TESIS DOCTORAL

TOMOGRAFÍA DE INTERNET: MEDICIÓN DE TRÁFICO Y SU RELACIÓN CON LA TOPOLOGÍA

AUTOR

ING. ESTEBAN CARISIMO

DIRECTOR Y CO-DIRECTOR DE TESIS

DR. ING. JOSÉ IGNACIO ALVAREZ-HAMELIN – FIUBA

DR. AMOGH DHAMDHERE – CAIDA/UC SAN DIEGO

JURADO DE TESIS

DR. CECILIA G. GALARZA – FIUBA
DR. ESTEBAN MOCSKOS – FCEN-UBA
DR. PABLO FIERENS – ITBA
DR. FABIÁN E. BUSTAMANTE – NORTHWESTERN UNIVERISTY (EE. UU.)

LUGAR DE TRABAJO

GRUPO DE REDES COMPLEJAS Y COMUNICACIÓN DE DATOS

INTENCIN (UBA-CONICET),

EN EL MARCO DE LA BECA DOCTORAL INTERNA DOCTORAL CONICET.

FACULTAD DE INGENIERÍA – UNIVERSIDAD DE BUENOS AIRES

BUENOS AIRES, ABRIL DE 2020

Tomografía de Internet: medición de tráfico y su relación con la topología

Ing. Esteban Carisimo

Índice general

Resumen **1**

Overview **3**

Agradecimientos **5**

1. Marco Teórico **7**

1.1. Nociones fundamentales del protocolo de Internet 7

1.1.1. Agotamiento IPv4 e IPv6 9

1.2. Border Gateway Protocol (BGP-4) 11

1.2.1. Elementos esenciales para la ejecución de BGP 12

1.2.2. Formato de mensajes BGP 13

1.2.2.1. Ejemplos de anuncios BGP en escenarios reales 15

1.2.2.2. Detalles de anuncios BGP en la práctica: MOAS 16

1.2.3. Algoritmo de selección de rutas y mecanismos de estímulo 16

1.2.3.1. Longest Prefix Match (LPM) 16

1.2.3.2. Algoritmo de selección de rutas 17

1.2.3.3. Método *AS-PATH prepend* 18

1.2.4. Servidor de rutas BGP 19

1.3. Topología de ASes 20

1.3.1. Relaciones comerciales entre ASes vecinos 21

1.3.1.1. Relaciones entre ASes en anuncios BGP 22

1.3.2. Clasificación de ASes según su rol 24

1.3.3. IXPs en el ecosistema de Internet 26

1.3.4. Redes de Distribución de Contenido (CDNs) 27

1.3.4.1. Arquitecturas de CDNs 29

1.3.4.2. Motivaciones a construir una CDN 29

1.3.5. Proyectos de medición de Internet 30

1.4. Mediciones activas en Internet 31

1.4.1. Mediciones activas a través de ICMP . . . 32
1.4.1.1. Ping . . . 32
1.4.1.2. Traceroute . . . 32
1.4.1.3. Aumento de la precisión a través de paris-traceroute . . . 35
1.4.2. Limitaciones . . . 35
1.4.3. Plataformas de mediciones activas . . . 38

2. Los CPs y transformación de la red **41**
2.1. Introducción . . . 42
2.2. Literatura relacionada . . . 44
2.3. Metodología . . . 46
2.3.1. Definiciones . . . 47
2.3.2. Evaluación de las métricas en el ecosistema de ASes . . . 48
2.3.3. Metodología propuesta . . . 51
2.4. Datos utilizados . . . 51
2.5. La evolución de los CPs en el núcleo de Internet . . . 53
2.5.1. Analizando los *sibling* ASes . . . 53
2.5.2. Monitoreando la evolución de los Big Seven en IPv4 . . . 55
2.5.3. Monitoreando la evolución de los Big Seven en IPv6 . . . 59
2.6. Evolución regional del núcleo de IPv4 . . . 62
2.6.1. Evolución geográfica de los Big Seven en IPv4 . . . 64
2.6.2. Vecinos locales . . . 67
2.7. El TOPcore más allá de los Big Seven . . . 68
2.7.1. Composición de los TOPcores . . . 68
2.7.2. Evolución de los adoptantes de dual-stack . . . 71
2.7.3. Tiempo para alcanzar el TOPcore . . . 71
2.7.4. Otros CPs destacados en el TOPcore . . . 73
2.8. Conclusions . . . 74

3. El impacto de los IXPs en Latinoamérica **77**
3.1. Introducción . . . 78
3.2. Datos utilizados . . . 80
3.3. Políticas Públicas en Latinoamérica e IXPs . . . 82
3.4. Evolución de los IXPs . . . 84
3.4.1. Topologías de las redes de IXP . . . 84
3.4.2. Miembros de las grandes de redes de IXPs . . . 85
3.4.3. ASes visibles . . . 86

3.4.3.1. Impacto doméstico . . . 87
3.4.3.2. Atracción de redes extranjeras . . . 90
3.4.4. Proveedores de tránsito . . . 90
3.4.5. Alcance geográfico de los IXPs . . . 92
3.5. IXPs y concentración . . . 94
3.6. Literatura relacionada . . . 96
3.7. Conclusiones . . . 97

4. La congestión persistente en Internet **99**
4.1. Introducción . . . 100
4.2. Otras distribuciones previamente propuestas . . . 101
4.3. La distribución estable . . . 102
4.4. Datos utilizados . . . 105
4.5. Firmas de congestión . . . 106
4.5.1. Desplazamiento de la distribución con la demanda diurna . . . 106
4.5.2. Evolución de las series temporales . . . 108
4.5.3. Firmas de congestión en pares de métricas . . . 111
4.6. Detección de la congestión con ML . . . 112
4.7. Conclusiones . . . 115

5. Conclusiones **117**

Bibliografía **121**

Índice alfabético **145**

Índice de figuras

1.1. Cintura del reloj de arena de la pila TCP/IP 8
1.2. Esquema de ejemplo de sesiones BGP. 13
1.3. Ejemplo de *AS-PATH prepending* 19
1.4. Comparación de número de sesiones BGP en presencia o ausencia de un servidor de rutas. 20
1.5. Comparación de anuncios ante relaciones c2p y p2p 22
1.6. Ejemplo de relaciones c2p entre ASes y los conos de clientes resultante. . 24
1.7. Diagrama de arquitectura de una Red de Distribución de Contenido. . . 28
1.8. Secuencia paso a paso del funcionamiento de un traceroute. 33
1.9. Posibles imprecisiones en un traceroute ejemplo debido al balance de carga. 36

2.1. Ejemplo de descomposición en k-núcleos de un grafo dado. 47
2.2. Visualización de la descomposición en k-núcleos del grafo de ASes en 2006 y 2016. 54
2.3. Evolución de los *Big Seven* en términos de cores en el grafo de ASes de IPv4. Todos ya han alcanzado el TOPcore. 56
2.4. Evolución de los *Big Seven* en términos de cores en el grafo de ASes de IPv6. Todos ya han alcanzado el TOPcore. 60
2.5. Evolución del shell-index de los *Big Seven* en cada RIR. La línea vertical punteada indica el inicio de lo registros de geolocalización. 63
2.6. Fracción de arribos de miembros del TOPcore ($k^* = 1$), acumulados a lo largo de los años. 66
2.7. Evolución del despliegue de dual-stack entre los ASes miembros de ambos TOPcores. 69
2.8. Evolución mensual de la fracción de CPs and proveedores de tránsito en el TOPcore. 69
2.9. Correlación entre la velocidad de crecimiento y la fecha de arribo al TOPcore. 70
2.10. Evolución en términos de cores de otros CPs notables en los grafos de ASes de IPv4 e IPv6. 72

3.1. Topología de la red de IXPs de CABASE y su política de anuncios. . . . 86
3.2. Número de ASes conectados a los IXPs regionales en IX.br, CABASE y PIT Chile en Julio de 2019. . . . 87
3.3. Fracción de loos ASN delegados al país y activos, que son visibles en los IXPs. . . . 88
3.4. Prevalencia de nacionalidades de los ASes visibles en los IXPs de Latinoamérica, África, Asia y Europa. . . . 89
3.5. Evolución del alcance geográfico de IX.br-SP, CABASE-BUE y PIT Chile-SCL. . . . 93
3.6. HHI para determinar la concentración del espacio de direcciones originado por país. . . . 96

4.1. Ejemplos de parametricaciones de las distribuciones estables. . . . 103
4.2. Histogramas por hora de los RTTs obtenidos del extremo lejano un enlace ente dominios. . . . 107
4.3. Comparación entre los parámetros δ y γ y la media y el desvío de las muestras $\log(RTT)$. . . . 107
4.4. Firma de latencia por pares en un enlace congestionado (izquierda) y no congestionado (derecha). . . . 109
4.5. Firmas de congestión para proveedores de acceso con diferentes tipos de redes adyacentes . . . 110
4.6. Exactitud, sensibilidad y precisión del clasificador ML para firmas de congestión. . . . 113

Índice de cuadros

1.1. Ejemplo selección de ruta por medio de LPM en un caso de superposición de prefijos 17

2.1. Comparación de métricas para estudio de despliegue de CPs 49
2.2. Porcentaje de vecinos locales en cada región. 65
2.3. Origen geográfico de acuerdo con lo reportando en WHOIS para los TOP-core ASes 67

3.1. Resumen de IXPs en Latinoamérica, sus patrocinadores y operadores. . . 83
3.2. Los cincos mayores upstream ASes en IX.br-SP, CABASE-BUE y PIT Chile-SCL en Julio de 2019. 91
3.3. Los dos mayores ASes origen por país. $*$ indica ASes de propiedad estatal. 96

Índice de listados de código

1.1. Ejemplo de anuncio recibido por el colector de Routeviews (oO19) ubicado en San Pablo, Brasil, el 1 de Marzo de 2017 entre las 00:00 y 00:01. El anuncio contiene una ruta hacia el prefijo 157.92.42.0/24, originado por la Universidad de Buenos Aires (AS3449). El acceso a esta información fue por medio de la herramienta BGPstream (OKG+16). 15

Resumen

Esta tesis se enfoca en el estudio de la topología de Internet, principalmente a nivel de Sistemas Autónomos, con especial interés en la relación que existen entre la estructura de la red de Internet y el tráfico que es transmitido. La red de Internet presenta un singular atractivo para su investigación: es el resultado de la interconexión de miles de organizaciones independientes, esparcidas alrededor del mundo, cada una con intereses y objetivos diferentes. Asimismo, la ausencia de un organismo central coordinador de la red, la reducida visibilidad de la topología a raíz de la dinámica de los protocolos y el resguardo de las organizaciones a revelar su estructura plantean el desafío de descubrir la topología de Internet y comprender las dinámicas por las cuales la red adopta tal forma.

Un gran número de factores nos motiva al estudio constante de la topología de Internet, sus cambios y su adaptación en función del tráfico. En primer lugar, la red se encuentra en constante crecimiento, debido a la incorporación de nuevos proveedores, nuevos usuarios y un mayor número de dispositivos conectados. En segundo lugar, ciertos eventos disruptivos, por ejemplo la aparición de los *smartphones*, generan periódicamente grandes cambios, que posteriormente también modifican la estructura de la red. Por último, el auge de nuevos servicios brindados a través de Internet, por ejemplo el significativo crecimiento de las plataformas de video en la última década, introducen cambios en los patrones de tráfico, la calidad de la experiencia y posiblemente también en la topología de la red.

En esta tesis presentaremos el estudio de la topología de la red y su vínculo con el tráfico transportado de manera integral, centrándonos en probar la siguiente hipótesis:

> *La irrupción de los proveedores de contenido a gran escala ha reconfigurado la topología de Internet, en particular alterando los principales puntos de entrega de tráfico*

La prueba la haremos recolectando evidencias desde tres perspectivas diferentes. El primer enfoque lo haremos a través del estudio de los proveedores de contenido, quienes en la actualidad son responsables de la generación de la mayor parte del tráfico transportado sobre Internet. Luego, nos centraremos en el estudio de los IXPs, ya que en estas estructuras se intercambia gran parte del tráfico de Internet, incluido el tráfico generado por los grandes proveedores de contenido. Finalmente, nos enfocaremos en examinar los efectos del marcado aumento del tráfico multimedia a través de puntos de la red con capacidad insuficiente para cumplir con tal demanda.

Nuestro aporte consiste en abordar diferentes eslabones en la generación y distribución de tráfico, y combinar el uso de herramientas matemáticas para caracterizar los

cambios observados en la topología, como también en métricas indirectas del tráfico y la congestión.

La tesis está organizada de la siguiente forma:

- En el Capítulo 1 introduciremos elementos teóricos de las redes de computadoras, tales como protocolos, terminología y otros conceptos afines al despliegue y operación de redes IP.
- En el Capítulo 2 estudiaremos a través de métricas de teoría de grafos cómo la consolidación de los proveedores de contenido multimedia transformó de la estructura de la red, llevando a que los proveedores de contenido se establezcan como organizaciones densamente conectadas.
- En el Capítulo 3 examinaremos la evolución y composición de los IXPs en Latinoamérica, ya que gran parte del tráfico de la región es presumiblemente intercambiado a través de estas estructuras.
- El Capítulo 4 abordará el estudio de la congestión persistente en los enlaces entre dominios, proponiendo el modelado matemático de la latencia para determinar el estado de los enlaces.
- Por último, el Capítulo 5 sintetizará el trabajo realizado a lo largo de la Tesis, poniendo el foco sobre las contribuciones realizadas.

Overview

This thesis focuses on the study of Internet's topology, mainly at AS level, with a special interesest in the existing relationship between network structure and traffic carried over it. Internet's network has a unique and attractive singularity: it consists of thousands of independent interconnected organizations, scattered all over the world, each one with different interests and goals. In addition to Internet's complexity, a number of technical limitations have set some challenges to understand network dynamics, such as the lack of a central coordinating organization, the limited visibility of the entire network due as well as organizations protecting commercial strategies by partially hiding their networks.

A large number of reasons motive the continious study of Internet topology, its changes and the way the network is reshaped to be capable of incorporating new traffic trends. First, Internet's network has been always growing driven by the birth of new providers, the raising penetration of broadband and the increasing number of connected devices. Second, periodic disruptive events, the latest one is the quick widespread of mobile devices, introduce big changes that afterwards impact on the network structure. Finally, the rise of online services, such as the large growth of on-demand video streaming platforms in the last decade, creates variations on traffic patterns, quality of experience and probably on the network topology as well.

This thesis presents a study of Internet's network and its relationship with the carried traffic in an integral manner by focusing on the the following hypothesis:

> *The large scale irruption of Content Providers have reshaped Internet's topology, particularly changing the main points for traffic delivery*

We proof the hypothesis by collecting evidence from three different perspectives. We begin by studying Internet Content Providers, which are currently responsible for generating the largest fraction of traffic carried over the network. Next, we investigate the role of IXPs in Latin American AS ecosystems, since large amounts of traffic are exchanged at IXPs all over the world, including traffic generated by large Content Providers. Finally, we focus on examining how raising multimedia demand affects quality of experience when underprovisioned network capacities cannot meet such demand.

Our contribution is to look at different parts of the traffic generation and traffic delivery path where we combine mathematical tools to model observed changes in traffic topology as well as indirect traffic and congestion metrics.

This thesis is structured as follows

- Chapter 1 introduces a brief high-level background of computer networks, such as protocols, terminology and other concepts related to rollout and network operations.
- In Chapter 2, we apply graph-theoretical metrics to study how Content Providers consolidation reshaped network structure in which Content Providers have turned out operating densely connected networks.
- In Chapter 3, we study the evolution and growth of Latin American IXPs, since most of the traffic of the regions is presumably exchanges at domestic IXPs.
- Chapter 4 presents the use of mathematical models to pinpoint interdomain links affected by persistent congestion during peak hours,
- Finally, Chapter 5 summarizes the work along this thesis, focusing on our contributions.

Agradecimientos

Esta tesis no hubiera sido posible sin la presencia, ayuda y compañía de muchas personas a lo largo de estos años. Antes de comenzar, quiero dedicarle algunas palabras a todos ellos.

En primer lugar, quiero agradecerle a Ignacio Alvarez-Hamelin por haberme acompañado a lo largo de todo este camino desde su lugar de director. Estoy sumamente agradecido a toda la dedicación y el tiempo que Ignacio dedicó a mi formación como Doctor. Quiero además de agradecerle un hecho que para mi es sumamente importante, y es la piedra fundamental para que hoy culmine mi doctorado: Ignacio insistió en que me inscribiera en el Doctorado, algo que para mi era impensado en ese momento.

También quiero agradecerle a Amogh Dhamdhere, quien me brindó la posiblidad de comenzar a trabajar junto a él a mitad de mi doctorado. Gracias a Amogh pude conocer nuevas formas de investigar, y aprendí de él una rigurosidad excepcional a la hora de llevar a cabo una investigación. Además agradezco a Amogh por el optimismo y el entusiamo que siempre mostró por cada uno de los trabajos que realizamos juntos.

También quiero dedicarle unas palabras a todas las personas junto a las que, de alguna u otra manera, investigué estos años. Quiero especialmente agradecerle a los coatuores de mis artículos, por las innumerables horas de dedicación para poder plasmar nuestros trabajos. Además quiero decirle gracias a todos aquellos que generosamente y desinteresadamente me ayudaron con mis investigaciones.

Finalmente, quiero agradecerles a mi familia y a mis amigos que siempre estuvieron conmigo, y me apoyaron constantemente durante esta etapa. No tengo palabras para agradecerle a mis viejos, Edgardo y Lula, y a mi hermano Agustín por el apoyo incondicional que me brindaron todos estos años. Agradecerle a mi tío *Orly*, por siempre estar a mi lado. También querido agradecerle infinitamente a todos mis amigos por su compañía, risas y apoyo.

Por último, quiero decirle gracias a mi abuela *Esther Nidia Wächter*, que ahora me acompaña desde otro lugar.

Capítulo 1

Marco Teórico

1.1. Nociones fundamentales del protocolo de Internet

La comunicación entre aplicaciones a través de Internet se lleva a cabo por la pila de protocolo TCP/IP, compuesta por cinco capas, donde cada una tiene una función específica para lograr la comunicación entre emisor y receptor. En particular, en esta tesis nos enfocaremos en la capa de red de la pila de protocolos TCP/IP.

La capa de red tiene como objetivo brindar el servicio de comunicación entre dispositivos, también denominados hosts o end-hosts. Esto hará que la capa de red sea la encargada de mover la información, en este caso los paquetes o datagramas, a través de la red de manera tal de poder comunicar hosts entre sí.

Dentro de la pila de protocolos TCP/IP, la capa de red tiene una característica excepcional: es la única capa en la cual no existe variedad de protocolos, siendo el protocolo de Internet (IP) el único dedicado al proceso de entrega de paquetes. Esto le otorga un papel central al protocolo IP, no sólo será el responsable de llevar a cabo al distribución de la información en la red, sino que será el único protocolo que estará obligatoriamente presente en todos los dispositivos. Además, ya que su rol es precisamente poder comunicar dispositivos, todos los dispositivos tendrán implementadas funciones de red. Esta centralidad y unicidad del protocolo IP, es usualmente expresada como la *cintura del reloj de arena*, tal como se muestra en la Figura 1.1. En la figura podemos notar la existencia de múltiples de protocolos en todas las capas, pero indefectiblemente, todos son encapsulados en datagramas IP.

La infraestructura a nivel de capa de red está compuesta por los dispositivos terminales (end-hosts), quienes son esencialmente los generadores y consumidores del tráfico de la red, y dispositivos intermediarios (ruteadores), encargados de interconectar las redes

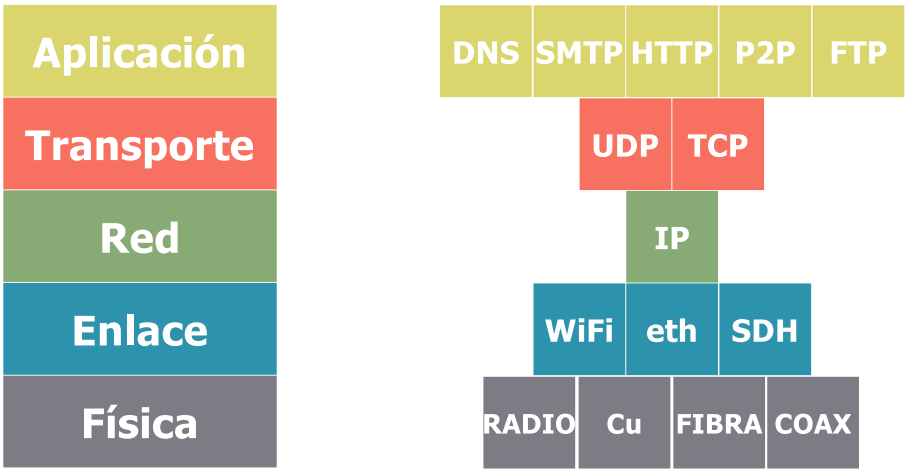


Figura 1.1: Definición de cada una de las capas presentes en la pila de protocolos TCP/IP. Existen una notable variedad de protocolos y tecnologías en todas las capas, excepto en la capa de red, siendo el protocolo IP *la cintura del reloj de arena*.

---

de dispositivos entre sí.

Centrándonos ahora exclusivamente en los ruteadores, su función requiere de dos instancias: la conexión a múltiples redes y el procesamiento de paquetes. Para poder cumplir con la misión de interconectar redes, los ruteadores deben lógicamente estar conectados a dos o más redes. Sin embargo, la distribución de los datagramas se concreta en el momento que los ruteadores reciben los datagramas, y procesan en qué dirección lo deben enviar. Las funciones vinculadas al procesamiento de paquetes en los ruteadores se dividen en dos categorías nítidamente definidas: *forwarding* y *routing*.

El *forwarding* (en español entendido como envío o re-envío) es la acción que ejecutan los ruteadores al recibir un paquete por una de sus interfaces. Esta consiste en enviar o re-enviar el paquete entrante por la interfaz de salida adecuada de forma tal de que este paquete tome la dirección correcta para alcanzar su destino. En cambio el *routing* (en español entendido como ruteo) es la acción de determinar, para un cierto ruteador, cuál es la interfaz adecuada para alcanzar un destino determinado. Es importante destacar que esta acción (determinar interfaces para cada destino), se lleva a cabo independientemente del arribo de paquetes para dichos destinos.

Las acciones relacionados con el proceso de *forwarding* suelen ser ejecutadas por medio de hardware presente en los dispositivos. A su vez, la puesta en práctica del forwarding depende exclusivamente del dispositivo, es decir, es inconsulta e independiente de las acciones de *forwarding* ejecutadas por los otros dispositivos de la red. Por el contrario, las acciones relativas al *routing* son implementadas en software, y requieren de coordinación, conocimiento de la topología y del estado del resto de los elementos que componen la

red. Tal procedimiento presente en el ruteo de la capa de red es conocido bajo el nombre de *algoritmos de ruteo*.

1.1.1. Agotamiento IPv4 e IPv6

En la actualidad, **todos** los dispositivos conectados a la red pública de Internet deben estar asociados de una dirección del protocolo de Internet versión 4 (IPv4), la cual debe ser única. El requerimiento de unicidad de cada dirección se debe a que las comunicaciones son entre dispositivos, y no pueden existir ambigüedades, ya que de otra manera, la información no sería fehacientemente entregada al destinatario. Aquí se destaca que no solo las direcciones son únicas, sino que adoptan el formato definido en la versión 4 de protocolo (ver RFC 791) ([Pos81b]), el cual fue masivamente adoptado por los dispositivos conectados a la red en el año 1981.

El formato de las direcciones definidas en la RFC 791 detalla que tienen una extensión fija de 32 bits. Esto significa que la cantidad máxima de direcciones únicas, y por ende de dispositivos, es 4.400 millones. Por más que esta cantidad era excesivamente abundante en 1981 cuando se introdujo la versión actual del protocolo, el grado de penetración de Internet en la sociedad ha llevado a que este espacio se torne insuficiente para conectar a todos los dispositivos.

Una de las razones por las cuales en la actualidad escasean las direcciones IPv4 es por fallas en el proceso de delegación de las direcciones. En la era de Internet primigenia, donde los participantes eran principalmente universidades (en su mayoría de EE.UU.), y otras entidades (donde se destaca la presencia del Departamento de Defensa de EE.UU.), se efectuaron grandes delegaciones ya que no se avizoraba la posibilidad de una escasez de recursos. La mayor parte de estas entidades recibió una o varias redes Clase “A” (bloques de 16 millones de direcciones), las cuales exceden ampliamente sus necesidades, dejando gran cantidad de direcciones sin uso ([DBK+16]).

A principios de los noventa, al notar la aceleración de la reducción del stock de direcciones, el Grupo de Trabajo de Ingeniería de Internet (del inglés Internet Engineering Task Force, IETF) y la Autoridad de Números Asignados en Internet (del inglés Internet Assigned Numbers Authority, IANA), entidades a cargo de la generación de estándares y delegación de recursos respectivamente, comenzaron a introducir cambios para hacer uso racional de los recursos. El primer cambio consistió en la reforma de la metodología de asignación de redes, sustituyendo *Classful Routing* ([MP85]) por *Classless Inter-Domain Routing (CIDR)* ([FLYV93; FL06]). Previo a que CIDR fuera introducido, *Classful Routing* permitía asignar redes de sólo tres rangos: Clase “A” (16 millones de hosts), Clase “B” (65 mil de hosts) y Clase “C” (254 hosts). El reducido número de clases y el marcado

aumento en el número de hosts entre clases contiguas (la clase subsiguiente contiene 256 veces más hosts), imposibilitaba ajustar las asignaciones con precisión a las necesidades de direccionamiento, dando lugar a redes subutilizadas. La migración a CIDR permitió una nueva estrategia de asignación de redes es más eficiente, lo cual redujo la tasa de consumo del stock de direcciones. Sin embargo, Classless Inter-Domain Routing no brinda herramientas para ampliar el previsiblemente escaso espacio de direcciones.

Ante la evidente falta de direcciones en un futuro próximo, a mediados de los noventa, el IETF comenzó a trabajar en diferentes alternativas. Sin embargo, esta labor contaba con una limitación fundamental: El protocolo IP es el único protocolo insustituible de toda la pila TCP/IP, sumado a que todos los dispositivos lo implementan. Es decir, un cambio en el protocolo requeriría que todos los dispositivos lo adopten para poder seguir interoperando. No obstante, al momento que planteó la necesidad de introducir cambios en el protocolo IP (precisamente IPv4), la cantidad de dispositivos conectados a la red ascendía a cientos de millones, lo cual tornaba imposible efectuar un *flag day*[1]. Frente a este escenario, el IETF trabajó simultáneamente en dos alternativas: i) *Network Address Translation(NAT)* ii) protocolo de Internet versión 6 (IPv6).

La primera alternativa es NAT (EF94), la cual mediante el uso de direcciones de red privadas (RMK+96) permite que múltiples dispositivos accedan a Internet compartiendo la misma dirección IP origen. Para poder llevar a cabo este proceso, dispositivos que corren NAT (en particular ruteadores), traducirán las direcciones privadas a públicas y viceversa cuando los paquetes pasen por ellos. Esta modificación elimina la necesidad de que cada dispositivo tenga una dirección IPv4, ya que multiples dispositivos (hasta el orden de los 10.000) pueden utilizar la misma dirección pública. Como contrapartida, esta modificación elimina el principio de horizontalidad de la red, donde *todos se pueden comunicar con todos* producto del hecho de compartir direcciones IP.

Las limitaciones que exhibe NAT no han presentado mayores inconvenientes en el paradigma actual de Internet, ya que los usuarios son esencialmente consumidores netos de contenido. La falta de barreras para la adopción de NAT permitió excepcional despliegue de esta tecnología (SHS+12). Aunque NAT haya reducido drásticamente el uso de direcciones IP, aún requiere el uso de direcciones públicas, consumiendo el stock de direcciones IPv4.

La segunda alternativa es IPv6 (DH98; DH17) la cual consiste en una serie de mejoras al protocolo IPv4, donde en el centro de la escena se destaca una expansión del espacio de direcciones. Habiendo experimentado que la caducidad de IPv4 es dará a causa de un subdimensionamiento en el espacio de direcciones, IPv6 propone expandir las direcciones

[1]Flag day: Día en el cual se adopta un cambio sustancial de versión de software lo que implica reiniciar todos las partes que componen el sistema.

de 32 a 128 bits, permitiendo albergar un total de $3{,}4 \cdot 10^{38}$ dispositivos. Por más que esta expansión elimine por completo la escasez de recursos, su éxito radica en la adopción de la nueva versión del protocolo. Como se mencionó anteriormente, no es posible llevar a cabo un *flag day*, por lo tanto, la solución es que los dispositivos utilicen dos direcciones, una IPv4 y otra IPv6, donde esta metodología se denomina *doble pila* (en inglés, *dual-stack).*

Aunque el estándar de IPv6 fue lanzado en 1998, su adopción fue casi nula hasta comienzos de 2010, presumiblemente por falta de incentivos concretos para su implementación. Sin embargo, más recientemente una serie de sucesos ha impulsado la adopción de IPv6, más precisamente de doble pila. El primer evento fue la última transferencia de bloques IPv4 de IANA a los Registros Regionales de Internet (en inglés Regional Internet registry, RIRs) (IAN11) en 2011, dejando la reserva central de direcciones en cero, y a la espera de que lo mismo suceda pronto en cada región. Un segundo efecto resonante fue el agotamiento total del stock de direcciones IPv4 del Registro Regional de Internet para Estados Unidos, Canadá y el Caribe (del inglés American Registry for Internet Numbers, ARIN) (ARI15) en Septiembre de 2015.

1.2. Border Gateway Protocol (BGP-4)

El ruteo, es decir el proceso de selección del camino a seguir por un paquete, esta dividido en dos niveles: *ruteo interno* y *ruteo externo*. Precisamente, *Border Gateway Protocol (BGP)* (RLH06) es, en la práctica, el único protocolo por el cual se lleva a cabo el ruteo externo en Internet.

Antes de comenzar con los detalles del protocolo BGP, es importante comprender en primer lugar el motivo detrás de la necesidad de desdoblar el proceso de ruteo en dos jerarquías. Primordialmente, la separación se centra en la necesidad de agrupar los ruteadores que componen la red en subconjuntos, donde los motivos se deben básicamente a dos principios: i) Escalabilidad ii) Independencia administrativa.

La necesidad de la escalabilidad surge del hecho que la red está compuesta por millones de ruteadores (CAI20), donde ciertos ruteadores no necesitan conocer en detalle la totalidad de la topología de la red. Más aún, ciertos protocolos de ruteo (interno) como OSPF (Moy98), ampliamente utilizado en Internet, basa su principio de comunicación entre ruteadores a través de mensajes broadcast. Debido a la escala de la red, sería imposible propagar por toda la red mensajes broadcast generados por millones de ruteadores al mismo tiempo. Entonces, la solución se genera a través de la reducción del número de ruteadores a los que se debe alcanzar por medio de mensajes de broadcast.

Por otra parte, la red pública de Internet es el resultado de la interconexión de miles de proveedores de servicios de Internet (en inglés *Internet Service Providers*, ISPs),

donde cada uno es propietario y administrador de una red de ruteadores. La voluntad de independencia administrativa se debe a que cada ISP puede tener la intención de utilizar diferentes protocolos de ruteo para su propia red de ruteadores. Sin embargo, dada la naturaleza comercial que actualmente se observa en Internet, los operadores pueden querer ocultar ciertos detalles de la vista de los demás, ya que estos pueden formar parte de su estrategia comercial. Entre los elementos más importantes a ser resguardados aparecen: la estructura de su propia red, acuerdos con otros ISPs o sus políticas de selección de rutas. Este último argumento lleva a que los conjuntos de ruteadores de cada operador tengan cierto grado de independencia de los ruteadores de los demás.

El resultado de esta necesidad de separar los ruteadores en conjuntos es la creación de una entidad denominada Sistema Autónomo (en inglés *Autonomous System*, AS). Un Sistema Autónomo entonces es una entidad la cual posee y administra ciertos ruteadores dentro de Internet, sobre los cuales tiene total autonomía administrativa. Cada AS se encuentra identificado por una clave única denominada Número de Sistema Autónomo (en inglés *Autonomous System Number*, ASN). Por lo tanto, para que un ISP disponga fácticamente de autonomía, éste debe contar con un ASN. Por lo general, cada ISP dispone de un ASN, sin embargo, existe evidencia empírica que demuestra que ciertas organizaciones disponen de más de un ASN (CHKW10).

Para concluir, presentaremos brevemente la función de BGP. El protocolo BGP, en su función de protocolo de ruteo externo, tiene como objetivo proveer el ruteo a nivel de Sistemas Autónomos (ASes), es decir calcular y difundir caminos de ASes para poder alcanzar un prefijo destino. Este concepto proyecta la importancia que posee BGP en el sistema, ya que este protocolo es el encargado de unir miles de ISPs independientes. Lógicamente, para que dos ASes puedan intercambiar información de ruteo, ambos deben implementar el mismo protocolo de ruteo, lo cual lleva a que BGP sea el protocolo de ruteo *de facto* de Internet.

1.2.1. Elementos esenciales para la ejecución de BGP

Antes de presentar los detalles de los mensajes del protocolo BGP, pondremos en contexto *dónde* y *cómo* ocurre la implementación del protocolo BGP en la red, basándonos en el esquema ilustrado en la Figura 1.2.

La Figura 1.2 nos presenta una imagen reducida de cómo es la estructura de la red de Internet, ya que en esta observamos la interconexión de dos ASes, donde cada uno se muestra como una red de ruteadores. Esto es ciertamente un recorte de como es la totalidad de Internet, ya que la red de Internet es el resultado de la interconexión de miles de ASes. Sin embargo, esta figura cuenta con todos los elementos para comprender

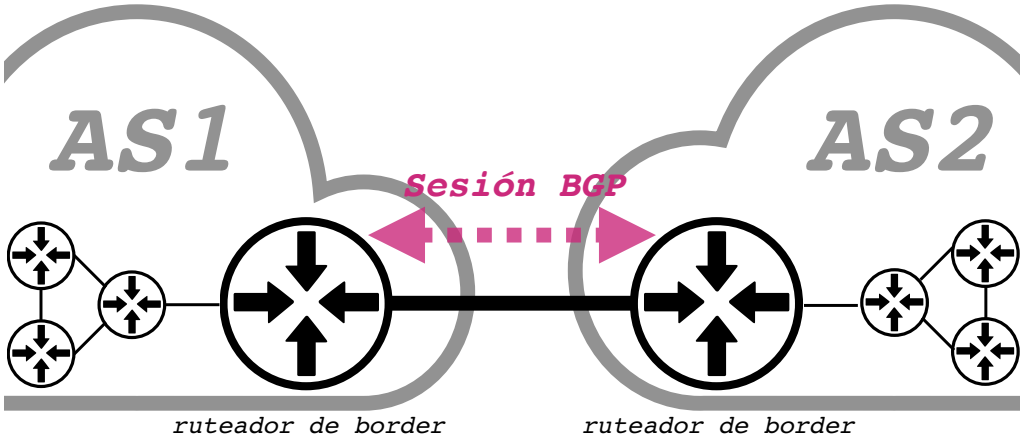


Figura 1.2: Ejemplo de interconexión entre dos ASes, donde se ilustra la puesta en funcionamiento de BGP entre Sistemas Autónomos. En este pequeño ejemplo se destaca la presencia de ruteadores de borde y el establecimiento de sesiones BGP.

---

la puesta en funcionamiento de BGP.

En primer lugar, la figura destaca la presencia de *ruteadores de borde*, que es donde específicamente se ejecuta el protocolo BGP dentro de un Sistema Autónomo. Su denominación es derivada de su ubicación en la topología del AS, ya que estos se colocan en la frontera del AS, conectándose con ruteadores de ASes vecinos. En el ejemplo de la figura, los ruteadores de borde se encuentran simplemente conectados a un solo AS vecino, pero esto responde a una limitación de la representación y no del protocolo. La figura además presenta otro requisito para la operación de BGP, cada AS debe disponer de un ruteador de borde en su frontera. Este hecho genera la presencia de un ruteador de borde a cada lado de los extremos del enlace entre ASes. Por última, el ejemplo también muestra que para poder llevar a cabo el intercambio de mensajes, los ruteadores de borde a cada lado del enlace establecerán una conexión TCP, la cual es conocida como *sesión BGP*.

1.2.2. Formato de mensajes BGP

El protocolo BGP dispone de cuatro tipos de mensajes *open, update, notification, keep alive*, de los cuales nos enfocaremos exclusivamente en el mensaje de actualización, *update*.

Los mensaje de *update* son centrales en el funcionamiento de BGP, ya que a través de esto se informa la alta y baja de rutas entre ASes. La estructura del mensaje de *update* cuenta con los siguientes elementos: 1) Prefijo[2] 2) Atributos 3) Rutas a descartar.

[2]En las especificaciones BGP el prefijo también se denomina *Network Layer Reachability Information (NLRI)*

El prefijo indica la red destino para la cual se está anunciando la ruta, donde el mensaje se expresa como una 2-tupla `(máscara, red)`, por ejemplo `(16, 157.92.0.0)`. Este campo tiene se presenta de la misma manera para aprender rutas nuevas (*updates*) como para descartar rutas (*withdrawals*).

Dado que un prefijo o una ruta puede tornarse inaccesible o inválida, los mensajes de *update* requieren poder indicar las rutas a revocar. En caso de ser necesario revocar rutas, el mensaje indicará estas rutas mediante el campo *Unfeasible Routes*. Si éste se encuentra vacío, no se llevará a cabo ningún descarte.

Los atributos son el último campo que conforman la estructura del mensaje BGP. Este campo es de largo variable, en donde se incluye una lista de atributos los cuales permiten indicar aspectos específicos de la ruta que está siendo anunciada. La lista de atributos está compuesta por tres tipos de elementos: obligatorios, discrecionales y opcionales.

Los atributos obligatorios son aquellos indispensables para poder definir una ruta, por lo tanto deben estar incluidos en todo anuncio. Estos elementos son `NEXT HOP` y `AS-PATH`. El campo `AS-PATH` indica la cadena de ASes que el paquete debe atravesar para alcanzar el prefijo destino indicado en el anuncio. Esta cadena de ASes no puede contener bucles, tal como lo impone la definición del protocolo BGP. La composición del `AS-PATH` surge de los anuncios de rutas entre ASes, donde un AS al anunciar una ruta, debe agrega su ASN en el extremo izquierdo de la cadena. De esta manera, `AS-PATH` definirá un camino hacia un prefijo, y además indicará la existencia de enlaces entre dominios entre ASNs contiguos. Continuando con esta lógica, el AS que se encuentra en el extremo derecho de la cadena es denominado *AS origen*, ya que es el AS donde se encuentra el prefijo anunciado.

Por su parte, el campo `NEXT HOP` indica la dirección IP del ruteador de borde del AS vecino por el cual se accede a la ruta indicada. Este campo es imprescindible ya que dos ASes vecinos, pueden contar con múltiples enlaces entre sí, y es necesario indicar precisamente por cual enlace se accede a la ruta. Más aún, podría surgir el caso donde distintos ruteadores en un AS vecino anuncien rutas diferentes. Además, el AS que recibe un anuncio puede tomar decisiones de selección de rutas en función de que `NEXT HOP` es más conveniente para alcanzar el destino.

Los atributos discrecionales son aquellos que pueden estar eventualmente presentes en un mensaje, como por ejemplo `LOCAL PREFERENCE (LOCAL PREF)`, el cual presentaremos con mayor detenimiento en la Sección 1.2.3.

```
{'as-path': '262750 28329 262589 262195 11058 3597 3449',
 'communities': set(['28329:2000',
                     '28329:2200',
                     '52376:1991',
                     '52376:991',
                     '64644:10100',
                     '64644:10101',
                     '64800:3333']),
 'next-hop': '187.16.217.14',
 'prefix': '157.92.42.0/24'}
```

Listado de código 1.1: Ejemplo de anuncio recibido por el colector de Routeviews (oO19) ubicado en San Pablo, Brasil, el 1 de Marzo de 2017 entre las 00:00 y 00:01. El anuncio contiene una ruta hacia el prefijo 157.92.42.0/24, originado por la Universidad de Buenos Aires (AS3449). El acceso a esta información fue por medio de la herramienta BGPstream (OKG+16).

Entre los atributos opcionales se encuentran las comunidades, cuya misión aumentar la información presente en los anuncios por medio de etiquetas (CTL96) Este mecanismo de señalización forma parte de las prácticas habituales de los operadores, y su difusión ha crecido sustancialmente durante los últimos años (SLB+18).

1.2.2.1. Ejemplos de anuncios BGP en escenarios reales

Para fortalecer la explicación de la estructura del mensaje de *update* de BGP y sus atributos, usaremos un ejemplo real de un anuncio observado en Internet. El listado de código 1.1 presenta el anuncio recibido por un router de borde perteneciente al proyecto Routeviews (oO19) (AS6447) ubicado en San Pablo, Brasil. Este anuncio fue recibido el 1 de Marzo de 2017 entre las 00:00 y 00:01 y contiene una ruta hacia el prefijo 157.92.42.0/24, originado por la Universidad de Buenos Aires (AS3449). El acceso y el extracción de la estructura de este anuncio fue realizado por medio de la herramienta BGPstream (OKG+16)

En el anuncio podemos observar la presencia de elementos indispensables para definir una ruta tales como el prefijo y los atributos `AS-PATH` y `NEXT HOP`. Además, este anuncio dispone de atributos opcionales como lo es la lista de comunidades. En particular, en este ejemplo podemos identificar que el prefijo corresponde a la Universidad de Buenos Aires, ya que el ASN que se encuentra más a la derecha del AS-PATH, es el 3449, el cual

fue delegado a la universidad.

1.2.2.2. Detalles de anuncios BGP en la práctica: Multiple Origin ASes

Concluiremos la presentación de los mensajes y anuncios de BGP presentando un escenario que es frecuentemente observado en la práctica: la presencia de *Multiple Origin ASes (MOAS)*.

En la Sección 1.1.1 argumentamos la necesidad de que cada dispositivo disponga de una dirección IP única, ya que de otra manera, esto generaría ambigüedades para el envío de datagramas. Este concepto puede ser extendido a nivel de redes, donde dos redes en Internet no pueden utilizar el mismo prefijo ya que esto conduciría nuevamente ambigüedades. Como hemos explicado recientemente, el último ASN en la cadena indica el AS donde se encuentra asignado el prefijo, por lo tanto para no incurrir en ambigüedades, los prefijos serán únicos siempre y cuando sean anunciados por el mismo AS origen. Sin embargo, esto en la práctica no es completamente cierto, ya que existe un gran número de prefijos originados por más de un ASN, no siendo esto a causa de errores de configuración ni secuestro de prefijos (en inglés*prefix hijacking*) (BFZ07). En la práctica esto se conoce como *Multiple Origin ASes* (MOAS) y en un gran número de oportunidades se debe a raíz de organizaciones que poseen múltiples ASNs (LJL+11), haciendo que el MOAS sea legítimo.

1.2.3. Algoritmo de selección de rutas y mecanismos de estímulo

Producto la propagación de los prefijos a través del protocolo BGP, los ASes pueden recibir múltiples rutas hacia un mismo prefijo. Ante esta multiplicidad, es necesario que el protocolo BGP determine cual es la ruta más conveniente para alcanzar el destino, lo cual se logra por medio del *algoritmo de selección de rutas*. Aunque los anuncios BGP, y su consolidación en las tablas de ruteo, se generan de manera automática de acuerdo con los cambios percibidos en la red, el algoritmo de selección de rutas también contempla entradas usualmente pre-configuradas de forma manual. Luego de este proceso, el algoritmo reportará la *mejor* ruta para cada prefijo.

1.2.3.1. Longest Prefix Match (LPM)

Los ruteadores seleccionan el puerto de salida por el que enviarán un datagrama basándose en una regla central en la operación de las redes IP: *Longest Prefix Match* (LPM) (DKT06; CR05). Mientras que las direcciones destino indicadas en la cabecera

RIB		Análisis de datagrama		
prefijo	interfaz	IP destino	¿Coincidencia?	¿Interfaz utilizada?
157.92.0.0/16	if0	157.92.48.31	Sí	No
157.92.48.0/22	if1	157.92.48.31	Sí	Sí

Cuadro 1.1: A la izquierda se muestra una tabla de ruteo con dos prefijos superpuestos pero con distintos puertos de egreso. A la derecha si muestra el análisis cuando la dirección destino es 157.92.48.31. La dirección se encuentra en el rango de ambos prefijos, pero se opta por aquel de máscara de mayor longitud.

---

del datagrama son únicas (también conocido como *unicast*), las entradas en las tablas de ruteo (también llamadas en inglés *Routing Information Base*, RIB) utilizan prefijos, es decir bloques de direcciones, como índices. Al usar bloques de direcciones, y al no existir restricciones para la superposición de prefijos, una dirección IP puede coincidir con más de una entrada en la tabla de ruteo. Por lo tanto, dado el caso de múltiples coincidencias, el ruteo IP plantea el mecanismo de LPM para efectuar el desempate seleccionando el prefijo con la máscara de mayor longitud.

Para aportar claridad al funcionamiento de LPM, el Cuadro 1.1 presenta una RIB con dos entradas superpuestas, y también describe el procedimiento de selección de prefijos cuando la dirección IP destino coincide con ambas entradas. En este ejemplo, la dirección destino 157.92.48.31 se encuentra incluida en el prefijo 157.92.0.0/16, como también en el 157.92.48.0/22. Entonces, al emplear el mecanismo de LPM se seleccionará el prefijo 157.92.48.0/22, al cual le corresponde la interfaz de salida `if1`.

1.2.3.2. Algoritmo de selección de rutas

Dado que un Sistema Autónomo, y en particular un ruteador puede contar con múltiples rutas para un mismo prefijo, el protocolo BGP tiene como objetivo seleccionar la *mejor* ruta para cada uno de los prefijos. El resultado de la selección es el producto de la evaluación ponderada de ciertas métricas vinculadas a los atributos de cada ruta recibida. Sin embargo, la mayor virtud del protocolo BGP es brindarle a los operadores la capacidad de seleccionar rutas de manera discrecional. De esta manera, por medio de BGP, los operadores pueden establecer factores arbitrarios, de forma tal que la ruta seleccionada sea aquella que obedezca motivos económicos o ligados a la política de la organización.

Toda implementación del algoritmo de selección de rutas, utiliza la siguiente jerarquía al momento de tomar decisiones, pudiendo también incluir ciertos pasos intermedios.

1. Seleccionar la ruta con el mayor `LOCAL PREFERENCE`

2. Seleccionar la ruta con el `AS-PATH` que incluya la menor cantidad de ASes

3. Seleccionar la ruta cuyo `NEXT HOP` se encuentre más cercano

El algoritmo de selección de rutas comenzará siempre evaluando el atributo `LOCAL PREFERENCE`, seleccionando la ruta con mayor valor de este atributo (LR01). El valor de `LOCAL PREFERENCE` se configura de manera local, es decir, cada AS puede establecer el valor de este atributo para las rutas recibidas. Dado que `LOCAL PREFERENCE` es la primer variable tenida en cuenta por el algoritmo, y que su valor es establecido por el mismo AS, esta es la variable a través de la cual se confeccionan las políticas de ruteo. Por ejemplo, por medio de este atributo un operador podría optar una ruta con un `AS-PATH` que incluye más ASes, pero que a cambio signifique mayor rédito económico para el AS.

En caso que dos o más rutas posean el mismo `LOCAL PREFERENCE`, ya sea porque han recibido la misma ponderación o porque no han recibido ninguna (`LocalPref=0`), el algoritmo seleccionará la ruta con menor cantidad de ASes en la traza. Esto obedece a que BGP es un protocolo de vector de distancias (LR01).

Finalmente, si aún persiste la igualdad de criterios, el algoritmo seleccionará la ruta que se acceda por medio del `NEXT HOP` más cercano. Esto quiere decir, que el Sistema Autónomo enviará el datagrama por medio del camino que conduzca al punto de egreso más próximo, haciendo de esta manera el menor uso de su infraestructura (capacidad de los enlaces, memoria en ruteadores, etc.). Esta política de ruteo suele ser conocida como *Hot Potato Routing* (TSGR04).

