TECHNISCHE UNIVERSITÄT BERLIN
FAKULTÄT FÜR ELEKTROTECHNIK UND INFORMATIK
LEHRSTUHL FÜR INTELLIGENTE NETZE
UND MANAGEMENT VERTEILTER SYSTEME

Revisiting the Interplay of Inter-Domain Traffic and Routing Policies

vorgelegt von
M.Sc. Thomas Jakob Krenc

von der Fakultät IV – Elektrotechnik und Informatik
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

DOKTOR DER NATURWISSENSCHAFTEN
–DR. RER. NAT.–

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender:	Prof. Georgios Smaragdakis, Ph. D., Technische Universität Berlin
Gutachterin:	Prof. Anja Feldmann, Ph. D., Max-Planck-Institut, Saarbrücken
Gutachter:	Prof. Thomas Zinner, Ph. D., Technische Universität Berlin
Gutachter:	Prof. Robert Beverly, Ph. D., Naval Postgraduate School
Gutachter:	Oliver Hohlfeld, Ph. D., RWTH Aachen

Tag der wissenschaftlichen Aussprache: 18. Januar 2019

Berlin 2019

To my family and friends.

Abstract

The Internet started out – with the inception of ARPANET in the 1960s, followed by NSFNET in 1986 – as a government-funded academic research network interconnecting universities and research facilities. Just a few years later, the World Wide Web was invented, and the first commercial service providers emerged. During the transition phase from a research to a commercial-purpose network, most of the crucial changes to protocols, infrastructure, and governance have occurred and continue to shape the Internet today.

In its current state, the Internet is a collection of tens of thousands interconnected networks which exchange traffic among each other. These heterogeneous networks, varying in size and type, are owned and operated by organizations with individual interests and goals. While some networks provide connectivity to consumers and companies, other networks focus on the distribution and delivery of content. Depending on their business model, traffic composition and volumes exchanged among these networks can vary significantly.

Network operators need to understand the composition of traffic in order to meet the quality expectations of their customers. However, Internet traffic is more diverse than ever. Many different applications, including video streaming, gaming, or file-sharing, dominate the dynamic composition of traffic, while live events such as the World Cup cause major temporal variations in traffic volume. Moreover, an increasing trend towards traffic encryption makes it increasingly difficult for network operators to obtain a holistic picture of the traffic landscape. Further, the manner in which traffic is exchanged between the networks makes it increasingly hard to reason about trends in the Internet. The exchange of traffic is governed by complex business relations among ISPs, or by traffic steering policies performed by CDNs and cloud providers who take advantage of network and path diversity by peering at IXPs.

Inter-domain routing and the associated traffic flow is steadily evolving. In order to keep track of developments in the Internet, it is vital to steadily revisit changes to infrastructure and policies and how they affect traffic flow. In this dissertation, we investigate the traffic flow and the underlying routing mechanisms. Specifically, we seek to gain a better understanding of heterogeneity and traffic asymmetries on inter-domain links and the global routing table growth. We dissect the composition of today's Internet traffic and the interactions of the involved parties and highlight the economic incentives that drive many of the main commercial players to deploy their servers deep within third-party networks. Further, we illuminate how hypergiants and complex business relationships impact the balance and distribution of ingress and egress in inter-domain traffic. Finally, we study the ramifications of these complexities on the global routing table growth by investigating one of the contributors to growth, namely prefix delegations, and how different parties use prefix delegations to influence path selection.

Through the analysis of the Internet from multiple points of view, we observe an increasing trend towards network heterogeneity and unconventional routing, which is increasingly diverging from our notion of a hierarchical Internet. Our findings contribute to advancing a new mental model for the Internet's ecosystem that goes beyond traffic agnostic AS-graph models and can support network operators in network planning and provisioning. Last, insights in the complexities of prefix delegations shed light on current protocol limitations and can inform protocol designers in the future.

Zusammenfassung

Das Internet begann – mit der Gründung von ARPANET in den 1960er Jahren, gefolgt von NSFNET im Jahr 1986 – als staatlich finanziertes akademisches Forschungsnetz, das Universitäten und Forschungseinrichtungen miteinander verband. Wenige Jahre später wurde das World Wide Web erfunden und die ersten kommerziellen Provider entstanden. Während der Übergangsphase von einem Forschungs- zu einem kommerziellen Netzwerk fanden die meisten wesentlichen Änderungen an Protokollen, Infrastruktur und *Governance* statt und prägen das Internet bis heute.

Das heutige Internet ist eine Ansammlung Zehntausender, miteinander verbundener Netzwerke, die untereinander Verkehr austauschen. Diese heterogenen Netzwerke unterscheiden sich in Größe und Typ, und werden von Organisationen mit individuellen Interessen und Zielen betrieben. Während einige Netzwerke Konnektivität für Verbraucher und Unternehmen bereitstellen, konzentrieren sich andere auf die Verteilung und Bereitstellung von Inhalten. Abhängig vom Geschäftsmodell, kann die Zusammensetzung und das Volumen des ausgetauschten Verkehrs erheblich variieren.

Netzbetreiber müssen die Zusammensetzung des Verkehrs verstehen, um die Qualitätserwartungen ihrer Kunden zu erfüllen. Der Internetverkehr ist jedoch vielfältiger als je zuvor: Verschiedene Anwendungen, z. B. Video-Streaming, Spiele oder Filesharing, dominieren die dynamische Zusammensetzung, während Live-Ereignisse, wie die Fußballweltmeisterschaft, große temporäre Volumenschwankungen verursachen. Der Trend zur Datenverschlüsselung und die Art des Verkehrsaustauschs zwischen den Netzen, machen es zunehmend schwerer für Betreiber ein Bild der Verkehrslandschaft zu erstellen und Trends zu verstehen. Dabei wird der Verkehrsaustausch von komplexen Geschäftsbeziehungen zwischen Providern oder von Verkehrssteuerung durch CDNs und Cloud-Anbietern, die die Netzwerk- und Pfadvielfalt in IXPs ausnutzen, beeinflusst.

Inter-Domain Routing, und der damit verbundene Verkehr, entwickeln sich kontinuierlich. Um Entwicklungen im Internet zu verfolgen, ist es wichtig, Änderungen an Infrastruktur und Richtlinien, und deren Auswirkungen auf den Verkehr ständig zu überprüfen. In dieser Arbeit untersuchen wir den Verkehrsfluss und die zugrundeliegenden Routing-Mechanismen. Insbesondere möchten wir ein besseres Verständnis der Heterogenität und der Verkehrsasymmetrien auf Inter-domain-Links, und des Wachstums der globalen Routing-Tabelle erlangen. Wir analysieren die Zusammensetzung des heutigen Verkehrs und die Interaktionen der beteiligten Akteure, und heben wirtschaftliche Anreize vieler kommerziell relevanter Akteure zum Einsatz von Servern in Dritt-Netzwerken hervor. Zudem beleuchten wir Auswirkungen von Hypergiants und komplexer Geschäftsbeziehungen auf die Balance und Verteilung von Ingress und Egress im Inter-Domain-Verkehr. Zum Schluss untersuchen wir die Auswirkungen dieser Komplexitäten auf das Wachstum der globalen Routing-Tabelle, indem wir einen der Wachstumsträger untersuchen, nämlich Präfixdelegationen, und wie sie von verschiedenen Akteuren verwendet werden, um die Pfadauswahl zu beeinflussen.

Durch die Analyse des Internets aus mehreren Blickwinkeln beobachten wir einen ansteigenden Trend hin zu Netzheterogenität und unkonventionellem Routing, der zunehmend von unserer Vorstellung eines hierarchischen Internets abweicht. Unsere Resultate tragen zu einem neuen Denkmodell für das Internetökosystem bei, das über verkehrs-unabhängige AS-Graph-Modelle hinausgeht, und Netzbetreiber bei der Netzwerkplanung und -bereitstellung unterstützen kann. Schließlich geben Einblicke in die Komplexität von Präfixdelegationen Aufschluss über aktuelle Protokolleinschränkungen und können Protokolldesigner in der Zukunft informieren.

List of Publications

Parts of this thesis are based on the following set of papers. These papers have been co-authored with other researchers. All my collaborators are among my co-authors and are acknowledged here. I thank them all for their valuable contribution.

Pre-published Papers

International Conferences

BGP Prefix Delegations: A Deep Dive.
THOMAS KRENC AND ANJA FELDMANN.
Proceedings of the ACM Internet Measurement Conference (IMC), 2016.

On the Benefits of Using a Large IXP As an Internet Vantage Point.
NIKOLAOS CHATZIS, GEORGIOS SMARAGDAKIS, JAN BÖTTGER,
THOMAS KRENC AND ANJA FELDMANN.
Proceedings of the ACM Internet Measurement Conference (IMC), 2013.

International Journals

An Internet Census Taken by an Illegal Botnet:
A Qualitative Assessment of Published Measurements.
THOMAS KRENC, OLIVER HOHLFELD AND ANJA FELDMANN.
ACM SIGCOMM Computer Communication Review Volume, 44(3), 2014

Under submission

International Conferences

Parts of this thesis are based on the following paper that is currently under submission.

On Traffic Volume Asymmetry.
THOMAS KRENC, BALAKRISHNAN CHANDRASEKARAN, ANJA FELDMANN,
OLIVER HOHLFELD, INGMAR POESE AND ENRIC PUJOL.

Contents

1 Introduction 1
- 1.1 Research Question . . . 3
- 1.2 Vantage Points and Datasets . . . 4
- 1.3 Contributions . . . 4
- 1.4 Overview & Roadmap . . . 5

2 Background 9
- 2.1 Design Principles . . . 9
- 2.2 Major Internet Players . . . 11
- 2.3 The Evolution of the Network Layer . . . 13
  - 2.3.1 Addressing . . . 13
  - 2.3.2 Address Structure . . . 14
  - 2.3.3 Allocation of Address Space . . . 15
  - 2.3.4 Routing vs. Forwarding . . . 16
  - 2.3.5 Routing Protocols . . . 17
  - 2.3.6 Hostnames . . . 20
- 2.4 Growing Pains . . . 21
  - 2.4.1 Registration of Internet Identifiers . . . 22
  - 2.4.2 IPv4 Addresss space exhaustion . . . 23
  - 2.4.3 Global Routing table growth . . . 24
- 2.5 Historical Overview . . . 26

3 IXP as a Vantage Point 27
- 3.1 Tracking Developements at IXPs . . . 27
- 3.2 IXP as a Rich Data Source . . . 29
  - 3.2.1 Available IXP-internal datasets . . . 29
  - 3.2.2 Methods for dissecting the IXP's traffic . . . 29
  - 3.2.3 Available IXP-external datasets . . . 32
  - 3.2.4 IP server meta-data . . . 33
- 3.3 Local yet Global . . . 33
  - 3.3.1 On the global role of the IXP . . . 34
  - 3.3.2 On the IXP's dual role . . . 36
  - 3.3.3 On the IXP's "blind spots" . . . 37
- 3.4 Stable yet Changing . . . 39
  - 3.4.1 Stability in the face of constant growth . . . 39
  - 3.4.2 Changes in face of significant stability . . . 43
- 3.5 Beyond the AS-level view . . . 44
  - 3.5.1 Alternative grouping of server IPs . . . 44
  - 3.5.2 New reality (I): ASes are heterogeneous . . . 46
  - 3.5.3 New reality (II): Links are heterogeneous . . . 47

3.6 Discussion and Caveats . . . 50
3.7 Chapter Summary . . . 52

4 Traffic Asymmetries on Inter-Domain Links 53
4.1 Traffic Volume Asymmetries . . . 53
4.2 A Peek at Traffic Asymmetry . . . 55
4.3 Dataset: Perspective of a Tier-1 ISP . . . 56
4.4 On Near and Far Neighbors . . . 57
4.5 The Interplay between Routing & Traffic Asymmetry . . . 59
4.6 The Role of Hypergiants . . . 62
4.6.1 On the Accuracy of Traffic Ratios in PeeringDB . . . 62
4.6.2 Hypergiants & Asymmetry . . . 63
4.7 Limitations . . . 64
4.8 Chapter Summary . . . 66

5 Prefix Delegations via BGP 67
5.1 Understanding BGP Prefix Delegations . . . 67
5.2 Background & related work . . . 68
5.3 Data sources . . . 69
5.4 Prefix delegations . . . 70
5.5 Delegations across 10 years . . . 72
5.6 AS business relationships . . . 73
5.7 Effects on path selection . . . 75
5.7.1 PA Prefixes from Provider to Customer . . . 76
5.7.2 Delegations from Customer to Provider . . . 76
5.7.3 Delegations among Non-Adjacent ASes . . . 77
5.8 Chapter Summary . . . 79

6 Internet-Wide Scans by a Botnet 81
6.1 Introduction . . . 81
6.2 Published Datasets . . . 83
6.3 Authenticity . . . 84
6.3.1 Reverse DNS . . . 85
6.3.2 Akamai IPs . . . 85
6.4 Looking Behind the Curtain . . . 86
6.4.1 Meta-data? Wrong! . . . 86
6.4.2 Data Quality . . . 87
6.5 Claims of the authors . . . 91
6.5.1 Finding Censuses . . . 91
6.5.2 Where are the Fast Scans? . . . 92
6.6 Discussion . . . 93
6.6.1 Robustness of the Data . . . 93
6.6.2 What's the News? . . . 94
6.6.3 Ethical Considerations . . . 94
6.7 Chapter Summary . . . 96

7 Conclusion 97
7.1 Summary . . . 97
7.2 Future Work . . . 98

1 Introduction

The Internet has become an essential infrastructure in many areas, e.g., e-commerce, education, entertainment. Bandwidth-heavy applications such as high-definition videos, software updates, as well as delay-sensitive real-time applications like gaming, video and audio communication are increasing in number and popularity. Moreover, services like cloud gaming, where video games are rendered in data centers, and the resulting video feed is sent to the client in a timely manner, are on the rise.

In order to accommodate the ever-growing demand for more bandwidth and less delay, the Internet undergoes constant structural changes. Figure 1.1 shows a simplified depiction of the Internet: Next to traditional *Internet service providers* (ISP) of different size which are interconnected in a tiered, hierarchical fashion, *Internet exchange points* (IXP) emerge – switching platforms where many networks meet to exchange predominantly server traffic directly. While originally build by Tier-2 networks to bypass costly transit via Tier-1 networks, IXPs developed into a competitive marketplace for all kinds of networks. Furthermore, *content distribution networks* (CDN) and cloud providers build their own infrastructure or deploy servers within existing networks to bring content closer to end-users. While the high-level purpose of CDNs and IXPs is to keep local traffic local and reduce delays, at the same time they render transit networks increasingly dispensable. This leads to a "flattening" of the Internet, i.e., a density of major players at single locations.

These changes impact the usage of *address space*, inter-domain *routing*, as well as *traffic* flow in a way that goes beyond existing textbook knowledge. For example, to accommodate users in different parts of the Internet, typical CDNs distribute content to caches placed in different networks and use existing technologies in unconventional ways. In order to redirect a user to the *closest* cache, CDNs perform application-level anycast using the *domain name system* (DNS). Additionally, the *border gateway protocol* (BGP) is used by ISPs and CDNs to implement strategic agreements, deaggregating and dele-

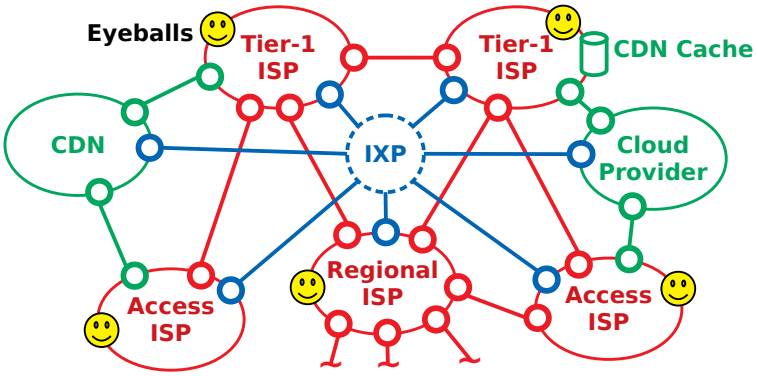


Figure 1.1: Simplified Internet

gating address space in order to, e.g., perform fine-grained traffic engineering; thereby further contributing to the complexity of the Internet.

One indicator of change in the Internet is the application mix, i.e., *the type of traffic that is exchanged* among the networks. For example, in 2008 traffic associated with file-sharing applications were reported to make up to 50% of the upstream traffic [27], which was a costly matter for ISPs since they had to pay transit costs for inter-domain connections induced by these applications. As a consequence, some ISPs started to block popular file-sharing applications [26, 33], while in the research community systems to support ISP-friendly peer selection algorithms were proposed [69, 43, 108]. With the emergence of video streaming platforms including Netflix and YouTube, BitTorrent traffic decreased and was considered to be a negligible contributor of the overall traffic mix [32]. Investigations in 2018, however, show that file-sharing traffic, in particular, BitTorrent, is responsible for 22% of all upstream traffic in the Internet, and with more than 31% the top contributor in Europe, and thus is on the rise again [34].

Another indicator of change is evident at the control plane, in particular, *how traffic is routed* among the networks. This change is, among other things, reflected by the global routing table growth. Somewhat counter-intuitively, the routing table size has continued to grow even after the exhaustion of the IPv4 address allocations in 2011 [28]. Studies investigating routing table inflation find that many of the routes are due to deaggregation, which is used for traffic engineering, load-balancing, and multihoming [72, 131].

A large distributed system like the Internet cannot be expected to grow in size as fast and dynamically as it did in the past decades without constant improvements. Major changes to the infrastructure have proven crucial to the continuous growth and the functioning of the Internet, while protocols have evolved to support these changes. In the early 1990s, the Internet's existence was threatened by several limitations that were introduced by its unexpected popularity and growth, e.g., *class-B exhaustion* or the *routing table explosion* [65]. These limitations led to a series of improvements in routing, allocation and addressing strategies over the course of time.

It is therefore vital to keep track of demands and the resulting changes in the Internet. Understanding how critical players shape the Internet and the traffic they carry sheds light on (i) scalability limits of the Internet, i.e., lack of functionality in existing protocols, and shortcomings in the traditional infrastructures, as well as (ii) flattening of the AS-level hierarchy, accompanied by changes in how traffic is exchanged among the networks.

1.1 Research Question

As pointed out inter-domain routing as well as the associated traffic flow in the Internet is constantly evolving. Thus, we have to constantly revisit which changes to the Internet infrastructure and policies are affecting the current traffic flows. Accordingly, our research challenge is:

How are current business decisions and traffic engineering policies affecting inter-domain traffic flows?

To address the above challenge we tackle the following questions:

A primary reason for our current inability to accurately track a constantly changing Internet is the lack of global vantage points where traffic from a sufficiently large portion of the Internet can be observed at a granularity that is sufficiently fine-grained. This raises the question as to whether or not such vantage points do indeed exist in today's Internet, and if so, *what exactly do they enable us to discern about the Internet as a whole as well as its individual constituents?*

In order to discern the make-up of today's Internet traffic and the interactions of the responsible parties a follow-up question is *how much traffic does each individual constituent contribute to the traffic?* Economic incentives drive many of the main commercial Internet players to deploy their servers deep within third-party networks. Therefore, traffic originating from, or destined to, a specific network might be misattributed to the wrong organization.

Traffic exchanged among networks is often subject to complex business relationships. Given a set of AS-level paths between two networks, the degree of traffic symmetry remains largely unexplored. *Does the ingress and egress traffic, originating from or destined for a specific network always traverse the same AS path?* If not, *what fraction of this traffic flows over alternate paths, and why?*

The continuous growth of the global routing table poses scaling problems in memory-limited routers in large networks. Moreover, it reflects the complexities involved in inter-domain routing, in particular by one if its contributors – prefix delegations – which always involve two ASes. *How do prefix delegations contribute to the routing table growth and what role do they play in the routing ecosystem?*

Given today's IPv4 address scarcity, it is important to understand which addresses are active and how they are used. Yet, measuring the entire address space actively not only requires a well-orchestrated measurement setup, but is also time-critical and susceptible to individual network and host configurations. In a supplementary study, we ask *what are typical pitfalls when conducting Internet-wide active measurements?*

To answer these questions we use multiple vantage points in this work that reflect the diversity of the Internet itself. In particular, we study traffic from a large ISP and a large IXP, as well as measurements obtained from distributed measurement platforms including BGP collectors and a botnet.

1.2 Vantage Points and Datasets

Throughout this work, we use several different vantage points to capture the current state of the Internet ecosystem. Each vantage provides a unique angle of view on the different facets of inter-domain routing and traffic, as well as address usage. In addition to data from these vantage points, we use various external datasets and active measurements to enrich our analyses.

We use data from a large European IXP, a switching platform interconnecting hundreds of networks, like service and cloud providers, or CDNs. This data consists of 17 consecutive weeks of uninterrupted anonymized sFlow records collected using a random sampling of 1 out of 16K. To complement our view on server-related traffic at the IXP we additionally (a) perform active measurements using a list of 25K recursive DNS resolvers seen from one of the largest commercial CDNs, and (b) use a proprietary dataset from a large European Tier-1 ISP, i.e., packet-level traffic traces from a point of presence at a subscriber network. We use the same Tier-1 ISP to investigate traffic asymmetries in the Internet. Thereby, we utilize one week of anonymized and sampled (1 out of 1K) NetFlow data exchanged at inter-domain links of the ISP; in particular to study traffic asymmetries over multiple AS paths induced by routing asymmetry, and the role of hypergiants.

To better understand how traffic is routed through the Internet, we study the underlying routing ecosystem by making use of publicly available routing information from BGP RIB dumps and updates provided by RIPE RIS and RouteViews; two well-known, worldwide distributed BGP collector projects. We use this BGP data to investigate the impact of IPv4 address space exhaustion and traffic engineering on the routing table growth, and how path selection is subsequently affected. Throughout this work, we use this BGP data as a general tool to assign IPs to prefixes and ASes, and to filter unrouted traffic.

Finally, we use the anonymously published measurements from the Carna botnet, including globally distributed ICMP probing of the IPv4 address space.

We use active and passive measurements to investigate the popularity and authenticity of the Carna datasets, and validate their suitability for sound measurement-based networking research. In particular, we reverse-engineer missing meta-data from the published results in order to characterize its hygiene, i.e., how carefully the anonymous authors checked the quality of the data.

1.3 Contributions

The contributions of this work are two-fold: First, we provide a better understanding of each of the vantage points, by quantifying and characterizing their visibility, and by showing what can and cannot be discerned from the corresponding datasets. Second, using the vantage points we investigate inter-domain traffic flow and study how it is affected by traffic engineering and other routing decisions. Moreover, we highlight scalability limits like address space or the size of the global routing table, and how both relate to each other. Our major contributions are:

⋆ **Network Heterogeneity**
We present an approach to characterize network heterogeneity in the Internet. We introduce available internal traffic data from one of the largest European IXPs and provide a method to dissect the traffic in order to distinguish between overall peering

traffic and server-related traffic. We elaborate on the dual role of the IXP as a local and as a global player and provide an understanding of what we can and cannot discern using this vantage point. In a longitudinal study, we investigate the stability of our observations as well as particular events. Additionally, using external meta-data like DNS information, URIs, and X.509 certificates we further characterize server IPs. By identifying server-based network infrastructures and classifying their ownership, we illuminate the network heterogenization and discuss the impact it has on the traditional AS-level view of the Internet.

⋆ **Traffic asymmetry on inter-domain links**
We take a first look at traffic volume asymmetries in the Internet. We make use of traffic data exchanged with other networks via inter-domain links of a large ISP and examine the prevalence of traffic asymmetries in relation to topological distance and the influence of hypergiants. To characterize the impact of routing (or path) asymmetry on traffic asymmetries, we introduce four classes of traffic pattern based on ingress and egress traffic over multiple links. Using seven days of observation time we investigate temporal changes in the different classes.

⋆ **Global Routing Table Growth**
We introduce a thorough analysis of the routing table growth with the focus on prefix delegations by making use of publicly available routing information from two major collector projects. We sanitize the BGP updates and snapshots and show what the individual sources contribute to the global routing table. To identify prefix delegations, we group prefixes based on their overlapping properties and originating AS. Based on AS paths, we further distinguish between different prefix delegation classes. We present a longitudinal analysis and show the evolution of the individual classes over several years. We add information, obtained from various external datasets, about business relationships between the involved networks and correlate them with prefix delegations. Further, we investigate the impact of prefix delegations on path selection by studying large-scale traceroute measurements. Using case studies we report on the diversity of prefix delegations and discuss its impact on the aggregability and the consequent inflation of the routing table.

⋆ **Pitfalls in Internet-scale Measurement studies**
We study Internet-scale measurements performed by an illegal botnet and published by anonymous authors. We introduce the different datasets including ICMP probes, services probes and traceroute measurements. In order to verify the authenticity of the data, we perform checks comparing reverse DNS results and service probes to server IPs. We highlight discrepancies between the description of the available datasets by the authors and our own assessment. We investigate the data quality in terms of probe distribution and activity. From our findings, we elaborate on the limitations of the botnet architecture and characterize the resulting scans of the address space. We attempt to verify the claims of the authors, i.e., the number of censuses allegedly performed and discuss the robustness of the data, the novelty of the measurement method, as well as ethical considerations.

1.4 Overview & Roadmap

In this thesis, we investigate the influence of major Internet players on the inter-domain traffic flow and the underlying routing mechanisms. Table 1.1 provides a brief overview of the main research questions we strive to answer in this thesis:

C.	Question	Work	Data	Novel	Implications
3	*What can be discerned from a global vantage point regarding traffic contribution of major players?*	Longitudinal visibility analysis; group server IPs by organization	IXP, ISP, BGP, DNS	Characterization external visibility of large IXP; scope & stability; algorithm to group server IPs	Observable trends, global events → network heterogenization; traffic to organizations purely based on AS not sufficient; TE; PM
4	*Does ingress/egress traffic always use same AS path?*	Per AS path direction/volume of network traffic	ISP	New traffic asymmetry classification	Asymmetry in traffic volumes → network planning and provisioning; TE; PM
5	*What is the role of prefix delegations and to what extent do they contribute to routing table growth?*	Identify prefix classes; longitudinal study; analysis of delegator/delegatee	BGP, CAIDA, traceroutes	New delegation classification; characterization; correlation with AS size/business relations	Increasing trend → impaired aggregatability of routing table; router design; deaggregation or filtering has impact on path selection; PM
6	*How to deal with datasets of unknown origin/quality?*	Validation of measurement-based networking research	Carna botnet	Study of illicit/poorly documented datasets; reverse-engineer meta-data	Reuse of datasets; ethical considerations

Table 1.1: *Brief overview of research questions order by chapter (For convenience we use the following abbreviations: TE = traffic engineering, PM = policy making).*

The remainder of this thesis is organized as follows:

Chapter 2: Background

We outline the design principles of the Internet and illuminate the evolution of the network layer protocols and functions, e.g., addressing and routing, and its constituents, e.g., registries and major players. In the process, we highlight the developments in the Internet from an educational to a commercial-purpose network, and show why changes to the Internet became necessary by discussing some of the limitations that the Internet was (and still is) facing.

Chapter 3: IXP as a Vantage Point

We investigate how new content distribution models and cloud infrastructure providers change the nature of content delivery. In particular, we are interested in how Internet players like ISPs, CDNs or cloud providers shape the traffic in a large and competitive ecosystem that is driven by a constantly increasing demand by applications for bandwidth. Typical challenges in tracking these developments involve identifying existing and emerging server infrastructures as well as finding the responsible organizations in order to properly attribute traffic.

By studying inter-domain traffic exchanged among hundreds of networks at a large IXP we are able to observe these developments in the Internet ecosystem. Coupled with our methodology to identify server infrastructures and grouping them by organizations, we observe a clear trend among many of the critical Internet players towards network heterogenization. That is, distributed network infrastructures are deployed and operated by today's commercial Internet players. Thus, networks often host servers of other, third-party networks deep within their own infrastructure which is generally not visible outside the corresponding ASes. Our observations contribute to advancing a new mental model for the Internet's ecosystem.

Chapter 4: Traffic Asymmetries on Inter-Domain Links

We investigate traffic volume asymmetries on inter-domain links and its susceptibility to routing asymmetry, i.e., the differences in the sequence of ASes in the two directions.

In particular, given a set of paths between a pair of ASes we study the balance of ingress and egress traffic across the different paths. Asymmetries in traffic volumes across a link have implications for network planning, provisioning, and traffic engineering.

Our analyses, based on traffic data spanning a period of one week from a Tier-1 ISP, highlights that traffic asymmetry is largely unaffected by routing asymmetry. We augment this characterization with some insights into the contribution of hypergiants towards this asymmetry and show that their traffic steering policies do not appear to affect the traffic asymmetry.

Chapter 5: Prefix Delegations via BGP

We investigate the global routing table growth and in particular one of its drivers – prefix delegations – and how they affect path selection. Since forwarding in the Internet is based on the destination address in the IP datagram, every network needs to maintain information about how to reach any other network in the Internet. The more networks that participate in the Internet, the more entries that must be stored in routing tables.

The routing table growth has counter-intuitive relation with address space depletion: On the one hand, the global routing table is increasing despite no more address space being available for allocation at the registries. On the other hand, a common practice for networks is to obtain addresses from their providers, a practice encouraged by registries to preserve increasingly scarce IPv4 addresses. This so-called *provider aggregatable* (PA) address space, can, in theory, be aggregated in the providers' routing table. Yet, while many providers do not aggregate, other providers cannot: Address space announced by multihomed customers can be aggregated by the delegating provider, but not by the other providers, thus multihoming in combination with PA addresses adds to the global routing table growth.

We observe approximately 20% of traffic associated with delegated prefixes at the IXP as well as the ISP. Understanding prefix delegations is essential to understand the complexities in inter-domain traffic. Using publicly available BGP snapshots and updates, we highlight that prefix delegations are more complex than commonly presumed and have a profound impact on the selection of paths on which traffic is traversing.

Chapter 6: Internet-Wide Scans by a Botnet

We perform an analysis on a set of Internet-wide activity measurements of the global address space. The measurements were performed by a presumably illegal botnet and published via BitTorrent by anonymous authors along with a report.

Typically, measurements of this kind and magnitude are hard to perform as they involve non-trivial resources and a careful design of the measurement setup. For example, scanning the complete address space should happen in a timely manner, e.g., due to periodic IP-reassignments to customers at ISPs. Incoherent or misaligned measurement campaigns can lead to a skewed assessment of IP address activity.

Given the discussion of IPv6 deployment and IPv4 address exhaustion [22], knowing which IP addresses are currently in use is of interest. Not surprisingly, the dataset was downloaded by many research and governmental institutions as well as service providers, as we observed by participating in the corresponding BitTorrent swarm. Since the performed measurement is unorthodox and its documentation rather superficial, we highlight the importance of adequate meta-data that enables the reuse of measurements and provide a discussion on its ethical implications.

Lastly, in **Chapter 7** we conclude our work.

2 Background

In this chapter, we outline the history of the Internet, from an educational to a commercial-purpose network. We illuminate some of the design principles and major players in the Internet, the evolution of the network layer functions and protocols, and finally discuss why changes to the Internet became necessary by highlighting some of the limitations that the Internet was (and still is) facing. At the end of this chapter, in Table 2.1, we provide a historical timeline presenting some of the major changes in the Internet from the creation of ARPANET to today's Internet.

2.1 Design Principles

This section briefly outlines the major design principles of the Internet, including packet-switching, layering, and the end-to-end argument.

Packet-Switching: The Internet is a packet-switching network, i.e., discrete chunks of data are exchanged on a shared infrastructure[1]. More precisely, before a stream of data is sent between two hosts, i.e., from source to destination, it is split into small chunks, each of which has a *header* containing information like the address of the destination. In a large network like the Internet, most of the hosts are not directly connected but are interconnected via a set of nodes (or routers) and links. Since forwarding in the Internet is destination based, each router has to know how to forward a packet from one link to the next, based on the destination address.

Packet-switched networks allow the same link to be used by many users at the same time. As in real-life networks like a shared road system, links in the Internet are not uniform, but consist of different technologies and thus provide different transmission

[1] As opposed to circuit-switching where data flows on a dedicated path between source and destination

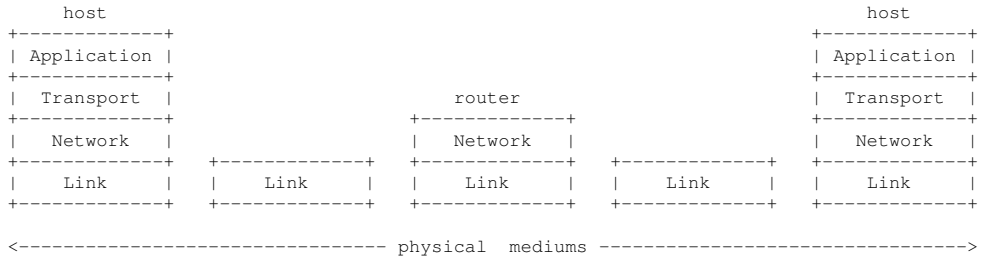


Figure 2.1: Layering and End-to-End Principle

speeds. This can lead to congestion, e.g., when packets enter a router on a fast link and need to be forwarded on a slower link. Since there is no dedicated connection between two hosts, the Internet is in principle a connection-less, best effort network. i.e., there is no guarantee that a packet actually arrives at its destination.

Layering: Internet protocols and function are organized into layers to manage the complexities of packet-switching. Typically, these layers are, from the top to the bottom, *application*, *transport*, *network*, and *link* layer, see Figure 2.1.

While this work mostly focuses on the network layer we give a brief overview of the functions and protocols used at each of the layers. Each layer makes use of the services provided by the respective layers below without the need to care about their specific functioning. This way of abstraction eases the development or improvement of protocols and functions at each individual layer. For example, software engineers developing an Internet application (application layer) do not need to worry about the specifics of the underlying transport or network layer protocols. Similarly, network layer protocols do not need to know the specific functions of link layer protocols, thus enabling the interconnection of networks using different link layer technologies. In the following, we provide example protocols and their essential functions in each of these layers.

- *Application:* Application layer protocols involve, e.g., *Hypertext Transfer Protocol* (HTTP) or *Domain Name System* (DNS) and enable many of today's most popular applications like video streaming, web browsing, emailing, online gaming, and peer-to-peer but also functions like name resolution or content delivery. Application layer protocols can be addressed by a port number in the header of the underlying transport layer protocols, e.g., 80 for HTTP, or 53 for DNS.
- *Transport:* The main functions of the transport layer are reliability, in-order transmission of bytes, flow- and congestion control. There are two major transport layer protocols in use today[2], *Transmission Control Protocol* (TCP) and *User Datagram Protocol* (UDP). Which transport layer protocol to use depends on the requirements of the application. For example, HTTP which relies on the complete and reliable transmission of data uses TCP while accepting the involved signaling overhead and state-keeping (e.g., induced by a handshake). DNS, on the other hand, uses the unreliable UDP since a DNS message usually fits into a single packet, thereby benefitting from UDP's simplicity. Possible packet-loss can be compensated by the application reissuing the request after a timeout.
- *Network:* On the network layer, the Internet can be viewed as a collection of networks (a network of networks) which are connected by routers [158]. The purpose of network layer protocols is to provide global addressing, forwarding traffic from

[2]QUIC is another relatively new transport layer protocol.

source to destination across networks, and routing. Devices within a network typically share the same address space, see Section 2.3.2. If operated by the same organization networks are aggregated in a so-called *autonomous system* (AS). The *Internet Protocol* (IP) is the principal network layer protocol in today's Internet. We elaborate more on IP in Section 2.3.1.

- *Link:* On the link layer, the Internet can be viewed as a collection of devices which are interconnected by communication links and packet switches. Link layer protocols are responsible for moving data on a link by providing an interface to the physical medium. Some of their functions involve detecting or avoiding collisions of signals, or their retransmission if necessary. How and which of these function are implemented depends on the medium, e.g., wired mediums use Ethernet, while wireless mediums use WiFi (IEEE 802.11). Other link technologies can be *Digital subscriber line* (DSL) or *Fiber to the home* (FTTH) used for residential broadband access, but also wide-area wireless technologies like 3G or LTE.

End-to-End Principle: The end-to-end principle is a design principle that guides the placement of functions in a distributed system like the Internet. Thereby, application-specific functions like the reliable transmission of packets should be implemented at communication endpoints rather than at intermediary nodes [148]. The rationale behind this concept is that any function at lower layers (e.g., link layer) like reliability will induce costs to applications that do not need this function or may be redundant if implemented at higher layers on an end-to-end basis. Also, lower layers may not have sufficient information to perform a function efficiently.

Also, modifications to vital functions like reliability can be made much easier at endpoints which typically require updating the operating system running on commodity hardware, compared to expensive core routers where functions are typically implemented in hardware. Another aspect is trust: functions in the network that guarantee, e.g., the integrity of data, leave the endpoints helpless if a network operator deviates from these functions.

This leads to a model of a dumb network (the routers) which needs to support only IP, a light-weight stateless protocol that moves datagrams unmodified across the network, and intelligent endpoints running, e.g., TCP, which provides heavier functions like error detection, retransmission, congestion, and flow control.

2.2 Major Internet Players

In 1986, NSFNET went online and replaced ARPANET, which was retired in 1990, as the backbone network. NSFNET, created by the US National Science Foundation as a government-funded academic research network, was subject to an *acceptable use policy* (AUP), i.e., it was restricted to non-commercial use. Organizations that wanted to connect to the NSFNET had to demonstrate that they serve the progress of science [112]. However, at that time many companies were already using IP networks and strived to interconnect, e.g., to exchange emails. During the late 80's, the first commercial *Internet service providers* (ISP) were founded, marking the transition from a government-funded academic network into what we know today as the Internet: In 1990, the World Wide Web was invented followed by the development of HTTP in 1992 and the first web browsers emerged[3]. Eventually, federal legislation lifted the AUP restrictions in 1993;

[3] Today, most of the traffic is HTTP, see Chapter 3.

and in 1995 NSFNET was decommissioned since it was no longer needed.

Internet Service Providers: The Internet consists mostly of ISPs of varying size and shape. They are run by organizations and can be regarded in a tiered hierarchy based on their topological relation. Large ISPs run backbone networks and connect with other large ISPs at different locations. Typically, they are in competition which each other but mostly agree to exchange traffic on a settlement-free basis. If they do not rely on other providers they are also referred to as *Tier-1 ISPs*, see Figure 1.1. *Regional ISPs* can operate in large regions, such as countries, metropolitan areas, but also in cities. They are called *Tier-2 ISPs* if they rely on upstream ISPs to reach other parts of the Internet while at the same time selling transit services. *Access ISPs*, also referred to as *Tier-3 ISPs*, buy transit services but do not sell transit services themselves.

ISPs that connect business customers or consumers to the Internet using last-mile technologies like DSL are also called *eyeball ISPs*. In a routing context, ISPs have one or more AS numbers and use BGP to express their contractual relations, see Section 2.3.5. Typically, these relations are considered sensitive and are therefore not disclosed.

Internet Exchange Points: Although exceptions exist, *Internet exchange points* (IXP) are large switching platforms and thus operate at the link layer (see Section 2.1). IXPs are the successors of *network access points* (NAP) which in turn were effective as a transitional strategy giving commercial providers a means to bridge the transition from a NSFNET to the modern Internet. In the years after the transition process, IXPs have become marketplaces to do business and sell services to customers. More than 300 IXPs are operated worldwide, and they experience increasing popularity considering the rising number of members, as well as the annual growth of 50-100% of exchanged traffic volume. Other indicators for their success is, e.g., the extremely dense mesh of interconnections which surpasses any number of interconnections outside IXPs [41], or the assembly of prominent members like large CDNs, cloud providers or ISPs.

