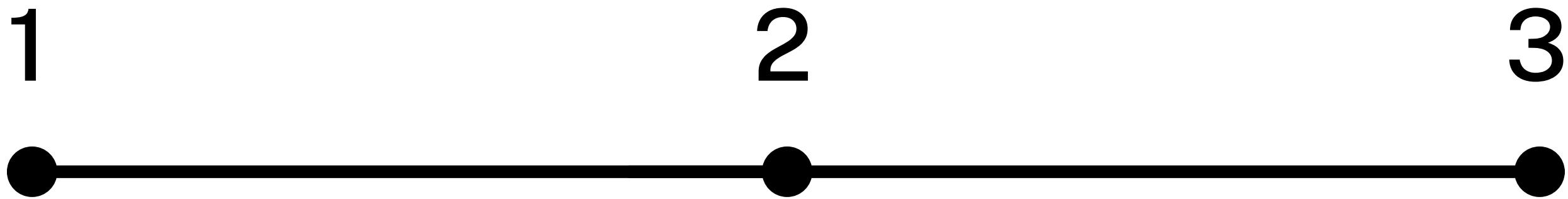
CloudFlare Radar

CloudFlare Radar, una plataforma de monitoreo de tráfico y ataques de internet, desarrollada y mantenida por CloudFlare

Objetivo del Trabajo



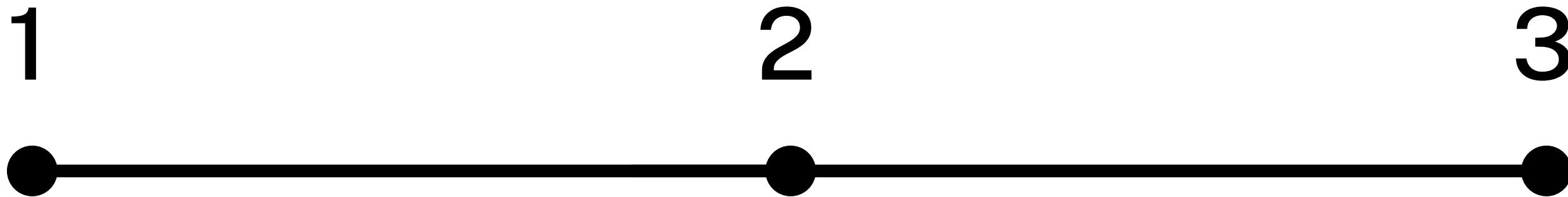
Objetivo del Trabajo



Dataset de Fallas

Recopilación manual y
clasificación de fallas

Objetivo del Trabajo



Dataset de Fallas

Recopilación manual y
clasificación de fallas

Dashboard de Visualización

Implementación sistemática de
visualizaciones de métricas para
todas las fallas de la lista

Objetivo del Trabajo

1

2

3

Dataset de Fallas

Recopilación manual y
clasificación de fallas

Dashboard de Visualización

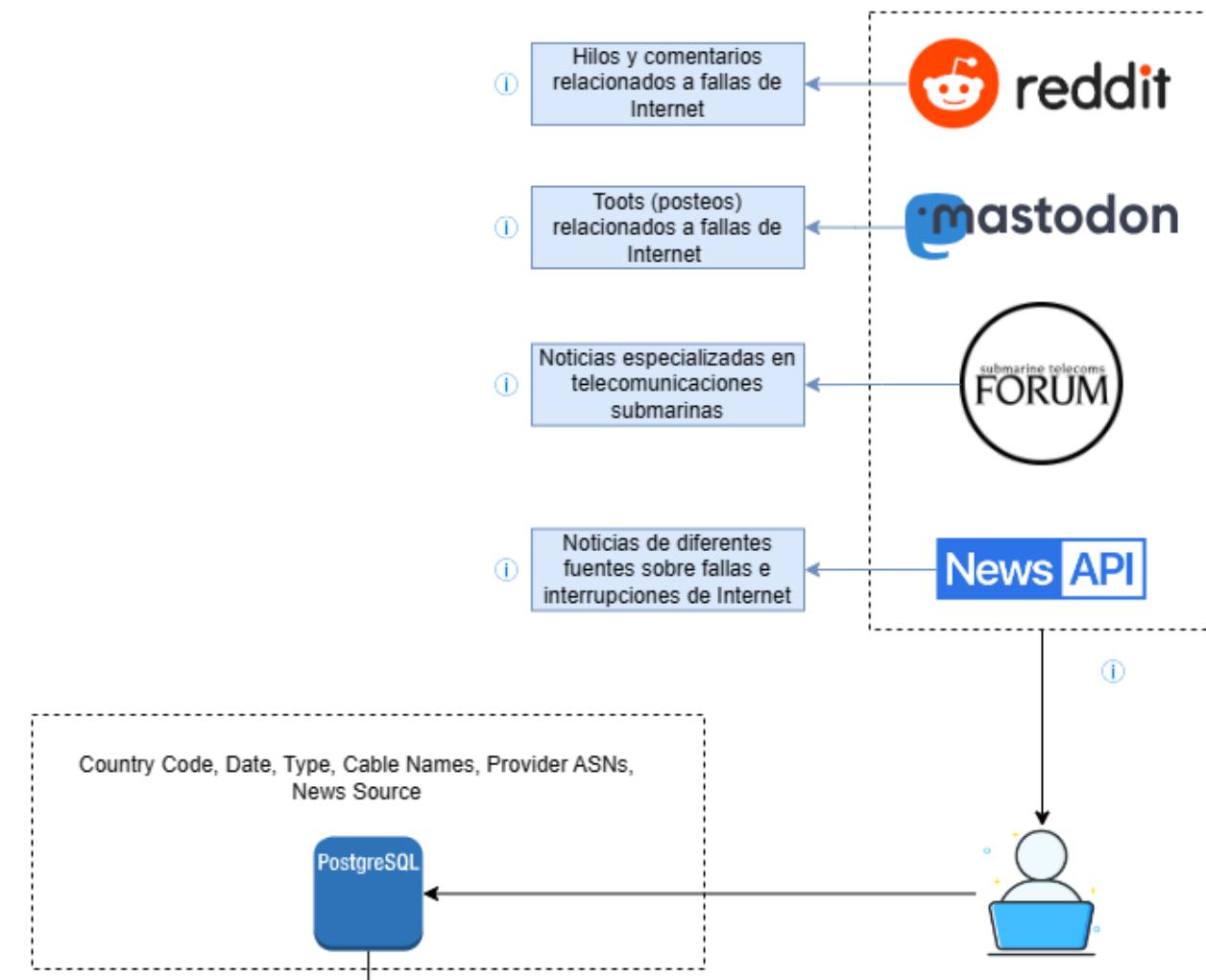
Implementación sistemática de
visualizaciones de métricas para
todas las fallas de la lista