1.2.3.3. Método *AS-PATH prepend*

Al ejecutar un anuncio, los ASes disponen de poca o nula influencia para determinar por cual ruta recibirán el tráfico entrante. Esto se debe a que el primer paso del algoritmo de selección de rutas únicamente considera el atributo `LOCAL PREFERENCE`, el cual no es fijado por quien anuncia la ruta. Sin embargo, los ASes anunciantes pueden llevar a cabo un estímulo, para aquellos casos en donde los receptores de los anuncios no apliquen `LOCAL PREFERENCE` para las rutas anunciadas. Este mecanismo de estímulo, o mejor dicho, de desaliento se conoce como *AS-PATH prepending*.

Considerando que BGP es un protocolo de vector de distancias, la técnica de AS-path prepending consiste en agregar sucesivos ASN replicados, extendiendo la longitud de `AS-PATH` de manera artificial. Esto se observa en la Figura 1.3, en donde la Universidad de Buenos Aires (AS3449), anuncia su prefijo 157.92.0.0/16 a dos de sus vecinos. Dado que en el ejemplo planteado la Universidad de Buenos Aires preferirá ser alcanzada a través de AS4270, al momento de generar los anuncios al AS3597, realizará *AS-path prepending* repitiendo sucesivas veces su ASN. De esta forma, cuando el resto de los

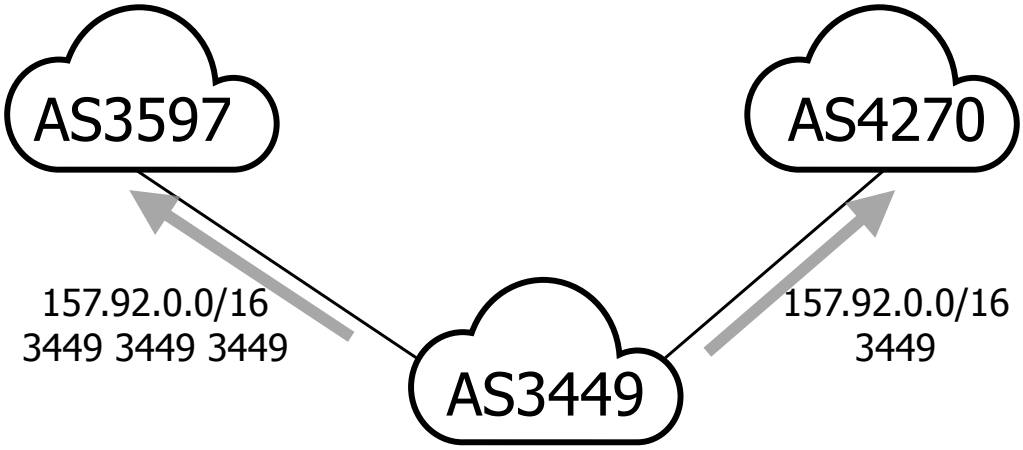


Figura 1.3: Ejemplo de técnica de desincentivo a través de *AS-PATH prepending*. El AS3449 origina el prefijo 157.92.0.0/16, y prefiere que el tráfico sea recibido por el AS4270. Para intentar lograr que se cumplan sus expectativas, el AS3449 extiende el AS-PATH en su anuncio al AS3597 replicando su ASN varias veces.

ASes de Internet reciban los anuncios del prefijo 157.92.0.0/16 tendrán dos rutas, donde la que incluye al AS4270 será la de menor longitud de camino.

1.2.4. Servidor de rutas BGP

Retomando lo explicado en la Sección 1.2.1, dos Sistemas Autónomos intercambiarán anuncios a través de sesiones BGP establecidas entre los ruteadores de borde de cada uno de los ASes. No obstante, un ruteador de borde podrá disponer de múltiples enlaces entre dominios, generando que el ruteador de borde deba establecer una sesión por cada AS vecino conectado a ese ruteador.

La creación y el mantenimiento de sesiones BGP conlleva un problema de escala en escenarios tales como los Puntos de Intercambio de Tráfico (PITs, en inglés Internet eXchange Points, IXP) (ver Secciones 1.3.3 y 3). Los IXPs reúnen en una misma ubicación física un gran número de ASes, brindando la posibilidad de que establezcan enlaces entre dominos los unos con los otros. Para tal fin, cada uno de los ASes participantes deberá colocar un ruteador de borde, y luego establecer una sesión BGP por cada uno de los ASes restantes. Sin embargo, un IXP puede reunir cientos o superar los mil ASes (CDFD+20), lo cual implicaría que cada ruteador de borde deberá establecer igual número de sesiones BGP, generado una enorme exigencia para las capacidades de hardware presentes en los ruteadores (RSF+14). Más aún, en un IXP que reúna n ASes, donde todos los ASes establecen sesiones entre sí, tendrá un total de $n \cdot (n-1)/2$ sesiones, lo cual implicaría un volumen de tráfico de orden $\mathcal{O}(n^2)$.

Un Servidor de Rutas (en inglés *Route Server*) es un hardware dedicado con el fin

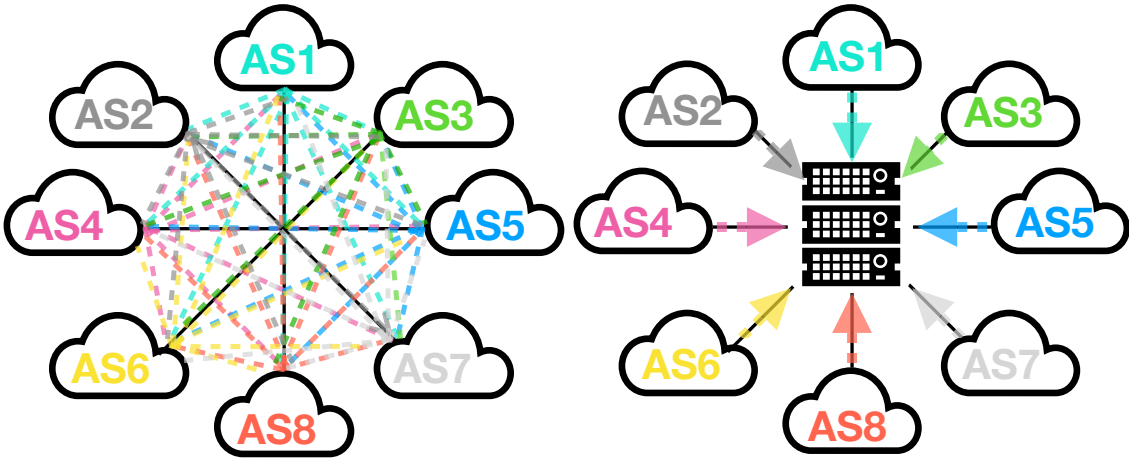


Figura 1.4: Comparación de cantidad de sesiones BGP atravesando el IXP en presencia y ausencia de un servidor de rutas. A la izquierda se muestra un escenario sin servidor de rutas, donde existe un denso número de sesiones BGP para poder comunicar los ocho ASes entre sí. A la derecha se observa un escenario con servidor de rutas donde alcanza con ocho sesiones para coordinar los anuncios de los ocho ASes presentes.

---

de reducir el impacto de tráfico y el número de sesiones en cada ruteador de borde. Ante la presencia de este servidor, los ruteadores de los ASes con presencia en el IXP realizarán sus anuncios únicamente a través de la sesión establecida con el servidor de rutas. Luego, será este servidor el encargado de, en nombre de cada uno de los ASes, redistribuir los anuncios hacia los restantes ASes. De esta manera, cada ruteador de borde pasará de mantener $n-1$ sesiones a una sola, el orden total de sesiones se reducirá de $\mathcal{O}(n^2)$ a $\mathcal{O}(n)$. Para poder cumplir con estos objetivos, y permitir que el sistema de anuncios continúe obedeciendo a la lógica del protocolo BGP, el servidor de rutas deberá contar con la capacidad suficiente para poder mantener tal cantidad de sesiones y anuncios en simultáneo. A modo de ejemplo, en la Figura 1.4 se muestra un esquema en el cual abundan las conexiones bilaterales (en el margen izquierdo) y luego se reducen sustancialmente al introducir el Servidor de Rutas (margen derecho).

1.3. Topología de ASes

Internet es el resultado de la interconexión de miles de organizaciones autónomas, donde cada una se presenta en la red con su propias necesidades y objetivos comerciales (LHD+13; CDZ97; LR01). Internet en particular, a diferencia de otras redes de telecomunicaciones predecesoras, por ejemplo la red telefónica, funciona de manera descentralizada, sin ninguna entidad que dicte el modo en el cual se estructura la red. El funcionamiento de la red está guiado por medio del protocolo BGP, el cual permite a

estas organizaciones, denominadas Sistemas Autónomos (ASes), lograr la interconexión y el alcance de todos los destinos de la red. Además, por medio de BGP se permiten incorporar variables que contemplen las necesidades específicas de cada organización (LHD+13; LR01).

La interconexión es el resultado de la necesidad fundamental con la que cuentan los ASes: deben permitir enviar y recibir el tráfico que involucra a sus usuarios (y también sus clientes). De esta forma, los ASes necesitarán interconectarse con otros, para luego poder por medio de esta interconexión anunciar sus prefijos, y por ende cursar tráfico. Los ASes también necesitarán aprender todos los prefijos existentes en la red, al igual que asegurarse que sus prefijos sean anunciados al resto de los ASes.

La interconexión de dos o más ASes se da a través de medios compartidos, los cuales interconectan ruteadores de borde de los ASes involucrados. Esta interconexión se puede dar a través de enlaces punto a punto, donde sólo se vinculan dos ASes, o por otras tecnologías de medio compartido, las cuales permiten interconectar simultáneamente múltiples ASes, tal como ocurre en los IXPs. Por lo tanto, la definición de interconexión en Internet, y específicamente a nivel de topología de ASes, requiere la presencia de elementos físicos para llevar a cabo tal interconexión. La disponibilidad o ausencia de medios físicos para llevar a cabo la interconexión con otros ASes, es lo que da lugar a celebrar acuerdos comerciales cómo parte del establecimiento de la interconexión.

El resultado de la interconexión de los ASes, y la falta de una entidad centralizada, nos motiva a conocer como es la estructura de Internet, cuáles son los motivos existentes para observar la interconexión de un par de ASes, y cómo es posible llevar a cabo esta exploración.

1.3.1. Relaciones comerciales entre ASes vecinos

La interconexión entre ASes, y por consiguiente su política de ruteo, se encuentra estrictamente regida por relaciones e intereses comerciales en los ASes que participan en la interconexión. Aunque existe una gran variedad de causas detrás de la interconexión, la literatura (LHD+13; LR01; DD10; DD08) ha clasificado los vínculos entre ASes en dos categorías: *cliente-proveedor* (en inglés *customer-to-provider relationship*, c2p) y *vínculos entre pares* (en inglés *peer-to-peer relationship*, o también conocido como *peering*, p2p).

Una relación c2p supone la existencia de un AS cliente y un AS proveedor, en donde el cliente le paga una tarifa al proveedor por el tráfico intercambiado entre los dos. A cambio, el cliente obtiene acceso a la red del proveedor, la cual incluye a su vez, los proveedores del proveedor. Estas relaciones, las cuales efectivamente existen, se basan sobre la hipótesis de la existencia de una estructura de red jerárquica. Ante esta premisa

(a) Propagación de anuncios bajo relaciones cliente-proveedor (c2p)

(b) Propagación de anuncios bajo relaciones entre pares (p2p)

Figura 1.5: Comparación de la propagación de prefijos en escenarios con vínculos cliente-proveedor (c2p) (Figura 1.5a) y vínculos entre pares (p2p) (Figura 1.5b).

los clientes aceptan la tarifa de sus proveedores, ya que estos disponen de redes con mayor visibilidad, es decir, cuentan con mayor número de rutas hacia su AS. Además, es presumible, que los proveedores dispongan de coberturas geográficas superiores a la de sus clientes, permitiendo, alcanzar otros clientes y otros proveedores. Por ejemplo, un proveedor de Internet de una pequeña ciudad del interior de la Argentina establece un vínculo c2p con un proveedor regional, que a su vez cuenta con una relación c2p con un proveedor nacional. Bajo el escenario planteado, donde existe una jerarquía, es plausible suponer que la mayoría (por no incluir a todos) de los ASes, tendrán al menos un vínculo c2p.

En contrapartida, una relación p2p le brinda a un AS el acceso a la red de su contraparte, sin que ninguno de los dos ASes exijan una tarifa por el tráfico intercambiado a través de los enlaces que conforman la interconexión. La ausencia de una tarifa se debe a que ambas partes consideran beneficiosa la interconexión, no pudiendo establecer una relación de preponderancia de un AS por sobre el otro. Por ejemplo, en un gran número de acuerdos p2p, aunque no en todos, el beneficio mutuo se observa por medio de flujos de tráfico simétricos entre los ASes (Ste11). Los acuerdos p2p no sólo permiten el alcance a la red de la contraparte, sino que también son extensivos en el acceso a las redes de sus clientes, pero excluyendo los proveedores de la contraparte.

1.3.1.1. Relaciones entre ASes en anuncios BGP

Comenzando por las relaciones c2p, el proveedor debe ofrecerle mayor visibilidad a sus clientes, y también acceso a sus proveedores. Para cumplir tales requisitos un proveedor deberá anunciar a sus proveedores **todas** las rutas aprendidas de sus clientes. Dado que todos los ASes (o más precisamente casi todos) contarán con al menos un enlace c2p, los sucesivos anuncios a través de los niveles jerárquicos de la red, darán como resultado el alcance global de cada un los de ASes conectados a la red.

De manera diferente, ante relaciones p2p, los ASes involucrados anunciarán las rutas aprendidas de sus pares **sólo** a sus clientes. Esta restricción es conocida como *valley-free*, donde un cliente jamás le anunciará a un proveedor rutas que no brinden un rédito económico concreto. Por lo tanto, las rutas aprendidas en relaciones p2p, las cuales no son tarifadas, no serán anunciadas a los proveedores, quienes tarifan el tráfico, ya que esto significaría egresos económicos, sin a cambio ninguna clase de egresos. Tampoco estas rutas serán anunciadasa otros ASes con los cuales se tiene relación p2p, ya que en tal caso intermediaría entre sus dos pares, brindando su infraestructura para tal fin, y sin recibir ninguna retribución económica. La selectividad y restricción de las rutas aprendidas hará que las relaciones p2p de los clientes sean invisibles desde los proveedores, y eventualmente por quienes reciban los anuncios de los proveedores. Luego, en la Sección 1.3.5 nos detendremos a analizar las consecuencias aparejadas con la discrecionalidad de los anuncios.

También se debe mencionar incluir la naturaleza comercial de estas relaciones c2p y p2p en el algoritmo de selección de rutas de BGP. Tres escenarios se derivan de las dos relaciones posibles: 1) El AS es proveedor y cobra por el tráfico intercambiado 2) El AS es un par por lo cual no existe rédito económico 3) El AS es cliente por lo cual debe pagar por el tráfico intercambiado. El orden en el cual se presentan los escenarios es el orden de preferencia en el cual se pretende seleccionar una ruta. Es decir, un AS siempre preferirá una ruta a través de un cliente que cualquier otra alternativa. En segunda instancia, un AS preferirá una ruta a través de un par (sin tarifa) por sobre una ruta que incluye un proveedor (debe pagar por el uso). Para que esto sea tenido en cuenta en el proceso de selección de rutas, los operadores deberán asignar `LOCAL PREFERENCE` de la siguiente manera

$$LocalPref(cli) > LocalPref(p2p) > LocalPref(prov) \tag{1.1}$$

Por último, presentaremos el concepto de *Cono de Clientes* (en inglés Customer Cone). Este concepto es derivado de la aplicación de las políticas de ruteo. Primero, retomemos la hipótesis de que Internet es una red jerárquica, donde los proveedores se conectan a otros proveedores de mayor envergadura. Sobre esta base, un proveedor anunciará a sus vecinos todas las rutas aprendidas de sus clientes. Por lo tanto, dado que un proveedor anuncia las rutas de sus clientes a sus proveedores, y luego este a sus proveedores y así sucesivamente, esto creará que los proveedores de mayor jerarquía anuncien las rutas de los clientes de sus clientes, de manera recursiva, hasta llegar a los estratos de mejor jerarquía. Entonces, el *Cono de Clientes* será el conjunto de ASes a los cuales un AS puede acceder a través de los sucesivos enlaces de proveedor a cliente que

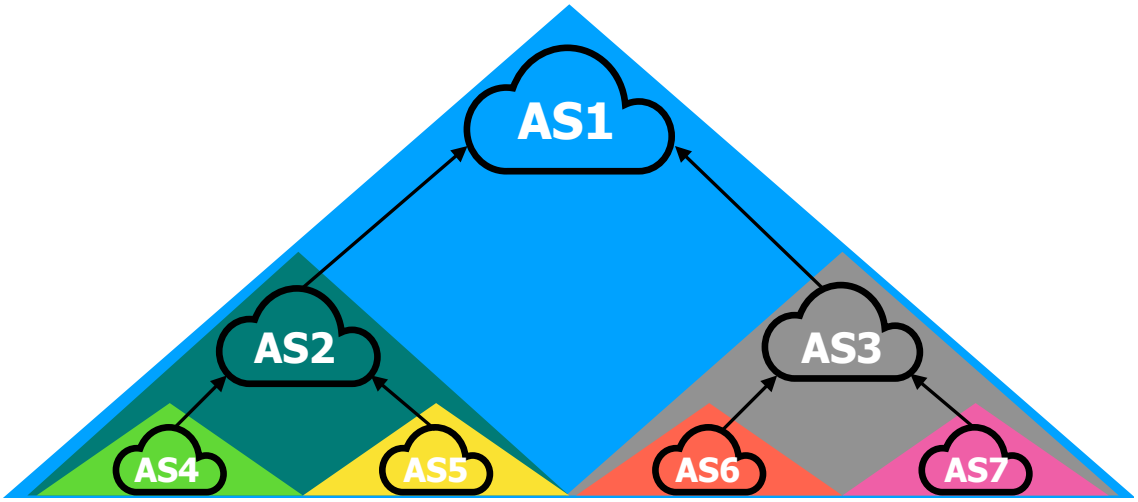


Figura 1.6: Ejemplo de relaciones c2p entre ASes y los conos de clientes resultante.

anuncian su clientes.

Para brindar mayor claridad, la Figura 1.6 presenta una pequeña topología de ASes, donde se muestran los vínculos (c2p) por medio de fechas y los correspondientes *conos de clientes* a través de figuras coloreadas. En este ejemplo, podemos ver como AS4 y AS5 tienen una relación c2p con AS2, al igual que AS6 y AS7 con AS3. Dadas estas relaciones, AS4 y AS5 pertenecen al *conos de clientes* de AS2, mientras que AS6 y AS7 al de AS3. Luego, AS2 es cliente AS1, al igual que AS3, por lo tanto ambos clientes deberá anunciar todos los rutas que aprendieron por medio de sus respectivos clientes. De esta manera, AS1 contará, por medio de sus clientes, con rutas p2c para alcanzar a AS4, AS5, AS6 y AS7, por lo que su *conos de clientes* estará conformado por AS2, AS3, AS4, AS5, AS6 y AS7.

1.3.2. Clasificación de ASes según su rol

Las relaciones entre ASes anteriormente presentadas son suficientes para llevar a cabo estudios enfocados en la estructura de Internet, su desempeño y evolución. También, las relaciones entre ASes nos permiten identificar el alcance de las rutas dependiendo de la naturaleza comerciales de los enlaces que atraviesa el tráfico. Aunque las relaciones c2p y p2p permiten una representación total de la relación entre ASes, es decir entre dos pares de ASes, estas no proveen categorías a los ASes de forma asilada, es decir por fuera de su vínculo con otro. Sin embargo, los ASes son frecuentemente identificados por adjetivos los cuales describen sus características. Mas aún, en un gran número de oportunidades los adjetivos derivan del conjunto de relaciones que tiene un AS con los demás.

Dada esta taxonomía *de-facto* existente para los ASes, a continuación presentaremos

las más destacadas. Es importante remarcar que la existencia categorías superpuestas y definiciones poco precisas.

- **Stub AS**: Así se define a un AS cuando no cuenta con ninguna relación c2p en la cual sea el *proveedor*.
- **AS de tránsito**: Un AS de tránsito es aquel AS que anuncia prefijos que no han sido originados en su Sistema Autónomo.
- **ASes de eyeballs**: Esta denominación se utiliza para los ASes en donde se encuentran los consumidores del contenido disponible en Internet (web, email, video).
- **ASes de acceso**: Estos son los ASes por los cuales los usuarios acceden a la red, por ejemplo los ASes correspondientes a los ISPs que ofrecen suscripciones domiciliarias mediante tecnologías de ADSL, DOCSIS o FTTH.
- **Proveedores de Contenidos**: Son aquellos ASes delegados a las compañías dedicadas a la prestación de servicios y la generación de contenido. En particular un AS se define como proveedor de contenidos cuando el contenido generado por la empresa es servido desde el AS en cuestión.
- **TIER-1**: La definición de ASes en el TIER-1 ha sido siempre utilizada a lo largo de la historia de Internet, pero no cuenta con una definición estricta. Específicamente, el TIER-1 esta conformado por el conjunto de ASes con la máxima jerarquía posible dentro de la red de ASes. Así como no existe una definición de TIER-1, tampoco es posible definir una única métrica para evaluar la jerarquía de los ASes, y así poder definir la máxima jerarquía. Nosotros adoptaremos la definición de Luckie *et al.* ([LHD+13]), donde el TIER-1 esta compuesto por el *cliqué de tránsito libre* (en inglés transit-free clique). Los integrantes de este conjunto serán aquellos ASes que por estar al tope de la jerarquía, no serán clientes de ningún AS. Además estos ASes contarán mayoritariamente con vínculos c2p con sus clientes, y tendrán vínculos p2p casi exclusivamente con los otros integrantes del *cliqué de tránsito libre*.
- **Sibling ASes**: Las organizaciones se identifican en los anuncios BGP por medio de su ASN, llevando a que cada organización necesite contar con su ASN propio. Por más que un único ASN es suficiente para el propósito de la empresa en BGP, un gran número de organizaciones cuenta con más de un ASN, donde en casos esta cifra asciende a decenas ([CHKW10]). En estos casos, cuando múltiples ASes pertenecen a una organización se denomina que estos ASes son *siblings* entre sí.

- **upstream y downstream ASes**: Upstream y downstream son denominaciones relativas entre ASes pertenecientes al mismo cono de clientes. Dado dos ASes dentro de un mismo cono de clientes, se denominará *upstream* a aquel que pueda acceder al otro extremo mediante una ruta que no incluya proveedores. En cambio, *downstream* identificará a aquellos ASes en el cono de clientes que requerirán utilizar enlaces hacia sus proveedores para acceder al otro extremo.

1.3.3. IXPs en el ecosistema de Internet

Los ASes siempre buscarán minimizar los egresos debidos a gastos operativos, y a su vez maximizar, según el caso, los ingresos o la calidad del servicio prestado (DD10). Con el objetivo de reducir los costos operativos, los ASes tienden a establecer el mayor número de relaciones p2p posibles, ya que de esta forma intercambian de tráfico libre de tarifas (ver Sección 1.3.1). Puntualmente, este ahorro se debe a que los enlaces p2p brindan rutas alternativas (sin costo) de las provistas por los proveedores de tránsito (con costo).

Los Puntos de Intercambio de Tráfico (PITs, del inglés Internet eXchange Points, IXP) son infraestructuras físicas que posibilitan la interconexión tanto pública como privada, donde pondremos énfasis en describir su la interconexión pública. Los ASes acceden a la infraestructura pública de los IXPs mediante el pago de una tarifa fija. A cambio los ASes obtienen acceso a nivel físico, es decir acceden a un medio compartido, al igual que el resto de los ASes miembros del IXP (ACF+12). Esta posibilidad de acceso físico les concede la posibilidad de intercambiar tráfico, y luego de haber establecido acuerdos de intercambio, con el resto de los miembros.

Específicamente a nivel de infraestructura, los IXPs brindan un conmutador de capa de enlace (en inglés switch), donde cada uno de los miembros es conecta un ruteador de borde a un puerto. Luego por medio de esta conexión, cada miembro podrá establecer sesiones BGP con el resto de los miembros, o con un servidor de rutas (ver Sección 1.2.4), en caso de que el IXP cuente con uno. Una vez establecidas las sesiones y anunciados los prefijos, los ASes intercambiarán tráfico a través del conmutador del IXP.

Los IXPs se presentan como una manera muy efectiva de establecer vínculos p2p, en particular desde el punto de vista de la infraestructura necesaria. Los ASes sólo requieren de un enlace para arribar al IXP e interconectarse con el resto de los miembros, mientras que ante la ausencia de IXPs necesitarían un enlace físico por cada uno de los ASes con los cuales desean interconectarse. Esto hace que los IXPs se presentan como una alternativa sencilla y económica a través de la cual se pueden establecer vínculos no tarifados con otros ASes. Esta significativa ventaja ha llevado a un rápido incremento del número

de los IXPs durante la década de los 2000s (DD10). Más aún, no sólo el número de IXPs ha crecido drásticamente, sino que también el número de miembros conectados y la cantidad de relaciones establecidas en ellos. Ya en 2012 Ager *et al.* (ACF+12) reportó más de 50.000 enlaces p2p presentes en un sólo IXP europeo.

La proliferación de IXPs y la oportunidad de establecer vínculos p2p en ellos generó a partir de los años 2000 un proceso de transformación en la red conocido como *aplanamiento* (en inglés flattening). Esta denominación surge debido al marcado aumento de los vínculos p2p reformó la antigua estructura jerárquica de la red, en una red más plana. Una serie de eventos fomentaron el proceso de *aplanamiento*. En primer lugar, a partir de los años 2000, una creciente fracción del tráfico de Internet comenzó a provenir de un número reducido de proveedores de contenidos y redes de distribución de contenidos (Google, YouTube, Facebook, Netflix, Akamai). Este suceso fue principalmente motorizado por el aumento de la adopción del video a través de HTTPS como reemplazo de las descargas a través de mecanismos de descarga peer-to-peer (BitTorrent, eMule). En segundo lugar, las redes de distribución de contenidos llevaron a cabo un despliegue global, ubicándose frecuentemente en IXPs locales (GALM08).

1.3.4. Redes de Distribución de Contenido (CDNs)

Las Redes de Distribución de Contenido (Content Delivery Networks, CDNs) consisten en estructuras descentralizadas desde las cuales se sirve el contenido a los usuarios (CFH+13; HWLR08). Desde estas plataformas, los proveedores de contenidos hacen disponible una gran variedad de contenido que incluye objetos web, imágenes, archivos, paquetes de software, y principalmente video.

Estas estructuras están compuestas por decenas o miles de servidores distribuidos en múltiples locaciones, incluyendo colocaciones en IXPs e incluso dentro de ISPs (HWLR08). Esta colección de servidores funciona de forma cooperativa y sincronizada, de manera tal de colocar en contenido en la ubicación óptima. Aunque la optimización puede incluir un gran número de variables, el objetivo primordial es hallar un punto desde el cual se sirva al usuario con mayor calidad de experiencia posible. Por ejemplo, a través de algoritmos sofisticados, las CDNs determinan dónde se encuentra el servidor que minimiza la latencia al usuario que generó la solicitud (CFH+13; HWLR08).

Otro componente central el sistema de distribución de contenido es una red de servidores de nombres (en inglés Domain Name Servers (Moc87)) autoritativos. Los usuarios peticionarán el contenido por medio de *Fully Qualified Domain Names (FQDNs)*, cuyos nombres deberá ser traducidos a direcciones IP para poder descargar el elemento

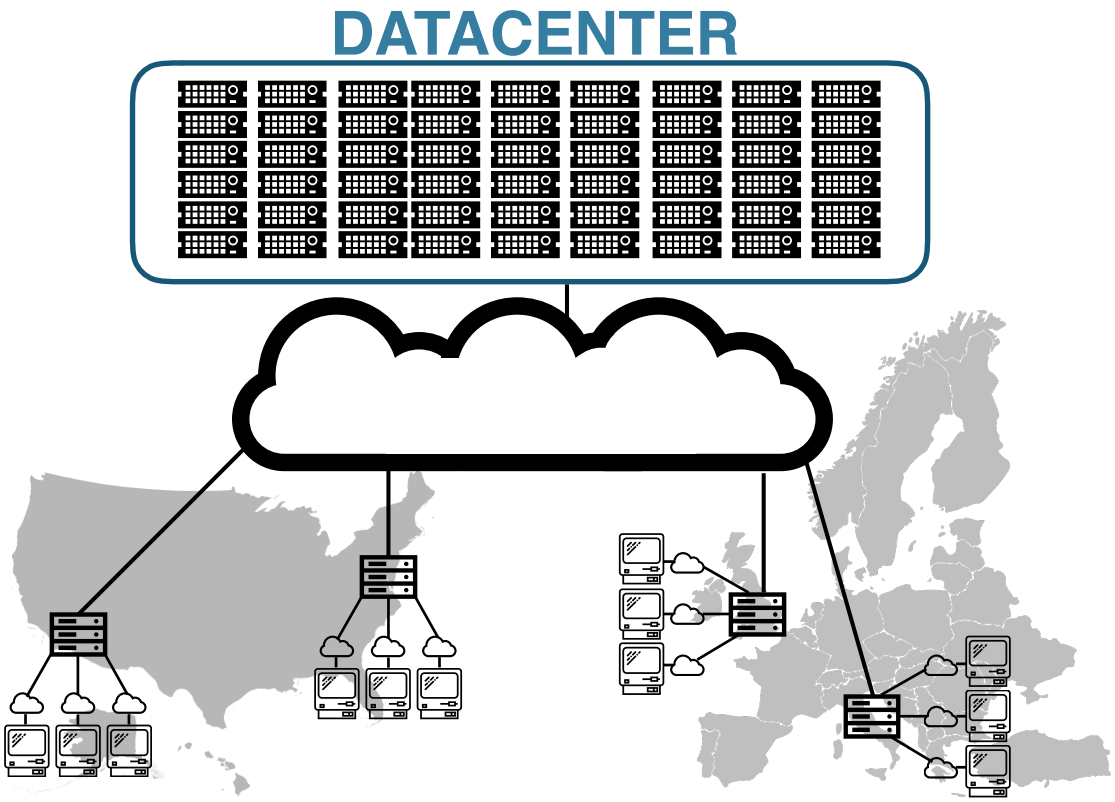


Figura 1.7: Diagrama de arquitectura de una Red de Distribución de Contenido. La réplica primaria del contenido se encuentra en el datacenter, desde donde se generan copias servidores distribuidos en diferentes ubicaciones en la proximidad a los usuarios. Los usuarios mejorarán la calidad de experiencia al reducir la latencia debido al acceso a servidores cercanos.

---

solicitado. Las CDNs aprovecharán el momento de la traducción para en este momento introducir la dirección IP del servidor destinado a servir la petición de usuario. Dependiendo de la ubicación del usuario, y de otras variables con las que la CDN evalúa la calidad, se determinará la mejor ubicación posible.

Por ejemplo, la Figura 1.7 presenta un ejemplo de una CDN que dispone de réplicas en cuatro ubicaciones: dos en EE.UU y dos en Europa. Probablemente, los usuario que soliciten el contenido desde EE.UU. serán direccionados hacia la réplica más cercana en este país, mientras los usuarios en suelo europeo accederán a réplicas en este continente.

De este funcionamiento se desprenden algunas conclusiones. En primer lugar, las CDNs ubican el contenido en réplicas, de manera tal de que ese contenido este disponible cerca de su base de consumidores (HBvR+13). En segundo lugar, las CDNs redireccionan las consultas de forma automática y transparente, evitando que los usuarios deban lidiar con el proceso de selección.

Finalmente, es necesario destacar que el método de selección de réplicas por medio de DNS no es el único disponible, y que redes menos sofisticadas emplean la distribución

a través de *Anycast* (CFKB+15).

La distribución a través de *Anycast* consiste en que cada una de las réplicas del contenido operen con la misma dirección IP, haciendo que cada una de las ubicaciones de la CDN anuncie em BGP el mismo prefijo. Esta multiplicidad de anuncios en múltiples destino generará, que el resto de los ASes reciba múltiples rutas hacia los servidores de la CDN. Sin embargo, la selección del servidor utilizado surgirá del proceso de selección de rutas de BGP a cargo del AS.

1.3.4.1. Arquitecturas de CDNs

La arquitectura de CDNs de divide a través de dos paradigmas: *enter deep* y *bring home*.

El paradigma *enter deep* plantea ubicar servidores de contenido dentro de los ISPs o tan próximo como se pueda, por ejemplo, colocado junto al ruteador de borde. Este paradigma busca que el contenido sea servido desde ubicaciones sumamente próximas (en términos de distancia física) a los usuarios, reduciendo la latencia y mejorando la tasa de transferencia de datos. Para lograr este cometido, la CDN deberá establecer servidores en cientos o miles de ISPs a lo largo del planeta, creando un esquema altamente distribuido. La contrapartida de estos beneficios será el aumento de la complejidad operativa de una red dispersa.

En cambio, el paradigma *bring home* busca crear grandes colecciones de servidores en pocas ubicaciones precisamente seleccionadas. Estas ubicaciones pueden ser por ejemplo grandes IXPs o sitios donde se dispone de una alta conectividad. A través de esta política, las CDNs pueden tener acceso a un gran número de ASes miembros del IXP desde una sola ubicación, reduciendo así los costos operativos.

1.3.4.2. Motivaciones a construir una CDN

Existe un gran número de razones por las cuales un proveedor de contenido se podría decidirse a llevar a cabo el despliegue y operación de una CDN. A continuación presentaremos los motivos planteados por Gill *et al.* (GALM08).

1. **Reducir dependencias**: Los proveedores de contenidos pueden estar interesados en no depender de terceros en la distribución. Esto se debe a que los proveedores de contenidos se verán también afectados en caso de que los intermediarios afronten dificultades técnicas o económicas.

2. **Reducir costos**: Al construir una CDN, los proveedores de contenido establecen relaciones p2p con los ISPs, lo que implica que será tráfico no tarifado. Esto implica

una ventaja mutua. Los proveedores de contenido no pagarán costos operativos por el tráfico cursado hacia proveedores por distribuir su tráfico. En el caso de los ISPs, al recibir el tráfico por medio de enlaces p2p, este será tráfico que no ingresará por medio de sus proveedores, lo que también redundará en un beneficio económico para éstos.

3. **Control de la calidad de experiencia**: El hecho de contar con intermediarios en el proceso de entrega del contenido hace que los proveedores tengan escasos recursos para controlar la calidad de la experiencia. Al servir el tráfico directamente desde sus CDN y por enlaces p2p, los proveedores de contenido podrán controlar precisas políticas de balances de carga.

1.3.5. Proyectos de medición de Internet

La coordinación del ruteo en Internet se hace de forma completamente descentralizada a través de BGP. Esto quiere decir, que la única forma para descubrir la topología, o más precisamente las rutas, es por medio de los anuncios BGP provenientes de los ASes adyacentes. Esto funcionamiento descentralizado y asincrónico lleva a que no exista ninguna entidad que cuente con la recopilación o control total de los anuncios del sistema.

Para llevar a cabo investigaciones centradas al estudio de las dinámicas presentes en la topología de Internet, se requiere de datos que permitan reconstruir la topología de Internet. En este caso en las rutas presentes en las tablas de ruteo BGP se presentan como la fuente principal de datos, ya que a través de éstas es posible reconstruir la interconexión de los ASes. Por lo tanto, para llevar a cabo las investigaciones se necesita disponer de tablas de ruteo.

Con el objetivo de adquirir tablas BGP, la primera alternativa que se puede pensar es que los ASes compartan voluntariamente sus tablas. Esta alternativa es no es factibles, ya que las tablas BGP incluyen secretos comerciales de los ASes. En algunas oportunidades los ISPs comparten sus tablas BGP con investigadores, asegurándose bajo acuerdos de confidencialidad la preservación de los datos y de información sensible.

Otra alternativa, y las más popular hoy en día en Internet, es crear ASes con fines científicos, que al operar recibirán anuncios BGP. Los anuncios recopilados serán luego compartidos de forma abierta a la comunidad, que a través de estos podrá estudiar la topología de Internet. Varios proyectos que recopilan tablas BGP a través de esta modalidad, pero en particular se destacan el proyecto Routeviews (oO19) de la Universidad de Oregon y RIPE RIS (NCC19b) perteneciente al RIR con competencia en Europa, Medio Oriente y Asia Central. Estos proyectos simplemente establecen sesiones BGP con ASes vecinos, de los cuales reciben los anuncios. Tanto Routeviews como RIPE RIS aseguran

a sus vecinos que el tráfico inyectado en sus redes será nulo, incentivando a sus vecinos a establecer sesiones con estos, ya que no representarán un costo por tráfico.

Las políticas de anuncios en BGP cuentan con una singularidad derivada de las relaciones comerciales entre ASes. Los ASes no cuentan con ningún incentivo económico de anunciar a sus proveedores las rutas aprendidas por medio de sus relaciones p2p con otros ASes (ver Sección 1.3.1). Esta selectividad en los anuncios hará que los enlaces p2p sean invisibles por fuera de los conos de clientes de los ASes involucrados en tal relación, generando opacidad en el sistema.

La opacidad genera que la topología derivada de la recolección de rutas sea incompleta, ya que hay rutas que no se anuncian. Más aún, dado que los vínculos p2p se observarán dentro del cono de clientes, al tomar aleatoriamente dos ASes, si estos se encuentran en conos de clientes no superpuestos, estos observarán otros vínculos p2p. La conclusión que se desprende es que la topología de Internet cambiará según desde que AS se la observe.

Los proyectos de recopilación de tablas BGP tienen el objetivo de poder capturar la topología de Internet de la manera más precisa y completa posible. Entonces, para lidiar con el problema de la opacidad, los proyectos de recopilación buscan colocar ruteadores de borde en el mayor número de ubicaciones posible. De esta forma, al aumentar la cobertura probablemente se incrementará el número de ASes con los que se establecen sesiones BGP, y por lo tanto maximizando el número de enlaces p2p capturados.

1.4. Mediciones activas en Internet

Las mediciones activas en Internet son aquellas que requieren la inyección de tráfico artificial para obtener información de la red. Estos experimentos basados en el envío de paquetes se enfocan mayoritariamente en adquirir métricas elementales: latencia, pérdida de paquetes, reordenamiento de paquete y recopilación de la topología (Luc10).

Las mediciones activas dan lugar básicamente en cuatro áreas de trabajo: topología de Internet, ruteo, performance y pruebas de carga (cla99). El área de topología de Internet se aboca a develar la estructura y el comportamiento de la red, por ejemplo recopilar rutas y medir su latencia. Los estudios de ruteo se centran en analizar las dinámicas presentes en los protocolos, tales como las actualizaciones y cambios en las tablas de ruteo. Los trabajos enfocados en performance analizan métricas que provean información de estado de la red entre dos destinos, de manera tal de poder diagnosticar fallas o proponer mejoras. Las pruebas de carga tiene como objetivo poder determinar la utilización de la infraestructura de la red, por ejemplo el porcentaje de utilización de la capacidad de los enlaces o las colas de espera en ruteadores y conmutadores.

Del amplio espectro de las mediciones activas, nosotros nos enfocaremos en particular de las mediciones basadas en la inyección o recolección de paquetes ICMP (Pos81a).

1.4.1. Mediciones activas a través de ICMP

El Protocolo de Mensajes de Control de Internet (del ingles Internet Control Message Protocol, ICMP) (Pos81a) , es un protocolo complementario al protocolo IP, utilizando por dispositivos finales (end-hosts) y ruteadores para el envío información de capa de red.

Este es un protocolo simple, que consta con un escaso número de mensajes posibles predeterminados en la especificación del protocolo. Los mensajes son definidos a través de los campos `tipo` y `código` de la cabecera del protocolo. La definición de la estructura de los mensajes ICMP también contiene un campo para la `suma de verificación` (en inglés checksum) y posibilidad de agregar datos útiles (en inglés payload), aunque es raramente necesario.

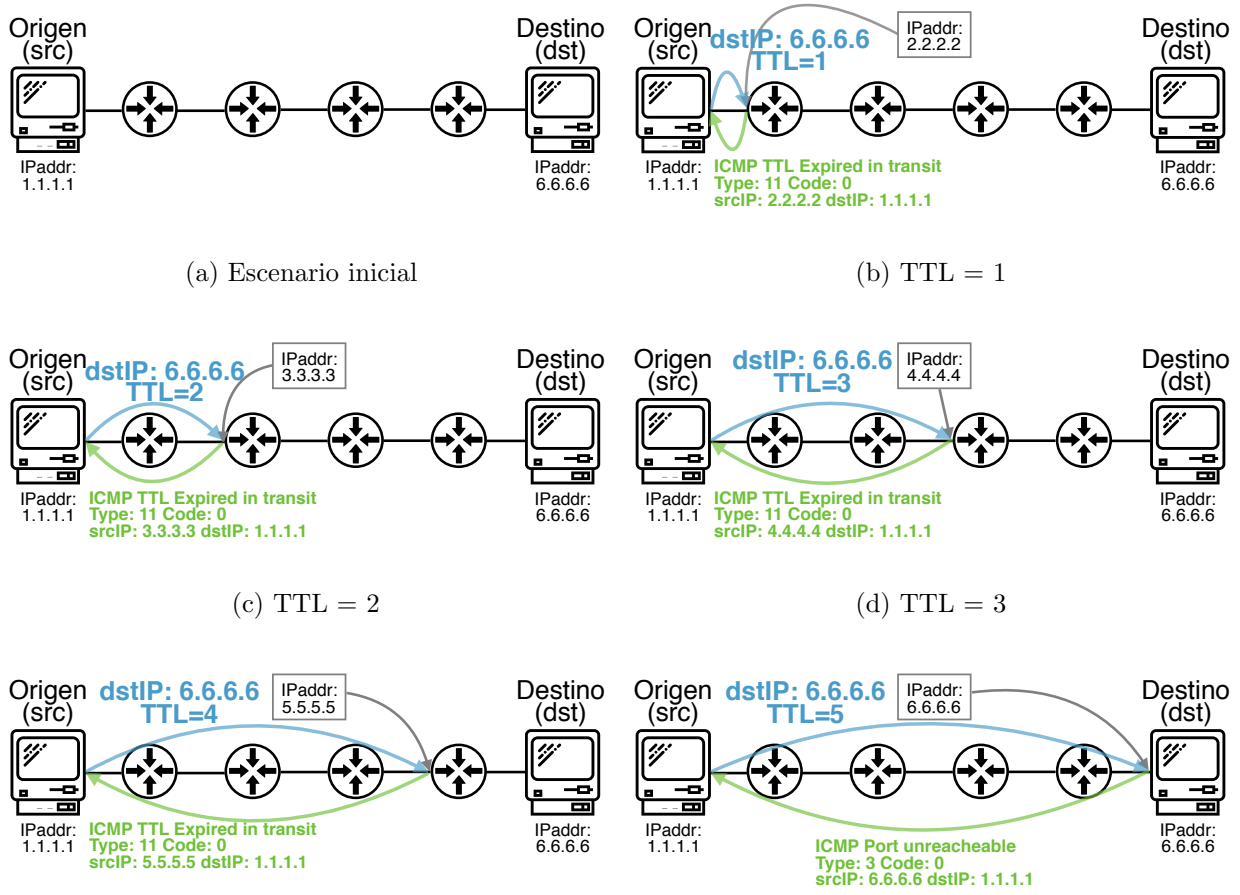
Usuarios, operadores e investigadores hacen uso frecuentemente ICMP para detectar y reportar errores en la comunicación. Sin embargo, la inyección de los mensajes ICMP en la red puede adicionalmente proporcionar mediciones de latencia, rutas (Jac89), uso de la capacidad (Dow99), congestión (DCGG+18) y otras características.

1.4.1.1. Ping

`ping` (Muu83) es una aplicación sumamente popular, presente en cualquier sistema operativo. Su propósito es dotar de un sonar o radar a los end-hosts, de manera de poder detectar la presencia y distancia del resto de los dispositivos en la red. El funcionamiento se basa en el intercambio de mensajes ICMP `echo request (tipo: 8, código: 0)` y `echo reply (tipo: 0, código: 0)`. La aplicación envía un mensaje `echo request` a una dirección IP objetivo específica, quien responderá la solicitud mediante un mensaje `echo reply`. Este intercambio de mensajes permite elaborar las siguientes conclusiones. Si el dispositivo objetivo esta disponible (encendido, conectado a la red, etc.) el solicitante obtendrá una respuesta. Además, la aplicación ping medirá el tiempo transcurrido entre la petición y la respuesta, a efectos de poder inferir la distancia a la que se encuentra el objetivo. En cambio, si el peticionante no obtiene una respuesta, podrá suponer que el dispositivo objetivo se encuentra inalcanzable.

1.4.1.2. Traceroute

Traceroute (Jac89) es otra aplicación popular basada en el uso de mensajes ICMP. Su objetivo es hallar las direcciones IP de las interfaces en los ruteadores del camino hacia



(a) Escenario inicial

(b) TTL = 1

(c) TTL = 2

(d) TTL = 3

(e) TTL = 4

(f) TTL = 5

Figura 1.8: Secuencia paso a paso del funcionamiento de un traceroute. El origen (1.1.1.1) utiliza el traceroute para descubrir las interfaces en el camino hacia el destino (6.6.6.6). Las figuras 1.8b, 1.8c, 1.8d, 1.8e y 1.8f muestran como los paquetes expiran producto de limitar el TTL en origen.

una dirección IP objetivo. Esta aplicación es frecuentemente utilizada por operadores de red e investigadores para poder diagnosticar problemas e inferir la topología de la red (ACO+06).

La herramienta generará datagramas IP con el objetivo que sea descartados antes de alcanzar el destino a causa de que su TTL alcanzó cero. Esta búsqueda se debe a que los ruteadores frecuentemente reporta al origen por medio de un mensaje ICMP cuando el paquete es descartado por una falla de TTL. Este reporte generará una interacción entre el origen y los ruteadores a lo largo de la traza hacia el destino.

El protocolo IP dispone del campo `TTL` para descartar datagramas que a causa de potenciales problemas de ruteo, han atravesado un gran número de de ruteadores en su camino. Para poder identificar que el paquete se vio afectado por una falla, los ruteadores al enviar un paquete decrementaran el TTL en uno, y lo descartarán en caso de que este contador arribe a cero. Aprovechando esta medida de protección del protocolo IP, el traceroute podrá precisar la distancia (en número de saltos) de la cual busca obtener la dirección del ruteador.

Para comprender la lógica de funcionamiento del traceroute utilizaremos la Figura 1.8, en la cual se explica paso a paso como la herramienta descubre las interfaces de los ruteadores en el camino desde el origen (1.1.1.1) hacia el destino (6.6.6.6).

El proceso comienza indicando la dirección de destino, en este caso 6.6.6.6, en el dispositivo origen (Figura 1.8a). Dado que el protocolo IP es *end-to-end*, el origen desconocerá las direcciones a lo largo del camino hacia el destino.

Para descubrir las direcciones IP en el camino hacia el destino, traceroute comenzará enviando un datagrama IP cuyo campo `TTL` portará el valor 1 (Figura 1.8b). Cuando el dispositivo origen envíe el datagrama, al arribar al primer ruteador este decrementará el TTL, alcanzando el valor 0, siendo descartado. Dependiendo de la configuración del ruteador, éste podrá enviarle un mensaje ICMP la origen (dispositivo con dirección 1.1.1.1), indicando que el paquete *"expiro en tránsito"* `(tipo: 11, código: 0)`. Entonces, al recibir el mensaje ICMP, en el origen podrá extraer la dirección IP del ruteador donde expiró el paquete del datagrama en el cual se encuentra encapsulado el mensaje ICMP. Más precisamente, el datagrama contendrá la dirección de la interfaz donde el paquete expiró. Finalizado esta primera iteración, el dispositivo origen (1.1.1.1) sabrá que en su camino hacia 6.6.6.6 tendrá la dirección 2.2.2.2 a distancia 1.