Originally, IXPs have been build by Tier-2 networks to bypass costly transit via Tier-1 networks which employ a volume-based metric (95th percentile). Typically, IXPs do not charge for exchanged traffic. Public peering links between members at IXPs can be settlement-free but also based on paid peering [66]. Placed in strategically significant locations, e.g., in metropolitan areas, the incentives of such peering points is to keep local traffic local and thereby reducing delays. Moreover, IXPs allow for an easy entry into the market: Everything that is needed is an AS number, a BGP enabled router and a port at the switching platform. Once connected, members can quickly initiate peering agreements by establishing a single connection with a free-to-use route server within the IXP. There can be one-time fees for creating a physical connection as well as monthly charges per port, which can vary with the bandwidth.

CDNs and Cloud Providers: Today's popular services, e.g., social networks, video streaming, are not served from single machine or location. *Content distribution networks* (CDN) are employed by various content providers to achieve fast and reliable content delivery. CDNs can be distinguished between shared CDNs (like Amakai or Limelight) and dedicated CDNs or *cloud providers* (like Google or Microsoft) [119]. CDNs deploy their server-based infrastructures, e.g., distributed caches or data centers, at strategic locations, like IXPs and metropolitan areas, to bring the content closer to the end-users. They employ IP- or Application layer anycast, for example via DNS, see Section 2.3.6. Due to their geographical distribution and typically large bandwidth usage, they are also referred to as *hypergiants*.

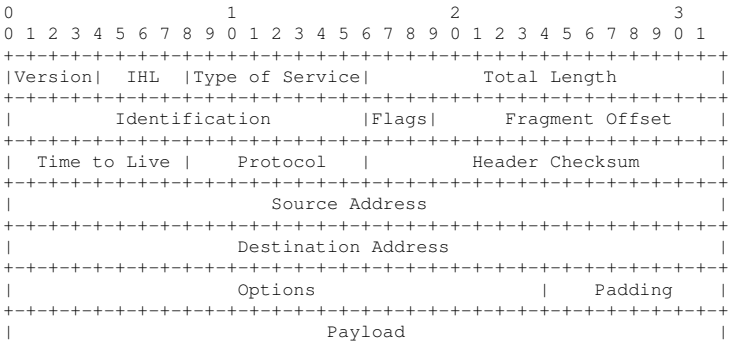


Figure 2.2: Example IP header (adopted from RFC791)

2.3 The Evolution of the Network Layer

In this section, we briefly introduce the core functionalities and protocols of the network layer. This involves IP addressing, the address structure, the allocation of address space, the distinction between routing and forwarding, and routing protocols. Also, we introduce the concept of hostnames and the domain name service; although not part of the network layer, it provides the translation of hostnames, a more memorable representation of IP addresses, which is a vital function in the Internet. For each of the functions and protocols, we outline the historical evolution to highlight necessary changes over the course of the maturing Internet.

2.3.1 Addressing

Internet Protocol: The *Internet Protocol* (IP) is the primary network layer protocol used for relaying data across the Internet. The main functions of IP are (i) addressing, e.g., for source and destination, and (ii) fragmentation to cope with varying datagram sizes of networks using a wide variety of communication technologies like Ethernet or WiFi.

An IP datagram (or packet) is a chunk of data that is exchanged between two hosts using IP. It consists of a header containing information on how to interpret it and the data it carries. Figure 2.2 shows an example IP header. Among other fields, the header consists of the *version* field that specifies the IP version of a datagram, i.e., IPv4 or IPv6. Also, it consists of the *destination address* based on which the datagram is relayed to the corresponding receiver, and the *source address* to identify the sender. The fields *identification*, *flags*, and *fragment offset* are used to support fragmentation in case the datagram is too large for a device on the path from source to destination. The *time to live* field is used to expire packets to avoid congesting the network, e.g., due to loops. Starting with an initial value, it is decreased by every router (hop) on the path and discarded when this value reaches zero.

IPv4: Today, the predominant IP version is referred to as *Internet Protocol version 4* (IPv4), specified in 1980 in RFC760 [16], obsoleted by and standardized in RFC791 in 1981 [18]. IPv4 provides 32 bits for addressing, meaning its address space spans around 4.3B numbers (2^{32}). They are usually represented in human-readable dot-decimal notation, i.e., *a.b.c.d*, whereby each letter represents subsequent blocks of 8 bits from left to right. For example, the address *00000001 00000010 00000100 00001000* can be repre-

sented as *1.2.3.4*.

IPv6: The newest version and direct successor of IPv4 is the *Internet Protocol version 6* (IPv6). It was first specified in 1995 in RFC1883 and became Internet Standard in 2017 (RFC8200). Compared to IPv4, IPv6 extends the address size from 32 bits to 128 bits. The resulting address space is more than $7.9*10^{28}$ larger than what IPv4 provides. Other improvements involve header format simplification in order to decrease processing and bandwidth cost of packet handling, as well as improved support for extensions, options, authentication, and privacy capabilities.

The preferred form of representing an IPv6 address is *x:x:x:x:x:x:x:x*, where each x denotes 16 of the 128 bits from left to right, and each 16-bit block is abbreviated by one to four hexadecimal digits, e.g., *2001:db8:0:0:8:800:200c:417a*. Successive 16-bit blocks of zeros can be compressed using a double colon (::), i.e., *2001:db8::8:800:200c:417a*.

An IPv6 deployment report released by the *Internet Society* (ISOC) in 2018 [36] states that around 25% of all networks announce IPv6 prefixes, while almost 50 countries deliver traffic via IPv6. Also, 28% of the Alexa top 1000 websites are reachable via IPv6[4].

2.3.2 Address Structure

IPs are assigned to network interfaces of hosts and routers. While hosts within a network typically have one interface, routers are designed to forward traffic and exchange route information between networks and thus have more than one interface, each connecting to a different network.[5]

An *IP network* consists of consecutive number of addresses. Initially, in the late 1970s, an IPv4 network, more precisely its *network number*, was denoted by the first 8 bits of an IP address and the remaining 24 bits denoted the local address [16]. Thus, only 256 networks were possible, each supporting ∼16M IPs. The first official network number assignments are documented in RFC750 in 1978 [13].

Classful Addressing: It was soon realized that there will be more than just 256 networks. In 1981, the IETF modified the addressing structure to support *classful addressing*, in particular, three classes of subnets, i.e., *class A* (supporting ∼16M hosts), *class B* (∼65K hosts), and *class C* (256 hosts) subnets [18]. The class was coded in the left-most bits of an address, i.e., class A addresses begin with a '0' bit, class B with '10', and class C with '110', which implicitly specified the respective address ranges. The remaining bits were again divided into a network part and a local address part. The first assignment of classful address blocks is specified in RFC790 in 1981 [17].

Along with classful addressing, the organization of networks changed as well. The increasing amount and diversification of networks made it necessary to split the Internet into a set of independent entities, i.e., ASes. Each AS is assigned an AS number. The first official allocation of AS numbers is specified in RFC820 in 1983 [19]. Private AS numbers are defined in RFC996.

Classless Inter-domain Routing: With the introduction of *Classless Inter-domain Routing* (CIDR) in 1992 in RFC1338 (became standard in RFC1519 in 1993 and obso-

[4] From the point of view of the ISP and IXP, IPv4 is the dominant protocol in terms of traffic and address usage.

[5] Exceptions are, e.g., hosts with more than one interface for the purpose of multi-homing, or mobile devices with multiple wireless technologies.

leted by RFC4632 in 2006), the IPv4 address space was restructured again. It replaced the fixed-length subnets of classful addressing in order to increase its lifespan [87] and is still in use today. Classless blocks of address space also referred to as prefixes, are assigned to networks. *IPv4 prefixes* are represented in dot-decimal notation, equal to IPv4 addresses, followed by a slash and the number of significant bits indicating the prefix length, e.g., *a.b.c.d*/*n*. Compared to three network sizes in classful addressing, network size in CIDR can be any power of 2, e.g., 1, 2, 4, ..., 256, 512, 1024, etc. For example, *1.2.3.4*/*22* denotes a network spanning 1024 hosts (2^{32-22}), ranging from *1.2.0.0* to *1.2.3.255*. IPv6 networks use the CIDR notation as well. For example, the IPv6 network *2001:db8::8:800:200c:417a*/*64* ranges from *2001:db8::* to *2001:db8::ffff:ffff:ffff:ffff* and spans more than 18 Quintillion (2^{128-64}) addresses.

Special Purpose Addresses: Not all of the address can be used for global addressing. Some special purpose address ranges are specifically used by network implementations or applications and are therefore assigned by *Internet Assigned Numbers Authority* (IANA) from reserved space. The most common special purpose addresses are Private-Use or Multicast.

Private address space is allocated to be used by, e.g., enterprise networks that require network layer connectivity between the hosts and are not intended to have access to the public Internet [78, 132]. It is only unique within these networks. The use of private address space does not require any coordination with IANA or an *Internet Registry* (IR), and thus can be used by many private networks. Still, communication between private and public networks can be achieved by middle-boxes like proxies, *network address translators* (NAT), etc. IANA has reserved the following address blocks for use in private networks:

```
10.0.0.0        -   10.255.255.255  (10/8 prefix)
172.16.0.0      -   172.31.255.255  (172.16/12 prefix)
192.168.0.0     -   192.168.255.255 (192.168/16 prefix)
```

The corresponding IPv6 address block (Unique local address) for private use is *fc00::*/*7*.

Multicasting is a technique to forward IP datagrams to a group of hosts and is handled by multicast routers. IPv4 Multicast addresses range from *224.0.0.0* to *239.255.255.255*. The IPv6 counterpart is *ff00::*/*8*.

2.3.3 Allocation of Address Space

Early Registry Function: Before any network in the Internet is qualified to use *Internet identifiers*, e.g., link, socket, port, protocol, and network numbers, they have to be assigned by a responsible authority. Such identifiers used in protocol implementations need to be unique and must not be used by several parties for different purposes. The function to ensure uniqueness exists since the beginning of the Internet and is referred to as IANA. Originally embodied by Jon Postel (back then a graduate student at UCLA, later USC-ISI), it was his responsibility to assign and maintain numeric identifiers that are vital for the operation of the Internet. In a rather informal process, networks which required, for example, address space simply asked for it.

Internet Registry: As the bulk of administrative tasks grew, the function to allocate and assign[6] various numeric identifiers was fulfilled by IANA to a single IR, and was per-

[6] Distinction between the allocation of IP addresses and the assignment of IP addresses: Addresses are allocated by regional registries to an ISP which in turn assigns addresses to its customer base [110].

formed among others by the *Defense Data Network Network Information Center* (DDN-NIC) and later by InterNIC. Also, due to the advancing globalization of the Internet, in 1990 first considerations to distribute the registration function on an international basis were made: Distributed regional registries would be empowered by the IANA and operated on continents, in particular, due to the experienced growth and maturity of the Internet in Europe, Central/South America, Pacific Rim areas (and later Africa) [61]. These plans were further advanced in 1992/1993 in [92, 93].

Regional Internet Registries: In 1996, RFC2050 (later obsoleted by RFC7020 in 2013) specifies the modern-day registry system for Internet numbers, with the following goals: (i) Limited address space should be distributed according to operational needs and should prevent stockpiling, (ii) it should be distributed in a hierarchical manner permitting aggregation to improve scalability of routing, and (iii) the distribution should be documented to ensure uniqueness and enable troubleshooting [110].

Today, IANA is responsible for global coordination of IP addresses (IPv4 and IPv6) as well as autonomous systems and forms the top of allocation hierarchies. While originally IANA directly managed all the IPv4 address space, currently it allocates available addresses to five *Regional Internet Registries* (RIR). The first RIR that was established is RIPE NCC (serving Europe, parts of Asia and the Middle East) and began its operations in 1992 and was followed by APNIC (serving parts of Asia and the Pacific region) in 1993. In 1997 ARIN (serving North America and parts of the Caribbean) was established and inherited all historical registrations by former global registries like DDN-NIC or InterNIC. Some of the historical registrations were transferred to RIPE NCC and APNIC in the scope of the *Early Registration Transfer* (ERX) project. LACNIC (serving Latin America and parts of the Caribbean) was established in 2002 followed by AfriNIC (serving Africa) in 2005 [107].

RIRs further allocate or assign address blocks to organizations within their region. These organizations can be registries of their own, e.g., *National Internet Registries* (NIR) which manage allocations at a national level, *Local Internet Registries* (LIR) like ISPs, or direct assignments to end users. Allocations performed by the five RIRs are made publicly available[7].

2.3.4 Routing vs. Forwarding

The core functions of a typical router are routing and forwarding. Routing involves computing the best route for each destination out of a set of route candidates, collected and stored in the *routing information base* (RIB). The computation is based on metrics, e.g., hop-count, delays, or bandwidth, and policies, which can differ depending on the scope and environment a router operates in. From the resulting best routes a *forwarding information base* (FIB) is constructed. Most routing in the Internet happens in a dynamic fashion as it should reflect changes in current topology, as opposed to static routing, which usually involves human intervention and is typically used in networks where no frequent changes occur.

Forwarding, on the other hand, is the process of forwarding IP datagrams based on information stored in the FIB. Typically, each destination network in the Internet is associated with an entry in the FIB of a router, along with the *next-hop* IP and the outgoing interface. Based on the destination address of an incoming IP datagram, the corresponding entry is searched in the FIB. If no entry is found, the datagram is forwarded via a

[7] E.g., https://ftp.ripe.net/ripe/stats/delegated-ripencc-latest for RIPE NCC

fallback route, i.e., the default route. The next-hop IP indicates the interface of the next router out of a series of routers towards the destination. If the next-hop IP is not set, the destination (final hop) is located in a directly attached network to which the datagram is directly relayed to.

The forwarding process in classful addressing involves identifying the correct class and masking the corresponding left-most bits of the destination IP (8 bits for class A, 16 bits for class B, and 24 bits for class C). The resulting destination network is then searched in the FIB. In CIDR, upon receiving a datagram, the destination IP is matched against entries in the FIB based on their prefix length. This process is referred to as *longest prefix match* and builds on the fact that an IP can match multiple prefixes due to the hierarchical addressing capabilities of CIDR.

In modern routers, the routing process is separated from the actual forwarding process so that the forwarding speed is not impaired during route updates and the subsequent computing of best routes. The separated functions are referred to as data plane or forwarding plane in case of forwarding, and control plane in case of routing [167]. Routing is typically done on general-purpose hardware components, i.e., processors and memory. Forwarding on the other hand, which involves time-sensitive per-packet processing, is usually performed by specialized network processors or application-specific integrated circuits (ASIC), while the FIB can be stored in expensive but fast memory, e.g., ternary content addressable memory (TCAM).

2.3.5 Routing Protocols

Dynamic routing: The goal of dynamic routing protocols is to calculate least-cost paths, react to dynamic changes in a network, e.g., link-failure, link utilization or policy, and exchange routing information. Usually, they are run on routers and employ one or a hybrid of the two routing algorithm classes: *distance vector* (DV) and *link-state* (LS).

LS algorithms use the global and complete knowledge about the state of a network, e.g., connectivity and link costs between nodes. An example LS algorithm is *Dijkstra's algorithm*. In DV algorithms, no node has complete information about the state of a network. Starting out with information about neighboring nodes, they iteratively learn, calculate and exchange least-cost paths to other nodes via their neighbors. An example of DV algorithms is the *Bellman-Ford algorithm*. Compared to LS algorithms, DV algorithms benefit from simplicity and a relatively small message overhead, but converge slower and are less robust.

Routing in the early Internet: Around the 1980's, the Internet consisted of a single network. ARPANET (and Satnet) directly interconnected computers at research centers and later local networks and was under a single routing administration. Routers acting as gateways between the IP networks used the *gateway-to-gateway protocol* (GGP), a distance vector protocol which was first documented in IEN30 in 1979 [14]. Route updates in GGP contained a list of 256 distances where the index of each distance corresponds to the network number that was assigned by Jon Postel (see Section 2.3.2). If a network was not reachable (or not active) the distance was infinite (or 127 in decimal notation). Thus, GGP only supported at most 256 networks.

Due to different implementations on diverse router hardware, it was impractical to make changes to this protocol on all the devices at the same time. Thus, the Internet was split into ASes: ARPANET became the *core* AS which served a backbone role. All other networks were called *stub* AS and interconnected via the core AS [22]. This event initiated

the separation into *intra-domain* and *inter-domain* routing. ASes can use and perform changes to individual routing protocols internally, while for interconnecting with other ASes they used a single standard protocol. An extended version of GGP (RFC823, 1982) which supported classful addressing was the first choice for distributing routing information within an AS, a so-called *Interior Gateway Protocol* (IGP). Today, typical protocols are RIP, OSPF, or IS-IS.

Exterior Gateway Protocol: In order to exchange routing information among ASes, the *Exterior Gateway Protocol* (EGP)[8] was developed; formally specified in RFC904 (1984). EGP relies on a tree-like hierarchical structure, with a single backbone AS to which many so-called stub-ASes connected, which in turn connected universities etc.

With the continued growth and evolution of the Internet, the ASes became increasingly heterogeneous, and more backbone ASes appeared. However, EGP was not designed to support complex topologies. Also, the requirements of the ASes changed towards *policy-based routing* which was not possible with EGP. There was no cost control possible so that is was not fit for the commercialization of the Internet.

Each route announcement in EGP is associated with a metric value, ranging from 0 to 255 and provides a handle for a preferred route. A network receiving announcements should select the route with the lowest cost/distance. However, metrics in EGP are not comparable since they provide no notion about the quality or monetary costs of a route; they can be set more or less arbitrarily. Moreover, EGP is a simple reachability protocol, comparable to distance vector protocols. The announcements provide no means to detect routing loops. While features like backup links could be implemented by carefully (and manually) configuring the routers, EGP basically relies on loop-free graphs, e.g., trees. That was in principle no problem in the early Internet as it consisted of a single backbone, i.e., ARPANET and later NSFNET.

Border Gateway Protocol: With the commercialization of the Internet and the formation of several backbone networks, it was quickly realized that a new inter-domain routing protocol needed a loop detection mechanism. The *Border Gateway Protocol* (BGP) is the successor of EGP and today's de-facto standard inter-domain routing protocol. It was first outlined in RFC1105 in 1989 and further developed to a draft standard in RFC4271 in 2006, commonly referred to as BGP-4. BGP was crucial for the positive trend of the early Internet as it was supposed to pave the way for the looming commercialization and break open limitations of EGP.

Like EGP, BGP is used by ASes to exchange routing information, or reachability information, with other ASes in form of BGP updates announcing network prefixes (see Section 2.3.2) along with so-called *path attributes*. Some of the attributes are *ORIGIN*, *NEXT_HOP*, or *LOCAL_PREF* [144]. The most important and distinctive feature compared to its predecessor is the *AS_PATH* attribute (or simply AS path), i.e., a sequence of ASes through which an announcement has passed; hence, BGP is also referred to as path vector protocol. Not only is it suitable to detect loops, but it also allows complex, general mesh topologies, which accommodates the rapid growth of the Internet to date. Moreover, BGP provides support for CIDR and thus allows to announce several prefixes aggregated into one, or to deaggregate prefixes into smaller ones, and announce them selectively to its neighbors.

Policies via the AS path: The AS_PATH attribute is iteratively augmented by each AS which forwards or originates a particular BGP update, by prepending its own AS number to the AS path. Since BGP is based on the destination-based forwarding paradigm,

[8] Not to be confused with EGP as class of exterior gateway protocols.

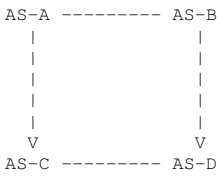


Figure 2.3: Example BGP configuration

every AS that announces BGP updates informally agrees to carry traffic toward the corresponding network prefix. Thereby, at the AS-level, some policy decisions may be enforced.

Consider the example configuration in Figure 2.3. AS_A and AS_B have a peering relationship, i.e., they exchange their traffic and that of their customers for free among each other. AS_A is provider of AS_C, and AS_B is the provider of AS_D, i.e., AS_C and AS_D pay their respective upstream providers to exchange traffic. Since AS_C and AS_D are peers as well, they decide to exchange traffic between themselves directly and for free while using the costly transit via AS_A and AS_B only as a backup solution in case their direct link fails.

Now consider AS_C sends a BGP update announcing its address space to its peer AS_D and to its upstream AS_A. AS_A will happily announce this update to AS_B, and AS_B will further announce it to AS_D, thus offering to carry traffic toward AS_C and generate revenue from their customer links. AS_D, on the other hand, will not announce the update to AS_B in order to prevent traffic being sent via its infrastructure towards AS_C for which AS_D will have to pay AS_B. In other words, there is no economic incentive for AS_D to act as a transit between AS_B and AS_C. Also, note that the loop prevention mechanism of BGP would prevent AS_D to forward the update from AS_B to AS_C.

Now in AS_D's routing table, there are two AS paths towards AS_C:

i. AS_C (direct connection)

ii. AS_B-AS_A-AS_C

In BGP's best path selection process, the path which is the shortest is always preferred, in this case, i., thus the direct connection is used. If the link between AS_C and AS_D fails for some reason, AS_D will notice that and remove i. from its routing table, leaving only ii. to reach AS_C. Thus, AS_D will use the costly upstream link to reach AS_C. Note, the entire scenario works analogously when AS_D announces its address space to AS_C and AS_B.

Other means to enforce policies in BGP is the *COMMUNITIES* path attribute which was added in 1996 [123]. BGP communities can be used to facilitate and simplify the control of routing information, e.g., by telling another BGP neighbor how to handle an announcement sent to it. Thereby, routes are tagged with 32-bit community values; typically they are split into two 16 bit parts, where the first specifies the originating AS and the second a community value based on which an action is performed. Also, an *extended community* attribute was introduced in 2006 which is a 64-bit value and provides not only an extended range but also structure for the community space [160].

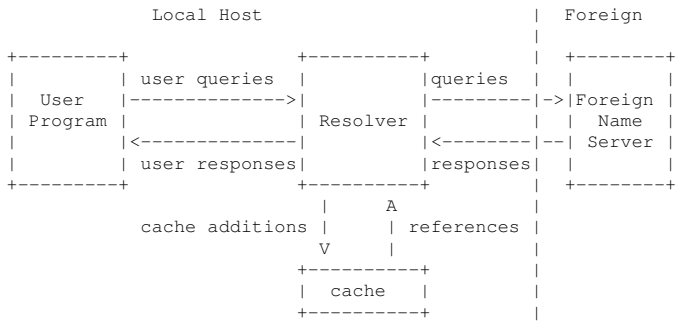


Figure 2.4: Common DNS configuration (adopted from RFC1035)

2.3.6 Hostnames

Numeric Internet addresses are difficult to remember. Their notation has no semantic meaning and they can frequently change. If one wants to connect to different hosts, this can get very bothersome. A host can be identified by a hostname which is a more memorable representation of IPs. In the very beginning of the Internet, each host maintained a file initially called *hosts*, a host table which contains mappings between hostnames and network addresses. This file was used by the local operating system for local hostname resolution.

The translation process was centralized[9] in 1973/74 which includes a machine-translatable ASCII text version and was accessible via the *File Transfer Protocol* (FTP). It was stored by the *network information center* (NIC) and secondary hosts for reliability. In 1982 the host table format was updated, and a query/response function for the translation of hostnames to Internet addresses was added. However, this approach was considered only an interim solution, further maintained by the NIC.

Domains: During the same year, the hierarchical naming convention was described in RFC819, based on the concept of domains and naming authorities [23]. A domain is a composition of one or more dot-separated names. A *fully qualified domain name* (FQDN) specifies all levels of hierarchy, e.g., *www.example.com.* while each dot-separated name has a designation and purpose. The root of the domain name hierarchy is depicted by a single dot and is set at the right of the *top-level domain* (TLD). *com* is a *generic top-level domain* (gTLD), like *net*, *gov*, etc. *Country code top-level domains* (ccTLD) are *us*, *de*, or *fr*. Under each TLD, an arbitrary number of hierarchical layers can be created. The DNS hierarchy further consists of a *second-level domain* (SLD), here *example*. SLDs are maintained by organizations which are accredited by registries responsible for the corresponding TLD. These organization can maintain *sub-domains*, here *www*, or further delegate authority for third-level domains.

Domain Name Registry: A *domain name registry* is a database maintained by the corresponding registry operator (sometimes called NIC) and contains information about domain names registered in a particular domain level. The distribution of registries for domain name registration was first described in RFC920 in 1984 [24]. Today, IANA coordinates and manages the critical root domain of the hierarchy, and delegates authority of TLDs to the corresponding domain name registries, such as DENIC for Germany. Most registries operate on the top-level, however, some registries introduce a second-level domain hierarchy, e.g., *co.uk*.

[9] RFC606, RFC608, RFC623, RFC625

Domain Name System: The *Domain Name System* (DNS) is a distributed, hierarchical database with the purpose of translating domain names to IP addresses. It was outlined back in 1982 in RFC830 and did not change much since then [21]. Each domain is served by (at least) one *authoritative nameserver* that has information about that domain and about nameservers that are responsible for any subordinate domain. For example, root nameservers have information about nameservers responsible for *com*, which in turn have information about nameservers responsible for, e.g., *example.com* and so on. Thus, in order to translate an FQDN, each of the nameservers in the hierarchy need to be queried.

The information returned by nameservers in response to queries are called resource records [25], the most relevant for domain name resolution being *NS*, *A*, and *CNAME*. An NS record contains the domain name of the authoritative nameserver responsible for the queried domain. An A (or AAAA) record contains the IPv4 (or IPv6) address of the queried domain. A CNAME record points to an arbitrary domain in the domain name space. Each of the records is valid for a specific time-to-live (TTL). Upon expiration, the information stored in the records can be reused.

Queries can be performed in two different ways, i.e., *iteratively* and *recursively*. Typically, a local host sends a recursive query to a local *resolver* which performs all necessary iterative queries involved in the translation process, see example configuration in Figure 2.4. The resolver stores the responses in a cache such that subsequent queries can be looked up there before a query is issued.

CDNs and DNS redirection: DNS is heavily used by CDNs to redirect users to the closest cache. Based on the IP address or network prefix of the requester (or the requesting resolver) for a CDN-ized website, the CDN derives the origin of the request and thus is able to assign an appropriate cache. This typically involves CNAME resource records: For example, upon receiving a request for a CDN-ized website (e.g., www.example.com) the corresponding authoritative nameserver returns, instead of an A record, a CNAME record pointing to a CDN domain (e.g., www.example.com.cdn.net); thus delegating the delivery of the content to the CDN. Many of the most popular websites use CDNs for content delivery. Today, more than 50% of the top-1000 Alexa sites utilize a CDN [99].

2.4 Growing Pains

The evolution of the Internet is driven by its popularity, from the creation of the ARPANET in the late 1960's to the present day which is dominated by commercial players. However, since the early Internet was conceived as a pure research-oriented infrastructure the rapid growth and commercialization of the Internet was not anticipated. As a consequence, it threatened to collapse under its own weight several times. Many constraints were due to the architectural history, address assignment strategies or router technology in the face of constant the growth of the Internet.

Due to the continued growth and internationalization of the Internet, in 1990 the *Internet Activities Board* (IAB) suggested to distribute the Internet identifier assignment, that is the assignment of network and autonomous system numbers, to delegated registries [61]; similar to the domain name registration, which at that time was already accommodating the distribution of its function since 1984 [24]. Morover, one year later the IAB outlined three dangers [65]. The Internet would soon experience (i) due to an extensive use the *exhaustion of class B subnets*, (ii) an altogether *depletion of IP addresses*, and (iii) given the number of networks the *growth of the routing table* beyond

what routers and routing protocols could handle at that time.

In the following years, these limitations were key to motivate the introduction of short-term solutions like supernetting, followed by CIDR [169, 88], or NAT [81]. In a parallel effort, allocation strategies for registries [92, 93] were developed which resulted in the creation of the first RIRs and the modern registry system we know today [110, 107]. CIDR turned out to be successful at reducing the consumption rate of IPv4 address space, and, since it outlasted its anticipated life-span, was re-evaluated from a short-term to a mid-term solution in 2006 [87]. Today, CIDR and NAT are still in use.

While the beforementioned solutions were meant to conserve address space and keep routability intact, it should be noted that conservation and routability are often conflicting goals [110]. For example, modifications to the address space which enables more networks with flexible network sizes, e.g., CIDR, can lead to the growth of the global routing table when networks follow their individual interests.

The key scaling issues or constraints that needed to be tackled can be divided into three different categories: (i) *Registration of Internet identifiers*, (ii) *IPv4 Address space exhaustion*, (iii) *Global Routing table growth*. In the following, we elaborate on each of them.

2.4.1 Registration of Internet Identifiers

Internet identifiers such as network numbers (later subnets and prefixes), AS numbers, or hostnames need to be unique in order to prevent two parties from doing two different things using the same identifier. Therefore, a registrar-type of administration became necessary. With the growth of the Internet, registries evolved into globally operating entities.

Domain Name Registry: A popular example for such registries involve the translation of hostnames to addresses. This function evolved from a single text-based file, individually maintained on each host during the times of ARPANET, and led to a hierarchical and globally distributed system, i.e., the DNS, see Section 2.3.6. Interestingly, the naming conventions used in DNS came out of the need to solve the complexity of relaying emails in 1982 [20]. Domain name registries accommodate the growth of the naming universe: As of today, the root domain contains more than 1.500 TLDs [35], and there are almost 340M domain name registrations across all TLDs [37].

Internet Registry: Another popular example is the Internet registry system. The allocation and assignment of address space and AS numbers to networks was initially maintained by Jon Postel, see Section 2.3.3. Starting with 21 assigned network numbers in 1978, it amounts to 40 in January 1981[10]. With the introduction of classful addressing, the number of assigned networks escalated: From 43 class A networks in September 1981 the number increased to a total of 18.781 networks in 1990 (comprising 34 class A, 2.533 class B, and 16.214 class C networks). Today, the assignment and allocation of address space is handled by a hierarchically distributed registry system. It consists of IANA as the top authority and world-wide distributed registries (RIRs, NIRs, and LIRs). From the 86% of IPv4 address space that is assigned to the RIRs, 99% is allocated. 14% is reserved by IANA or IETF.

[10] Jon Postel maintained and updated lists of network numbers in RFC750, RFC755, RFC758, RFC762, RFC770, and RFC776 from 1978 to 1981. Classful addressing was maintained in RFC790, RFC820, RFC870, RFC900, RFC923, RFC943, RFC960, RFC990, RFC997, RFC1020, RFC1062, RFC1117, and RFC1166 from 1981 to 1990.

Not only does the registry system reduce the administration overhead of handling hundreds of thousand networks and AS numbers, but by global coordination, it also facilitates the fair and efficient allocation of address blocks according to the requirements of the individual regions. This became particularly important, e.g., when in 1992 the utilization of class A and class B subnets reached a critical level. RFC1466 suggested in 1993 that the allocation of class A and B should be restricted [93]. While the allocation of class A subnets was at the discretion of IANA, class B subnets were allocated to the regional registries which had to ensure that a requesting organization can justify its need. The address spaces spanned by class C was divided into divisions and allocated to the distributed regional registries.

Transfer Markets: Since the exhaustion of the IPv4 address pool in 2011 [28], IPv4 addresses turned from a free resource to a commodity. In so-called IPv4 transfer markets, RIRs carefully manage the trading of IPv4 blocks. Transfer markets have been legitimized by policies by the RIRs; next to intra-RIR transfers also inter-RIR transfers are possible. Buyers or sellers of IPv4 address blocks send a request to the corresponding RIR, but can also involve third-party participants, i.e., IPv4 brokers [125].

2.4.2 IPv4 Addresss space exhaustion

Wasteful Allocation: Address space, as per definition, is a limited resource. Initially, it was informally given out in generous amounts, i.e., from a pool of 256 network numbers, each spanning ∼16M addresses. From today's perspective clearly a scalability limitation. As a first reaction to this rigid and wasteful assignment of address space was the modification to support classful addressing in order to accommodate different network sizes. However, classful addressing turned out to be inflexible, since class A subnets were too large for most organizations, while class C subnets were usually too small; leaving class B subnets as the preferred choice.

Further optimizations were made with the standardization of CIDR in 1983, which not only increases the number of possible networks but also allows for more flexibility in terms of network sizes. Thereby, the smallest allocation is typically a /24 prefix (former class C block) [88]. CIDR is still today's standard of address space structure.

When CIDR was deployed it was considered a transitional technology while a new protocol with a larger address space is developed, i.e., IPv6. It is supposed to give the Internet breathing space with regard to IPv4 exhaustion [112, 169, 88]. However, the adoption of IPv6 which was proposed in RFC1883 in 1995, turned out to be slower than originally expected and was thus re-evaluated to be a long-term solution. Short-term solutions were suggested in order to conserve the currently available IPv4 address space.

Conserve Address Space: Due to a limited pool of allocatable IPv4 address blocks, guidelines for ISPs were published in RFC2050 in 1996 [110]. In order to efficiently utilize the available address space, among others, it suggests that an ISP or organization should request so-called *provider-aggregatable* (PA) address space from its upstream provider or LIR, if possible. This process is often referred to as *prefix delegation* [72, 153]. Only if justified, can they obtain *provider-independent* (PI) address space directly from the RIR.

RFC2050 also discourages the use of static IPv4 address assignments to dial-up users by ISPs given the current consumption rate. In static assignments, an IP address is permanently reserved for a host independent of whether it is actually active or not, thus wasting this resource. Using a dynamic assignment of IP addresses, on the other

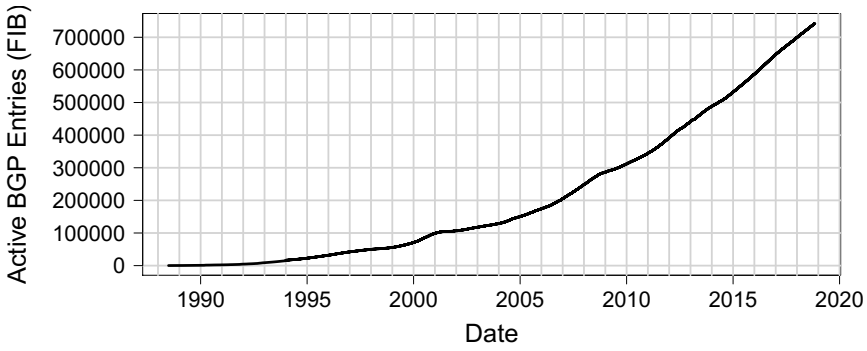


Figure 2.5: Growth of the global routing table for 30 years (data from www.cidr-report.org).

hand, an IP can be re-assigned from an inactive to an active host. Roughly 450M IPv4 addresses, which corresponds to ∼15% of routed IPv4 addresses in 2015 were reported to be potentially unused [145].

Moreover, private networks or organizations which are not connected to the global Internet, or have no intention to ever do so, should not use public IPv4 address space, but instead, resort to private-use addressing [132]. Another short-term solution is *Network Address Translation* (NAT) which provides a mechanism to connect private networks using private address space to public networks using globally unique addresses. Thereby a translation table is used. In 1999, Carrier-Grade NATs (CGN) are proposed which are deployed in ISPs (as opposed to NAT which is deployed in subscriber networks) [106]. RFC6598 requests the allocation of private address space for use in CGN deployments [164].

2.4.3 Global Routing table growth

In order for a network (or AS) to be globally reachable, every default-route-free router needs to know at least one route towards it. Therefore, every network that announces its address space to the Internet contributes to the global routing table growth. The Internet went through several stages of routing protocol development and improvements in router hardware in order to support the growing number of networks. Thereby, the global routing table growth is considered a constant scalability issue and comes with different ramifications.

Design limitations: Protocol and hardware designers did not anticipate the rapid increase of networks. An example of limitations in protocol design is the restriction to 256 networks in the early GGP from 1979, see Section 2.3.5. As a consequence, GGP was adapted to support classful addressing and conjoined with EGP in order to implement intra- and inter-domain routing. However, since router components were very expensive they were provisioned to accommodate only a subset of possible networks, at the time deemed sufficient. For example, consider a fragmented IP datagram containing the complete routing table. If the routing table exceeded a certain size, e.g., 4.000 networks, some old router's network interfaces had to reallocate buffers if too many (e.g., four) fragments were to be received. This could lead to the loss of parts of the routing table and subsequently to routing instabilities [112].

Other recent examples involve limitations induced by limited TCAMs in modern router devices. When the number of entries in the FIB exceeds a certain amount, some of the

processing had to be outsourced to software, e.g., the central CPU and memory, which is considerably slower. With the growing number of networks, network operators have to make provision to accommodate enough prefixes for some years to come. However, since memory is scarce and expensive the FIB size was limited. Router vendors like Cisco gradually increased the amount of TCAM in different generations of hardware. Upgrades were necessary, e.g., when in 2008 the routing table size hit 256K entries, or 512K later in 2014, see Figure 2.5. Yet, outages due to insufficient FIB size occurred [31].

The sheer amount of networks: The more networks exist the more reachability information needs to be stored in memory. Much worse, a proportional amount of route updates are triggered when changes occur, however, each router has a limited amount of processing power. Once that limit is exceeded, unprocessed updates are queued up and may not necessarily reflect the current topology. Thus, the overwhelming amount of routing updates may rapidly turn into oscillation, due to inconsistent views. Converging into a stable state can take a significant amount of time and can noticeably impair the performance of (global) services in the Internet, or lead to outages.

Routing table aggregation: Stopping the routing table growth is not feasible since IPv4 networks (and in particular IPv6 networks) will keep joining the Internet in the future. However, it can be slowed down and give some breathing space to (i) router vendors to increase the capacity in affordable router technology and to (ii) network operators between necessary overhaul cycles.

In order to slow down the rate of this growth a mechanism for the aggregation of routing information is necessary. However, due to the distributed nature of the Internet, routing table aggregation requires address assignment in a coordinated fashion where contiguous address space is assigned hierarchically. Beginning at the highest level, blocks of addresses are allocated to various registries responsible for different regions, i.e., continents [93, 107], see Section 2.3.3. Furthermore, sufficient address blocks are allocated to ISPs which can assign or allocate subsets of these blocks (PA address space) to their customer networks, e.g., smaller service providers, data centers, organizations, etc [88]. This implies that customer networks using their service provider's address space will be routed via that service provider and thus allows for easy aggregation at multiple levels of the hierarchy.

Yet, we observe the dramatic increase of routing table entries. Despite the fact that the last remaining address blocks have been allocated by IANA to the five RIRs in 2011 [28], the rapid growth continues. Currently, the global routing table contains more than 700.000 entries and has been growing at an annual rate of roughly 50k, see Figure 2.5. Today, there are two commonly known reasons for the continuation of the routing table growth: Traffic engineering and delegation of provider aggregatable (PA) address space[11].