Ingestor de Fallas Nuevas

Automatización del proceso de
agregado de fallas nuevas al
dataset y al dashboard de
visualizaciones

2. Diseño del Sistema

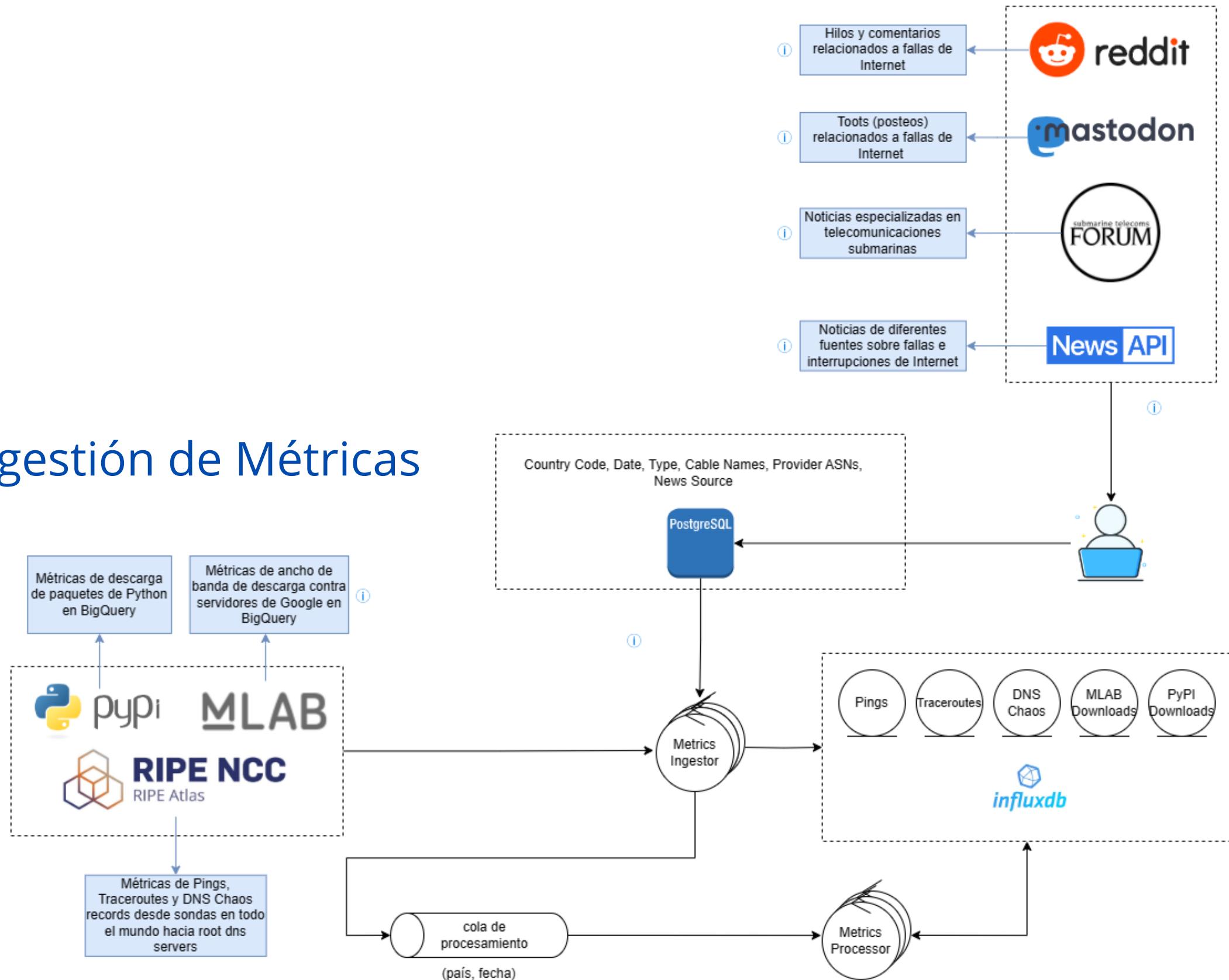
1 Recopilación de Fallas



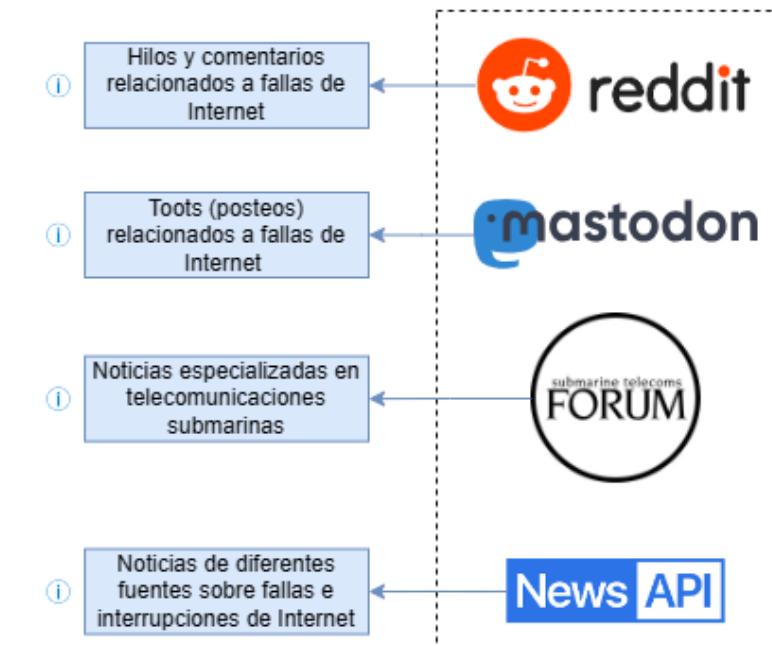
Arquitectura Básica

1 Recopilación de Fallas

2 Ingestión de Métricas

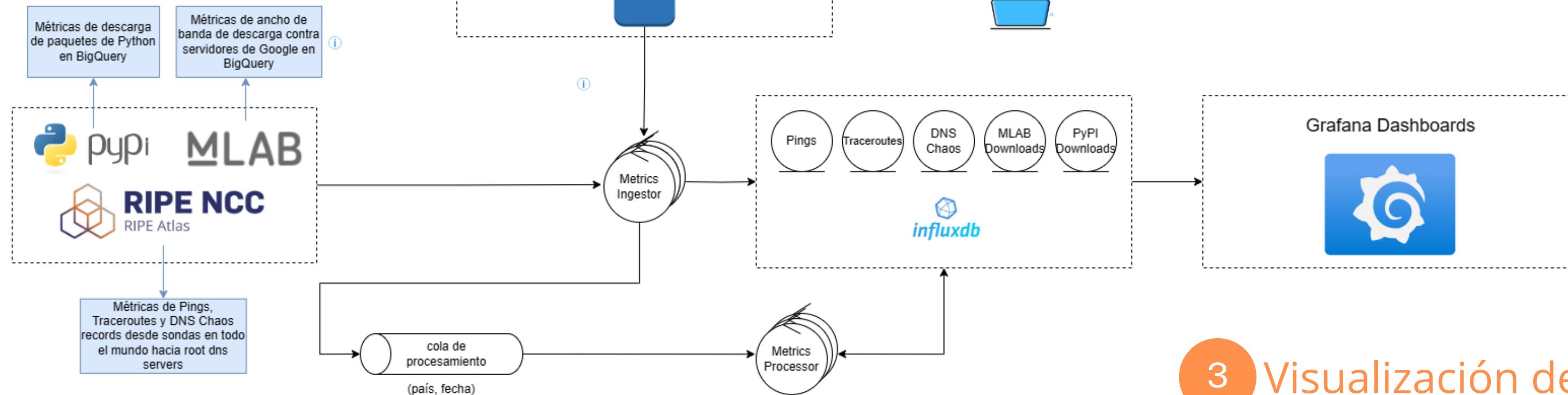


Arquitectura Básica



1 Recopilación de Fallas

2 Ingestión de Métricas



3 Visualización de Métricas

2.1 Recopilación de Fallas

Fuentes de Información de Fallas



Reddit

Fuentes de Información de Fallas



Reddit



Subtel Forum

Fuentes de Información de Fallas



Reddit



Subtel Forum



Mastodon

Fuentes de Información de Fallas



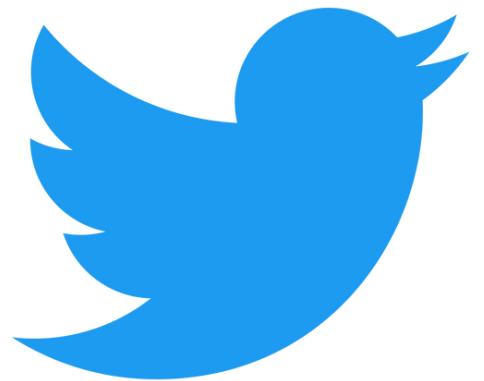
Reddit



Subtel Forum



Mastodon



Twitter

Fuentes de Información de Fallas



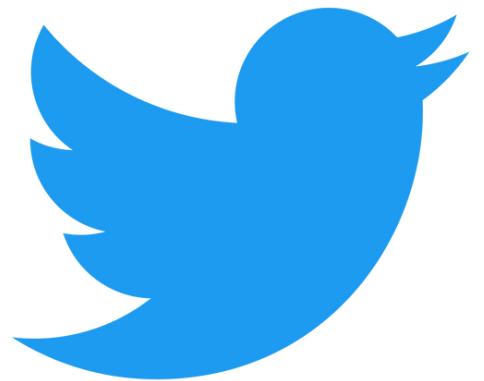
Reddit



Subtel Forum



Mastodon



Twitter