Este proceso se repetirá sucesivas veces aumentando el TTL de a 1, como se observa en la Figuras 1.8c, 1.8d, 1.8e, 1.8f, hasta alcanzar el destino. Al alcanzar el destino, este responderá con un mensaje ICMP *puerto inaccesible* `(tipo: 3, código: 0)`, siempre y cuando los datagramas enviados por el origen envíen datagramas UDP dirigidos a puertos destinos de alta numeración (LHH08).

1.4.1.3. Aumento de la precisión a través de paris-traceroute

Aunque algoritmo del traceroute es capaz de develar rutas, cuenta con imprecisiones frente a *balanceadores de carga* a lo largo del camino.

El *balance de carga* es una técnica por la cual se habilita el uso activo de múltiples rutas que conducen a un mismo destino. La distribución del tráfico en múltiples rutas permitirá un uso de toda la infraestructura disponible, generando un aumento de la capacidad agregada y una disminución de la utilización en cada enlace. Una amplia variedad de protocolos de ruteo permiten efectuar balance de carga, por ejemplo OSPF (Moy98), bajo la modalidad *múltiples camino del mismo costo* (en inglés Equal Cost Multipath, ECMP). En particular, las técnicas de balance de carga se dividen en tres categorías: por paquete, por flujo y por política de destino.

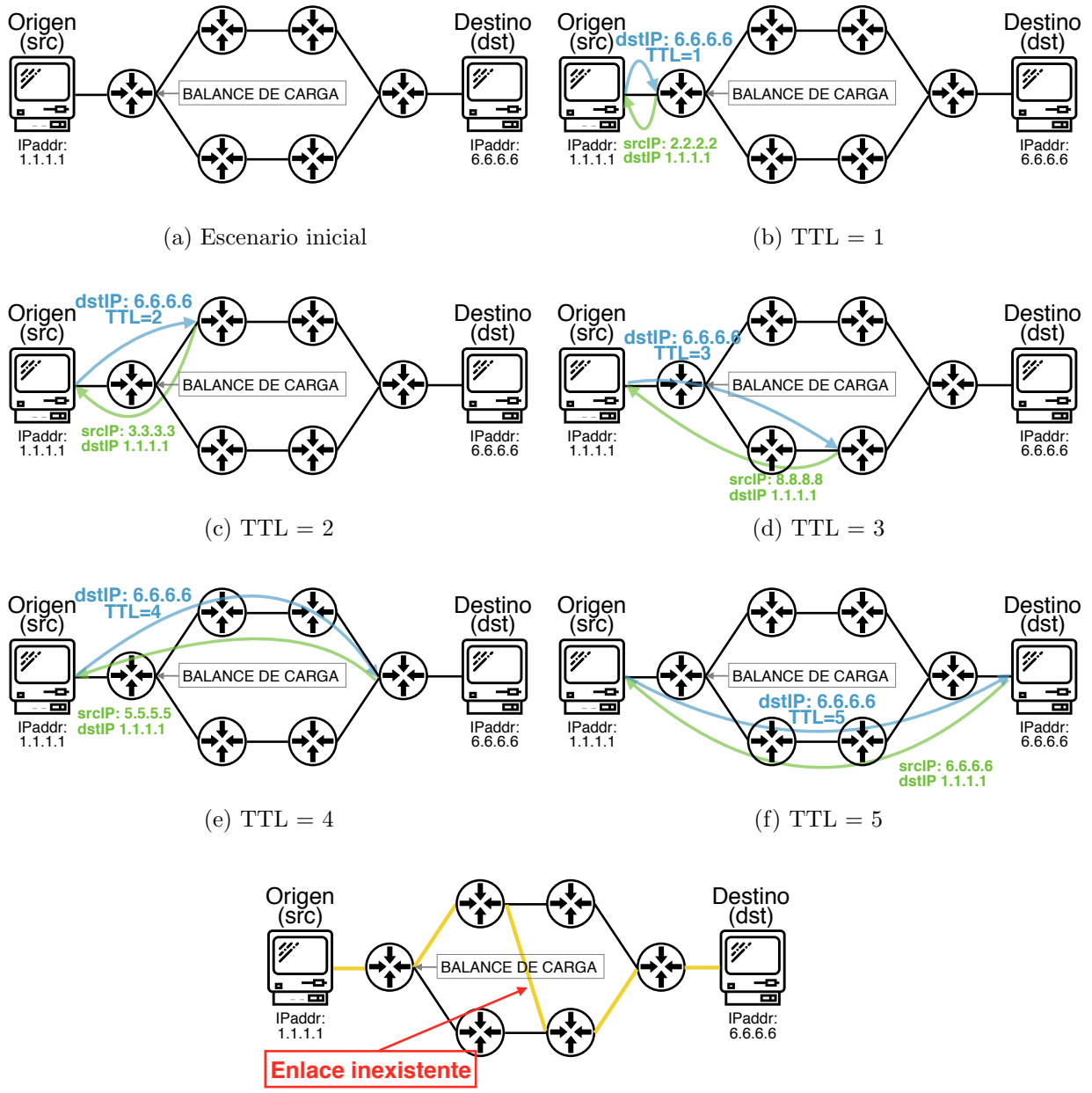
La imprecisión del traceroute se debe a que el procedimiento utiliza una series de paquetes, más precisamente un paquete por cada salto a descubrir. Por su parte, los balanceadores eligen un camino por cada uno de los paquetes recibido. Entonces, esta decision paquete a paquete posibilitará que cada paquete generado por el traceroute sea enviado por un camino diferente.

La Figura 1.9 muestra un ejemplo de las imprecisiones a las que se puede enfrentar un traceroute ante la presencia de un balanceador de carga. En el escenario inicial que se presenta en la Figura 1.9a, es similar al del ejemplo en la Figura 1.8a, salvo qué en este nuevo caso se presentan dos rutas hacia el mismo destino. Como observamos en la secuencia de Figuras 1.9b, 1.9c, 1.9d, 1.9e y 1.9f, luego de atravesar el balanceador, cada paquete alterna la ruta hacia el destino. Esto generará una incosistencia, tal como se ve en la Figura 1.9g, ya que supone la existencia de un enlace (*fake link*) entre los ruteadores del centro, lo cual no se corresponde con la realidad.

Frente a la presencia fehaciente de balanceadores en la infraestructura actual de Internet, y dada la necesidad de contar con un traceroute confiable, Augustin *et al.* (ACO+06) presentaron una sustancial modificación llamada `paris-traceroute`. Esta modificación hace que ante la presencia de balanceadores de carga *por flujo*, los sucesivos paquetes enviados por la herramienta sean enviados por el mismo camino.

1.4.2. Limitaciones

Las herramientas basadas en el protocolo ICMP brindan información certera de las dinámicas y estructuras de la red. Sin embargo, un gran número de barreras surgen a la hora de recolectar información a través de métodos basados en ICMP. A continuación presentaremos tres causas detrás de estas barreras: bloqueo ICMP, imprecisiones en le cumplimiento del protocolo IP y el uso de direcciones de terceros.



(a) Escenario inicial

(b) TTL = 1

(c) TTL = 2

(d) TTL = 3

(e) TTL = 4

(f) TTL = 5

(g) Topología recopilada

Figura 1.9: Secuencia paso a paso del funcionamiento de un traceroute y sus posibles imprecisiones debido al balance de carga. El origen (1.1.1.1) utiliza el traceroute para descubrir las interfaces en el camino hacia el destino (6.6.6.6). Las figuras 1.9b, 1.9c, 1.9d, 1.9e y 1.9f muestran como los paquetes expiran producto de limitar el TTL en origen. Esta secuencia también exhibe la alternancia de caminos.

Bloqueo ICMP: El protocolo ICMP es un protocolo complementario en el funcionamiento de la pila TCP/IP, ya que la comunicación se puede realizar perfectamente ante la ausencia de la implementación del protocolo ICMP. No obstante, es importante destacar que el IETF establece que los dispositivos debe responder a los pings (**?** Bra89). A pesar la posibilidad lograr diagnósticos por medio de ICMP, su implementación puede exponer a la red y sus dispositivos a una serie de vulnerabilidades.

En primer lugar, existen innumerables registros de *ataques de denegación de servicio distribuido* (en inglés Distributed Denial of Service, DDoS) llevados a cabo por medio de pings. La modalidad de este ataque consiste en tener el control de un gran número de dispositivos desde los cuales se ejecutan `ICMP Echo Request` en simultáneo hacia un mismo objetivo. El abrumador número de mensajes ICMP tiene como propósito agotar el ancho de banda o el procesador de la dirección atacada, haciendo que esta se vuelva inaccesible (Kum06).

En segundo lugar, la generación de mensajes ICMP en los ruteadores ocurre en un sector de dispositivo denominado *camino lento* (en inglés slow path). Dado que la contestación de los mensajes ICMP no es una prioridad dentro de las tareas del camino lento, los operadores imponen restricciones a la tasa de paquetes ICMP a responder (RUB15). Esta restricción se minimiza la posibilidad de sobrecargar al camino lento con mensajes ICMP, y eventualmente exponerlo a un ataque a través de paquetes ICMP. Sin embargo, esto puede conducir a que los ruteadores a no respondan a solicitudes provenientes de pings y traceroute *benignos*.

Por último, los responsables de los ASes querer ocultar información de su topología a sus competidores o posibles atacantes (LDH+16). Esta intranquilidad genera un desincentivo a brindar de manera transparente respuestas ICMP a través de las cuales se revele la estructura interna de los ASes.

Para de mitigar esta exposición, un gran número de dispositivos opta por no resguardase de pings y traceroutes, no respondiendo a estas consultas. La forma de resguardo puede ser: i) no implementar el protocolo ICMP ii) sofisticados mecanismos de defensa mediante *firewalls* (**?**).

Imprecisiones en le cumplimiento del protocolo IP: El protocolo IP establece se debe descartar un datagrama cuando el valor de su TTL alcance cero (Pos81b). Sin embargo, se ha observado que la existencia de ruteadores que reenvían datagramas con valores de TTL igual a cero (ACO+06). Esto se presenta como un inconveniente para descubrir las interfaces a lo largo de un camino por medio de traceroute, ya que al generar un reenvío en lugar de un descarte, el ruteador no enviará el mensaje ICMP *expiró en tránsito*. Entonces, al no recibir el mensaje ICMP correspondiente, la herramienta no

podrá conocer la interfaz del ruteador.

Uso de direcciones de terceros: Dado que un ruteador dispone de múltiples interfaces, cada una con una dirección IP diferente, el IETF recomienda que los datagramas producidos por los ruteadores tengan como dirección IP de origen, la dirección asociada a la interfaz que serializará el datagrama (Bak95). Cumpliendo estos lineamientos, cuando un paquete expira en un ruteador, el mensaje ICMP derivado de esa caducidad deberá tener como dirección IP origen a la dirección en donde ingresó el paquete que expiró. Sin embargo, se ha recolectado evidencia empírica la cual revela que en ciertas oportunidades los rutadores responden con direcciones IP asociadas a otras interfaces (LDH+16).

Este fenómeno denominado *direcciones de terceros* (en inglés third-party addresses, genera severas imprecisiones al momento de develar una ruta mediante traceroutes. Ante esta circunstancia, el resultado del traceroute incluirá como dirección IP en el camino hacia el destino, una dirección IP que no corresponde al camino real.

1.4.3. Plataformas de mediciones activas

La enorme escala de la red de Internet y su permanentemente crecimiento, genera un gran desafío desde el punto de vista técnico y metodológico (Luc10) a la hora de efectuar mediciones. Durante los últimas décadas un gran número de recursos y plataformas han surgido con el fin de brindar a las comunidades científicas y técnicas la posibilidad de explorar el funcionamiento de Internet. Dentro de estas plataformas podemos destacar principalmente dos proyectos: Archipiélago (CAIa) y RIPE Atlas (NCC19a).

La plataforma Archipiélago desarrollada desde 2007 por CAIDA, cuenta en Enero de 2020 con aproximadamente 200 sondas, en 140 ASes, en un total de 53 países a lo largo de los cinco continentes. Esta plataforma esta compuesta por medio de dispositivos controlados por CAIDA corriendo sobre Raspberry Pis. El acceso a la plataforma no es abierto, es decir, los usuarios no pueden diseñar ni planificar los experimentos y mediciones que corren sobre esta plataforma, sino que esto se encuentra estrictamente restringido a CAIDA. En la plataforma se ejecutan sistemáticamente diferentes campañas de medición con el objetivo de recopilar la topología de Internet. En ciertos casos se disponen de mediciones complementarias, como es el caso de la medición de la congestión en Internet (DCGG+18). Una vez ejecutadas estas campañas, los datos recopilados son de acceso libre.

La plataforma RIPE Atlas desarrollada por RIPE NCC dispone en Enero de 2020 de 10933 sondas en 1517 ASes, en un total de 176 países. La plataforma esta compuesta por pequeñas sondas corriendo sobre un pequeño dispositivo construido por RIPE ex-

clusivamente para este propósito. Esta plataforma es de acceso libre para los usuarios registrados en el sistema, donde la cantidad de mediciones se encuentra limitada bajo un sistema de colaboraciones y contribuciones. Los usuarios pueden diseñar sus propios experimentos, cuyos resultados quedarán disponibles para todos los usuarios de la plataforma.

Capítulo 2

Los CPs en la transformación de la estructura de la red

Las Redes de Distribución de Contenido (en inglés *Content Delivery Networks*, CDNs) se han transformado en infraestructuras esenciales de la Internet contemporánea. Las CDNs tienen el objetivo de reducir la latencia y aumentar la calidad de experiencia (en inglés *Quality of Experiencie*, QoE) percibida por los usuarios al acceder al contenido multimedia. Su importancia se debe a que de estas variables dependen la fidelidad y la permanencia en el sitio de los consumidores (FDT+13). Para lograr su objetivo, las CDNs están compuestas por cientos de servidores desplegados alrededor del mundo, ubicados en la proximidad de los de los usuarios, para poder servir contenido con baja latencia (CFH+13).

Las CDNs implementan diferentes procedimientos de optimización dependiendo de si el contenido es *estático* o *dinámico*. En los casos donde el contenido es *estático*, por ejemplo una imagen o un video, los servidores de proximidad cuentan con una réplica de la información, es decir operan como caches. La proximidad física del servidor reducirá el tiempo de propagación generando una disminución en la latencia. Además el uso de caches locales reducirá el tráfico dentro de la CDN, ya evitará que todas las peticiones se concentren a la réplica primaria de la información. En casos donde el contenido solicitado es *dinámico*, por ejemplo una consulta a un motor de búsqueda o una compra online, los servidores de proximidad se brindarán como *front-ends*. Su misión será intermediar entre los usuarios y los centros de cómputos (grandes y distantes datacenters), donde la optimización de la QoE surgirá a través de implementar split TCP (HWLR08).

Existe evidencia reciente de que el núcleo de Internet, entendido como el punto de mayor interconectividad de la red, que anteriormente se encontraba dominado por grandes proveedores de tránsito, se ha reconfigurado hacia una red orientada a la prestación de servicios multimedia, primero por el arribo de CDNs de uso general y ahora por CDNs

privados. En este trabajo nosotros usamos k-núcleos, un elemento de la teoría de grafos, para definir qué ASes componen el núcleo de Internet y para monitorear la evolución del núcleo desde 1999. Específicamente, investigamos si los grandes jugadores del ecosistema de CDNs y proveedores de contenido pertenecen al núcleo y, de ser así, desde cuándo. Además, examinamos las diferencias entre los núcleos IPv4 e IPv6. Luego investigamos las diferencias regionales en la evolución de los grandes proveedores de contenido. Finalmente, mostramos que el núcleo de Internet ha incorporado un número creciente de contenidos ASes en los últimos años. Para permitir la reproducibilidad de este trabajo, proporcionamos un sitio web que permite el análisis interactivo de nuestros conjuntos de datos para detectar, por ejemplo, ASes "emergentes" mediante consultas personalizadas.

2.1. Introducción

La estructura de la red de Sistemas Autónomos (*Autonomous Systems*, ASes) ha ido cambiando a lo largo de los años debido a cambios disruptivos en la red de Internet (LIJM+10). Durante la era de la NSFNET, Internet tenía una red troncal monolítica implementada en los Estados Unidos para interconectar las instituciones educativas y de investigación (kcB93). Luego de que el gobierno de los Estados Unidos desmanteló el NSFNET, los cambios en la estructura de la red entre dieron lugar a la era de tránsito, en la cual la red presentaba una estructura jerárquica (LIJM+10; Cha13). Más recientemente, Internet se ha transformado en una red multimedia, impulsada por altas demandas de ancho de banda y bajos requisitos de latencia, lo que resulta en una era del Contenido (Hus16).

Las Redes de Distribución de Contenido (*Content Delivery Networks*, CDNs) han desempeñado un papel decisivo en la evolución hacia una red multimedia (GALM08) y el resultante *aplanamiento* [1] de Internet (LIJM+10; ACF+12) (ver Sección 1.3.4). Las CDNs son infraestructuras descentralizadas que proporcionan front-ends cerca de los usuarios para reducir la latencia, maximizar el rendimiento y evitar la entrega de paquetes a través de prolongadas rutas, aumentando la latencia y la probabilidad de experimentar congestión (Lei09). Las CDNs suelen establecer una gran cantidad de acuerdos de pares (peer-to-peer agreement, p2p) con ASes en donde se encuentran los consumidores de su contenido ("eyeballs" [2]). Es importante destacar que no es necesario que cada Proveedor de Contenido (Content Provider, CP) deba desplegar su propia CDN. Un número de *CDNs de uso general* proporcionan servicios de entrega de contenido sin ser generadores

[1] Conocido en inglés como: *flattening*

[2] *eyeballs*: Terminología de la comunidad de operadores de red en idioma inglés, a través de la cual se define a los usuarios, cuya conducta se basa en el consumo de contenido online.

de contenido, como por ejemplo Akamai y LimeLight. Sin embargo, es evidente que varios CPs se han transformado en CDNs privadas con cobertura mundial en lugar de servir su contenido a través de proveedores de tránsito o CDN de uso general debido a una variedad de razones técnicas, económicas y legales (Wir16; Jai13; Dan16; SKC+17; Su,09; ET12).

Una extensa cantidad de información anecdótica, sugiere la notable transición hacia las CDNs privadas. En una entrevista concedida a WIRED en 2012 (WIR12), Craig Labovitz, fundador de Deepfield Networks[3], mencionó, que en ese momento, al menos 40 compañías se encontraban trabajando en el despliegue de sus propias CDNs. En particular, esta transición hacia CDNs privadas se fundamentaba en los mecanismos por los cuales los CPs son capaces de monetizar sus servicios. Este fenómeno también fue observado por Geoff Huston (Hus17), quien reportó un marcado cambio en el método de financiamiento del tendido de fibra óptica transoceánica. Mientras esta infraestructura era previamente solventada por consorcios de proveedores de tránsito, actualmente el tendido de nuevos enlaces es financiado por CPs, y usualmente en solitario.

Además de las CDNs, los Puntos de Intercambio de Tráfico (Internet Exchange Points, IXPs) han sido cruciales para transformar la estructura jerárquica de la red de ASes, transformándola en una red *plana (flat)* (DD10) (ver Sección 1.3.3). La disponibilidad de IXP es crítica para las CDNs, que prefieren tener relaciones directas de interconexión con tantos ASes como puedan (DD08). Los IXPs también se interesan en alojar CDNs para de esta forma proporcionar una vía rentable para que los miembros de IXPs accedan al contenido (Far07).

Más recientemente, los ASes han incorporado lentamente la accesibilidad IPv6 para lidiar con el agotamiento de direcciones IPv4. Un conjunto de hitos, como la última transferencia de bloque IPv4 de IANA (IAN11), el lanzamiento mundial de IPv6 (Int12) y el agotamiento total del stock de direcciones IPv4 de ARIN (ARI15) han fomentado la adopción de IPv6. Se ha observado de que el hecho que IPv6 sea incompatible su antecesor IPv4 ha llevado a que las rutas entre orígenes y destinos puedan diferir dependiendo de la versión del protocolo (DLH+12).

En este capítulo usamos el término "núcleo" de la red para referirnos al subconjunto de ASes que se encuentran densamente conectados. En el pasado, el "núcleo" de la red consistía principalmente en redes denominadas TIER-1, en el cual se encontraban grandes proveedores de tránsito internacional. Estos proveedores tenían la característica de estar conectados a todos las demás ASes TIER-1, es decir formaban un cliqué, y además, no requerían contar con proveedores de tránsito. Los CPs, así como las redes de "eyeballs", que eran los destinatarios del tráfico generado por los CPs, se ubicaban en la periferia de la red. Sin embargo, los CPs y las CDNs de uso general han estado construyendo

[3]Deepfield Networks: `https://www.nokia.com/networks/solutions/deepfield/`

redes troncales intercontinentales, además de haber realizado miles de acuerdos de pares en los últimos años. La importancia creciente de los CPs ha llevado a la discusión y la especulación sobre si los CPs son ahora los jugadores dominantes en el ecosistema de Internet (Hus16).

El objetivo de este capítulo es investigar qué papel juegan ahora los CPs en el ecosistema de Internet y, en particular, si los CPs son ahora parte del "núcleo" de Internet. Específicamente, motivamos este trabajo con las siguientes preguntas: ¿Cómo podemos identificar si un CP pertenece o no al núcleo de Internet? Si el núcleo de la red incluye los CPs, ¿quiénes son? Dado que la adopción general de IPv6 ha sido lenta, ¿notamos ese retraso en la evolución del núcleo de IPv4 e IPv6? Dado que el ecosistema de ASes ha mostrado diferencias notables según las regiones geográficas (DD08), ¿se observan también diferencias geográficas en el papel de los CPs y su presencia en el "núcleo " de las estructuras regionales de Internet? Finalmente, a medida que más CPs despliegan sus CDN privadas, ¿podemos detectar las CDNs "emergentes " que actualmente no están en el núcleo de la red pero que probablemente lo estén en el futuro?

Utilizamos el concepto de *k-núcleos* para analizar la estructura de la red de ASes en las últimas dos décadas. Primero nos centramos en siete grandes CPs y confirmamos que todas están actualmente en el núcleo de Internet. Luego, profundizamos en la evolución de estos grandes actores para correlacionar las características topológicas observadas con prácticas comerciales documentadas que pueden explicar cuándo y por qué estas redes ingresaron al núcleo. A continuación, repetimos la metodología pero utilizando el conjunto de datos IPv6 para comparar y contrastar la evolución de los CPs en ambas redes. Con base en los resultados, investigamos razones comerciales y técnicas por las cuales los CPs comenzaron a implementar la conectividad IPv6. Por último, tomamos una visión más amplia, caracterizando el conjunto de ASes en el núcleo de Internet IPv4 en términos de tipo de negocio y geografía. Nuestro análisis revela que un número creciente de CPs se encuentra ahora en el núcleo de Internet. Finalmente, demostramos que el análisis de k-núcleo tiene el potencial de revelar el aumento de CPs "emergentes". Para fomentar la reproducibilidad de nuestros resultados, hamos puesto a disposición nuestros conjuntos de datos a través de un sistema de consulta interactivo en `http://cnet.fi.uba.ar/TMA2018/`.

2.2. Literatura relacionada

La creciente importancia de las CDNs en el ecosistema de Internet ha producido una vasta literatura sobre este tema, que comparte ciertos de los objetivos con los presentes en este capítulo. Varios artículos han estudiado la estructura interna de las

CDNs (DMP+02; PV06; HWLR08; PBV08), donde el foco se posa sobre los beneficios económicos y técnicos de las CDNs, la necesidad de replicación de datos, estrategias de distribución de contenido y actualizaciones de los cachés, y ubicación geográfica de los cachés. La literatura respecto de las CDNs también ha reconocido la creciente importancia de las CDNs privadas. De hecho, han surgido varios estudios sobre las CDNs privadas más relevantes. La CDN de Google se ha estudiado desde muchos puntos de vista: la expansión geográfica de la infraestructura en los últimos años (CFH+13), la Calidad de la Experiencia de Usuario (en inglés *Quality of Experience*, QoE) (CDF+14), el balance de carga interno (Dan16), la ingeniería de tráfico por su WAN a través de Redes Definidas por Software (en inglés *Software Defined Networking*, SDN) (Jai13) y así sucesivamente. La CDN de Facebook se estudiaron desde el punto de vista de la replicación de datos (HBvR+13), la administración de la red (STWZ16) y de su SDN (SKC+17). Bottger *et al.* (BCT+18) estudiaron la infraestructura de Netflix, llamada Open Connect, debido a su arquitectura notablemente diferente de otras CDNs, así como a la contribución de Netflix al tráfico total agregado de Internet. Calder *et al.* analizaron la CDN de Microsoft, conocido como Azure, como un ejemplo representativo de una CDN alcanzable por medio de Anycast (CFKB+15).

Los IXPs también han recibido mucha atención en la comunidad científica durante la última década. Durante la década de los 2000s, los IXPs fueron en parte responsables de una *peering revolution*, ofreciendo puntos neutrales para que los ASes establecieran acuerdos p2p (no tarifados) entre sí. Los IXPs fomentan la interconexión para mantener el tráfico local y evitar alcanzar vecinos locales (ASes ubicados en la misma región geográfica) a través de enlaces de tránsito tarifados o prolongadas rutas (Cha13). Un fenómeno observado y documentado por Dhamdhere *et al.* (DD10) es que la proliferación de los IXPs ha contribuido a un *aplanamiento (flattening)* de Internet, por medio de cientos de IXPs, independientes unos de otros, repartidos por todo el mundo facilitando la interconexion entre miles de ASes en una misma ubicación. En la literatura, varios artículos han estudiado la anatomía de grandes IXPs (ACF+12), así como el papel de los IXPs en las regiones en desarrollo (GCF+14; FFA15; CDFD+20).

Recientemente, Geoff Huston observó cómo el aplanamiento de la estructura de Internet dio lugar al aumento de los CPs (Hus16). Huston sugiere que el aumento de tanto IXPs como CPs están marginando el papel de los proveedores de tránsito, calificándolo como *"La muerte de los proveedores de tránsito" ("The Death of Transit").*

IPv6 ha cobrado más atención en los últimos años debido al agotamiento del espacio de direcciones IPv4 (IAN11) y al aumento de los adoptantes de IPv6 (Hus18). El crecimiento de la accesibilidad IPv6 y su incompatibilidad con IPv4 han alentado a estudiar las diferencias entre ambas redes, como la longitud de las rutas, el rendimiento y las

perspectivas en el sistema de ruteo (DLH+12).

La previsible coexistencia a largo plazo de ambos protocolos, sumado al escaso número de prefijos IPv4 no asignados, ha dado lugar a la compra de bloques IPv4 entre organizaciones, la cual es permitida por algunos RIRs[4], conocida como *Mercado de Transferencias (Transfer market)* (Liv13; LED17). A pesar de la gran cantidad de transferencias que se han firmado desde 2009, recientemente se ha observado un patrón peculiar: los proveedores de contenido están adquiriendo grandes bloques de direcciones de universidades estadounidenses. Esto se evidenció cuando Google y Amazon obtuvieron bloques IPv4 que anteriormente pertenecían a Merit (AS237) y MIT (AS3) (Int17a; Int17b). Habiendo pagado un promedio de 10 dólares por dirección, se refuerza la importancia que IPv4 aún tiene sobre IPv6.

Existe una vasta literatura sobre la aplicación de conceptos de teoría de grafos para estudiar la estructura de la red de ASes. Algunos ejemplos son los trabajos que han utilizado la descomposición en k-núcleos para estudiar las propiedades de la red (AHDBV08; DGM06; OGLK14). Estos trabajos toman principalmente una perspectiva matemática sobre la estructura de la topología de la red de ASes. En cambio en este capítulo utilizaremos la técnica de descomposición en k-núcleos de la teoría de grafos para estudiar el papel específico de los CPs en Internet a lo largo de los años. En particular nuestro interés es determinar si la descomposición en k-núcleos es capaz de indicar la veloz expansión de las redes de distribución de contenidos, de los proveedores de contenidos más conocidos de Internet. Para poder verificar que lo observado por medio de esta métrica tiene concordancia con el despliegue de estas redes, incorporaremos conocimientos propios del campo de Internet, las estrategias y acciones documentadas por los propios CPs, lo que proporciona más contexto y explicación del fenómeno observado. Si efectivamente la descomposición en k-núcleos revela información de la expansión de las CDNs, esta métrica nos podrá indicar el surgimiento de nuevos y futuros actores relevantes en la estructura de Internet.

2.3. Metodología

El objetivo de este capítulo es estudiar los cambios en la estructura del ecosistema de Internet de nivel ASes desde la perspectiva de los proveedores de contenido y las CDNs. Específicamente, el interés se centra en determinar si los grandes CPs ahora son parte del núcleo de la red, y además, cuándo ocurrió tal transición. Para este propósito, es

[4]RIR: Regional Internet Registry. Organización regional que administra los recursos de números de Internet, tales como Números de Sistemas Autónomos, prefijos IPv4 e IPv6, etc.

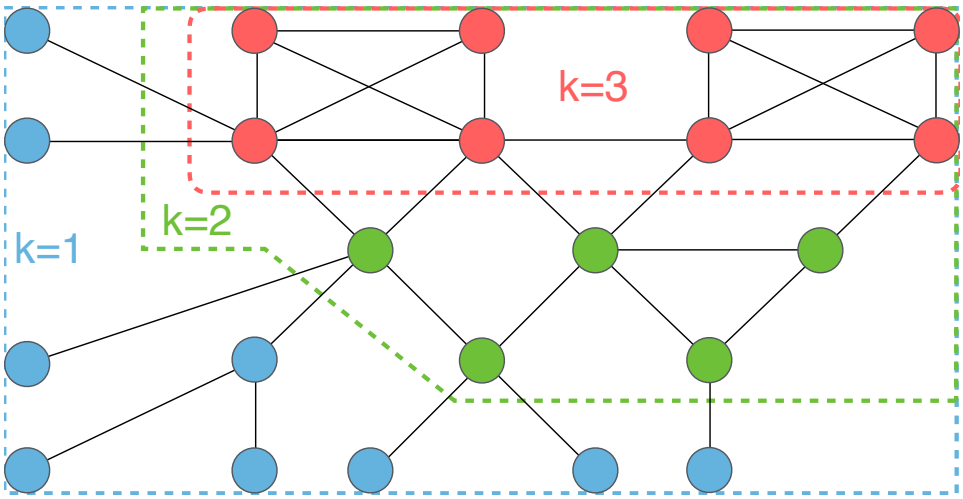


Figura 2.1: Ejemplo de descomposición en k-núcleos de un grafo dado. En color rojo, verde y celeste se muestran los nodos pertenecientes al núcleo con shell-index igual a 3, 2 y 1 respectivamente.

necesario definir una metodología para determinar qué ASes son parte del núcleo de la red.

Dado que buscamos demostrar que los CPs se han conectado tan densamente como los proveedores de tránsito, la metodología elegida debe determinar la conectividad de los ASes en función del número de enlaces: i) con otros ASes (vecinos) ii) de sus vecinos (vecinos de sus vecinos). Por lo tanto, evaluaremos diferentes métricas de teoría de grafos con el objetivo de encontrar una que pueda capturar y contrastar la capilaridad de las conexiones de proveedores de tránsito y de contenido.

2.3.1. Definiciones

Para comenzar, examinamos un conjunto de métricas de teoría de grafos para determinar cuál es capaz de indicar ASes que tienen el mismo nivel de conectividad. Las métricas estudiadas fueron grado de nodo (node degree, k), grado de tránsito (transit degree, T_d), grado promedio de vecinos más cercanos (Average Nearest-Neighbor Degree, $ANND$) y el shell-index de los k-núcleos. La complejidad matemática y computacional de estas métricas es sumamente diferente y el significado de la métrica también lo es.

Grado de nodo (k) Dentro de la teoría de grafos, esta es la métrica más simple para evaluar la conectividad de un nodo. Su valor indica la cantidad de aristas a las cuales está asociado un nodo. Sin embargo, esta métrica no tiene en cuenta ninguna propiedad de los vecinos de un nodo dado. Sin embargo, la mayor limitación de esta métrica se encuentra dado por la subrepresentación del conjunto de aristas del grafo de Internet

tanto en conjuntos de datos recolectados por medio de técnicas activas y pasivas (ver Sección 1.3.5). La marcada opacidad de los vínculos p2p llevará a que ciertos ASes con un presumible alto número de aristas p2p sean subrepresentados, como es el caso de los CPs

Grado de tránsito (T_d) Se define por el número de vecinos únicos que aparecen a ambos lados de un AS en un AS-PATH (LHD+13). Tomando todos los trigramas o 3-gramas a partir de AS-PATHS recolectados en los anuncios BGP, esta métrica contará en cuantos trigramas únicos un ASN aparece en el centro de secuencia. Debido a esta definición, el grado de tránsito es una métrica que mide la relevancia de *intermediación* de un AS.

Grado promedio de vecinos más cercanos ($ANND$) Tal como lo definen el Pastor-Satorras *et al.* (PSVV01), $ANND$ se calcula como el grado promedio de los vecinos de un AS. En comparación con *grado de nodo*, esta métrica tiene en cuenta la relevancia de los vecinos, sin embargo, también se verá sustancialmente afectada ante imprecisiones por sesgos de medición.

Shell-index de los *k*-núcleos Un k-núcleo de un grafo $\mathcal{G}$ es el subgrafo inducido maximal [5] en el que todos los vértices tienen al menos un grado k (ver (BZ11)). Un vértice o nodo que pertenece a un k-núcleo tiene al menos k vecinos que tienen un grado de al menos k. Además, un nodo que pertenece al k-núcleo k también pertenece a cualquier k-núcleo $j < k$, por lo tanto, el shell-index viene dado por el k-núcleo máximo al que pertenece un nodo. La figura 2.1 muestra k-núcleos usando un pequeño ejemplo de gráfico donde los nodos están coloreados para indicar su shell-index. Como muestra la figura, el shell-index viene dado por el grado del nodo y el grado de los vecinos en el grafo inducido. Esto se puede ver en el ejemplo donde algunos nodos de grado cuatro están en el núcleo 2 mientras que los nodos de grado 3 están en el núcleo 3. Además, los grafos ASes son *core-conectados* (Mar08), lo que significa que hay k rutas diferentes a nivel de enlaces (arista-disjuntas)entre dos ASes del mismo k-núcleo.

2.3.2. Evaluación de las métricas en el ecosistema de ASes

Luego de definir un conjunto de métricas candidatas a determinar qué ASes son los más densamente conectados en Internet, a continuación analizamos el desempeño estas

[5] Sea $G = (V, E)$ un grafo cualquiera G, donde V es el conjunto de vertices y E el conjunto de aristas asociados a los vértices V y W subconjunto de vértices $W \subset V$. Un grafo inducido es un subgrafo de G cuyos nodos están dados por W y su conjunto aristas es el subconjuto de E cuyos nodos asociados pertenece a W.

métricas. Para hacer este análisis, elegimos ASes que probablemente se encuentren densamente conectados pero que tengan un propósito comercial diferente. Entre esos ASes, incluimos los TOP10 ASes en el AS-RANK ([CAI18]) de CAIDA, también conocidos como *TIER-1 Transit Providers*, y siete proveedores de contenido populares. Los ASes de tránsito que encabezaban el AS-RANK al momento de hacer esta investigación, en marzo de 2018, son Level3 (AS3356), Telia (AS1299), NTT (AS2914), GTT (AS3257), Telecom Italia (AS6762), HE (AS6939), TATA (AS6453), PCCW (AS3491) y Level3 (anteriormente GBLX) (AS3549). Este último actualmente, en marzo de 2020, se encuentra fuera del TOP10, en el puesto número 12, relegado por Zayo (AS6461) y Vodafone (AS1273). Nuestra selección de CPs se basa en los siete mayores *hypergiant ASes* identificados por Bottger *et al.* ([BCU18]), considerando *hypergiants* según las capacidades del puertos, la huella geográfica y el perfil de tráfico informado en PeeringDB. Estos siete ASes, los que nos referimos como *Big Seven*, son: Akamai (AS20940), Amazon (AS16509), Apple (AS714), Facebook (AS32934), Google (AS15169), Yahoo! (AS10310) y Netflix (AS2906).

El cuadro 2.1 muestra las cuatro métricas para los 17 ASes en estudio: 10 Transit ASes y 7 CPs, para el grafo de ASes en IPv4. Esta tabla no contiene información sobre IPv6 porque solo nos enfocamos en analizar qué propiedades de la red pueden capturar estas métricas en lugar de comparar resultados en redes IPv4 e IPv6.

	Nombre del AS	ASN	AS Rank	k	T_d	$ANND$	shell-index
ASes de tránsito	Level 3 Comm.	3356	1	5601	5374	51	133
	Telia Company AB	1299	2	2010	1947	106	133
	Cogent Comm.	174	3	6080	6062	42	130
	NTT America	2914	4	1737	1989	110	126
	GTT Comm.	3257	5	2010	1732	106	119
	TELECOM ITALIA	6762	6	545	527	248	115
	Hurricane Electric.	6939	7	8173	8079	69	133
	TATA Comm.	6453	8	715	717	177	117
	PCCW Global, Inc	3491	9	701	674	211	117
	Level 3 Comm.	3549	12	2196	2135	52	94
CPs	Apple	714	6360	270	263	926	133
	Netflix	2906	4918	282	196	955	130
	Yahoo!	10310	687	286	249	879	120
	Google	15169	747	356	255	803	133
	Amazon	16509	3444	313	187	893	133
	Akamai	20940	1925	478	411	667	133
	Facebook	32934	4056	344	215	839	133

Cuadro 2.1: Grado de tránsito (T_d), grado de nodo (k), grado promedio de vecinos más cercanos ($ANND$) y shell-index para los TOP10 ASes en el AS-RANK y siete proveedores de contenido conocidos en el grafo de ASes de IPv4. Los métricas corresponden a los datos de Marzo de 2020.

Este cuadro indica que el grado de nodo (k) varía significativamente entre los ASes de tránsito en el TOP10 del AS-RANK, así como también entre los *Big Seven*. Por ejemplo, mientras que el grado de nodo observado para Level3 (AS3356) es 5601 para Telecom Italia (AS6762) es 545. Además, el grado de nodo para los CPs es mucho menor que para los ASes de tránsito ya que muchos enlaces de interconexión son a menudo invisibles en los datos donde se calcularon estas métricas (OPW+10).

El grado de tránsito (T_d) está destinado a medir la intermediación de tránsito, por lo tanto, se espera que los proveedores de tránsito tengan el mayor grado de tránsito. Al comparar ambas tablas, el grado de tránsito para los ASes de tránsito suele ser un orden de magnitud mayor que para los CPs. Los proveedores de contenido generalmente se encuentran al final del AS-PATH, por lo que el T_d tiende a ser bastante pequeño.

El rango en el cual varía $ANND$ es ampliamente diferente entre los proveedores de tránsito y los proveedores de contenido según el cuadro 2.1. En el grafo de la red de ASes, las aristas entre los proveedores de tránsito y sus clientes son siempre visibles. Además, una elevada fracción de los clientes de los proveedores de tránsito son ASes que tienen un bajo grado de nodo (1 o 2), por lo tanto, $ANND$ tiende a ser bajo para los proveedores de tránsito. Por otro lado, los proveedores de contenido también se vinculan con una gran cantidad de ASes de grado 1 o 2, sin embargo, es probable que esos enlaces sean enlaces de interconexión y permanezcan invisibles en nuestros datos (OPW+10; CSR+15). Además, los CPs no tienen clientes que afecten su valor de $ANND$. Por lo tanto, los CPs solo son visibles a través de un pequeño subconjunto de tránsitos TIER-1, lo que hace que los CPs tengan un valor $ANND$ bastante elevado.

El cuadro 2.1 muestra que los grandes proveedores de contenido y los proveedores de tránsito TIER-1 tienen el mismo (o casi el mismo) shell-index. Aunque las definiciones $ANND$ y k-núcleos son aparentemente similares, en realidad no lo son. Y se define por el grado promedio de los vecinos de un nodo, mientras que el shell-index de un nodo dice que un nodo tiene k vecinos de grado k en el subgrafo inducido. Usando k-núcleos podemos ver que tanto los CPs como los TIER-1 se encuentran conectados densamente, donde todos los ASes bajo análisis cuentan con valores similares de shell-index. Por lo tanto, el cuadro 2.1 revela que el desempeño del shell-index con los datos recolectados es satisfactorio para poder capturar y medir la conectivdad de CPs y proveedores de tránsito por igual. Es importante destacar que cuando este trabajo fue publicado en Junio de 2019 (CSAHD19), todos los ASes contaban con el mismo shell-index, $k = 82$. En la actualidad vemos los valores ha crecido significativamente, mostrando un crecimiento más heterogéneo ente los TIER-1 que los CPs.

2.3.3. Metodología propuesta

Para resumir el análisis de las métricas propuestas, el shell-index de la descomposición en k-núcleos es la única métrica observada que indicó valores similares para proveedores de contenido y proveedores de tránsito. Esto se debe a que la definición de la descomposición en k-núcleos establece que un AS está densamente conectado *si y solo si* está conectado a ASes que están tan densamente conectados este. Esta restricción es tan estricta que solo los CPs grandes y los tránsitos TIER-1 la pueden cumplir y, por lo tanto, utilizaremos esta definición para calcular el análisis de la evolución de los proveedores de contenido. Nos referiremos al *núcleo* de la red como el subconjunto de AS que están densamente conectados.

Aplicando la descomposición en k-núcleos, la parte central de la red está compuesta por ASes que pertenecen al núcleo máximo k_{max}. En nuestro análisis estudiamos la evolución del shell-index de los CPs. Sin embargo, los k_{max} así como el resto de los valores (k) del shell-index del grafo de ASes varían con el tiempo. Por esta razón, normalizamos los valores de k en cada muestra por su índice k_{max}, lo que lleva a un k normalizado con valores entre 0 y 1, referidos como k^*. Por ahora, TOPcore se referirá a $k^* = 1$. Para calcular la descomposición en k-núcleos en cada muestra del grafo de ASes, utilizamos dos herramientas, LaNet-vi (Mar08), que además proporciona una visualización de red, y NetworkX, una biblioteca de Python.

2.4. Datos utilizados

Para aplicar la descomposición en k-núcleos de manera longitudinalmente en el grafo de Internet, se necesita contar con una colección con muestras periódicas (por ejemplo: semanales, mensuales, etc) de la topología de la red de ASes en IPv4 e IPv6.

Nuestro estudio se basa en las muestras de la topología de la red de ASes provistas por CAIDA [6], las cuales son de dominio público. CAIDA mantiene una curaduría de datos de la topología de ASes derivada tanto de fuentes BGP como de traceroutes. Los datos proporcionan las relaciones entre ASes en la red IPv4[7] fue inferido por medio de tablas BGP obtenidas por los colectores de RouteViews y RIPE RIS (LHD+13) desde 1998 hasta el presente, y contienen enlaces ASes observados por los colectores BGP junto con la relación comercial inferida. Los datos que brindan las relaciones entre ASes para la red IPv6 [8] se crearon con las mismas entradas pero basado en anuncios de BGPv6 (GLHc15)

[6] CAIDA: Center for Applied Internet Data Analysis. https://www.caida.org

[7] Datos BGP serial-1 de CAIDA: http://data.caida.org/datasets/as-relationships/serial-1/

[8] Conjunto de datos BGP IPv6 AS-REL de CAIDA: http://data.caida.org/datasets/

desde 2004 hasta la actualidad.

Utilizamos un segundo conjunto de datos IPv4 compuesto por enlaces entre ASes inferidos a través de traceroutes generados desde la plataforma Archipiélago de CAIDA (Ark) (CAIa). Cada sonda de la plataforma genera periódicamente traceroutes hacia cada prefijo /24 [9] presente en las tablas de ruteo BGP(también conocidas como RIBs)[10]. Ambos conjuntos de datos IPv4 pueden proporcionar vistas ligeramente diferentes de la topología de nivel ASes de Internet. Si bien el número de aristas en cada muestra extraída de datos BGP es mayor que en las muestras derivadas de traceroute, los traceroute suelen revelar enlaces entre pares (p2p) que no son visibles en los colectores BGP (YHkc03). Para obtener la imagen más completa de la conectividad IPv4 a nivel de AS, elegimos combinar los datasets BGP y Ark, a los que nos referimos como el conjunto de datos "Ark + BGP ". Este conjunto de datos consta de muestras mensuales que datan de 1998 hasta el presente, que es lo suficientemente extenso como para detectar la evolución de los CPs. Desafortunadamente, no contamos datos similares basados en traceroute para ampliar los datos de relaciones en IPv6, derivado de colectores BGP. Además, hemos generado una visualización de la descomposición en k-núcleos de Internet, con la cual se creó una galería disponible a través de un sitio web[11].

Una limitación de nuestra metodología es que los CPs también sirven contenido de cachés ubicados dentro de los ISPs (Su,09; BCT+18), que no son visibles como enlaces entre ASes en los datasets BGP o por medio de traceroutes. Sin embargo, incluso los CPs que siguen la estrategia de almacenamiento en cachés dentro de la red de los ISPs, por lo general, necesitan un contar con enlaces (usualmente p2p) con los ISPs que no están dispuestos a alojar cachés en sus redes, llenar los cachés y servir contenido dinámico que no se puede almacenar en caché. En este trabajo solo estudiamos la evolución de la conectividad a nivel de AS de los CPs; aunque un análisis de la infraestructura de caché es importante para interiorizarnos con la forma en que se sirve el contenido, consideramos que dicha tarea está fuera del alcance de este capítulo y ha quedado pendiente para futuros trabajos.

`2015-asrank6-data-supplement/`

[9] Un prefijo /24 es una dirección de red anunciada sobre BGP, la cual tiene una máscara de red con 24 1s. Otra notación esta máscara es 255.255.255.0.

[10] Los datos de Ark (CAIa) se fusionaron con los datos de skitter (CAId) `http://data.caida.org/datasets/topology/skitter-aslinks/`.

[11] Sitio web con la visualización de la descomposición en k-núcleos de Internet: `http://cnet.fi.uba.ar/TMA2018/`.

2.5. La evolución de los CPs en el núcleo de Internet

Una tendencia que ha sido documentada en la evolución de Internet es que el conjunto de ASes responsables de generar la mayor parte del tráfico se ha estado reduciendo. Estudios recientes han demostrado que solo unas pocas decenas de ASes en conjunto generan la mayor parte del tráfico, mientras que en el pasado esta cifra era del orden de miles (”St16; LIJM+10). Dada esta tendencia de consolidación del tráfico, monitorearemos la evolución en términos de cores de siete grandes jugadores, a los que nos referimos como *Big Seven*: Akamai (AS20940), Amazon (AS16509), Apple (AS714), Facebook (AS32934), Google (AS15169), Yahoo! (AS10310) y Netflix (AS2906).

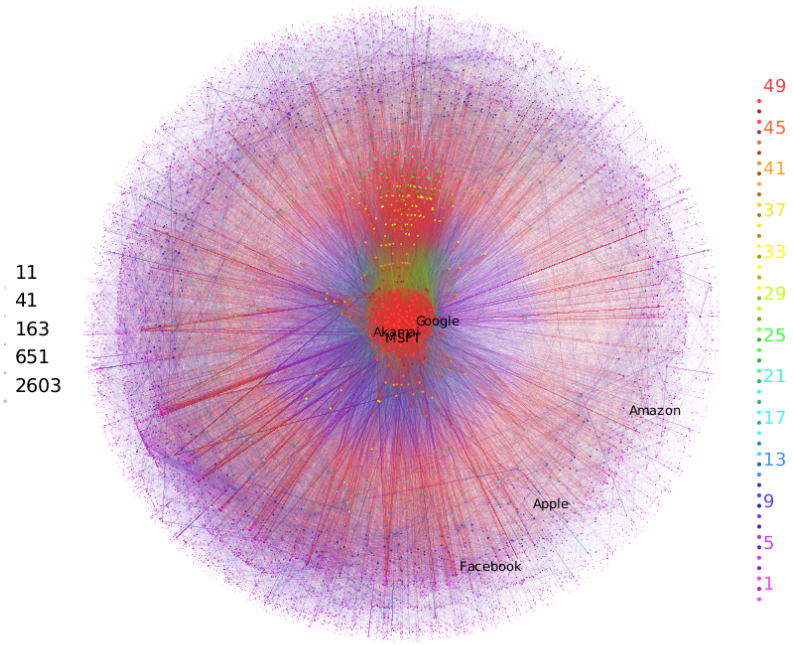
Nuestra hipótesis *a priori* es que todos estos CPs actualmente pertenecen al TOPcore. Verificamos si nuestra hipótesis es cierta, y si es así, *cuándo* y *qué tan rápido* llegaron al TOPcore. Luego intentamos profundizar en las razones por las que observamos estos CPs en TOPcore y correlacionamos con factores externos como disputas legales, expansiones de mercado, mejoras de QoE, lanzamientos de servicios, etc. para explicar porqué los CPs se unieron al TOPcore en cierto momento.