Traffic Engineering: Aggregation of prefixes is considered crucial for reducing the routing table size as well as the rate of BGP updates. There are multiple reasons to filter but also not to filter deaggregated (or more specific) BGP routes. Some of them include traffic engineering, enforcing contract compliance, and memory preservation. However, arbitrarily filtering more specifics can lead to unexpected traffic flow as described in RFC7789 [60]. Moreover, operators rely on deaggregation and often tolerate the consequent bloat of the routing table to enable services such as multi-homing and realize traffic engineering [72, 56, 131]. Also, in the forwarding process, the more specific prefix always wins (see Section 2.3.4) and thus attracts traffic. In order not to *lose* traffic

[11] In Chapter 5 we study the routing table growth and show that prefix delegations are more complex than commonly presumed.

to competitors, they follow suit and announce more specifics as well [115]. We confirm that deaggregation is increasingly contributing to the growth of the routing table [72], i.e., around half the global routing table entries are due to deaggregation.

2.5 Historical Overview

Year	Description
1969:	*inception of ARPANET, first nodes were brought online [74]*
1974:	TCP (RFC675), obsoleted by RFC793 (1981, Internet standard), RFC7805 (2016)
1978:	IPv4 (IEN54), RFC760 (1980), RFC791 (1981, Internet standard, *classful addressing*)
1979:	GGP (IEN30 and IEN109)
1979:	IAB (Internet Architecture Board)
1982:	EGP (RFC827, conceptually discussed), RFC904 (1984, formally developed)
1983:	*transition from NPC to TCP/IP standard protocol suite in ARPANET, flag day Jan. 1st*
1983:	DNS (RFC882), obsoleted by RFC1034 and RFC1035 (1987, Internet standard, CNAME)
1984:	"End-to-End arguments in system design"; Saltzer, Reed, Clark
1986:	*NSFNET goes online*
1986:	IETF (Internet Engineering Task Force)
1987:	*first commercial ISP was founded [74]*
1989:	BGP-1 (RFC1105), obsoleted by BGP-2 RFC1163 (1990) and BGP-3 RFC1267 (1991)
1990:	*ARPANET was retired, NSFNET became backbone*
1990:	*Sir Tim Berners-Lee invents the World Wide Web [74]*
1992:	Supernetting / CIDR (RFC1338), obsoleted by RFC1519 (1993, proposed standard) and RFC4632 (2006, best current practice), *classless addressing*, *allocation of "blocks of C"*
1992:	Guidelines for Management of IP Address Space (RFC1366), obsoleted by RFC1466 (1993), RFC2050 (1996), RFC7020 (2013)
1992:	IESG Deliberations on Routing and Addressing (RFC1380): *Class B network number exhaustion*, *Routing table explosion*, *IP address space exhaustion*
1992:	*first RIR founded*
1992:	HTTP/0.9
1993:	An Architecture for IP Address Allocation with CIDR (RFC1518)
1994:	BGP-4 (RFC1654, proposed standard) support of CIDR, obsoleted by RFC1771 (1995, draft standard) and RFC4271 (2006, draft standard)
1995:	*NSFNET was decommissioned and replaced by backbones operated by several commercial Internet Service Providers.*
1995:	IPv6 (RFC1883, proposed standard), obsoleted by RFC2460 (1998, draft standard), RFC8200 (2017, Internet standard)
1996:	HTTP/1.0 (RFC1945)
1997:	HTTP/1.1 (RFC2068, proposed standard), obsoleted by RFC2616 (1999, draft standard) and RFC7230-7235 (2014, proposed standard)
2011:	*allocation of last 5 /8s from IANA to RIRs*
2015:	HTTP/2.0 (RFC7540, proposed standard)

Table 2.1: Historical Overview of Major Changes in the Internet

3

IXP as a Vantage Point

The Internet has grown in size and complexity over the past decades. Understanding current practices employed, and how in particular addressing, routing, and traffic flow is affected by them is of general interest. In this chapter we begin our quest to identify and characterize these changes by considering one of today's most relevant, if not the most relevant constituent in the Internet's ecosystem — the Internet exchange point (IXP).

3.1 Tracking Developements at IXPs

Due to the ever-growing demand for Web-based traffic [70, 120] (e. g., HD video and other streaming media, e-commerce services), together with the proliferation of new Internet-enabled devices, new content distribution models and cloud infrastructure providers are radically transforming the nature of content delivery in today's Internet. These features are also having a profound impact on how some of the main Internet players (e. g., ISPs, CDNs, Web hosting companies, and content providers) operate in such a dynamic environment and do business in an increasingly competitive marketplace. Unfortunately, carefully tracking these developments to obtain an accurate picture of how this critical cast of players shapes and impacts much of the Internet and its traffic has become an increasingly daunting task. In the past, attempts at painting such a picture had limited success because they typically relied on piecing together incomplete and often inaccurate information from many different sources of varying quality [77, 136, 149] or using various types of hard-to-get (i. e., proprietary) datasets [120] or hard-to-justify estimates of difficult-to-measure (e. g., inter-AS traffic matrix) quantities [64].

The large European IXP considered in this work also featured prominently in the recent work by Ager et al. [41]. However, while that study focused squarely on the discovery

of a surprisingly rich peering fabric among the member ASes of that IXP, in this work, we are mainly concerned with mining the traffic seen at this IXP to determine the IXP's visibility into the Internet. Put differently, while [41] exploited the IXP measurements to obtain an accurate picture of the "inside" of this IXP (i. e., its member ASes, their peerings, and the IXP-specific traffic matrix), this work mines recent traffic data to obtain a view of the "outside" of the IXP; that is, the larger Internet beyond the boundary formed by the members of the IXP. In terms of results, while [41] highlighted the severe level of incompleteness of the commonly-studied AS maps of the Internet, this work establishes and provides concrete evidence for why and in what sense this traditional AS-level view— although still useful for exploring and understanding various connectivity- or reachability-related aspects— is largely inept for accounting for critical elements of the networks that make up today's Internet. Thus, representing two largely complementary efforts, the combined findings of [41] and of this work take the observations of the study by Labovitz et al. [120] to the next level. In the process, we identify and outline an alternative and largely orthogonal perspective to the traditional AS-level view that centers around organizations or companies and their server-based infrastructures that are spread across many networks and countries and defy traditional network and geographic boundaries.

In order to track a constantly changing Internet we make use of a large European IXP which carries traffic from a large portion of the Internet. In numbers, we observe week-in and week-out traffic of more than 10 PB (daily average) from around 42K ASes and around 450K routed networks at a fine-grained granularity that allows us to discern the make up of today's Internet traffic and the interactions of the responsible parties. Our major findings in this chapter are:

1. The visibility of our vantage point provides a global view as we can observe almost 230M IPv4 addresses from almost all countries in the world in a single week. Moreover, we identify almost 1.5M server IPv4 addresses (from around 20K ASes) which are responsible for more than 70% of the overall traffic (after filtering). These numbers provide the first of its kind assessment of IPv4 addresses actively involved in traffic exchange that was observed at a single vantage point to that date. Considering the existence of other European IXPs of similar size, according to released press info at the respective Web-sites, these results outline the importance of IXPs as a truly rich data source. Yet we note that the role as a global vantage point comes with caveats.

2. Our weekly observations reveal a steady picture of traffic volumes exchanged at that vantage point, indicating that any weekly snapshot provides more or less the same information. In particular, considering server-related traffic in each of the 17 weeks of observation, we see that out of the 1.5M server IPs around 30% are always visible (from around 70% respecitvely stable ASes). Thereby, these stable IPs/ASes are responsible for more than 60% of traffic. Yet, through public information we note that the IXP exhibits a constant growth of new members and upgrades to higher port-speeds of existing members. Studying consecutive snapshots, we are able to observe e.g. the increase of HTTPS traffic, expansion of cloud services, or the influence of major events like large-scale outages.

3. When grouping the identified server IPs by organizations, we are able observe a clear trend towards heterogeneous networks and network interconnections. Thereby, organizations take advantage of network diversity and purposefully deploy their server-based infrastructures accross mutliple networks. We find that the traditional AS-level view cannot capture the heterogenizations in its entirety for

two reasons. First, some organizations operate without an AS number which leads to traffic go unnoticed or misattributed to a different organization. Second, caches from members like Akamai can be placed in/behind other members as well. We find that 11% of Akamai's traffic stems from non-Akamai members (which involves 15K out of 21K Akamai servers). While these findings are not specific at this IXP they argue that future studies of business strategies and relationships have to move beyond the largely traffic-agnostic AS-level view.

3.2 IXP as a Rich Data Source

In this section, we describe the IXP measurements that are at our disposal for this study and sketch and illustrate the basic methodology we use to identify the traffic components relevant for our work.[12] We also list and comment on the different IXP-external datasets that we rely on throughout this work to show what we can and cannot discern from the IXP-internal measurements alone.[13]

3.2.1 Available IXP-internal datasets

The work reported in this work is based on traffic measurements collected between August 27 (beginning of week 35) and December 23 (end of week 51) of 2012 at one of the largest IXPs in Europe. At the beginning of the measurement period in week 35, this IXP had 443 member ASes that exchanged on average some 11.9 PB of traffic per day over the IXP's public peering infrastructure (i. e., a layer-2 switching fabric distributed over a number of data centers within the city where the IXP is located). During the measurement period, the IXP added between 1-2 members per week. Specifically, the measurements we rely on consist of 17 consecutive weeks of uninterrupted anonymized sFlow records that contain Ethernet frame samples that were collected using a random sampling of 1 out of 16K. sFlow captures the first 128 bytes of each sampled frame. This implies that in the case of IPv4 packets the available information consists of the full IP and transport layer headers and 74 and 86 bytes of TCP and UDP payload, respectively. For further details about the IXP infrastructure itself as well as the collected sFlow measurements (e. g., absence of sampling bias), we refer to [41]. In the following, we use our week 45 data to illustrate our method. The other weekly snapshots produce very similar results and are discussed in more detail in Section 3.4.

3.2.2 Methods for dissecting the IXP's traffic

Peering traffic

Figure 3.1 details the filtering steps that we applied to the raw sFlow records collected at this IXP to obtain what we refer to as the "peering traffic" component of the overall traffic. As shown in Figure 3.1, after removing from the overall traffic, in succession, all non-IPv4 traffic (i. e., native IPv6 and other protocols; roughly 0.4% of the total traffic, most of which is native IPv6), all traffic that is either not member-to-member or stays

[12] For an overview of the importance of IXPs for today's Internet, we refer to [66].

[13] We anonymize IPs by using a prefix-preserving function on-the-fly. That is, at no time we store the IXP and ISP dataset with complete IP address information in it. To perform active measurements, we augment the prefixes with necessary information.

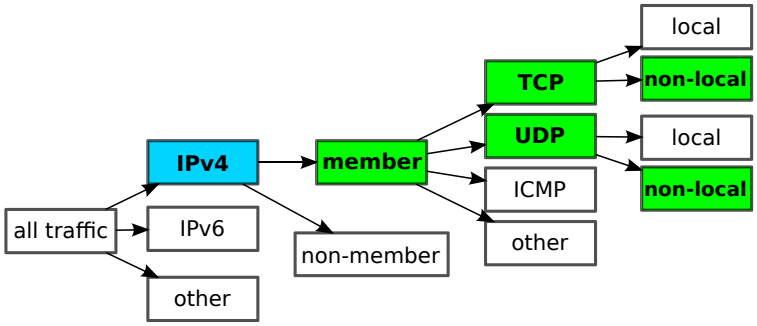


Figure 3.1: Traffic filtering steps

local (e. g., IXP management traffic; about 0.6%), all member-to-member IPv4 traffic that is not TCP or UDP (i. e., ICMP and other transport protocols; less than 0.5%), this peering traffic makes up more than 98.5% of the total traffic. As an interesting by-product, we observe that 82% of the peering traffic is TCP and 18% is UDP.

Web server-related traffic

We next identify the portion of the peering traffic that can be unambiguously identified as Web server-related traffic. Our motivation is that Web servers are generally considered to be the engines of e-commerce, which in turn argues that Web server-related traffic is, in general, a good proxy for the commercial portion of Internet traffic. Accordingly, we focus on HTTP and HTTPS and describe the filtering steps for extracting their traffic.

To identify HTTP traffic, we rely primarily on commonly-used string-matching techniques applied to the content of the 128 bytes of each sampled frame. We use two different patterns. The first pattern matches the initial line of request and response packets and looks for HTTP method words (e. g., GET, HEAD, POST) and the words HTTP/1.{0,1}. The second pattern applies to header lines in any packet of a connection and relies on commonly used HTTP header field words as documented in the relevant RFCs and W3C specifications (e. g., Host, Server, Access-Control-Allow-Methods). Using these techniques enables us to identify which of the IP endpoints act as servers and which ones act as clients. When applied to our week 45 data, we identify about 1.3 million server IPs together with roughly 40 million client IPs. Checking the port numbers, we verify that more than 80% of the server IPs use the expected TCP ports, i. e., 80 and 8080. Some 5% of them also use 1935 (RTMP) as well as 443 (HTTPS). Note that by relying on string-matching, we miss those servers for which our sFlow records do not contain sufficient information; we also might mis-classify as clients some of those servers that "talk" with other servers and for which only their client-related activity is captured in our data.

With respect to HTTPS traffic, since we cannot use pattern matching directly due to encryption, we use a mixed passive and active measurement approach. In a first step, we use traffic on TCP port 443 to identify a candidate set of IPs of HTTPS servers. Here, we clearly miss HTTPS servers that do not use port 443, but we consider them not to be commercially relevant. However, given that TCP port 443 is commonly used to circumvent firewalls and proxy rules for other kinds of traffic (e. g., SSH servers or

VPNs running TCP port 443), in a second step we rule out non-HTTPS related use by relying on active measurements. For this purpose, we crawl each IP in our candidate set for an X.509 certificate chain and check the validity of the returned X.509 certificates. For those IPs that pass the checks of the certificate, we extract the names for which the X.509 certificate is valid and the purpose for which it was issued. In particular, we check the following properties in each retrieved X.509 certificate: *(a) certificate subject*, *(b) alternative names*, *(c) key usage* (purpose), *(d) certificate chain*, *(e) validity time*, and *(f) stability over time*. If a certificate does not pass any of the tests, we do not consider it in the analysis.

We keep only the IPs that have a certificate subject and alternative names with valid domains and also valid country-code second-level domains (ccSLD) according to the definition in [135]. Next, we check if the key usage explicitly indicates a Web server role. In the certificate chain we check if the delivered certificates do really refer to each other in the right order they are listed up to the root certificate, which must be contained in the current Linux/Ubuntu white-list. Next, we verify the validity time of each certificate in the chain by comparing it to the timestamp the certificate fetching was performed. Lastly, we perform the active measurements several times and check for changes because IPs in cloud deployments can change their role very quickly and frequently. Ignoring validity time, we require that all the certificates fetched from a single IP have the same properties. In the case of our week 45 data, starting with a candidate set of approximately 1.5M IPs, some 500K respond to repeated active measurements, of which 250K are in the end identified as HTTPS server IPs.

When combined, these filtering steps yield approximately 1.5M different Web server IPs (including the 250K HTTPS server IPs). In total, these HTTP and HTTPS server IPs are responsible for or "see" more than 70% of the peering traffic portion of the total traffic. Some 350K of these IP addresses appear in both sets and are examples of multi-purpose servers; that is, servers with one IP address that see activity on multiple ports. Multi-purpose servers are popular with commercial Internet players (e.g., Akamai which uses TCP port 80 (HTTP) and TCP port 1935 (RTMP)), and their presence in our data partially explains why we see a larger percentage of Web server-related traffic than what is typically reported in the literature [82, 91, 128], but is often based on a strictly port-based traffic classification [70, 120].

Among the identified HTTP and HTTPS server IPs, we find some 200K IPs that act both as servers and as clients. These are responsible for some 10% of the server-related traffic. Upon closer inspection of the top contributors in this category, we find that they typically belong to major CDNs (e. g., EdgeCast, Limelight) or network operators (e. g., Eweka). Thus, the large traffic share of these servers is not surprising and reflects typical machine-to-machine traffic associated with operating, for example, a CDN. Another class of IPs in this category are proxies or clients that are managed via a server interface (or vice versa).

In the context of this work, it is important to clarify the notion of a server IP. Throughout this work, a server IP is defined as a publicly routed IP address of a server. As such, it can represent many different real-world scenarios, including a single (multi-purpose) server, a rack of multiple servers, or a front-end server acting as a gateway to possibly thousands of back-end servers (e.g., an entire data center). In fact, Figure 3.2 shows the traffic share of each server IP seen in the week 45 data. It highlights the presence of individual server IPs that are responsible for more than 0.5% of all server–related traffic! Indeed, the top 34 server IPs are responsible for more than 6% of the overall server traffic. These server IPs cannot be single machines. Upon closer examination,

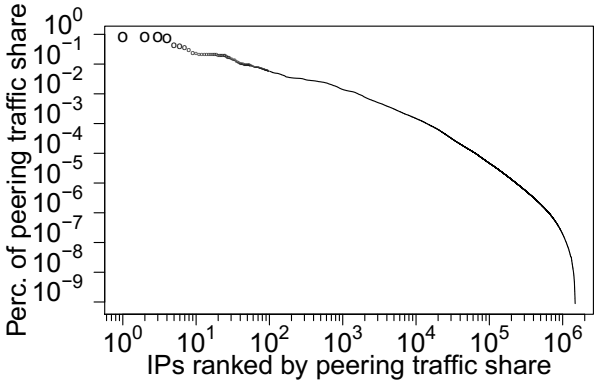


Figure 3.2: Traffic per server IP sorted by traffic share

they are identified as belonging to a cast of Internet players that includes CDNs, large content providers, streamers, virtual backbone providers, and resellers, and thus represent front-end servers to large data centers and/or anycast services. Henceforth, we use the term server to refer to a server IP as defined above.

3.2.3 Available IXP-external datasets

When appropriate and feasible, we augment our IXP-based findings with active and passive measurements that do not involve the IXP in any form or shape and are all collected in parallel to our IXP data collection. Such complementary information allows us to verify, check, or refine the IXP-based findings.

One example of a complementary IXP-external dataset is a proprietary dataset from a large European Tier-1 ISP consisting of packet-level traffic traces.[14] With the help of the network intrusion detection system Bro [140] we produce the HTTP and DNS logs, extract the Web server-related traffic and the corresponding server IPs from the logs, and rely on the resulting data in Section 3.3.

For another example, we use the list of the top 280K DNS recursive resolvers—as seen by one of the largest commercial CDNs—as a starting set to find a highly distributed set of DNS resolvers that are available for active measurements such as doing reverse DNS lookups or performing active DNS queries. From this initial list of DNS servers, we eliminate those that cannot be used for active measurements (i.e., those that are not open, delegate DNS resolutions to other resolvers, or provide incorrect answers) and end up with a final list of about 25K DNS resolvers in some 12K ASes that are used for active measurements in Section 3.3.3.

Other examples of IXP-external data we use in this work include the publicly available lists of the top-1M or top-1K popular Web sites that can be downloaded from `www.alexa.com`. We obtained these lists for each of the weeks for which we have IXP data. We also utilized blogs and technical information found on the official Web sites of the various technology companies and Internet players. In addition, we make extensive use of publicly available BGP-based data that is collected on an ongoing basis by Route-Views [12], RIPE RIS [10], Team Cymru [11], etc.

[14] For this trace we anonymized the client information before applying the analysis with the network intrusion detection system Bro. We always use a prefix preserving function when anonymizing IPs.

3.2.4 IP server meta-data

Our efforts in Section 3.5 rely on certain meta-data that we collect for server IPs and that is obtained from DNS information, URIs, and X.509 certificates from HTTPS servers.

Regarding DNS information, obvious meta-information is the hostname(s) of a server IP. This information is useful because large organizations [100] often follow industrial standards in their naming schema's for servers that they operate or host in their own networks. Another useful piece of meta-data is the Start of Authority (SOA) resource record which relates to the administrative authority and can be resolved iteratively. This way one can often find a common root for organizations that do not use a unified naming schema. Note that the SOA record is often present, even when there is no hostname record available or an ARPA address is returned in the reverse lookup of a server IP.

Next, the URI as well as the authority associated with the hostname give us hints regarding the organization that is responsible for the content. For example, for the URI `youtube.com`, one finds the SOA resource record `google.com` and thus can associate Youtube with Google.

Lastly, the X.509 certificates reveal several useful pieces of meta-data. First, they list the base set of URIs that can be served by the corresponding server IP. Second, some server IPs have certificates with multiple names that can be used to find additional URIs. This is typically the case for hosting companies that host multiple sites on a single server IP. In addition, it is used by CDNs that serve multiple different domains with the same physical infrastructure. Moreover, the names found in the certificates can be mapped to SOA resource records as well.

Overall, we are able to extract DNS information for 71.7%, at least one URI for 23.8%, and X.509 certificate information for 17.7% of the 1.5M server IPs that we see in our week 45 data. For 81.9% of all the server IPs, we have at least one of the three pieces of information. For example, for streamers, one typically has no assigned URI, but information from DNS. Before using this rich meta-data in Section 3.5, we clean it by removing non-valid URIs, SOA resource records of the Regional Internet Registries (RIRs) such as `ripe.net`, etc. This cleaning effort reduces the pool of server IPs by less than 3%.

3.3 Local yet Global

The main purpose of this section is to show that our IXP represents an intriguing vantage point, with excellent visibility into the Internet as a whole. This finding of the IXP's important global role complements earlier observations that have focused on the important local role that this large European IXP plays for the greater geographic region where it is located [41], and we further elaborate here on its dual role as a local and as a global player. Importantly, we also discuss what we can and cannot discern about the Internet as a whole or its individual constituents based on measurements taken at this vantage point. The reported numbers are for the week 45 measurements when the IXP had 452 members that exchanged some 14 PB of traffic per day and are complemented by a longitudinal analysis in Section 3.4.

		week 45	educated guesses of ground-truth
Peering Traffic	IPs	232,460,635	(routed) approx. 2.6B
	#ASes	42,825	approx. 43K
	Subnets	445,051	450K+
	countries	242	250
Server Traffic	IPs	1,488,286	unknown
	#ASes	19,824	unknown
	Subnets	75,841	unknown
	Countries	200	250

Table 3.1: IXP summary statistics—week 45

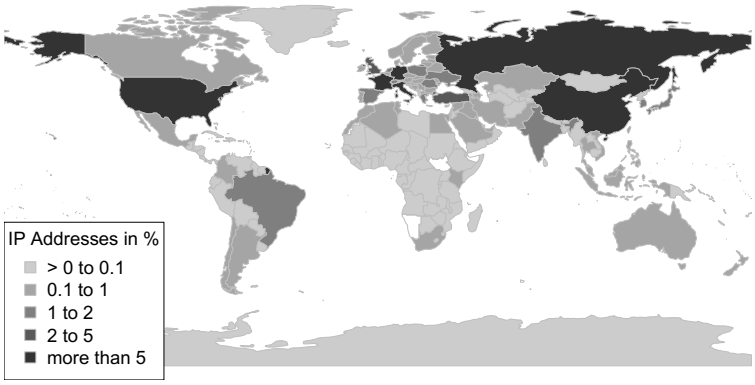


Figure 3.3: Percentage of IPs per country—week 45

3.3.1 On the global role of the IXP

By providing a well-defined set of steps and requirements for establishing peering links between member networks, IXPs clearly satisfy the main reason for why they exist in the first place – keeping local traffic local. To assess the visibility into the global Internet that comes with using a large European IXP as a vantage point, we focus on the peering traffic component (see Section 3.2.2) and summarize in Table 3.1 the pertinent results.

First, in this single geographically well-localized facility, we observe during a one-week period approximately 230M+ unique IPv4 addresses (recall that the portion of native IPv6 traffic seen at this IXP is negligible). This number corresponds to approximately 10% of the advertised address space[15] which suggests that this IXP "sees" a significant fraction of the ground truth. The global role of this IXP is further illuminated by geo-locating all 230M+ IP addresses at the country-level granularity [142] and observing that this IXP "sees" traffic from every country of the world, except for places such as Western Sahara, Christmas Islands, or Cocos (Keeling) Islands. This ability to see the global Internet from this single vantage point is visualized in Figure 3.3, where the different countries' shades of gray indicate which percentage of IPs a given country contributes to the IPs seen at this IXP.

Second, when mapping the encountered 230M+ IP addresses to more network-specific

[15]The number of total address space advertised in week 45 amounts to 2.6B IPs according to http://bgp.potaroo.net/as2.0/.

entities such as subnets or prefixes and ASes, we confirm the IXP's ability to "see" the Internet. More precisely, in terms of subnets/prefixes, this IXP "sees" traffic from 445K subnets; that is, from essentially all actively routed prefixes. Determining the precise number of actively routed prefixes in the Internet in any given week remains an imprecise science as it depends on the publicly available BGP data that are traditionally used in this context (e.g., RouteViews, RIPE). The reported numbers vary between 450K-500K and are only slightly larger than the 445K subnets we see in this one week. With respect to ASes, the results are very similar – the IXP "sees" traffic from some 42.8K actively routed ASes, where the ground truth for the number of actively routed ASes in the Internet in any given week is around 43K [5] and varies slightly with the used BGP dataset.

Lastly, to examine the visibility that this IXP has into the more commercial-oriented Internet, we next use the Web server-related component of the IXP's peering traffic (see Section 3.2.2). Table 3.1 shows that this IXP "sees" server-related traffic from some 1.5M IPs that can be unambiguously identified as Web server IPs. Unfortunately, we are not aware of any numbers that can be reliably considered as ground truth of all server IPs in the Internet in any given week. Even worse, available white papers or reports that purportedly provide this information are typically very cavalier about their definition of what they consider as "Web server" and hence cannot be taken at face value [77, 149].

To indirectly assess how the roughly 1.5M Web server IPs seen at this IXP stack up against the unknown number of Web server IPs Internet-wide, we use an essentially orthogonal dataset, namely the HTTP and DNS logs from a large European Tier-1 ISP that does not exchange traffic over the public switching infrastructure of our IXP. Applying the method as described in Section 2, we extract the Web server IPs from this ISP dataset and find that of the total number of server IPs that are "seen" by this ISP, only some 45K are not seen at the IXP. Importantly, for the server IPs seen both at the IXP and the ISP, those we identified as server IPs using the IXP-internal data are confirmed to be indeed server IPs when relying on the more detailed ISP dataset. In any case, mapping the 1.5M server IPs from the IXP to prefixes, ASes, and countries shows that this IXP "sees" server-traffic from some 17% of all actively routed prefixes, from about 50% of all actively routed ASes, and from about 80% of all the countries in the world.

Source Address Spoofing: We acknowledge that the sFlow data we used in this chapter may contain spoofed traffic. Spoofed traffic can bias our visibility of peering traffic at the IXP, in particular, our inferences on the numbers of visible ASes, subnets, and countries may be overstated. A single spoofed packet may add to any of these numbers. Also, the number of visible IPs (230M+) may be influenced. However, we assume that, e.g., DDoS attacks with spoofed packets would amount to many more IPs than those reported. Assume an attacker would forge the source IPs by using addresses from the entire address space. This would lead to around 2.6B visible IPs at the IXP, which corresponds to the number of total address space advertised in week 45. Moreover, the sFlow data is randomly sampled (1 out of 16K) which should lessen the impact of spoofed traffic. Finally, we believe that the prevalence of spoofed traffic is generally negligible. In a study on spoofed IP addresses, the authors show that the amount of spoofed traffic in the case of a large IXP is low [124]. Further steps to lessen the impact of spoofed traffic can be (i) filtering by unallocated address space using official allocation information from RIRs, (ii) filtering by unrouted address space using public routing information, or (iii) using customer cone information to remove traffic from an invalid source [124]. We believe that spoofed traffic should only have a negligible impact on the remaining analyses.

	rank	All IPs Country	Server IPs Country	All IPs Network	Server IPs Network
IPs	1	US	DE	Chinanet	Akamai
	2	DE	US	Vodafone/DE	1&1
	3	CN	RU	Free SAS	OVH
	4	RU	FR	Turk Telekom	Softlayer
	5	IT	GB	Telecom Italia	ThePlanet
	6	FR	CN	Liberty Global	Chinanet
	7	GB	NL	Vodafone/IT	HostEurope
	8	TR	CZ	Comnet	Strato
	9	UA	IT	Virgin Media	Webazilla
	10	JP	UA	Telefonica/DE	Plusserver
Traffic	1	DE	US	Akamai	Akamai
	2	US	DE	Google	Google
	3	RU	NL	Hetzner	Hetzner
	4	FR	RU	OVH	VKontakte
	5	GB	GB	VKontakte	Leaseweb
	6	CN	EU	Kabel Deu.	Limelight
	7	NL	FR	Leaseweb	OVH
	8	CZ	RO	Vodafone/DE	EdgeCast
	9	IT	UA	Unitymedia	Link11
	10	UA	CZ	Kyivstar	Kartina

Table 3.2: Top 10 contributors—week 45

3.3.2 On the IXP's dual role

Visuals such as Figure 3.3 illustrate that by using this IXP as a vantage point, we are able to see peering traffic from every country and corner of the world or from almost every AS and prefix that is publicly routed. However, such figures do not show whether or not certain countries or corners and ASes or prefixes are better visible than others in the sense that they are responsible for more traffic that is exchanged over the public switching fabric of the IXP. In particular, we would like to know whether the importance of the local role that this IXP plays for the larger geographic region within which it is situated is more or less recovered when considering the peering or server-related traffic that the IPs or server IPs are responsible for, respectively. To this end, we show in Table 3.2 the top-10 countries in terms of percentage of IP addresses (and associated traffic) and percentage of server IPs (and associated traffic). In addition, we show the top-10 networks. While the role of the IXP for the European region becomes more dominant when we change from peering to server-related traffic, there are still prominent signs of the IXP's global role, even with respect to the commercial Internet, and they reflect the relative importance of this IXP for countries such as USA, Russia, and China or ASes such as 20940 (Akamai), 15169 (Google), and 47541 (VKontakte).

For a somewhat simplified illustration of the IXP's dual role as a local as well as global player, we divide the set of all actively routed ASes into three disjoint sets, $A(L)$, $A(M)$, and $A(G)$. $A(L)$ consists of the member ASes of the IXP; $A(M)$ consists of all ASes that are distance 1 (measured in AS-hops) from a member AS; and $A(G)$ is the complement of $A(L) \cup A(M)$ and contains those ASes that are distance 2 or more from the member ASes. Intuitively, the set $A(L)$ captures the importance of the local role of the IXP, whereas the set $A(G)$ is more a reflection of the IXP's global role, with A(M) covering some middle ground. Table 3.3 shows the breakdown of the IPs, prefixes, and ASes for peering traffic and Web server-related traffic, respectively, for the three sets. It basically

		Member AS $A(L)$	Distance 1 $A(M)$	Distance > 1 $A(G)$
Peering Traffic	IPs	42.3%	45.0%	12.7%
	Prefixes	10.1%	34.1%	55.8%
	ASes	1.0%	48.9%	50.1%
	Traffic	67.3%	28.4%	4.3%
Server Traffic	IPs	52.9%	41.2%	5.9%
	Prefixes	17.2%	61.9%	20.9%
	ASes	2.2%	61.5%	36.3%
	Traffic	82.6%	17.35%	0.05%

Table 3.3: IXP as local yet global player—week 45

confirms our above observation that there is a general trend towards the set $A(L)$ as we move from IPs and the peering traffic they are responsible for to server IPs and their traffic. Note, while the relative importance of the IXP's local role over its global role with respect to the commercial Internet (i.e., server-related traffic) makes economic sense and is well-captured by this cartoon picture, in reality, there is potentially significant overlap between the sets $A(L)$, $A(M)$, and $A(G)$, e. g., due to remote peerings, IXP resellers, and non-European networks joining the IXP for purely economic reasons. But this is unlikely to invalidate our basic findings concerning the IXP's dual role.

3.3.3 On the IXP's "blind spots"

While the IXP "sees" traffic from much of the Internet, taking measurements exclusively at this single vantage point can tell us only so much about the network as a whole or its individual constituents. Hence, knowing what we can discern about the network with what sort of accuracy is as important as understanding what we cannot discern about it, and why.

We show in Section 3.3.1 how the use of an essentially orthogonal IXP-external dataset (i. e., the HTTP and DNS logs from the large European Tier-1 ISP) enables us to indirectly assess how the approximately 1.5M server IPs seen at the IXP in a given week compare to the unknown number of server IPs network-wide. In the following, we discuss additional examples where the use of IXP-external data, either in the form publicly available measurements, active or passive measurements, or proprietary information, enables us to check, validate, or refine what we can say with certainty when relying solely on IXP measurements.

To examine in more detail how the approximately 1.5M server IPs seen at the IXP in a given week compare to all server IPs in the Internet, we now use a more extensive combination of IXP-external measurements. To start, using the list of the top-1M Web sites available from `www.alexa.com` and based on the URIs retrieved from the limited payload part of the sampled frames at the IXP, we recover about 20% of all the second-level domains on Alexa's top-1M list of sites; this percentage increases to 63% if we consider only the top-10K list and to 80% for the top-1K. Note that many hostnames on these lists are dynamic and/or ephemeral. Next, to assess how many additional server IPs we can identify using the approximately 80% of domains we cannot recover using the URIs seen at the IXP, we rely on active measurements in the form of DNS queries to those uncovered domains using our set of 25K DNS resolvers across 12K ASes (see Section 2.3). From this pool of resolvers, we assign 100 randomly-selected resolvers to each URI. This results in approximately 600K server IPs, of which more then 360K are

already seen at the IXP and identified as servers.

To provide insight into the remaining 240K server IPs that are not seen as a server at the IXP, we classify them into four distinct categories. First, there are servers of CDNs that are hosted inside an AS and serve exclusively clients in that AS ("private clusters"). These servers reply only to resolvers of that AS for content that is delivered by the global footprint of those CDNs. Traffic to these servers should not be observable at the IXP as it should stay internal to the AS. Second, there are servers of CDNs or cloud/hosting providers that are located geographically far away from the IXP. If these networks have a global footprint and distribute content in a region-aware manner, it is unlikely that these server IPs are seen at the IXP. The third group includes servers that some ASes operate for the sole purpose of handling invalid URIs. Finally, the last category contains those servers of small organizations and/or universities in parts of the world that are geographically far away from the IXP. These IPs are typically not visible at the IXP. In terms of importance, the first two categories account for more than 40% of the 240K servers not seen at the IXP.

For a concrete example for illustrating "what we know we don't know", we consider Akamai. In our week-long IXP dataset, we observe some 28K server IPs for Akamai in 278 ASes (for details, see Section 3.5). However, Akamai publicly states that it operates some 100K servers in more than 1K ASes [44]. The reasons why we cannot see this ground truth relying solely on our IXP-internal data are twofold and mentioned above. First Akamai is known to operate "private clusters" in many third-party networks which are generally not visible outside those ASes and therefore cannot be detected at the IXP. Second, we cannot expect to uncover Akamai's footprint in regions that are geographically far away from the IXP, mainly because Akamai uses sophisticated mechanisms to localize traffic [121, 138]. Akamai's large footprint makes discovering all of its servers difficult, but by performing our own diligently chosen IXP-external active measurements [109] that utilize the URIs collected in the IXP and the open resolvers discussed in Section 3.2.3, we were able to discover about 100K servers in 700 ASes. Thus, even for a challenging case like Akamai, knowing what our IXP-internal data can and cannot tell us about its many servers and understanding the underlying reasons is feasible.

Regarding our assumption that server-related traffic is a good proxy for the commercial portion of Internet traffic, there are clearly components of this commercial traffic that are not visible at the IXP. For example, the recently introduced hybrid CDNs (e. g., Akamai's NetSession [40]) serve content by servers as well as by end users that have already downloaded part of the content. Since the connections between users are not based on a HTTP/HTTPS server-client architecture but are P2P-based, we may not see them at the IXP. However, while the traffic of these hybrid CDNs is increasing (e. g., the service is mainly used for large files such as software downloads), the overall volume is still very low [143].

Lastly, by the very definition of an IXP, any traffic that does not pass through the IXP via its public-facing switching infrastructure remains entirely invisible to us. For example, this includes all traffic that traverses the IXP over private peering links. IXPs keep the private peering infrastructure separate from its public peering platform, and we are not aware of any kind of estimates of the amount of private peering traffic handled by the IXPs.

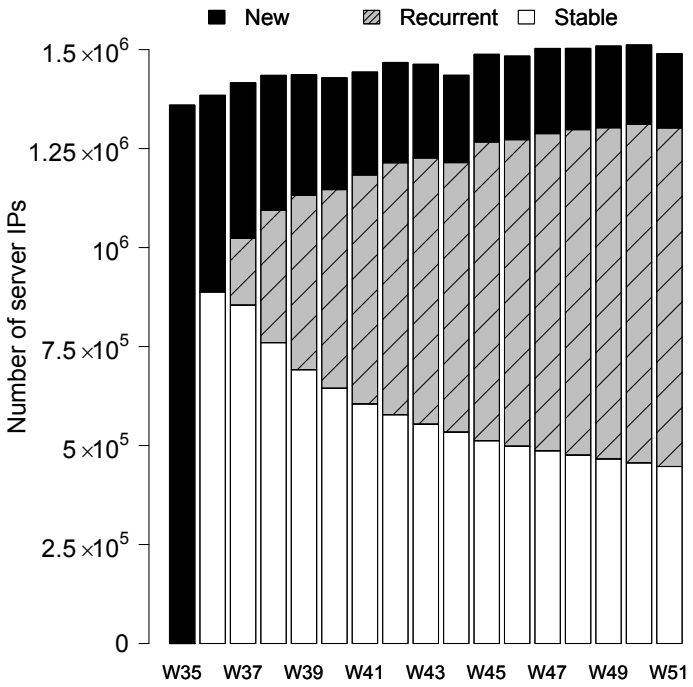


Figure 3.4: Churn of server IPs

3.4 Stable yet Changing

In this section, we report on a longitudinal analysis that covers 17 consecutive weeks and describes what using our large IXP as a vantage point through time enables us to say about the network as whole, about some of its constituents, and about the traffic that these constituents are responsible for.

3.4.1 Stability in the face of constant growth

Publicly available data shows that during 2012, this IXP has experienced significant growth, increasing the number of member ASes by 75 and seeing the average daily traffic volume grow by 0.1%. In terms of absolute numbers, we see in week 35 a total of 443 IXP member ASes sending an average daily traffic volume of 11.9 PB over the IXP's public-facing switching infrastructure. By week 51, the member count stood at 457, and the average traffic volume went up to 14.5 PB/day. For what follows, it is important to note that these newly added member ASes are typically regional and local ISPs or organizations and small companies outside of central Europe for which membership at this IXP makes economic sense. To contrast, all the major content providers, CDNs, Web hosting companies, eyeball ASes, and Tier-1 ISPs have been members at this IXP for some time, but may have seen upgrades to higher port speeds since the time they joined.