Google

Fuentes de Información de Fallas



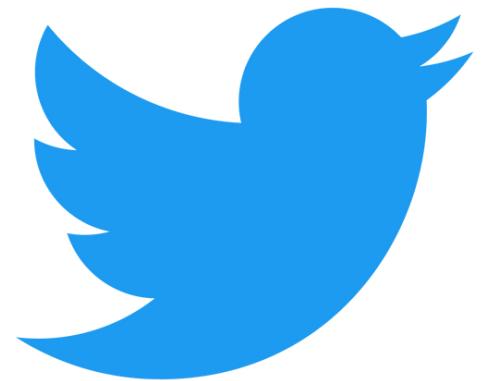
Reddit



Subtel Forum



Mastodon



Twitter



Google



NewsAPI

Información recopilada de cada falla

Campo

- País
- Fecha
- Tipo de falla
- Cables afectados
- Proveedores afectados
- Visibilidad en otras plataformas
- Fuente

Ejemplo

Argentina
2017-08-02
Soft-Submarine
ATLANTIS-2
LEVEL3
No
Portal de noticias

Fallas Recopiladas

84

Fallas Soft

54

Fallas Hard

32

Paises Afectados

41

Período Comprendido

Septiembre de 2014 a
Abril de 2023

2.2 Ingestión de Métricas

Fuentes de Métricas



RIPE Atlas

Una plataforma abierta, global y distribuida de medición de Internet, compuesta por miles de dispositivos de medición distribuidos en múltiples países.



MLab

Un consorcio de investigadores, empresas y socios de interés público que proporciona un ecosistema para la medición abierta y confiable del rendimiento global de la red.