2.5.1. Analizando los *sibling* ASes

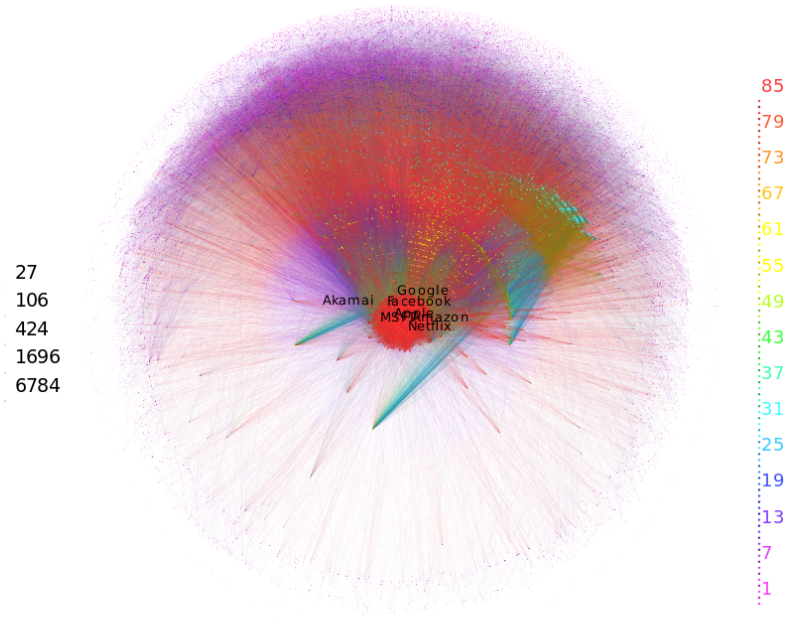
Es presumible que organizaciones, como los CPs, tengan múltiples Números de Sistema Autónomo (del inglés Autonomous System Numbers, ASNs), donde los ASNs que pertenecen a una misma organización generalmente se denominan *siblings*. Hay una gran variedad de razones detrás del hecho de que las organizaciones posean múltiples ASNs, como el ASNs heredados luego de fusiones o adquisiciones, o disponer de múltiples ASNs para diferentes propósitos. Sin embargo, las organizaciones con múltiples ASNs tienden a tener un ASN *primario*, que presumiblemente es más visible que el resto.

Nuestro interés es seguir la evolución de la conectividad de los proveedores de contenido como organizaciones, por lo tanto, necesitamos encontrar todos los ASNs que pertenecen a cada uno de los *Big Seven*. Hallamos los *sibling* ASes de los *Big Seven* usando el los datos de `AS-to-organization` de CAIDA (CAIc), el cual provee una lista basada en registros de WHOIS que une el ASN con el campo `org_id`. En primer lugar, utilizamos los ASNs primarios de los *Big Seven* para obtener los `org_ids` de estas organizaciones. Este paso requiere del conocimiento previo de los ASNs primarios de los *Big Seven*, los cuales se encuentran públicamente disponibles en las entradas de PeeringDB. Luego, a través de los `org_ids` identificamos los restantes ASNs de la organización, siendo estos los *sibling* ASes.

Luego de realizar este procedimiento, descubrimos que 39 ASNs correspondientes a *sibling* ASes de los *Big Seven*, donde Akamai posee 17 ASN, Apple 3, Amazon 3, Google 7,



(a) 2006.



(b) 2016

Figura 2.2: Visualización de la descomposición en k-núcleos del grafo de ASes en 2006 (2.2a) y 2016 (2.2b). Las etiquetas indican la ubicación de los cores de los *Big Seven*, estando ubicadas en el centro las redes de mayor shell-index.

Facebook 2, Yahoo! 5 y Netflix 2. Entre los *siblings* encontramos ciertos ASNs populares y frecuentemente mencionados por la literatura y los operadores, como el AS36040 de Google (anteriormente ASN de YouTube) y el AS6185 de Apple.

Monitoreamos la evolución de los 39 ASNs a lo largo de los años y descubrimos que nunca ha habido un *sibling* tan relevante en términos de shell-index como el ASN primario en el grafo de ASes de IPv4. Mientras que los ASNs primarios pertenecen al TOPcore, los ASNs secundarios han estado a lo sumo en cores a medio camino entre el núcleo 1 y el TOPcore.

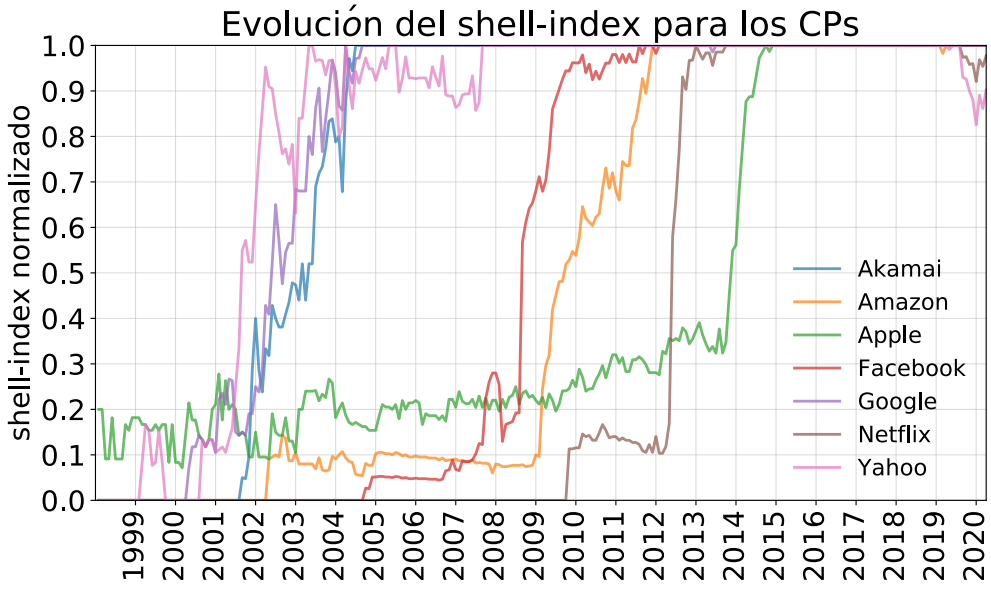
Después de realizar este análisis, podemos concluir que para IPv4 es suficiente monitorear sólo el ASN principal de los *Big Seven*, y luego correlacionar los cambios en esos ASNs con estrategias comerciales. Como trabajo futuro planeamos extender este trabajo a nivel de organizaciones, sin embargo, creemos que par tal fin se debería mejorar el estado del arte de los métodos de agrupamiento de *siblings*.

2.5.2. Monitoreando la evolución de los Big Seven en IPv4

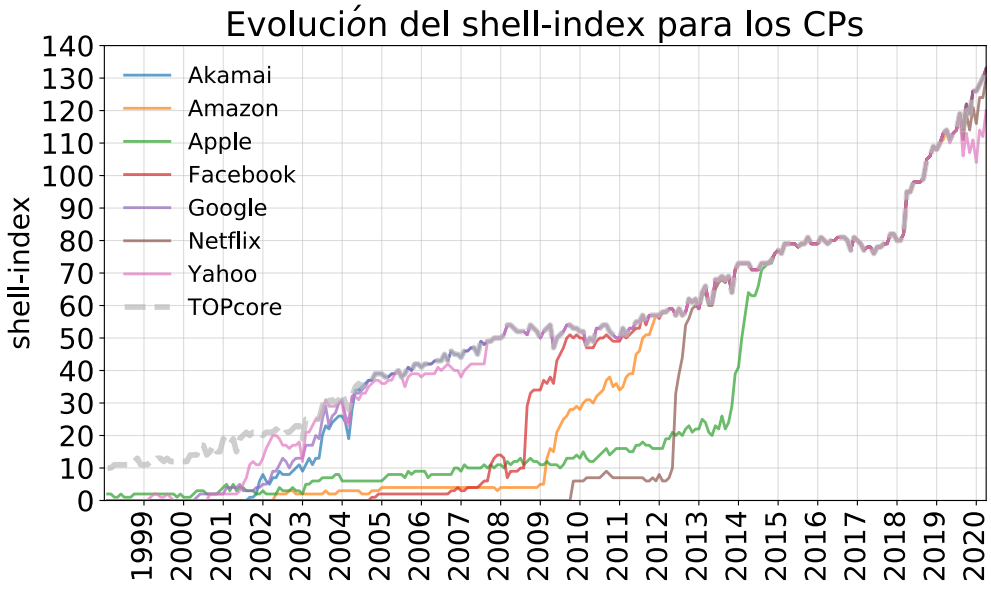
Comenzamos analizando la figura 2.2, en donde se muestra la descomposición en k-núcleos para dos instantes de la topología de la red de ASes, en 2006 (Fig. 2.2a) y 2016 (Fig. 2.2b). Esta comparación nos presenta dos conclusiones. La primera es que durante la década 2006-2016, Facebook, Apple, Amazon y Netflix (no posee etiqueta en la imagen) han hecho esfuerzos para incorporarse al TOPcore. La segunda conclusión, es que la presencia en el TOPcore de Google, Akamai y Microsoft (no presente en los *Big Seven* pero de similar magnitud de acuerdo con Bottger *et al.* (BCU18)) data de varias décadas.

La Figura 2.3 muestra la evolución mensual de los núcleos de los CPs en el dataset Ark + BGP IPv4, donde la Figura 2.3a está normalizada y la Figura 2.3b no. Una primera observación es que a fines de 2017, todos los CPs estudiados se han unido al TOPcore, lo que se indica por el hecho de que el valor central normalizado para cada CP es 1. Con respecto a nuestra publicación de Junio de 2019 (CSAHD19), durante el último año, Netflix y Yahoo! no podido sostener su crecimiento, llevado a su salida del TOPcore. En el caso de Netflix y Yahoo se encuentran en los núcleos 130 y 120 respectivamente, siendo que el TOPcore actualmente tiene un valor de 133.

Parece haber dos grupos entre los CPs estudiados, uno compuesto por Akamai, Google y Yahoo! que alcanzaron el TOPcore en 2005, y otro compuesto por Amazon, Apple, Facebook y Netflix, que se convirtieron en miembros de TOPcore entre 2010 y 2015. Los CPs en el primer grupo están posiblemente más establecidos y han brindado una variedad



(a) Normalizado



(b) No normalizado

Figura 2.3: Evolución de los *Big Seven* en términos de cores en el grafo de ASes de IPv4. Todos ya han alcanzado el TOPcore.

de servicios en línea durante muchos años. El segundo grupo consiste en CPs que en algún momento decidieron implementar su infraestructura propia y dejar de servir contenido utilizando CDNs de terceros como Akamai, ya que el contenido multimedia comenzó a dominar la participación en el tráfico de Internet (San11). Además, la transición desde los cores inferiores a los cores superiores entre los miembros del último grupo es más rápida que en el primer grupo. Es probable que la rápida evolución de los cores de Amazon, Apple, Facebook y Netflix haya sido alentada por la gran cantidad de instalaciones para efectuar vínculos p2p (en inglés peering facilities) que aparecieron durante la última década (Cha13; SSLB17).

De acuerdo con la Figura 2.3b, el TOPcore ha estado siempre creciendo linealmente, a pesar de una pequeña disminución a fines de 2017. En la muestra inicial en 1999, ningún CPs del *Big Seven* pertenecía al TOPcore, y el núcleo máximo era 10. El TOPcore alcanzó su máximo durante 2016, donde el valor de core era 90. Por ejemplo, en 2012 Netflix hizo la transición de 71 núcleos (del núcleo 6 al 77) para poder alcanzar el TOPcore. Por otro lado, esto puede no haber sido realmente una dificultad debido a la expansión de la infraestructura de interconexión (DD10).

A continuación, profundizamos en la evolución de los CPs individualmente. Específicamente, intentamos correlacionar las características topológicas de los CPs (su núcleo) con estrategias comerciales, adquisiciones u otros factores que podrían explicar porqué el CP ingresó al TOPcore.

Akamai Akamai ha estado en el TOPcore desde 2005. Akamai es un pionero en la entrega de contenido, y dado que su modelo de negocio se basa en proporcionar hosting de alta disponibilidad y baja latencia en lugar de generar contenido, siempre han apuntado a tener una gran cantidad de vínculos p2p. Además, Akamai adquirió Speedera Networks (Aka05), una CDN rival, en 2005 para consolidar su posición en el mercado y para ampliar su plataforma. De acuerdo con la Figura 2.3a, Akamai ya había alcanzado el TOPcore cuando compró las Speedera Networks.

Amazon El despliegue de la infraestructura de Amazon parece haber ocurrido en dos pasos, de acuerdo con la Figura 2.3a, que se corrobora con información disponible públicamente en el sitio web de Amazon (Ama17a). En 2009, Amazon estableció su datacenter en el norte de California, que coincide con el primer crecimiento. Entre 2010 y 2012, Amazon estableció datacenters en varias partes del mundo, lo que coincide con el segundo crecimiento acelerado de 2010 a 2012. Además de la implementación de los datacenters,

Amazon estableció docenas de Puntos de Presencia (en inglés *Point of Presence*, PoPs)[12] en todo el mundo para impulsar la expansión, lo que se correlaciona con su ascenso al TOPcore.

Apple Encontramos que el AS de Apple alcanzó el TOPcore en 2015 después de un rápido crecimiento. Según la información pública, Apple ha estado quitando sostenidamente su contenido de Akamai y migrándolo hacia su propia CDN desde 2013 (App16). El porcentaje de tráfico de Apple ha crecido rápidamente en los últimos años debido a las actualizaciones de software, como las nuevas versiones del sistema operativo (ARS14) y los parches de seguridad. Además, la compañía ha anunciado recientemente que está planeando ingresar al mercado de la televisión, produciendo programas de televisión propios, que serán servidos desde la CDN de Apple (LA 17).

Facebook El AS32934 de Facebook se aproximó al TOPcore en 2010 después de un rápido crecimiento de su k^* (núcleo normalizado) entre 2008 y 2010. El número de usuarios en Facebook creció exponencialmente de 12 millones en diciembre de 2006 a 350 millones a finales de 2009 (Yah12), que coincide con el período de expansión de Facebook y su ascenso al TOPcore. Aunque Facebook continuó creciendo exponencialmente desde entonces, el crecimiento masivo durante ese período alentó a Facebook a establecer múltiples acuerdos de pares (p2p). Este marcado aumento de acuerdos p2p visible en los archivos de relaciones entre ASes provistos por CAIDA, es lo que le permitió a Facebook alcanzar el TOPcore. Además, el registro de WHOIS para `fbcdn.net`, alias por el cual se identifica Facebook CDN, se creó en 2007 cuando estaba sucediendo la expansión de Facebook.

Google Google se lanzó en septiembre de 1997 y en solo un par de años se convirtió en el motor de búsqueda más popular (Hor). Con el tiempo, cuando Google comenzó a servir grandes volúmenes de tráfico de video mediante la adquisición de YouTube en 2006 (New), expandió su CDN para acercarse lo más posible a las redes de *eyeballs*. Incluso antes de desplegar su CDN, entre 1999 y 2003, de acuerdo a los archivos de relaciones entre ASes, Google tenía acuerdos de pares un gran número de proveedores de tránsito TIER-1 como Level3 (AS3549), TATA (AS6453), Telstra (AS4637), NTT (AS2914), Zayo (AS6461), Qwest (AS209), GTT (AS3257) y Cogent (AS174). La cantidad de vínculos con una serie de grandes proveedores de tránsito hicieron que Google se convirtiera en parte del mismo núcleo que esos proveedores de tránsito.

[12]Punto de presencia: es una instalación en la cual un proveedor de Internet dispone de equipos de comunicaciones. En estas instalaciones se suelen llevar a cabo el conexionado de enlaces entre ASes.

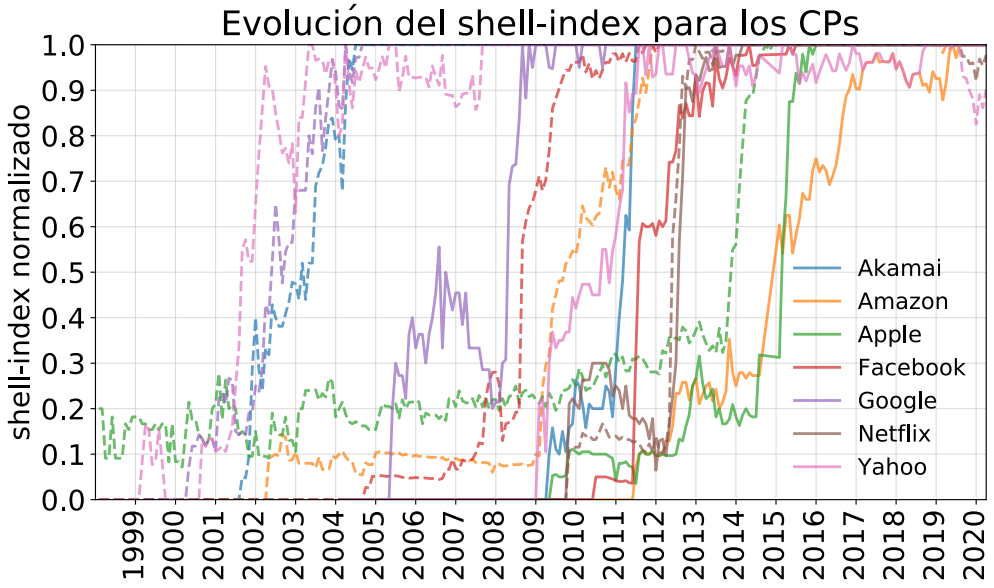
Yahoo! La burbuja de las *puntocom* a principios de la década de 2000 motivó a Yahoo a construir su propia infraestructura WAN para evitar depender de tránsitos por dos razones: i) para reducir la dependencia de entrega de contenido en redes intermediarias entre ellas y redes de *eyeballs* ii) reducir los costos operativos (GALM08). De hecho, el crecimiento en términos de cores de Yahoo coincide con el final de la burbuja de las *puntocom* en 2002.

Netflix En 2012, Netflix tardó menos de un año en pasar del núcleo $k^* = 0{,}1$ al TOPcore. Netflix comenzó a ofrecer transmisión de video en 2007 utilizando CDN tercerizadas y proveedores de tránsito. Con la creciente popularidad del servicio y los crecientes volúmenes de tráfico, la compañía trasladó contenido a su propia plataforma, Open Connect (Net12a), en 2012, que se evidencia como un fuerte aumento en sus valores de cores normalizados entre 01/2012 y 09/2012 como se muestra en la Figura 2.3a. El veloz despliegue la CDN de Netflix, también tuvo su impacto en Internet y su desempeño. Por ejemplo, al comenzar a servir contenido desde *Open Connect*, un gran número de enlaces comenzó a ser afectado por congestión persistente (Qua14). El agotamiento de la capacidad instalada desembocó en conflictos legales entre Netflix y las redes de *eyeballs* para determinar cual de las partes debía hacerse cargo de la inversión para reestablecer la calidad del servicio (Qua14).

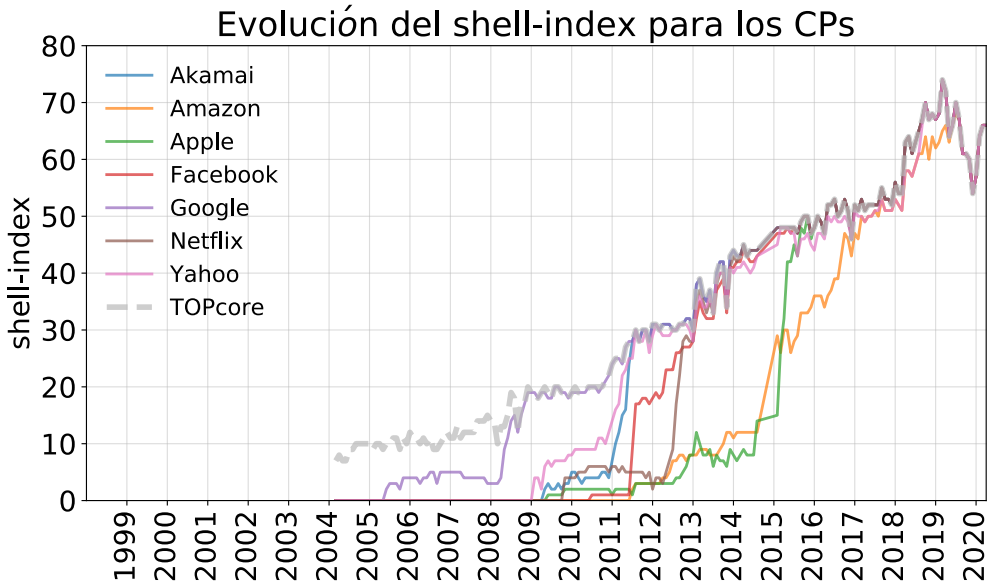
En resumen, todos los CPs estudiados pasaron de CDN tercerizadas a CDN privados y accedieron al TOPcore. En particular, Apple, Facebook y Netflix retiraron significativamente su contenido de Akamai. Estos cambios provocaron una pérdida sustantiva de ingresos para Akamai y una caída en el precio de sus acciones (See15). A pesar de perder clientes importantes, la Figura 2.3a muestra que Akamai todavía permanece en TOPcore, lo que muestra que los acuerdos de pares de Akamai no dependen exclusivamente de estos grandes clientes.

2.5.3. Monitoreando la evolución de los Big Seven en IPv6

La Figura 2.4 muestra la evolución mensual de los cores de los CPs en el conjunto de datos BGP IPv6 comparado con IPv4, donde la Figura 2.4a está normalizada y la Figura 2.4b no. La figura 2.4a también compara la evolución de los cores normalizada para los *Big Seven* en los cores IPv4 e IPv6. Esta cifra confirma que todos estos ASes están actualmente presentes en ambos TOPcores, sin embargo, la fecha de llegada al TOPcore de IPv6 es significativamente diferente al de IPv4. Dos factores al parecer impulsaron la adopción de IPv6 en estas compañías: el lanzamiento mundial de IPv6 en



(a) Evolución de los *Big Seven* en términos de cores en el grafo de ASes de IPv6. Las líneas punteadas indican la evolución del AS en el grafo de IPv4 mientras que las sólidas representan la trayectoria en IPv6.



(b) Evolución de los *Big Seven* en los cores sin normalizar

Figura 2.4: Las Figuras 2.4a y 2.4b, muestran la evolución de los *Big Seven* en términos de cores en el grafo de ASes de IPv6. La Figura 2.4a compara la evolución a lo largo del tiempo de los ASes en las redes IPv4 e IPv6. La Figura 2.4b presenta la evolución de los cores en IPv6 sin normalización mes a mes. Ambas figuras muestran que en el presente todos los miembros del *Big Seven* ya han alcanzado el TOPcore.

2012 (Int12) y el agotamiento del stock de direcciones IPv4 de ARIN en 2015 (ARI15).

A pesar de que han pasado varios años desde el *World IPv6 Launch*, y que el grupo de los *Big Seven* alcanzaron el IPv6 TOPcore, el tamaño (cantidad de nodos y aristas) del grafo IPv6 a nivel de ASes sigue siendo significativamente menor que IPv4 (Hus08; Hus18). Como se muestra en las Figuras 2.3b y 2.4b, donde los cores no están normalizados, el máximo shell-index en la última muestra es 83 para IPv4 mientras que es 52 para IPv6.

A continuación, presentaremos algunas estrategias comerciales que, a nuestro juicio, consideramos que pueden haber impulsado a los miembros del *Big Seven* a incorporarse a la red IPv6.

Akamai Después de un leve crecimiento entre 2009 y 2011, Akamai alcanzó rápidamente el IPv6 TOPcore en 2011. Este crecimiento coincide con la campaña que llevó a cabo la empresa para captar los primeros clientes de IPv6 en 2011 (Chr11). Para tal fin, la compañía seleccionó un subconjunto de clientes para comenzar a servir su contenido a través de dual-stack.

Amazon La implementación de IPv6 de Amazon siguió una tendencia casi idéntica a la de su crecimiento de IPv4, pero se retrasó unos años en comparación con la otra versión del protocolo IP. Amazon incorporó progresivamente países donde el servicio de dual-stack ya se encontraba disponible (Ama11; Ama17b), similar a su expansión mundial de IPv4. A fines de 2014, se produjo un crecimiento notable en el valor del núcleo de IPv6 de Amazon. Al examinar detenidamente las muestras mensuales del grafo de ASes correspondientes a este período, observamos que Amazon comenzó a establecer acuerdos de pares con un gran número de ASes brasileños sobre IPv6 a través de IX.br-SP, IXP brasileño ubicado en San Pablo.

Apple Apple desplegó su propia CDN en 2015 y, como se ve en la Figura 2.4a, tanto sus cores en IPv4 como en IPv6 crecieron al mismo ritmo. Justo después de que Apple desplegara su CDN, Apple comenzó a implementar la preferencia IPv6 (Dyn15), lo que podría verse como un indicador de porqué Apple lanzó la accesibilidad IPv6.

Google Este CP fue el primero de los *Big Seven* en alcanzar el TOPcore en IPv6. Google comenzó a probar su accesibilidad IPv6 utilizando el dominio `ipv6.google.com` durante IETF72 en marzo de 2008. Poco después, en enero de 2009, Google comenzó a estar públicamente disponible a través de IPv6 (Lor09).

Facebook Facebook alcanzó el IPv6 TOPcore después de dos grandes pasos, uno en 2011 y el otro en 2012. El prefijo IPv6 de Facebook `2a03:2880::/29` se asignó, según los registros de WHOIS, en agosto de 2011. Luego, Facebook fue uno de los participantes del lanzamiento mundial de IPv6 en junio de 2012 (Fac09) y durante este evento, la compañía alcanzó el IPv6 TOPcore.

Yahoo! La compañía ha respaldado la adopción de IPv6, y se unió y patrocinó el *The World IPv6 Launch* y *The IPv6 World day* en 2011 y 2012 respectivamente (Int12). Yahoo! alcanzó el TOPcore en IPv6 unos meses antes de *The World IPv6 Launch* en 2011.

Netflix Netflix implementó su CDN (llamada Open Connect) en IPv4 e IPv6 simultáneamente en 2012. Como se muestra en la Figura 2.4a, Netflix creció rápidamente en ambos núcleos, y al mismo ritmo. Aunque el núcleo máximo en IPv4 e IPv6 fueron diferentes en 2012 (Figuras 2.3b y 2.4b), la estrategia agresiva de Netflix de establecer vínculos entre pares le permitió conectarse densamente en ambos núcleos al mismo tiempo. Además, según la información de Netflix, el video se entrega a través de IPv6 siempre que sea posible (Net12b; Net16).

2.6. Evolución regional del núcleo de IPv4

También investigamos si los CPs pertenecen al TOPcore de IPv4 en cada región geográfica, tomando la defección de región de la división de los *Regional Internet Registries (RIR)*. Repetimos el análisis de velocidad y fecha de arribo realizado en la Sección 2.5.2 para cada CP en cada RIR con un enfoque en la detección de diferencias por región, especialmente retrasos sistemáticos cuando ciertos CPs cobraron importancia en regiones específicas.

Para determinar en qué regiones está presente un AS, utilizamos la base de datos de geolocalización NetAcuity (Dig) para geolocalizar cada prefijo anunciado por un AS en una muestra dada. Para este análisis, nos centraremos en la evolución del shell-index de IPv4 debido a la falta de datos de geolocalización para los prefijos IPv6. Las imprecisiones de las bases de datos de geolocalización ha sido ampliamente estudiada (PUK+11). Sin embargo, un trabajo previo presentó evidencias que la base de datos de NetAcuity es altamente confiable para la geolocalización a nivel de país (Gha17). Utilizamos granularidad a nivel de RIR en este análisis, por lo que creemos que no se verá afectado por imprecisiones en la geolocalización. Luego de geolocalizar los ASes, combinamos las

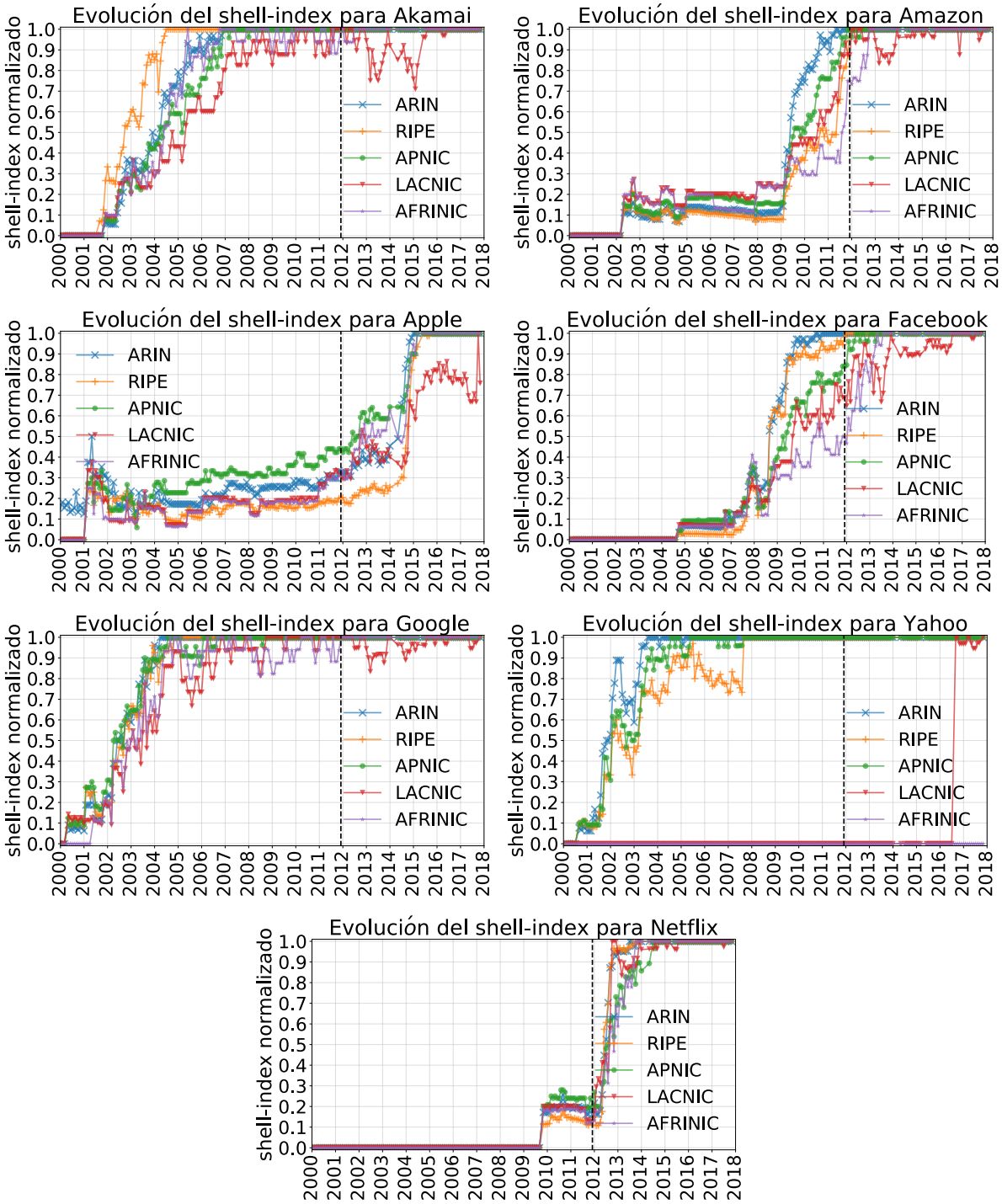


Figura 2.5: Evolución del shell-index de los *Big Seven* en cada RIR. La línea vertical punteada indica el inicio de lo registros de geolocalización.

muestras mensuales de la topología presente en el dataset "Ark + BGP " con el mapeo entre ASes y RIRs para crear sub-grafos mensuales de RIRs.

Hay dos inconvenientes con esta metodología elemental que necesitamos tener en cuenta. Primero, necesitamos información de geolocalización AS a lo largo de la duración del conjunto de datos "Ark + BGP". Sin embargo, la bases de datos con la que se trabajó solo contaba con registros de NetAcuity desde noviembre de 2011, mientras que nuestro dataset topológico comienza en enero de 1998. Segundo, parece que existe en un lapso de tiempo entre que un prefijo comienza a estar activo en una nueva ubicación, y cuando esta nueva ubicación se consolida en la base de datos. Por ejemplo, NetAcuity comenzó a informar la presencia de Netflix en la región de LACNIC en diciembre 2016, mientras que Wayback Machine (Mac19) muestra que en junio de 2015 [13] Netflix ya era miembro de IX.br-SP. Como nuestro objetivo es hacer un seguimiento histórico de la evolución, es necesario contar con el AS en el sub-grafo a nivel RIR cuando los cambios realmente están sucediendo y no una vez que ya sucedieron. Para dar cuenta de estos problemas, realizamos dos modificaciones a la metodología elemental.

1. Asumimos que los 7 CPs bajo análisis siempre han tenido presencia en cada RIR. Sin embargo, al construir el subgrafo del RIR, solo incluimos la conectividad observada entre los CP y otros ASes geolocalizados al RIR.

2. Asumimos que antes de noviembre de 2011 (el inicio de nuestro conjunto de datos NetAcuity), los ASes tenían las mismas ubicaciones que tenían en noviembre de 2011.

Si bien esta metodología nos permite crear sub-grafos a nivel de RIR, no podemos inferir dónde ocurre realmente la conexión entre dos AS cuando esos AS tienen presencia en múltiples RIR. Por ejemplo, Google y Level3, que actualmente están presentes en cada sub-grafo RIR, pueden no tener un enlace físico en cada RIR.

2.6.1. Evolución geográfica de los Big Seven en IPv4

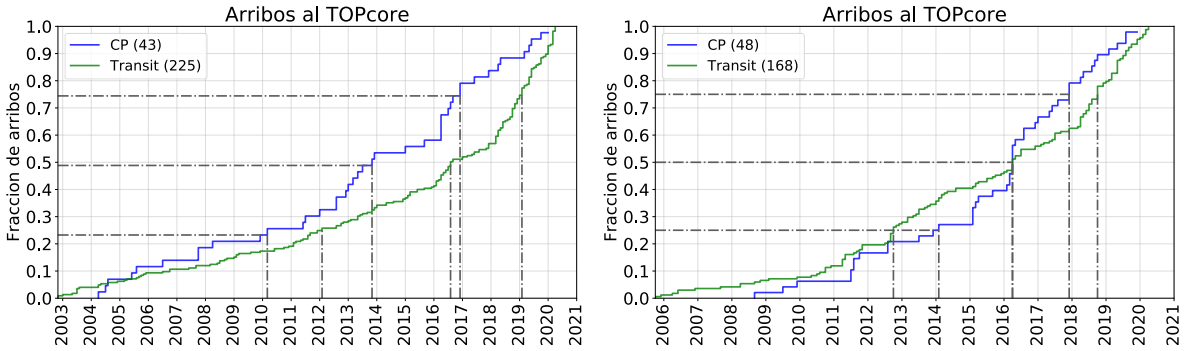
La figura 2.5 muestra la evolución de cada CP en cada RIR. Encontramos que todos los CPs han alcanzado el TOPcore de IPv4 en cada RIR, aunque la fecha de llegada varía según el CP y el RIR. De acuerdo con las tendencias, los gráficos en la Figura 2.5 se pueden clasificar de la misma manera que en la Sección 2.5.2, donde hay un grupo compuesto por Akamai, Google y Yahoo! y otro formado por Amazon, Apple, Facebook y Netflix.

[13] Miembros del IXP de Brasil en la captura de Wayback Machine de Junio de 2015: `http://web.archive.org/web/20150617231252/http://ix.br/particip/sp`

		ARIN	RIPE	APNIC	LACNIC	AFRINIC
Akamai	2007	0.33	0.75	0.21	0.0	0.05
	2012	0.45	0.74	0.45	0.11	0.0
	2017	0.41	0.71	0.47	0.56	0.23
Amazon	2007	1.0	0	0	0	0
	2012	0.49	0.75	0.24	0.35	0.0
	2017	0.40	0.68	0.37	0.53	0.03
Apple	2007	0.73	0	0.40	0	0
	2012	0.60	0.15	0.29	0	0
	2017	0.42	0.67	0.4	0.17	0.07
FB	2007	0.51	0.68	0.21	0.11	0.11
	2012	0.49	0.75	0.37	0.36	0.05
	2017	0.44	0.73	0.37	0.51	0.14
Google	2007	1.0	0	0	0	0
	2012	0.43	0.81	0.27	0.07	0.0
	2017	0.39	0.70	0.38	0.56	0.07
Yahoo!	2007	0.7	0.45	0.15	0	0
	2012	0.57	0.72	0.44	0	0
	2017	0.53	0.73	0.46	0.6	0
Netflix	2007	0	0	0	0	0
	2012	0.86	0.14	0	0	0
	2017	0.39	0.77	0.39	0.57	0.10

Cuadro 2.2: Porcentaje de vecinos locales en cada región.

Amazon y Facebook muestran diferencias entre RIR en su crecimiento a fines de la década de 2000 y principios de 2010. Amazon primero, antes de 2009, estableció datacenters y PoPs en los EE.UU., y luego expandió su presencia a Singapur (APNIC) en 2010, Brasil (LACNIC) en 2011 y varias ubicaciones en Europa (RIPE) en 2011 (Ama17a). La figura 2.5 muestra que las principales tendencias de Amazon siguen el despliegue de infraestructura reportado por la empresa. Facebook, que ha sido parte del TOPcore de IPv4 mundial desde 2009, se quedó atrás en APNIC, LACNIC y AFRINIC, donde llegó al TOPcore varios años después de ARIN y RIPE. Facebook llegó al TOPcore de IPv4 en ARIN en agosto de 2010, APNIC en agosto de 2012, LACNIC en agosto de 2013 y AFRINIC en marzo de 2013. En RIPE, Facebook ha estado en los cores superiores ($k^* \geq 0{,}9$) desde principios 2010, sin embargo, finalmente alcanzó el TOPcore de IPv4 en enero de 2012. Facebook reconoció públicamente su falta de presencia en las regiones en vías de desarrollo y tomó medidas para corregir a fin de mejorar la QoE de los usuarios en esas regiones (Qui16).



(a) IPv4 (b) IPv6

Figura 2.6: Fracción de arribos de miembros del TOPcore ($k^* = 1$), acumulados a lo largo de los años. A la izquierda (Fig. 2.6a) se muestra el arribo al TOPcore IPv4, mientras que a la derecha (Fig. 2.6b) se muestra al TOPcore IPv6. Ambas Figuras presentan un análisis por separados de los proveedores de tránsito y de contenido.

Como los *Big Seven* son empresas de origen estadounidense, es esperable que primero hayan alcanzado el TOPcore de IPv4 en ARIN, y luego se hayan expandido a regiones en vías de desarrollo como LACNIC y AFRINIC. A continuación, investigamos esta hipótesis, al tiempo que observamos que el análisis que sigue es específico de estas empresas y puede no generalizarse a otros proveedores de contenido o regiones. La figura 2.5 muestra, sin embargo, que Akamai y Google mostraron diferencias no significativas entre RIR a principios de la década del 2000, lo que no coincide con la información reportada sobre el despliegue de sus CDNs. Por ejemplo, Google estableció un PoP en Argentina solo en 2011 (Gal13). La razón de esta discrepancia es que Akamai y Google contaban con un gran número de vínculos de interconexión con proveedores de tránsito TIER-1 presentes en esas regiones, lo que hizo que los CPs también estuvieran en el TOPcore de esas regiones. Una mirada a las relaciones entre pares a principios de la década del 2000 confirma esta hipótesis: Google no estaba presente en la región de LACNIC, sin embargo, se asoció con Level3 (AS3549), TATA (AS6453) y Qwest (AS209), que estaban presentes en LACNIC. Confirmamos que los TIER-1 estaban presentes en LACNIC porque establecieron conexiones físicas con los dos ISP argentinos más grandes, Cablevisión (AS10318) y Telecom Argentina (AS4926), que solo estaban presentes en Argentina en ese momento. Descubrimos que ambos ISP argentinos tenían conexiones con Level3 (AS3549), TATA (AS6453) y Qwest (AS209). Por lo tanto, la presencia de Google en el TOPcore de LACNIC probablemente se debe al hecho de que tenía varios acuerdos con proveedores de tránsito activos en LACNIC.

Al igual que Google y Akamai, Yahoo! tuvo tendencias similares en la evolución a nivel de núcleos en ARIN, RIPE y APNIC a principios de la década de 2000. Sin embargo,

Yahoo tuvo un retraso significativo en la región de LACNIC donde la compañía alcanzó el TOPcore en 2016 cuando se unió al IXP brasileño en San Pablo [14]

		ARIN	RIPE	APNIC	LACNIC	AFRINIC	Unknown
COREv4	Content	14	27	2	0	0	-
	Transit	14	160	29	10	9	3
COREv6	Content	26	20	2	0	0	-
	Transit	23	105	22	9	8	1

Cuadro 2.3: Origen geográfico de acuerdo con lo reportando en WHOIS para los TOPcore ASes

Netflix y Apple fueron los últimos en ingresar al TOPcore de IPv4 mundial, así como al TOPcore de IPv4 en cada RIR. Netflix se ubicaba en los núcleos inferiores ($k^* < 0{,}3$) en cada RIR en enero de 2012. En enero de 2014 se trasladó al TOPcore en cada RIR. El crecimiento de Apple fue similar — en junio de 2014 se encontraba en núcleos inferiores a 0,5. Un año después estaba en el TOPcore de IPv4 en cada RIR, excepto LACNIC, donde alcanzó el TOPcore en enero de 2017. El crecimiento simultáneo en cada RIR podría estar relacionado con asuntos comerciales y económicos, sin embargo, el crecimiento exponencial de la cantidad de IXPs en el mundo, así como el creciente número de miembros en cada uno de ellos, podrían haber permitido que Netflix y Apple implementen sus plataformas más rápido.

2.6.2. Vecinos locales

El análisis de la sección anterior mostró que la evolución de los núcleos no refleja necesariamente la expansión geográfica de los CPs. Ahora presentamos un análisis complementario. El cuadro 2.2 muestra el porcentaje de vecinos locales en cada sub-grafo, donde entendemos por locales a los ASes que se encuentran registrados en mismo RIR (de acuerdo con los registros de WHOIS) de donde se los geolocalizó. Por ejemplo, Google tenía el 38 % de los pares locales en APNIC en 2017, lo que significa que el 38 % de los enlaces de Google con ASes presentes en APNIC estaban con ASes registrados en APNIC, mientras que el 62 % restante estaban con ASes presentes en APNIC pero registrados en otra región del mundo. Esta métrica proporciona información sobre cuándo un CPs arribó por primera vez a una región, ya que eso conduciría intuitivamente a un aumento en la métrica de vecinos locales en el sub-grafo a nivel de RIR. Esto es intuitivo,

[14] Wayback Machine de miembros del IX.br-SP. 09/2016 http://web.archive.org/web/20160904012004/http://ix.br/particip/sp

ya que cuando un CP se une a un IXP en una nueva región, se conectará con una gran cantidad de pequeños ASes locales que nunca tendrán presencia fuera de esa región.

El cuadro 2.2 muestra que Akamai, Google y Yahoo! aumentaron significativamente su número de pares locales en Latinoamérica (LACNIC) entre 2012 y 2017. APNIC también ha mostrado un crecimiento en el número de pares locales, pero más lento que en LACNIC. En contraste con la Figura 2.5 donde todos los CPs pertenecen a cada TOPcore, el cuadro 2.2 muestra un número aún bastante reducido de vecinos locales para estos CPs en AFRINIC. A partir de 2017, Akamai tenía la fracción más notoria con 0,23, Facebook en segundo lugar con 0,14 y el resto estaba por debajo de 0,10.

Si bien el porcentaje de vecinos locales de CPs aumenta con los años en regiones donde inicialmente tenían una pequeña fracción de pares locales, ARIN muestra la tendencia opuesta. Esto es probable porque los CPs bajo análisis son compañías estadounidenses. En consecuencia, su número de vecinos locales en ARIN se satura, mientras que el número de vecinos no locales aumenta a medida que las empresas extranjeras despliegan infraestructura en ARIN, específicamente en territorio de los Estados Unidos, y crean vínculos con los CPs.

2.7. El TOPcore más allá de los Big Seven

Concluimos el análisis de este capítulo observando otros ASes en TOPcore. Específicamente, investigamos cuatro aspectos relacionados con este conjunto de AS: i) Composición de los TOPcores (Sec. 2.7.1) ii) Evolución de los adoptantes de dual-stack (Sec. 2.7.2) iii) Tiempo requerido para alcanzar el TOPcore (Sec. 2.7.3) iv) Tendencias de otros CPs notables que no se incluyeron en el *Big Seven* (Sec. 2.7.4).

Para identificar los ASes incluidos en el TOPcore, usamos el criterio de que un AS debe estar en $k^* > 0{,}975$ en cualquier momento, y que además $k^* \geq 0{,}95$ durante los últimos seis meses de nuestro dataset (de Octubre de 2019 a Marzo de 2020). Se debe tener en consideración que esta definición del TOPcore es más amplia que la utilizada en el resto del capítulo donde el criterio para pertenecer al TOPcore era $k^* = 1$.

2.7.1. Composición de los TOPcores

Ahora nos enfocamos en conocer cuántos ASes hay en TOPcore, qué categoría (tránsito o contenido) de ASes son, y qué fracción de los ASes en el TOPcore corresponde a redes de distribución de contenido.

Según la actual definición TOPcore, encontramos 268 ASes en TOPcore el IPv4 —

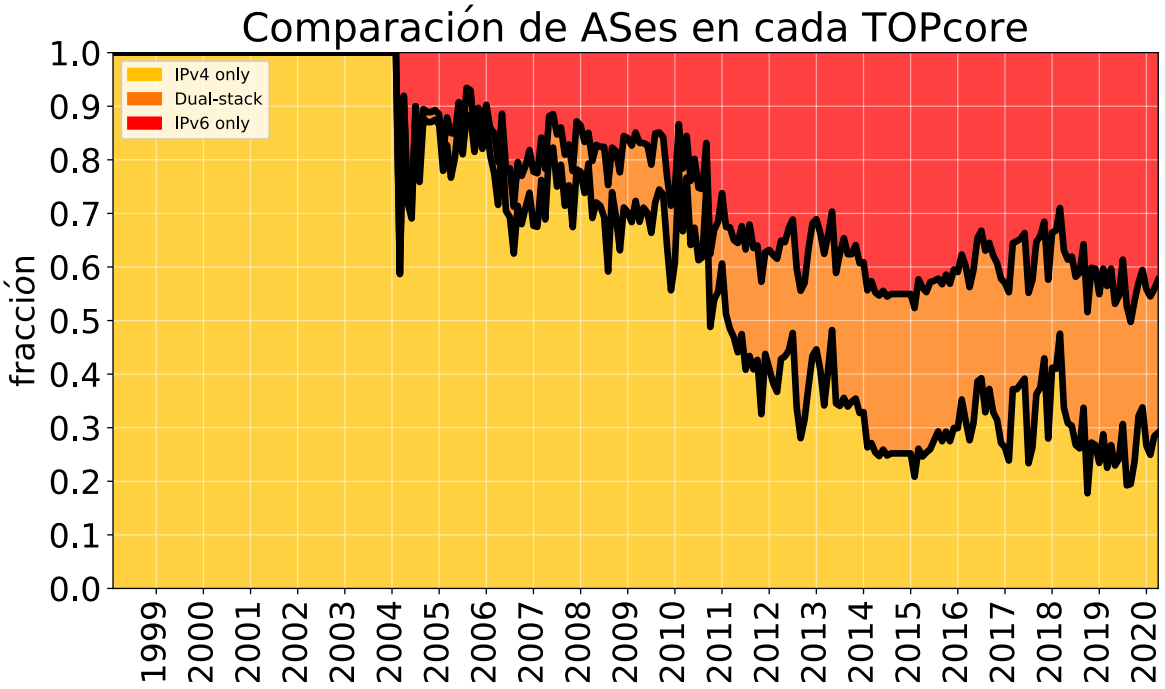


Figura 2.7: Evolución del despliegue de dual-stack entre los ASes miembros de ambos TOPcores.