Given our interest in the commercial Internet and knowing (see Section 2) that the server-related traffic is more than 70% of the peering traffic seen at the IXP, we focus in the rest of this chapter on the server-related portion of the IXP traffic. The initial

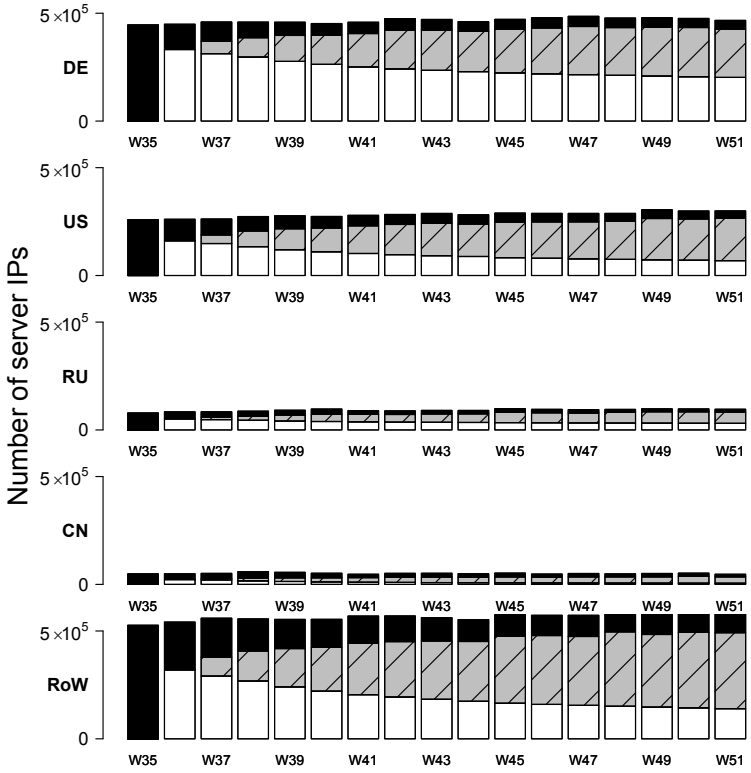


Figure 3.5: Churn of server IPs per region

set of findings from our longitudinal analysis paints an intriguingly stable picture of the commercial Internet as seen from our vantage point. In particular, analyzing in detail each of the 17 weekly snapshots shows that during every week, we see server-related traffic at this IXP from about 20K (i. e., about half of all) actively routed ASes, some 75K or approximately 15% of all actively routed prefixes, and from a pool of server IPs whose absolute size changes only so slightly but tends to increase in the long term.

This last property is illustrated in Figure 3.4 when focusing only on the absolute heights of the different bars that represent the total number of server IPs seen in a given week. When considering the full version of this figure, including the within-bar details, Figure 3.4 visualizes the weekly churn that is inherent in the server IPs seen at the IXPs. To explain, the first bar in Figure 3.4 shows the approximately 1.4M unique server IPs that we see in week 35. The next bar shows that same quantity for week 36, but splits it into two pieces. While the lower (white) piece reflects the portion of all week 36 server IPs that were already seen during week 35, the upper (black) piece represents the set of server IPs that were seen for the first time during week 36. Starting with week 37, we show for each week $n \in \{37, 38, \ldots, 51\}$ snapshot a bar that has three pieces stacked on top of one another. While the first (bottom, white) piece represents the server IPs that were seen at the IXP in each one of the week k snapshot $(k = 35, 36, \ldots, n)$, the second (grey-shaded) piece shows the server IPs that were seen at the IXP in at least one previous week k snapshot $(k = 35, 36, \ldots, n-1)$, but not in all; the third (top black) piece represents all server IPs that were seen at the IXP for the first time in week n.

A key take-away from Figure 3.4 is that there is a sizable pool of server IPs that is seen at the IXP during each and every week throughout the 17-week long measurement

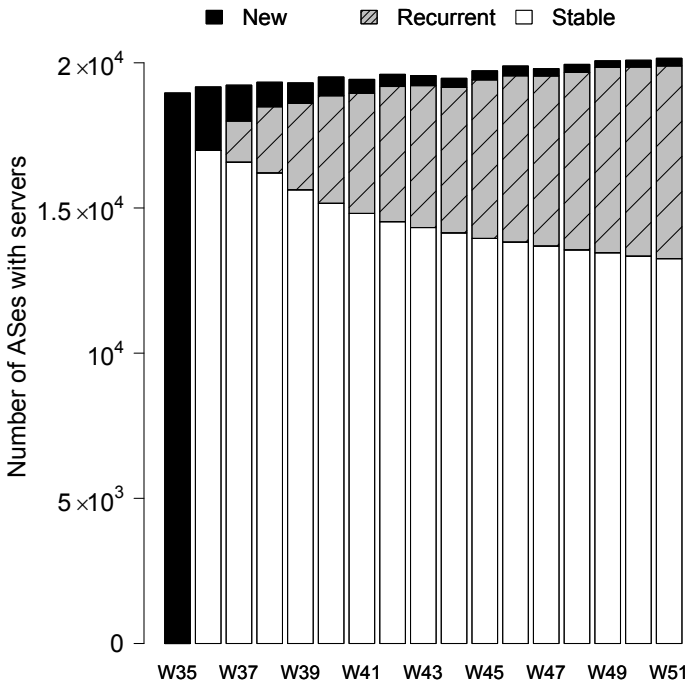


Figure 3.6: Churn of ASes with servers

period. In fact, this stable portion of server IPs that is seen at the IXP week-in and week-out is about 30% as can be seen by looking at the bottom (white) portion of the week 51 bar. Instead of requiring for a server IP to be seen in each and every week, we also consider a more relaxed notion of stability called recurrence. This recurrent pool of server IPs consists of all server IPs that, by week 51, have been seen at the IXP during at least one previous week (but not in each and every previous week), is represented by the grey-shaded portion of the week 51 bar, and consists of about 60% of all server IPs seen in week 51. Note that the number of server IPs seen for the first time in week n (top black portion) decreases over time and makes up just about 10% of all server IPs seen in week 51.

To look in more detail at the stable and recurrent pools of server IPs and examine their churn or evolution during the 17-week long measurement period, we rely on the GeoLite Country database [129] to geo-locate the server IPs to the country level and group them by geographic "region" as follows: DE, US, RU, CN, RoW (rest of world). Figure 3.5 is similar to Figure 3.4, but shows for each week the portions of IPs for each of these five regions and visualizes the make-up of these server IPs in the same way as we did in Figure 3.4. Note that the shown region-specific stable portions in week 51 add up to the 30% number observed in Figure 3.4, and similarly for the region-specific recurrent portions in week 51 (their sum adds up to the roughly 60% portion of the recurrent pool shown in Figure 3.4). Interestingly, while the stable pool for DE is consistently about half of the overall stable pool of server IPs seen at the IXP, that pool is vanishing small for CN, slightly larger for RU. This is yet another indication of the important role that this IXP plays for the European part of the Internet.

An even more intriguing aspect of stability is seen when we consider the server-related traffic that the server IPs that we see at the IXP are responsible for. For one, we find that

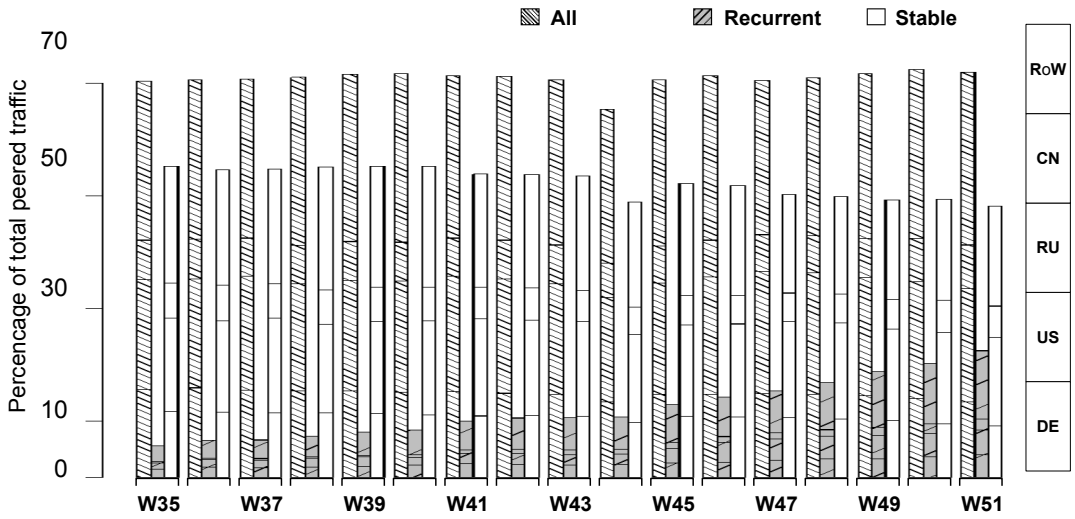


Figure 3.7: Churn of server traffic by region—weeks 35-51

the stable pool of server IPs is consistently contributing more than 60% of the server-related traffic. That is, of the server IPs that this IXP "sees" on a weekly basis, more than 30% of them are not only seen week after week, but they are also responsible for most of the server-related traffic that traverses the IXP each week. When considering the weekly recurrent pools of server IP (grey-shaded segments in Figure 3.4), their traffic portions keep increasing, but only to less than 30% of all server traffic. To examine the make-up of the server-related traffic attributed to the stable and recurrent pools of server IPs, respectively, Figure 3.7 shows for each week n three bars, each with five segments corresponding to the five regions considered earlier. The first bar is for the server-related traffic portion of all peering traffic that all server IPs see at the IXP in week n; the second bar reflects the server-related traffic portion in week n attributed to the recurrent pool of server IPs in that week, while the third bar shows the server-related traffic portion in week n that the stable pool of server IPs is responsible for. From Figure 3.7, we see that while the stable and recurrent pools of server IPs from China are basically invisible at the IXP in terms of their traffic, both US and Russia have the property that the stable pool of server IPs is responsible for much all the server-related traffic seen from those regions at the IXP.

In addition to examining the churn and evolution of the server IPs seen at the IXP, it is also instructive to study the temporal behavior of the subnets and ASes that the encountered server IPs map into. To illustrate, we only consider the ASes and show in Figure 3.6 for ASes what we depicted in Figure 3.5 for server IPs. The key difference between server IPs and ASes is that the stable pool of ASes represented by the white portion of the week 51 bar is about 70% compared to the 30% for the stable pool of server IPs. Thus, a majority of ASes with server IPs is seen at the IXP during each and every week, and the number of ASes that are seen for the first time becomes miniscule over time. In summary, the stable pool of server IPs (about 1/3 of all server IPs seen at the IXP) gives rise to a stable pool of ASes (about 2/3 of all ASes seen at the IXP and have server IPs) and is responsible for much of the server-related traffic seen at the IXP.

3.4.2 Changes in face of significant stability

One benefit of observing a significant amount of stability with respect to the server-related portion of the overall peering traffic seen at the IXP is that any weekly snapshot provides more or less the same information. At the same time, subsequent weekly snapshots that differ noticeably may be an indication of some change. Next, we briefly discuss a few examples of such changes that we can discern about the Internet as a whole and some of its individual constituents when we have the luxury to observe and measure the network at this IXP for a number of consecutive weeks.

The first example is motivated primarily by our ability described in Section 3.2.2 to specifically look for and identify HTTPS server IPs, but also by anecdotal evidence or company blogs [84, 101] that suggest that due to widespread security and privacy concerns, the use of HTTPS is steadily increasing. To examine this purported increase, we extract for each weekly snapshot all HTTPS server IPs and the traffic that they contribute. When comparing for each week the number of HTTPS server IPs relative to all server IPs seen in that week and the weekly traffic associated with HTTPS server IPs relative to all peering traffic, we indeed observe a small, yet steady increase, which confirms that the Internet landscape is gradually changing as far as the use of HTTPS is concerned.

For a different kind of example for using our IXP vantage point, we are interested in tracking the recently announced expansion of Netflix using Amazon's EC2 cloud service [38] into a number of Scandinavian countries [137]. To this end, we relied on publicly available data to obtain Amazon EC2's data center locations [45] and the corresponding IP ranges [46]. We then mined our 17 weeks worth of IXP data and observed for weeks 49, 50, and 51 a pronounced increase in the number of server IPs at Amazon EC2's Ireland location, the only data center of Amazon EC2 in Europe. This was accompanied by a significant (but still small in absolute terms) increase in Amazon EC2's traffic. All this suggests that it may be interesting to watch this traffic in the future, especially if the observed changes are in any way related to Netflix becoming available in Northern Europe towards the end of 2012.

Yet another example concerns the detection of regional or national events at this IXP. For example, considering in more detail week 44, which shows up as a clear dip in, say, Figure 3.4, we notice that this week coincides with Hurricane Sandy that had a major impact on the US East Coast region. To examine its impact, we use the IXP vantage point to discern this natural disaster from traffic that we see at the IXP from a particular Internet constituent, a major cloud provider. Using publicly available information about the cloud platform's data centers and corresponding IP ranges, we look in our data for the corresponding server IPs and find a total of about 14K. A detailed breakdown by data center location for weeks 43-45 shows a drastic reduction in the number of server IPs seen at the IXP from the US East Coast region, indicating that the platform of this major cloud provider faced serious problems in week 44, with traffic dropping close to zero. These problems made the news, and the example highlights how a geographical distant event such as a hurricane can be discerned from traffic measurements taken at this geographically distant IXP.

Lastly, we also mention that an IXP is an ideal location to monitor new players such as "resellers". Resellers are IXP member ASes, and their main business is to provide and facilitate access to the IXP for smaller companies that are typically far away geographically from the IXP. For IXPs, the emergence of resellers is beneficial as they extend the reach of the IXP into geographically distant regions and thereby the potential member-

ship base. For example, for a particular reseller at our IXP, we observed a doubling of the server IPs from 50K to 100K in four months, suggesting that this reseller has been quite successful in attracting new networks with significant server-based infrastructures as its customers.

3.5 Beyond the AS-level view

To illustrate the benefits and new opportunities for Internet measurements that arise from being able to use our IXP as a vantage point with very good visibility into the Internet, we describe in this section an approach for identifying server-based network infrastructures and classifying them by ownership. In the process, we report on a clear trend towards more heterogeneous networks and network interconnections, provide concrete examples, and discuss why and how this observed network heterogenization requires moving beyond the traditional AS-level view of the Internet.

3.5.1 Alternative grouping of server IPs

To this point, we have followed a very traditional approach for looking at our IXP data in the sense that we measured the IXP's visibility into the Internet in terms of the number of actively routed ASes or subnets seen at the IXP. However, there exist Internet players (e. g., CDN77, a recently launched low-cost no-commitment CDN; Rapidshare, a one-click hosting service; or certain meta-hosters that utilize multiple hosters) that are not ASes in the sense that they do not have an assigned ASN. Thus, as far as the traditional AS-level view of the Internet is concerned, these players are invisible, and the traffic that they are responsible for goes unnoticed, or worse misattributed to other Internet players. Yet, being commercial entities, these companies actively advertise their services, and in the process often publish the locations and IP addresses of their servers. This then suggests an alternative approach to assessing the IXP's ability to "see" the Internet as a whole—group servers according to the organization or company that has the administrative control over the servers and is responsible for distributing the content. While this approach is easy and works to perfection for companies like CDN77 that publish all their server IPs, the question is what to do if the server IPs are not known.

Accordingly, our primary goal is to start with the server IPs seen at the IXP and cluster them so that the servers in one and the same cluster are provably under the administrative control of the same organization or company. To this end, we rely in parts on methods described by Plonka et al. [141] for traffic and host profiling, Bermudez et al. [54] for discerning content and services, and Ager et al. [42] for inferring hosting infrastructures from the content they deliver. We also take advantage of different sets of meta-data obtained from assorted active measurement efforts or available by other means as discussed in Section 3.2.4. Recall that this meta-data includes for every server IP seen in the IXP data the corresponding URIs, the DNS information from active measurements, and, where available, the list of X.509 certificates retrieved via active measurements. In the rest of this section the reported numbers are for week 45.

The clustering proceeds in three steps. First, we focus on those server IPs for which we have a SOA resource record and consider a first category of clusters that have the property that all server IPs assigned to a given cluster have the IP and the content managed by the same authority. We identify those clusters by grouping all server IPs

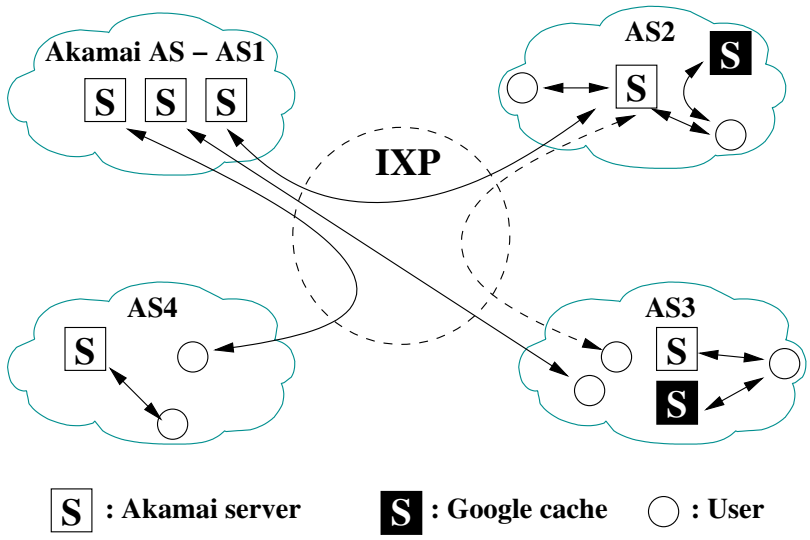


Figure 3.8: Heterogeneity of ASes and AS links

where the SOA of the hostname and the authority of the URI lead to the same entry. Prominent organizations that fall into this first category are Amazon and big players like Akamai and Google when they are located in their own ASes or when they are in third-party ASes but have assigned names to their own servers. 78.7 % of all our server IPs are clustered in this first step.

In a second step, we consider clusters with the property that for the server IPs in a given cluster, most of the server IPs and most of the content are managed by the same authority. This can happen if the SOA is outsourced (e. g., to a third-party DNS provider) and is common property among hosters and domains served by virtual servers. In these cases, to group server IPs, we rely on a majority vote among the SOA resource records, where the majority vote is by (i) the number of IPs and (ii) the size of the network footprint. This heuristic enables us to group some server IPs together with organizations inferred in the previous step and also applies to meta-hosters such as Hostica. 17.4 % of all our server IPs are clustered in this second step. Lastly, for the remaining 3.9 % of server IPs that have been seen in our IXP data and have not yet been clustered, we only have partial SOA information. This situation is quite common for parts of the server infrastructure of some large content providers and CDNs such as Akamai that have servers deployed deep inside ISPs. In this case, we apply the same heuristic as in the second step, but only rely on the available subset of information.

To validate our clustering that results from this three-step process, we manually compare the results by (1) checking against the coverage of the public IP ranges that some organizations advertise (see Section 3.4.2), (2) utilizing the information of certificates that point to applications and services, and (3) actively downloading either the front page (e. g., in the case of Google, it is always the search engine front page) or requested content that is delivered by a CDN (e. g., in the case of Akamai, any content is delivered by any of its servers [162]). Our method manages to correctly identify and group the servers of organizations with a small false-positive rate of less than 3%. Moreover, we observe that the false-positive rate decreases with increasing size of the network footprint. However, there are false-negatives in the sense that our methodology misses some servers due to the “blind spots” discussed in Section 3.3.3.

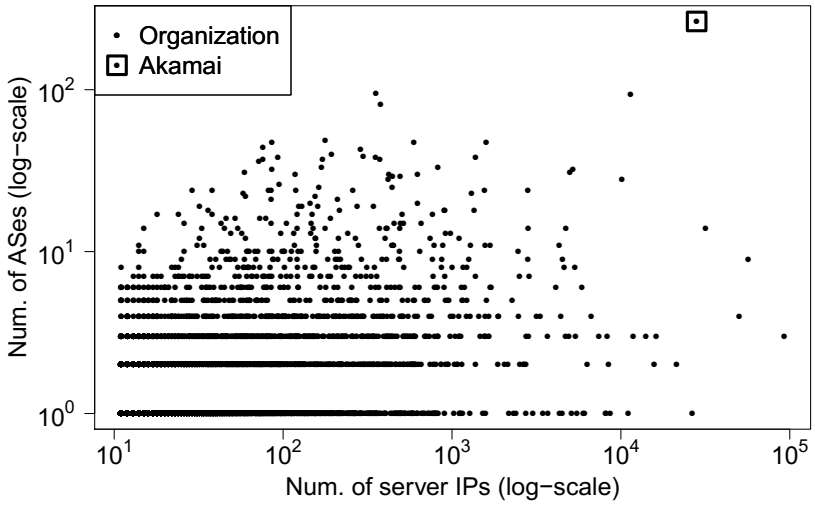


Figure 3.9: Scatter plot of number of server IPs vs. the number of ASes per organization

3.5.2 New reality (I): ASes are heterogeneous

Equipped with an approach for grouping server IPs by organizations, we examine next to what extent this grouping is orthogonal to the prevailing AS-level view of the Internet. The issues are succinctly illustrated in Figure 3.8 where we augment the traditional AS-level view (i. e., a number of different ASes exchanging traffic over (public) peering links at an IXP) with new features in the form of AS-internal details (i. e., the third-party servers that the ASes host). Note that while the traditional view that makes a tacit homogeneity assumption by abstracting away any AS-internal details may have been an adequate model for understanding some aspects of the Internet and the traffic it carries at some point in the past, things have changed, and we assert that the cartoon picture in Figure 3.8 captures more accurately the current Internet reality; that is, a trend towards distributed network infrastructures that are deployed and operated by today's commercial Internet players.

To quantify how much closer the cartoon Figure 3.8 is to reality than the traditional AS-level view, we apply our clustering approach to the 1.5M server IPs seen in our week 45 IXP data and obtain some 21K clusters, henceforth referred to organizations or companies. Among them are the well-known big players like Akamai with 28K active server IPs, Google with 11.5K server IPs, and several large hosters, each with more than 50K server IPs (e. g., AS92572 with 90K+ server IPs; AS56740 and AS50099, both with more than 50K server IPs). Indeed, of the 21K identified organizations, a total of 143 organizations are associated with more than 1000 server IPs and more than 6K organizations have more than 10 servers IPs. For the latter, Figure 3.9 shows a scatter plot of the number of server IPs per organization vs. the number of ASes that they cover. More precisely, every dot in the plot is an organization, and for a given organization, we show the number of its server IPs (x-axis) and the number of ASes that host servers from that organization (y-axis).[16] We observe that operating a highly diverse infrastructure is commonplace in today's Internet and is not limited to only the Internet's biggest players,

[16] While in a few isolated cases, the ASes that host servers from a given organization are part of that organization (e. g., see [58]), hand-checking the 143 organizations with more than 1000 servers confirmed that in almost all cases, these ASes are genuine third-party networks that are run and operated independently from the organization whose servers they host.

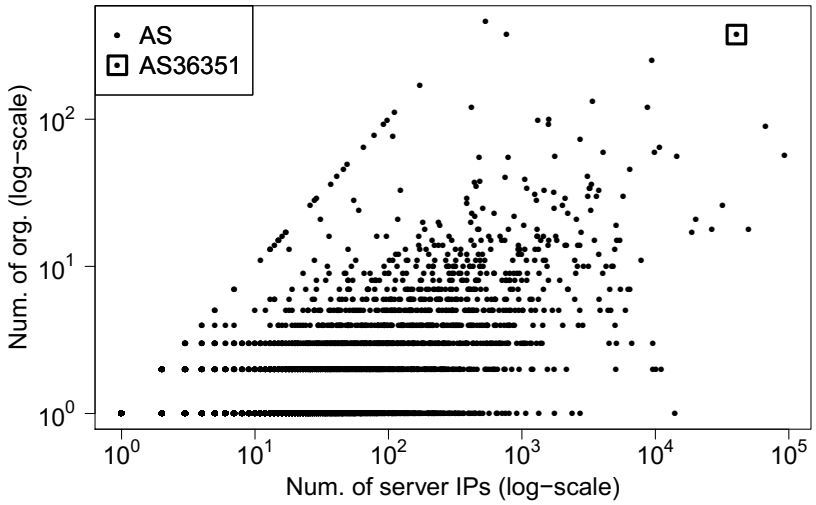


Figure 3.10: Scatter plot of number of organizations vs. the number of server IPs for each AS

but reporting on the bewildering array of scenarios we encountered when examining the extent of the different organizations and the networks they partner with is beyond the scope of this work.

The realization that many organizations operate a server infrastructure that is spread across many ASes implies that the complementary view must be equally bewildering in terms of diversity or heterogeneity. This view is captured in Figure 3.8 by focusing on, say AS1, and examining how many third-party networks host some of their servers inside that AS. Thus, yet another way to quantify how much closer that cartoon figure is to reality than the traditional AS-level view with its implicit homogeneity assumption concerning the administrative authority of servers hosted within an AS is shown in Figure 3.10. Each dot in this figure represents an AS, and the number of organizations a given AS hosts is given on the y-axis while the number of identified server IPs is shown on the x-axis. As before, the figure only shows organizations with more than 10 servers. We observe that many ASes host a sizable number of server IPs that belong to many organizations; there are more than 500 ASes that host servers from more than five organizations, and more than 200 ASes that support more than 10 organizations.

Indeed, this observation is again fully consistent with public announcements [1, 8] and content providers' efforts [7] to install their own brand of single-purpose CDNs inside various ISPs. The end effect of such developments is a clear trend towards more heterogeneous eyeball ISP networks by virtue of such ASes hosting more servers from an increasing number of interested third-party networks. In view of similar announcements from key companies such as Google [83, 59, 86], Amazon [2], or Facebook [6], the challenges of studying, leave alone controlling, such increasingly intertwined networks and traffic are quickly becoming daunting. As an example, consider a large Web hosting company (AS36351), for which we identified more than 40K server IPs belonging to a total more than 350 different organizations (highlighted in Figure 3.10 with a square).

3.5.3 New reality (II): Links are heterogeneous

In Section 3.5.2, we show that organizations take advantage of network diversity and purposefully spread their infrastructure across multiple networks. This development

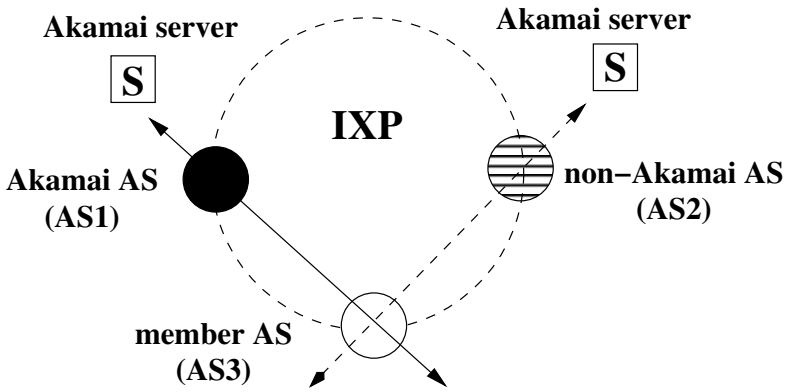


Figure 3.11: Observing traffic from a direct and a non direct link of Akamai

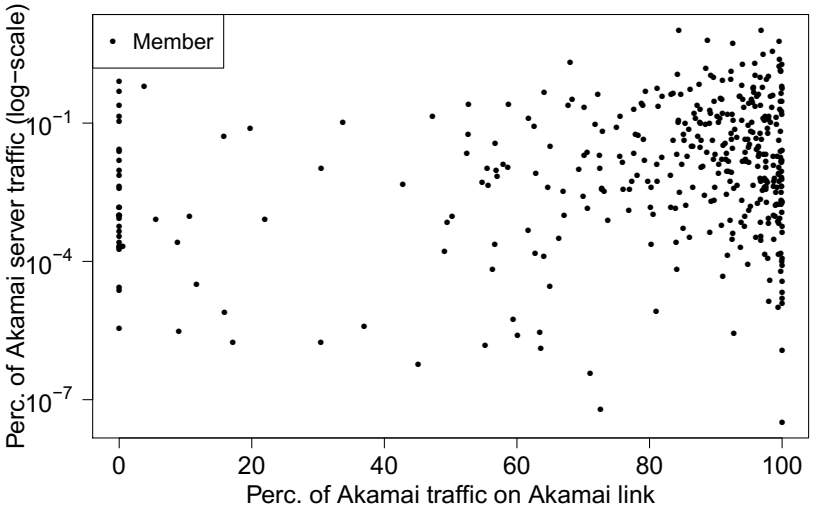


Figure 3.12: Perc. of Akamai traffic vs. perc. of Akamai traffic via direct link

creates very fluid and often transparent network boundaries, which in turn causes havoc when trying to attribute the right traffic to the right network. The issues are illustrated in the cartoon Figure 3.11. The figure shows the traditional AS-level perspective, whereby Akamai is a member AS (AS1) of this IXP, and so are a generic AS3 and another generic (non-Akamai) AS2, and the Akamai AS peers at this IXP with AS3 which, in turn, peers also with AS2. This traditional AS perspective is enriched with member-specific details that specify that there is an Akamai server behind/inside (non-Akamai) AS2 and behind/inside the Akamai AS. Note that in terms of the traditional AS-level view, the question of how much Akamai traffic is seen at this IXP is clear-cut and can be simply answered by measuring the traffic on the peering link between AS3 and the Akamai AS. However, when accounting for the fact that there is an Akamai server behind/inside the non-Akamai member AS2, answering that same question becomes more involved. It requires measuring the traffic on the (Akamai) peering link between AS3 and the Akamai AS as well as accounting for the Akamai traffic on the (non-Akamai) peering link between AS3 and (non-Akamai) AS2.

Clearly, for accurately attributing traffic to the responsible parties in today's network, the trend towards network heterogenization creates problems for the traditional AS-

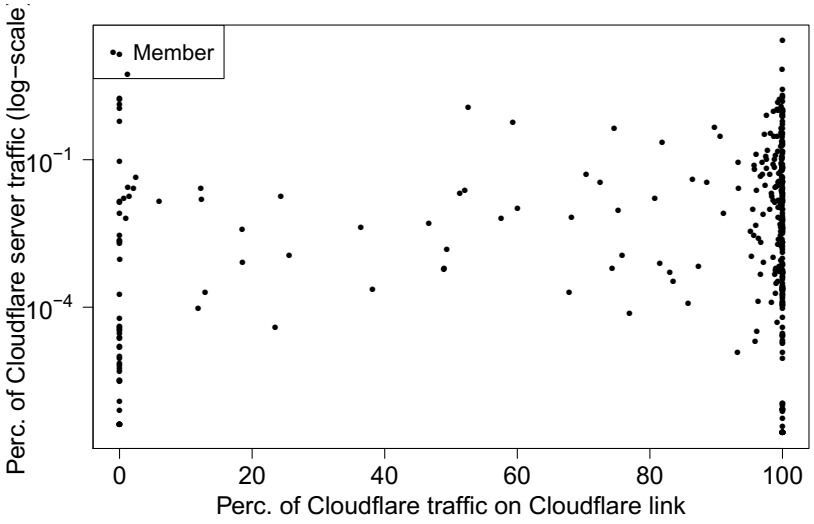


Figure 3.13: Perc. of CloudFlare traffic vs. perc. of CloudFlare traffic via direct link

level view of the Internet. To illustrate the extent of these problems, we show in Figure 3.12 what we observe at the IXP for Akamai. Recall that Akamai (AS20940) is a member of the IXP and peers with some 400 other member ASes. In the traditional view, accounting for Akamai traffic traversing the IXP simply means capturing the traffic on all the peering links between Akamai and those member ASes. Unfortunately, this simple view is no longer reflecting reality when Akamai servers are hosted inside or "behind" (non-Akamai) IXP member ASes. To capture this aspect, Figure 3.12 shows for each IXP member that peers with Akamai (indicated by a dot) the percentage of Akamai traffic on the direct peering link to Akamai (x-axis) vs. the percentage of total Akamai-server traffic for this member AS (y-axis). Under the traditional assumption, all dots would be stacked up at x=100, reflecting the fact that to account for Akamai-related traffic, all that is needed is to measure the Akamai peering links. However, with Akamai servers being massively deployed in third-party networks, including many of the other member ASes of the IXP, we observe that some members get all their Akamai-related traffic from ASes other than the (member) Akamai AS (x=0), even when that traffic is sizable (y>>0). Moreover, the scattering of dots across Figure 3.12 succinctly captures the diverse spread of traffic across the direct peering link vs. the other member links. In terms of numbers, Akamai sends 11.1% of its traffic not via its peering links with the member AS. Put differently, traffic from more than 15K out of the 28K Akamai servers that we identified in our IXP data is seen at the IXP via non-IXP member links to Akamai. The same holds true for other major CDNs but also for relatively new players such as CloudFlare. Figure 3.13 shows the same kind of plot as Figure 3.12 for CloudFlare. It demonstrates that despite adhering to very different business models (i. e., Akamai deploys servers inside ISPs vs. CloudFlare operates its own data centers), the two CDNs have similar usage patters as far as their peering links are concerned.

Looking beyond Akamai, we observe that different services from the same organization use their servers differently resulting in different usage patterns of the peering links. For example, for Amazon CloudFront, Amazon's "CDN part", almost all traffic is send via the IXP's Amazon links. However, for Amazon EC2, the "cloud part", a sizable fraction comes via other IXP peering links. We also noticed that for most cases where we see the use of the non-IXP member links, the percentage of traffic in those links increases during peak times. This may be due to reasons such as load balancing, performance

improvement, or cost savings. Lastly, how our view of the usage of the IXP's public peering links is impacted by private peerings that may be in place between member ASes of the IXP remains unexplored.

3.6 Discussion and Caveats

We are not the first to try and uncover the footprints of the infrastructures of commercial Internet players. One group of prior studies targets specific Internet companies (e. g., Akamai [163, 109, 157], Youtube [39, 94, 62], Netflix [38]), or one click hosters [50]). Other work is more concerned with inferring Web hosting infrastructures by relying on content only [42]. Our approach differs from these earlier works. For one, we rely on a unique vantage point in the form of one of the largest European IXPs to supply us with a weekly pool of some 230M IPs from which we diligently extract some 1.5M server IPs. Next, we rely exclusively on publicly available data[17] to group these servers by organizations that have the administrative authority over them and are responsible for their content. In doing so, we are inspired by earlier studies such as [54, 141]. Lastly, the methodology we develop for grouping servers by their organization is general in the sense that it applies equally well to content providers, CDNs, hosting companies, cloud infrastructure providers, eyeball ASes, or other Internet players.

The difference in perspective between the more traditional AS-level view of the Internet and our perspective that centers around organizations and companies and their heterogeneously deployed server-based infrastructure becomes evident when comparing our approach to the recent work by Cai et al. [58] on mapping ASes to organizations. For one, the starting point for [58] is the traditional AS-level view of the Internet, and two ASes are grouped into two different organizations if neither of the organizations is a subsidiary of the other (i. e., majority-owned by the other). While such a top-down ownership-based grouping of ASes captures one aspect of how ASes are inter-related, it is oblivious to how network infrastructures get used and deployed in today's Internet. In particular, while the method described in [58] may succeed in clustering all Akamai-owned ASes under the umbrella organization Akamai, the publicly known fact that Akamai has more than 100K servers deployed in hundreds of different third-party non-Akamai ASes [138] cannot be accounted for at all by that approach.

Our work relies critically on the sFlow records provided by one of the largest IXPs in Europe, and it can be argued that for many researchers, access to such data cannot be taken for granted. However, it is important to note that some of these largest IXPs in Europe generally welcome collaborations with researchers and are supportive of research efforts that make explicit use of their data (see for instance [47]). Once access to data collected from such unique and powerful vantage points is established, the opportunities for researchers are plentiful.

After presenting evidence for the kind of visibility into the Internet that comes with using one of these largest European IXPs as a vantage point, we highlight in this work some of the benefits that arises from having access to such a vantage point. However, despite its impressive capabilities, our IXP and the measurements it collects can only tell us so much about the network's "state", and many important issues remain concerning our knowledge about what exactly we can and cannot discern about the Internet

[17] Note that our use of the set of DNS resolvers from a large commercial CDN in Section 3.2.3 is a shortcut. A similar list could also have been obtained by relying on publicly available data only [157, 163, 109], e. g., via active scanning or from DNS logs.

as whole and its individual constituents. While we have identified a number of "blind spots", much remains to be done in terms of identifying and collecting IXP-external information that can be brought to the table for either checking, validating, or refining the findings obtained from the use of IXP-internal data only. The question of how to appropriately fuse selective IXP-external data with IXP-internal measurements to obtain a picture of the global network and its individual constituents that is unprecedented in terms of its accuracy, details, and insight looms as an important open research problem.

3.7 Chapter Summary

In this chapter, we studied the visibility of a large IXP and observed the traffic exchange among hundreds of networks.

By analyzing in detail the traffic that traverses the physical infrastructure of one of the largest IXPs in Europe, we provided evidence that the large European IXPs such as AMS-IX, DE-CIX, and LINX represent global Internet vantage points that "see" week-in and week-out traffic from hundreds of millions of IPs, from almost all routed prefixes and from all routed ASes, and from essentially every country around the world. We also illustrated these IXPs' dual role as a global and a local player within the Internet's ecosystem and caution that despite their outstanding visibility into the Internet, their use as global Internet vantage points comes with caveats (e. g., having "blind spots").

When concentrating on the Web server-related portion of the IXP traffic and performing a longitudinal analysis over a 17-week long period, we observed significant stability – of all the server IPs for which traffic is observed at the IXP during this 17-week period, some 30% are seen at the IXP week-in and week-out and are responsible for around 60% of all server-related traffic in each and every week. At the same time, the traffic seen at the IXP does exhibit differences from one week to the next, and we illustrated with some examples what different types of changes enable us to say about the network as a whole or some of its individual constituents. In this sense, the traffic seen at these global Internet vantage points can be used as complementary source of information for recent efforts analyzing Internet events such as large-scale outages due to censorship [76], natural disasters [68], etc.

To illustrate the kind of benefits that arise from having access to a global Internet vantage point in the form or our large European IXP, we confirmed a feature of today's Internet that is well-known among experts but remains largely under-reported in the networking research literature—a tendency of certain Internet players to either host servers from third-party networks within their own network infrastructures or deploy their own servers in strategically-chosen third-party ASes. More importantly, we presented a methodology for discovering an organization's servers, whether they are deployed within the organization's own AS (or ASes) or inside some third-party network's infrastructure, and used it to systematically assess the *extent* of this network heterogenization and study its impact on the usage of peering links at IXPs by these increasingly more heterogeneous member ASes. However, we want to stress that our AS-links usage-related findings are not IXP-specific (i. e., public peering links), but apply to any AS-link in the Internet, pointing towards serious challenges when trying to attribute the right traffic to the right party.

4

Traffic Asymmetries on Inter-Domain Links

In the previous chapter, we have shown how major Internet players, like CDNs or cloud providers, shape the traffic exchange at a large IXP. While the analyses so far focused on the heterogeneity of networks and inter-domain links induced by practices performed by these players, in this chapter take a deeper look at how the resulting complexities affect the traffic volume asymmetries. By using a large Tier-1 ISP we provide a first look into the contribution of traffic volumes in both directions between ASes.

4.1 Traffic Volume Asymmetries

The Internet is composed of tens of thousands of autonomous systems (ASes) connected to one another via a complex network of interconnections. Traffic exchanges across these interconnections are often subject to complex business relationships that heavily influence what traffic flows over the interconnection. Stated differently, policy (not performance) typically dictates the network path over which traffic flows in the Internet. Traffic patterns and flows, however, change over time: newer applications or use cases emerge that might significantly alter existing traffic patterns; business relationships between any two networks might change leading to a significant churn in the feasible paths between various networks. Of the various factors, the symmetry in network traffic, or lack thereof, has serious implications for business relationships; for instance, for settlement-free peering relationships between two adjacent ASes to be economically feasible, the ingress and egress traffic exchanged between the two networks across their peering links should be more or less balanced.