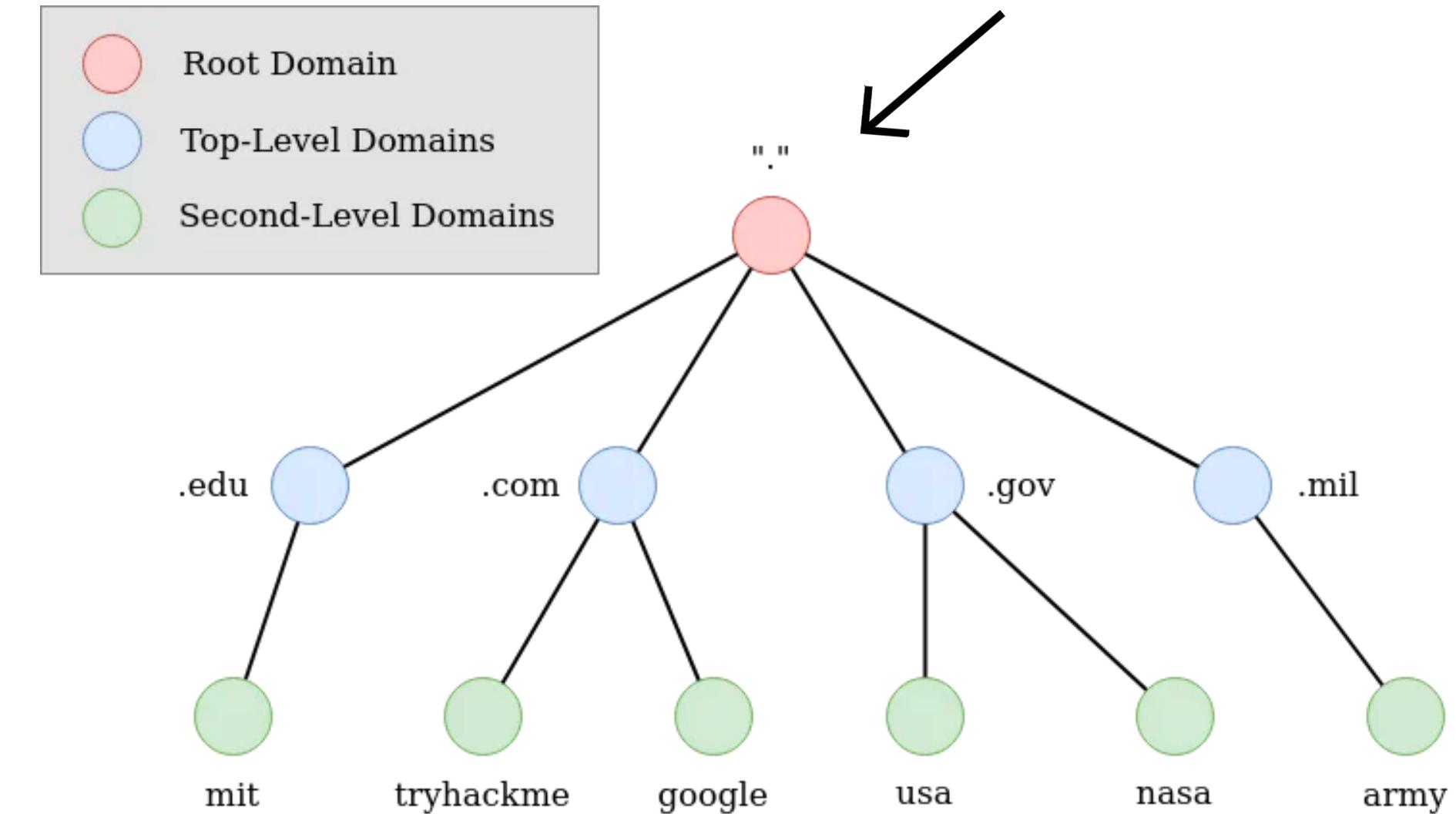
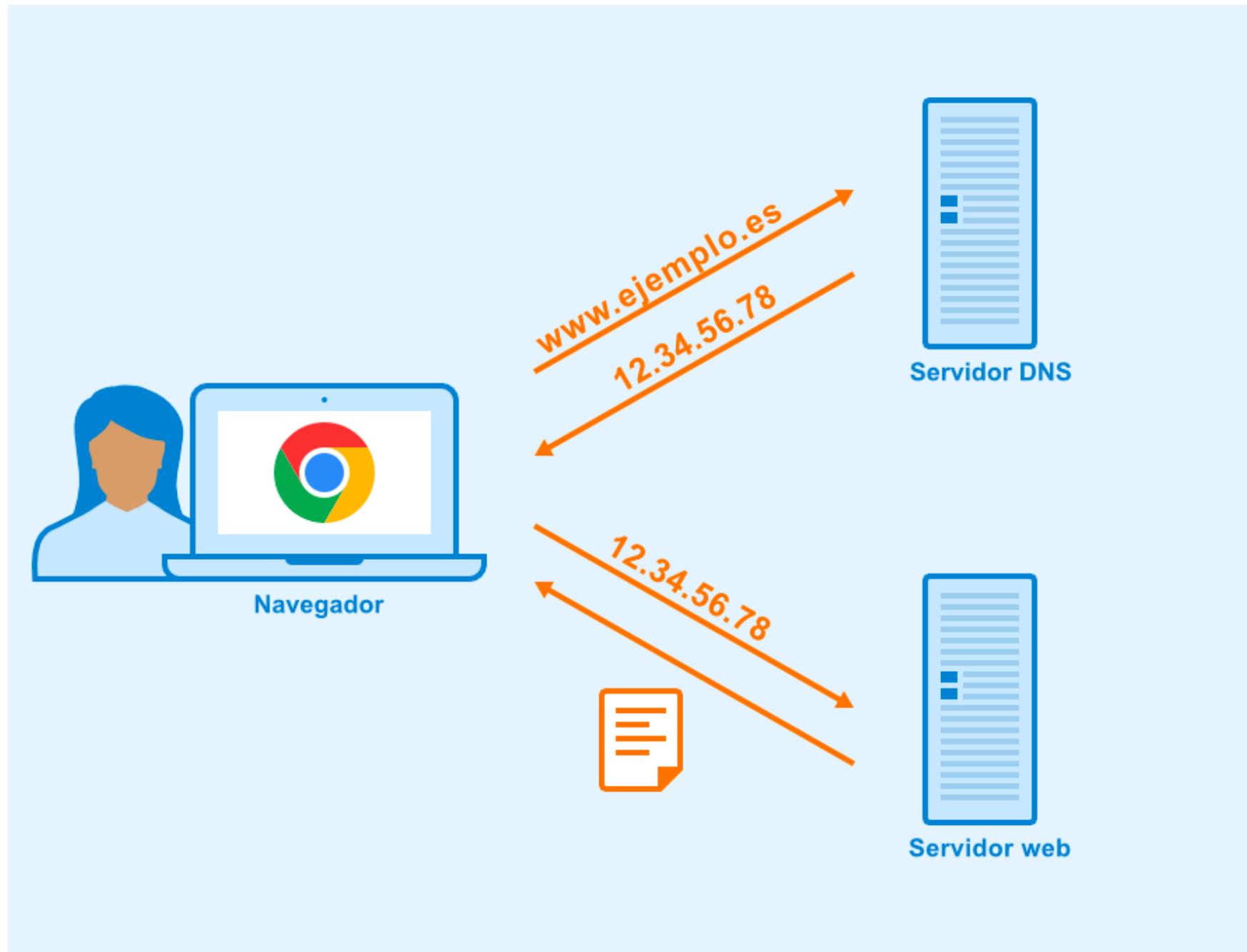


PyPI Downloads

El repositorio estándar de paquetes de Python, que ofrece públicamente sus estadísticas de descargas a través de BigQuery.