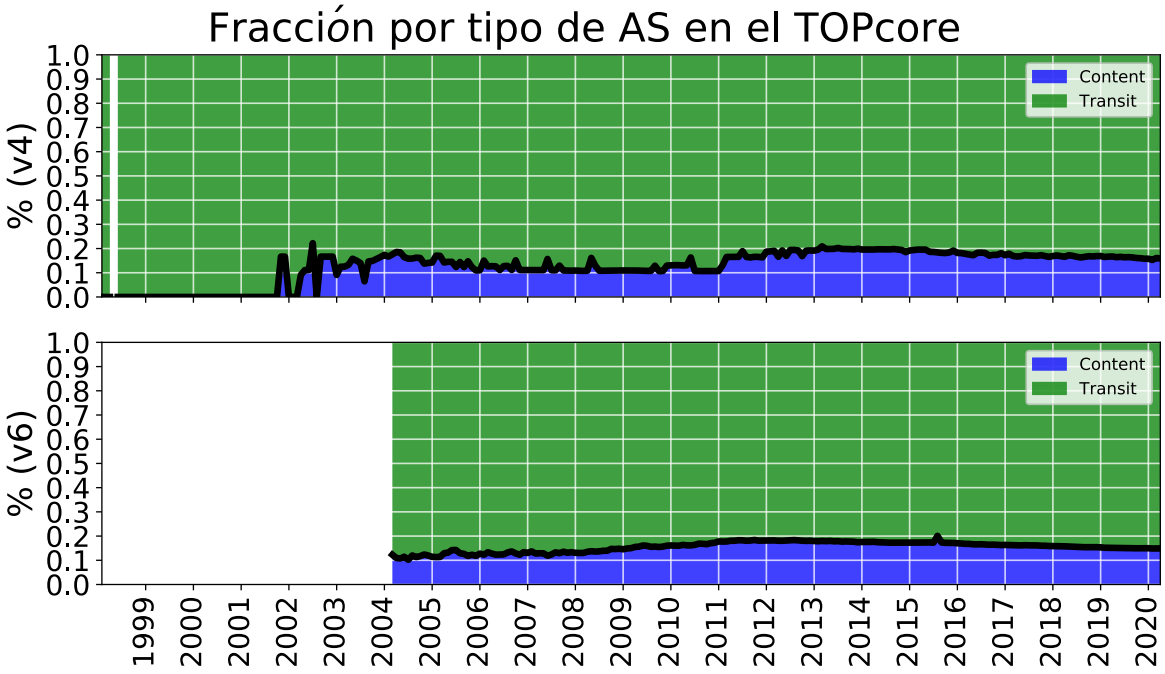
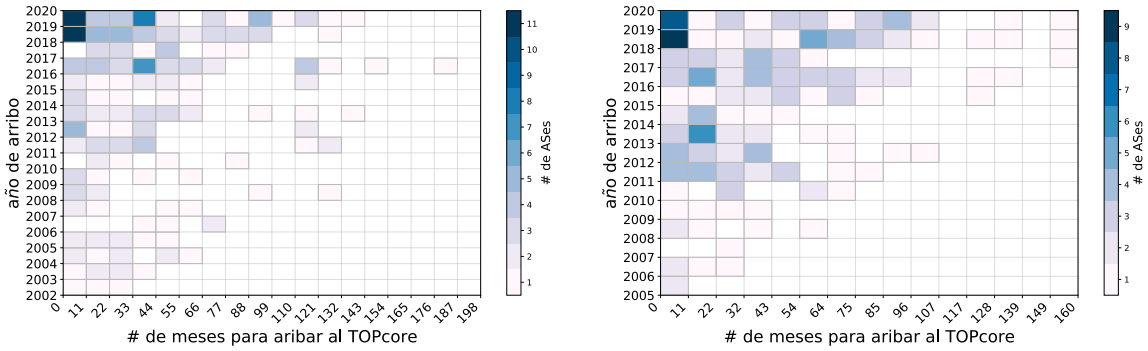


Figura 2.8: Evolución mensual de la fracción de CPs and proveedores de tránsito en el TOPcore.



(a) IPv4 (b) IPv6

Figura 2.9: Correlación entre la velocidad de crecimiento y la fecha de arribo al TOPcore.

43 proveedores de contenido y 188 proveedores de tránsito/acceso de acuerdo con la clasificación AS proporcionada por CAIDA (CAIb). En el TOPcore de IPv6 hallamos 216 ASes, donde 48 son proveedores de contenido y 148 proveedores de tránsito/acceso. Nos referimos al conjunto de ASes en el TOPcore de IPv4 e IPv6 como COREv4 y COREv6, respectivamente.

La figura 2.6a muestra la fracción de COREv4 de Marzo de 2020 (separada en Contenido y Tránsito) que alcanzó el TOPcore de IPv4 en función del tiempo. Este gráfico muestra claramente que con el tiempo, más CPs se han unido al TOPcore. Llamativamente, el 75 % de los CPs en el conjunto estudiado ingresó por primera vez al TOPcore después de 2010. Además, vemos dos fases distintas en la curva de los CPs: la tasa a la que llegaron los CPs al TOPcore ha aumentado desde 2011. El arribo de los proveedores de tránsito, por otro lado, parece constante a lo largo de los años.

La Figura 2.6b muestra el mismo análisis que en la Figura 2.6a pero para el TOPcore IPv6. Mientras que la tendencia para los proveedores de tránsito en la Figura 2.6a aumentó linealmente durante los años, la Figura 2.6b muestra un punto de inflexión a principios de 2011 que coincide con el momento en que IANA anunció la transferencia de su último /8 a los RIRs (IAN11). Con respecto a los proveedores de contenido, el 75 % de los CPs en COREv4 alcanzaron TOPcore después de 2011, mientras que más del 90 % de los CPs en COREv6 alcanzaron el TOPcore en el mismo período. La llegada de proveedores de tránsito y contenido al COREv6 presenta una aceleración, especialmente para los CPs, después de que ARIN anunció que su stock de direcciones IPv4 llegó a cero en septiembre de 2015 (ARI15).

El Cuadro 2.3 muestra la distribución geográfica de ASes en los TOPcores. Vemos que los CPs en los COREv4 y COREv6 son principalmente de ARIN y RIPE (con la excepción de 2 de APNIC en los COREv4 y COREv6). Sin embargo, entre los proveedores

de tránsito, RIPE tiene significativamente más ASes en COREv4 y COREv6 que otras regiones. AFRINIC y LACNIC tienen una presencia escaza o nula en ninguna de las categorías. Por otra parte, APNIC cuenta con un número considerable de proveedores de tránsito en COREv4 y COREv6, pero pocos CPs. Al comparar la composición geográfica de COREv4 y COREv6 por categoría, ambos tienen exactamente la misma distribución. Por lo tanto, la distribución geográfica de ASes densamente conectados es invariable a los cambios en el protocolo IP.

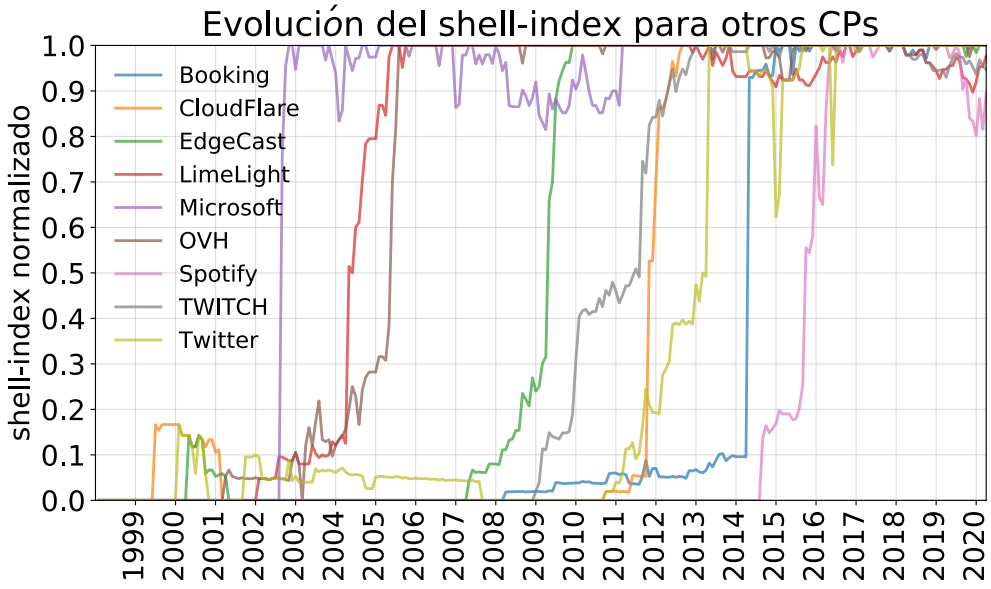
2.7.2. Evolución de los adoptantes de dual-stack

A continuación, analizamos la fracción de ASes que pertenecen simultáneamente a ambos TOPcores en cada muestra de la de los grafos de ASes desde 1999. La figura 2.7 muestra la fracción de ASes en el TOPcore de IPv4, el TOPcore de IPv6 y en ambos en simultáneo. Desde 2004, cuando comienzan los datos de IPv6, la fracción de ASes que solo pertenecen al TOPcore de IPv6 ha crecido constantemente. El número de ASes en ambos TOPcores simutáneamente ha crecido sostenidamente desde 2004 hasta 2014, en donde desde entonces se ha mantenido constante en un 30 %. En marzo de 2020, la red indicaba que aproximadamente el 70 % de los ASes del TOPcore son accesibles a través de IPv6. La constante expansión del TOPcore de IPv6 y el estancamiento del dual-stack nos permite concluir que cada vez es más marcada la diferencia entre los ASes más conectados en IPv4 e IPv6. El crecimiento sostenido de la fracción de IPv6 no es más que el resultado de ASes que pueden alcanzar un alto nivel de conectividad en IPv6, una red más reducida, pero no lo pueden alcanzar en IPv4.

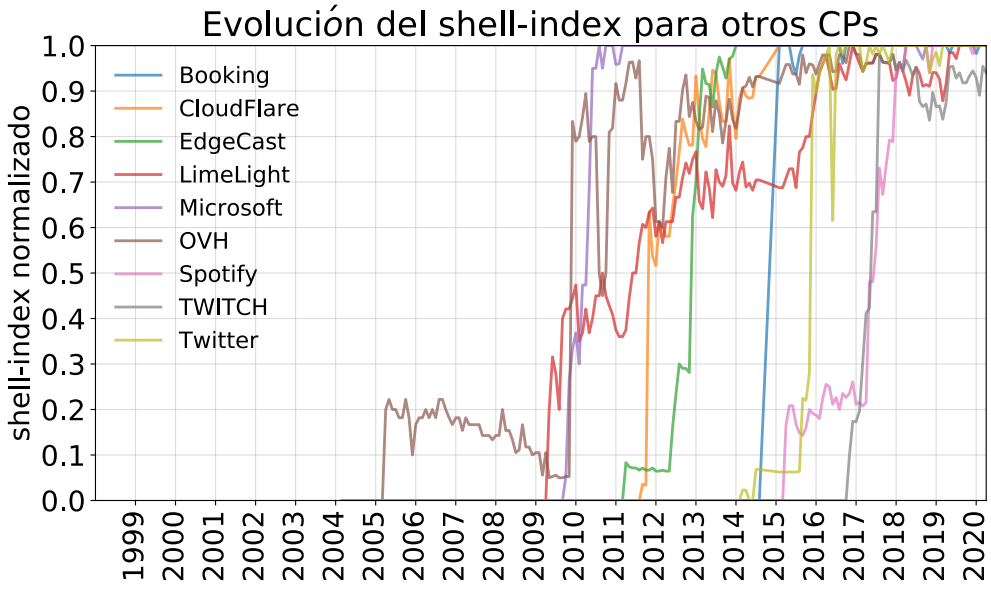
Luego investigamos la composición de ASes en TOPcore a lo largo del tiempo. En la Figura 2.8, aplicamos el criterio para pertenecer al TOPcore para determinar qué AS pertenecen al TOPcore cada mes, y luego clasificar estos ASes en el TOPcore como proveedores de contenido o tránsito. Hallamos que la fracción de CPs en ambos TOPcores ha aumentado constantemente hasta mediados de 2013; alcanzando el valor máximo de 22 %. Luego de alcanzar este pico, tanto el COREv4 como el COREv6 ha reducido levemente su fracción de CPs.

2.7.3. Tiempo para alcanzar el TOPcore

También estamos interesados en analizar *cuán rápido* los ASes arribaron al TOPcore. Las figuras 2.9a y 2.9b muestran un mapa de calor de la cantidad de ASes que arribaron al TOPcore en un momento determinado y a una cierta *velocidad*. Definimos *speed* como



(a) IPv4



(b) IPv6

Figura 2.10: Evolución en términos de cores de otros CPs notables en los grafos de ASes de IPv4 e IPv6.

el número de meses para pasar de $k^* = 0{,}3$ a $k^* = 0{,}975$, y esta definición se basa en las transiciones de los cores inferiores a superiores que se ven en la Figura 2.3a. La figura 2.9a muestra que 214 ASes de COREv4 se unieron a TOPcore entre 2011 y 2018 y la mayoría de ellos se trasladaron desde cores inferiores en sólo unos pocos meses, donde el tiempo promedio requerido para unirse a COREv4 fue de 32 meses. La figura 2.9b muestra la contraparte de IPv6 donde 193 ASes de COREv6 se unieron al TOPcore entre 2011 y 2020. El tiempo promedio requerido para la transición de cores inferiores al TOPcore de IPv6 fue en promedio de 30 meses, que conformado por un número menor de ASes que su contraparte en la red IPv4.

2.7.4. Otros CPs destacados en el TOPcore

Para finalizar, estudiamos la evolución de los cores de otros nueve CPs notables que pertenecen al TOPcore pero que no se incluyeron en los *Big Seven*. Siete de los nueve ASes seleccionados son los ASes restantes en el TOP15 de Bottger *et al.* (BCU18), excepto Hurricane Electric (AS6939) que no consideramos como CP ya que está etiquetado como tránsito/acceso en la clasificación de ASes de CAIDA (CAIb). Estos siete ASes son OVH (AS16276), LimeLight (AS22822), Microsoft (AS8075), Twitter (AS13414), Twitch (AS46489), CloudFlare (AS13335) y EdgeCast (AS15133). Los otros dos ASes son Booking.com (AS43996) y Spotify (AS8403). Curiosamente, Booking.com o Spotify normalmente no se consideran entre los principales CPs, sin embargo, pertenecen a ambos TOPcores.

Las Figuras 2.10a y 2.10b muestran la evolución de los nueve CPs (diferentes de los Big Seven), en donde se observa que se han unido a los TOPcores de IPv4 e IPv6 en los últimos años. Las figuras también indican que muchos transicionaron rápidamente de cores inferiores a superiores.

Twitch es otro CP notable en esta lista, que tal vez no es tan popular como los *Big Seven*, sin embargo, es extremadamente popular entre la comunidad *gamer*. Twitch es una plataforma de transmisión de video que permite a sus usuarios transmitir en vivo la partida de videojuego que se encuentran jugando. El servicio es responsable de ser la cuarta fuente de tráfico pico en EE.UU. (blo14) y su audiencia es incluso mayor que las emisoras de medios tradicionales (ins18). La transmisión de video en vivo es servida exclusivamente por la infraestructura de servicio de Twitch (AS46489) que se extiende por 21 ubicaciones (cada ubicación es identificada por un código IATA) y 12 países (DTCU17). Además, analizando detenidamente los registros de Twitch en PeeringDB, el CP tiene presencia en 47 IXPs alrededor del mundo (DB18). El despliegue

de la CDN de Twitch se observa claramente en la Figura 2.10a, donde rápidamente alcanzó el TOPcore en 2014. Twitch también está presente en COREv6 como se muestra en la Figura 2.10b. Vale la pena señalar que, según esta figura, el lanzamiento de Twitch IPv6 ocurrió en 2017.

Encontramos tendencias similares en la Figura 2.10 y en la Figura 2.4a (*Big Seven*). Para empezar, los ASes que alcanzaron el TOPcore de IPv4 a principios de la década de los 2000s, como LimeLight en la Figura 2.10 o Akamai en la Figura 2.4a, pospusieron el despliegue de IPv6. Por otro lado, también notamos que los ASes que desplegaron sus CDNs en los últimos años son los que tienen menos, o incluso ningún, retraso entre la evolución de cores en IPv4 e IPv6. Si bien Netflix evidencia este patrón en la Figura 2.4a, también lo hace Booking.com en la Figura 2.10.

2.8. Conclusions

En este capítulo demostramos que los CPs han desempeñado un papel decisivo en el ecosistema de ASes, donde siete grandes empresas proveedoras de contenido en Internet se han movido hacia el núcleo de la red. Al analizar la evolución de los CPs en términos de k-núcleos, pudimos identificar posibles razones relacionadas con las prácticas comerciales, estrategias y expansión geográfica que explican el ascenso de estas redes al TOPcore. Además, demostramos que el núcleo de la red ha ido incorporando rápidamente proveedores contenidos a lo largo de los años.

También demostramos que la mayoría de los CPs y proveedores de tránsito alcanzaron el TOPcore de IPv6 varios años después de alcanzar el TOPcore de IPv4, lo que coincide con el hecho de que muchos ASes pospusieron el despliegue de IPv6. Sin embargo, los ASes fueron más veloces en alcanzar el TOPcore de IPv6 ya que la infraestructura física ya estaba disponible para ese entonces.

Confiamos que el análisis de la evolución de los cores puede ser una herramienta para identificar ASes que están aumentando en importancia, los llamados CPs "emergentes". Como un aporte a la comunidad, hemos creado un sitio web para replicar nuestros resultados y poder así monitorear la aparición de CPs "emergentes": `http://cnet.fi.uba.ar/TMA2018/`.

Por último, como resultado de esta investigación se han publicado dos artículos, uno en conferencia con proceedings y otro en una revista, siendo los artículos mencionados los siguientes

1. Carisimo, E., Selmo, C., Alvarez-Hamelin, J.I. and Dhamdhere, A., 2018, June. Studying the Evolution of Content Providers in the Internet Core. In 2018 *Network*

Traffic Measurement and Analysis Conference (TMA) (pp. 1-8). IEEE. (CSAD18)

2. Carisimo, E., Selmo, C., Alvarez-Hamelin, J.I. and Dhamdhere, A., 2019. Studying the evolution of content providers in IPv4 and IPv6 internet cores. *Computer Communications.* (CSAHD19)

Capítulo 3

El impacto de los IXPs en el ecosistema Latinoamericano

Los Puntos de Intercambio de Tráfico (PITs, en inglés *Internet eXchange Points*, IXP) son infraestructuras públicas a través de los cuales los ASes miembros pueden establecer acuerdos de pares bilaterales (p2p) o multilaterales (ver Sección 1.3.3). Específicamente, los IXPs son instalaciones, con presencia en un único o múltiples puntos geográficos, en los cuales los ASes colocan ruteadores de borde de manera tal de intercambiar tráfico con otros ASes presentes en el IXP. La presencia de los ASes se debe principalmente a incentivos económicos. En primer lugar, los ASes pueden establecer múltiples acuerdos de pares con otros miembros, en una única instalación. Es decir, mediante un único enlace se podrá tener acceso a N miembros, cuando por fuera del IXP, esto requeriría N enlaces, lo que implicaría más costos de instalación y operación. En segundo lugar, los acuerdo multilaterales en los IXPs son no tarifados. Esta característica beneficia a los ASes miembros, ya que todo el tráfico intercambiado en el IXP será tráfico que se retiró de los enlaces tarifados p2c, generando un ahorro en la operación (DD10).

Los réditos comerciales, sumado también a mejoras en la QoE, dieron lugar a una masiva aparición de IXPs alrededor del mundo durante la década de los 2000s. Debido a que estos beneficios se deben estrictamente a cambios en la topología de Internet (surgimiento de nuevos vínculos entre ASes), el notable arribo múltiples IXPs al ecosistema de ASes dio lugar a un drástico cambio en la topología conocido como el *aplanamiento* de Internet (DD10).

En este capítulo investigaremos los IXPs implementados en Latinoamérica. Presentaremos que un gran número de estados latinoamericanos ha participado activamente y enérgicamente en el desarrollo de IXPs en sus territorios. Además, daremos pruebas de la existencia de una correlación entre el éxito de un IXP nacional y la ausencia de ASes locales monopólicos los cuales concentran la major parte del espacio de direcciones

IPv4 del país. En particular, tres IXPs latinoamericanos han dado señales de éxito nivel doméstico: IX.br-SP, CABASE-BUE y PIT Chile-SCL. Además, compararemos estos IXPs con otros fuera de Latinoamérica. Expondremos que, en las regiones en desarrollo, los IXPs han tenido un crecimiento similar en los últimos años y donde se encuentran principalmente compuestos por ASes regionales. Este último punto contrastará claramente con los IXPs europeos conocidos internacionalmente cuyos miembros abarcan múltiples regiones.

3.1. Introducción

Latinoamérica cubre 20 millones de km^2 (Ban16) y comprende 20 países: justo detrás de Norteamérica, tiene la mayor tasa de población urbana (Ban18). Además, Latinoamérica es el hogar de 652 millones de personas (Nat17) y tiene tres de las cuatro áreas metropolitanas más grandes de América (San Pablo, Ciudad de México y Buenos Aires con poblaciones de 21.3M, 21.2M y 15.3M de habitantes respectivamente) (Nat16). Latinoamérica también tiene números atractivos con respecto a Internet: en Julio de 2019, 8661 de los 10171 ASNs delegados a LACNIC aparecen actualmente en las tablas de ruteo BGP. Además, de los 65438 ASes activos, 6458 han sido delegados por NIC.br (NIR brasileño) a organizaciones con sede en Brasil. Sin embargo, pocos estudios de Internet se han centrado en Latinoamérica, y mucho menos en sus IXPs.

Latinoamérica también participó en la irrupción masiva de Puntos de Intercambio de Tráfico (IXPs) que comenzó a principios de la década de 2000 y que contribuyó a *aplanar* Internet (DD10): alberga 119 de los 967 IXPs presentes en todo el mundo (Hou19). Muchas razones sugieren porqué los IXPs también se han extendido en Latinoamérica. Primero, los IXPs nacionales en Latinoamérica son esenciales para evitar que el envío paquetes entre end-hosts locales sea a través de rutas que implican desvíos internacionales de miles de kilómetros (Gal16). De hecho, la capacidad de establecer vínculos entre pares localmente en los IXPs no solo reduce los caminos, sino que también reduce la latencia (Gal16). En segundo lugar, Latinoamérica posee varias megalópolis densamente pobladas, donde reside una gran base de clientes de servicios y aplicaciones online. Esto atrae a CDNs que, como una forma efectiva de llegar a los *eyeballs*, se unen los IXPs para generar vínculos entre pares, en simultáneo con múltiples ASes (DD08). A su vez, los IXPs también están interesados en incorporar CDNs como miembros para proporcionar al resto de sus miembros un acceso económico al contenido (Far07).

En comparación con regiones como Norteamérica y Europa, *Latinoaamérica tiene escasos recursos para efectuar mediciones de Internet*. Por ejemplo, Routeviews (oO19) (RVs) y RIPE RIS (NCC19b) solo tienen dos y un colectores BGP en Latinoamérica,

respectivamente. La falta de colectores solo permite recrear una representación bastante incompleta del ecosistema latinoamericano de ASes (LBCX03). Por otro lado, sólo se pueden realizar reducidos análisis derivados de medición activa en Latinoamérica (por ejemplo, para descubrir rutas desde/hacia proveedores de contenido), debido a una escasa disponibilidad de sondas en la región. Por ejemplo, en Julio de 2019, RIPE Atlas y Ark CAIDA, las dos plataformas de acceso público para llevar a cabo campañas de traceroutes, contaban con 311 de 10.209 y 12 de 190 sondas en la región respectivamente.

En este capítulo, analizamos puntualmente los IXPs latinoamericanos. Estamos interesados en conocer si se llevaron a cabo políticas públicas para la creación, crecimiento y desarrollo a lo largo del tiempo de los IXPs en la región, y el papel que desempeñan estas infraestructuras en su propio ecosistema AS nacional. En particular, en la Sección 3.2, presentamos los datos que reunimos para llevar a cabo nuestro análisis:

1. Identificamos múltiples colectores de BGP del proyecto Packet Clearing House (PCH) que proporcionan datos valiosos del ecosistema latinoamericano de ASes. Por otra parte, extendimos manualmente la visibilidad de BGP en Brasil aprovechando varios *Looking Glasses* (LG) disponibles y distribuidos en la red brasileña de IXPs, IX.br.

2. Utilizamos los datos de CAIDA que proporcionan las relaciones entre ASes en la red IPv4, archivos de delegación de los RIRs y asignación de prefijos para derivar métricas que ayudan a cuantificar el crecimiento de los IXPs y comprender mejor el papel de los proveedores de tránsito en los IXPs.

Las contribuciones presentes en este capítulo son:

1. Brindamos información en la Sección 3.3 sobre cómo las políticas públicas de los países han alentado el desarrollo de IXPs en América Latina.

2. Proponemos varias métricas en las Secciones 3.4 y 3.5 que permiten tener en cuenta cómo los IXPs han ido cobrando cada vez más importancia desde su creación y cómo este fenómeno se correlaciona con la presencia de un ecosistema ASes doméstico equilibrado, es decir, la ausencia de ASes de tránsito o acceso monopólicos.

3. Comparamos IXPs implementados en varios continentes y encontramos que los IXPs en regiones en vías de desarrollo comparten propiedades similares.

4. Publicamos el código que permite obtener los datos disponibles públicamente que utilizamos y replicar nuestros resultados [1]. Además, ponemos a disposición del

[1] Repositorio del proyecto: `https://github.com/CoNexDat/latam-ixp-obs`

público los datos recopilados a través de los LGs en Brasil [2].

3.2. Datos utilizados

Los datos utilizados en este capítulo se basan principalmente en descargas de tablas BGP (en inglés BGP table dumps, BGP-TD) obtenidas de colectores desplegados en varios países de Latinoamérica. También recolectamos manualmente BGP-TD de Looking Glasses disponibles en Brasil. Además, utilizamos los archivos de delegación de los RIRs, las relaciones comerciales entre ASes inferidas por CAIDA, y los bases de datos que vinculan prefijos con sus ASes origen (*pfix2as*), servicios de geolocalización, PeeringDB y otros documentos digitalizados. A continuación, detallamos cada una de estas fuentes de datos.

BGP-TDs: Utilizamos BGP-TDs del colector de RVs en San Pablo, Brasil (BR)[3]. Las primeras muestras recolectadas por este colector datan del año 2011. También descargamos BGP-TDs recolectadas por PCH, las cuales se encuentran disponibles en el repositorio llamado *"IPv4 Daily Snapshot"* [4]. Este recurso nos permitió acceder a una colección suficiente extensión temporal para llevar acabo análisis longitudinales, ya que en algunos casos los registros comienzan en el año 2010. En particular, la presencia de PCH en la región se ubica en Argentina (AR), Belice (BZ), Chile (CL), Costa Rica (CR), Ecuador (CE), Haití (HT), Honduras (HN), México (MX), Paraguay (PY) y Trinidad y Tobago (TT). Es importante destacar que PCH también cuenta con presencia en el IXP boliviano de Megalink, aunque este es un IXP sin miembros (Mar13), por lo que no fue considerado en nuestro análisis. De esta manera, con 15 monitores ubicados en IXPs en diversos países de Latinoamérica, PCH era, en Julio de 2019, el proyecto de recolección de tablas BGP con la mayor presencia en la región.

Además, para poner nuestros resultados en contexto, también descargamos BGP-TDs de colectores de PCH situados en IXPs de otras regiones: France-IX (París, Francia), DE-CIX (Frankfurt, Alemania), JINX (Johannesburgo, Sudáfrica) y BKNIX (Bangkok, Tailandia). Elegimos estos IXPs porque ellos mismos, o los países donde se implementan, comparten propiedades con los implementados en Latinoamérica: cuentan con las poblaciones más grandes de su región (por ejemplo, Francia, Alemania y Brasil), edades similares (por ejemplo, BKNIX y el IXP chileno fueron creados recientemente , mientras que DE-CIX y los IXPs argentinos, CABASE, han estado operando durante más de dos

[2]LG dumps: https://cnet.fi.uba.ar/latam-ixp-obs/lg-ribs/
[3]Colector de Routeviews en San Pablo: http://routeviews.org/route-views.saopaulo
[4]PCH "IPv4 Daily Snapshot" https://www.pch.net/resources/Routing_Data/IPv4_daily_snapshots/

décadas) y valores actuales comparables de Producto Interno Bruto (PIB) per cápita (por ejemplo, el sudeste asiático, Sudáfrica y Latinoamérica) (Ban19).

Todos los BGP-TDs de RV y PCH se recolectaron el primer día de cada mes. Observamos que algunos ASes comparten sus tablas completas, y creemos que esto no es lo que realmente se anuncia en los IXPs, es decir, siguiendo los principios de Gao-Rexford (LR01), ningún AS ofrecería tránsito gratuito a través de sus proveedores ascendentes. En consecuencia, al analizar cada IXP, solo tomamos en las entradas proporcionadas por su servidor de rutas (en inglés, route server): en estos casos, las rutas recolectadas generalmente contienen a los ASes miembros, y a sus clientes, al menos de forma parcial. Finalmente, se removió el `AS-PATH prepend` (ver Sección 1.2.3.3) de todos los BGP-TDs y se decartaron entradas con AS sets, las cuales corresponden a menos del 1 % de las entradas. Si bien es posible que los BGP-TD no puedan capturar de forma completa la topología de la red de ASes, es posible completar parcialmente las aristas faltantes por medio de campañas de traceroutes (HFU+10). No obstante, la región cuenta con limitados sitios para llevar a cabo tales mediciones (consulte la Sección 3.1).

Finalmente, expandimos el dataset BGP en Brasil utilizando los LGs presentes en IX.br (BSF+16), los cuales son públicamente accesibles a través de `telnet`. Desafortunadamente, IX.br no mantiene un repositorio con BGP-TDs históricos de los LGs. Al ejecutar "`show ip bgp paths`", obtuvimos BGP-TDs en los 31 IXPs regionales de IX.br en julio de 2019. A pesar de que solo se pueden obtener BGP-TDs parciales en San Pablo y Curitiba (BSF+16), esto no afecta nuestro análisis, como se explica en la Sección 3.4.2.

Archivos de delegación de los RIRs: Descargamos los archivos de delegación de LACNIC para determinar el conjunto de ASes delegados a cada país. Sin embargo, debe tenerse en cuenta que la nacionalidad en los archivos de delegación de RIR [5] no indica fehacientemente que un AS opera en el país al que se delegó el ASN, pero sí muestra que la organización que posee el ASN tiene actividades económicas en ese país. Además, nuestro objetivo no es determinar con precisión la ubicación de ASes, sino más bien de dónde provienen las empresas que se unen a los IXPs.

Relaciones comerciales entre ASes y los archivos *pfix2as:* [6] mientras que los primeros se utilizaron para identificar los ASes *activos* cada mes, es decir, ASes con al menos una relación inferida, estos últimos se utilizaron para calcular el espacio de direcciones originado por cada AS.

[5] Archivos de delegación de LACNIC: `ftp://ftp.lacnic.net/pub/stats/lacnic`
[6] data.caida.org/datasets

Geolocalización de stub ASes: Utilizamos la API del AS-RANK (CAI18; CAI19) para obtener la longitud y latitud de los *stubs* ASes. Si bien muchos estudios han reportado ciertas imprecisiones en las bases de datos de geolocalización (PUK+11), hemos acotado nuestro análisis a los stubs ASes bajo la premisa que los registros serán más precisos para este tipo de ASes.

La confiabilidad del servicio de geolocalización de AS-RANK, se debe en primer lugar, a su proveedor de datos de geolocalización. Las coordenadas indicadas por el AS-RANK son derivadas de la geolocalización a través de de NetAcuity (Dig), de la cual se ha probado su confiabilidad a nivel de país (Gha17).

Además nuestra intención es poder obtener la ubicación en la cual se distribuyen las direcciones IPs en control de los ASes, es decir, originadas por éstos. En consecuencia, hemos estudiado el procedimiento por el cual AS-RANK determina la latitud y longitud de cada AS. Este se basa en calcular el promedio de la latitud y longitud de todas las direcciones IP anunciadas por cada AS, siendo estas direcciones las originadas por el mismo AS como así también por cualquiera de sus *downstream* ASes. Por lo tanto al restringirnos a los *stub* ASes, las coordenadas serán exclusivamente de las direcciones originadas por el AS. Finalmente, tomaremos la hipótesis que los stub ASes tienen una expansión geográfica más acotada que aquellos que brindan tránsito.

PeeringDB: utilizamos PeeringDB (Pee) para obtener los ASNs de los servidores de rutas de los IXPs y validar las inferencias.

Documentos digitalizados: recolectamos documentos digitalizados sobre las políticas públicas acerca de Internet aplicadas por los gobiernos de Latinoamérica, siendo por ejemplo: documentos legales, diarios, sitios web, presentaciones.

3.3. Políticas Públicas en Latinoamérica e IXPs

A continuación investigamos si han existido políticas públicas detrás de la creación de los IXPs en Latinoamérica. Para esto, utilizamos en el dataset de documentos digitalizados que recolectamos. El Cuadro 3.1 muestra las organizaciones que actualmente operan estos IXPs y que fomentaron su creación. Una mirada general de la tabla nos muestra que, de los 16 países que cuentan actualmente con IXPs en Latinoamérica, los gobiernos participaron en la creación de más del 55 % de ellos.

En particular, es de nuestro interés conocer detalladamente cómo los estados nacionales se involucraron en la creación de los IXPs. El presidente de Costa Rica firmó

	País	AR	BO	BR	BZ	CL	CO	CR	CU	EC	HT	HN	MX	PA	PY	PE	TT
	Patrocinador	CABASE	Ley	CGI	PUC	PIT CL	CCIT	Decreto	Estado	IXP.EC	AHTIC	CONATEL	IFT	SENACYT	SENATICS	NAP.PE	TTIX
	Operador	CABASE	Estado	NIC.br	UoBZ	PIT CL	CCIT	NIC.cr	NAP.CU	IXP.EC	AHTIC	UNAH	CITI	InteRED	NIC.py	NAP.PE	TTIX
BGP TDs	Monitor	PCH	X	RVs/LGs	PCH	PCH	X	PCH	X	PCH	PCH	PCH	PCH	X	PCH	X	PCH
	#Miemb.	127		1156	6	72		28		5	4	4	6		15		5
	#Agr. IPs	7.9M		26M	67K	19.4M		401K		28K	102K	131K	795K		1.5M		196K

Cuadro 3.1: IXPs en Latinoamérica. Los colores azul, amarillo y magenta representan agencias estatales, organizaciones sin fines de lucro y universidades, respectivamente. #AggIPs se calcula en el espacio de direcciones anunciado por los miembros del IXP (excluyendo sus *downstream* ASes y prefijos repetidos debido a Multi-Origin ASes (MOAS)). La tabla no incluye países de la región sin IXPs y territorios europeos de ultramar en Latinoamérica.

---

una Decreto (dIJ14; Ric19) mientras que el Congreso de Bolivia aprobó una ley (Gal16) dando lugar a la creación de sendos IXPs. Además, agencias federales como Senatics en Paraguay (Hor15), PUC en Belice (Tel16) y SENACYT en Panamá (Int19) fomentaron la creación de sus IXPs nacionales. Los entes reguladores participaron en México (Instituto Federal de Telecomunicaciones, IFT) (dT16), Honduras (Comisión Nacional de Telecomunicaciones de Honduras, CONATEL-HN) (TEC16) y Paraguay (Comisión Nacional de Telecomunicaciones del Paraguay, CONATEL-PY) (Nic16). En Brasil, el Comité Gestor de Internet (CGI), una junta de múltiples partes interesadas con varios representantes estatales, fue responsable de crear el IX.br, la red de IXPs (Asc15) brasileños. Por otro lado, el Cuadro 3.1 también indica que, de manera similar al modelo europeo de IXPs (Cha13), en Latinoamérica, un gran número de organizaciones sin fines de lucro crearon y administraron IXP. Este es el caso de CABASE (AR), CCIT (CO), PIT Chile (CL), IXP.EC (EC), AHTIC (HT), NAP.PE (PE), TTIX (TT). En particular, CABASE (AR) y CCIT (CO) son operados por organizaciones relacionadas con asociaciones de ISPs locales como ocurre en IXPs fuera de la región, por ejemplo, en DE-CIX (DE) (DC19) y JINX (ZA) (JIN19). Además, Belice (University of Belize, UoBZ), Honduras (Universidad Nacional Autónoma de Honduras, UNAH) y Paraguay (NIC.py-Univerisdad Nacional de Asunción, UNA) han delegado la operación de sus IXPs a universidades. Finalmente, la presencia de regulaciones estatales también influyó en el desarrollo de instalaciones para efectuar vínculos p2p (*peering facilities*) en Chile. La subsecretaría de telecomunicaciones emitió la Resolución 1483 (Sub99) en 1999 que obligó a los ISPs chilenos a intercambiar el tráfico doméstico sin abandonar el territorio nacional. Para cumplir con este requisito, los ISPs se unieron rápidamente a NAP Chile, el primer IXP chileno. Más recientemente, en 2016, PIT Chile se estableció sobre la densa infraestructura de interconexión del NAP Chile, aunque trajo cambios significativos en el ecosistema de ASes chileno: mientras que NAP Chile estaba estrictamente limitado

a ASes nacionales, PIT Chile fue creado como un IXP neutral permitiendo también la presencia de ASes no nacionales, en particular abriendo la puerta a CDNs extranjeras.

La notoria presencia estatal en la creación y operación de IXPs en la región tiene una gran variedad de motivos. En el caso de Chile, la Resolución 1483 fue creada en el marco de asegurar la calidad de servicio de Internet ofrecida a los abonados (Sub99) . En casos como los de Bolivia y Paraguay, la iniciativa de impulsar un IXP tuvo como objetivo reducir los costos de tránsito internacional, y así reducir también el costo de los abonos, permitiendo el crecimiento en la penetración de Internet en la población.

3.4. Evolución de los IXPs

Un gran número de los IXPs en Latinoamérica han estado funcionando durante años. En consecuencia, nuestro objetivo es comprender si estos IXPs han podido consolidarse en su región, al igual que otros IXPs que operan fuera de Latinoamérica. El siguiente análisis de los IXPs se enfocará en los siguientes aspectos: i) topología de red ii) miembros, es decir, ASes conectados, iii) ASes conectados a través de miembros (ASes visibles), iv) rol de los proveedores de tránsito y v) el alcance geográfico de los IXPs.

La mayoría de los países que alojan un colector BGP (ver Cuadro 3.1) tienen IXPs pequeños, es decir, que cuentan con menos de 30 ASes conectados cuyo espacio de direcciones total es menor a 2 millones de direcciones IP únicas. Como esto limita las conclusiones que se pueden extraer de ellos, nuestro análisis se centra principalmente en los IXPs más grandes, que son Argentina, Brasil y Chile.

3.4.1. Topologías de las redes de IXP

Utilizamos PeeringDB, documentos digitalizados disponibles en las webs de los IXPs y conocimientos previos, para buscar organizaciones que operen múltiples IXPs (también conocidas como redes de IXPs) en Latinoamérica. Encontramos que para Julio de 2019, IX.br operaba 31 IXPs regionales en Brasil, CABASE 28 en Argentina y PIT Chile 6 en Chile. A continuación, estudiaremos cómo estas organizaciones coordinan e interconectan sus IXPs. En CABASE, los IXPs regionales como CABASE-BUE (AS11058) o CABASE-COR (AS52374) son independientes y tienen sus propios ASNs. Además, cada IXP regional está conectado a un nodo central, *CABASE Ruteo Central* (CABASE-RCN, AS52376), cuya función es interconectar los IXPs regionales (no es un IXP regional que tenga miembros). A través de CABASE-RCN, se aplica una *Política de Anuncios Multilateral Obligatoria* (en inglés, *Mandatory Multilateral Peering Policy*, MMPP): los prefijos

anunciados en un IXP regional son anunciados por el nodo central en todos los IXP regionales, como se puede ver en la Figura 3.1 para CABASE-BUE y CABASE-COR. Además, esto fue verificado por medio de las BGP-TDs de recolectadas por PCH en IXPs regionales de CABASE, como por ejemplo CABASE-ROS, CABASE-PSS, CABASE-NQN y CABASE-COR. Por otro lado, PIT Chile está estructurado como CABASE: los IXPs regionales también están conectados a un nodo central, PIT Chile-SCL (AS61522), pero este a su vez es un IXP regional que cuenta con miembros. Si bien los IXPs regionales chilenos son visibles como miembros de PIT Chile-SCL, dado que PIT Chile solo cuenta con un colector en Santiago y no impone ninguna política de pares, no podemos asegurar si la reciprocidad también es válida. Más precisamente, desde el colector situado en PIT Chile-SCL no podemos conocer si los IXPs regionales tienen alcance unos con otros. Además, tampoco podemos determinar si los IXPs regionales están anunciando todos sus miembros en Santiago. Finalmente, IX.br utiliza un sólo ASN (AS26162) y no tiene una topología centralizada, ni tampoco una red troncal que interconecte los IXPs regionales.

3.4.2. Miembros de las grandes de redes de IXPs

Para identificar a los miembros de los IXPs o *redes conectadas* de cada IXP regional, utilizamos BGP-TDs descargadas en julio de 2019. En particular, para CABASE-BUE y PIT Chile-SCL los obtuvimos de PCH, y para IX.br por medio de sus LGs. Es necesario recordar que en CABASE y PIT Chile se utilizó un único colector en cada caso, mientras que para IX.br se usó una muestra por cada IXP regional. Esto se debe al hecho de que los dos primeros IXPs tienen un nodo central en su red (ver Sección 3.4.1) Mientras que en CABASE utilizamos tablas de CABASE-BUE, que no es el nodo central, pero ve todos los anuncios debido a la imposición de MMPP, para PIT Chile los obtuvimos de PIT Chile-SCL, su nodo central. Por otro lado, dado que IX.br no tiene un nodo central, utilizamos un LG por IXP regional. Finalmente, los miembros de IXPs se infirieron como el primer AS encontrado en cada AS-PATH después de los ASNs de los IXPs (por ejemplo, servidores de rutas, IXPs regionales).

Por ejemplo, para el caso de CABASE, desde CABASE-BUE podemos inferir los miembros de CABASE-DLC (IXP en el Partido de la Costa) ya que estos serán los ASNs que se encuentren antecedidos por la siguiente sucesión: `11058-52376-52370`, donde el primer elemento de la cadena corresponde al ASN de CABASE-BUE, el segundo a CABASE-RCN y el tercero a CABASE-DLC. De esta manera podemos inferir que la presencia de la *Cooperativa Eléctrica Servicios y Obras Publicas de San Bernardo Ltda.* (AS52419) en CABASE-DLC.

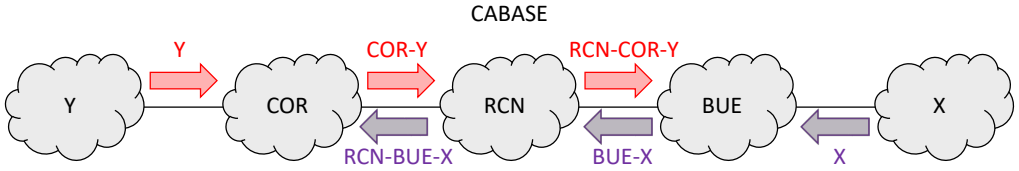


Figura 3.1: Topología de CABASE (para 2 IXPs regionales) y su *Política de Anuncios Multilateral Obligatoria* (en inglés, *Mandatory Multilateral Peering Policy*, MMPP). Las flechas indican anuncios de BGP y sus respectivos AS-PATHS. RCN es un nodo central que interconecta IXPs regionales (por ejemplo, BUE, COR) y reenvía todos los anuncios a todos los IXPs regionales.

Además, verificamos que, a pesar de que las BGP-TDs de los LG en San Paulo (SP) y Curitiba (PR) son parciales (ver Sección 3.2), el número de miembros parece no estar comprometido: mientras RV observa 1156 pares en IX.br-SP, el LG en el mismo IXP regional informa 1164.

La Figura 3.2 muestra el número de redes conectadas en cada IXP regional en IX.br, CABASE y PIT Chile. Algunos IXPs en IX.br y PIT Chile fueron excluidos ya que las BGP-TDs no mostraron miembros. En las tres redes de IXPs, el IXP regional más grande tiene un orden de magnitud mayor que el segundo:

- Brasil: San Pablo: 1156, Río de Janeiro: 245
- Argentina: Buenos Aires: 127, Córdoba: 21
- Chile: Santiago de Chile: 72, Arica: 3

La población de las áreas metropolitanas donde se ubican los IXPs regionales parece tener un impacto en este resultado, San Pablo con 21.3 millones de habitantes, Río de Janeiro con 6.3, Buenos Aires con 15.3, Córdoba con 1.8 y Santiago de Chile con 5.6. Teniendo en cuenta que estos IXPs latinoamericanos atraen principalmente ASes locales (ver Sección 3.4.3), la cantidad de ASes delegados y activos en cada país (en Julio de 2019), Brasil con 6458, Argentina con 791 y Chile con 241, también podría explicar la diferencia de tamaño entre ellos.

3.4.3. ASes visibles

Los ASes conectados a través de miembros, llamados *ASes visibles*, corresponden al conjunto de ASes vistos en los BGP-TDs, es decir, que aparecen en los AS-PATHS de los prefijos anunciados en el IXP. Esta métrica es relevante ya que, a pesar de que algunos ASes podrían no ser miembros del IXP, aún así podrían beneficiarse indirectamente de él.

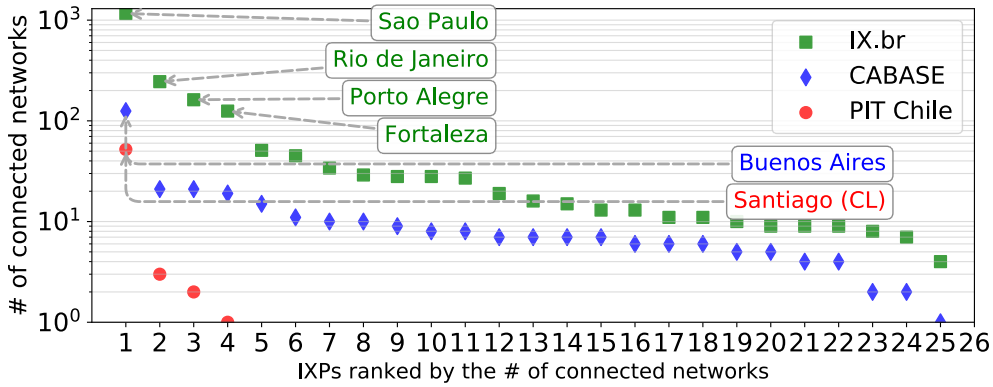


Figura 3.2: Número de ASes conectados a los IXPs regionales en IX.br, CABASE y PIT Chile en Julio de 2019.

Estamos interesados en conocer el impacto de los IXPs en su región, y también determinar el número de ASes extranjeros han sido atraídos por los IXPs latinoamericanos. Además, queremos entender si los IXPs en otras regiones muestran comportamientos similares. Para realizar este análisis, utilizamos los BGP-TDs recolectados por PCH para todos los IXPs, excepto en IX.br, donde utilizamos los datos recolectados por el colector de RV. Además, utilizamos los archivos de delegación de los RIRs para determinar el conjunto de ASNs delegados a cada país.