The notion of symmetry (or asymmetry), for the most part, has been used only in the

context of routing: path (or routing) symmetry captures whether the forward and backward path between any two endpoints in the network consist of the same sequence of intermediate ASes. There exists a rich body of prior work on characterizing path asymmetry [139, 103, 79] and identifying their root causes [139, 90, 159, 155]. Virtually all of these prior studies focus on only one notion of asymmetry—the routing asymmetric (or asymmetric in network paths). Paxson's study of the end-to-end routing behavior is one of the earliest known work to analyze routing asymmetry and discuss the root causes [139]. More studies have since looked at measuring routing asymmetry, e.g., [103] and [79]. Gao et al. [90] and Tangmunarunkit et al. [159] analyzed the impact of policy-based routing on asymmetry. A characterization of traffic asymmetry and an analysis of the interplay between routing and traffic asymmetry has remained largely unexplored. [114] is most relevant to our work; unlike us John et al., however, limit themselves to quantifying asymmetric (or symmetric) traffic over specific links and do not address the problem of quantifying the share of traffic over all observed AS-level paths.

The notion of network *traffic* asymmetry, however, has largely remained unexplored. Consider two ASes that exchange traffic; a consumer network sending small requests and a content network sending large responses. Naturally, one would expect the traffic profile of the consumer network to be inbound-heavy and that of the content network to be outbound-heavy, thus the ingress and egress traffic volumes are asymmetric. However, this scenario is simple and does not reflect the complexities in today's Internet ecosystem. Usually, ASes exchange traffic with hundreds of other diverse ASes; some exchange traffic directly, others via transit ASes. Thus, given an AS, its overall symmetry (or asymmetry) is coined by the sum of exchanged ingress and egress traffic with other ASes. Additionally, routing configurations can be very complex and lead to – sometimes unexpected – traffic exchange via multiple AS-links, e.g., due to local routing policies. Therefore, the symmetry (or asymmetry) of a specific AS is determined by the sum of traffic it sends or receives over one or more AS-links.

In this chapter, we take the first steps toward a high-level characterization of network traffic asymmetry by utilizing a large European ISP as a vantage point. We characterize the traffic profile of our vantage point and how it is coined by other ASes. Also, we attempt to show the interplay between routing asymmetry and traffic asymmetries. Specifically, we quantify what fraction of the ingress and egress traffic between the ISP and other ASes use the same AS-link and what fraction is affected by, e.g., routing asymmetry. We augment our characterization of traffic asymmetry by also discussing the contribution of hypergiants to the asymmetry. Our major findings in this chapter are:

1. From the perspective of our vantage point, the total observed traffic is mostly ingress while there is a variation of around 12% over a one-week study period. This asymmetry follows a diurnal cycle and reflects typical end-user behavior in the Internet. Our observations are consistent with regard to the type of our vantage point, i.e., an eyeball (and transit) ISP, and the type of applications, like video streaming, which dominate the application mix of that ISP. We find an inconsistency, however, in the variance of traffic asymmetries when looking at near ASes (those that are directly connected to the ISP) and far ASes (at least one AS-hop away from the ISP) individually. Moreover, peaks in traffic asymmetries are not always aligned, which can be, e.g., due to a difference in time zones between near and far ASes.
2. Using our AS-link-based traffic classification, we find an insignificant amount of unidirectional traffic. Most of the traffic flows bidirectionally, i.e., ingress and

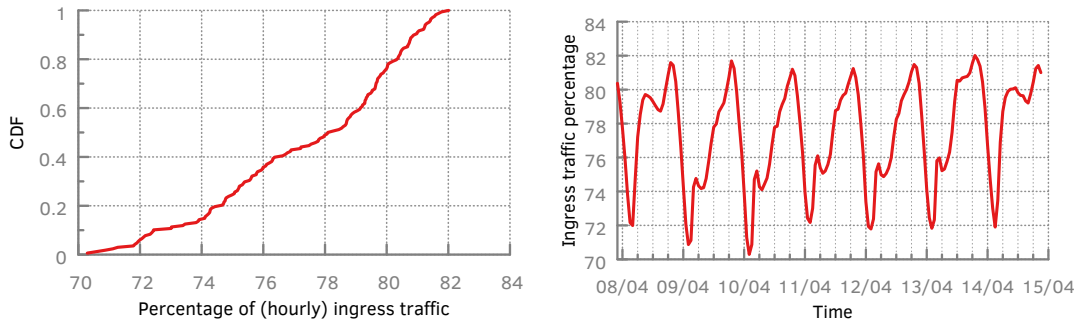


Figure 4.1: Traffic volume asymmetry. (a) Traffic is highly asymmetric, with ingress share being typically close to 78% of the overall traffic volume, and (b) variations in asymmetry (or ingress traffic share) exhibit clear diurnal patterns.

egress traffic uses the same AS-link. Moreover, a surprisingly high fraction of ingress and egress traffic in the mixed class use the same AS-link, indicating that the majority of traffic seen by the ISP flows bidirectionally. We find that bidirectional and mixed are highly correlated for near and far ASes: a drop in the bidirectional traffic is often accompanied by a peak in the mixed class, which can be due to practices employed by content providers. Looking at the asymmetry of individual ASes in the mixed class, we find that far ASes have a higher share of non-bidirectional traffic compared to near ASes.

3. To study the impact of hypergiants on traffic asymmetry, we first compare self-reported traffic ratios from PeeringDB with traffic ratios observed at the ISP, and find that *mostly outbound* or *outbound-heavy* (which include most of the hypergiants) are the most reliable labels. When removing traffic from hypergiants from our observations, we see a wide variety of traffic profiles for the remaining ASes, indicating that traffic from the these ASes is also highly asymmetric in nature.

The rest of the chapter is organized as follows. We present a motivating example in Section 4.2 and then follow it up with a discussion of the dataset. We explore some of the factors affecting traffic asymmetry (in Sections 4.4 and 4.5) and conclude with a discussion of the contribution of hypergiants in Section 4.6.

4.2 A Peek at Traffic Asymmetry

Traffic asymmetry typically refers to an imbalance (or skew) in the contribution of ingress and egress traffic to the total traffic. Figure 4.1 reveals the prevalence of this traffic asymmetry and its temporal characteristics. We calculate, for each hour of traffic observed by a large European Tier-1 ISP, the total ingress and egress traffic volumes. We plot, in Figure 4.1(a), the CDF of the percentage of ingress traffic in the total traffic, computed once for each hour of observed traffic over the one-week study period. The share of egress traffic is simply the complement of this CDF, and our choice of plotting the ingress share simply reflects the dominant characteristic of the dataset: traffic observed by the ISP is mostly ingress (with a median ingress traffic of roughly 78%) and the percentage of ingress traffic varies from (a very high minimum of) 70% to a maximum of 82%. The 12% of variation in ingress traffic share, although appearing smaller,

contributes to a substantial volume of traffic.

Figure 4.1(b) shows the temporal characteristics of traffic asymmetry, with time along the X-axis and hourly ingress traffic share along the Y-axis. The plot asserts that the asymmetry follows a diurnal cycle (consistently over the entire week) similar to that exhibited by end-user Internet traffic. Most of the end-user applications (e.g., Web browsing and video streaming) are highly asymmetric, i.e., size of requests are at least an order of magnitude smaller than that of responses. With such applications accounting for the majority of the network traffic, the traffic asymmetry we observe (along with its temporal characteristics) is not surprising. This high-level characterization of traffic asymmetry in Figure 4.1, nevertheless, leads to several follow-up questions. We enumerate a small subset of these questions that have huge implications for network planning and operations.

- Who contributes most to the traffic asymmetry: is the traffic originating from and destined for ASes that directly peer with the ISP more asymmetric than that compared to others?
- What is the interplay between routing and traffic asymmetry: with more than one path between two ASes, how does traffic flow over these paths? From the viewpoint of either endpoint, for instance, is traffic always ingressing and egressing over the same AS path?
- Are hypergiants the biggest contributor of traffic asymmetry? Are their traffic steering policies and mechanisms the primary reason behind ingress and egress traffic flowing over different paths?

In the remainder of this chapter, we deviate from the typical volume-based characterization of traffic asymmetry and attempt to answer the above questions.

4.3 Dataset: Perspective of a Tier-1 ISP

In an effort to provide as comprehensive a view as feasible on traffic volume asymmetry, we exploit the network perspective of a large Tier-1 European ISP; the ISP serves as both an eyeball network as well as a transit provider. We gather sampled (1 out of 1K) and anonymized (prefix-preserving) *NetFlow* statistics of traffic observed on the inter-domain links, i.e., ingress and egress traffic, of the ISP. We map the source and destination subnet (or prefix) of the NetFlow records into their respective autonomous system numbers (ASNs) using publicly available BGP data—RIPE RIS [10] and Route Views [12]. Lastly, we aggregate the traffic observed into hourly bins based on source and destination ASes, direct peer (or AS) through which the ISP sends or receives traffic, and the direction of traffic (i.e., ingress or egress). Table 4.1 describes the fields in the ISP's dataset, which covers a period of one week from April 7, 2017 (2300 hours UTC) through April 14, 2017. Over any particular day, the ISP observes traffic from $500+$ directly connected neighbor ASes (or simply, direct peers). Some of these direct peers ($\sim$20) neither originate any traffic nor is any traffic destined for them, but simply provide transit for other networks. The ISP also sends traffic to and receives traffic from 59K ASes that are not directly connected to the ISP. Lastly, for determining whether an AS is a direct peer we use ground truth on peering relationships (or connectivity information) from the ISP.

Although we focus on traffic that is either originating from or destined for the ISP, we note that the dataset includes a non-trivial amount of transit traffic (i.e., traffic for which

Table 4.1: Structure of the aggregated NetFlow statistics dataset gathered from the ISP.

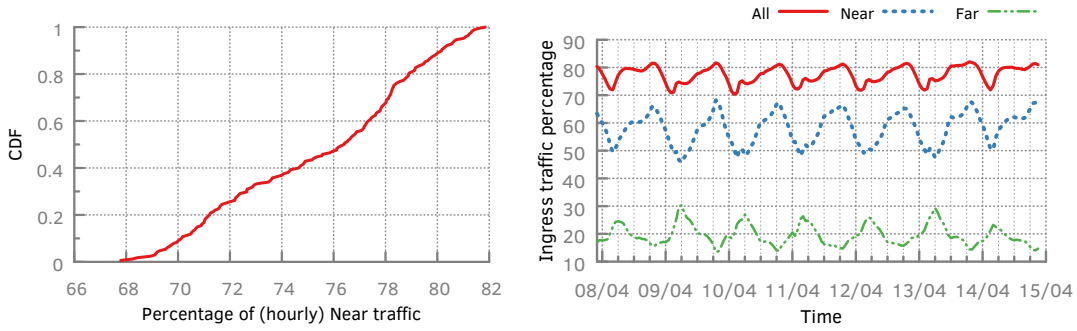
Field	***Description***
Timestamp	Time (in UTC) rounded to the nearest hour
Ingress?	'0' for egress traffic and '1' for ingress traffic
Transit?	'0' for non-transit traffic and '1' for transit traffic
src. or dst. prefix	Subnet of the traffic source or destination
src. or dst. ASN	ASN of the traffic source or destination
direct peer	Directly connected peer AS over which we receive or send the traffic
Traffic volume	Volume of traffic in bytes

the ISP is neither the source nor the destination). Transit traffic accounts (based on analyses during peak hours of traffic) for approximately 15% of ingress and egress traffic volumes. Since transit, by definition, is traffic traversing the network, the volume of transit affects the ingress and egress equally and does not affect the inferences based on volumetric analyses.

4.4 On Near and Far Neighbors

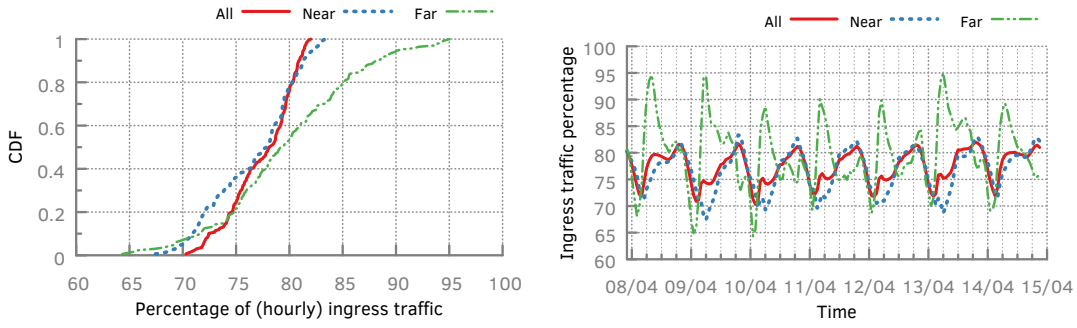
A first look at the overall asymmetry, in Figure 4.1, shows an inbound-heavy traffic profile, i.e., the sum of ingress traffic received from ASes outweighs the sum of egress traffic sent to ASes. In this section, we investigate the impact of distance in terms of AS-level hops. Our motivation behind this approach is that local policies may have a global impact, e.g., resulting in routing asymmetries, contributing to the overall traffic asymmetry of our vantage point. To evaluate this argument we classify the neighbor ASes of the ISP into two broad categories: *near* and *far* neighbors (or ASes). Near neighbors are ASes that are one hop away, i.e., directly connected to the ISP. Far neighbors, in contrast to near ASes, are ASes that are at least two hops away from the ISP, i.e., not directly connected to the ISP and having one or more transit ASes in the path between the concerned neighbor and the ISP. For convenience reasons, in the remainder of this chapter, we refer to near and far neighbors simply as neighbors, if possible. We tag traffic that is either originating from or destined for the near neighbors as *near traffic* and the remainder (i.e., the traffic originating from or destined for the far neighbors) as *far traffic*.

First, we look at the contribution of near and far traffic to the overall asymmetry. Most of the ISP's traffic falls under the near traffic category, as shown in Figure 4.2(a), with the near traffic typically (i.e., in the median) accounting for 76% of the total observed traffic. Adjusting the observed traffic asymmetry in the near and far categories by their respective share in the overall traffic reveals that the contribution of far traffic towards asymmetry is relatively small compared to that of near traffic, see Figure 4.2(b). Interestingly, the peaks of near and far traffic are not aligned. Despite both classes showing fluctuations in asymmetry, the misalignment of peaks or, more precisely, the respective alignment of peaks and valleys in near and far traffic during the first quarter of each day leads to a lower overall variation in asymmetry over the week. Thereby, the difference in overall ingress traffic contributed by near and far ASes can range between $\sim$20% at 6:00h and $\sim$50% at 18:00h every day. We conjecture that this misalignment can be due to timezone differences, or due to policy changes based on the time of day.



(a) (b)

Figure 4.2: Weighted share of near and far traffic. (a) Most of the ISP's traffic belongs to the near traffic category, and (b) weighting near and far traffic by their respective share in the total traffic shows that the overall asymmetry in traffic is mostly explained by the near traffic.



(a) (b)

Figure 4.3: Near vs Far traffic. (a) Traffic originating from as well as destined for far ASes show larger variation in asymmetry than that associated with the near neighbors. (b) Diurnal patterns in traffic asymmetry as a function of time, showing that variations in near and far traffic are not always correlated.

Next, we look at the asymmetries of near and far traffic independently. Figure 4.3 shows the CDF of the asymmetry in terms of percentage of ingress traffic in near and far traffic as well as the temporal patterns in the asymmetry.[18] The CDF (and variation) of asymmetry in near traffic, in Figure 4.3(a), more closely resembles that in the total traffic ("All") observed by the ISP. The far traffic, besides being different from "All", shows much higher variation in asymmetry than the near traffic: over the one-week period the span between the extremes is 30% for far traffic and is almost twice that of the near traffic. Figure 4.3(b) shows that these characteristics not necessarily follow the daily cycle. While the variations between peak and valley are rather constant in near traffic, they are more volatile in far traffic. We see a maximum ingress share of ∼95% traffic only at the 8th, 9th and 13th of April, and a minimum share of ∼65% traffic only at the 9th and 10th of April.

Overall, we conclude that, besides the traffic volumes, there are visible differences in how near and far ASes contribute to the asymmetric traffic profile of the ISP. Thus, we are motivated to separate these two in the following analyses.

[18]The percentage of ingress traffic in the total traffic ("All" line) is reproduced from Figure 4.1 for convenience.

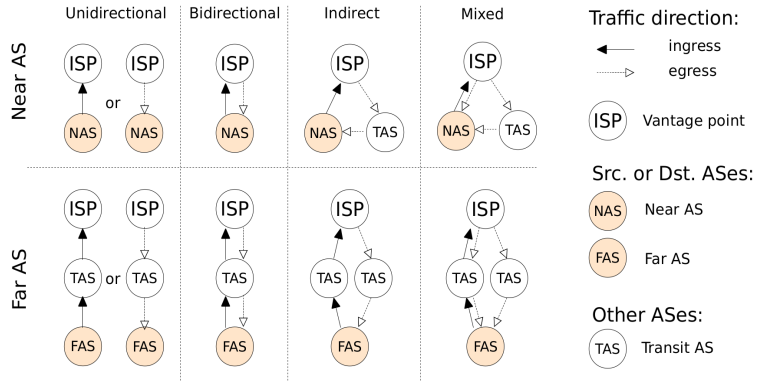


Figure 4.4: Categorizing traffic based on AS-links over which the traffic flows.

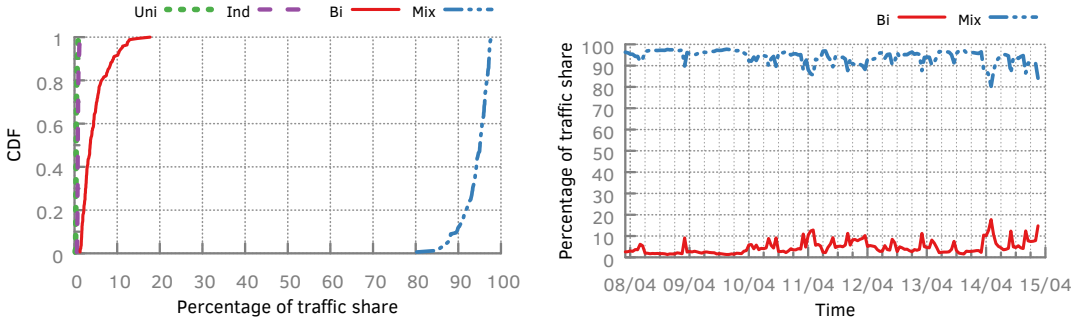


Figure 4.5: Traffic share per class over time. (a) Most of the observed traffic belongs to the bidirectional ("Bi") and mixed ("Mix") classes, and (b) traffic in these two classes appear highly correlated.

4.5 The Interplay between Routing & Traffic Asymmetry

So far, we have investigated the overall traffic asymmetry observable at the ISP and how it is shaped by asymmetries in near and far traffic. While the insights gained only reflect the overall traffic asymmetry, in this section we focus on the individual neighbors and how their traffic, whether symmetric or asymmetric, distributes over multiple AS-links, e.g., due to routing asymmetries. We classify the traffic into four classes based on the AS-link over which the traffic is observed, illustrated in Figure 4.4: "Unidirectional", "Bidirectional", "Indirect", and "Mixed". This classification captures the impact of routing on traffic asymmetry, and, more importantly, allows us to characterize how traffic flows over an available set of AS-level interconnections. For a given neighbor, the *unidirectional* category is where on a set of AS-links we observe traffic in just one direction, i.e., either ingress traffic from or egress traffic towards that neighbor. *Bidirectional* represents the category where on any AS-link we observe both, ingress and egress traffic, i.e., for a given neighbor, there is no AS-link with unidirectional traffic. In the case of *indirect*, we observe both, ingress and egress traffic, but never on the same AS-link. *Mixed* captures the case where we see at least one link with bidirectional traffic and at least one link with unidirectional traffic.

In the following, we characterize the distribution of traffic per asymmetry class. Fig-

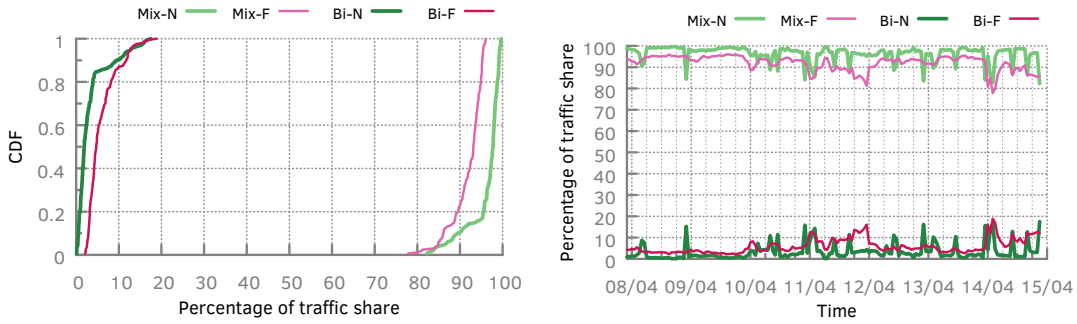


Figure 4.6: Traffic share per class for near and far ASes over time. (a) "Mix-F" and "Bi-F" account for a slightly lesser share of the overall (far) traffic (b) consistent correlation between mixed and bidirectional in near and far traffic.

ure 4.5(a) plots the CDF of the fraction of traffic in the different categories over the one-week period. There is little volume of traffic in the unidirectional ("Uni") and indirect ("Ind") categories; most of the traffic falls under the bidirectional ("Bi") and mixed ("Mix") categories. More specifically, over the one-week period, the traffic share of bidirectional traffic ranges from less than 1% to ∼18%, and the share of mixed ranges from ∼80% to ∼98%. Consequently, we ignore the unidirectional and indirect categories and plot the percentage of traffic contributed by the other two classes as a function of time, in Figure 4.5(b). The figure shows that bidirectional and mixed classes are correlated: a drop in the mixed ratio is often accompanied by a peak in the bidirectional ratio. Supported by the two figures, we highlight that the peaks are rather the exception than the regular case: 80% of the time the share of bidirectional and mixed traffic is <5% and >95%, respectively.

Next, we combine these four traffic classes with the previous classification of near and far traffic, and we plot the fraction of bidirectional and mixed classes in the overall near and far traffic categories. The CDFs of traffic shares in the various classes, in Figure 4.6(a), shows that the mixed and bidirectional categories of far traffic ("Mix-F" and "Bi-F", respectively) account for a slightly lesser share of the overall (far) traffic compared to that in the near traffic case. In fact, far traffic exhibits a higher, however, still negligible share unidirectional and indirect traffic compared to near traffic. The plot of the traffic shares against time, in Figure 4.6(b), however, shows that the similar behavior of mixed and bidirectional classes is consistent across both near and far traffic cases. We conclude that, even if routing asymmetry, induced by the distance between the ISP and far neighbors exist, it does not affect the behavior of mixed and bidirectional traffic. Moreover, some peaks in near and far traffic appear to be in sync, e.g., during midnight. This can be due to video-on-demand providers to improve the distribution of traffic over multiple (near and far) upstream providers.

From the observations so far we see that up to 18% of the traffic flows bidirectionally, i.e., we observe ingress and egress traffic on the same AS-link. Also, a vast majority (up to 98%) of the traffic flows in a mixed fashion. Thereby, a subset of the mixed traffic flows bidirectionally as well, but additionally includes some amount unidirectional traffic. In the following, we take a closer look at the composition of mixed traffic. In particular, we investigate the share of bidirectional traffic over the one-week measurement period. Figure 4.7(a) shows the CDF of the bidirectional portion of the mixed traffic. Over the one-week measurement period, a surprisingly high fraction of the ingress and

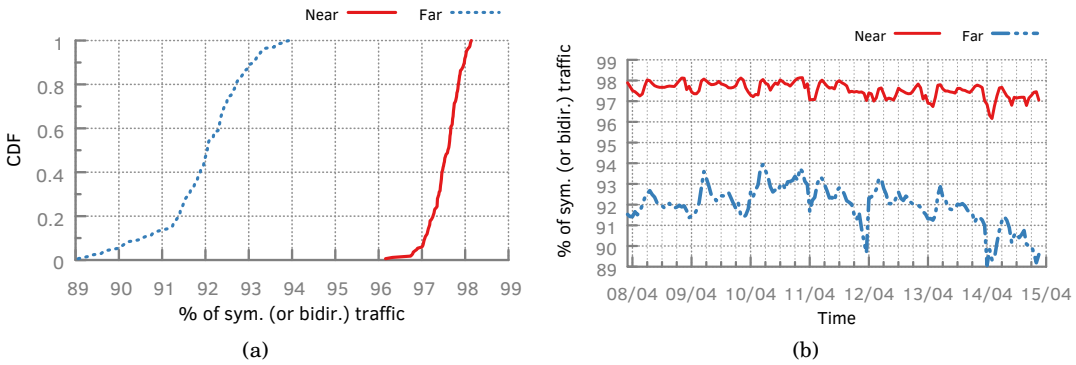


Figure 4.7: Share of bidirectional traffic in mixed class (a) high fraction of bidirectional traffic in the mixed class (b) timeline of the fraction of bidirectional traffic in the mixed class.

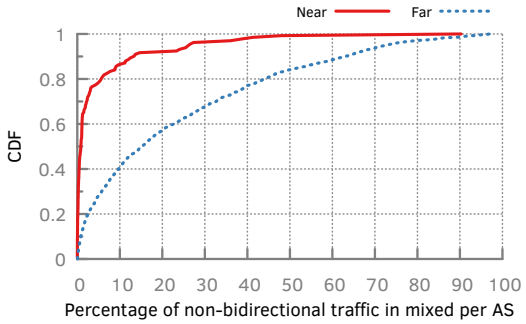
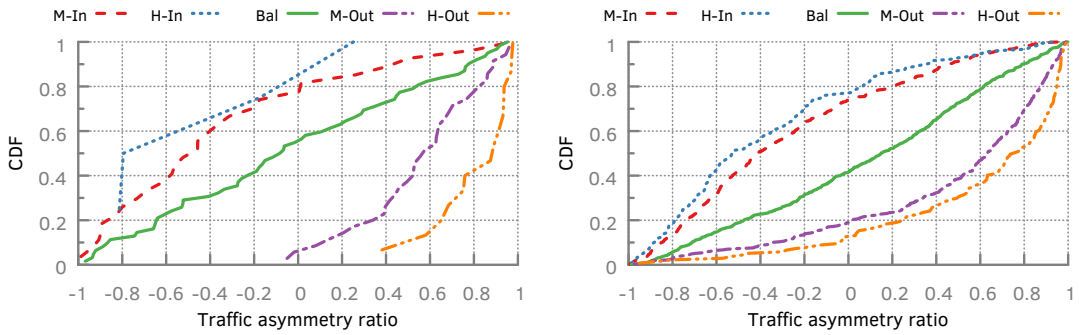


Figure 4.8: CDF over non-bidirectional (unidirectional) traffic ratio in mixed class per near and far AS, aggregated over 24 hours.

egress traffic in the mixed class flows over the same AS-link. In particular, the share of bidirectional in near mixed traffic is between 97% and 98%, in ∼80% of the time, while in far mixed traffic the variation is between 89% and 94%. Looking at the share as a function of time, in Figure 4.7(b), we see that the share of bidirectional traffic in near is more stable than in far. The drop towards the weekend (Saturday, 15th of April) in far traffic (and slightly in near) can be due to, e.g., cache updates induced by an increased usage of end-users.

Finally, we look at the share of bidirectional traffic of the individual near and far ASes in the mixed class. Thereby, we consider only one day (April 9, 2017) of data. We note that the results are consistent throughout the one-week measurement period. Figure 4.8 shows a CDF over non-bidirectional (i.e., unidirectional) share of the overall mixed traffic per AS. We observe that many far ASes have a higher share of non-bidirectional traffic compared to near ASes. More specifically, while 50% of far ASes have a share of roughly 15% or less, it is 1% or less for near ASes (in case of the 90% quantile it is <65% for far and <15% for near ASes).

Overall, we summarize that the majority of the traffic seen by the ISP is bidirectional, i.e., ingress and egress traffic flows over the same AS-link. Moreover, regarding individual ASes, we show that the share of non-bidirectional traffic in mixed is usually higher for far ASes than near ASes.



(a) Near ASes (b) Far ASes

Figure 4.9: Comparison of empirically observed asymmetry in (a) near and (b) far traffic with that of self-reported "Traffic Ratio" labels from PeeringDB; ratios less than zero imply ingress heavy traffic. Among the self-reported labels, "balanced" seems the most misleading.

4.6 The Role of Hypergiants

Hypergiants are one of the key entities of the Internet ecosystem, largely because a significantly large fraction of the Internet's traffic is associated with content (e.g., Netflix, Youtube, and Facebook) or cloud hypergiants (e.g., Amazon, Google, Akamai, and Cloudflare). With the volume of traffic that hypergiants (such as CDNs) exchange with other networks [71] and their well-known traffic steering policies [168, 150], we investigate the role of hypergiants in our observation of bidirectional and mixed traffic. Since there is no clear definition for what constitutes a hypergiant, we follow up with a recent work that attempts to characterize hypergiants using publicly available data from PeeringDB [55]. Specifically, we use the 15 hypergiants (refer Table 2 in [55]) that the authors identified, remove the traffic associated with them from our dataset, and quantify the extent to which the remaining ASes (i.e., the non-hypergiants) contribute to traffic asymmetry. To address the concerns of a characterization based only on PeeringDB, we first evaluate the accuracy of the self-reported traffic ratios (or "traffic profiles" according to Bottger et al. [55]) against our empirical observations.

4.6.1 On the Accuracy of Traffic Ratios in PeeringDB

PeeringDB [9] is a well-known, publicly available dataset to the networking community and many research efforts have relied on this dataset, e.g., for mapping the physical interconnections in the Internet [97, 133, 52], IP geolocation [63], understanding the business relationships between networks [126], characterizing autonomous systems [55], and detecting infrastructure issues [95]. Since we utilize the hypergiants identified by Bottger et al. [55], we focus on the "Traffic Ratio" field that they utilize for their classification. As with most data in PeeringDB, this field is a self-reported measure of network traffic asymmetry and its accuracy might vary widely. This *qualitative* measure takes one of six different labels—"Not Disclosed", "Heavy Outbound", "Mostly Inbound", "Balanced", "Mostly Inbound", and "Heavy Inbound". In this section, we compare our empirical analyses to these categorical values for estimating the accuracy of the field.

We look up the ASes in our traffic dataset (i.e., ASes which sent traffic to or received traffic from the ISP) against PeeringDB to map each AS to its self-reported traffic ratio. We

group these ASes by the traffic ratio field, one group for each of the five different labels, and compute the relative difference between ingress and egress traffic volumes. These relative differences capture the asymmetry in traffic ranging from -1 to 1, whereby -1 represents ASes with only ingress traffic and 1 represents ASes with only egress traffic. Everything in between captures a mixture of ingress and egress traffic. For the purpose of providing a measure for comparing the relative differences with the traffic rations from PeeringDB, we consider values within the range of -1 and -0.2 to indicate ingress-heavy ASes, and values within 0.2 and 1 to indicate egress-heavy ASes. All ASes with values between -0.2 and 0.2 are suggestive of symmetric or balanced ASes. Figure 4.9 plots the CDF of traffic (asymmetry) ratios of the different ASes grouped by the traffic ratio labels and further divided based on whether it is a near or far neighbor. For this analysis, we consider only one day (April 9, 2017) of data from our week-long study period. Our inferences, however, are consistent across the entire study period.

Per Figure 4.9(a), most near ASes (nearly 70%) in the "Mostly Inbound" ("M-In") and "Heavy Inbound" ("H-In") categories are indeed ingress-heavy ASes, with ingress traffic being at least 20% of the total traffic. The remaining 30% in the two categories "H-In" and "M-In" do not fulfill our above definition for ingress-heavy ASes as their asymmetry ratio ranges somewhere between -0.2 and 1. The "Mostly Outbound" ("M-Out") and "Heavy Outbound" ("H-Out") labels, in contrast to the inbound-related labels, tend to be much more *reliable*: nearly all of the ASes in these two labels are egress-heavy, thus matching the advertised traffic profiles. The "Balanced" ("Bal") category appears to be the most *unreliable*; only approximately 20% of the ASes with the self-reported balanced traffic ratio are indeed balanced according to empirical observations. Our observations are similar in the case of far ASes except for two differences. The "M-Out" and "H-Out" labels appear less reliable for far ASes than for near ASes, with around 20% of the ASes contradicting the labels (i.e., they are balanced or inbound heavy). Similar applies to the labels "M-In" and "H-In". The reliability of the "Bal" label remains the same.

We note that our definition of traffic ratio ranges, e.g, -0.2 and 0.2 for balanced ASes, may deviate from the notion of the individual network operators, and may also be impaired by the visibility of our vantage point. However, our focus in this section is on hypergiants which are typically heavy-outbound. As we have shown, the visibility of our ISP with regard to "M-Out" and "H-Out" labels seems not to be affected. Moreover, of the 15 hypergiants identified in [55], 13 are near ASes and contribute 75% of the total near traffic; the remaining 2 in far ASes account for only 1.5% of the far traffic.

4.6.2 Hypergiants & Asymmetry

In the following, we investigate the impact of hypergiants on traffic asymmetry and on the use of multiple AS-links by looking at non-hypergiant ASes. Of the 15 hypergiants identified in [55], only two are associated with the "Balanced" traffic ratio; the remaining 13 are associated with either "Mostly Outbound" or "Heavy Outbound". That the "Balanced" label is not as reliable characterization as the latter two should not affect our inferences since we remove all the traffic associated with these 15 hypergiants from our dataset for the rest of the analyses.

We compute, for each remaining AS (i.e., any non-hypergiant), the asymmetry in the traffic associated with that AS. We further separate the ASes into near and far neighbors (or ASes) as done before. Per Figure 4.10, the CDF of asymmetry in traffic volumes for these non-hypergiants reveals a wide spectrum of traffic profiles: The distribution of the ASes on this spectrum appears roughly uniform, with approximately 41% of the near

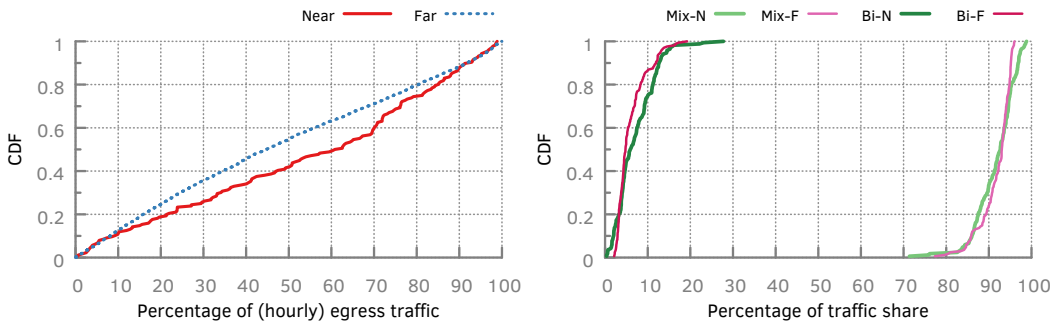


Figure 4.10: (a) traffic from non-hypergiant ASes is highly asymmetric in nature (b) fraction of mixed and bidirectional classes of non-hypergiant traffic remains largely unaffected.

ASes being ingress-heavy; far ASes are even more uniformly distributed than the near ASes, with nearly 52% of the ASes being ingress-heavy. We conclude that traffic from these ASes is highly asymmetric in nature. Interestingly, due to the uniform distribution, traffic the overall traffic profile of the ISP is surprisingly balanced, i.e., ingress and egress traffic volumes are similar. In other words, hypergiants with their heavy-egress traffic profiles appear to be the main reason for the ingress-heavy traffic profile of the ISP.

Finally, in Figure 4.10(b) we plot the fraction of bidirectional and mixed classes in the near and far traffic categories excluding traffic from hypergiants. Compared to Figure 4.6(a) which contain all the far and near traffic, it shows that even after removing the traffic associated with hypergiants the CDF of the fraction of mixed and bidirectional classes of far traffic remains largely unaffected. Mixed and bidirectional classes in near traffic, on the other hand, exhibit a slight convergence in terms of traffic volumes and resembles those of far traffic. We conclude that even without hypergiants we observe a similar distribution of mixed and bidirectional traffic. Thus, hypergiants are not the sole reason behind the existence of mixed traffic.

4.7 Limitations

The characterization presented in this chapter has three broad limitations. First, in analyzing the interplay between routing and path asymmetry, we focused only on the AS-level path (or interconnections) and could not pursue the more complex task of tracking traffic over individual links. Stated differently, even if we observe traffic to be traversing the same AS-path (i.e., belonging to the "bidirectional" category) the paths could be different (or asymmetric) at the link (or layer-2) level. Gathering the ground truth of the ISP's topology and accurately tracking the traffic over various link-level paths are, however, non-trivial challenges. Second, the role of hypergiants, in particular, the effect of their traffic steering policies on the observed asymmetry, deserves a more detailed look than presented here. Third, the Internet is a highly heterogeneous ecosystem and, unsurprisingly, analyzing the complex traffic dynamics from a single vantage point does not inspire much confidence.

Source Address Spoofing: We acknowledge that the NetFlow data we used in this chapter may contain spoofed traffic. Spoofed traffic can contribute to our volumetric

traffic analysis and bias our AS-link-based traffic classification. However, we believe that it should have a negligible impact on our overall results. The NetFlow data is sampled (1 out of 1K) which should lessen the impact of spoofed traffic. As discussed already in Section 3.3, we believe that the prevalence of spoofed traffic is generally negligible [124]. Remind that we used public routing information to map prefixes to AS numbers, and thereby removed all unrouted traffic, see Section 4.3, which is one step to deal with spoofed traffic. Further steps to lessen the impact of spoofed traffic can be filtering unallocated prefixes using official allocation data from RIRs, or using customer cone information to remove traffic from an invalid source [124].

4.8 Chapter Summary

In this chapter, we leveraged the vantage point of a large Tier-1 ISP to characterize network traffic asymmetry and investigated its susceptibility to routing asymmetries and traffic steering by hypergiants. Therefore, we collected one week of NetFlow statistics and assigned hourly aggregated traffic volumes in both directions, ingress and egress, to the corresponding ASes.

As a motivating example, we have provided a high-level characterization of traffic volume asymmetry from the perspective of our vantage point. We showed that most of the traffic is ingress and that it varies by up to 12% over the week. We then evaluated the role of routing in traffic asymmetry. Therefore, we classified the ASes into two broad categories, i.e., near and far ASes. Accordingly, we tagged the corresponding traffic as near and far traffic and analyzed temporal patterns in the asymmetry as well variations. We showed that, throughout the week, the variation of ingress traffic for far ASes is almost 30%, almost twice as much as the variation for near ASes. Thereby, the diurnal cycles of both, near and far traffic, are not aligned. This might be due to timezone differences or traffic steering involving CDN cache updates after peak hours.

To better understand the use of alternative paths, e.g., due to routing asymmetries, and its impact on traffic asymmetry we developed a method to classify traffic asymmetry. Thereby, we grouped the traffic into four classes reflecting the combination of traffic direction (ingress and egress) and the involved AS-level paths: *unidirectional* (single direction), *indirect* (both directions, each on a different paths), *bidirectional* (both directions on same path), and *mixed*, which is bidirectional traffic plus some unidirectional traffic. We combined these four traffic classes with the previous classification of near and far traffic. We attributed approximately 98% of the traffic to the bidirectional and mixed class. We further analyzed the composition of traffic in the mixed class and found that most of it flows bidirectionally; there is just a small amount of traffic in the unidirectional class. Far ASes, however, show a slightly lower share of bidirectional traffic in the mixed class, compared to near ASes. Assuming that alternative paths, e.g., unidirectional or indirect, are caused by routing asymmetries, we conclude that they barely affect ingress and egress traffic.