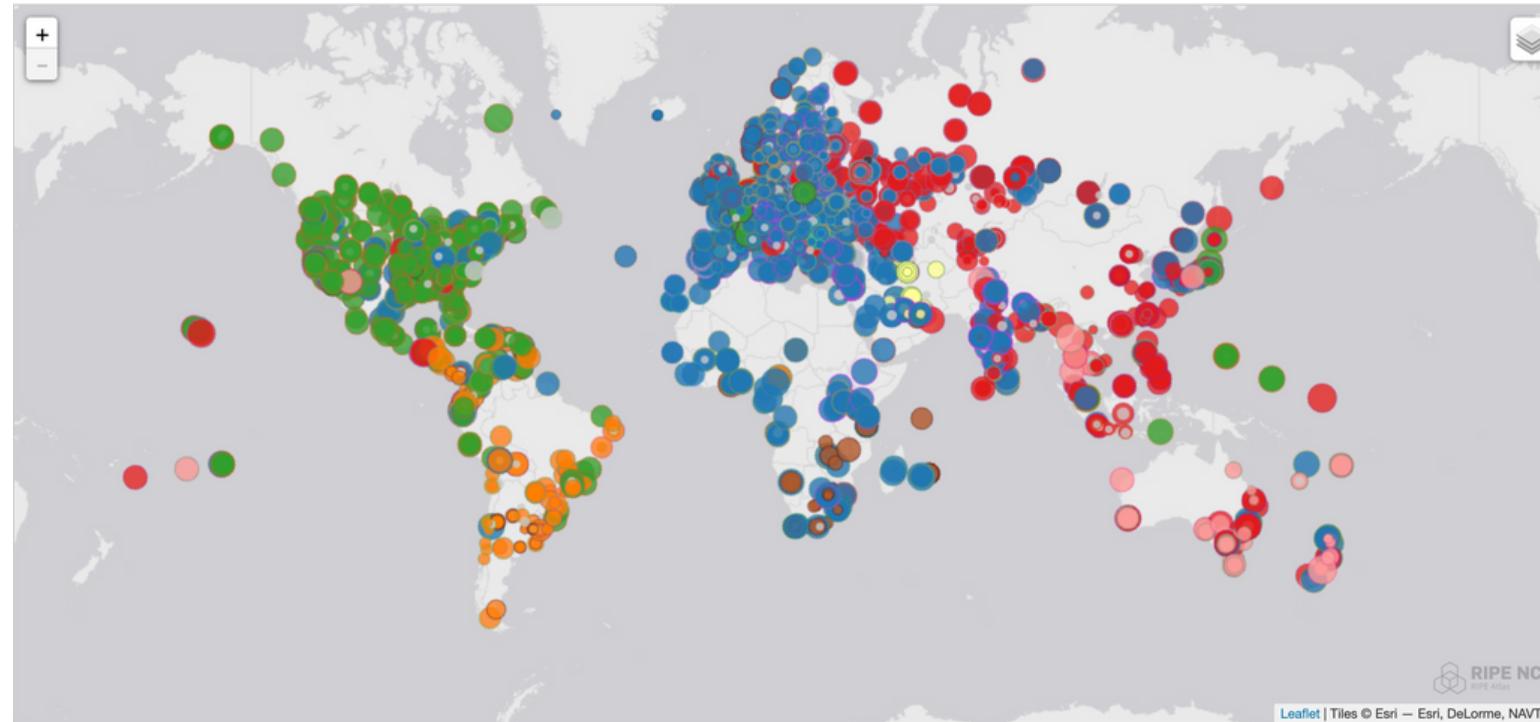
2.2.1 RIPE Atlas

Domain Name System (DNS)



Crítico para el funcionamiento de Internet

Distribución de Root DNS Servers y Sondas



Instancias de Root DNS servers

~1.800 instancias de 12 operadores independientes



Sondas de RIPE Atlas

~10.000 nodos mantenidos por investigadores y voluntarios

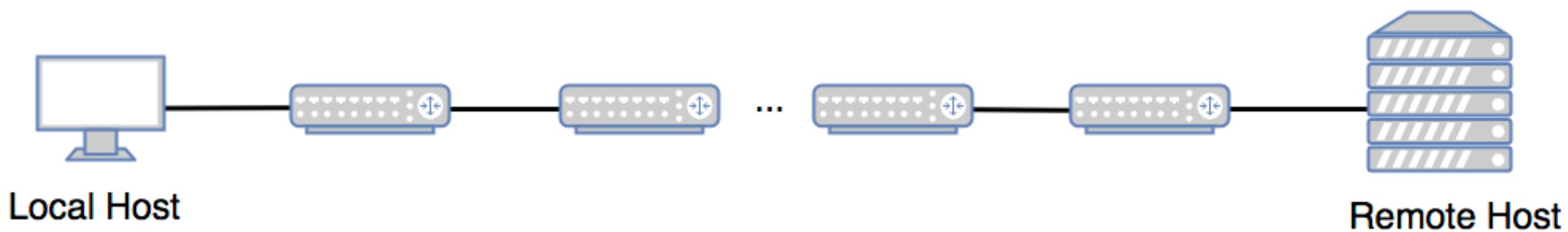


Fuente: RIPE Atlas

1 Pings

- Es una prueba de conectividad enviada desde un dispositivo a otro a través de internet. En el caso de RIPE Atlas, enviado desde una sonda a un Root DNS Server.

Realizado cada **240** segundos.



1 Pings

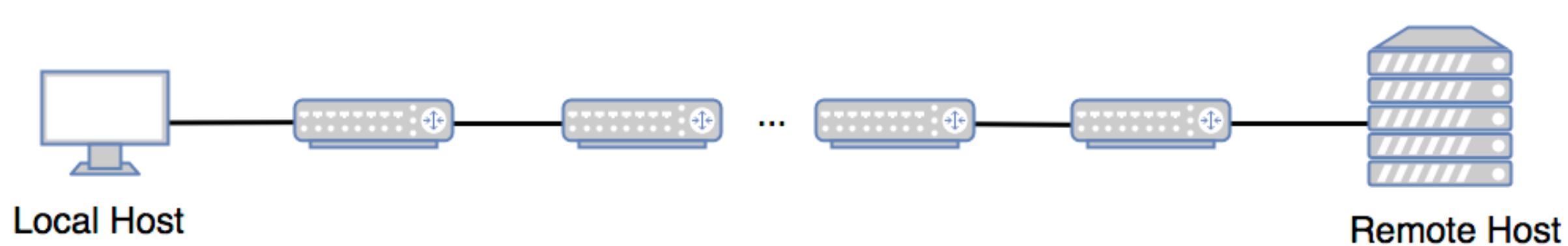
- Es una prueba de conectividad enviada desde un dispositivo a otro a través de internet. En el caso de RIPE Atlas, enviado desde una sonda a un Root DNS Server.