3.4.3.1. Impacto doméstico

Primero, utilizamos las relaciones comerciales entre ASes inferidas por CAIDA y los archivos de delegación para determinar *todos* los ASNs delegados y activos para cada país y, por lo tanto, para cada IXP. Para esto, simplemente filtramos los ASNs delegados pero inactivos, es decir, ASNs correspondientes a ASes sin relaciones inferidas con otros ASes. Luego buscamos ASes que: i) son visibles en cada IXP producto de observar los AS-PATHS en los BGP-TDs y; ii) son locales, es decir, poseen un ASN delegado en el país donde se ubica el IXP. La figura 3.3 muestra la fracción de ASes visibles locales de todos los ASN delegados y activos para los IXPs más grandes de Latinoamérica: IX.br-SP, CABASE-BUE y PIT CL-SCL. Además, la figura también muestra resultados para France-IX, DE-CIX, JINX y BKNIX (ver Sección 3.2).

La figura 3.3 revela que el 80 % de los ASNs delegados y activos de Brasil y Argentina son visibles en IX.br-SP y CABASE-BUE, respectivamente. Esta fracción es similar a la observada en DE-CIX (Frankfurt) y mucho más grande que la observada en France-IX (París), a pesar de la gran brecha en términos de riqueza (es decir, el PIB per cápita) entre la Unión Europea y Latinoamérica (Ban19). De hecho, a pesar de que Latinoamérica abarca una extensión geográfica más grande, los IXPs de la región aún han logrado

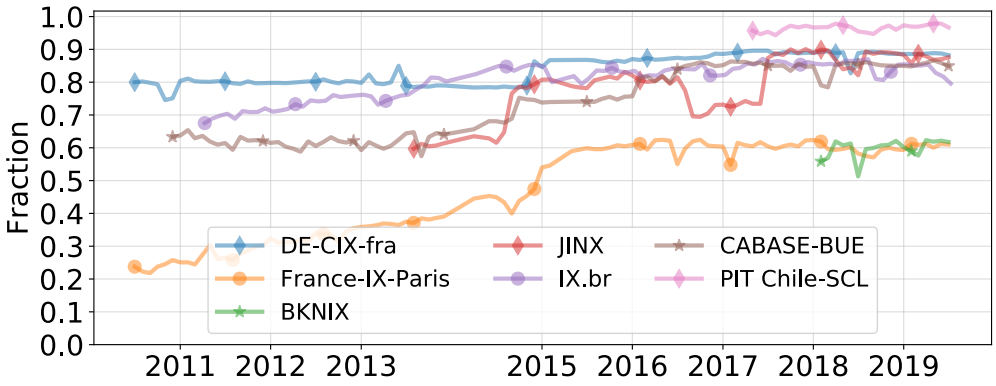


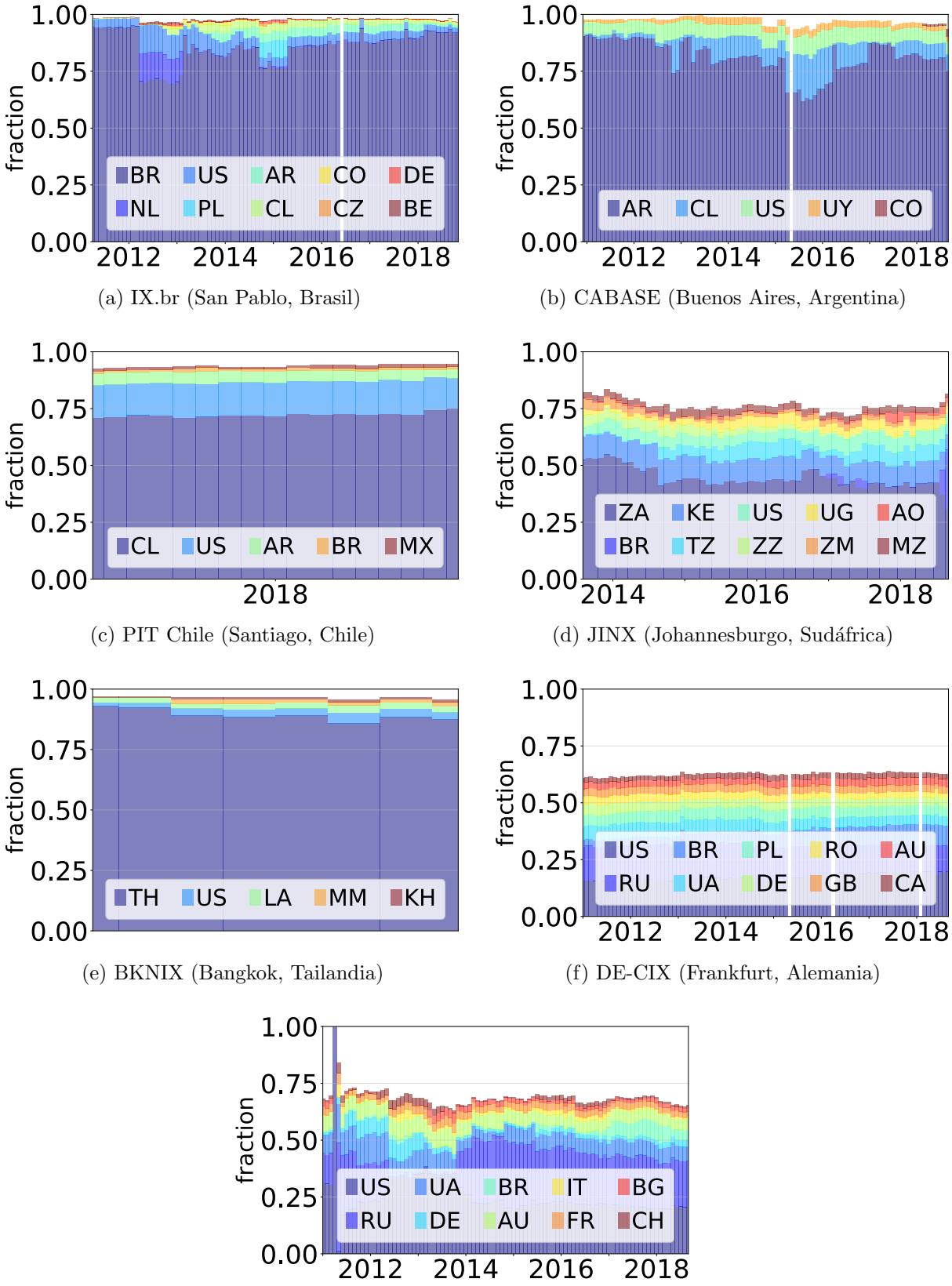
Figura 3.3: Fracción de loos ASN delegados al país y activos, que son visibles en los IXPs.

---

desplegar una infraestructura que les permite incluir una gran fracción de sus ASes locales. Además, aunque DE-CIX se ha estancado en este valor de fracción desde 2011, CABASE-BUE y IX.br-SP han estado creciendo constantemente desde el comienzo de la década, cuando apenas tenían alrededor del 60 %. El crecimiento de la red brasileña de IXPs en la última década fue impulsado por las inversiones en telecomunicaciones para organizar la Copa Mundial de la FIFA 2014 y los Juegos Olímpicos de 2016 (L.08; dB14). Por otro lado, la fracción de ASes visibles en CABASE, así como el número de IXPs regionales, ha aumentado desde que Google se unió al IXP a fines de 2011.

Además, la figura 3.3 también muestra que PIT Chile-SCL, que comenzó a operar en 2016, tiene una fracción sorprendente del 90 % incluso desde la primera muestra que obtuvimos del colector de PCH en 2017. Este es el valor histórico más elevado en Latinoamérica, y de hecho, sumamente elevado para un IXP recientemente creado: por ejemplo, BKNIX, que se lanzó en 2015, cubre solo el 60 % de los ASNs delegados y activos en Tailandia. Para crecer rápidamente, PIT Chile aprovechó las políticas públicas chilenas (ver Sección 3.3).

Finalmente, es importante tener en cuenta que JINX, el IXP en Sudáfrica, también ha aumentado la fracción de ASNs delegados y activos visibles en el país a lo largo del tiempo. Las similitudes con los IXPs de Brasil y Argentina en términos del mismo 20 % de aumento y el hecho de que los tres IXPs hayan alcanzado un valor comparable a un IXP grande como DE-CIX, permite especular sobre un proceso de madurez que se replica en todos los continentes: después de muchos años, las regiones sub-representadas en Internet parecen haber sido capaces de atraer tantos ASes locales como algunos IXPs consolidados en Europa.



(a) IX.br (San Pablo, Brasil)

(b) CABASE (Buenos Aires, Argentina)

(c) PIT Chile (Santiago, Chile)

(d) JINX (Johannesburgo, Sudáfrica)

(e) BKNIX (Bangkok, Tailandia)

(f) DE-CIX (Frankfurt, Alemania)

(g) France-IX (París, Francia)

Figura 3.4: Prevalencia de nacionalidades de los ASes visibles en los IXPs de Latinoamérica, África, Asia y Europa.

3.4.3.2. Atracción de redes extranjeras

La Figura 3.4 muestra la prevalencia de las nacionalidades de los ASes en cada IXP, es decir, de todos los ASes visibles en un IXP, cuántos provienen de cada país. Para este análisis, filtramos los prefijos anunciados por Hurricane Electric (HE, AS6939), ya que este se presenta como un caso particular: es un proveedor de tránsito global, que anuncia una gran cantidad de prefijos y utiliza política abierta (*open peering policy*) (GLHc15). Debido a los desafíos que implica comprender la dinámica de HE, filtramos sus anuncios en IX.br, JINX, DE-CIX y France-IX.

Como se puede ver en la Figura 3.4, los tres IXPs más grandes de Latinoamérica brindan principalmente beneficios a su ecosistema local: la fracción más grande de ASes visibles, alrededor del 75 % en todos los casos, provienen de los países donde se emplazan los IXPs. Sin embargo, estos IXPs también pueden extenderse a otros países de la región, que generalmente suman la mayor parte de la fracción restante en la Figura 3.4. En el caso particular de IX.br, dada su envergadura, éste cuenta con presencia de proveedores de tránsito internacionales lo que permite encontrar prefijos delegados a países europeos, tales como Polonia (PL) o República Checa (CZ), pero aún así la presencia de redes europea es menor que de latinoamericanas. Estos resultados son similares a los observados en BKNIK y JINX. De hecho, todos estos IXPs no cuentan con una notable presencia internacional, es decir, los ASes que albergan provienen de menos de 50 países diferentes en todos los casos. Todo esto contrasta claramente con lo que sucede en los IXPs europeos que actúan como centros internacionales: no solo el número de nacionalidades visibles es mayor que 100 para France-IX y más de 200 para DE-CIX, sino también la mayoría de sus ASes visibles no domésticos con respecto a donde se emplazan los IXP. A pesar de estas diferencias, es notable que EE. UU. esté siempre dentro de las cinco nacionalidades AS más frecuentes[7] para todos los IXPs: esto probablemente se deba al anuncio de prefijos de empresas estadounidenses relevantes en la generación de contenidos, por ejemplo, Google, Facebook, Netflix, CloudFlare, Fastly. Efectivamente, el hecho de que las CDNs encuentren en los IXPs una manera de estar cerca de sus clientes y ofrecerles un mejor servicio, donde esta premisa es también es particularmente cierta en Latinoamérica, Asia y África.

3.4.4. Proveedores de tránsito

Estamos interesados conocer cómo el tráfico es transportado desde/hacia los IXPs latinoamericanos por los proveedores de tránsito, es decir, ASes intermediarios entre los

[7]Por nacionalidad nos referimos a un AS que se ha delegado a EE. UU.

IX.br-SP	ASN	16735	262589	7049	61832	28329
	#	903	381	218	209	207
CABASE-BUE	ASN	3549	52361	7049	19037	11664
	#	219	113	100	82	81
PIT Chile-SCL	ASN	7004	22661	52280	19228	14259
	#	88	87	70	57	57

Cuadro 3.2: Los cincos mayores upstream ASes en IX.br-SP, CABASE-BUE y PIT Chile-SCL en Julio de 2019. El orden de los upstream ASes está dado por el número de downstream ASes que anuncian en cada uno de los IXPs, donde este número se indica en las filas marcadas con (#).

IXPs y ASes de origen vistos en esos IXPs. Más precisamente, dado que los ASes en Latinoamérica podrían estar potencialmente dispersos en vastas extensiones geográficas, nuestro foco se posa en identificar si los proveedores de tránsito han contribuido a la consolidación de los IXPs en su país local, y en tal caso, cuáles han sido estos ASes. Por lo tanto, observamos el tamaño del conjunto de ASes visibles *por upstream AS*, es decir, el conjunto de ASes únicos que aparecen a derecha de cada AS en los AS-PATHS (ver Sección 1.3.2). Este concepto, y su métrica desprendida, suele ser conocido como *Customer Cone* (LHD+13). Para esto, utilizamos BGP-TDs capturados en julio de 2019 por PCH y RV.

El Cuadro 3.2 muestra para IX.br-SP, CABASE-BUE y PIT Chile-SCL, los cinco upstraem ASes que anunciaron los conjuntos de ASes visibles más numerosos. Los resultados muestran un ecosistema ASes más rico en Brasil: Algar (AS16375) solo anuncia más *downstream* ASes en IX.br-SP que todos los ASes visibles presentes en CABASE-BUE y en PIT Chile-SCL. Por otro lado, al observar la nacionalidad de los top 5 upstream ASes en cada IXP, vemos principalmente proveedores de tránsito nacionales. Sin embargo, existen excepciones: Internexa (AS262589, Colombia (CO)) y Silica (AS7049, AR) en IX.br, Level3-GBLX (AS3549, EE. UU.) en CABASE-BUE; Internexa (AS52880, CO) en PIT Chile-SCL.

Además, el Cuadro 3.2 muestra que Level3 (AS3549) es el mayor upstream AS en CABASE-BUE y, aunque no se muestra en el Cuadro 3.2, también ocupó el sexto lugar en PIT Chile-SCL (AS21838, ASN heredado producto de la adquisición de IMPSAT (La 06)). Investigamos aún más el papel de Level3 en ambos IXPs y determinamos que este ISP de EE. UU. en realidad actúa como proveedor de tránsito nacional en Latinoamérica: el conjunto de downstream ASes de Level3 anunciado en CABASE-BUE está compuesto por 204 ASes argentinos de un total de 209 ASes anunciados. De igual manera sucede en PIT Chile-SCL donde 37 de los 43 ASes anunciados por Level3 fueron delegados por LACNIC a Chile.

Finalmente, el cuadro 3.2 también revela la presencia de ISPs de propiedad estatal entre los mayores upstream ASes: Internexa (AS262589, AS262195) y ARSAT (AS52361). Internexa es un proveedor colombiano parcialmente estatal en cual el Ministerio de Hacienda y Crédito Público de Colombia posee el 51 % de las acciones, mientras que el municipio de Medellín (Colombia) posee otro 10 % (ISA19). Por otro lado, ARSAT (AS52361) es un proveedor de tránsito argentino totalmente estatal (dGdMdA19). Es importante destacar que si bien el servicio de tránsito de ARSAT se centra en Argentina, la huella de tránsito de Internexa comprende países extranjeros, como Argentina, Brasil y Chile.

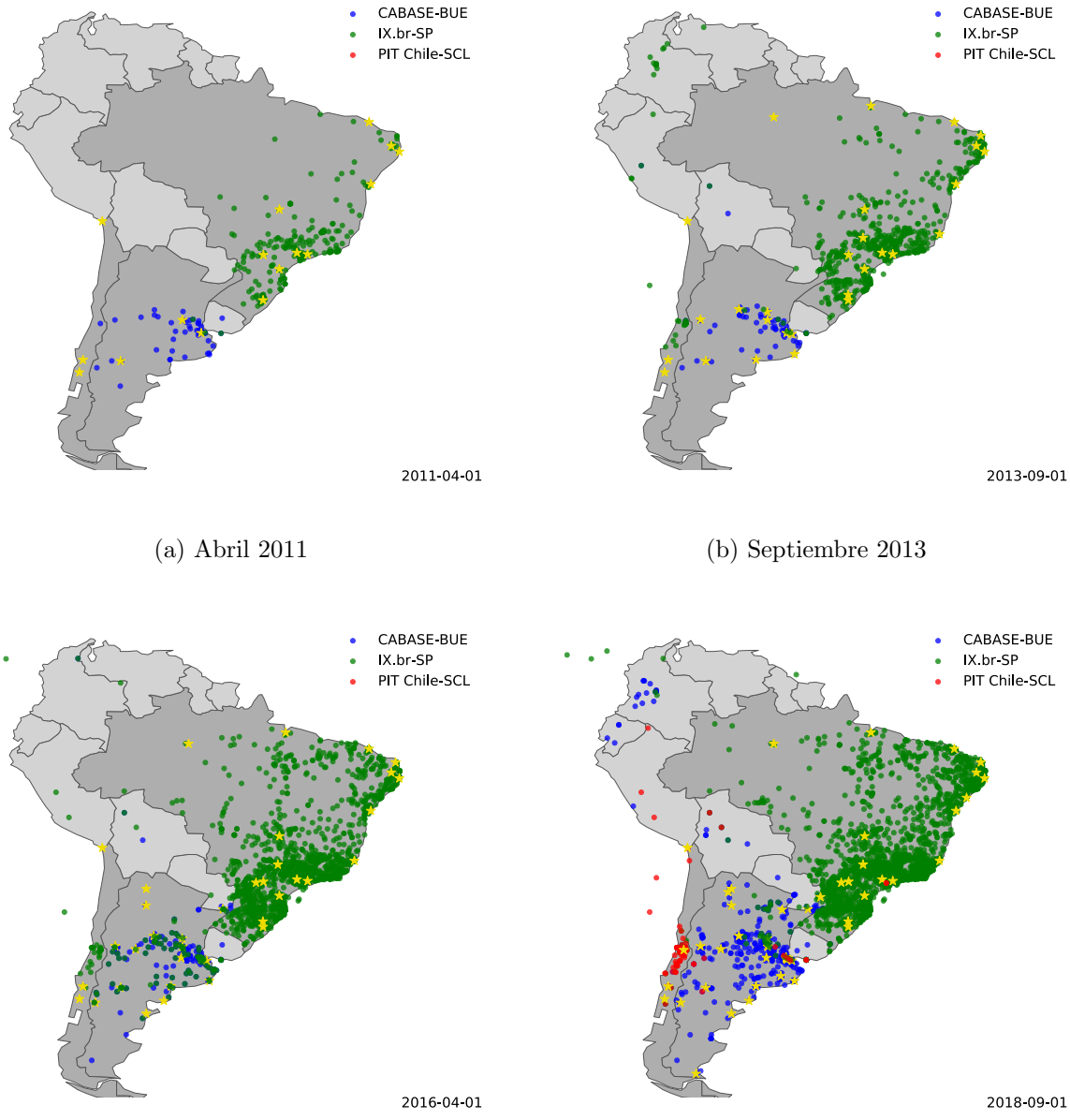
3.4.5. Alcance geográfico de los IXPs

La misión de los IXPs suele estar definida a través el lema *mantener local el tráfico local*, es decir, su propósito es intercambiar tráfico de ASes ubicados en una misma región. Sin embargo, como vimos anteriormente, los IXPs más relevantes de Argentina, Brasil y Chile son capaces de reunir directa o indirectamente casi la totalidad de los ASes domésticos. Esto demuestra que los IXPs más relevantes trascienden el lema *mantener local el tráfico local*, ya que aquí se pueden acceder a ASes distantes, en particular en el caso Latinoamericano.

Nuestro interés ahora se centra en visualizar la expansión del alcance geográfico de IX.br-SP, CABASE-BUE y PIT Chile-SCL durante su proceso de consolidación como centros de intercambios regionales. Por consiguiente, en esta sección investigaremos la ubicación de los ASes visibles en los IXPs. Esto lo llevaremos a cabo por medio de la utilización de datos de geolocalización provistos por la Interfaz de Programación de Aplicaciones (del inglés *Application Programming Interface*, API) de AS-RANK. En particular, nuestro análisis se limitará a los stub ASes visibles, de manera tal de mitigar imprecisiones presentes en los datos de geolocalización y remover la influencia del *customer cone* para determinar las coordenadas del AS (ver Sección 3.2).

La Figura 3.5 presenta por medio de círculos la ubicación (latitud y longitud) de los stubs ASes visibles en IX.br, CABASE-BUE y PIT Chile, indicando por medio de colores colores verde (IX.br), azul (CABASE) y rojo (PIT Chile) en qué IXP es visible el AS localizado. También, por medio de estrellas de color amarillo se indican las ubicaciones de los IXPs. Esta figura detalla la evolución a lo largo de la última década por medio de en cuatro capturas en momentos diferentes: Abril de 2011 (Fig. 3.5a), Septiembre de 2013 (Fig. 3.5b), Abril de 2016 (Fig. 3.5c), Septiembre de 2018 (Fig. 3.5d).

Observamos que durante la última década, donde sucedió la consolidación de los IXPs en Argentina, Brasil y Chile, el alcance geográfico también tuvo una notable expansión. Mientras que IX.br-SP contaba con una rala presencia en la región sudeste de Brasil



(a) Abril 2011

(b) Septiembre 2013

(c) Abril 2016

(d) Septiembre 2018

Figura 3.5: Evolución del alcance geográfico de IX.br-SP, CABASE-BUE y PIT Chile-SCL en cuatro momentos durante la última década.

en 2011 (región más poblada del país), esta vio dotada de significativos avances ya para la muestra obtenida cinco años más tarde (Abril 2016). Además, durante el período Abril 2016-Septiembre 2018 IX.br-SP incrementó sustancialmente su alcance en la región Nordeste. De manera similar sucedió en CABASE-BUE, donde en 2011 sólo contaba con una decena de stubs en la región centro de Argentina, mientras que en la última muestra (09/2018), su presencia abarca desde Calafate (Santa Cruz, Argentina), hasta Colombia. Finalmente, la expansión de PIT Chile-SCL no ha podido ser capturada debido a que este es un IXP que comenzó a operar en 2016. Sin embargo, en el corto período de tiempo en el cual hemos podido monitorear su evolución, hemos visto que su alcance ha podido trascender las fronteras, alcanzando ASes en Buenos Aires (Agentina) y Rio de Janeiro (Brasil).

La Figura 3.5 también muestra la presencia de círculos fuera de la superficie del continente. Analizando cada uno de estos puntos, pudimos detectar que éstos corresponden a dos tipos de ASes. El primer tipo corresponde a ASes con gran extensión territorial, donde el hecho de resumir la ubicación de la red en un único punto genera una ubicación ficticia, y en este caso por fuera del terreno del continente. El segundo tipo corresponde a ASes ubicados en islas, donde las bases de geolocalización aparentan ser más imprecisas. Estas conclusiones nuevamente refuerzan la necesidad de abordar métodos más certeros para poder indicar la ubicaciones de los ASes.

3.5. IXPs y concentración

Nuestra hipótesis *a priori* que la presencia de ASes monopólicos puede ser un factor fundamental para desalentar el despliegue/crecimiento de los IXPs. Por lo tanto, observamos si el espacio de direcciones IPv4 delegado a los países latinoamericanos está distribuido de manera equitativa, es decir, si ningún AS posee la mayoría de los prefijos IP asignados a un país.

Para este análisis, consultamos los archivos *pfix2as* de CAIDA de julio de 2019 y los archivos de delegación de LACNIC. Mientras que los primeros se usaron para determinar el conjunto de prefijos *activos* (vistos en las tablas de ruteo) y los ASes que los originaron, los últimos permitieron verificar los países a los que se le delegaron estos bloques de red. Al final, la combinación de ambos conjuntos de datos genera una base de datos que indica, para cada país latinoamericano, todos los prefijos activos y los ASes que los originan. Sin embargo, reconocemos cuatro limitaciones de esta metodología, que detallaremos a continuación

1. Prefijos delegados por otros RIRs (no LACNIC) podrían estar activos en Lati-

noamérica, por ejemplo, prefijos delegados por ARIN a Level3 (AS3356-AS3549) que se están siendo utilizados en Latinoamérica (ver Sección 3.4.4).

2. No podemos determinar cuáles de las direcciones anunciadas se encuentran realmente en uso (DBK+16), es decir, de la totalidad de un prefijo anunciado, cuántas direcciones realmente corresponden a dispositivos activos.

3. Los prefijos delegados por LACNIC a los ASes con personería jurídica en algún país de Latinoamérica pueden estar asociados a interfaces de dispositivos situados fuera de la región.

4. La presencia de Carrier Grade NAT (CGN) (RWVR+16) (ver Sección 1.1.1) podría llevar una sub-representación de los ASes que, aunque originan pequeños espacios de direcciones, tienen una gran cantidad de abonados, en especial, operadores móviles.

Si bien el uso de bases de datos de geolocalización puede mitigar estos problemas, se ha estudiado que estas fuentes son inexactas en muchos casos (PUK+11). En consecuencia, el perfeccionamiento de la metodología seguida para detectar prefijos activos en cada país se deja como trabajo futuro.

Utilizamos nuestra base de datos para calcular el Índice Herfindahl e Hirschman (*Herfindahl-Hirschman Index*, HHI), una medida estadística que informa concentración, la cual varía de 1 (origen monopólico) a 0. Esta métrica es utilizada por el Departamento de Justicia de los Estados Unidos para aplicar las regulaciones antimonopolio (Rho93) y en ecología para medir la diversidad (conocido como *Simpson's Diversity Index*). La Figura 3.6 muestra HHI para países latinoamericanos con más de 1 millón de direcciones IP delegadas. El extremo derecho muestra países con baja tasa de concentración, como Brasil, Chile y Argentina. De hecho, estos países son los que albergan las redes de IXPs más grandes. Por el contrario, el lado izquierdo incluye países como Uruguay, República Dominicana y Venezuela, que no tienen IXPs, y Paraguay, Costa Rica y México, todos con un HHI de más de 0,3.

Tomamos Uruguay, Venezuela, Costa Rica y México como casos de estudio y mostramos en el Cuadro 3.3 los primeros y segundos ASes dominantes que concentran la mayoría de las direcciones IP delegadas a estos países. En todos los casos, el AS más dominante no solo se origina entre 55 % a 90 % de su respectivo espacio nacional de direcciones, sino que también posee al menos 47 % más que el segundo. En particular, los países dominados por grandes proveedores estatales como Venezuela (CANTV) y Uruguay (ANTEL) ni siquiera planean crear un IXP (Fre18; dTdU19). Costa Rica es el ejemplo opuesto: mientras el estado posee ICE (AS11830), el ISP principal que origina

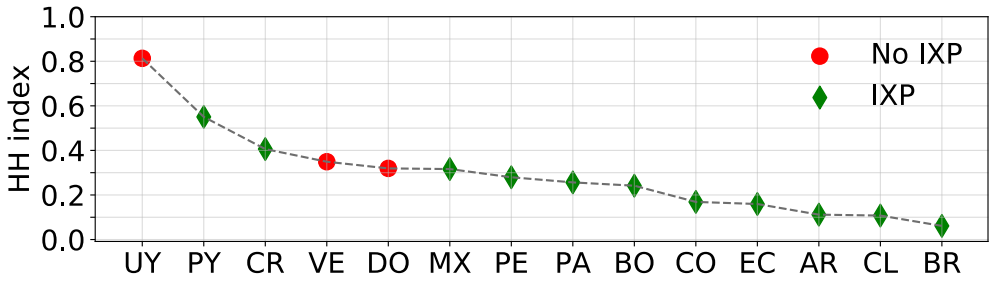


Figura 3.6: Índice Herfindahl-Hirschman (HHI) para determinar la concentración del espacio de direcciones originado en países a los que se les ha delegado más de 1 millón de direcciones IP.

	UY		VE		CR		MX	
ASN	6057*	19422	8048*	6306	11830*	52228	8151	13999
ip-cnt$_{cc}$	2.38M		5.15M		2.42M		24.9M	
ip-cnt	2.15M	90.1k	2.84M	629k	1.52M	197k	13.7M	2.05M
ip-frac	0.90	0.04	0.55	0.14	0.63	0.08	0.55	0.08

Cuadro 3.3: Los dos mayores ASes origen por país. $*$ indica ASes de propiedad estatal.

el 63 % del espacio nacional de direcciones, el primer IXP nacional fue creado por Decreto presidencial en 2014 (ver Sección 3.3). Sorprendentemente, ICE nunca se ha sido miembro del IXP (Nac15). México es otro país con elevado HHI cuyo IXP solo tiene 6 miembros. Sospechamos que, a pesar de que la creación del IXP fue patrocinada por el gobierno mexicano, en 2014, luego de la recomendación de la OCDE (Cas14), la ausencia de Telmex (AS8151) (ITU16), por lejos el AS dominante en el país, desalentó el crecimiento de IXP.

3.6. Literatura relacionada

Aunque Latinoamérica se encuentra sub-representada en los proyectos de medición de Internet, algunos trabajos han analizado específicamente esta región. Berenguer *et al.* (BCAHP16) estudiaron cómo se pueden aumentar las colecciones BGP de RIPE RIS y RouteViews en Latinoamérica mediante el uso adicional de BGP-TDs recolectados en LGs en la región. Kiedanski *et al.* (KGAH18) utilizaron en RIPE Atlas para determinar el estado del despliegue de IPv6 en la región de LACNIC, mientras que el estudio de este capítulo no se centra en IPv6, prevemos extender el análisis como trabajo futuro. Brito *et al.* (BSF+16) estudiaron cuidadosamente la composición e interconexión de la red de IXPs públicos de Brasil en tres momentos, y luego compararon el tamaño del IXP brasileño en términos de ASes conectadas y prevalencia de políticas de interconexión

con IXPs en otras regiones. Formoso *et al.* (FC16) utilizaron las sondas RIPE Atlas en Latinoamérica para crear una matriz de latencia entre países como una forma de detectar rutas fuertemente asimétricas y países defectuosamente interconectados. Más recientemente, Bottger *et al.* (BAF+18) estudiaron el impacto de un gran número de IXPs en el ecosistema ASes durante la última década, donde también cubrió Latinoamérica. Sin embargo, este estudio sólo se basó en muestras de PeeringDB y mediciones activas, que proporcionan una visión acotada de la región debido a los escasos recursos presentes en Latinoamérica durante ese período de tiempo.

Además de los artículos que se centran en Latinoamérica, existe una extensa literatura que estudió IXPs. Dhamdhere *et al.* (DD10) estudiaron cómo los IXPs contribuyeron al aplanamiento del ecosistema de ASes, mientras que Augustin *et al.* (AKW09) midieron cuidadosamente el número de enlaces de p2p observados en IXPs alrededor del mundo. Otros artículos también analizaron la anatomía de grandes IXPs europeos (AKW09), como así también el papel de los IXPs en el ecosistema africano de ASes (FFA15; FVD17; Fan18).

3.7. Conclusiones

Este capítulo aporta cuatro hallazgos con respecto a la investigación de topología de Internet. Primero, descubrimos que los estados latinoamericanos han estado involucrados en la creación de IXPs nacionales de varias maneras: legislación, regulación, patrocinio, financiamiento, operaciones y servicio de tránsito desde/hasta los IXPs. En segundo lugar, descubrimos tres IXPs consolidados, IX.br-SP, CABASE-BUE y PIT Chile-SCL, que reúnen principalmente ASes locales pero también regionales. En tercer lugar, comparamos estos IXPs con otros desplegados en otros continentes y descubrimos que algunos IXPs en regiones en desarrollo no solo han tenido un crecimiento similar en los últimos años, sino que también parecen haber alcanzado la madurez, es decir, han podido atraer la mayor cantidad de ASes locales, al igual que algunos IXPs consolidados en Europa. Sin embargo, los IXPs europeos también han logrado reunir miembros de diferentes regiones, un mercado que podría ser explotado en el futuro por los IXPs menos conocidos, y hasta ahora con enfoques locales en América Latina, Asia y África. Cuarto, estudiamos la correlación entre la existencia de espacios de direcciones concentrados en pocos ASes, y el desarrollo y consolidación de IXPs. De hecho, en varios países latinoamericanos, la existencia de ASes monopólicos, algunos de propiedad estatal, parecen haber desalentado la proliferación de IXPs.

Este capítulo sugiere varias direcciones prometedoras a futuro. Primero, el trabajo podría extenderse estudiando el despliegue de las CDNs en Latinoamérica y su presencia

en los IXPs. En segundo lugar, nos gustaría comparar y contrastar las políticas de *peering* en los IXP de Latinoamérica y del resto del mundo. Tercero, nos interesaría investigar el despliegue de IPv6 en América Latina y el papel de los IXPs en dicho proceso.

Por último, se han publicado dos artículos, en conferencia con proceedings y otros en una revista, derivados de nuestra continua investigación acerca del ecosistema latinoamericano de Internet. Los artículos mencionados son los siguientes,

1. Berenguer, S.S., Carisimo, E., Alvarez-Hamelin, J.I. and Pintor, F.V., 2016, August. Hidden internet topologies info: Truth or myth?. In *Proceedings of the 2016 workshop on Fostering Latin-American Research in Data Communication Networks* (pp. 4-6). ACM.

2. Carisimo, E., Del Fiore, J.M., Dujovne, D. , Pelsser, C. and Alvarez-Hamelin, J. I. 2020. A first look at Latin American IXPs. In *ACM SIGCOMM Computer Communication Review (CCR)*. 50 (18-24).

Capítulo 4

La congestión persistente en Internet

En este capítulo estudiaremos la congestión persistente en Internet, la cual es frecuentemente observada en enlaces entre ASes en los Estados Unidos. Aquí presentaremos una continuación del trabajo existente en la materia, donde las publicaciones previas ya corroboraron la existencia de congestión persistente en Internet e identificaron precisamente los enlaces donde sucede este fenómeno. Por lo tanto, nuestro aporte será analizar el problema desde una perspectiva innovadora por medio del modelado de la latencia en muestras recolectadas en enlaces congestionados. Si bien el modelado de la latencia podría ser en sí mismo una temática a abordar por nuestro campo de investigación, nosotros creemos que el uso de modelos paramétricos puede introducir un nuevo abordaje para identificar enlaces congestionados. Nuestra propuesta se basa en poder detectar la congestión por medio del uso de parámetros derivados del ajuste de nuestros modelos paramétricos, ya que confiamos que dichos parámetros son capaces de resumir fielmente la naturaleza del proceso.

A continuación investigamos el uso de distribuciones estables para modelar muestras de latencia a través de un enlace interdominio obtenidas por medio de mediciones activas desde el borde de la red. Descubrimos que las distribuciones estables son sorprendentemente eficaces en la captura de las características de la distribución de la latencia en muestras pertenecientes a intervalos de tiempo de 10 minutos. Aún más sorprendente, el ajuste paramétrico de distribuciones estables de tales muestras produce valores de parámetros capaces de diferenciar efectivamente los enlaces congestionados de los no congestionados. Luego, exploramos la utilidad práctica de este resultado aplicando un algoritmo de aprendizaje automático (en inglés, *Machine Learning*) para la distribución conjunta de estos parámetros a lo largo del tiempo, y descubrimos que para nuestro dataset (ciertamente limitado) de 125 enlaces entre dominios de los tres principales ISPs en

los EE. UU., el detector identifica enlaces congestionados con una precisión del 80 %. A pesar de la notable complejidad que traen las distribuciones estables, y las relaciones de compromiso requeridas para usarla para este propósito, su capacidad para capturar valores extremos brinda un enfoque prometedor para analizar el comportamiento de latencia de Internet.

4.1. Introducción

La congestión persistente en los enlaces entre dominios de Internet puede potencialmente degradar la Calidad de la Experiencia (del inglés *Quality of Experience*, QoE) experimentada al acceder a servicios populares brindados sobre Internet. La carencia de datos empíricos para ilustrar tal congestión ha inspirado trabajos recientes a detectar y analizar enlaces entre dominios sin acceso directo a ellos[1] (LDC+14; FVD17; DCGG+18). La intuición detrás del método utilizado en estos estudios es que al observar un incremento en la latencia a través de un enlace – más precisamente, mayor latencia del lado del enlace que corresponde al otro dominio sin experimentar tal incremento en el extremo del enlace correspondiente al AS origen – refleja que la cola de ruteador se encuentra llena y por lo tanto proporciona una señal indirecta de congestión en ese enlace.

Dhamdhere *et al.* (DCGG+18) describieron la implementación y uso de un sistema que operacionalizó esta metodología a escala. Los autores utilizaron un conjunto de sondas de la plataforma Ark de CAIDA (CAIa) para develar todos los enlaces de interconexión, a nivel IP, con ASes directamente adyacentes (LDH+16), y luego continuamente (cada 5 minutos durante años) midieron la latencia en el lado cercano y lejano de cada enlace entre dominios. Este método fue batuizado *Time Series Latency Probes* (TSLP) (LDC+14). Para inferir la congestión, Dhamdhere *et al.* basaron su método en la autocorrelación de las muestras de latencia para así detectar varios días sucesivos de demoras elevadas alrededor de las mismas horas a causa de la demanda diurna. Sin embargo, este método para detectar la congestión requiere cierto nivel de inspección manual. El tamaño de Internet inspira dos preguntas: i) ¿podemos automatizar este proceso de inferencia?, ii) ¿podemos hacerlo aprovechando modelos matemáticos que brinden una mayor comprensión de la naturaleza de la latencia de Internet? Un enfoque de modelado matemático podría reformular estas preguntas de la siguiente manera: i) ¿Podemos modelar con precisión la distribución de la latencia, dada su no estacionariedad, debido a la congestión diurna? ii) ¿Puede tal modelado ayudar a clasificar automáticamente los enlaces como congestionados o no congestionados?

En este capítulo presentamos los resultados de un primer experimento en esta direc-

[1] *Acceso directo* hace referencia a acceso a los equipo tal como lo tendría el propietario

ción, explorar el uso de distribuciones estables (Sección 4.3) para modelar el tiempo de ida y vuelta (en inglés *Round Trip Time*, RTT) [2] recopilada mediante mediciones activas hacia enlaces entre dominios. Mostramos que las distribución estables son superiores (en precisión) a otros enfoques para el modelado de la distribución de la latencia (Sección 4.4). Analizamos las características de las distribuciones estables ajustadas a los RTTs, y descubrimos que los parámetros derivados de este ajuste permiten obtener información capaz de diferenciar efectivamente enlaces congestionados y no congestionados (Sección 4.5). Explicamos la intuición detrás de hallazgo, y exploramos la utilidad práctica de este conocimiento mediante el uso de un algoritmo de aprendizaje automático (del inglés *Machine Learning*, ML) para clasificar los enlaces como congestionado o no congestionado (Sección 4.6). Obtuvimos una precisión del 83 % en un dataset relativamente pequeño de 125 enlaces entre dominios que conectan los tres principales ISPs de banda ancha de EE.UU. a otros ASes. Vemos el trabajo presente en este capítulo como un paso en la larga trayectoria en el desafío de aplicar modelos matemáticos para comprender la dinámica del tráfico de Internet, y la esperanza que otros encuentren este estudio útil para explorar otras distribuciones u otros enfoques para automatizar la inferencia de la congestión en los enlaces entre dominios.

4.2. Otras distribuciones propuestas para modelar la latencia

Un modelo estadístico paramétrico se basa en un número fijo de parámetros para caracterizar una distribución de probabilidad. Nuestro objetivo es poder interpretar los fenómenos que ocurren en la red por medio de los parámetros que definen a la distribución de probabilidad. Sin embargo, la inferencia de los parámetros que definen una distribución de probabilidad, por ejemplo las distribuciones estables, pueden ser una tarea sumamente compleja. Esto plantea si realmente es necesario contar con complejos modelos parámetricos para estudiar la red en lugar de utilizar métodos estadísticos simples tales como la media, el desvío estándar o los percentiles. También plantea el desafío de encontrar una distribución que sea capaz de representar con precisión los datos. Pero para la latencia, el modelado paramétrico puede ofrecer ventajas significativas. Por ejemplo, las métricas estadísticas como la media y el desvío estándar requieren que los datos tengan una media o varianza finita Ademas, la media y el desvío estándar son sensibles (Hub64) ante la presencia de valores atípicos (en inglés *outliers*) que aparecen en las colas pesadas de muchas distribuciones de latencia. Otro inconveniente registrado

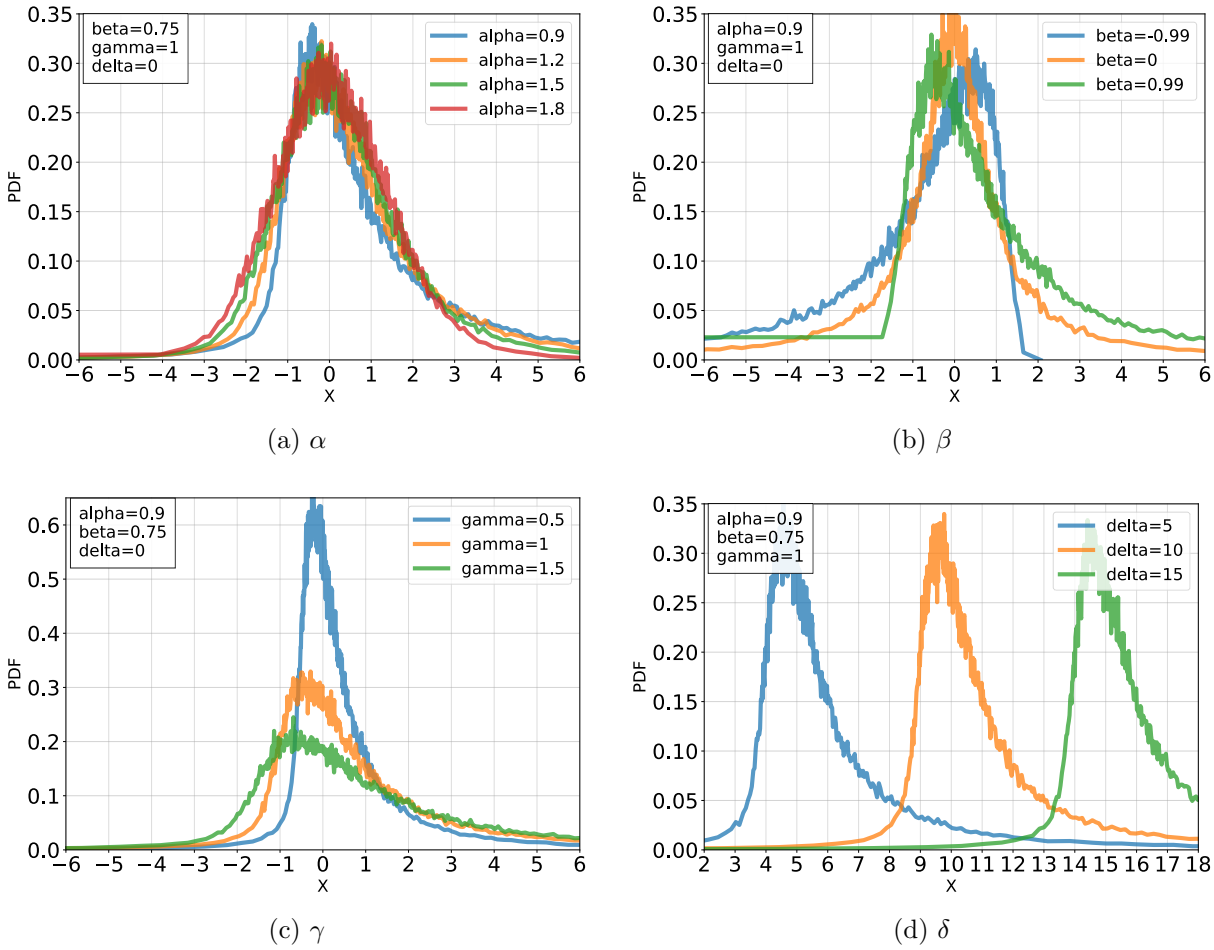
[2] Aclaración: A lo largo del capítulo utilizaremos indistintamente los términos latencia y RTT.

en el caso de la latencia es que la varianza no cuenta con convergencia aparente a medida que crece el número de muestras (CB97).

Varios estudios han modelado las distribuciones de la latencia de Internet utilizando modelos paramétricos. Fontugne *et al.* (FMF15) utilizaron una mezcla de variables aleatorias log-normales para describir la distribución de RTTs de los flujos TCP provenientes de diferentes continentes. Papagiannaki *et al.* (PMF+03) descubrieron que la distribución de Weibull capturaba la demora de la cola (medida pasivamente) en un router central troncal. Hernández *et al.* (HP06) utilizaron una mezcla de variables aleatorias de Weibull para modelar mediciones de retardo unidireccionales de monitores de RIPE. Estos estudios encontraron éxito con diferentes enfoques de modelado para el mismo tipo de fenómeno (RTT) tal vez debido a las diferencias en los métodos de medición, muestreo o agregación. Ninguno de estos estudios se enfocó en las alteraciones de la distribución de la latencia para inferir la dinámica del tráfico. Nuestra hipótesis es que la distribución estable puede capturar con mayor precisión las colas pesadas y los sesgos (asimetría) observados en los datos de latencia, así como la naturaleza aleatoria de RTT en las mediciones de sondeo activas. A través de una representación más precisa de la naturaleza de la latencia podremos identificar más certeramente muestras correspondientes a períodos de congestión y de no congestión.

4.3. La distribución estable

La distribución estable es una familia de distribuciones de cuatro parámetros. $S(\alpha, \beta, \gamma, \delta)$, donde $\alpha \in (0, 2]$ es el parámetro característico que regula la caída de la cola, $\beta \in (-1, 1)$ es el parámetro de inclinación o asimetría, $\gamma \in \mathbb{R} > 0$ es el parámetro de escala y $\delta \in \mathbb{R}$ es el parámetro de ubicación. El parámetro de escala regula el ancho (la dispersión aunque no es el desvío estandar) de la distribución. Ecuación 4.1 muestra la función característica $g(k)$ de las distribución estables, entre varias parametrizaciones posibles disponibles para definirla.



(a) α (b) β

(c) γ (d) δ

Figura 4.1: Ejemplos de parametricaciones de las distribuciones estables. Cada una de las figuras muestra como afecta la variación de uno de los parámetros cuando el resto permanece invariante.

$$g(k) = \exp\{\delta[ik\gamma - |k|^{\alpha}\omega(k;\alpha,\beta)]\} \tag{4.1}$$

$$\omega(k;\alpha,\beta) = \begin{cases} \exp[-i\beta\Phi(\alpha)\text{sign}(k)], & \alpha \neq 1 \\ \pi/2 + i\beta \log|k|\text{sign}(k), & \alpha = 1 \end{cases} \tag{4.2}$$

$$\Phi(\alpha) = \begin{cases} \alpha\pi/2 & \alpha < 1 \\ (\alpha-2)\pi/2 & \alpha > 1 \end{cases} \tag{4.3}$$

Para poder familiarizarnos con las distribuciones estables en la la Figura 4.1 presentaremos algunos ejemplos. Además aprovecharemos para explicar como cada uno de los parámetros modifican las distribuciones. La Figura 4.1 muestra cuatro gráficos en donde cada uno de ellos presenta las variaciones en las funciones de distribución empíricas

al variar uno de los parámetros de las distribuciones estables. Cada una de las curvas presentadas fue construida a través de la generación de muestras aleatorias de las parametrizaciones indicadas en los gráficos. En la Figura 4.1a muestra la variación de la función de densidad al variar α, parámetro que regula la caída de la cola. Aquí se muestra cómo al reducir el valor de α, la cola de la función de densidad reduce su pendiente. La Figura 4.1b muestra como la cola de la distribución es a izquierda o derecha, dependiendo de si β toma valores negativos o positivos. El espesor del lóbulo de la distribución se encuentra determinado por el parámetro γ, y tal como se muestra en la Figura 4.1c, el espesor se reduce cuando gamma toma valores más cercanos a cero. Por último, la Figura 4.1d muestra como incrementos de δ generan una traslación de la función de densidad.