In order to investigate the impact of hypergiants and traffic steering practices on traffic asymmetries, we made use of publicly available data from PeeringDB. First, we checked the compliance of reported traffic profiles, e.g., *inbound-heavy*, *mostly inbound*, *balanced*, *mostly outbound*, and *outbound-heavy*, with the actual traffic volume ratio observed at the ISP. We then showed that almost all ASes from the outbound-heavy class, which include most of the hypergiants identified by [55], match our observations. Next, we removed all traffic associated with hypergiants and computed, for each remaining non-hypergiant AS, the asymmetry in the traffic. Our results indicate that traffic associated with most of these ASes is highly asymmetric in nature, but in sum, they make up a balanced ingress and egress ratio. Indeed, the hypergiants are significantly contributing to the overall ingress-heavy traffic profile of the ISP. On the other hand, we found that hypergiants have almost no influence on the distribution of traffic asymmetry classes.

5

Prefix Delegations via BGP

In the previous chapters we have shown how Internet players, like ISPs and CDNs make use of network diversity by deploying their server based infrastructures in multiple networks. The resulting complex traffic patterns, measured at a large IXP, provide evidence for a continuous heterogenization of networks and their inter-domain links. Moreoever, measuring traffic asymmetries on inter-domains links at a large ISP shows that ingress and egress points for traffic of a given AS can differ significantly.

In this chapter we provide a detailed analysis of the underlying routing mechanisms that involve address space deaggregation in order to achieve sophisticated traffic engineering. In particular, the focus of our analysis involves the delegation of more specific address blocks. Therefore, we make use of a globally distributed vantage point, namely BGP collectors.

5.1 Understanding BGP Prefix Delegations

Today's heavy use of deaggregation—by some considered abuse—renders the routing table more and more un-aggregatable. At the same time it shows the lack of alternative means to satisfy the needs of today's Internet routing system. A common explanation/observation for deaggregated prefixes leaking into the routing system is the delegation of PA prefixes to a multi-homed AS. Even though the PA prefix can be aggregated by the delegating provider it cannot be aggregated by other providers. Thus, the prefix adds to the routing table. In this work we focus on such prefix delegations by mining ten years of publicly available BGP data. Using traffic data from one of the largest European *Internet exchange points* (IXP), we find that more than 14% of its traffic is originating from delegated prefixes while 5% of the traffic is addressed to them. To better understand prefix delegations, we subclassify delegated prefixes into four categories

based on the AS path of prefix announcements. We then use large-scale traceroute measurements to quantify the impact on the actual traffic flow. To our surprise, we find a variety of prefix delegations including from-customer-to-provider or delegations among ASes that have no apparent topological relation.

This chapter contributes to the understanding of the global routing table growth, a scalability factor since ever. In particular, our analyses focus on the mostly neglected practice of prefix delegations. The contributions of this chapter can be summarized as follows:

1. We provide a first of its kind yet simple classification of prefix delegations, solely based on the overlapping properties of two prefixes announced by two different ASes, as well as the existence and order of these ASes in observed AS paths collected by BGP collectors. We investigate prefix delegations over a timespan of ten years, and find that delegations have been present since then. While prefix delegations are not the fastest growing type of prefixes, they contribute almost 15% to the global routing table size.
2. We enhance our resulting dataset with business relationship inferences from CAIDA and find that delegations of provider aggregatable address space from provider to customer (53%), which is the only commonly presumed delegation type in liturature, is indeed not the only way delegations are performed. We find that 10% of delegations are also performed from customer to provider, which make the fastest growing delegation type. Moreover, 34% of delegations also happen between ASes that are not directly connected, or which have no apparent topological relation at all.
3. While we show that delegations are performed by various ASes of shape and size, we highlight in case studies the extensive use of delegations by large CDNs and ISPs. We show that the delegatees differ significantly in size, which indicate that the usage of prefix delegations can serve diverse purposes, e.g., sophisticated traffic engineering. For examples, while ISPs delegate prefixes to their customers, CDNs delegate prefixes to large ISPs in order to support their content distribution strategies. Moreover, we show in case of a globally operating hotel company that delegations are not exclusive to big players.
4. Finally, we use traceroute measurements obtained from CAIDA and correlate it with our dataset. The results clearly indicate that prefix delegations have a profound impact on traffic flow in the Internet.

5.2 Background & related work

AS-level Internet and BGP: The Internet consists of more than 50,000 interconnected *autonomous systems* (AS). An AS is a network which is operated as a single administrative entity. Usually, neighboring ASes have complex contractual agreements that govern their routing policies. Common agreements include: *provider*/*customer* and *peering*. Customers pay their provider for the traffic (volume) that they exchange and get access to Internet routes. In a peering relationship both ASes exchange their traffic and that of their customers on a settlement-free basis. In *sibling* relationships the routing policies are mixtures and the two involved ASes often belong to the same administrative entity. Although provider/customer and peering make up the majority of business relationships in the Internet, more complex relationships exist and become more common-place [96].

To exchange routing information ASes use the *Border Gateway Protocol* (BGP), the de-

facto standard inter-AS routing protocol. With BGP ASes selectively originate and forward prefix announcements to neighboring ASes. BGP is a path vector protocol, i.e., whenever an announcement is forwarded each AS appends its own AS number to the *AS path* attribute. The AS path, i.e., a sequence of AS numbers, is used for loop detection and as distance metric. The AS that originates the announcement is the *origin AS*.

Since BGP configuration is error prone, e.g., [57], it is critical for operators to view the routes they announce. Thus, BGP collectors, e.g., RIPE RIS and RouteViews, have collected *BGP updates*, i.e., prefix announcements and withdrawals, over many years. Although the collectors do not cover the whole AS topology, e.g., [147, 146, 102], the data is used for many other studies as it is sufficient to infer the AS graph, e.g., [89, 56, 131, 72, 127, 116, 96].

Business relationship inference: Usually, ASes do not disclose their specific routing policies. As a consequence, business relationships have to be inferred. Most inference algorithms rely on the assumption that there is no economic incentive to forward traffic across two peering links or via a customer. This is captured by the valley free property, e.g., see [89]. An AS path is *valley free* if a provider-to-customer (p2c) or peer-to-peer (p2p) edge is not succeeded by a customer-to-provider (c2p) or peer-to-peer edge.

One approach, presented by Luckie et al. [127], relies on the assumption that (i) there is a clique of peering ASes at the top of the hierarchy, (ii) most customers enter a transit agreement to be globally reachable and (iii) cycles of p2c links should not exist to enable routing convergence. In the same work the authors present multiple methods to infer the customer cone of an AS; a set of ASes that can be reached by traversing p2c links only.

Related work: McDaniel et al. [130] use a delegation tree to study the feasibility of origin authentication. While studying the growth of the Internet, Sriraman et al. [156] use allocation data from RIRs and BGP prefixes to construct a delegation hierarchy. Motivated by the increasing routing table size, Cittadini et al. [72] analyze overlapping prefixes, by relying on BGP data. While their focus is on deaggregation, they infer a widespread use of AS path prepending and scoped advertisements. Similarly, Bu et al. and Meng et al. [56, 131] use overlapping prefixes and the AS path to infer load balancing and multi-homing and their impact on the routing table growth.

5.3 Data sources

In this work we use a variety of datasets. We first obtain and clean publicly available BGP data, in order to identify and classify delegated prefixes. Then we use several datasets from CAIDA to study the relationships among the involved ASes and how delegations affect the traffic flow.

BGP: We download BGP data from BGP collectors maintained by RouteViews [12] (*rviews*) and RIPE RIS [10] (*ripe*). Since BGP is a routing protocol and not a measurement tool the collected data suffers from misconfigurations, errors, etc. It contains artifacts such as unallocated AS numbers, unallocated prefixes and poisoned paths, e.g., AS paths with loops. We clean our dataset by removing such announcements. We also remove announcements from beacon ASes or compatibility AS numbers. In addition, we remove prepended ASes in the AS paths. Contrary to other BGP studies, we keep all prefixes independent of the prefix length. We further clean the dataset by removing announcements that are not suitable/compatible for the delegation classification, see

source	# prefixes RIB	# updates	total
ripe	635k (-4%)	590k (-3%)	644k (-5%)
rviews	667k (-4%)	597k (-2%)	668k (-4%)
both	686k (-6%)	610k (-4%)	696k (-7%)

Table 5.1: Overview of BGP dataset d_{2016}. In parentheses relative decrease after cleaning.

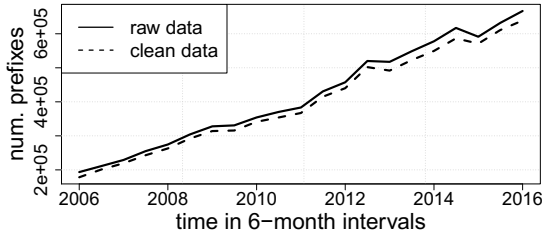


Figure 5.1: Overview of BGP dataset d_{hist}.

Section 5.4. This includes those with an AS path length less than two, with AS sets, and with ambiguous information, e.g., prefixes with multiple origins. The latter affects around 1% of the overall prefixes before cleaning. Table 5.1 and Figure 5.1 summarize our datasets before and after cleaning. By merging the data from *ripe* and *rviews* the total number of prefixes increases only slightly.

Adding the updates does not substantially increase the number of prefixes either. While data cleaning does remove some information it does not drastically reduce the number of prefixes in both datasets. For a longitudinal study we use routing table dump snapshots which contain the routing information base (RIB) from 2006 to 2016 every 6 months for one day (d_{hist}). To study prefix delegations in January 2016 in more detail we also include BGP updates (d_{2016}).

CAIDA: We augment our data with AS business relationships using the data from Luckie et al. [127] ($d_{business}$). It labels AS pairs as either p2c or p2p. Furthermore, we use the corresponding customer cone dataset [127] (d_{cone}). We enhance the above datasets with complex AS relationships from Giotsas et al. [96] ($d_{complex}$) as well as with additional AS links in IXPs from Giotsas et al. [98] (d_{mlp}). We further use CAIDA's AS-to-organization dataset which provides unique organization identifiers mapped to AS numbers [4] (d_{org}). Finally, we use large-scale traceroute measurements taken by CAIDA's Archipelago infrastructure [3] (d_{trace}). We obtained these datasets from the same time period as d_{2016}.

5.4 Prefix delegations

In this section we describe how we identify and classify prefix delegations. Hereby, we use the following notation: For a prefix P_α, AS_α refers to the AS that originates the prefix. $Path(P_\alpha)$ refers to the set of paths that are recorded for the prefix.

Consider the example of a multi-homed environment where AS_x is a customer of AS_y and AS_z, see Figure 5.2. AS_z announces a *deaggregated* prefix. More specifically, $P_{z'}$ is deaggregated from P_z and both are announced by the same AS_z. AS_x announces a *delegated* prefix, namely P_x which is a subset of P_y. Thus, P_x is deaggregated from P_y and delegated to/announced by AS_x.

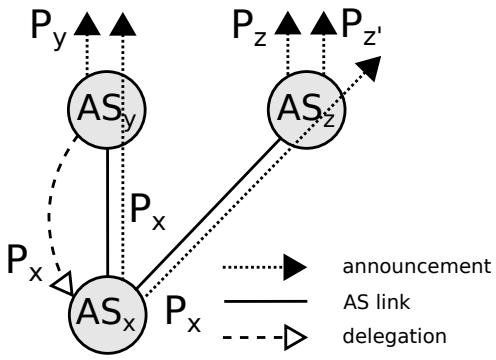


Figure 5.2: Multi-homing example

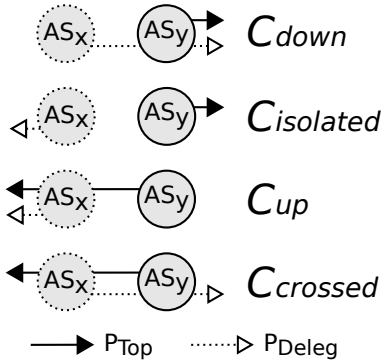


Figure 5.3: Delegation classification

Prefix classification: To identify delegations we classify prefixes based on their overlapping properties and the origin AS. More specifically, we first check if any two prefixes are subsets of each other, i.e., $P_x \subset P_y$. If positive, we further check if they are announced by the same AS, i.e., $AS_x = AS_y$. Inspired by Cittadini et al. [72] we distinguish the following prefix classes:

$P_{lonely} : \nexists\ P$ where $P_{lonely} \subset P \vee P \subset P_{lonely}$

$P_{top} : \nexists\ P, \exists\ P'$ where $P_{top} \subset P \wedge P' \subset P_{top}$

$P_{deagg} : \exists!\ P_{top}$ where $P_{deagg} \subset P_{top} \wedge AS_{deagg} = AS_{top}$

$P_{deleg} : \exists!\ P_{top}$ where $P_{deleg} \subset P_{top} \wedge AS_{deleg} \neq AS_{top}$

In words, prefixes in P_{lonely} do not overlap with any other prefix. Prefixes in P_{top} are always the less specific of two overlapping prefixes. The more specific prefixes are either in P_{deagg} or P_{deleg}. If two prefixes P and P' overlap and P is in P_{top} then P' must be in P_{deagg} or P_{deleg}. It is in P_{deagg} if both prefixes are announced by the same AS. If the prefixes are announced by different ASes P' is in P_{deleg}. We refer to the AS that is originating P as AS_{top} and the AS that is originating P' either as AS_{deleg} or AS_{deagg}. Hereby, AS_{deagg} is the same as AS_{top}.

Delegation classification: After using the above method to identify delegations, we next subclassify them into four different classes. We analyze AS paths of each prefix in P_{deleg} and of the correspondent less specific prefix in P_{top}. Considering the example of AS_y delegating address space to AS_x, as shown in Fig. 5.2, we check for two properties in the respective AS paths: (i) AS_x announces the more specific prefix via AS_y and (ii) while AS_y announces the less specific one, the announcement does not pass through AS_x. Delegations with these properties are in C_{down}[19]. While this is the most intuitive, other combinations exist as well, see Figure 5.3. Note, this classification does not require the existence of an AS link between both ASes. We distinguish the following four delegation classes:

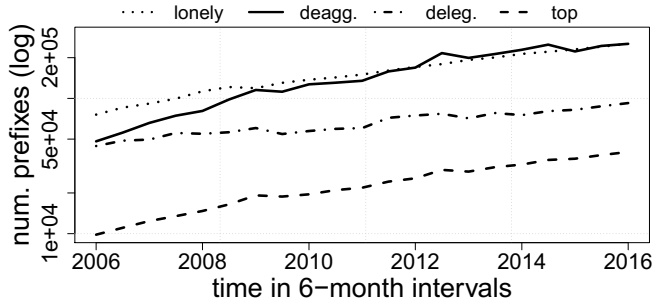
$C_{down} : AS_{top} \in Path(P_{deleg}) \wedge AS_{deleg} \notin Path(P_{top})$.

$C_{isolated} : AS_{top} \notin Path(P_{deleg}) \wedge AS_{deleg} \notin Path(P_{top})$.

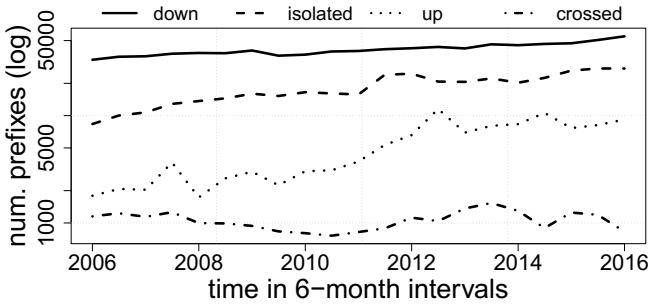
$C_{up} : AS_{top} \notin Path(P_{deleg}) \wedge AS_{deleg} \in Path(P_{top})$.

$C_{crossed} : AS_{top} \in Path(P_{deleg}) \wedge AS_{deleg} \in Path(P_{top})$.

[19]The notation C_{down} implies the delegation going *down* the AS level hierarchy, e.g., from provider to customer.



(a) Prefix classes over time.



(b) Delegation classes over time.

Figure 5.4: Longitudinal study, d_{hist}: Jan. 2006 to Jan 2016.

In effect, the class C_{up} is the opposite of C_{down}. In C_{up} the delegator announces the less specific prefix via the delegatee and the delegatee announces the more specific prefix but not via the delegator. $C_{isolated}$ and $C_{crossed}$ cover the two remaining cases. In $C_{isolated}$ both the delegator and the delegatee announce their prefix separately. In $C_{crossed}$ the delegator announces the less specific prefix via the delegatee and the delegatee announces the more specific prefix via the delegator.

5.5 Delegations across 10 years

Using d_{hist}, Figure 5.4(a) shows how each class of prefixes: P_{top}, P_{lonely}, P_{deagg} and P_{deleg} evolved. All prefix classes have grown at almost constant rate over the past ten years. In 2006 the number of prefixes in P_{deagg} was about the same as those in P_{deleg}. In fact, Cittadini et al. [72] show that before 2006 P_{deleg} had an even bigger share of the routing table than P_{deagg}. The authors conjectured that the increased popularity of PI address space led to the changes in these classes. Today, P_{deagg} is the fastest growing class—by a factor of 5.3— and is among the largest classes since 2009. In January 2016 P_{deleg} make up 14.6% of the routing table size which corresponds to more than 93k prefixes, see Table 5.2.

Even though P_{deleg} is not growing as fast as some of the other classes, we consider *delegated prefixes* the most intriguing class as they reflect the large complexity of BGP; its numbers have doubled in the past decade. Using d_{hist}, Figure 5.4(b) shows how much each of the delegation classes contribute to this increase. As expected, C_{down} has and has

Prefix class	#Prefixes	%Prefixes	#ASes
P_{lonely}	252,917	39.2%	38,971
P_{deagg}	257,244	39.9%	11,558
P_{deleg}	**93,754**	**14.6%**	**13,689**
P_{top}	40,475	6.3%	12,647

Table 5.2: Overview of prefix classes in d_{2016}.

Delegation class	#Prefixes	%Prefixes	#AS pairs
C_{down}	**56,294**	**60.0%**	**12,427**
$C_{isolated}$	27,016	28.8%	5,748
C_{up}	9,546	10.2%	1,183
$C_{crossed}$	898	1.0%	127

Table 5.3: Overview of delegation classes in d_{2016}.

had the largest share with well over 50%. However, $C_{isolated}$ is a substantial contributor with more than 20% in the past decade. The largest increase is seen in the class C_{up}; it has grown by a factor of 5.1 over the past ten years. Considering d_{2016}, we confirm that C_{down} is the most common case with a share of 60%. However, $C_{isolated}$ and C_{up} are substantial with more than 29% and 10%, respectively. The smallest class, with a contribution of only 1%, is $C_{crossed}$.

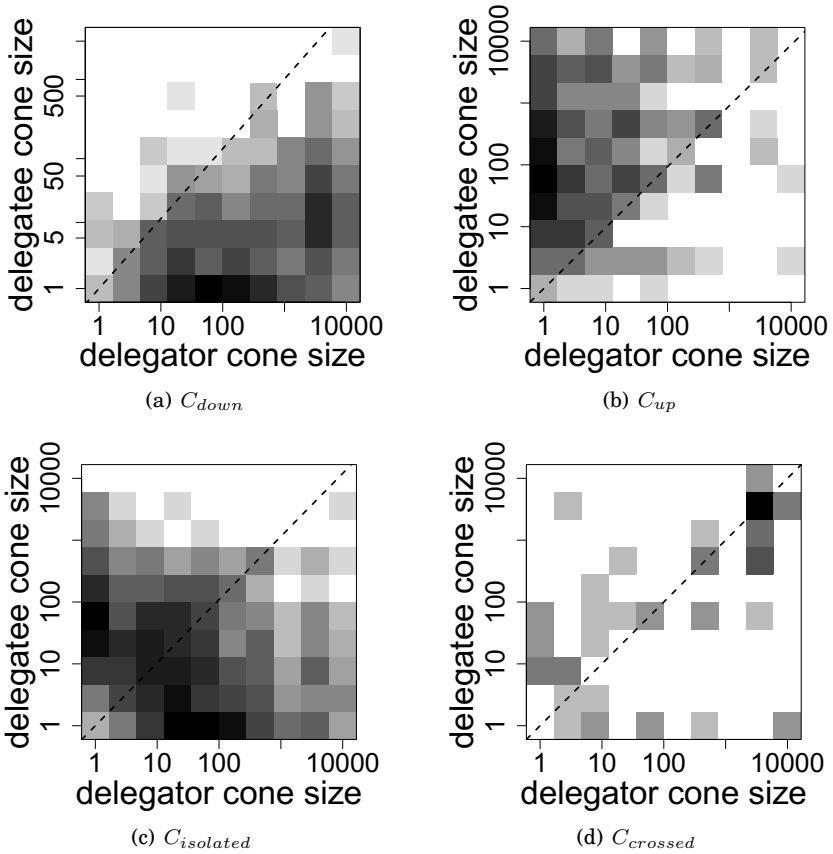
5.6 AS business relationships

The delegations from d_{2016} involve more than 16k delegators as well as delegatees, emphasizing that delegations are common practice. Also, they involve more than 19k AS pairs (delegator to delegatee), see Table 5.3. We observe that an AS pair can be involved in more than one delegation type.

Delegation vs. AS size: We explore to which extent the four delegation classes align with the relative size of the two involved ASes. Hereby, we use the customer cone size as a proxy for the AS size using d_{cone}. Figure 5.5 shows four heat maps—one for each delegation class—with the cone size of the delegator (x-axis) vs. the cone size of the delegatee (y-axis).

All delegation classes include ASes of varying AS sizes (from 1 to 10k+) both as delegatee and as delegator. We notice substantial differences. In C_{down} more than 99.5% of all delegators have a larger cone size than the delegatee. For C_{up} we see the opposite—in 93% the delegator has a smaller cone size than the delegatee. Thus, the delegated prefix is either originated by (C_{up}) or announced via (C_{down}) the AS with the larger cone size. The heat map for $C_{isolated}$ is not that focused on either the upper or lower half; we see a mixture. Some delegations take place between ASes with large cones to those with smaller cones and the other way around. $C_{crossed}$ shows a dense spot of AS pairs which have large customer cones. Examples include delegations between NTT America and Cogent, two ASes of Level 3, AT&T and Qwest, Qwest and Verizon, and AT&T and Level3. These mainly involve major ISPs and content delivery networks. Some of the delegations may be artifacts of mergers or internal network practices.

Business relationships: We next correlate prefix delegations with business relations using $d_{business}$. For each of the 19k AS pairs involved in prefix delegations, i.e., delegator and delegatee, we assign either a c2p, p2c, p2p relationship or label it with x if no rela-



(a) C_{down} (b) C_{up}

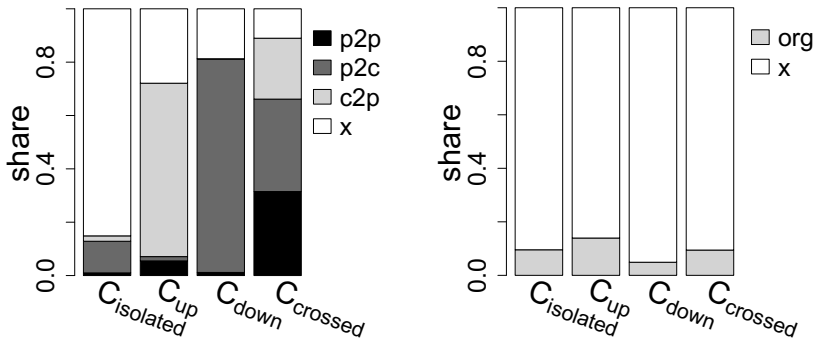
(c) $C_{isolated}$ (d) $C_{crossed}$

Figure 5.5: Heat map per delegation class: delegator cone size vs. delegatee cone size. The gray scale indicates the density in log scale.

tionship is included in the data. In any of the *x2y*-like assignments, *x* is the delegator and *y* the delegatee. Figure 5.6 shows barplots of the classification for each of the different delegation classes. Interestingly, we find that in 39.5% of the AS pairs, the delegator is not adjacent to the delegatee (34.4% of all delegations). In $C_{isolated}$ 85% of AS pairs are unclassified while the fraction in C_{up}, C_{down} and $C_{crossed}$ is less: 28% / 19% / 11%.

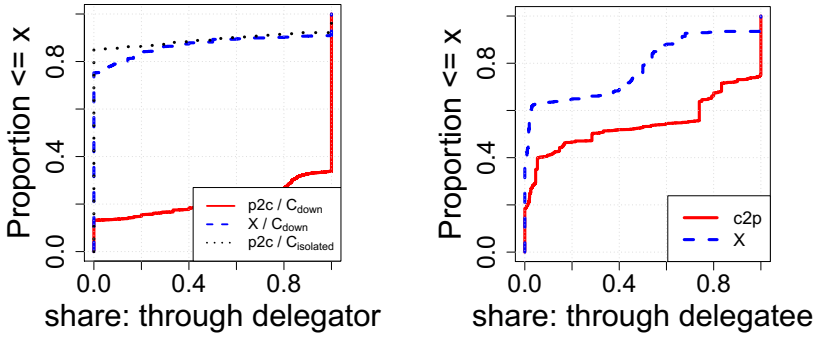
If we only consider delegations between adjacent ASes we find that 99% AS pairs in C_{down} have a p2c relationship. This is what one may expect and is consistent with our previous observations regarding the customer cone size differences. These are the ones that fall below the diagonal in Figure 5.5(a). Similar observations hold for C_{up}. 90.2% have the expected c2p relationship and indeed these are the ones that fall above the diagonal in Figure 5.5(b). The others are mainly p2p (7.6%) with only a small fraction of p2c (2.2%). $C_{crossed}$ includes the largest fraction of p2p relationships. This hints at mutual agreements between the two involved ASes which can result in such apparently unusual routing arrangements. It is not surprising that this class includes AS pairs where both have large customer cones. Overall, these results show that each delegation class involves a distinct variety of business relations among the ASes.

We acknowledge that some of the *x*-labeled AS pairs might be caused by AS links that are not visible in this dataset. To mitigate the impact of such missing links, we use additional datasets, d_{mlp} and $d_{complex}$. However, we note that they only provide a minute number of additional AS links. In particular, out of the links in $d_{complex}$ we only find 38



(a) business relationship (b) organizations

Figure 5.6: Barplot: delegations by business relationship and organization.



(a) C_{down} & $C_{isolated}$ (b) C_{up}

Figure 5.7: CDF: path selection

and in d_{mlp} 49 additional links which corresponds to 1% of our x-labeled AS pairs. We discuss delegations among non-adjacent ASes in detail in Section 5.7.

Organizations: We also use the AS-to-organization mapping from d_{org} to check if most of the delegations occur between ASes of the same organization. The resulting barplot in Figure 5.6(b) shows that 10-15% of the AS pairs are within the same organization. The largest fraction are in C_{up} the smallest in C_{down}. However, this is by far not the majority.

5.7 Effects on path selection

While BGP routing information provides multiple paths towards a destination the actual traffic follows the *best* path. To better understand how prefix delegations affect the traffic flow, we complement our analysis of prefix delegation classes with large-scale traceroute measurements from d_{trace}. We select two sets of traceroutes: those that target the delegated prefix P_{deleg}, and those that target non-overlapping parts of the associated/less specific P_{top}. We exclude traceroute results that do not reach the destination AS. For that we map IPs from d_{trace} to ASes by applying the longest prefix match using prefixes from d_{2016}. This results in traceroutes towards 56,543 (60%) of the delegated prefixes and towards 7,458 (70%) of the less specific prefixes. We refer to this set of traceroutes as d_{flow}. For each delegation covered by d_{flow} we determine the ratio of traceroutes going through the delegator / delegatee. We refer to it as the *pass-through*

rate (*ptr*). Figure 5.7 shows the *empirical cumulative distribution* (CDF) of the resulting *ptrs*. In both plots we see that for some delegations the traceroutes always use the path through the delegator / delegatee (*ptr* is 1), for other delegations this is never the case (*ptr* is 0). In the following we analyze some of the delegation scenarios using this additional information.

5.7.1 PA Prefixes from Provider to Customer

A common example of a delegation is in a p2c relationship, where the provider AS delegates PA address space to its customer AS. This delegation scenario falls into C_{down} if the provider announces both, the more and the less specific prefix. We find traceroutes for 29.1k of those delegations in d_{flow}. The solid line (p2c/C_{down}) in Figure 5.7(a) shows the corresponding *ptrs*. For 65% of those delegations the traceroutes always go via the delegating provider to the customer. A more detailed analysis shows that one third of the delegatees appear to be single-homed (cone size of 1). This contradicts best current practices as stated, e.g., in[153]. However, we also see the opposite: For roughly 15% the traceroutes never go through the delegating provider, but via an alternative one. We conclude, given the number of single-homed customers, there is significant potential for further address aggregation.

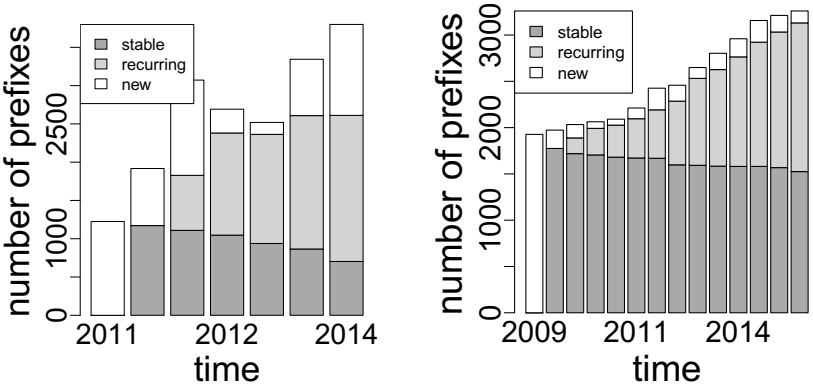
If however the provider only announces the less specific, it falls into $C_{isolated}$. Because opposite to the previous case, here the delegating provider aggregates the PA prefix with its own. Note that the customer must be multi-homed because despite aggregation, we can observe the delegated prefix. We find traceroutes for 1.3k of those delegations in d_{flow}. The dotted line (p2c/$C_{isolated}$) in Figure 5.7(a) shows the corresponding *ptrs*. For the majority (around 85%) of delegations the traceroutes towards the customer never pass through the delegating provider. However, in 10% of the delegations they always do, despite aggregation. This hints at either limited propagation of the more specific prefix or uncommon routing policies.

Next, we check if for delegations in C_{down} with non-adjacent ASes traceroute data provides additional information. Using d_{flow}, we find traceroutes for 2.6k delegations of this type. For around 75% we find no traceroute which goes through the delegator, see dashed line (*x*/C_{down}). Less than 10% always go through the delegator. Note that in this case 15% of the delegatees are in the customer cone of the delegator. This supports our claim that there are indirect business relationships among two involved ASes.

5.7.2 Delegations from Customer to Provider

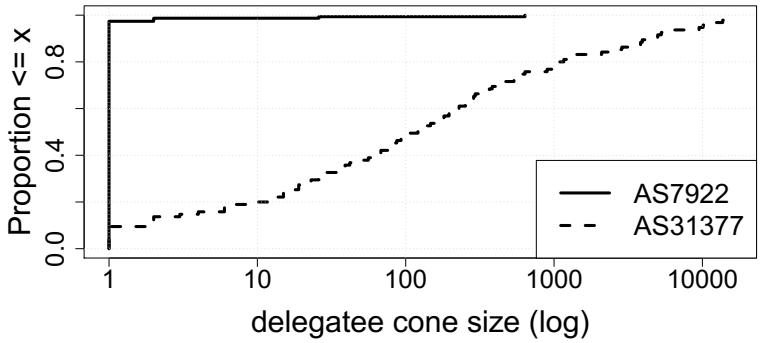
Next, we explore the case where the customer AS delegates prefixes to its provider. This involves 80.4% of the delegations in C_{up}, 31.4% in $C_{crossed}$ and 3.7% in $C_{isolated}$. These findings contrast previous work [153] where the authors state that delegations from customer to provider are unlikely to be found in the Internet. While we find many customer ASes delegating single prefixes to providers, we also see some ASes delegating hundreds of prefixes. The latter involves, e.g., delegations within organizations or CDNs. Recall, the number of delegations in this class has grown the most over the past ten years indicating the need for such services.

We find traceroutes for 5.2k of those delegations (c2p/C_{up}) in d_{flow}. The solid line in Figure 5.7(b) shows the CDF of the corresponding *ptrs*: For 57% some traceroutes go through the delegatee and others do not, i.e., the *ptr* is between 0 and 1. In only 25%



(a) Barplot: AS31377

(b) Barplot: AS7922



(c) CDF: delegatee customer cone size distribution

Figure 5.8: Case studies—Large ISP/large CDN: Churn and cone size distribution.

of the delegations the traceroutes always pass through the provider. For the remaining 18%, none of the traceroutes do. Using d_{flow} again, we find traceroutes for further 1.3k C_{up} delegations among non-adjacent ASes (x/C_{up}). Here we observe a similar behavior: For more than 60% of those delegations the corresponding *ptr* is between 0 and 1, see dashed line in Figure 5.7(b).

Compared to the delegations of PA prefixes from provider to customer, here the path selection of the traceroutes is less consistent. While for roughly 20% of the delegations in Figure 5.7(a) the *ptr* is between 0 and 1, it is roughly 60% of the delegations in Figure 5.7(b). We conclude that depending on the type of delegations, i.e., how prefixes are announced by the delegator and delegatee, the path selection of traceroutes is noticeably affected.

5.7.3 Delegations among Non-Adjacent ASes

34% of the delegations involve AS pairs without any AS link (recall Section 5.6). The majority is in $C_{isolated}$ where no announcements pass through the involved ASes. This is confirmed by the traceroute data, i.e., in around 90% of $C_{isolated}$ delegations, the traceroute never goes via the delegator or the delegatee (plot not shown). In order to underline the diversity of prefix delegations, we present case studies comparing a large ISP (Comcast) and a large CDN (Akamai). Also, we look at a small organization operating at global scale.

Often organizations use several ASes, e.g., AT&T. Yet, it appears that only a single AS (or a small number of ASes) is used to delegate prefixes to other organizations. For example: between 2011 and 2014 the Akamai AS_{31377} delegated more than 2,000 prefixes to several non-adjacent Tier-1 ASes world-wide (C_{up}) and to other Akamai ASes. After its disappearance from the routing system in 2014 AS_{31377} was replaced by AS_{35994}. In January 2016 AS_{35994} delegated more than 4,000 prefixes to more than 100 ASes. We find that these delegations are volatile. Figure 5.8(a) shows the churn in delegations over time for AS_{31377} using d_{hist}. Whenever we observe a delegation for the first time we label it *new*. As long as a prefix keeps being announced we label it *stable* until it disappears. If it reappears it is in *recurrent*. Supported by the high increase of delegations and the irregular but high growth of new delegations, AS_{31377} often delegates and revokes delegations. This is in contrast to AS_{7922} (Comcast) where the number of stable prefixes decreases slower over time, see Figure 5.8(b).

Next, we compare the customer cone size distribution of delegatees of AS_{31377} and AS_{7922}, using d_{cone} data from 2014. We consider only delegations to delegatees in organizations different from the delegator. While AS_{7922} delegates 170 prefixes to 154 unique delegatees (all US organizations), AS_{31377} delegates 2,062 prefixes to 95 delegatees, worldwide. Both ASes mainly delegate to non-adjacent delegatees, i.e., 93.5% for AS_{7922} and 97.9% for AS_{31377}. While for AS_{7922} most delegations are in C_{down} for AS_{31377} all are in $C_{isolated}$ or C_{up}. Figure 5.8(c) shows the CDF of the delegatee cone size for both ASes. We observe that for AS_{7922} more than 97% have a cone size of 1. The distribution for AS_{31377} differs significantly as roughly 50% have a cone size larger than 100.

In order to show that these versatile delegations are not limited to big organizations, we next study a hotel company: Fairmont Hotels & Resorts Inc. owns a /16 according to WHOIS and its address space is maintained by 'Q9 Networks - Canada's data centre'. They use a conglomeration of different delegation strategies: The main Q9 AS (AS_{12188}) appears to delegate all prefixes to its upstream providers, i.e., Cogent, Qwest, Bell Ca, and Shaw (C_{up}). In addition, we find intra-organization delegations to other ASes, which are connected to yet another set of large ASes including Tinet SpA. Furthermore, there are a lot of $C_{isolated}$ delegations to ASes in several locations, e.g., Canada, US, Mexico, Bermuda, Singapore. Other delegatees are Level3 and Verizon.

5.8 Chapter Summary

To better understand the growth of the global routing table in this chapter we studied one of its constituents: delegated prefixes. Therefore we used publicly available routing information from RIPE RIS and RouteViews from a timespan of ten years.

We first cleaned the available BGP data from routing artifacts and information that we did not need for our analysis. We presented the combined view of both collector projects and contrasted the impact of our cleaning process over time. While merging RIB dumps and updates from both, *ripe* and *rviews*, only slightly increases the number of records, our cleaning efforts did not reduce our base dataset significantly either. Also, we augmented our dataset with various datasets provided by CAIDA, some of which include business relation inferences or large-scale traceroute measurements.

By applying a methodology inspired by Cittadini et al. [72], we identified delegated prefixes from BGP routing information Thereby, each delegated prefix is tied to a less specific prefix, while both being announced by two different ASes. We then introduced our own unique classification methodology which enabled us to exclusively identify four delegations classes, based on the existence and order of the two origin ASes in any visible AS path of the corresponding prefix announcements.

We quantified the share of delegations in the global routing table as well as each of the four delegation classes over time. While delegations are not the fastest growing prefix type it contributes to almost 15% to the global routing table. C_{down} delegations, i.e., down the AS-level hierarchy, represent with 60% the highest share among the delegations, while C_{up} delegations (up the AS-level hierarchy) represent the fastest growing delegation class.

Considering business relations, we were able to characterize a diverse ecosystem: We identify delegations from provider to customer, but also delegations from customer to provider, as well as delegations between AS with no apparent topological relation. Also, we found that only around 10% of delegations happen among ASes belonging to the same organization. Adding also traceroute measurements into consideration, we observed that prefix delegations impact the traffic flow. For example, C_{down} delegations from provider to customer, resulted in most of the corresponding probes flowing through the delegating provider to the customer. In the case of $C_{isolated}$ delegations from provider to customer, most of the corresponding probes did not flow through the delegating provider, but via a different provider to the customer. We were thus able to observe the effectiveness of multihomed ASes selecting one provider over the other.

Finally, we presented case studies of a large ISP, a large CDN, and a globally operating hotel company. We investigated the churn of delegations over time and found that the ISP's delegations are much more stable in terms of visibility as compared to the CDN. Moreover, we showed that the delegatees of both, ISP and CDN, differ in size. While ISPs mostly delegate to small ASes, delegatees of CDNs are of various size among which are large ASes.

6 Internet-Wide Scans by a Botnet

In the previous chapters we studied traffic at a large IXP and ISP, and routing information obtained from BGP collectors. We observed complex traffic patterns and traffic asymmetries on inter-domain links, leading to an overall network heterogenization. Also, we investigated some of the underlying routing mechanisms, more precisely prefix delegations, that are employed and its affect on the global routing table size.