Realizado cada **240** segundos.

2 Traceroutes

- Traceroute es una herramienta que rastrea y muestra la ruta que los paquetes de datos toman desde una computadora hasta un destino, revelando los saltos intermedios y los tiempos de respuesta de cada uno.

Realizado cada **1800** segundos.



1 Pings

- Es una prueba de conectividad enviada desde un dispositivo a otro a través de internet. En el caso de RIPE Atlas, enviado desde una sonda a un Root DNS Server.

Realizado cada **240** segundos.



Local Host

2 Traceroutes

- Traceroute es una herramienta que rastrea y muestra la ruta que los paquetes de datos toman desde una computadora hasta un destino, revelando los saltos intermedios y los tiempos de respuesta de cada uno.

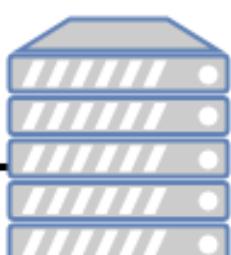
Realizado cada **1800** segundos.



3 Chaos DNS

- Es similar a Ping, con el adicional de que la instancia de Root DNS server incluye un campo identificatorio en su respuesta, el cual nos permite deducir su ubicación.

Realizado cada **1800** segundos.



Remote Host

1

Pings

3 mediciones de RTT (Round Trip Time)

tiempo de respuesta de ida y vuelta del paquete

Pérdida de paquetes (PLR)

cuántos paquetes no retornan de los enviados

```
{  
    "result": [  
        {  
            "rtt": 164.083626  
        },  
        {  
            "rtt": 164.111375  
        },  
        {  
            "rtt": 166.205875  
        }  
    ],  
    "rcvd": 3,  
    "sent": 3,  
    "msm_id": 1009,  
    "prb_id": 55438,  
    "timestamp": 1628208042,  
    "msm_name": "Ping"  
}
```

2 Traceroutes

Hops que realizan los paquetes

*routers que se encuentran en el camino
desde la sonda hasta el Root DNS Server.*

*Algunos sistemas autónomos o proveedores ocultan
esta información

```
{  
    "result": [  
        {  
            "hop": 1,  
            "result": [  
                {"from": "192.168.1.1", "ttl": 64, "size": 68, "rtt": 0.853},  
                {"from": "192.168.1.1", "ttl": 64, "size": 68, "rtt": 0.799},  
                {"from": "192.168.1.1", "ttl": 64, "size": 68, "rtt": 0.741},  
            ],  
        },  
        {  
            "hop": 2,  
            "result": [  
                {"from": "45.171.36.241", "ttl": 63, "size": 68, "rtt": 2.52},  
                {"from": "45.171.36.241", "ttl": 63, "size": 68, "rtt": 1.825},  
                {"from": "45.171.36.241", "ttl": 63, "size": 68, "rtt": 2.969},  
            ],  
        },  
        {  
            "hop": 3,  
            "result": [  
                {"from": "172.16.80.1", "ttl": 62, "size": 68, "rtt": 2.963},  
                {"from": "172.16.80.1", "ttl": 62, "size": 68, "rtt": 2.709},  
                {"from": "172.16.80.1", "ttl": 62, "size": 68, "rtt": 3.768},  
            ],  
        },  
    ],  
    "msm_id": 1009,  
    "prb_id": 55438,  
    "timestamp": 1628208263,  
    "msm_name": "Traceroute",  
}
```

3 CHAOS

Código IATA o UN/LOCODE

*información sobre dónde se encuentra el
Root DNS Server*

**No está completamente estandarizado el formato de
respuesta*

```
{  
  "result": {  
    "rt": 164.743,  
    "answers": [  
      {  
        "TYPE": "TXT",  
        "NAME": "id.server",  
        "RDATA": [  
          "EZE.cf.f.root-servers.org"  
        ]  
      }  
    ]  
  },  
  "msm_id": 10404,  
  "prb_id": 55438,  
  "timestamp": 1628208604,  
  "msm_name": "Tdig"  
}
```

Jitter

Permite entender cómo varía la latencia respecto del tiempo

Se calcula la diferencia entre mediciones de latencia (Round Trip Time) consecutivas.

Grafos de ASN

Permite entender qué Sistemas Autónomos atraviesan las mediciones de traceroutes, y cómo varían entre mediciones consecutivas

Se toman ventanas de 1h de tiempo para las cuales se realiza un mapeo de IP a ASN utilizando [bdrmapIT](#) las mediciones de traceroutes. A partir de ese mapeo, se confecciona un grafo con nodos representando a los números de los Sistemas Autónomos, y aristas representando los enlaces dirigidos entre ellos.

Distancias de traceroute a nivel IP

Permite entender cuánto variaron los caminos que realizan los paquetes entre las sondas y los Root DNS Server en un período de tiempo.

Se toman ventanas de 12hs de tiempo para las cuales se definen matrices de transición empíricas utilizando todos los traceroutes disponibles, y luego se calcula la distancia entre matrices correspondientes a ventanas de tiempo consecutivas, siendo dicha distancia la métrica plasmada en el dashboard.