La complejidad de las distribuciones estables crean algunos desafíos: ni las distribuciones ni las densidades de la distribución estable se puede expresar en términos de funciones elementales (polinomios, logaritmos, etc.) (UZ99). Solo dos casos excepcionales de la familia de distribución estable pueden expresarse en términos de funciones elementales, y en particular se trata de funciones de probabilidad ampliamente conocidas: la distribución normal y la distribución de Cauchy. La distribución estable produce una distribución normal si $\alpha = 2$, $\beta = 0$, en cuyo caso γ y δ corresponden a los parámetros σ y μ (varianza y media). Recientemente, Julian *et al.* (JMdVG+17) desarrolló un método paralelo basado en el cálculo en GPU que puede ajustar de forma rápida y precisa la distribución estable a los datos empíricos.

A pesar de la complejidad inherente de la distribución estable, una gran cantidad de literatura en otros campos (PMH+93; SGGR+13; ABT01), ha explorado la distribución estable como un enfoque prometedor para tratar con colas pesadas. Los economistas lo han utilizado durante décadas para estudiar patrones económicos, por ejemplo, la variabilidad en los rendimientos de los activos financieros (Man63; Fam63; MS95). Antes de que Mandelbrot introdujera esta distribución en ese campo, los economistas modelaron los rendimientos de los activos financieros $R(t)$ ajustando una distribución normal a las diferencias en muestras consecutivas (Bac00), tal como se muestra en la Ecuación 4.4.

$$R(t) = X(t+T) - X(t) \tag{4.4}$$

La Ecuación 4.4 determina el rendimiento $R(t)$ calculado en función de muestras de un proceso estocástico $X(t)$, separadas entre sí por un intervalo de tiempo T . Sin embargo, la presencia de colas pesadas en las muestras del proceso estocástico llevará a observar el mismo fenómeno en el rendimiento $R(t)$, por lo cual la distribución normal es incapaz de capturar fehacientemente este fenómeno Al observar las imprecisiones del

modelo, Mandelbrot descubrió que las distribuciones estables se ajustaban correctamente al comportamiento aleatorio de los rendimientos financieros si éstos se modelaban mediante, $\log(X(t+T)) - \log(X(t))$, en lugar de la fórmula de la Ecuación 4.4. Inspirado por su amplio éxito en economía, consideramos su capacidad para capturar colas pesadas en mediciones de latencia; también adoptaremos la idea de Mandelbrot de modelar el $\log(RTT)$ en lugar de RTT directamente capturado en las mediciones.

4.4. Datos utilizados

La precisión del ajuste de cualquier modelo paramétrico depende del número de muestras y, desafortunadamente, la frecuencia de muestreo de 1 cada 5 minutos utilizada en estudios previos (FVD17; DCGG+18) es demasiado escasa para ajustar una distribución de probabilidad a los datos de latencia. Para explorar la efectividad del ajuste de la distribución en los datos de latencia utilizamos muestreos de alta frecuencia, 1 muestra por segundo, pero dirigimos estas campañas de mediciones solo a los enlaces que muestran evidencia de patrones de congestión diurna. La reducción del alcance de nuestra campaña buscar evitar sobrecargar ruteadores con consultas a ruteadores no congestionados, y acotar el consumo del ancho de banda en la red donde se encuentra el punto de medición. Tal evidencia utilizada fue adquirida por medio de la metodología TSLP (ver Sección 4.1). Dado que los ruteadores son capaces de imponer límites a la tasa de respuestas a nuestras peticiones, sorteamos esta dificultad enviando paquetes limitados por TTL que expiran en el ruteador objetivo, en lugar de simplemente enviar la petición directamente al ruteador.

Nuestro conjunto de datos contiene aproximadamente 416 millones de muestras de RTT sin procesar, recolectadas durante 2017 de 5 puntos de medición (en inglés *Vantage Point*, VP), los cuales los nombraremos como VP1, VP2, ..., VP5, en tres de los principales ISPs de banda ancha de EE.UU. La muestras fueron recolectada producto de ejecutar mediciones activas hacia 16 ASes vecinos, los cuales son nombrados como CP1, CP2, CP3 o Transit1, Transit2, ... Transit13, dependiendo si el destino es un proveedor de contenidos (CP) o de tránsito (Transit). Esta colección cuenta con un total de 125 enlaces entre dominios a nivel IP y 1667 tuplas únicas `(punto de medición, día, dirección del ruteador)` .

Aplicamos estimadores de máxima verosimilitud para ajustar la distribución estable (`libstable`[3] (JMdVG+17)), la distribución de Weibull (`python scipy`) y la distribución lognormal (`python scipy`) a aproximadamente 800,000 ventanas de datos de diez

[3] https://github.com/hpcn-uam/libstable-opencl

minutos, con un máximo de 600 muestras por ventana, desde un punto de observación hacia un enlace objetivo.

Al igual que previos trabajos en otros campos que ajustaron las distribuciones estables a datos empíricos, como por ejemplo Salas-Gonzalez *et al.* (SGKR09), medimos la bondad de ajuste para cada distribución utilizando la divergencia Kullback-Leibler (D_{KL}) entre datos empíricos (RTT) y muestras generadas mediante el uso de parámetros de máxima verosimilitud (una divergencia más baja implica un mejor ajuste). Para nuestros datos, la distribución estable tiene el menor D_{KL} de los tres modelos en el 99.87 % de las ventanas de 10 minutos. Entre esos casos, el valor medio de D_{KL} entre la distribución estable y los datos empíricos es 0.017, mientras que el valor medio de D_{KL} para Weibull y log-normal son 1.18 y 1.36 respectivamente.

Sumado a las mediciones a través de la divergencia Kullback-Leibler (D_{KL}), una ventana de tiempo de 10 minutos cuenta con un número suficiente de muestras para tener un error bastante bajo en los parámetros estimados, y, además, asumimos que la latencia es casi estacionaria dentro de ese período. Al desarrollar este método computacional para ajustarse a la distribución estable, Julian *et al.* (JMdVG+17) presentó que 1000 muestras fueron suficientes para contar con un ajuste preciso, por lo que también disponemos de cierta confianza para asumir que 600 muestras son suficientes.

4.5. Firmas de congestión

La latencia que medimos es la suma de fenómenos determinísitcos y aleatorios en cada ruteador que el paquete atraviesa en la ruta directa e inversa. El retraso en la cola es un componente de la latencia, y corresponde al retraso que experimenta un paquete entre su llegada a un ruteador y cuando este es finalmente serializado. A continuación, ajustamos la distribución estable a las muestras del RTT y examinamos la evolución de los parámetros de la distribución estable a lo largo del tiempo.

4.5.1. Desplazamiento de la distribución con la demanda diurna

Los enlaces que bajo el método de autocorrelación+TSLP (DCGG+18) fueron clasificados como congestionados durante las horas pico mostraron una diferencia significativa de sus RTTs durante la hora pico y fuera de ella. La figura 4.2 presenta los histogramas de las muestras del RTT en ventanas de una hora recolectadas por un punto de observación hacia el extremo más alejado de un enlace entre dominios. Tanto la mayor parte de la distribución como su moda cambian durante las horas pico (entre las 16:30 y las 21:30).

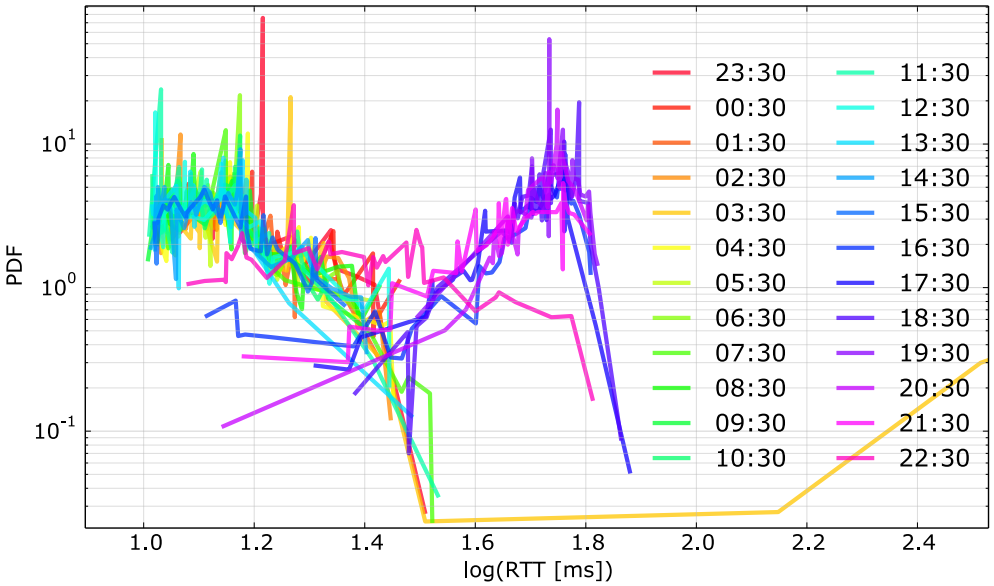
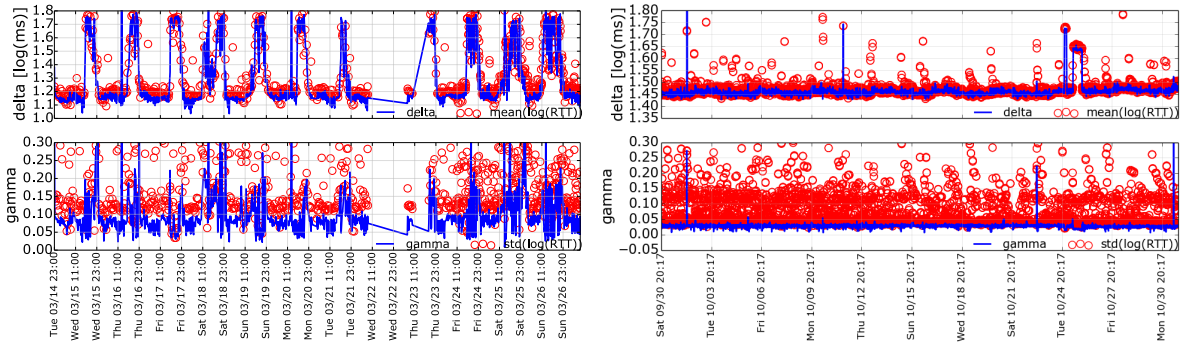


Figura 4.2: Histogramas por hora de los RTTs obtenidos del extremo lejano un enlace ente dominios, entre un gran proveedor de acceso y gran proveedor de contenido (17 de marzo de 2017). La distribución de latencia se desplaza hacia valores mayores durante las horas *pico*, reflejando la no estacionariedad del proceso. Las distribuciones de latencia en las horas de menor demanda se exhiben asimétricas a la derecha. La ausencia de muestras en el lado izquierdo de la distribución se produce por la restricción del tiempo de propagación, generando la distribución asimétrica a derecha.



(a) VP1 hacia CP1 (14/03/17 - 26/03/17, conges-tionado)

(b) VP2 hacia Transit1 (30/09/17 - 30/10/17, no congestionado)

Figura 4.3: Las líneas azules continuas muestran la evolución de δ y γ (ver Ecuación 4.1), mientras que los puntos rojos representan la media y el desvío de las muestras $\log(RTT)$. Cada punto representa datos en una ventana de tiempo de 10 minutos. La media y δ siguen una tendencia similar, aumentando durante condiciones de congestión (a). Por el contrario, el desvío estándar aumenta durante las horas de menor actividad, mientras que γ permanece bajo.

Las distribuciones de latencia más típicas (no congestionadas) exhiben colas pesadas y sesgadas a la derecha. El cambio de ubicación de estos histogramas por hora demuestra que cuando se observa en el transcurso de un día típico, la latencia es un proceso no estacionario. Desafortunadamente, la estimación de máxima verosimilitud supone una variable aleatoria estacionaria. Para lidiar con la no estacionariedad, ajustamos las distribuciones estable a las muestras usando `libstable` [4] (JMdVG+17) en ventanas de tiempo de 10 minutos en las que suponemos que la latencia es un proceso estacionario. Por lo tanto, por cada ventana de 10 minutos desde un punto de observación hacia a un enlace destino, extraemos un conjunto de parámetros para la distribución estable que mejor se ajustan a los datos en esa ventana. Esta elección de una ventana de 10 minutos representa una relación de compromiso entre el número de muestras, y por ende bondad de ajuste, y la estacionariedad. Probamos ventanas de tiempo de 5 minutos y 60 minutos; mientras que el primero tuvo un ajuste significativamente peor (mayor D_{KL}) posiblemente por la escasez de muestras, el segundo combinó ventanas con un comportamiento diferente, debido a la no estacionariedad del proceso. Esta comprobación empírica ratifica nuevamente la elección del uso ventanas de 10 minutos.

4.5.2. Evolución de las series temporales

Lo más interesante que descubrimos es el comportamiento de dos parámetros, γ y δ, de las distribuciones estables, y cómo reflejan las características de latencia y la dinámica de tráfico asociada, a lo largo del tiempo (Figura 4.3). Los otros dos parámetros, α y β, son menos interesantes, como explicamos a continuación. Las restricciones de velocidad de la luz sobre el retraso de propagación, es decir, que ante la ausencia de cualquier otra fuente de retraso, el paquete demorará en alcanzar su destino y retornar un tiempo mínimo determinado por la distancia física entre el VP y el destino. Dado que el RTT nunca podrá ser menor que el tiempo de propagación, y que la presencia ocasional de retraso en las colas es siempre mayor a cero, implican que una distribución de latencia siempre se desplace (β) hacia la derecha. Descubrimos que la tendencia en α es menos reveladora que γ y δ, y observamos que los economistas que usan la distribución Estable para estudiar la volatilidad del índice S&P[5] descubrieron que γ era más sensible a la volatilidad que α (MS95).

Es importante recordar que la distribución estable se reduce a la distribución normal en los casos en que $\alpha = 2$ y $\beta = 0$, en cuyo caso γ y δ son el desvío estándar y la media, respectivamente. Por lo tanto, comparamos estos dos pares de parámetros: δ

[4] `https://github.com/hpcn-uam/libstable-opencl`
[5] S&P: Standard & Poor's Index. `https://www.standardandpoors.com/en_US/web/guest/home`

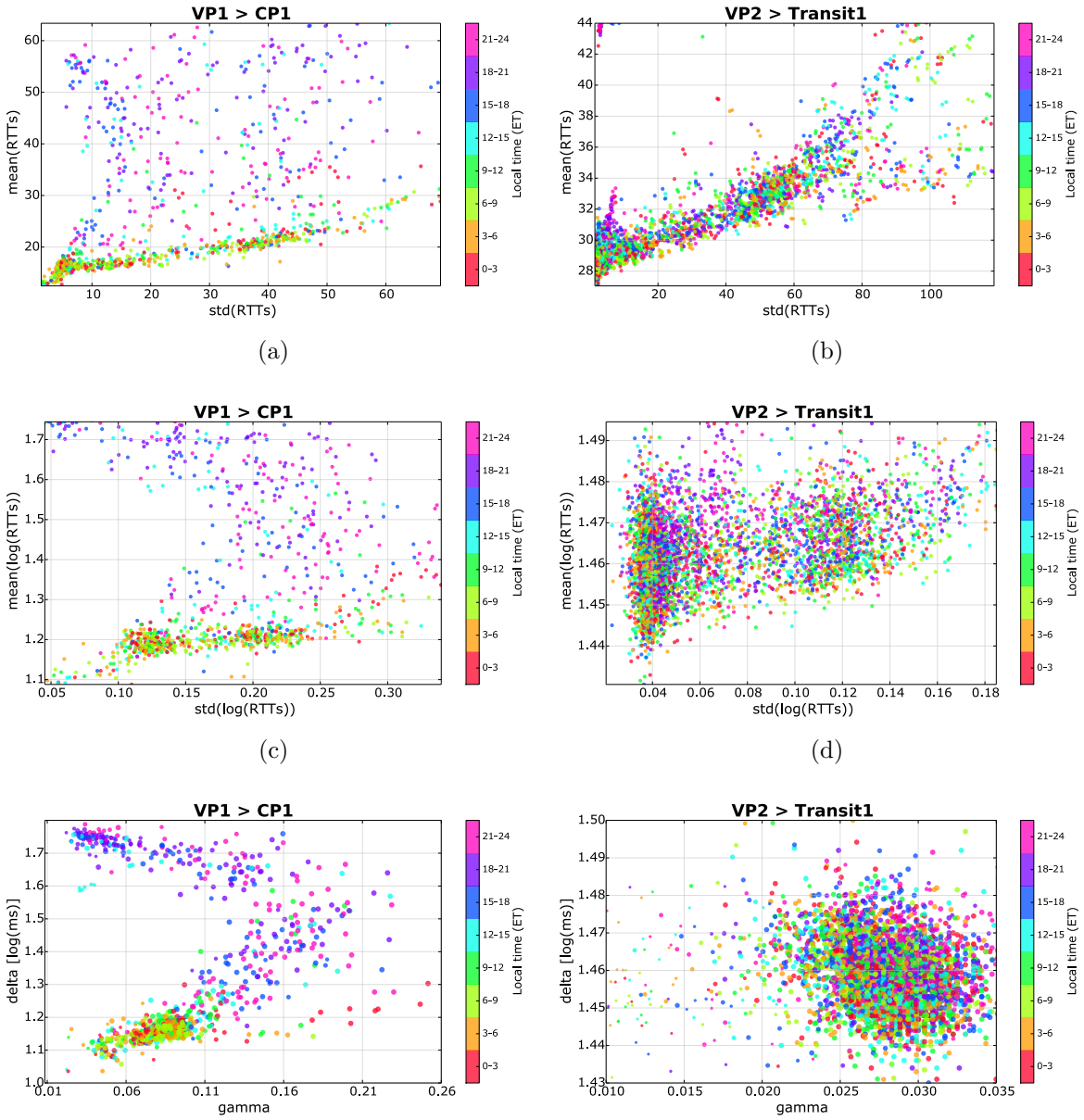
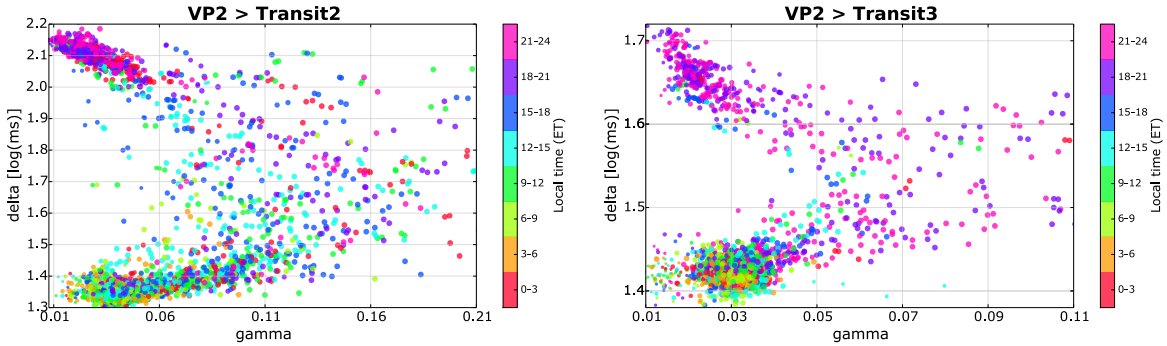
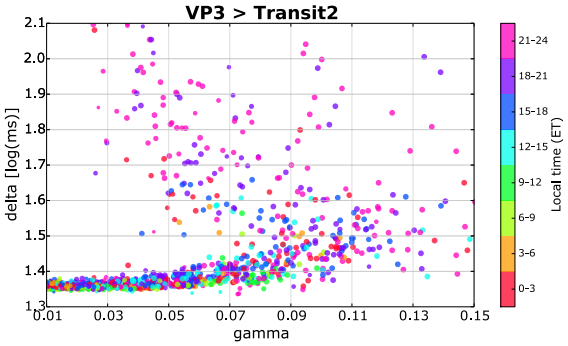


Figura 4.4: Firma de latencia por pares en un enlace congestionado (izquierda) y no congestionado (derecha). Los puntos representan pares $(std, mean)$ para muestras de RTT sin procesar (4.4a, 4.4b), pares $(std, mean)$ para muestras de $\log(RTT)$ (4.4c, 4.4d) y pares de parámetros de la distribución estable (γ, δ) (4.4e, 4.4f), cada uno para una ventana de 10 minutos. Las figuras 4.4e y 4.4f muestran una estructura clara. Las estadísticas más simples (4.4a,4.4b,4.4c,4.4d)) se dispersan y no se correlacionan con el estado de congestión del enlace, pero los parámetros Estable muestran diferencias dramáticas según si un enlace muestra patrones de congestión diurna. La figura 4.4e (un enlace congestionado) traza una forma *boomerang*, con γ en un rango estrecho durante las horas de menor actividad. El enlace no congestionado no muestra ese patrón.



(a)

(b)



(c)

Figura 4.5: Formas de boomerang para (γ, δ) indican congestión para enlaces que conectan proveedores de acceso con diferentes tipos de redes adyacentes (por ej., Proveedores de contenido, proveedores de tránsito, etc).

versus la media, y γ versus el desvío estándar. Esta comparación la llevamos a cabo en la Figura 4.3 donde se presentan dos enlaces, uno congestionado y otro no congestionado. En el caso del enlace congestionado, la media y δ (gráficos superiores) muestran variaciones significativas durante las horas pico. Por el contrario, ni media ni δ (gráficos superiores) muestran un patrón diurno para el enlace no congestionado (Figura 4.3b).

Mientras que δ y la media capturan algún aspecto de la ubicación de una distribución, γ y el desvío estándar capturan su dispersión. Para la distribución estable, γ regula la concentración de la densidad de la distribución de una variable aleatoria que tiene la distribución estable. El panel inferior de la Figura 4.3a muestra cómo estas variables difieren en lo que capturan. La figura 4.3a muestra aumentos de γ por largos períodos durante las horas pico (congestionadas) pero no durante las horas nocturnas no congestionadas; el desvío estándar tiene valores altos en la mayoría de las ventanas de 10 minutos. El escenario no congestionado (Figura 4.3b) es aún más interesante: el desvío estándar tiene un rango enorme, en contraste con γ. De esta manera, concluimos que γ proporciona más información sobre la dinámica del tráfico subyacente que el desvío estándar.

Por último, observamos en la Figura 4.3b cuatros picos en la serie temporal de δ, siendo el último sostenido por varios minutos. Aunque no podemos contamos con herramientas para poder validar nuestra hipótesis, suponemos que debido a los valores que tomó δ, esto pudo ser a causa de un pequeño incremento de la carga del *slow path*.

4.5.3. Firmas de congestión en pares de métricas

A continuación, exploramos las diferencias entre las tendencias de tres pares de métricas calculadas sobre los RTTs en cada ventana de 10 minutos: 1) media y desvío estándar calculados en los RTTs sin procesar $(raw(std, mean))$, 2) media y desvío estándar calculados en el logaritmo de los RTT $(\log(std, mean))$ y 3) (γ, δ) obtenidos al ajustar la distribución estable en el logaritmo de los RTTs (log(std, mean)) (Figura 4.4). La columna izquierda (4.4a, 4.4c, 4.4e) corresponde al enlace entre dominios congestionado en la Figura 4.3a, y la columna derecha ((4.4b, 4.4d, 4.4f)) a la no congestionada (ver Figura 4.3b). Las columnas muestran los tres pares de métricas, con cada punto se encuentra coloreado de acuerdo a la hora local de la medición.

Observamos una estructura clara en los gráficos (γ, δ) (Figuras 4.4e y 4.4f), y una alta sensibilidad e inestabilidad de el desvío estándar (Figura 4.3), mientras que γ y δ abarcan un rango estrecho durante las horas de menor actividad. Estos gráficos ilustran que la distribución estable, y en particular estos dos parámetros, captura valores extremos sin distorsión, en contraste con la media y el desvío estándar.

Un análisis más detallado de la forma generada por los parámetros del enlace congestionado (Figura 4.4e), forma que llamaremos *boomerang*, revela alguna explicación de porqué estos dos parámetros en tándem se correlacionan con el estado de congestión en el enlace. Los puntos en la esquina inferior izquierda del boomerang tienen los γ y δ más bajos, y (reflejados en el color) corresponden exclusivamente a las horas de menor actividad. Los puntos de hora pico (15 a 24 ET [6]) abarcan un rango mucho más amplio para δ (1.3-1.75) y γ (0.05-0.2). La región más interesante es la concentración más densa de muestras de horas pico, en la esquina superior izquierda, con el valor más alto de δ ($\approx$ 1,7) pero el más bajo de γ, con valores similares a las horas de menor actividad. Este patrón refleja los retrasos en la cola durante las horas pico: en cierto punto, la cola está tan persistentemente llena, y la latencia es elevada (reflejada a través de δ). Producto de la ocupación de la cola, esta no permite variabilidad entre muestras, lo que se reflejada a través de la disminución de γ. Contraste estos valores de parámetros con los del enlace no congestionado (Figura 4.4f): ni δ ni γ aumentan significativamente, y los colores (Figura 4.3b) no sugieren una indicación similar de congestión persistente.

De los 125 enlaces entre dominios en nuestro conjunto de datos, los enlaces congestionados casi siempre presentaban patrones (γ, δ) como el *boomerang* de la Figura 4.4e, los enlaces no congestionados no presentaban ninguna estructura. Sin embargo, la variedad de proveedores de acceso y redes adyacentes representadas en nuestros datos produjo una variedad de diferentes tipos de boomerangs, ilustrados en la Figura 4.5.

4.6. Detección de la congestión por medio de Aprendizaje Automático

Para investigar la utilidad práctica de este resultado, probamos un enfoque de aprendizaje automático supervisado usando los parámetros (γ, δ) para detectar automáticamente cuándo la firma de latencia de un enlace refleja la congestión de las horas pico. Para entrenar un detector a través de ML, utilizamos como evidencias de campo las etiquetas de congestión generadas a través el método de autocorrelación presentado en Dhamdhere *et al.* (DCGG+18), en donde los autores validaron estas afirmaciones por medio cuatro fuentes de datos independientes: pérdida de paquetes, rendimiento de video y mediciones de rendimiento, y datos de tráfico de los operadores. También estos autores inspeccionaron manualmente los datos en busca de enlaces los cuales corroboraron como congestionados para evitar falsas asignaciones de congestión.

[6] ET: Huso Horario del Este de los Estados Unidos

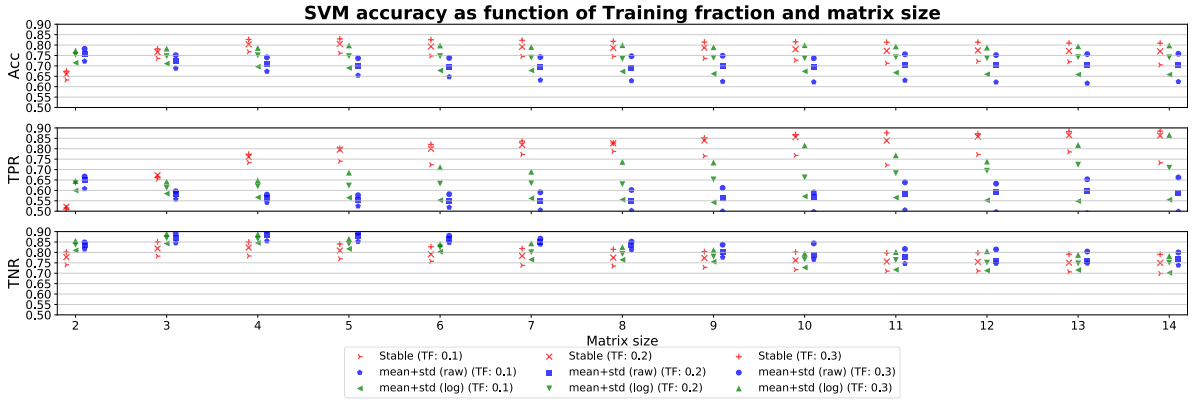


Figura 4.6: Exactitud (en inglés, accuracy), sensibilidad (en inglés recall, o True Positive Rate, TPR) y precisión (en inglés precision, o True Negative Rate, TNR) de tres clasificadores utilizando como entrada raw(std + mean), log(std + mean) y (γ, δ). El clasificador basado en la distribución estable es más preciso para la misma fracción de entrenamiento, su precisión alcanza el máximo en 83 %. Este clasificador también tiene la sensibilidad más alta (casi 90 %) (TPR).

Dado que queremos que nuestro detector infiera la congestión basada en la estructura de diagrama de dispersión (γ, δ), construimos grillas cuadradas donde contamos la cantidad de puntos inscriptos dentro de los limites de cada celda de la grilla. Estas grillas, las cuales las representaremos por medio de matrices, fueron luego normalizadas por el número total de muestras, a efectos de reportar la fracción de muestras en una región dada del diagrama de dispersión, lo que permite la comparación entre muestras con diferentes números de pares pero de forma similar. Dado que el tamaño de la matriz puede afectar la fase de entrenamiento del detector, así como la forma del patrón, exploramos diferentes tamaños de matriz de 2 a 15. La firma de congestión diaria tiene solo 144 pares (cada uno para una ventana de 10 minutos). Para evitar histogramas 2D ralos, no extendemos el análisis más allá de matrices de 15 x 15 por dimensión. Podemos resumir este proceso en tres pasos: (1) separamos una serie temporal de mediciones de RTTs desde un punto de medición hacia un enlace entre dominios en ventanas de 10 minutos; (2) ajustamos la distribución estable a cada ventana; (3) utilizamos los cuatro parámetros de la distribución Estable de cada ventana para construir una firma de congestión diaria, es decir, una matriz que captura la forma del patrón.

Para implementar el aprendizaje supervisado, utilizamos *Support Vector Machine* (SVM) ([CV95]), un clasificador lineal binario que crea un hiperplano para discernir de qué lado está el vector de entrada, en nuestro caso *congestionado* y *no congestionado*.

Utilizamos una fracción de nuestros 1667 casos para entrenar el detector y los casos restantes para probar su rendimiento. Consideramos cada tupla única `(punto de medición, día, dirección del ruteador)` como un caso distinto. Nuestros datos pre-

sentaron dos desafíos. Primero, este contiene más enlaces no congestionados (1021) que congestionados (646), lo que puede conducir a una detección inexacta (CJK04). Como nuestros enlaces ya estaban clasificados de antemano, seleccionamos al azar el mismo número de casos para cada categoría. En segundo lugar, el conjunto de entrenamiento es relativamente pequeño, lo que puede dificultar el rendimiento de clasificación (especialmente la solidez a valores extremos), por lo que utilizamos *data augmentation* (TW87) para expandir el conjunto de entrenamiento creando copias ligeramente modificadas de los datos de entrada. Utilizamos un factor de aumento de 10, es decir, generamos 9 copias de cada entrada original, introduciendo ruido gaussiano blanco aditivo en pares (γ, δ), donde las perturbaciones tenían una relación señal a ruido de 0,05. Finalmente, volvimos a escalar cada entrada del detector aplicando la estandarización, que consiste en restar la media y dividirla por el desvío estándar.

Desempeño del clasificador: Para evaluar el rendimiento del detector, utilizamos una fracción de los casos como entrada de entrenamiento y el resto para evaluar su rendimiento. Para evitar sesgos, probamos el detector SVM 100 veces con diferentes conjuntos de entrenamiento generados aleatoriamente y utilizando diferentes tamaños de fracciones de conjuntos de entrenamiento que van desde 0.1 a 0.3.

Utilizamos tres métricas para medir el desempeño de nuestro clasificador: la exactitud (en inglés, *Accuracy*, Acc), sensibilidad (en inglés *recall*, o *True Positive Rate*, TPR) y precisión (en inglés *precision*, o *True Negative Rate*, TNR). Estas se definen a través de las siguientes fórmulas

$$Acc = \frac{vp + vn}{vp + vn + fp + fn} \tag{4.5}$$

$$TPR = \frac{vp}{vp + fn} \tag{4.6}$$

$$TNR = \frac{vp}{vp + fp} \tag{4.7}$$

Las ecuaciones 4.5, 4.6 y 4.7 son se definen a través del uso de cuatro tipos de muestras: verdaderos positivos (vp), verdaderos negativos (vn), falso positivo (fp) y falso negativo (fn). Los primeros dos casos (vp y vn) corresponde a que el valor real de la muestra y el valor precedido concuerdan, mientras que los otros dos (fp y fn) dos casos presentan oposición entre el valor real de la muestra y de su predicción.

La figura 4.6 muestra la exactitud, sensibilidad y precisión del detector SVM, para los tres tipos diferentes de gráficos de dispersión (raw(std + mean), log(std + mean) y (γ, δ)) utilizados como entrada, tres fracciones de entrenamiento diferentes y un número variable

de dimensiones de nuestra matriz. Nuestro enfoque de ML usando patrones (γ, δ) siempre tuvo una mayor exactitud y rendimiento que los otros dos métodos, para cada fracción de entrenamiento y tamaño de matriz, con la única excepción de las matrices de 2x2. La exactitud varía de 0.63 a 0.84 para patrones (γ, δ), de 0.71 a 0.79 para log(std+mean) y de 0.62 a 0.75 para raw(std + mean). Además, los patrones (γ, δ) alcanzan los valores TPR más altos: este método es capaz de detectar certeramente un enlace congestionado en casi el 90 % de los casos. Sin embargo, los patrones (γ, δ) generalmente exhiben las tasas de TNR levemente más bajas, aunque es similar al TNR de los otros métodos y es mayor a 0.8 en la mayoría de los casos.

4.7. Conclusiones

Este capítulo aporta tres nuevas ideas y técnicas a la investigación de la medición en Internet. Primero, consideramos relevante introducir la distribución estable a nuestro campo, en nuestro caso, modelar latencias de enlaces en el contexto de inferir patrones diurnos de congestión de enlaces. En segundo lugar, aunque la distribución estable no es un modelo definitivo en estudio de la congestión, hemos demostrado que la distribución conjunta de sus parámetros (γ, δ) distingue claramente los enlaces con patrones persistentes de congestión, creados por colas completas en los buffers de los ruteadores. En tercer lugar, utilizamos un algoritmo de aprendizaje automático para construir un detector capaz de clasificar un enlace como congestionado o no congestionado para patrones de (γ, δ) en un día determinado. Nuestro clasificador aún no tiene una precisión del 100 % en la detección de enlaces congestionados, pero incluso una precisión de más del 80 % puede reducir en gran medida la necesidad de inspección humana de las inferencias de congestión.

El trabajo presentado en este capítulo sugiere varias direcciones prometedoras. Primero, planeamos expandir el número de puntos de observación, la duración del período de medición y utilizar metadatos (AS anfitrión, ubicación), todo lo cual mejorará nuestra capacidad de entrenar clasificadores de ML. También consideraremos diferentes transformaciones en la entrada y otros algoritmos de ML, como las redes neuronales. Analizaremos el uso del historial de latencias de enlace durante varios días, en lugar de considerar cada día de forma independiente. Finalmente, según nuestro hallazgo de que durante la congestión, δ crece pero γ sigue siendo pequeño, por lo que el seguimiento de la evolución de los pares (δ, γ) puede permitir la detección de congestión casi en tiempo real en futuros sistemas de medición.

Finalmente, queremos mencionar que el trabajo que fue presentado en este capítulo está será próximamente enviando a alguna conferencia del área de mediciones de Inter-

net.

Capítulo 5

Conclusiones

En esta tesis estudiamos diferentes aspectos de la topología de Internet, y el vínculo de cada uno de ellos con respecto a la distribución y generación de tráfico.

En el Capítulo 2 demostramos a través de la descomposición en k-núcleo el ascenso de los proveedores de contenido al núcleo de Internet. Hemos demostrado cómo debido a la consolidación del tráfico multimedia transportado sobre la red, un gran número de proveedores de contenido han optado por desplegar y operar su propias Redes de Distribución de Contenido. En este capítulo hemos puesto especial atención a los motivos detrás del despliegue de la infraestructura de los mayores actores en la generación de tráfico. Además, hemos notado como cada vez más proveedores de contenido crean redes densamente conectadas, lo cual genera a su vez un cambio en la topología de Internet.

En el Capítulo 3 investigamos la estructura y evolución de los IXPs en Latinoamérica, una región poco explorada por la literatura. Nuestro aporte ha sido en dos planos diferentes: presentamos nuevos conjuntos de datos y llevamos a cabo un riguroso estudio de la evolución de los mayores IXPs de la región. Nuestro primer aporte fue evidenciar la presencia de colecciones de tablas BGP recolectadas por colectores ubicados en los IXPs, aunque pertenecientes a fuentes de datos raramente consultados. Al darle entidad a estos datos, pudimos avanzar con nuestro segundo aporte, el cual se centró en caracterizar la relevancia de los IXPs más grandes en Latinoamérica. Llevado a cabo este estudio observamos que los sistemas de IXPs en Argentina, Brasil y Chile reúnen casi la totalidad de Sistemas Autónomos de cada unos de esos países. Habiendo alcanzado niveles de consolidación similares a los vistos en grandes IXPs europeos, los grandes IXPs de Latinoamérica se presentan como atractivos puntos de intercambio de tráfico entre proveedores de tránsito, de acceso y de contenido.

Finalmente en el Capítulo 4 estudiamos la congestión persistente en los enlaces entre dominios a causa de falta de aprovisionamiento de los enlaces. Hemos demostrado como la congestión se manifiesta a través del sostenido aumento de la latencia, el cual puede

durar hasta 12 horas. Nuestra contribución al campo fue el uso de la distribución estable para el modelado de la latencia para muestras recolectadas en intervalos de 10 minutos. Hemos demostrado que la distribución estable se ajusta mejor que otras distribuciones propuestas anteriormente. También hemos presentado que a través de los parámetros obtenidos del ajuste de la distribución, y haciendo uso de métodos de aprendizaje automático, podemos correctamente detectar si un enlace se encuentra afectado por la congestión.

Para concluir, retomamos la hipótesis planteada al principio de esta Tesis, la cual enuncia lo siguiente

> *La irrupción de los proveedores de contenido a gran escala ha reconfigurado la topología de Internet, en particular alterando los principales puntos de entrega de tráfico*

A lo largo de esta Tesis hemos demostrado esta hipótesis por medio de evidencias de diversas índoles. El Capítulo 2 presentó los CPs montaron sus propias CDNs desde la cuales sirven su propio tráfico, brindando evidencia de cambios en la estructura de Internet y en las formas de servir el tráfico. En el Capítulo 3 mostramos como los IXPs en Latinoamérica, fomentados por el arribo de CPs, permitieron pequeños y medianos a ISPs acceder al contenido a bajo costo, siendo la consolidación de los IXPs en la región una prueba de los cambios en el acceso al contenido. Por último, en el Capítulo 4 recolectamos evidencias que indican que el ascenso del tráfico multimedia ha generado un estrés significativo a la red en Estado Unidos.

Los resultados presentados en cada uno de los capítulos fueron publicados en las siguientes conferencias y revistas internacionales, destacando que el material del último capítulo aún resta por ser publicado.

1. **Carisimo, E.**, Del Fiore, J.M., Dujovne, D. , Pelsser, C. and Alvarez-Hamelin, J. I. 2020. A first look at Latin American IXPs. In *ACM SIGCOMM Computer Communication Review (CCR)*. 50 (18-24).

2. **Carisimo, E.**, Selmo, C., Alvarez-Hamelin, J. I., y Dhamdhere, A. 2019. Studying the evolution of content providers in IPv4 and IPv6 internet cores. En *Computer Communications*, 145, 54-65. Elsevier.

3. **Carisimo, E.**, Selmo, C., Alvarez-Hamelin, J. I., y Dhamdhere, A. (2018, June). Studying the evolution of content providers in the internet core. En *2018 Network Traffic Measurement and Analysis Conference (TMA)* (pp. 1-8). IEEE.

4. Berenguer, S. S., **Carisimo, E.**, Alvarez-Hamelin, J. I., y Pintor, F. V. (2016, August). Hidden internet topologies info: Truth or myth?. En Proceedings of the *2016 workshop on Fostering Latin-American Research in Data Communication Networks (LANCOMM)* (pp. 4-6).

Bibliografía

[ABT01] A. Achim, A. Bezerianos y P. Tsakalides. «Wavelet-based ultrasound image denoising using an alpha-stable prior probability model». En «Proceedings 2001 International Conference on Image Processing (Cat. No.01CH37205)», volumen 2, páginas 221–224 vol.2. IEEE, Oct de 2001. ISSN null. 104

[ACF+12] Bernhard Ager, Nikolaos Chatzis, Anja Feldmann, Nadi Sarrar, Steve Uhlig y Walter Willinger. «Anatomy of a Large European IXP». En «Proceedings of the ACM SIGCOMM 2012 Conference», SIGCOMM '12, páginas 163–174. ACM, New York, NY, USA, 2012. ISBN: 9781450314190. URL http://doi.acm.org/10.1145/2342356.2342393. 26, 27, 42, 45

[ACO+06] Brice Augustin, Xavier Cuvellier, Benjamin Orgogozo, Fabien Viger, Timur Friedman, Matthieu Latapy, Clémence Magnien y Renata Teixeira. «Avoiding traceroute anomalies with paris traceroute». En «Proceedings of the 2006 Internet Measurement Conference», IMC '06, páginas 153–158. Association for Computing Machinery, New York, NY, USA, 2006. ISBN: 1595935614. URL https://doi.org/10.1145/1177080.1177100. 34, 35, 37

[AHDBV08] J. Ignacio Alvarez-Hamelin, Luca Dall'Asta, Alain Barrat y Alessandro Vespignani. «K-core decomposition of Internet graphs: hierarchies, self-similarity and measurement biases». *Networks and Heterogeneous Media*, vol. 3 nº 2, páginas 371–293, 2008. 46

[Aka05] Akamai. «Akamai To Acquire Speedera Networks». https://www.akamai.com/uk/en/about/news/press/2005-press/akamai-to-acquire-speedera-networks.jsp, 2005. 57

[AKW09] Brice Augustin, Balachander Krishnamurthy y Walter Willinger. «IXPs: mapped?». En «IMC 2009», páginas 336–349. ACM, ACM, New York, NY, USA, 2009. 97

[Ama11] Amazon. «Elastic Load Balancing Announces Support for IPv6, Zone Apex Support and Security Group Integration». https://aws.amazon.com/es/about-aws/whats-new/2011/05/24/elb-ipv6-zoneapex-securitygroups/, 2011. 61

[Ama17a] Amazon. «AWS global infrastructure». https://aws.amazon.com/es/about-aws/global-infrastructure/, 2017. 57, 65

[Ama17b] Amazon. «AWS IPv6 Update: Global Support Spanning 15 Regions & Multiple AWS Services». https://aws.amazon.com/es/blogs/aws/aws-ipv6-update-global-support-spanning-15-regions-multiple-aws-services/, 2017. 61

[App16] Apple Insider. «Apple's in-house CDN efforts spell trouble for Akamai as infrastructure biz warns of losses». http://appleinsider.com/articles/16/02/10/apples-in-house-cdn-efforts-spell-trouble-for-akamai-as-infrastructure-b, 2016. 58

[ARI15] ARIN. «ARIN IPv4 Free Pool Reaches Zero». https://www.arin.net/vault/announcements/2015/20150924.html., 2015. 11, 43, 61, 70

[ARS14] ARS Technica. «Apple's multi-terabit, $100M CDN is live - with paid connection to Comcast». https://arstechnica.com/information-technology/2014/07/apples-multi-terabit-100m-cdn-is-live-with-paid-connection-to-comcast/, 2014. 58

[Asc15] Eduardo Ascenco. «Peering in Brazil». https://ix.br/doc/nic.br.ptt.br.ais-sandiego.20150405-02.pdf, 2015. 83

[Bac00] Louis Bachelier. «Théorie de la spéculation». Gauthier-Villars, 1900. 104

[BAF+18] Timm Böttger, Gianni Antichi, Eder Leão Fernandes, Roberto di Lallo, Marc Bruyere, Steve Uhlig y Ignacio Castro. «The elusive internet flattening: 10 years of IXP growth». *CoRR*, vol. abs/1810.10963, 2018. URL http://arxiv.org/abs/1810.10963. 97

[Bak95] F. Baker (Ed.). «Requirements for IP Version 4 Routers». RFC 1812 (Proposed Standard), junio de 1995. ISSN 2070-1721. URL https://www.rfc-editor.org/rfc/rfc1812.txt. Updated by RFCs 2644, 6633. 38

[Ban16] World Bank. «World Development Indicators: Rural environment and land use». http://wdi.worldbank.org/table/3.1, 2016. 78

[Ban18] World Bank. «The World Bank data: Urban population». "https://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS?page=1", 2018. 78

[Ban19] World Bank. «GDP per capita in LAC, EU, East Asia, TH, ZA». "https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?locations=ZJ-EU-Z4-TH-ZA&year_high_desc=false", 2019. 81, 87

[BCAHP16] S. S. Berenguer, E. Carisimo, J.I. Alvarez-Hamelin y F. V. Pintor. «Hidden internet topologies info: Truth or myth?». En «LANCOMM 2016», páginas 4–6. ACM, ACM, New York, NY, USA, 2016. 96

[BCT+18] Timm Bottger, Felix Cuadrado, Gareth Tyson, Ignacio Castro y Steve Uhlig. «Open Connect Everywhere: A Glimpse at the Internet Ecosystem Through the Lens of the Netflix CDN». *ACM SIGCOMM Computer Communication Review*, vol. 48 nº 1, páginas 28–34, apr de 2018. ISSN 0146-4833. URL http://doi.acm.org/10.1145/3211852.3211857. 45, 52

[BCU18] Timm Böttger, Felix Cuadrado y Steve Uhlig. «Looking for Hypergiants in peeringDB». *ACM SIGCOMM Computer Communication Review*, vol. 48 nº 3, páginas 13–19, sep de 2018. ISSN 0146-4833. URL http://doi.acm.org/10.1145/3276799.3276801. 49, 55, 73

[BFZ07] Hitesh Ballani, Paul Francis y Xinyang Zhang. «A study of prefix hijacking and interception in the Internet». *ACM SIGCOMM Computer Communication Review*, vol. 37 nº 4, páginas 265–276, 2007. 16

[blo14] Twitch blog. «Twitch is 4th in Peak US Internet Traffic». https://blog.twitch.tv/twitch-is-4th-in-peak-us-internet-traffic-90b1295af358, 2014. 73

[Bra89] R. Braden (Ed.). «Requirements for Internet Hosts - Communication Layers». RFC 1122 (Internet Standard), octubre de 1989. ISSN 2070-1721. URL https://www.rfc-editor.org/rfc/rfc1122.txt. Updated by RFCs 1349, 4379, 5884, 6093, 6298, 6633, 6864, 8029. 37

[BSF+16] Samuel Henrique Bucke Brito, Mateus A. S. Santos, Ramon dos Reis Fontes, Danny A. Lachos Perez y Christian Esteve Rothenberg. « Dissecting the Largest National Ecosystem of Public Internet eXchange Points in Brazil». En Thomas Karagiannis y Xenofontas Dimitropoulos (Eds.), «Passive and Active Measurement», páginas 333–345. Springer International Publishing, Cham, 2016. ISBN: 9783319305059. 81, 96

[BZ11] Vladimir Batagelj y Matjaž Zaveršnik. «Fast algorithms for determining (generalized) core groups in social networks». *Advances in Data Analysis and Classification*, vol. 5 nº 2, páginas 129–145, 2011. 48

[CAIa] CAIDA. «Archipelago (Ark) Measurement Infrastructure». https://www.caida.org/projects/ark/. 38, 52, 100

[CAIb] CAIDA. «CAIDA's AS classification list». http://data.caida.org/datasets/as-classification/. 70, 73

[CAIc] CAIDA. «Mapping Autonomous Systems to Organizations: CAIDA's Inference Methodology». https://www.caida.org/research/topology/as2org/. 53

[CAId] CAIDA. «What exactly is skitter?». https://www.caida.org/projects/ark/what-is-skitter.xml. 52

[CAI18] CAIDA. «AS Rank». http://as-rank.caida.org/, 07 de 2018. 49, 82

[CAI19] CAIDA. «AS Rank APIv2». https://api.asrank.caida.org/v2/docs, 2019. 82

[CAI20] CAIDA. «Internet topology at router- and AS-levels, and the dual router and AS Internet topology generator». https://www.caida.org/research/topology/generator/, 2020. 11

[Cas14] Carlos Casasús. «Inauguración del Primer IXP Mexicano». http://www.ixp.mx/noticias/14_04_30_inaguracion.php, 2014. CITI. 96