This chapter deals with large and public dataset containing diverse types of Internet-wide active measurements, like ICMP pings, reverse DNS, and traceroutes. However, the dataset was published anonymously, and more important, it was generated by a botnet that uses hundreds of thousands unsecured home routers. Considering this extensive collection of measurement data, it is — not surprisingly — of significant interest to research institutions and network operators, which we were able to observe downloading via BitTorrent. Yet, little is known about it. Therefore, in this chapter we try to reverse engineer as much meta-data as possible, review the claims made by the author(s), and evaluate its aptitude for third-party usage.

6.1 Introduction

Anonymous authors released an Internet census report on March 17, 2013, together with the underlying dataset via a mailing list typically used for disclosure of security information [151]. The release contains a report anonymously hosted on BitBucket and GitHub, as well as 568GB of compressed data (9TB uncompressed) released via BitTorrent. In the Internet census report the authors claim to have conducted multiple scans of the entire IPv4 address space within 24 hours, using a large botnet which they call *Carna*. Primarily, these scans were directed at hosts via ICMP ping, at open ports and services, the reverse DNS tree, and some traceroutes. Part of these scans have been con-

firmed with CAIDA's Internet telescope which was scanned by the botnet as well [117]. Ironically enough, the anonymous authors build their botnet, supposedly consisting of 420k hosts, by exploiting default passwords. Note, using system resources without user permission is a violation of any reasonable terms of use. Thus, based on academic standards, their study is not only unorthodox but has to be considered *unethical*.

Although extensive ICMP censuses [104], port scans [80, 122, 105], and traceroutes [152, 67, 73] have been conducted before and even at a larger scale, the nature as well as the scale of the Internet census resulted in a media buzz [161, 111, 154, 165, 51, 166, 75, 170], an investigation by the Australian Computer Response Team [48], and in the creation of an Internet Census 2012 Search engine [30]. These responses, and the easy availability of the dataset have attracted many hundred downloaders world-wide. By participating in the BitTorrent swarm during the days immediately after the release, we observed more than 470 peers located in 38 countries, predominantly in China, USA, and Germany. Further, by mapping peers to ASes and to reverse DNS hostnames, we identify among the downloaders more than 30 universities and research facilities, 20+ ISPs, several infrastructure providers, as well as governmental and security organizations.

Whether unethical or not, the interest is evident. However, in particular considering the level of attraction, there is almost no knowledge about the authenticity and the quality of the published data. This is further exacerbated by the fact that the authors are anonymous and the left-behind dataset description is superficial or not existent at all. Without any kind of documentation or meta-data of the data *consumers* can easily misuse the datasets, as they do not know if the data quality is suitable for answering their questions in the first place. Often consumers simply assume that the data is of good enough quality for their purpose.

Due to many uncertainties that exist in and around the dataset, we challenge the claims made by the authors of the Internet census. The major contributions of this chapter can be summerized as followed:

1. In our effort to establish the authenticity of the released dataset, we reproduce some of the non-time-sensitive measurements and compare the results based on string-matching. In particular, we use some reverse-DNS measurements performed shortly before the release of the Carna dataset, which contains hostnames for around 70M IPs across 177 /8s. We confirm a match of more than 95% of the records from our dataset. Moreover, we attempt to validate nmap service probes for 4M Akamai IPs on port 80 returning an Akamai-specific user agent string. In order to do so, we scan these IP addresses ourselves shortly after the release of the Carna data, and compare the resulting user agent strings, again by string comparison. We confirm that around 84% of the results match those of the the Carna dataset. In conclusion, there is reason to believe that the published dataset is authentic. This outcome concurs with observations made by CAIDA using telescope data.

2. In order to verify the claims of the authors of the Internet census, we reverse engineer missing meta-data as best as possible and assess the quality of the data. Contrary to what the report states, we find less probes in the data. In particular, the target address space of the scans span around 3.7B IPs in 221 /8s which corresponds to the allocated address space. Also, the reported measurement periods of ICMP scans do not match the data. While we identify seperate measurement campains, we note that each exposes a different scanning rate, ranging from 220M to more than 2000M probes per day. Looking in more detail, we observe that scanning

was performed in groups of /8s, which however exhibit a temporal misalignment of more than 15 hours. Moreover, not a single /8 was fullly probed within 24 hours, some /8s are probed at least three times as often as other /8s, and some prefixes are probed up to twice as fast as others. In the search of clean censuses, we observe several iterations over the address space which however are overlapping and probably include test runs. Some iterations span up to six weeks while none of them reach the 100% of the target address space. Given all the discrepancies and weaknesses in the data, it is impossible for us to identify individual censuses.

3. Indeed, contrary to the authors' claims, we find that the dataset contains at most *one* census, as we will demonstrate in this chapter. In a parallel effort to the analysis, this work serves as example on how to validate measurement-based networking research, based on a methodology proposed by Krishnamurthy et al. [118]. We examine the data hygiene, i.e., how carefully the quality of the datasets was checked by the authors of the Internet census report, by analyzing the provided meta information. Furthermore, to drive our analysis we ask specific questions to ensure that certain requirements are fulfilled to reuse the data, e.g., questions that aim to uncover likely reasons for errors in the data. We conclude this chapter with a discussion on whether the adequate rigor was used by the authors to estimate the size of the Internet, considering the available data. Finally, we elaborate on the novelty of the conducted measurements as well as the public reactions and ethical considerations.

6.2 Published Datasets

In this section we introduce the datasets, their file organization, along with the data structure within the files. The Internet Census 2012 announcement [151] points to a Web site [49] containing the report as well as the datasets, available to everyone for download and analysis—in principle a great service to the community. The data spans several archives, 568GB compressed/9TB uncompressed, and is offered via BitTorrent, which is how we obtained the data. It includes the following datasets: *i) ICMP ping* reachability and latency information, *ii) nmap* port scans for open ports and per-port service information (host and service probes), *iii) nmap* TCP/IP fingerprints and IP ID sequence information, *iv) reverse DNS* records, and *v) traceroute* records. Each dataset is subdivided into smaller files, grouping all the probes into /8 blocks, based on the respective destination IP. In those /8 block files, each tab-separated probe record includes the destination IP, timestamp, and the probing result (e.g., ICMP ping result, list of open ports, etc.). Regarding the service probes, separate /8 block files are provided for each probed service, e.g., port 80/http. Finally, the downloaded data includes some (wallpaper) images, some data for the website, along with the source code of the website, a modified nmap tool, and a Hilbert graphic generator.

Having compiled the most basic description of the published data, we first want to understand the properties of the measurements. First, we find that the data collection started in April 3rd, 2012 and lasted until December 18th, 2012[20]. For this measurement period, in Table 6.1, we summarize the number of total ICMP probes, host probes, reverse DNS queries, and traceroutes records available in the data as well as the number of total probes that were stated in the report. Surprisingly, there are various mismatches in what is claimed in the report to what is in the actual dataset. For example the report

[20]Two timestamps date back to 1978 which obviously is outside the range of the Internet census.

dataset	total probes		probed hosts
	data	report	
ICMP Ping	49.5B	52B	3,706,585,088
Host probes	19.7B	19.5B	3,705,342,574
Reverse DNS	10.5B	10.5B	3,700,481,860
Traceroute	68.7M	68M	64,666,758

Table 6.1: High-level statistics for some of the datasets.

states that there are 2.5B (5%) ICMP ping probes more than we count. We elaborate more on the inconsistencies in Section 6.4.1.

Further, we analyze the targeted address space and count the number of unique hosts that were probed. Since this information was not provided by the *producers* of the data, we see it as our responsibility to fill the missing information. Except for the traceroute records, the number of unique hosts is more than 3.7B which corresponds to the currently allocatable address space [113]. Indeed, we did not see any probes for IP blocks that are listed as reserved by IANA, private address space, as well as multicast space. So in total, the datasets comprise of probes launched towards 221 /8 blocks. Notably, in the case of ICMP, considering the overall and the unique probed hosts, the data allows for at most 13 censuses.

We note that all datasets but the traceroute records do not include the source IP address, i.e., the IP of the probing host. This is problematic and challenges the use of the data as some results depend on the location of both the source as well as the destination, e.g., the ICMP latencies. But there are also more subtle problems that are not addressed by the authors, e.g., is the destination IP behind a firewall or a proxy that may alter the reachability results? We note that including information about the operating conditions of any involved network during the measurement periods can be crucial to properly interpret the data.

Finally, we observe that the measurement periods of the individual datasets are of varying lengths, irregular, and only partially overlap. In fact, we cannot recognize any reasonable kind of measurement schedule in the data as well as in the documentation.

6.3 Authenticity

We continue our analysis by asking the most fundamental question: is the data authentic or manufactured—an April fool hoax? We aim to answer this question by reproducing some of the measurements. This is non-trivial, as network conditions in the Internet are subject to constant changes and therefore decrease the reproducibility over time. However, if we are able to reproduce some of the measurements, it can be a strong indication of authentic data. We note that CAIDA has confirmed that the scanning took place, using a combination of telescope data and the census data [117]. In this section, we add to this by taking a closer look at those parts of the datasets that are less time dependent, e.g., the reverse DNS records and the server IP addresses—for which we happen to have comparable datasets.

6.3.1 Reverse DNS

We start with the reverse DNS dataset which we compare to a separate, external dataset of reverse-resolved IP addresses captured in November 2012. Note, November 2012 is just shortly after the reverse DNS data collection of the Internet census ended. Our dataset contains 70.6M IP addresses across 177 /8s for comparison, while the Internet census data contains 3.7B in 221 /8s.

We check the datasets for consistency via string comparison of the hostnames from both datasets using the external dataset as basis. We find exact hostname matches for 95.2% of the tested IPs. For 3.1% of the IPs the external dataset finds a reverse name, which is not reported in the census data. A closer look at these 3.1% shows that the unsuccessful lookups are due to DNS lookup errors (86.5%), non-existing reverse DNS entries (8.2%), and timeouts (5.2%). For the remaining 1.7% of the IPs, we do not find exact matches in the hostnames. The reasons for this are different hostnames (61.6%), even though the domain name and top level domain match, differences in capitalization (3.5%), or multiple reverse entries with different reverse entries in each of the datasets. The latter requires manual checking.

Overall, our test finds that almost all entries match in principle (>96%) for a dataset that was unknown to the authors of the Internet census. This indicates that the data is unlikely to be artificially manufactured.

6.3.2 Akamai IPs

The Internet census report states that 5% (4M) of all web servers on port 80 return the *AkamaiGHost* user agent string [49]. This user agent string is announced by Akamai CDN caches when requesting content that is not hosted by Akamai, e.g., as seen by nmap service probe scans during the Internet census. Similar to reverse DNS, we want to reproduce the measurements to verify the authenticity of the data.

Therefore, we collect another dataset by probing all 4M IPs from the Internet census a single time from a single IP address from our local university network in July 2013. For our probes we use an in depth understanding of Akamai's caching infrastructure. It was shown that any CDNized object is accessible from any Akamai cache [109, 162]. We exploit this by downloading two image objects hosted by Akamai (one from a major social network, another from a major car manufacturer), and one non-Akamai object. The latter download lets us distinguish open proxies from Akamai caches. Our script validates the SHA-1 hash and HTTP return code of all retrieved objects. We consider an IP address to belong to an Akamai cache, iff all three tests passed, i.e., if the hash and HTTP return code for the two CDNized objects match, and the non-CDNized object cannot be retrieved.

Out of the 4M IPs, 84.2% pass all tests and thus are consistent with Akamai caches. The remaining 15.8% fall into two categories. 10.5% of the IPs were unreachable, e.g., because of firewalled hosts that cannot be reached from the public Internet. The remaining 5.3% did not pass one or two of our tests, e.g., due to timeouts. However, 84% served at least one Akamai hosted object correctly and thus *appear* to be valid Akamai caches. Overall, the large number of validated Akamai caches again shows that presumably the data is not manufactured.

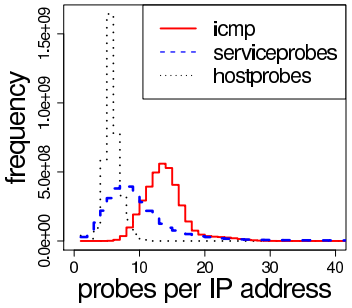


Figure 6.1: Probing Frequency: Distribution of how often each IP probed was probed for three kinds of probes.

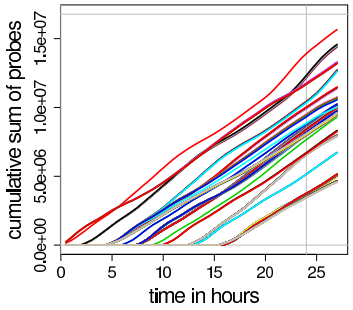


Figure 6.2: Misalignment of timestamps within 24 hours: Evolution of cumulative sum of ICMP probes for each /8 prefix in $icmp_3$.

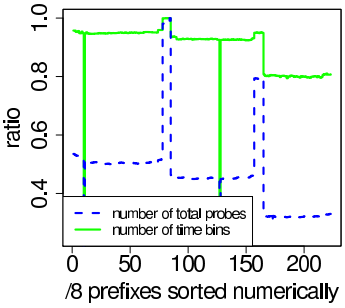


Figure 6.3: Scan diversities: ordered /8 prefixes plotted against the #probes/total (dashed) and #timebins/total (solid) ratios in $icmp_2$.

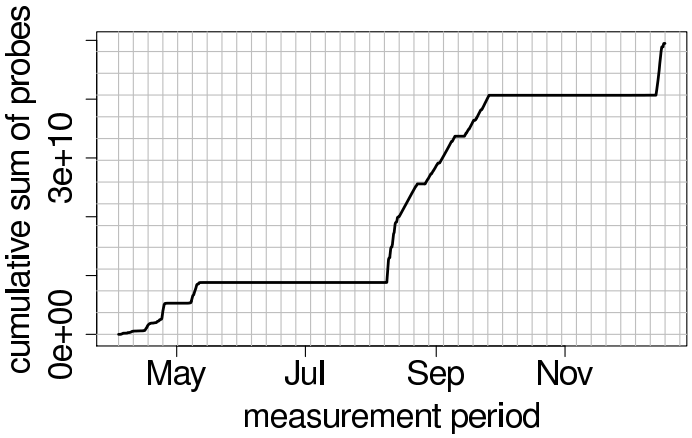


Figure 6.4: Measurement periods: Cumulative sum of all ICMP probes over entire measurement period (horizontal lines: 3.7B target address space).

6.4 Looking Behind the Curtain

Given that the data appears to be authentic, we want to validate the claims of the authors, e.g., claims that they have performed several censuses, among which are fast scans sweeping the complete Internet address space within 24 hours. However, we note that there is no meta-data in the report that relates to censuses, except for the description of two ICMP ping scanning methods, i.e., *"long term scan [...] for 6 weeks on a rate of [a complete scan] every few days"* or *"fast scans [...] probed the IPv4 address space within a day"*. In order to verify those claims, the ultimate goal of this section is to reverse engineer as much meta-data as possible, starting from what we are able to uncover in Section 6.2, to identify the censuses in the ICMP dataset.

6.4.1 Meta-data? Wrong!

As a first step, we examine as much information as possible from the meta-data reported by the authors, in order to check the reusability of the datasets. Regarding the Internet

census report, one would expect a detailed description of the measurement tool (botnet) and measurement data. However, while the Internet census report contains some rather superficial information about the measurement methodology using a large botnet, the dataset documentation itself lacks detailed information about the measured data. It gets even worse, when checking the consistency between the report and the data. For instance, the report mentions the ICMP measurement period spans *"from June 2012 to October 2012"*. However, in June and July no probes are reported in the dataset. This is confirmed by Figure 6.4 which shows the cumulative number of ICMP probes over the whole measurement period. The horizontal support lines correspond to 3.7B IP addresses, supposedly the base line of the probed address space, while the vertical support lines correspond to weeks. (The plot appears to be consistent with the interactive plot included in the report.)

Further inconsistencies between the data and the report concern the probed address space, and the number of samples. For instance, while the report states 52B ICMP probes, the dataset only contains 49.5B ICMP probes, see Table 6.1. Also, while the report refers to scans of *"all 3.6 billion IP addresses of the Internet"* or *"240k sub-jobs, each [...] scanning approximately 15 thousand IP addresses"*, the dataset reports roughly 3.7B probed IP addresses, see Table 6.1. When evaluating the completeness of the censuses, we therefore assume that the *target address space includes at least 3.7B IP addresses*.

From a hygiene perspective, a well maintained documentation of the data, e.g., meta-data, includes as much information as possible in order to allow any consumer to reuse the data adequately. The low level of documentation we find here, however, is problematic: While we cannot rely on the documentation as it is at best superficial and inconsistent with the published measurements, it is the only source of information given by the anonymous authors.

6.4.2 Data Quality

In the previous section, we find that the Internet census report only includes limited and partially inconsistent information about the data and how it was collected. In order to verify the claims of the authors, we need to reverse engineer as much meta-data as possible. Throughout this analysis, we focus on the ICMP dataset only, since from the report we assume it contains the censuses in question.

Probing Distribution

We attempt to reverse engineer the meta-data step by step to find the missing information, e.g., when does a census start and when does it end? Part of filling in the missing information requires knowing how each IP address was probed. This can help us to understand 1) how many censuses we can expect, and 2) what other data, except for clean censuses, are included in the dataset. For example, are there scans of different types? Do they overlap in time? Do we miss probes or see reprobes, e.g., due to bot failures?

In this section, we investigate the distribution of probes by counting the number of probes per IP address. Ideally, assuming our hypothesis in Section 6.2 is correct, we should find that each IP address was probed 13 times, resulting in 13 censuses. Figure 6.1 shows a histogram of the number of probes per IP address for three different kinds of probes: ICMP pings, host probes, and service probes[21]. We find that all are

[21] Due to resource constraints, we only focus on service probes directed at well-known ports.

name	period	total probes	probed hosts	days
$icmp_1$	Apr.-May	8.8B	3,682,182,938	40.0
$icmp_2$	Aug.-Oct.	31.8B	3,706,583,819	49.5
$icmp_3$	December	8.8B	3,704,509,119	4.3

Table 6.2: ICMP measurement periods overview.

highly skewed. Regarding ICMP pings, while most IP addresses are probed between 6 and 25 times, some IP addresses are probed more than 600 times and others only once. Indeed, the latter is highly problematic, as strictly speaking the data can thus only contain a *single complete census*, contrary to what the authors claim.

Due to the skewed distribution we cannot assume that the data consists of clean censuses. Indeed, there are many possible explanations for these different probing frequencies. One explanation is that beside the censuses, there is additional data included in the ICMP ping dataset. Since the meta-data description is poor, we cannot reject the hypothesis that census data is mixed with other unreported data. Another explanation can be problems with the bots. Failures by a subset of the botnet, which is a worldwide distributed set of workers and aggregation nodes, can severely impact the measurement data.

Probing Activities

Our first attempt to determine the number of censuses from the number of probes per IP has failed. Instead of 13 clean censuses we find a skewed distribution, which indicates that the alleged censuses may be mixed with other data collected for different purposes. In this section, we classify the measurement activity periods, to identify potentially separate experiments. This may enable us to not mix the results from incoherent scans and determine their individual purposes. Therefore, we analyze the ICMP probing activities, i.e., determining when and for how long the IPv4 based Internet was probed. Recall, Figure 6.4 shows the probing activity for the overall measurement period which spans more than eight months. Even a cursory glance at the plot indicates that the probing intensity varies significantly over time, thus it make sense to separate these periods. Initially, there appears to be some initialization period, then some scans, then a break, another scanning period, another break, and a final scanning period. Although not documented, we find three major activity periods separated by longer inactivity periods. Thus, we split the data into three subsets: $icmp_{1-3}$. Table 6.2 reports the number of the total number of probed IPs, unique IPs, as well as the measurement duration.

However, the purpose of those activity periods is not immediately apparent. The relatively slow scanning rate of $icmp_1$ (220M probes/day on average) and its irregular scanning behavior (short activity burst and a two-week break) suggest that it contains test runs while gathering experience with using the botnet as a measurement tool. Further, Table 6.2 shows that $icmp_1$ contains less than 3.7B uniquely probed hosts, contrary to the other periods. Test runs may explain the skewed distribution from Section 6.4.2. While $icmp_2$ (642Mp/d) consists of several stable scans separated by small breaks, including two 5-days breaks, the probing behavior in the very beginning is rather steep. Together with $icmp_3$ (2047Mp/d), a steep, short and stable period, these two periods appear to be potential candidates for fast scans.

Due to the surprisingly different and thus noteworthy characteristics that $icmp_{1-3}$ expose, we, in the remainder of this chapter, report our findings using these measurement

periods. Note, that there is no related description available in the Internet census report.

Botnet Architecture

As discussed in Section 6.4.2, the architecture of the botnet can be one reason for the skewed distribution of probes per IP address. For example, the challenges to be addressed by the data collection are handling failures both at the worker level as well as at the aggregation node level. Does the controller start the job from scratch at the same or another intermediate node? What happens with the results of the workers? Can these be stopped or reintegrated into the process? We assume that, due to the distributed nature of botnets, failures that relate to the orchestration of aggregation nodes and workers are reflected in the probing frequency of particular IP groups. For example, if one aggregation node fails, it can miss all the data that was measured and transmitted by the workers, while other aggregation nodes keep collecting data, leading to a skewed distribution. Thus, we need to understand at which granularity the jobs on the aggregator level are delegated and collected. This view enables us to see whether the botnet infrastructure causes some IP groups to be probed differently than others.

Since the measurement data is organized in /8 block files, we begin with checking if /8 is also the granularity for which an aggregation node is responsible. Accordingly, for the first day of $icmp_3$, Figure 6.2 plots the cumulative number of probes for each individual /8 prefix against the timestamps. Note the horizontal support line, indicating the /8 address space of around 16.7M IPs, and the vertical support line boxing the first day. We choose $icmp_3$, because it is a short and stable measurement period, thus our observations will most probably not be biased by different probing activities, as argued in the previous section. Surprisingly, with regard to the timestamps, the plot highlights a temporal misalignment of the start of the probing for all /8s which spans a time period of more than 15 hours. Similar observations also hold for the other measurement periods, i.e., $icmp_1$ and $icmp_2$. However, for smaller aggregation levels, e.g., /16 within the /8s, this misalignment is *not* present. In addition, in Figure 6.2 we observe that several prefix groups show similar characteristics. We therefore conjecture that *the scans are organized by /8 prefixes*. Finally, we note that from the first started prefixes not a single one received more than 16.7M probes within the first 24 hours.

A likely reason for the temporal misalignment seems to be the use of the local time at the aggregation nodes, in order to specify when a probe is launched (or a response is received). As the timestamps are not addressed in the report, we do not know how to interpret them. In which time zone are timestamps reported? How is time normalized (e.g., to UTC)? The logical presumption is that the experiment uses a single reference timezone, and that all timestamps are accordingly normalized. Otherwise, the dataset should contain timezone information which is not the case.

Probing Characteristics

Now that we can assume that the jobs are organized per /8s, we wonder whether each /8 was probed equally or not, in order to eventually find the explanation for the skewed probing distribution, as shown in Section 6.4.2. Throughout this section, we address this task by contrasting the individual /8 prefixes from $icmp_2$, but note that similar behavior holds for /16 and /24 prefixes as well as $icmp_1$ and $icmp_3$. Figure 6.3 plots for each /8 (in sorted order) (a) the ratio (dashed) of probes towards this /8 vs. the maximum number

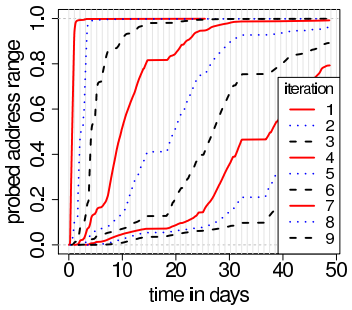


Figure 6.5: Overlapping iterations: First nine iterations over the probed address range in $icmp_2$.

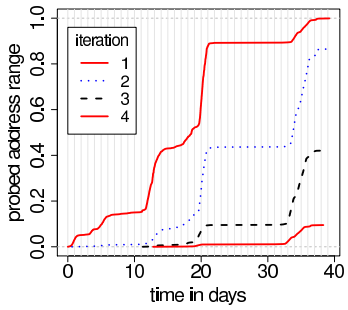


Figure 6.6: Overlapping iterations: First four iterations over the probed address range in $icmp_1$.

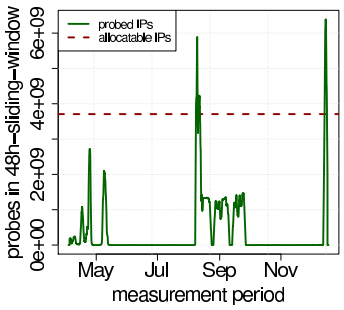


Figure 6.7: Finding *fast scans*: Sum of probes in 48h-sliding window over entire measurement period.

of probes any /8 got and (b) the ratio (solid) of time bins with probes towards this /8 vs. the maximum number of time bins seen for any /8. This way of plotting ensures that at least one prefix will have value 1 for both metrics. We note that there are no probes for 10/8 and 127/8. For time bin granularity we choose 30 minutes, as it is the maximum accuracy available in the data.

From the plot we can identify five different prefix groups when focusing on the ratio of probes at 100%, 80%, 50%, 45% and 35%. This implies that prefixes from one group are probed in a similar fashion. Moreover, prefixes from one group are IP-wise adjacent when sorted numerically. We point out the significant differences in probing frequency across the groups. Some /8s are probed at least three times as often as other /8s. This hints at some problems with the control flow of the experiment, or the distributed nature of the botnet. Thus, we conclude that how /8 prefixes are probed differs across /8s.

Also, for the second ratio which is normalized by the number of time bins, we again notice at least five prefix groups. This underlines the above observations that the probing is not done in a uniform manner across the /8s. Moreover, if we consider the relationships of the two ratios for the same /8 the different probing rates for the /8s, i.e., how fast the probing is done, become apparent. For example, we observe prefixes 1/8 to 165/8 to have a similar time bin ratio between 90 and 100%. However, as shown before, the ratio of probes in the same prefix range shows drastic differences, in particular for the prefixes 75/8 to 80/8, or 160/8 to 165/8. We conclude that the /8 prefixes are probed with different probing rates. More specifically, some prefixes are probed up to twice as fast as others.

Unfortunately, we find that the probing characteristics per /8 (/16 as well as /24) differ significantly both in terms of number of probes, as well as in terms of probing rate. We attribute these probing diversities to failures in the botnet architecture, which eventually seem to be responsible for the bad data quality. Therefore, as conclusion of our analysis, we point out that significant meta information remains unknown and the gained insights are not sufficient to verify the claims of the authors.

6.5 Claims of the authors

After having pointed out some inconsistencies between the Internet census report and data, and commented on the lack of meta information, as well as the data quality, we now turn our attention to two central claims of the authors related to the ICMP ping dataset. Concretely, the authors claim to have conducted several censuses. Some of them—the fast scans—supposedly are done within 24 hours, while the long term scan spans a period of six weeks. Thus, in this section we try to find all the censuses, and identify the fast scans.

6.5.1 Finding Censuses

One of our main problems in validating these claims is the lack of meta information. With regards to finding censuses, this is particularly relevant, as the dataset description does not state when a census starts, or when it ends. Therefore, in this section we try to uncover that missing information, and give an estimation of the number of censuses.

Scan Iterations

In the Internet census report, the authors claim to have scanned the IPv4 address space multiple times. When performing multiple scans in parallel, it is crucial to ensure that they do not overlap in time or, if they do overlap, to be able to separate the resulting datasets. To distinguish different scans of the address space, regardless of their duration, we use the concept of iterations: The first time an address is scanned belongs to the first iteration. The next scan belongs to the second iteration, etc. Should the data contain two full censuses then the first and the second iteration would cover the full IP address range. Should the data contain two full censuses, and some reprobing in order to do error recovery, the data should contain a full first and second iteration of the IP address range, a partial third iteration, and maybe even partial fourth and fifth iterations. We note, that reprobing is almost a necessity in order to recover from failures among the measurement bots. If two censuses are properly separated in time, namely non-overlapping, then the concept of iterations over limited time periods lets us separate censuses, since we can distinguish reprobings from the next census.

Figure 6.5 plots the CDF for the first nine iterations for time period $icmp_2$. The probed address range (y-axis) is premised on the respective number of probed hosts in Table 6.2. We notice that even the first iteration takes more than 6 weeks in order to cover the full address range. The other eight iterations also start within the first few hours, but do not reach 100%, indicating that there is reprobing and that there are no non-overlapping censuses in $icmp_2$. The second through fifth iterations reach 96-99%, while the sixth iteration does not even reach 90% of the IP address range. Moreover, it highlights the different probing frequencies of the different IP address ranges. Similar observations hold for the other measurement periods, e.g., see Figure 6.6 for $icmp_1$, or smaller time periods. We conclude that we cannot distinguish the scans, and therefore cannot count the number of individual and clean censuses in the ICMP dataset.

Censuses vs. Test Runs

When comparing Figure 6.6 for $icmp_1$ to Figure 6.5 for $icmp_2$, we notice a significant difference. While in $icmp_2$ the CDF for iteration one is very steep in the beginning, the first iteration in $icmp_1$ is rather flat and irregular for more than five weeks. We conjecture that the latter observation is due to test runs. Next, we take a closer look at Figure 6.6: We observe reoccurring probing activities at days 0-5, 11-13, and 17-21 for $icmp_1$. We note that the probing activities are not comprehensive, i.e., they only contain partial probing of the address space, which is 3.6B IPs, according to Table 6.2. Moreover, except for the first, each of those probing activities include restarts or reprobes of previous activities which is reflected in the emergence of new iterations. Overall, 87% of the IPs are probed twice ($\sim$42% are probed three times, and $\sim$10% four times). Thus, from the irregular and incomprehensive scans, as well as the restarts, we conclude that $icmp_1$ includes test runs.

How many Censuses are there?

Since the iterations overlap throughout the entire measurement period, we are not able to count how many censuses the dataset contains. Thus, we cannot verify the claims of the authors to have conducted several scans. Therefore, we elaborate on the number of censuses based on our findings so far.

If we are strict with regard to the findings in Section 6.4.2, i.e., if we require that each census contains records for all IP addresses that were probed throughout the entire measurement period, then the data can contain at most a single census. The same holds for the individual measurement periods, as we have shown in Section 6.5.1 that only the first iterations are the most complete. Thus, the number of censuses would be three, one for each measurement period. If we are satisfied with partial censuses in the sense of scanning up to 99.5% of the IP address space, then $icmp_1$ and $icmp_3$ contain one census each, while $icmp_2$ contains three censuses. Finally, if we look at each time period, and the respective numbers in Table 6.2 separately, there is room for two censuses in $icmp_1$ and $icmp_3$, as well as eight censuses in $icmp_2$. However, we need to keep in mind that $icmp_1$ probably includes test runs, and we are not able to separate and count censuses in the dataset.

6.5.2 Where are the Fast Scans?

In the previous sections we were not able to identify the censuses, or to find out how many there are in total. Given that there is at least one census in the ICMP data, in this section we try to identify the fast scans that were advertised by the authors, i.e., scans of the entire IPv4 address space completed within 24 hours. For this we resort to the "typical" approach of using a sliding window to count the number of unique IPs within 24 hours. In principle, a sliding window of 24 hours length should suffice for this analysis. However, given that the timestamps in the dataset are misaligned and scattered over almost an entire day (see Section 6.4.2), we use a 48-hour sliding window as conservative approach.

Figure 6.7 shows the number of probes per 48-hours sliding window across time. We added a supporting dotted line at 3.7B IPs. This corresponds to the number of currently allocatable IP addresses. We see that only for two measurement periods the number of probes exceeds the required number of probes, whereby the extraordinary high number

indicates overlapping iterations. Within these periods, we find that each candidate window contains probes of all 221 probed /8 prefixes. However, none of the /8s was probed completely. The number of missing IPs per /8 ranges from 2,522 IPs for the most frequently probed prefix to 1,060,415 IPs for the least frequently probed prefix. We thus conclude that we are unable to find any complete fast scans, even when we use a 48-hour sliding window. Thus, we are unable to verify the claims of the authors.

6.6 Discussion

In this section, we examine the size of the Internet, as it was determined by the authors of the Internet census report and highlight related, typical pitfalls. Also, we comment on the novelty of the census as well as the public reactions that followed the release of the datasets. We close by discussing ethical considerations and concerns.

6.6.1 Robustness of the Data

Part of the Internet census report deals with estimating the size of the Internet. As part of our validation efforts, we ask whether the proper rigor was used to estimate the size of the Internet, considering the bad data quality, e.g., due to different probing rates.

Size of the Internet

The authors of the Internet census report to have used data *"from June 2012 to October 2012"* to estimate the size of the Internet. Note that in Section 6.4.1 we have already reported that there is no data collected in June and July. Still, using such a long time range can significantly bias the results. Consider the following thought experiment: A customer uses the Internet once a day and is assigned a new IP address by its ISP every time it connects. Then, considering the five months of measurements, this single customer is responsible for 150 IP addresses. Thus, mixing incoherent measurement periods together may exaggerate the size of the Internet. However, measurement failures and probe drops may underestimate the size.

In their final remarks the authors add up numbers from all the different datasets. They assume if they have any indication that an IP address might have had any activity, then it needs to be included in the calculation of the size of the Internet. This may be problematic, as they consider an IP used, if it has a reverse DNS entry in the reverse DNS tree. However, reverse DNS entries are not mandatory to be assigned to the IP addresses by their owners. Furthermore, network operators sometimes automatically prepopulate entire (large) address ranges. This way a reverse DNS entry can be assigned to an otherwise unused IP address.

Typical Pitfalls

Getting a good grip on the size of the overall Internet is definitely an interesting research challenge. However, whether the number of IP addresses "in use" is a good proxy for the size of the Internet is debatable, since today a single IP address is rarely assigned to a single real person or machine. Rather it is an any-to-any relationship. Among the culprits for this are NAT gateways, which allow many hosts to share a single IP, Proxy

servers, load-balancers, etc. Moreover, many hosts have multiple network interfaces and may, therefore, have multiple IPs. In addition, services such as anycast and multicast are used frequently in the Internet.

Still, given the discussion of IPv6 deployment and IPv4 address exhaustion [113], knowing which IP addresses are currently in use is of interest. However, this is a highly dynamic process. One example is dynamic address assignment in residential access networks and companies. Customers are assigned an IP address when they use the Internet; once they are offline their address can be reused. Indeed, many ISPs assign a different IP address to their customers whenever they connect. Also, allocations of complete IP blocks, and their usage can change drastically, e.g., infrastructure providers can renumber each host when allocations are changed.

Moreover, IP addresses that do not respond to probes are not necessarily "unused". Rather, they may have been configured to ignore any kind of probe. In addition, some probes can be dropped along the way, e.g., due to ICMP rate limiting. Furthermore, there are some devices that are configured to respond to any probe, even if the destination IP address is actually not in use. For example, part of the nmap probes, i.e., TCP acknowledgments on port 80, did not reach the CAIDA telescope [117]. They were intercepted and replied by proxies in networks where some of the bots were located, such that some IPs from the telescope address space were reported active.

6.6.2 What's the News?

The anonymous authors announced the availability of the Internet census in March 2013 with the statement: *"This project is, to our knowledge, the largest and most comprehensive IPv4 census ever"* [151]. The media picked up these statements and claimed: *"[t]he Most Detailed, GIF-Based Map Of The Internet"* [111], *"the Most Detailed Picture of the Internet Ever"* [134], *"one of the most comprehensive surveys ever"* [51], *"remarkable academic paper"* [165], etc. What is behind this buzz? Is the Internet census 2012 really unique?

In our view what makes this census so unique is not necessarily the data itself, but rather the unethical measurement methodology. ICMP censuses have been captured by Heidemann et al. since 2003 [104]. Moreover, extensive and Internet-wide *nmap*-based port scans and service probes have been conducted in the past [80, 122, 105]. With regards to traceroutes, this Internet census dataset provides 68M records. More extensive studies have been conducted, e.g., by Shavitt et al. with 230M [152], Chen et al. with 541M [67], and Claffy et al. with 2.1B [73] sample records. Thus, what remains unique and novel is the combined dataset that was allegedly captured using a large distributed network of 420k bots. The technical contribution, however, appears to be overrated by the press.

6.6.3 Ethical Considerations

One of the most fundamental questions for researchers given the availability of the Internet census data is whether it can/should be used for publications. While answering this question is beyond the scope of this work, we raise concerns on the ethical validity of the data. While using traceroute, ping, and nmap for measurement purposes is in principle legitimate, using resource of end-users without permission is not only unorthodox but a violation of the terms of use. Thus, based on academic standards, the study as well

as the data has to be considered *unethical*.

When discussing the dataset with members of the community, we observed a diverse set of opinions ranging from *never touch such data* to *why not?* Such controversy raises the question of whether we as community need better ethical guidelines for networking and security research. While ethical standards in medical research are well defined (see e.g., the Belmont Report [15]), similar standards for Internet research are still not clearly defined. However, first steps exist: For example, the Menlo Report [53] as equivalent to the Belmont Report, or the Internet Measurement Conference. The IMC enforces adherence to its ethical standards as specified in the call for paper [29]. However, the interpretation is up to the individual PC members.

6.7 Chapter Summary

In this chapter, we studied a published dataset performed by the Carna botnet containing various measurements, among which are Internet-wide censuses, as it is claimed by the anonymous authors. Immediately after the release via BitTorrent, we participated in the corresponding swarm and observed a high interest in this dataset, predominantly by research facilities, as well as governmental and security organizations. Motivated by this high interest and the potential benefit for research, we performed a series of test to analyze its usability. Thereby, we approached this dataset from a high-level view to more detailed analyses.

We first provided an overview of the published dataset, their file organization and data structure of the individual measurements. Since not sufficiently documented by the creators of the dataset, we presented high-level statistics, e.g., the targeted address space or the number of probed unique hosts, as an entry point for a general understanding of the dataset's composition. We then reproduced some of the non-time-sensitive measurements to validate genuineness and to rule out forgery. In particular, we did a string comparison between the provided data and our own measurements of reverse-DNS entries of 70M IPs as well as HTTP responses from 4M Akamai IPs and found that the data is likely to be authentic.

Since the appealing claims made by the authors are not backed by a reasonable documentation or meta-data, we reverse-engineered as much meta-data as possible. Through detailed analyses we were able to identify three potentially separate measurement campaigns. Moreover, we found that the scans have been performed in groups of /8s, whereby the probing frequency and probing rate differs across the groups. We pointed out significant disparities presumably caused by the architecture of the botnet.

In a final effort to verify the claims of the authors, we turned our attention to the ICMP ping dataset. Due to the skewed measurements and the lack of meta-information, e.g., when does a census start, when does it end, we pragmatically grouped the ICMP probes based on the iterations each of the IPs were scanned. Since the iterations show a significant overlap we were not able to distinguish the scans, let alone count the number of individual censuses. We elaborated on the potential number of censuses based on our findings.

Finally, we provided a discussion on the usefulness of the dataset. In particular, we asked whether the proper rigor was used by the authors of the dataset in order to estimate the size of the Internet. In the process, we highlighted typical pitfalls when dealing with this and similar questions. Also, we commented on the novelty of the census by hinting at related work, but also we challenged the public reactions that followed the release of the dataset. We discussed ethical considerations concerning third-party use of datasets of unknown origin and quality in publications.