2.2.2 M Lab

¿Qué guardamos? MLab

- MLAB permite realizar mediciones de ancho de banda
- Resultados públicos y se almacenan en BigQuery
- Obtenemos métricas de ancho de banda de descarga y pérdida de paquetes, así como la ubicación del servidor

```
{  
  "mean_download_mbps": "1.5327591956750981",  
  "min_rtt": "0.04850099999999996",  
  "loss_rate": "0.0",  
  "ts": "2023-09-29 23:04:01.288123 UTC",  
  "client_ip": "190.225.213.12",  
  "client_city": "Tafi Viejo",  
  "client_latitude": "-26.8602",  
  "client_longitude": "-65.2408",  
  "client_cidr": "190.224.0.0/15",  
  "client_asn_id": "7303",  
  "client_asn_name": "Telecom Argentina S.A.",  
  "svr_ip": "181.199.209.88",  
  "svr_cc": "CL",  
  "svr_city": "Santiago",  
  "svr_latitude": "-33.3928",  
  "svr_longitude": "-70.7856",  
  "svr_cidr": "181.199.209.64/26",  
  "svr_asn_id": "12956",  
  "svr_asn_name": "TELEFONICA GLOBAL SOLUTIONS SL"  
}
```

2.2.3 PyPI

¿Qué guardamos? PyPI

- PyPI es un índice de paquetes de Python
- Información sobre descargas y tráfico disponible en BigQuery
- Lo utilizamos para visualizar cantidad de descargas por hora

```
[{  
    "country_code": "AR",  
    "num_downloads": "5075",  
    "time": "2022-07-08 00:00:00.000000 UTC"  
, {  
    "country_code": "AR",  
    "num_downloads": "5935",  
    "time": "2022-07-08 01:00:00.000000 UTC"  
, {  
    "country_code": "AR",  
    "num_downloads": "3808",  
    "time": "2022-07-08 02:00:00.000000 UTC"  
}]
```

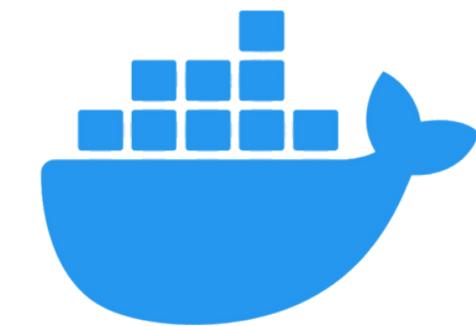
3. Tecnologías Usadas

Tecnologías Usadas



Python

Lenguaje de programación utilizado para prototipar y para orquestación de la ingestión.



docker®

Docker y Docker-Compose

Herramienta de containerización

Tecnologías Usadas - Almacenamiento



InfluxDB

Almacenamiento de datos provenientes de RIPE,
MLab y PyPI



Postgres

Almacenamiento de información acerca de las fallas
y estadísticas de ingestión del sistema

Tecnologías Usadas - Runtime



dagster

Dagster

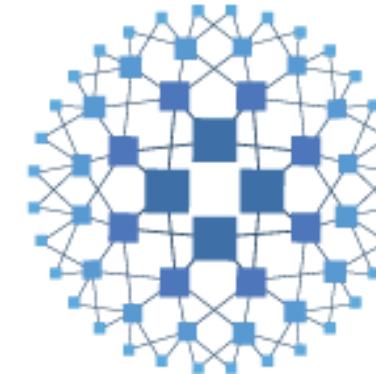
Herramienta de
orquestación del pipeline
(DAG) de ingestión



Grafana

Grafana

Herramienta de construcción de
dashboards y visualizaciones de
métricas de fallas



HAProxy

HAProxy

Proxy que ofrece capacidades de
rate limiting, seguridad y
balanceo de carga

4. Demo

5. Conclusiones

Conclusiones

1. Las fallas en cables submarinos impactan en la latencia, las rutas y/o la pérdida de paquetes.
2. Las soft failures son distinguibles de las hard failures.
3. Los mecanismos de redundancia hacen que sea más preponderante la degradación de la conexión a la pérdida total de la conexión.
4. No todas las soft failures tienen el mismo impacto.

Conclusiones

5. Una misma falla puede afectar múltiples países de manera diferente.
6. Una falla puede afectar a un sistema autónomo en un país, pero no afectar a otro del mismo país.
7. Las fallas en cables submarinos no suelen impactar en la descarga de paquetes de PyPI.
8. Una falla en un cable submarino puede generar congestión recurrente a partir de la misma.

6. Trabajo Futuro

Trabajo Futuro

1. Automatizar el proceso de detección de pistas posibles fallas en portales de noticias y otras fuentes mediante el uso de scraping y procesamiento de lenguaje natural.
2. Desarrollo de mecanismos estadísticos de detección de fallas a partir de las métricas existentes.
3. Ingestión en tiempo casi real de métricas de probes para detectar fallas en casi tiempo real y alertar sobre las mismas.
4. Incorporación de nuevas fuentes de métricas, en particular alguna que pueda dar cuenta del tráfico de red del usuario promedio.
5. Desarrollo de taxonomía formal de fallas según su impacto en las distintas variables.

¿Preguntas?

¡Muchas Gracias!

ALUMNOS

Aguerre, Nicolás Federico

naguerre@fi.uba.ar

Klein, Santiago

sklein@fi.uba.ar

TUTORES

Dr. Beiró, Mariano Gastón

mbeiro@fi.uba.ar

Dr. Carisimo, Esteban

ecarisimo@fi.uba.ar