[CB97] M. E. Crovella y A. Bestavros. «Self-similarity in world wide web traffic: evidence and possible causes». *IEEE/ACM Transactions on Networking*, vol. 5 nº 6, páginas 835–846, Dec de 1997. ISSN 1558-2566. 102

[CDF+14] P. Casas, A. D'Alconzo, P. Fiadino, A. Bar, A. Finamore y T. Zseby. «When YouTube Does not Work: Analysis of QoE-Relevant Degradation

in Google CDN Traffic». *IEEE Transactions on Network and Service Management*, vol. 11 nº 4, páginas 441–457, Dec de 2014. ISSN 1932-4537. 45

[CDFD+20] Esteban Carisimo, Julán Martín Del Fiore, Diego Dujovne, Cristel Pelsser y J. Ignacio. Alvarez-Halemin. «A first look at Latin American IXPs». *ACM SIGCOMM Computer Communication Review*, vol. 50 nº 1, páginas 1–6, 2020. 19, 45

[CDZ97] Kenneth L. Calvert, Matthew B. Doar y Ellen W. Zegura. «Modeling internet topology». *IEEE Communications magazine*, vol. 35 nº 6, páginas 160–163, 1997. 20

[CFH+13] Matt Calder, Xun Fan, Zi Hu, Ethan Katz-Bassett, John Heidemann y Ramesh Govindan. «Mapping the Expansion of Google's Serving Infrastructure». En «Proceedings of the 2013 Internet Measurement Conference», IMC '13, páginas 313–326. ACM, New York, NY, USA, 2013. ISBN: 9781450319539. URL http://doi.acm.org/10.1145/2504730.2504754. 27, 41, 45

[CFKB+15] Matt Calder, Ashley Flavel, Ethan Katz-Bassett, Ratul Mahajan y Jitendra Padhye. «Analyzing the Performance of an Anycast CDN». En «Proceedings of the 2015 Internet Measurement Conference», IMC '15, páginas 531–537. Association for Computing Machinery, New York, NY, USA, 2015. ISBN: 9781450338486. URL https://doi.org/10.1145/2815675.2815717. 29, 45

[Cha13] Chatzis, Nikolaos and Smaragdakis, Georgios and Feldmann, Anja and Willinger, Walter. «There is More to IXPs Than Meets the Eye». *ACM SIGCOMM Computer Communication Review*, vol. 43 nº 5, páginas 19–28, nov de 2013. ISSN 0146-4833. URL http://doi.acm.org/10.1145/2541468.2541473. 42, 45, 57, 83

[CHKW10] Xue Cai, John Heidemann, Balachander Krishnamurthy y Walter Willinger. «Towards an AS-to-Organization Map». En «Proceedings of the 2010 Internet Measurement Conference», IMC '10, páginas 199–205. Association for Computing Machinery, New York, NY, USA, 2010. ISBN: 9781450304832. URL https://doi.org/10.1145/1879141.1879166. 12, 25

[Chr11] Christian Kaufmann. «Akamai's V6 Rollout Plan and Experience from a CDN Point of View». *MENOG9*, vol. , 2011. 61

[CJK04] Nitesh V Chawla, Nathalie Japkowicz y Aleksander Kotcz. «Special issue on learning from imbalanced data sets». *ACM Sigkdd Explorations Newsletter*, vol. 6 nº 1, páginas 1–6, 2004. 114

[cla99] kc claffy. «Internet measurement and data analysis: topology, workload, performance and routing statistics». En «Proc. NAE'99 Workshop, Los Angels, CA, USA, March», 1999. 31

[CR05] Matthew Caesar y Jennifer Rexford. «BGP routing policies in ISP networks». *IEEE network*, vol. 19 nº 6, páginas 5–11, 2005. 16

[CSAD18] E. Carisimo, C. Selmo, J. I. Alvarez-Hamelin y A. Dhamdhere. «Studying the evolution of content providers in the internet core». En «2018 Network Traffic Measurement and Analysis Conference (TMA)», páginas 1–8. IEEE, June de 2018. ISSN null. 75

[CSAHD19] Esteban Carisimo, Carlos Selmo, J. Ignacio Alvarez-Hamelin y Amogh Dhamdhere. «Studying the evolution of content providers in IPv4 and IPv6 Internet cores». *Computer Communications*, vol. , 2019. 50, 55, 75

[CSR+15] Yi-Ching Chiu, Brandon Schlinker, Abhishek Balaji Radhakrishnan, Ethan Katz-Bassett y Ramesh Govindan. «Are We One Hop Away from a Better Internet?». En «Proceedings of the 2015 Internet Measurement Conference», IMC '15, páginas 523–529. ACM, New York, NY, USA, 2015. ISBN: 9781450338486. URL http://doi.acm.org/10.1145/2815675.2815719. 50

[CTL96] R. Chandra, P. Traina y T. Li. «BGP Communities Attribute». RFC 1997 (Proposed Standard), agosto de 1996. ISSN 2070-1721. URL https://www.rfc-editor.org/rfc/rfc1997.txt. Updated by RFC 7606. 15

[CV95] Corinna Cortes y Vladimir Vapnik. «Support-vector networks». *Machine Learning*, vol. 20 nº 3, páginas 273–297, 01 de Sep de 1995. ISSN 1573-0565. URL https://doi.org/10.1007/BF00994018. 113

[Dan16] Daniel E. Eisenbud and Cheng Yi and Carlo Contavalli and Cody Smith and Roman Kononov and Eric Mann-Hielscher and Ardas Cilingiroglu and Bin Cheyney and Wentao Shang and Jinnah Dylan Hosein. «Maglev:

A Fast and Reliable Software Network Load Balancer». En «13th USENIX Symposium on Networked Systems Design and Implementation (NSDI 16)», páginas 523–535. USENIX Association, Santa Clara, CA, 2016. ISBN: 9781931971294. URL https://www.usenix.org/conference/nsdi16/technical-sessions/presentation/eisenbud. 43, 45

[dB14] Governo do Brasil. «R$ 1.8 Billion in Telecommunications Investments for 2014 FIFA World Cup». https://bit.ly/2OA86jp, 2014. 88

[DB18] Peering DB. «TWITCH (AS46489) entry on Peering DB». https://www.peeringdb.com/net/1956, 2018. 73

[DBK+16] A. Dainotti, K. Benson, A. King, B. Huffaker, E. Glatz, X. Dimitropoulos, P. Richter, A. Finamore y A. C. Snoeren. «Lost in Space: Improving Inference of IPv4 Address Space Utilization». *IEEE Journal on Selected Areas in Communications*, vol. 34 nº 6, páginas 1862–1876, June de 2016. ISSN 1558-0008. 9, 95

[DC19] DE-CIX. «History of DECIX». https://www.de-cix.net/Files/d4167da7aafe0da34f384181606db6b8572cf2c1/DE-CIX_From-the-interconnection-of-three-ISP-to-the-worlds-leading-IX.pdf, 2019. 83

[DCGG+18] Amogh Dhamdhere, David D. Clark, Alexander Gamero-Garrido, Matthew Luckie, Ricky K. P. Mok, Gautam Akiwate, Kabir Gogia, Vaibhav Bajpai, Alex C. Snoeren y kc claffy. «Inferring Persistent Interdomain Congestion». En «Proceedings of the ACM SIGCOMM 2018 Conference», SIGCOMM '18, páginas 1–15. Association for Computing Machinery, New York, NY, USA, 2018. ISBN: 9781450355674. URL https://doi.org/10.1145/3230543.3230549. 32, 38, 100, 105, 106, 112

[DD08] Amogh Dhamdhere y Constantine Dovrolis. «Ten Years in the Evolution of the Internet Ecosystem». En «Proceedings of the 2008 Internet Measurement Conference», IMC '08, páginas 183–196. ACM, New York, NY, USA, 2008. ISBN: 9781605583341. URL http://doi.acm.org/10.1145/1452520.1452543. 21, 43, 44, 78

[DD10] Amogh Dhamdhere y Constantine Dovrolis. «The Internet is Flat: Modeling the Transition from a Transit Hierarchy to a Peering Mesh». En «Proceedings of Conference on emerging Networking EXperiments and

Technologies», CoNEXT '10, páginas 21:1–21:12. ACM, New York, NY, USA, 2010. ISBN: 9781450304481. URL http://doi.acm.org/10.1145/1921168.1921196. 21, 26, 27, 43, 45, 57, 77, 78, 97

[dGdMdA19] Jefatura de Gabinete de Ministros de Argentina. «Mapa del Estado». https://mapadelestado.jefatura.gob.ar/organismos.php, 2019. 92

[DGM06] S. N. Dorogovtsev, A. V. Goltsev y J. F. F. Mendes. «k-Core Organization of Complex Networks». *Phys. Rev. Lett.*, vol. 96, página 040601, Feb de 2006. URL https://link.aps.org/doi/10.1103/PhysRevLett.96.040601. 46

[DH98] S. Deering y R. Hinden. «Internet Protocol, Version 6 (IPv6) Specification». RFC 2460 (Draft Standard), diciembre de 1998. ISSN 2070-1721. URL https://www.rfc-editor.org/rfc/rfc2460.txt. Obsoleted by RFC 8200, updated by RFCs 5095, 5722, 5871, 6437, 6564, 6935, 6946, 7045, 7112. 10

[DH17] S. Deering y R. Hinden. «Internet Protocol, Version 6 (IPv6) Specification». RFC 8200 (Internet Standard), julio de 2017. ISSN 2070-1721. URL https://www.rfc-editor.org/rfc/rfc8200.txt. 10

[Dig] Digital Element. «NetAcuity». https://www.digitalelement.com/solutions/. 62, 82

[dIJ14] Sistema Costarricense de Informacion Juridica. «Decreto Ejecutivo 38388». "http://www.pgrweb.go.cr/scij/Busqueda/Normativa/Normas/nrm_texto_completo.aspx?param1=NRTC&nValor1=1&nValor2=77230", 2014. 83

[DKT06] S. Dharmapurikar, P. Krishnamurthy y D. E. Taylor. «Longest prefix matching using bloom filters». *IEEE/ACM Transactions on Networking*, vol. 14 nº 2, páginas 397–409, April de 2006. ISSN 1558-2566. 16

[DLH+12] Amogh Dhamdhere, Matthew Luckie, Bradley Huffaker, kc claffy, Ahmed Elmokashfi y Emile Aben. «Measuring the Deployment of IPv6: Topology, Routing and Performance». En «Proceedings of the 2012 Internet Measurement Conference», IMC '12, páginas 537–550. ACM, New York, NY, USA, 2012. ISBN: 9781450317054. URL http://doi.acm.org/10.1145/2398776.2398832. 43, 46

[DMP+02] John Dilley, Bruce Maggs, Jay Parikh, Harald Prokop, Ramesh Sitaraman y Bill Weihl. «Globally Distributed Content Delivery». *IEEE Internet Computing*, vol. 6 nº 5, páginas 50–58, sep de 2002. ISSN 1089-7801. URL https://doi.org/10.1109/MIC.2002.1036038. 45

[Dow99] Allen B. Downey. «Using pathchar to estimate Internet link characteristics». *ACM SIGCOMM Computer Communication Review*, vol. 29 nº 4, páginas 241–250, 1999. 32

[dT16] Instituto Federal de Telecomunicaciones. «Consulta Pública». https://bit.ly/2OxPdxB, 2016. 83

[DTCU17] Jie Deng, Gareth Tyson, Felix Cuadrado y Steve Uhlig. «Internet Scale User-Generated Live Video Streaming: The Twitch Case». En Mohamed Ali Kaafar, Steve Uhlig y Johanna Amann (Eds.), «Passive and Active Measurement», páginas 60–71. Springer International Publishing, Cham, 2017. ISBN: 9783319543284. 73

[dTdU19] Cámara de Telecomunicaciones de Uruguay. «El monopolio de Antel». http://www.telecomunicaciones.org.uy/index.php/el-monopolio-de-antel/, 2019. 95

[Dyn15] Dyn Blog. «IPv6: One Operating System at a Time». https://dyn.com/blog/ipv6-one-operating-system-at-a-time-2/, 2015. 61

[EF94] Kjeld Egevang y Paul Francis. «The IP Network Address Translator (NAT)». RFC 1631 (Informational), mayo de 1994. ISSN 2070-1721. URL https://www.rfc-editor.org/rfc/rfc1631.txt. Obsoleted by RFC 3022. 10

[ET12] Nicholas Economides y Joacim Tåg. «Network neutrality on the Internet: A two-sided market analysis». *Information Economics and Policy*, vol. 24 nº 2, páginas 91–104, 2012. 43

[Fac09] Facebook Engineering. «Adding :face: to every IP: Celebrating IPv6's one-year anniversary». https://www.facebook.com/notes/facebook-engineering/adding-face-to-every-ip-celebrating-ipv6s-one-year-anniversary/10151492544578920/, 2009. 62

[Fam63] Eugene F. Fama. «Mandelbrot and the stable Paretian hypothesis». *The journal of business*, vol. 36 nº 4, páginas 420–429, 1963. 104

[Fan18] Fanou, Roderick and others. «A System for Profiling the IXPs in a Region and Monitoring their Growth: Spotlight at the Internet Frontier». *IJNM*, vol. , 2018. 97

[Far07] Peyman Faratin. «Economics of overlay networks: An industrial organization perspective on network economics». En «Proceedings of the NetEcon+ IBC workshop», página 1, 2007. 43, 78

[FC16] Agustin Formoso y Pedro Casas. «Looking for network latency clusters in the lac region». En «LANCOMM 2016», páginas 10–12. ACM, ACM, New York, NY, USA, 2016. 97

[FDT+13] Tobias Flach, Nandita Dukkipati, Andreas Terzis, Barath Raghavan, Neal Cardwell, Yuchung Cheng, Ankur Jain, Shuai Hao, Ethan Katz-Bassett y Ramesh Govindan. «Reducing Web Latency: The Virtue of Gentle Aggression». En «Proceedings of the ACM SIGCOMM 2013 Conference on SIGCOMM», SIGCOMM '13, páginas 159–170. Association for Computing Machinery, New York, NY, USA, 2013. ISBN: 9781450320566. URL https://doi.org/10.1145/2486001.2486014. 41

[FFA15] Rodérick Fanou, Pierre Francois y Emile Aben. « On the Diversity of Interdomain Routing in Africa». En Jelena Mirkovic y Yong Liu (Eds.), «Passive and Active Measurement», páginas 41–54. Springer International Publishing, Cham, 2015. 45, 97

[FL06] V. Fuller y T. Li. «Classless Inter-domain Routing (CIDR): The Internet Address Assignment and Aggregation Plan». RFC 4632 (Best Current Practice), agosto de 2006. ISSN 2070-1721. URL https://www.rfc-editor.org/rfc/rfc4632.txt. 9

[FLYV93] V. Fuller, T. Li, J. Yu y K. Varadhan. «Classless Inter-Domain Routing (CIDR): an Address Assignment and Aggregation Strategy». RFC 1519 (Proposed Standard), septiembre de 1993. ISSN 2070-1721. URL https://www.rfc-editor.org/rfc/rfc1519.txt. Obsoleted by RFC 4632. 9

[FMF15] R. Fontugne, J. Mazel y K. Fukuda. «An empirical mixture model for large-scale RTT measurements». En «2015 IEEE Conference on Computer Communications (INFOCOM)», páginas 2470–2478. IEEE, April de 2015. ISSN 0743-166X. 102

[Fre18] Freedom House. «Freedom on the Net 2018: Venezuela». https://freedomhouse.org/report/freedom-net/2018/venezuela, 2018. 95

[FVD17] Rodérick Fanou, Francisco Valera y Amogh Dhamdhere. «Investigating the causes of congestion on the african ixp substrate». En «Proceedings of the 2017 Internet Measurement Conference», IMC '17, páginas 57–63. Association for Computing Machinery, New York, NY, USA, 2017. ISBN: 9781450351188. URL https://doi.org/10.1145/3131365.3131394. 97, 100, 105

[Gal13] Hernán Galperín. «Connectivity in Latin America and the Caribbean: The role of internet exchange points». *Internet Society, November*, vol. , 2013. 66

[Gal16] Hernán Galperín. «Localizing Internet infrastructure: Cooperative peering in Latin America». *Telematics and Informatics*, vol. 33 nº 2, páginas 631–640, 2016. 78, 83

[GALM08] Phillipa Gill, Martin Arlitt, Zongpeng Li y Anirban Mahanti. «The Flattening Internet Topology: Natural Evolution, Unsightly Barnacles or Contrived Collapse?». En Mark Claypool y Steve Uhlig (Eds.), «Passive and Active Network Measurement», páginas 1–10. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008. ISBN: 9783540792321. 27, 29, 42, 59

[GCF+14] Arpit Gupta, Matt Calder, Nick Feamster, Marshini Chetty, Enrico Calandro y Ethan Katz-Bassett. «Peering at the Internet's Frontier: A First Look at ISP Interconnectivity in Africa». En Michalis Faloutsos y Aleksandar Kuzmanovic (Eds.), «Passive and Active Measurement», páginas 204–213. Springer International Publishing, Cham, 2014. ISBN: 9783319049182. 45

[Gha17] Gharaibeh, Manaf and Shah, Anant and Huffaker, Bradley and Zhang, Han and Ensafi, Roya and Papadopoulos, Christos. «A Look at Router Geolocation in Public and Commercial Databases». En «Proceedings of the 2017 Internet Measurement Conference», IMC '17, páginas 463–469. ACM, New York, NY, USA, 2017. ISBN: 9781450351188. URL http://doi.acm.org/10.1145/3131365.3131380. 62, 82

[GLHc15] Vasileios Giotsas, Matthew Luckie, Bradley Huffaker y kc claffy. «IPv6 AS Relationships, Cliques, and Congruence». En Jelena Mirkovic y Yong Liu (Eds.), «Passive and Active Measurement», páginas 111–122. Springer International Publishing, Cham, 2015. ISBN: 9783319155098. 51, 90

[HBvR+13] Qi Huang, Ken Birman, Robbert van Renesse, Wyatt Lloyd, Sanjeev Kumar y Harry C. Li. «An Analysis of Facebook Photo Caching». En «Proceedings of the Twenty-Fourth ACM Symposium on Operating Systems Principles», SOSP '13, páginas 167–181. ACM, New York, NY, USA, 2013. ISBN: 9781450323888. URL http://doi.acm.org/10.1145/2517349.2522722. 28, 45

[HFU+10] Hamed Haddadi, Damien Fay, Steve Uhlig, Andrew Moore, Richard Mortier y Almerima Jamakovic. «Mixing biases: Structural changes in the AS topology evolution». En «TMA workshop», páginas 32–45. Springer, 2010. 81

[Hor] Tom Hormby. «The Rise of Google: Beating Yahoo at Its Own Game». http://lowendmac.com/2013/the-rise-of-google-beating-yahoo-at-its-own-game/. 58

[Hor15] Última Hora. «Proyecto de Senatics ayudará a abaratar acceso a internet». https://www.ultimahora.com/c864692, 2015. 83

[Hou19] Packet Clearing House. «Internet Exchange Directory». "https://www.pch.net/ixp/dir", 2019. 78

[HP06] José-Alberto Hernández y Iain W Phillips. «Weibull mixture model to characterise end-to-end Internet delay at coarse time-scales». *IEE Proceedings-Communications*, vol. 153 nº 2, páginas 295–304, 2006. 102

[Hub64] Peter J. Huber. «Robust estimation of a location parameter». *The annals of mathematical statistics*, vol. , páginas 73–101, 1964. 101

[Hus08] Geoff Huston. «Measuring IPv6 Deployment». https://meetings.ripe.net/ripe-56/presentations/Huston-Measuring_IPv6_Deployment.pdf, 2008. 61

[Hus16] Geoff Huston. «The death of Transit». https://blog.apnic.net/2016/10/28/the-death-of-transit/, 2016. 42, 44, 45

[Hus17] Geoff Huston. «The death of Transit». https://labs.apnic.net/presentations/store/2017-05-25-death-of-transit.pdf, 2017. 43

[Hus18] Geoff Huston. «What Drives IPv6 Deployment?». https://labs.ripe.net/Members/gih/what-drives-ipv6-deployment, 2018. 45, 61

[HWLR08] Cheng Huang, Angela Wang, Jin Li y Keith W. Ross. «Understanding Hybrid CDN-P2P: Why Limelight Needs Its Own Red Swoosh». En «Proceedings of the 18th International Workshop on Network and Operating Systems Support for Digital Audio and Video», NOSSDAV '08, páginas 75–80. ACM, New York, NY, USA, 2008. ISBN: 9781605581576. URL http://doi.acm.org/10.1145/1496046.1496064. 27, 41, 45

[IAN11] IANA. «The IANA IPv4 Address Free Pool is Now Depleted». https://www.arin.net/vault/announcements/2011/20110203.html, 2011. 11, 43, 45, 70

[ins18] Business insider. «Amazon's streaming service Twitch is pulling in as many viewers as CNN and MSNBC». https://www.businessinsider.com/twitch-is-bigger-than-cnn-msnbc-2018-2, 2018. 73

[Int12] Internet Society. «World IPv6 Launch». www.worldipv6launch.org/, 2012. 43, 61, 62

[Int17a] Internet Society. «Google buys a /12 IPv4 Address Block». https://www.internetsociety.org/blog/2017/05/google-buys-a-12-ipv4-address-block/, 2017. 46

[Int17b] Internet Society. «MIT Goes on IPv4 Selling Spree». https://www.internetsociety.org/blog/2017/05/mit-goes-on-ipv4-selling-spree/, 2017. 46

[Int19] InteRED. «InteRED». http://intered.org.pa/intered/, 2019. 83

[ISA19] ISA. «Composición accionaria». http://www.isa.co/es/nuestra-compania/Paginas/quienes-somos/composicion-accionaria.aspx, 2019. 92

[ITU16] ITU. «IXP Mexico». https://bit.ly/2R5PCZL, 2016. 96

[Jac89] Van Jacobson. «Traceroute». vol. , 1989. 32

[Jai13] Jain, Sushant and Kumar, Alok and Mandal, Subhasree and Ong, Joon and Poutievski, Leon and Singh, Arjun and Venkata, Subbaiah and Wanderer, Jim and Zhou, Junlan and Zhu, Min and Zolla, Jon and Hölzle, Urs and Stuart, Stephen and Vahdat, Amin. «B4: Experience with a Globally-deployed Software Defined WAN». En «Proceedings

of the ACM SIGCOMM 2013 Conference», SIGCOMM '13, páginas 3–14. ACM, New York, NY, USA, 2013. ISBN: 9781450320566. URL http://doi.acm.org/10.1145/2486001.2486019. 43, 45

[JIN19] JINX. «About INX-ZA». https://www.inx.net.za, 2019. 83

[JMdVG+17] Guillermo Julián-Moreno, Jorge E López de Vergara, Iván González, Luis de Pedro, Javier Royuela-del Val y Federico Simmross-Wattenberg. «Fast parallel \alpha α-stable distribution function evaluation and parameter estimation using OpenCL in GPGPUs». *Statistics and Computing*, vol. 27 nº 5, páginas 1365–1382, 2017. 104, 105, 106, 108

[kcB93] George C. Polyzos kc claffy y Hans-Wener Braun. «Traffic characteristics of the T1 NSFNET backbone». En «IEEE Conference on Computer Communications (INFOCOM)», volumen 2, páginas 885–893, Jan de 1993. 42

[KGAH18] Diego Kiedanski, Eduardo Grampín y J. Ignacio Alvarez-Hamelin. «The Atlas Vision of IPv6 in Latin America: Topology and Latency». En «LANC '18», páginas 40–47. ACM, New York, NY, USA, 2018. ISBN: 9781450359221. URL http://doi.acm.org/10.1145/3277103.3277122. 96

[Kum06] Sanjeev Kumar. «PING attack–How bad is it?». *Computers & Security*, vol. 25 nº 5, páginas 332–337, 2006. 37

[L.08] Julimar L. «Resultados Copa do Mundo». ftp://ftp.registro.br/pub/gter/gter41/02-IX.br-update.pdf, 2008. 88

[La 06] La Nación. «GBLX adquirió Impsat». https://www.lanacion.com.ar/economia/global-crossing-adquirio-impsat-por-us-336-millones-nid853038, 2006. 91

[LA 17] LA Times. «Apple's original TV production to begin small: 'We are just starting out'». http://beta.latimes.com/business/hollywood/la-fi-ct-apple-television-strategy-planet-apps-20170214-story.html, 2017. 58

[LBCX03] A. Lakhina, J. W. Byers, M. Crovella y P. Xie. «Sampling biases in ip topology measurements». En «IEEE INFOCOM 2003. Twenty-second Annual

Joint Conference of the IEEE Computer and Communications Societies (IEEE Cat. No.03CH37428)», volumen 1, páginas 332–341 vol.1. IEEE, March de 2003. ISSN 0743-166X. 79

[LDC+14] Matthew Luckie, Amogh Dhamdhere, David Clark, Bradley Huffaker y kc claffy. «Challenges in Inferring Internet Interdomain Congestion». En «Proceedings of the 2014 Conference on Internet Measurement Conference», IMC '14, páginas 15–22. Association for Computing Machinery, New York, NY, USA, 2014. ISBN: 9781450332132. URL https://doi.org/10.1145/2663716.2663741. 100

[LDH+16] Matthew Luckie, Amogh Dhamdhere, Bradley Huffaker, David Clark y kc claffy. «Bdrmap: Inference of borders between ip networks». En «Proceedings of the 2016 Internet Measurement Conference», IMC '16, páginas 381–396. Association for Computing Machinery, New York, NY, USA, 2016. ISBN: 9781450345262. URL https://doi.org/10.1145/2987443.2987467. 37, 38, 100

[LED17] Ioana Livadariu, Ahmed Elmokashfi y Amogh Dhamdhere. «On IPv4 transfer markets: Analyzing reported transfers and inferring transfers in the wild». *Computer Communications*, vol. 111, páginas 105–119, 2017. 46

[Lei09] Tom Leighton. «Improving performance on the internet». *Commun. ACM*, vol. 52 nº 2, páginas 44–51, feb de 2009. ISSN 0001-0782. URL http://doi.acm.org/10.1145/1461928.1461944. 42

[LHD+13] Matthew Luckie, Bradley Huffaker, Amogh Dhamdhere, Vasileios Giotsas y kc claffy. «AS Relationships, Customer Cones, and Validation». En «Proceedings of the 2013 Internet Measurement Conference», IMC '13, páginas 243–256. ACM, New York, NY, USA, 2013. ISBN: 9781450319539. URL http://doi.acm.org/10.1145/2504730.2504735. 20, 21, 25, 48, 51, 91

[LHH08] Matthew Luckie, Young Hyun y Bradley Huffaker. «Traceroute Probe Method and Forward IP Path Inference». En «Proceedings of the 2008 Internet Measurement Conference», IMC '08, páginas 311–324. Association for Computing Machinery, New York, NY, USA, 2008. ISBN: 9781605583341. URL https://doi.org/10.1145/1452520.1452557. 34

[LIJM+10] Craig Labovitz, Scott Iekel-Johnson, Danny McPherson, Jon Oberheide y Farnam Jahanian. «Internet Inter-domain Traffic». En «Proceedings of the ACM SIGCOMM 2010 Conference», SIGCOMM '10, páginas 75–86. ACM, New York, NY, USA, 2010. ISBN: 9781450302012. URL http://doi.acm.org/10.1145/1851182.1851194. 42, 53

[Liv13] Livadariu, Ioana and Elmokashfi, Ahmed and Dhamdhere, Amogh and claffy, kc. «A first look at ipv4 transfer markets». En «Proceedings of Conference on emerging Networking EXperiments and Technologies», CoNEXT '13, páginas 7–12. ACM, New York, NY, USA, 2013. ISBN: 9781450321013. URL http://doi.acm.org/10.1145/2535372.2535416. 46

[LJL+11] DK Lee, Keon Jang, Changhyun Lee, Gianluca Iannaccone y Sue Moon. «Scalable and systematic Internet-wide path and delay estimation from existing measurements». *Computer Networks*, vol. 55 nº 3, páginas 838–855, 2011. 16

[Lor09] Lorenzo Colitti. «IPv6 at Google». https://www.ripe.net/participate/meetings/roundtable/february-2009/LorenzoIPv6atGoogle.pdf, 2009. 61

[LR01] Lixin Gao y J. Rexford. «Stable Internet routing without global coordination». *IEEE/ACM Transactions on Networking*, vol. 9 nº 6, páginas 681–692, Dec de 2001. ISSN 1063-6692. 18, 20, 21, 81

[Luc10] Matthew Luckie. «Scamper: A Scalable and Extensible Packet Prober for Active Measurement of the Internet». En «Proceedings of the 2010 Internet Measurement Conference», IMC '10, páginas 239–245. Association for Computing Machinery, New York, NY, USA, 2010. ISBN: 9781450304832. URL https://doi.org/10.1145/1879141.1879171. 31, 38

[Mac19] Wayback Machine. «Wayback Machine». http://waybackmachine.org, 2019. 64

[Man63] Benoit Mandelbrot. «The Variation of Certain Speculative Prices». *The Journal of Business*, vol. 36 nº 4, páginas 394–419, 1963. ISSN 00219398, 15375374. URL http://www.jstor.org/stable/2350970. 104

[Mar08] Mariano G. Beiró and J. Ignacio Alvarez-Hamelin and Jorge R. Busch. «A

low complexity visualization tool that helps to perform complex systems analysis». *New J. Phys*, vol. 10 nº 12, página 125003, 2008. 48, 51

[Mar13] Mario Durán Chuquimia. «Seminario sobre el PIT». http://desarrollotics.blogspot.com/2013/07/preparando-un-seminario-tecnico-sobre.html, 2013. 80

[Moc87] P.V. Mockapetris. «Domain names - implementation and specification». RFC 1035 (Internet Standard), noviembre de 1987. ISSN 2070-1721. URL https://www.rfc-editor.org/rfc/rfc1035.txt. Updated by RFCs 1101, 1183, 1348, 1876, 1982, 1995, 1996, 2065, 2136, 2181, 2137, 2308, 2535, 2673, 2845, 3425, 3658, 4033, 4034, 4035, 4343, 5936, 5966, 6604, 7766. 27

[Moy98] J. Moy. «OSPF Version 2». RFC 2328 (Internet Standard), abril de 1998. ISSN 2070-1721. URL https://www.rfc-editor.org/rfc/rfc2328.txt. Updated by RFCs 5709, 6549, 6845, 6860, 7474, 8042. 11, 35

[MP85] J.C. Mogul y J. Postel. «Internet Standard Subnetting Procedure». RFC 950 (Internet Standard), agosto de 1985. ISSN 2070-1721. URL https://www.rfc-editor.org/rfc/rfc950.txt. Updated by RFC 6918. 9

[MS95] Rosario N Mantegna y H Eugene Stanley. «Scaling behaviour in the dynamics of an economic index». *Nature*, vol. 376 nº 6535, página 46, 1995. 104, 108

[Muu83] Mike Muuss. «The story of the ping program». *Retrieved March*, vol. 30, página 2010, 1983. 32

[Nac15] La Nacion. «ICE rechaza unirse a sistema para agilizar Internet a usuarios». https://www.nacion.com/el-pais/servicios/ice-rechaza-unirse-a-sistema-para-agilizar-internet-a-usuarios/VPE4YCNSWZEYDOTIJJXRFFGIZM/story/, 2015. 96

[Nat16] United Nations. «The World's Cities in 2016». "http://www.un.org/en/development/desa/population/publications/pdf/urbanization/the_worlds_cities_in_2016_data_booklet.pdf", 2016. 78

[Nat17] United Nations. «World Population Prospects 2017». "https://population.un.org/wpp/Download/Standard/Population/", 2017. 78

[NCC19a] RIPE NCC. «RIPE Atlas: Percentage of connected probes per country». "https://atlas.ripe.net/results/maps/density/", 2019. 38

[NCC19b] RIPE NCC. «RIPE Routing Information Service (RIS)». https://www.ripe.net/analyse/internet-measurements/routing-information-service-ris, 2019. 30, 78

[Net12a] Netflix Media Center. «"Announcing the Netflix Open Connect Network"». https://media.netflix.com/en/company-blog/announcing-the-netflix-open-connect-network, 2012. 59

[Net12b] Netflix Tech Blog. «Enabling Support for IPv6». https://medium.com/netflix-techblog/enabling-support-for-ipv6-48a495d5196f, 2012. 62

[Net16] Netflix Tech Blog. «Building fast.com». https://medium.com/netflix-techblog/building-fast-com-4857fe0f8adb, 2016. 62

[New] New York Times. «Google to Acquire YouTube for $1.65 Billion». http://www.nytimes.com/2006/10/09/business/09cnd-deal.html. 58

[Nic16] Nicolás Evers. «Consultoria - IXP Paraguay». https://bit.ly/2Lay76N, 2016. IV Foro Regional sobre Interconectividad 11 y 12 de agosto Tegucigalpa – Honduras SesioÌn 6. 83

[OGLK14] Chiara Orsini, Enrico Gregori, Luciano Lenzini y Dmitri Krioukov. «Evolution of the Internet K-dense Structure». *IEEE/ACM Transactions on Networking*, vol. 22 nº 6, páginas 1769–1780, dec de 2014. ISSN 1063-6692. URL http://dx.doi.org/10.1109/TNET.2013.2282756. 46

[OKG+16] Chiara Orsini, Alistair King, Danilo Giordano, Vasileios Giotsas y Alberto Dainotti. «BGPStream: A Software Framework for Live and Historical BGP Data Analysis». En «Proceedings of the 2016 Internet Measurement Conference», IMC '16, páginas 429–444. Association for Computing Machinery, New York, NY, USA, 2016. ISBN: 9781450345262. URL https://doi.org/10.1145/2987443.2987482. IX, 15

[oO19] University of Oregon. «RouteViews». http://www.routeviews.org/, 2019. IX, 15, 30, 78

[OPW+10] Ricardo Oliveira, Dan Pei, Walter Willinger, Beichuan Zhang y Lixia Zhang. «The (in)Completeness of the Observed Internet AS-level Structure». *IEEE/ACM Transactions on Networking*, vol. 18 nº 1, páginas 109–122, feb de 2010. ISSN 1063-6692. URL http://dx.doi.org/10.1109/TNET.2009.2020798. 50

[PBV08] Mukaddim Pathan, Rajkumar Buyya y Athena Vakali. «Content delivery networks: State of the art, insights, and imperatives». *Content Delivery Networks*, vol. , páginas 3–32, 2008. 45

[Pee] PeeringDB. «https://www.peeringdb.com». 82

[PMF+03] K. Papagiannaki, S. Moon, C. Fraleigh, P. Thiran y C. Diot. «Measurement and analysis of single-hop delay on an IP backbone network». *IEEE Journal on Selected Areas in Communications*, vol. 21 nº 6, páginas 908–921, Aug de 2003. ISSN 1558-0008. 102

[PMH+93] C-K Peng, J Mietus, JM Hausdorff, Shlomo Havlin, H Eugene Stanley y Ary L Goldberger. «Long-range anticorrelations and non-Gaussian behavior of the heartbeat». *Physical review letters*, vol. 70 nº 9, página 1343, 1993. 104

[Pos81a] J. Postel. «Internet Control Message Protocol». RFC 792 (Internet Standard), septiembre de 1981. ISSN 2070-1721. URL https://www.rfc-editor.org/rfc/rfc792.txt. Updated by RFCs 950, 4884, 6633, 6918. 32

[Pos81b] J. Postel. «Internet Protocol». RFC 791 (Internet Standard), septiembre de 1981. ISSN 2070-1721. URL https://www.rfc-editor.org/rfc/rfc791.txt. Updated by RFCs 1349, 2474, 6864. 9, 37

[PSVV01] Romualdo Pastor-Satorras, Alexei Vázquez y Alessandro Vespignani. «Dynamical and Correlation Properties of the Internet». *Phys. Rev. Lett.*, vol. 87, página 258701, Nov de 2001. URL https://link.aps.org/doi/10.1103/PhysRevLett.87.258701. 48

[PUK+11] Ingmar Poese, Steve Uhlig, Mohamed Ali Kaafar, Benoit Donnet y Bamba Gueye. «IP Geolocation Databases: Unreliable?». *ACM SIGCOMM Computer Communication Review*, vol. 41 nº 2, páginas 53–56, apr de 2011. ISSN 0146-4833. URL http://doi.acm.org/10.1145/1971162.1971171. 62, 82, 95

[PV06] George Pallis y Athena Vakali. «Insight and Perspectives for Content Delivery Networks». *Commun. ACM*, vol. 49 nº 1, páginas 101–106, jan de 2006. ISSN 0001-0782. URL http://doi.acm.org/10.1145/1107458.1107462. 45

[Qua14] Quartz Media. «"The inside story of how Netflix came to pay Comcast for internet traffic"». https://qz.com/256586/the-inside-story-of-how-netflix-came-to-pay-comcast-for-internet-traffic 2014. 59

[Qui16] James Quinn. «Being Open: How Facebook Got It's Edge». *NANOG68*, vol. , 2016. 65

[Rho93] S. A. Rhoades. «The Herfindahl-Hirschman Index», 1993. 95

[Ric19] MICITT (Costa Rica). «IXP Costa Rica: Una oportunidad estrategica». https://micit.go.cr/index.php?option=com_content&view=article&id=6329, 2019. 83

[RLH06] Y. Rekhter (Ed.), T. Li (Ed.) y S. Hares (Ed.). «A Border Gateway Protocol 4 (BGP-4)». RFC 4271 (Draft Standard), enero de 2006. ISSN 2070-1721. URL https://www.rfc-editor.org/rfc/rfc4271.txt. Updated by RFCs 6286, 6608, 6793, 7606, 7607, 7705, 8212. 11

[RMK+96] Y. Rekhter, B. Moskowitz, D. Karrenberg, G. J. de Groot y E. Lear. «Address Allocation for Private Internets». RFC 1918 (Best Current Practice), febrero de 1996. ISSN 2070-1721. URL https://www.rfc-editor.org/rfc/rfc1918.txt. Updated by RFC 6761. 10

[RSF+14] Philipp Richter, Georgios Smaragdakis, Anja Feldmann, Nikolaos Chatzis, Jan Boettger y Walter Willinger. «Peering at Peerings: On the Role of IXP Route Servers». En «Proceedings of the 2014 Internet Measurement Conference», IMC '14, páginas 31–44. Association for Computing Machinery, New York, NY, USA, 2014. ISBN: 9781450332132. URL https://doi.org/10.1145/2663716.2663757. 19

[RUB15] R. Ravaioli, G. Urvoy-Keller y C. Barakat. «Characterizing ICMP rate limitation on routers». En «2015 IEEE International Conference on Communications (ICC)», páginas 6043–6049. IEEE, June de 2015. ISSN 1938-1883. 37

[RWVR+16] Philipp Richter, Florian Wohlfart, Narseo Vallina-Rodriguez, Mark Allman, Randy Bush, Anja Feldmann, Christian Kreibich, Nicholas Weaver y Vern Paxson. «A Multi-perspective Analysis of Carrier-Grade NAT Deployment». En «Proceedings of the 2016 Internet Measurement Conference», IMC '16, páginas 215–229. ACM, New York, NY, USA, 2016. ISBN: 9781450345262. URL http://doi.acm.org/10.1145/2987443.2987474. 95

[San11] Sandvine. «Global internet phenomena report Spring 2011», 2011. 57

[See15] Seeking Alpha. «Apple, Microsoft And Facebook Bring More Traffic To In-House CDNs, Impacting Akamai's Media Business». https://seekingalpha.com/article/3613736-apple-microsoft-facebook-bring-traffic-house-cdns-impacting-akama 2015. 59

[SGGR+13] Diego Salas-González, JM Górriz, Javier Ramírez, M Schloegl, Elmar Wolfgang Lang y Andrés Ortiz. «Parameterization of the distribution of white and grey matter in MRI using the α-stable distribution». *Computers in biology and medicine*, vol. 43 nº 5, páginas 559–567, 2013. 104

[SGKR09] Diego Salas-Gonzalez, Ercan E Kuruoglu y Diego P Ruiz. «Modelling and assessing differential gene expression using the alpha stable distribution». *The International Journal of Biostatistics*, vol. 5 nº 1, 2009. 106

[SHS+12] Justine Sherry, Shaddi Hasan, Colin Scott, Arvind Krishnamurthy, Sylvia Ratnasamy y Vyas Sekar. «Making middleboxes someone else's problem: network processing as a cloud service». *ACM SIGCOMM Computer Communication Review*, vol. 42 nº 4, páginas 13–24, 2012. 10

[SKC+17] Brandon Schlinker, Hyojeong Kim, Timothy Cui, Ethan Katz-Bassett, Harsha V. Madhyastha, Italo Cunha, James Quinn, Saif Hasan, Petr Lapukhov y Hongyi Zeng. «Engineering Egress with Edge Fabric: Steering Oceans of Content to the World». En «Proceedings of the ACM SIGCOMM 2017 Conference», SIGCOMM '17, páginas 418–431. ACM, New York, NY, USA, 2017. ISBN: 9781450346535. URL http://doi.acm.org/10.1145/3098822.3098853. 43, 45

[SLB+18] Florian Streibelt, Franziska Lichtblau, Robert Beverly, Anja Feldmann, Cristel Pelsser, Georgios Smaragdakis y Randy Bush. «BGP Communities:

Even More Worms in the Routing Can». En «Proceedings of the Internet Measurement Conference 2018», IMC '18, páginas 279–292. Association for Computing Machinery, New York, NY, USA, 2018. ISBN: 9781450356190. URL https://doi.org/10.1145/3278532.3278557. 15

[SSLB17] Volker Stocker, Georgios Smaragdakis, William Lehr y Steven Bauer. «The growing complexity of content delivery networks: Challenges and implications for the Internet ecosystem». *Telecommunications Policy*, vol. 41 nº 10, páginas 1003–1016, 2017. 57

["St16] "Stefan Meinders". «The New Internet». ENOG11, 2016. 53

[Ste11] Richard A. Steenbergen. «A Guide to Peering on the Internet». En «NANOG 51». NANOG, 2011. 22

[STWZ16] Yu-Wei Eric Sung, Xiaozheng Tie, Starsky H.Y. Wong y Hongyi Zeng. «Robotron: Top-down Network Management at Facebook Scale». En «Proceedings of the 2016 ACM SIGCOMM Conference», SIGCOMM '16, páginas 426–439. ACM, New York, NY, USA, 2016. ISBN: 9781450341936. URL http://doi.acm.org/10.1145/2934872.2934874. 45

[Su,09] Su, Ao-Jan and Choffnes, David R. and Kuzmanovic, Aleksandar and Bustamante, Fabián E. «Drafting Behind Akamai: Inferring Network Conditions Based on CDN Redirections». *IEEE/ACM Transactions on Networking*, vol. 17 nº 6, páginas 1752–1765, dec de 2009. ISSN 1063-6692. URL http://dx.doi.org/10.1109/TNET.2009.2022157. 43, 52

[Sub99] SubTel. «Resolucion 1483». https://www.subtel.gob.cl/images/stories/articles/subtel/asocfile/res_1483conexiones_entre_isp.pdf, 1999. 83, 84

[TEC16] GESTION DE TECNOLOGÍA. «Lanzan IXP-HN». https://blogs.unah.edu.hn/degt/lanzamiento-del-punto-de-intercambio-de-trafico-de-internet-de-honduras-
2016. 83

[Tel16] TeleGeography. «Belize gets internet exchange point; BIXP becomes twelfth such facility in Caribbean». https://bit.ly/2sxgisc, 2016. 83

[TSGR04] Renata Teixeira, Aman Shaikh, Tim Griffin y Jennifer Rexford. «Dynamics of Hot-Potato Routing in IP Networks». En «Proceedings of the Joint In-

ternational Conference on Measurement and Modeling of Computer Systems», SIGMETRICS '04/Performance '04, páginas 307–319. Association for Computing Machinery, New York, NY, USA, 2004. ISBN: 1581138733. URL https://doi.org/10.1145/1005686.1005723. 18

[TW87] Martin A Tanner y Wing Hung Wong. «The calculation of posterior distributions by data augmentation». *Journal of the American statistical Association*, vol. 82 nº 398, páginas 528–540, 1987. 114

[UZ99] Vladimir V. Uchaikin y Vladimir M. Zolotarev. «Chance and stability: stable distributions and their applications». Walter de Gruyter, 1999. 104

[WIR12] WIRED. «Google and Netflix Make Land Grab On Edge Of Internet». https://www.wired.com/2012/06/cdn/, 2012. 43

[Wir16] Wired. «Google and Netflix Make Land Grab On Edge Of Internet». https://www.wired.com/2012/06/cdn/, 2016. 43

[Yah12] Yahoo Finance. «Number of active users at Facebook over the years». https://finance.yahoo.com/news/number-active-users-facebook-over-years-214600186--finance.html, 2012. 58

[YHkc03] Andre Broido Young Hyun y kc claffy. «Traceroute and BGP AS Path Incongruities». Informe técnico, Cooperative Association for Internet Data Analysis (CAIDA), Mar de 2003. 52

Índice alfabético

k-núcleos, 41
stub ASes, 74

ADSL, 21
algoritmos de ruteo, 5
Anycast, 24, 39
API, 74
aprendizaje automático, 93
Archipiélago, 34
AS origen, 10
AS sets, 73
AS-PATH prepend, 73
AS-PATH prepending, 14
AS-PATHS, 41
AS-RANK, 42
ASN, 8, 47
ataques de denegación de servicio distribuido, 33
Autonomous System, 8
Autonomous System Number, 8
Azure, 39

balance de carga, 30
BGP, 7, 41
Border Gateway Protocol, 7
bring home, 25
broadcast, 7

c2p, 17
CAIDA, 34, 42
camino lento, 33
CDNs, 23, 35, 36
CIDR, 5
Classful Routing, 5
Classless Inter-Domain Routing, 5
cliente-proveedor, 17
cliqué de tránsito libre, 21
conmutador, 22
Cono de Clientes, 19
Content Delivery Networks, 23, 35
Content Provider, 36
Customer Cone, 19
customer-to-provider relationship, 17

data augmentation, 105
DDoS, 33
direcciones de terceros, 34
Distributed Denial of Service, 33
DOCSIS, 21
dual-stack, 7

ECMP, 32
end-hosts, 3
enter deep, 25
Equal Cost Multipath, 32
eyeballs, 36

forwarding, 4
front-ends, 36
FTTH, 21

gamer, 65
Google, 38
grado de nodo, 40

hosts, 3

Hot Potato Routing, 14

IANA, 5
ICMP, 27
IETF, 5, 32
Internet Control Message Protocol, 27
IP, 3
IPv4, 5
IPv6, 6
ISPs, 7
IXP, 15, 22, 69

Latinoamérica, 70
LG, 71
Longest Prefix Match, 12
Looking Glasses, 71
LPM, 12

Machine Learning, 93
Mandatory Multilateral Peering Policy, 76
Microsoft, 39
ML, 93
MMPP, 76
MOAS, 12
Multiple Origin ASes, 12

Número de Sistema Autónomo, 8
Números de Sistema Autónomo, 47
NAT, 6
Network Address Translation, 6
NSFNET, 36

outliers, 93

p2p, 17
Packet Clearing House, 71
PCH, 71
peer-to-peer relationship, 17
peering, 17
ping, 28
PITs, 15, 22, 69
Política de Anuncios Multilateral Obligatoria, 76
protocolo de Internet, 3
proveedores de servicios de Internet, 7
Puntos de Intercambio de Tráfico, 15, 22, 69

QoE, 39, 92
Quality of Experience, 92

Redes de Distribución de Contenido, 23, 35
RIB, 12
RIPE Atlas, 34
RIRs, 39
Round Trip Time, 93
Route Server, 15
route server, 73
Routeviews, 11
routing, 4
Routing Information Base, 12
RTT, 93
ruteadores, 3
ruteo, 39

servidor de rutas, 73
shell-index, 41
Sistema Autónomo, 8
slow path, 33
Software Defined Networking, 39
split TCP, 35
Support Vector Machine, 105
SVM, 105
switch, 22

TCP/IP, 3
third-party addresses, 34
tiempo de ida y vuelta, 93
TIER-1, 21, 37

TOPcore, 44
Traceroute, 28
transit degree, 40
transit-free clique, 21
Twitch, 65

unicast, 12

vínculos entre pares, 17
valley-free, 18
Vantage Point, 97

WAN, 39