7 Conclusion

The Internet is a constantly evolving ecosystem in which traffic engineering and other routing decisions affect inter-domain traffic flow. In particular, major commercial players shape the landscape of the Internet by introducing changes to the infrastructure and engaging in complex business relationships. Thereby, economic incentives drive many of them, e.g., to deploy their servers deep within third-party networks or to use prefix delegations to perform fine-grained traffic engineering. By using a variety of global vantage points we were able to track current developments in infrastructural changes and the interactions between the responsible parties.

7.1 Summary

Heterogeneity: We reported on the existence of single, well-localized physical locations or vantage points within the Internet infrastructure where one can "see" much of the global Internet. Mining the data collected at a large IXP reveals a network that teems with heterogeneity whichever way one looks. In response to this observed heterogeneity, the work also contributes to Internet topology research by advancing a new mental model for the Internet's ecosystem that accounts for the observed network heterogenization. It also points towards measurements that reveal and keep track of this ongoing heterogenization process, and is rich and flexible enough to adapt to a constantly changing Internet environment. Doing so only scratches the surface of a new and rich problem space, and our efforts reported in this work that focus less on the Internet's connectivity structure and more on how traffic flows over this connectivity structure are just a first step towards exploring that space.

Traffic Asymmetries: Motivated by the increasing diversification of networks and inter-domain links we subsequently investigated the prevalence of traffic asymmetries. We presented a first look into Internet traffic volume asymmetries, i.e., the balance of

ingress and egress traffic, revealing the extent and consistency of these inequalities. Our analyses, based on traffic data spanning a one-week period from a Tier-1 ISP, highlighted that traffic asymmetry is largely unaffected by routing asymmetry. To a large extent, this asymmetry is characterized by the directly connected peers. We also showed that, while hypergiants significantly coin the overall traffic profile or the ISP (inbound-heavy), the traffic steering policies or routing asymmetries induced by them do not affect the traffic asymmetry more than other ASes. We concluded that traffic associated with networks is mostly asymmetric in nature.

Prefix Delegations: Since policies typically dictate the network path over which traffic flows, we consequently investigated the routing practices in the Internet. Our analysis of publicly routing information provided by BGP collectors revealed the popularity of prefix delegations. They are common-place among ASes of any size and type. By distinguishing between four delegations classes, our results highlighted that delegations are not limited to the preservation of IPv4 address space, i.e., by the delegation of PA address space. On the contrary, they indicated the delegation of address space up and across the AS level hierarchy, often with different objectives. Our analyses revealed the existence of indirect business relationships among the involved ASes, e.g., they involve CDNs which pursue their content distribution strategies. Finally, we showed that delegations significantly influence path selection, thus filtering these more specific prefixes can have a negative impact on traffic engineering strategies.

Carna: Given the discussion of IPv6 deployment and IPv4 address exhaustion, knowing which IP addresses are currently in use is of interest. The novel, but unethical way the Internet census 2012 was performed attracted many different reactions, e.g., the technical contribution that was overrated by the press, or ethical discussions that came up in the community about using the data for publications. We showed that the provided measurement data seems to be authentic, based on some spot tests. However, analyzing the quality of the data revealed qualitative problems that are caused by methodological flaws, or the lack of meta information. These problems render the data unusable for many further analyses as well as conclusions drawn from the measurements, e.g., the size of the Internet estimated by the authors and recited in the press. Finally, we believe that our analysis provided educationally useful hints about pitfalls in Internet-scale measurements and large data analysis.

7.2 Future Work

Since the Internet is a continuously evolving ecosystem, constantly tracking changes and developments are necessary. In the context of this work, we answered just a subset of questions given the large problem space we tackled. In the quest of identifying and tracking global developments and trends in the Internet, there are many challenges that need to be addressed in future work.

Given that economic incentives drive many of the main commercial Internet players to either host third-party servers in their own network infrastructures or deploy their own servers, often in massive numbers, in strategically selected third-party networks, we expect the observed trend to accelerate. Especially, in view of the growing importance of cloud providers, we believe to see increasingly heterogeneous networks and network links in the future Internet. As an interesting consequence of more servers being deployed close to the end users, we also expect that IXPs in the future will “see” less end user-to-server traffic but an increasing amount of server-to-server traffic. The resulting density of major players at single locations as well as the consolidation of the respec-

tive traffic further pushes the Internet towards a flat hierarchy, and renders transit ISPs more and more dispensable. This also has implications for business relationships. Therefore, IXPs will continue to play an important part in the future. Continuous efforts to track current developments at IXPs is key to provide insights about the fast and dynamically evolving Internet ecosystem as a whole. In particular, since we were able to classify *only* around 80% of server IPs, follow-up research at the IXP needs to improve the detection of server infrastructures and expand its focus on infrastructures serving IPv6 traffic.

By analyzing prefix delegations which are not necessarily aligned with hierarchical AS relationships, we show that prefix announcements go beyond simple reachability. Thus, supported by our historical analysis, we conclude that the routing table will continue to grow as an increasing number of announcements become unaggregatable. We also give reason to believe that delegations are an important tool in the operator scene that is not going to disappear soon, and at the same time reflects the lack of functionalities in existing protocols. The complexities in routing, in particular, due to prefix delegations, which are more complex and diverse than commonly presumed, give rise to questions regarding the routing table growth. Despite many warnings in the past, deeming the routing table growth a scalability problem, and occasional outages due to too large tables, overall the size of the routing table, which currently contains more than 700K entries, does not appear to be a threat for the functioning of inter-domain routing (anymore). Current router technology is able to handle 1M entries. Given its current growth rate, which shows no signs of slowing down, an increase of entries up to 16M, assuming a maximum prefix length of 24, does not seems impossible. Considering the complexities in routing, future work needs to (i) re-evaluate the applicability of current aggregation strategies, and those proposed by the IETF, e.g., LISP [85], (ii) the impact of IPv4 transfer markets on delegations, and (iii) a detailed analysis of prefix delegations in IPv6. Continuous analyses of the composition of the routing table constituents can inform router vendors and protocol designers about ongoing routing practices.

Today's traffic in the Internet is mostly coined by bandwidth-heavy rich media content, such as high-definition videos, and distributed by CDNs or cloud providers. The application used for delivery has a significant influence on traffic asymmetries, i.e., the balance between ingress and egress traffic. For example, while client-server applications, such as HTTP, tend to generate asymmetric traffic, peer-to-peer traffic is likely to be more symmetric. The asymmetry of traffic, in turn, has serious implications on relationships between two ASes. For example, ASes enter a peering relationship when the amount of traffic they send is more or less balanced. Otherwise they can resort to paid-peering or customer-provider relationship. Given this economical impact that applications and the resulting traffic asymmetries have of on interconnections, this can lead to complex traffic steering policies and peering disputes. This behavior is often undesired as it exacerbates debugging, traffic analysis or traffic classification. Network operators need to understand the composition of traffic in order to meet the quality expectations of their customers. As of today, many aspects about traffic asymmetries still remain unexplored. Future research in this area needs to (i) further investigate the extent of traffic asymmetries on a large scale using multiple vantage points, as well as the long-term impact of popular applications, and (ii) pursue the more complex task of tracking traffic asymmetries at the prefix-level as well as over individual links, at the link-layer. Understanding traffic asymmetries and the root causes can help network operators in network planning, provisioning, and traffic engineering.

To conclude, due to the fast and dynamic evolution of the Internet, continuous evaluation of its ecosystem, i.e., infrastructure, traffic, and routing, is necessary.

Acknowledgements

This thesis is the final chapter of an intense and rewarding journey. I would like to thank my dissertation committee members, supervisors, and collaborators for giving me this opportunity, and everybody who took this journey with me.

First and foremost, I would like to express my immense appreciation to my supervisor Anja Feldmann. From the first meeting, through the writing of papers, to the completion of my doctorate, I will always remember the time and effort you invested in me. Without your care and friendship, it would have been impossible to reach my goals. Thank you for your incredible trust. I am especially indebted to Thomas Zinner for his willingness to give his time so generously, for a door that was always open, and encouragement towards the finish line of my Ph.D. I consider myself very lucky to know you, as a supervisor, and above all, as a friend. Georgios Smaragdakis, I will always remember the day during my master studies when we were first talking about me pursuing a Ph.D. For your encouragement, support, and an inexhaustible pool of creative ideas, I would like to thank you. Thanks also to Oliver Hohlfeld with whom I had and still have the pleasure to collaborate with. And thank you, Robert Beverly, for your feedback, support, and the runs.

During my Ph.D., I never lacked the necessary tools to conduct my research. Therefore, I would also like to extend my thanks to the technical staff of INET for providing and maintaining the physical infrastructure. In particular, I would like to thank Sarah, Sabet, Simon, and Rainer for their time and support.

Also, I would like to thank all colleagues I had the pleasure to meet and work with including but not limited to Lars, Enric, Matthias, Susi, Niklas, Damien, Theresa, Philipp, Felix, Thorben, Mirko; thanks for many enjoyable days in and around INET. Moreover, I would like to thank Ingmar Poese, my very first supervisor during my studies, for his good heart and dark souls. I am particularly grateful for the assistance given by Stefan Wahl by providing me with technical know-how and feedback, and Tobias Fiebig for giving me the right push at the right moment. I would also like to extend my thanks to all the students that assisted me during my Ph.D.

Finally, I would like to express my deepest gratitude to my family and friends, for always staying behind my decisions. Dear parents, none of this would have ever happened without your courageous decisions in life. Thank you! Anna and Rainer, thanks for taking my mind off things during my visits and the fun times we enjoyed together. A huge thank you goes to my friends Holger, Jan, Antje, Vivi, Huy, Pascal, Alex, Toppa, Marco, Falko, and Virginia, for bearing with me over the last couple of years. I am happy to have you all in my life. I would like to especially thank Oliver K. for his helpful support towards the defense, Tobias J. and Sebastian, friends and fellow students with whom it would have not been possible to finish my studies. And finally, Rachel for her continuous support during tough times, and the entire Makkabi soccer team for the fun Fridays.

There is no one mentioned above that did not contribute to the person I am today. Thank you all.

I am grateful to the funding received through the Leibniz Price project funds of DFG/ German Research Foundation (FKZ FE 570/4-1).

Bibliography

[1] Akamai and AT&T Forge Global Strategic Alliance to Provide Content Delivery Network Solutions. http://www.akamai.com/html/about/press/releases/2012/press_120612.html. [Last accessed: November 4, 2018].

[2] Amazon CloudFront - Amazon Web Services. http://aws.amazon.com/cloudfront/. [Last accessed: November 4, 2018].

[3] Caida. Archipelago measurement infrastructure. http://www.caida.org/projects/ark/. [Last accessed: November 4, 2018].

[4] Caida. Mapping Autonomous Systems to Organizations. https://www.caida.org/research/topology/as2org/. [Last accessed: November 4, 2018].

[5] CIDR Report. http://www.cidr-report.org/.

[6] Like Netflix, Facebook is boosting its edge network. http://gigaom.com/2012/06/21/like-netflix-facebook-is-planning-its-own-cdn/. [Last accessed: November 4, 2018].

[7] Netflix Open Connect. https://signup.netflix.com/openconnect. [Last accessed: November 4, 2018].

[8] Orange and Akamai form Content Delivery Strategic Alliance. http://www.akamai.com/html/about/press/releases/2012/press_112012_1.html. [Last accessed: November 4, 2018].

[9] PeeringDB. https://www.peeringdb.com. [Last accessed: October 16, 2018].

[10] RIS - RIPE Network Coordination Centre. http://ris.ripe.net/. [Last accessed: November 4, 2018].

[11] Team Cymru. http://www.team-cymru.org/. [Last accessed: November 4, 2018].

[12] University of Oregon RouteViews project. http://www.routeviews.org/. [Last accessed: November 4, 2018].

[13] Assigned numbers. RFC 750, Sept. 1978.

[14] Gateway Routing - An Implementation Specification. https://www.rfc-editor.org/ien/ien30.pdf, 1978. [Last accessed: October 31, 2018].

[15] Ethical Principles and Guidelines for the Protection of Human Subjects of Research. http://www.hhs.gov/ohrp/humansubjects/guidance/belmont.html, 1979. [Last accessed: November 4, 2018].

[16] DoD standard Internet Protocol. RFC 760, Jan. 1980.

[17] Assigned numbers. RFC 790, Sept. 1981.

[18] Internet Protocol. RFC 791, Sept. 1981.

[19] Assigned numbers. RFC 820, Aug. 1982.

[20] Computer mail meeting notes. RFC 805, Feb. 1982.

[21] Distributed system for Internet name service. RFC 830, Oct. 1982.

[22] Exterior Gateway Protocol (EGP). RFC 827, Oct. 1982.

[23] The Domain Naming Convention for Internet User Applications. RFC 819, Aug. 1982.

[24] Domain requirements. RFC 920, Oct. 1984.

[25] Domain names - implementation and specification. RFC 1035, Nov. 1987.

[26] Comcast Throttles BitTorrent Traffic, Seeding Impossible. https://torrentfreak.com/comcast-throttles-bittorrent-traffic-seeding-impossible/, 2007. [Last accessed: October 31, 2018].

[27] ipoque Internet Study 2007: P2P File Sharing Still Dominates the Worldwide Internet. https://www.ipoque.com/news-media/press-releases/2007/ipoque-internet-study-2007-p2p-file-sharing-still-dominates, 2007. [Last accessed: October 31, 2018].

[28] Free Pool of IPv4 Address Space Depleted. https://www.nro.net/ipv4-free-pool-depleted/, 2011. [Last accessed: November 4, 2018].

[29] IMC 2013 Call for Papers. http://conferences.sigcomm.org/imc/2013/cfp.html, 2013. [Last accessed: November 4, 2018].

[30] Internet Census 2012 Search. http://www.exfiltrated.com/querystart.php, 2013. [Last accessed: November 4, 2018].

[31] BGP Routing Table Size Limit Blamed for Tuesday's Website Outages. https://www.datacenterknowledge.com/archives/2014/08/13/bgp-routing-table-size-limit-blamed-for-tuesdays-website-outages, 2014. [Last accessed: October 31, 2018].

[32] The Long, Slow Decline of BitTorrent. https://www.plagiarismtoday.com/2017/06/01/the-long-slow-decline-of-bittorrent/, 2017. [Last accessed: October 31, 2018].

[33] Bad ISPs. http://wiki.vuze.com/w/Bad_ISPs, 2018. [Last accessed: October 31, 2018].

[34] Global Internet Phenomena Preview: File sharing on the internet reverses a downward trend. https://www.sandvine.com/blog/global-internet-phenomena-preview-file-sharing-reverses-a-downward-trend, 2018. [Last accessed: October 31, 2018].

[35] List of Top-Level Domains. https://www.icann.org/resources/pages/tlds-2012-02-25-en, 2018. [Last accessed: October 31, 2018].

[36] State of IPv6 Deployment 2018. https://www.internetsociety.org/resources/2018/state-of-ipv6-deployment-2018/, 2018. [Last accessed:

November 4, 2018].

[37] The Verisign Domain Name Industry Brief. https://www.verisign.com/en_US/domain-names/dnib/index.xhtml, 2018. [Last accessed: October 31, 2018].

[38] V. Adhikari, Y. Guo, F. Hao, M. Varvello, V. Hilt, M. Steiner, and Z.-L. Zhang. Unreeling Netflix: Understanding and Improving multi-CDN Movie Delivery. In *IEEE INFOCOM*, 2012.

[39] V. K. Adhikari, S. Jain, Y. Chen, and Z.-L. Zhang. Vivisecting YouTube: An Active Measurement Study. In *IEEE INFOCOM*, 2012.

[40] P. Aditya, M. Zhao, Y. Lin, A. Haeberlen, P. Druschel, B. Maggs, and B. Wishon. Reliable Client Accounting for Hybrid Content-Distribution Networks. In *NSDI*, 2012.

[41] B. Ager, N. Chatzis, A. Feldmann, N. Sarrar, S. Uhlig, and W. Willinger. Anatomy of a Large European IXP. In *ACM SIGCOMM*, 2012.

[42] B. Ager, W. Mühlbauer, G. Smaragdakis, and S. Uhlig. Web Content Cartography. In *ACM IMC*, 2011.

[43] V. Aggarwal, A. Feldmann, and C. Scheideler. Can ISPs and P2P Users Cooperate for Improved Performance? In *ACM SIGCOMM CCR*, 2007.

[44] Akamai. Facts and Figures: Network Deployment. http://www.akamai.com/html/about/facts_figures.html. [Last accessed: November 4, 2018].

[45] Amazon. AWS Dashboard. http://status.aws.amazon.com/. [Last accessed: November 4, 2018].

[46] Amazon. EC2 Public IP ranges. https://web.archive.org/web/20130116082028/https://forums.aws.amazon.com/ann.jspa?annID=1701. [Last accessed: November 4, 2018].

[47] AMS-IX. AMS-IX hosts BGP-Mux research project. https://www.ams-ix.net/newsitems/82. [Last accessed: November 4, 2018].

[48] Anonymous. Carna Botnet Scanning of all IPv4 Addresses. https://web.archive.org/web/20130503090242/http://www.auscert.org.au/render.html?it=17258, 2013. [Last accessed: November 4, 2018].

[49] Anonymous. Internet Census 2012: Port scanning /0 using insecure embedded devices. https://web.archive.org/web/20130402235845/http://internetcensus2012.bitbucket.org/paper.html, 2013. [Last accessed: November 4, 2018].

[50] D. Antoniades, E. Markatos, and C. Dovrolis. One-click Hosting Services: A File-Sharing Hideout. In *ACM IMC*, 2009.

[51] arstechnica.com. Guerilla researcher created epic botnet to scan billions of IP addresses. With 9TB of data, survey is one of the most exhaustive—and illicit—ever done. https://arstechnica.com/security/2013/03/guerilla-researcher-created-epic-botnet-to-scan-billions-of-ip-addresses/, 2013. [Last accessed: November 4, 2018].

[52] B. Augustin, B. Krishnamurthy, and W. Willinger. IXPs: Mapped? In *ACM IMC*, 2009.

[53] M. Bailey, D. Dittrich, E. Kenneally, and D. Maughan. The Menlo Report. In *IEEE Security & Privacy*, 2012.

[54] I. Bermudez, M. Mellia, M. Munafà, R. Keralapura, and A. Nucci. DNS to the Rescue: Discerning Content and Services in a Tangled Web. In *ACM IMC*, 2012.

[55] T. Böttger, F. Cuadrado, and S. Uhlig. Looking for Hypergiants in PeeringDB. In *ACM SIGCOMM CCR*, 2018.

[56] T. Bu, L. Gao, and D. Towsley. On characterizing BGP routing table growth. *Computer Networks*, 2004.

[57] M. Caesar and J. Rexford. BGP routing policies in ISP networks. *IEEE Network*, 2005.

[58] X. Cai, J. Heidemann, B. Krishnamurthy, and W. Willinger. Towards an AS-to-Organization Map. In *ACM IMC*, 2010.

[59] M. Calder, X. Fan, Z. Hu, E. Katz-Bassett, J. Heidemann, and R. Govindan. Mapping the Expansion of Google's Serving Infrastructure. In *ACM IMC*, 2013.

[60] C. Cardona, P. Francois, and P. Lucente. Impact of BGP Filtering on Inter-Domain Routing Policies. RFC 7789, Apr. 2016.

[61] V. G. Cerf. IAB recommended policy on distributing internet identifier assignment and IAB recommended policy change to internet "connected" status. RFC 1174, Aug. 1990.

[62] M. Cha, H. Kwak, P. Rodriguez, Y. Y. Ahn, and S. Moon. I Tube, You Tube, Everybody Tubes: Analyzing the World's Largest User Generated Content Video System. In *ACM IMC*, 2008.

[63] B. Chandrasekaran, M. Bai, M. Schoenfield, A. Berger, N. Caruso, G. Economou, S. Gilliss, B. Maggs, K. Moses, D. Duff, K.-C. Ng, E. G. Sirer, R. Weber, and B. Wong. Alidade: IP Geolocation without Active Probing. Technical Report CS-TR-2015-001, Duke University, January 2015.

[64] H. Chang, S. Jamin, Z. M. Mao, and W. Willinger. An Empirical Approach to Modeling Inter-AS Traffic Matrices. In *ACM IMC*, 2005.

[65] L. Chapin, D. D. Clark, R. T. Braden, R. Hobby, and V. G. Cerf. Towards the Future Internet Architecture. RFC 1287, Dec. 1991.

[66] N. Chatzis, G. Smaragdakis, A. Feldmann, and W. Willinger. There is More to IXPs Than Meets the Eye. In *ACM SIGCOMM CCR*, 2013.

[67] K. Chen, D. R. Choffnes, R. Potharaju, Y. Chen, F. E. Bustamante, D. Pei, and Y. Zhao. Where the Sidewalk ends: Extending the Internet as Graph using Traceroutes from P2P users. In *ACM CoNEXT*, 2009.

[68] K. Cho, C. Pelsser, R. Bush, and Y. Won. The Japan Earthquake: the Impact on Traffic and Routing Observed by a Local ISP. In *ACM SWID*, 2011.

[69] D. R. Choffnes and F. E. Bustamante. Taming the Torrent: A Practical Approach to Reducing Cross-ISP Traffic in Peer-To-Peer Systems. In *ACM SIGCOMM*, 2008.

[70] Cisco. Visual Networking Index (VNI) and Forecast. http://www.cisco.com/en/US/netsol/ns827/networking_solutions_sub_solution.html. [Last accessed: November 4, 2018].

[71] Cisco. The Zettabyte Era: Trends and Analysis. https://bit.ly/2h3jXbJ, June 2017. [Last accessed: October 16, 2018].

[72] L. Cittadini, W. Muhlbauer, S. Uhlig, R. Bush, P. Francois, and O. Maennel. Evolution of Internet Address Space Deaggregation: Myths and Reality. In *IEEE J-SAC*, 2010.

[73] k. Claffy, Y. Hyun, K. Keys, M. Fomenkov, and D. Krioukov. Internet Mapping: from Art to Science. In *IEEE CATCH*, 2009.

[74] M. Crovella and B. Krishnamurthy. *Internet Measurement: Infrastructure, Traffic and Applications*. John Wiley & Sons, Inc., New York, NY, USA, 2006.

[75] Daily Mail. This is what the internet looks like: Spectacular image created by computer hacker captures every iota of online data for 'day in the life' of the web. http://www.dailymail.co.uk/sciencetech/article-2299936/This-internet-looks-like-Spectacular-image-created-hacker-captures-iota-online-data-day-life-web.html, 2013. [Last accessed: November 4, 2018].

[76] D. Dainotti, C. Squarcella, E. Aben, K. Claffy, M. Chiesa, M. Russo, and A. Pescapé. Analysis of Country-Wide Internet Outages Caused by Censorship. In *ACM IMC*, 2011.

[77] Data Center Knowledge. Who Has the Most Web Servers? http://www.datacenterknowledge.com/archives/2009/05/14/whos-got-the-most-web-servers. [Last accessed: November 4, 2018].

[78] G. J. de Groot, D. Karrenberg, Y. Rekhter, and R. Moskowitz. Address Allocation for Private Internets. RFC 1597, Mar. 1994.

[79] W. de Vries, J. J. Santanna, A. Sperotto, and A. Pras. How Asymmetric Is the Internet? In *AIMS*, 2015.

[80] P. Eckersley and J. Burns. An observatory for the SSLiverse. https://www.eff.org/files/DefconSSLiverse.pdf, 2010. [Last accessed: November 4, 2018].

[81] K. B. Egevang and P. Francis. The IP Network Address Translator (NAT). RFC 1631, May 1994.

[82] J. Erman, A. Gerber, M. Hajiaghayi, D. Pei, and O. Spatscheck. Network-aware Forward Caching. In *WWW*, 2009.

[83] J. B. F. Streibelt, N. Chatzis, G. Smaragdakis, and A. Feldmann. Exploring EDNS-Client-Subnet Adopters in your Free Time. In *ACM IMC*, 2013.

[84] Facebook blog. A Continued Commitment to Security. https://web.archive.org/web/20121213155827/http://www.facebook.com/blog/blog.php?post=486790652130. [Last accessed: November 4, 2018].

[85] D. Farinacci, V. Fuller, D. Meyer, and D. Lewis. The Locator/ID Separation Protocol (LISP). RFC 6830, Jan. 2013.

[86] T. Flach, N. Dukkipati, A. Terzis, B. Raghavan, N. Cardwell, Y. Cheng, A. Jain, S. Hao, E. Katz-Bassett, and R. Govindan. Reducing Web Latency: the Virtue of Gentle Aggression. In *ACM SIGCOMM*, 2013.

[87] V. Fuller and T. Li. Classless Inter-domain Routing (CIDR): The Internet Address Assignment and Aggregation Plan. RFC 4632, Aug. 2006.

[88] V. Fuller, T. Li, K. Varadhan, and J. Yu. Classless Inter-Domain Routing (CIDR): an Address Assignment and Aggregation Strategy. RFC 1519, Sept. 1993.

[89] L. Gao. On Inferring Autonomous System Relationships in the Internet. In *IEEE/ACM ToN*, 2001.

[90] L. Gao and F. Wang. The Extent of AS Path Inflation by Routing Policies. In *IEEE GLOBECOM*, 2002.

[91] A. Gerber and R. Doverspike. Traffic Types and Growth in Backbone Networks. In *OFC/NFOEC*, 2011.

[92] E. P. Gerich. Guidelines for Management of IP Address Space. RFC 1366, Oct. 1992.

[93] E. P. Gerich. Guidelines for Management of IP Address Space. RFC 1466, May 1993.

[94] P. Gill, M. F. Arlitt, Z. Li, and A. Mahanti. Youtube Traffic Characterization: A View From the Edge. In *ACM IMC*, 2007.

[95] V. Giotsas, C. Dietzel, G. Smaragdakis, A. Feldmann, A. Berger, and E. Aben. Detecting Peering Infrastructure Outages in the Wild. In *ACM SIGCOMM*, 2013.

[96] V. Giotsas, M. Luckie, B. Huffaker, and k. claffy. Inferring Complex AS Relationships. In *ACM IMC*, 2014.

[97] V. Giotsas, G. Smaragdakis, B. Huffaker, M. Luckie, and k. claffy. Mapping Peering Interconnections to a Facility. In *ACM CoNEXT*, 2015.

[98] V. Giotsas, S. Zhou, M. Luckie, and k. claffy. Inferring Multilateral Peering. In *ACM CoNEXT*, 2013.

[99] GlobalDots. Multi Content Delivery Network Explained. https://www.globaldots.com/multi-content-delivery-network-explained/, 2017. [Last accessed: November 4, 2018].

[100] Google. What is 1e100.net? http://support.google.com/bin/answer.py?hl=en&answer=174717. [Last accessed: November 4, 2018].

[101] Google Search Blog. Making Search More Secure. http://googleblog.blogspot.com/2011/10/making-search-more-secure.html. [Last accessed: November 4, 2018].

[102] E. Gregori, A. Improta, L. Lenzini, L. Rossi, and L. Sani. On the Incompleteness of the AS-level Graph: A Novel Methodology for BGP Route Collector Placement. In *ACM IMC*, 2012.

[103] Y. He, M. Faloutsos, S. Krishnamurthy, and B. Huffaker. On routing asymmetry in the Internet. In *IEEE GLOBECOM*, 2005.

[104] J. Heidemann, Y. Pradkin, R. Govindan, C. Papadopoulos, G. Bartlett, and J. Bannister. Census and Survey of the Visible Internet. In *ACM IMC*, 2008.

[105] N. Heninger, Z. Durumeric, E. Wustrow, and J. A. Halderman. Mining your Ps and Qs: Detection of Widespread Weak Keys in Network Devices. In *USENIX Security Symposium*, 2012.

[106] M. Holdrege and P. Srisuresh. IP Network Address Translator (NAT) Terminology and Considerations. RFC 2663, Aug. 1999.

[107] R. Housley, J. Curran, G. Huston, and D. R. Conrad. The Internet Numbers Registry System. RFC 7020, Aug. 2013.

[108] C.-H. Hsu and M. Hefeeda. ISP-friendly Peer Matching Without ISP Collaboration. In *ACM CoNEXT*, 2008.

[109] C. Huang, A. Wang, J. Li, and K. Ross. Measuring and Evaluating Large-scale CDNs. In *ACM IMC*, 2008.

[110] K. Hubbard, D. J. Postel, M. A. Kosters, D. Karrenberg, and D. R. Conrad. Internet Registry IP Allocation Guidelines. RFC 2050, Nov. 1996.

[111] Huffington Post. The Most Detailed, GIF-Based Map Of The Internet Was Made By Hacking 420,000 Computers. http://www.huffingtonpost.com/2013/03/22/internet-map_n_2926934.html, 2013. [Last accessed: November 4, 2018].

[112] C. H. Huitema. *Routing in the Internet*. Prentice-Hall, 2000.

[113] IANA. The Allocation of Internet Protocol Version 4 (IPv4) Address Space. http://www.iana.org/assignments/ipv4-address-space/ipv4-address-space.xml, 2013. [Last accessed: November 4, 2018].

[114] W. John, M. Dusi, and K. C. Claffy. Estimating Routing Symmetry on Single Links by Passive Flow Measurements. In *ACM IWCMC*, 2010.

[115] C. Kalogiros, M. Bagnulo, and A. Kostopoulos. Understanding Incentives for Prefix Aggregation in BGP. In *Workshop on Re-architecting the Internet*, 2009.

[116] A. Khan, H.-c. Kim, T. Kwon, and Y. Choi. A Comparative Study on IP Prefixes and Their Origin Ases in BGP and the IRR. In *ACM SIGCOMM CCR*, 2013.

[117] A. King and A. Dainotti. Carna botnet scans confirmed. http://blog.caida.org/best_available_data/2013/05/13/carna-botnet-scans/, 2013. [Last accessed: November 4, 2018].

[118] B. Krishnamurthy, W. Willinger, P. Gill, and M. Arlitt. A Socratic Method for Validation of Measurement-Based Networking Research. *Computer Communications*, 2011.

[119] J. F. Kurose and K. W. Ross. *Computer Networking: A Top-Down Approach (6th Edition)*. Pearson, 6th edition, 2012.

[120] C. Labovitz, S. Iekel-Johnson, D. McPherson, J. Oberheide, and F. Jahanian. Internet Inter-Domain Traffic. In *ACM SIGCOMM*, 2010.

[121] T. Leighton. Improving Performance on the Internet. *Communications of the ACM*, 2009.

[122] D. Leonard and D. Loguinov. Demystifying Service Discovery: Implementing an Internet-Wide Scanner. In *ACM IMC*, 2010.

[123] T. Li, R. Chandra, and P. S. Traina. BGP Communities Attribute. RFC 1997, Aug. 1996.

[124] F. Lichtblau, F. Streibelt, T. Krüger, P. Richter, and A. Feldmann. Detection, Classification, and Analysis of Inter-domain Traffic with Spoofed Source IP Addresses. In *ACM IMC*, 2017.

[125] I. Livadariu, A. Elmokashfi, A. Dhamdhere, and k. claffy. A First Look at IPv4

Transfer Markets. In *ACM CoNEXT*, 2013.

[126] A. Lodhi, N. Larson, A. Dhamdhere, C. Dovrolis, and k. claffy. Using peeringDB to Understand the Peering Ecosystem. In *ACM SIGCOMM CCR*, 2014.

[127] M. Luckie, B. Huffaker, A. Dhamdhere, V. Giotsas, and k. claffy. AS Relationships, Customer Cones, and Validation. In *ACM IMC*, 2013.

[128] G. Maier, A. Feldmann, V. Paxson, and M. Allman. On Dominant Characteristics of Residential Broadband Internet Traffic. In *ACM IMC*, 2009.

[129] Maxmind. GeoLite Country. http://dev.maxmind.com/geoip/legacy/geolite. [Last accessed: November 4, 2018].

[130] P. McDaniel, W. Aiello, K. Butler, and J. Ioannidis. Origin Authentication in Interdomain Routing. *Computer Networks*, 2006.

[131] X. Meng, Z. Xu, B. Zhang, G. Huston, S. Lu, and L. Zhang. IPv4 Address Allocation and the BGP Routing Table Evolution. *ACM SIGCOMM CCR*, 2005.

[132] R. Moskowitz, D. Karrenberg, Y. Rekhter, E. Lear, and G. J. de Groot. Address Allocation for Private Internets. RFC 1918, Feb. 1996.

[133] R. Motamedi, B. Chandrasekaran, B. M. Maggs, R. Rejaie, and W. Willinger. On the geography of X-Connects. Technical Report CIS-TR-2014-1, Oregon University, 2014.

[134] Motherboard. This Is the Most Detailed Picture of the Internet Ever (and Making it Was Very Illegal). https://web.archive.org/web/20130807145010/http://motherboard.vice.com/blog/this-is-most-detailed-picture-internet-ever, 2013. [Last accessed: November 4, 2018].

[135] Mozilla Foundation. PublicSuffix.org. http://publicsuffix.org/. [Last accessed: November 4, 2018].

[136] Netcraft. January 2013 Web Server Survey. http://news.netcraft.com/archives/2013/01/07/january-2013-web-server-survey-2.html. [Last accessed: November 4, 2018].

[137] Netflix Nordics blog. Netflix Launches Today in Sweden, Denmark, Norway, Finland. http://nordicsblog.netflix.com/2012/10/. [Last accessed: 2012].

[138] E. Nygren, R. K. Sitaraman, and J. Sun. The Akamai Network: A Platform for High-performance Internet Applications. *ACM SIGOPS Operating Systems Review*, 2010.

[139] V. Paxson. End-to-end routing behavior in the Internet. In *IEEE/ACM ToN*, 1997.

[140] V. Paxson. Bro: A System for Detecting Network Intruders in Real-Time. In *USENIX Security Symposium*, 1998.

[141] D. Plonka and P. Barford. Flexible Traffic and Host Profiling via DNS Rendezvous. In *SATIN*, 2011.

[142] I. Poese, S. Uhlig, M. A. Kaafar, B. Donnet, and B. Gueye. IP Geolocation Databases: Unreliable? In *ACM SIGCOMM CCR*, 2011.

[143] Private communication.

[144] Y. Rekhter, S. Hares, and T. Li. A Border Gateway Protocol 4 (BGP-4). RFC 4271,

Jan. 2006.

[145] P. Richter, G. Smaragdakis, D. Plonka, and A. Berger. Beyond Counting: New Perspectives on the Active IPv4 Address Space. In *ACM IMC*, 2016.

[146] M. Roughan, S. J. Tuke, and O. Maennel. Bigfoot, Sasquatch, the Yeti and Other Missing Links: What We Don'T Know About the As Graph. In *ACM IMC*, 2008.

[147] M. Roughan, W. Willinger, O. Maennel, D. Perouli, and Y. Bush. 10 Lessons from 10 Years of Measuring and Modeling the Internet's Autonomous Systems. In *IEEE J-SAC*, 2011.

[148] J. H. Saltzer, D. P. Reed, and D. D. Clark. End-to-end Arguments in System Design. *ACM Trans. Computer Systems*, 1984.

[149] Sandvine. Global Internet Phenomena Report. https://web.archive.org/web/20130602133924/http://www.sandvine.com/news/global_broadband_trends.asp. [Last accessed: November 4, 2018].

[150] B. Schlinker, H. Kim, T. Cui, E. Katz-Bassett, H. V. Madhyastha, I. Cunha, J. Quinn, S. Hasan, P. Lapukhov, and H. Zeng. Engineering Egress with Edge Fabric: Steering Oceans of Content to the World. In *ACM SIGCOMM*, 2017.

[151] seclists.org. Port Scanning /0 Using Insecure Embedded Devices. http://seclists.org/fulldisclosure/2013/Mar/166, 2013. [Last accessed: November 4, 2018].

[152] Y. Shavitt and E. Shir. DIMES: let the Internet measure itself. *ACM SIGCOMM CCR*, 2005.

[153] J. a. L. Sobrinho, L. Vanbever, F. Le, and J. Rexford. Distributed Route Aggregation on the Global Network. In *ACM CoNEXT*, 2014.

[154] Spiegel Online International. Mapping the Internet: A Hacker's Secret Internet Census. http://www.spiegel.de/international/world/hacker-measures-the-internet-illegally-with-carna-botnet-a-890413.html, 2013. [Last accessed: November 4, 2018].

[155] N. Spring, R. Mahajan, and T. Anderson. The Causes of Path Inflation. In *ACM SIGCOMM*, 2003.

[156] A. Sriraman, K. R. Butler, P. D. McDaniel, and P. Raghavan. Analysis of the IPv4 Address Space Delegation Structure. In *Computers and Communications*, 2007.

[157] A. Su, D. Choffnes, A. Kuzmanovic, and F. Bustamante. Drafting Behind Akamai. In *ACM SIGCOMM*, 2006.

[158] A. S. Tanenbaum and D. J. Wetherall. *Computer Networks*. Prentice Hall Press, Upper Saddle River, NJ, USA, 5th edition, 2010.

[159] H. Tangmunarunkit, R. Govindan, S. Shenker, and D. Estrin. The Impact of Routing Policy on Internet Paths. In *IEEE INFOCOM*, 2001.

[160] D. Tappan, S. S. Ramachandra, and Y. Rekhter. BGP Extended Communities Attribute. RFC 4360, Feb. 2006.

[161] The Register. Researcher sets up illegal 420,000 node botnet for IPv4 internet map. http://www.theregister.co.uk/2013/03/19/carna_botnet_ipv4_internet_map/, 2013. [Last accessed: November 4, 2018].

[162] S. Triukose, Z. Al-Qudah, and M. Rabinovich. Content Delivery Networks: Protection or Threat? In *Computer Security – ESORICS*, 2009.

[163] S. Triukose, Z. Wen, and M. Rabinovich. Measuring a Commercial Content Delivery Network. In *WWW*, 2011.

[164] J. Weil, V. Kuarsingh, C. Donley, C. Liljenstolpe, and M. Azinger. IANA-Reserved IPv4 Prefix for Shared Address Space. RFC 6598, Apr. 2012.

[165] wired.co.uk. Botnet-generated map of Internet gathered data 'unethically'. https://web.archive.org/web/20131005232409/http://www.wired.co.uk/news/archive/2013-05/16/internet-census, 2013. [Last accessed: November 4, 2018].

[166] Yahoo News. Watch 24 hours of internet activity around the world in 8 seconds. http://news.yahoo.com/watch-24-hours-internet-activity-around-world-8-113700364.html, 2013. [Last accessed: November 4, 2018].

[167] L. Yang, T. A. Anderson, R. Gopal, and R. Dantu. Forwarding and Control Element Separation (ForCES) Framework. RFC 3746, Apr. 2004.

[168] K.-K. Yap, M. Motiwala, J. Rahe, S. Padgett, M. Holliman, G. Baldus, M. Hines, T. Kim, A. Narayanan, A. Jain, V. Lin, C. Rice, B. Rogan, A. Singh, B. Tanaka, M. Verma, P. Sood, M. Tariq, M. Tierney, D. Trumic, V. Valancius, C. Ying, M. Kallahalla, B. Koley, and A. Vahdat. Taking the Edge off with Espresso: Scale, Reliability and Programmability for Global Internet Peering. In *ACM SIGCOMM*, 2017.

[169] J. Yu, T. Li, K. Varadhan, and V. Fuller. Supernetting: an Address Assignment and Aggregation Strategy. RFC 1338, June 1992.

[170] ZDNet. What 420,000 insecure devices reveal about Web security. http://news.cnet.com/8301-1009_3-57574919-83/what-420000-insecure-devices-reveal-about-web-security/, 2013.