On Measuring Internet Topology & its Applications

Area Exam Report
March - 2019
Bahador Yeganeh

***Abstract*—Given the importance of the Internet, it is crucial to assess its key characteristics (e.g. performance, stability, and resiliency) through measurement as it expands and evolves over time. Measuring different characteristics of the Internet is challenging mainly due to its scale and heterogeneity. Capturing and characterizing Internet topology offers the critical insight not only for understanding the physical infrastructure of the Internet but also for examining the impact a wide range of more subtle characteristics that depend on the topology such as routing, end-to-end performance, and resiliency to attacks or disruptions.**

This area exam reviews a large body of recent studies on capturing and characterizing various aspects of the Internet topology as well as studies that explore implications of Internet topology on other real-world problems. To this end, we organize the prior studies on Internet topology based on their considered resolution into four groups as follows: (i) AS-level, (ii) router-level, (iii) PoP-level, and (iv) physical-level. For each group of studies, we discuss proper measurement tools and techniques, common datasets, relevant characteristics, related challenges and main findings at that resolution. We also broadly categorize studies on the implications of Internet topology based on whether they focus on performance, resiliency, or network peering relationship aspects of the Internet. We primarily describe how topology information with a particular scope and specific resolution serve as input to study more subtle aspects of the Internet. Finally, we present how the increasing popularity of cloud services in recent years have led to significant changes in Internet topology that motivate further measurement-based studies.

I. Introduction

The Internet since its inception as a network for interconnecting a handful of academic and military networks has gone through constant evolution throughout the years and has become a large scale distributed network spanning the globe that is intertwined with every aspect of our daily lives. Given its importance, we need to study its health, vulnerability, and connectivity. This is only made possible through constant network measurements. Researchers have conducted measurements in order to gain a better understanding of traffic routing through this network, its connectivity structure as well as its performance. Our interest and ability to conduct network measurements can vary in both scopes with respect to the number or size of networks under study as well as the resolution with regards to focusing on networks as a single unit or paying attention to finer network elements such as routers.

The topology of the Internet is a key enabler for studying routing of traffic in addition to gaining a better understanding of Internet performance and resiliency. Capturing Internet topology is challenging due to many factors namely, (i) scale: the vast scale of the Internet as a network spanning the globe limits our abilities to fully capture its structure, (ii) visibility: our view of the Internet is constrained to the perspective that we are able to glean from the limited number of vantage points we are able to look at it, (iii) dynamic: the Internet as an ever-evolving entity is under constant structural change added to this the existence of redundant routes, backup links, and load-balanced paths limits our ability to fully capture the current state of the Internet's topology, and (iv) tools: researchers have relied on tools which were originally designed for troubleshooting purposes the protocol stack of Internet lacks any inherent methods for identifying topology.

Despite these challenges, for the past couple of decades, the network measurement community have collected data, devised tools, and expanded their test beds to infer new information and conduct measurements at different scale and resolutions. The obtained insight from these studies has informed network designers, engineers, ISPs, and application developers to address issues on the performance, resiliency, and scalability of the Internet.

This area exam explores a collection of prior studies for various aspects of Internet measurement to gain insight into the topology of the Internet as well as its implications in designing applications. For Internet measurement, we focus on recent studies regarding the simulation and characterization of Internet topology. Furthermore, we organize these studies based on the resolution of the uncovered topology with an emphasis on the utilized datasets and employed methodologies. On the second part, we focus on various implications of Internet topology on the design and performance of applications. These studies are organized in accordance with the implication of topology on performance or resiliency of the Internet. Furthermore we emphasis on how various resolutions of Internet topology allow researchers to conduct different studies. The collection of these studies present a handful of open and interesting problems regarding the future of Internet topology with the advent of cloud providers and their centrality within today's Internet.

The rest of the document is organized as follows. First, in Section II we present a primer on the Internet and introduce the reader with a few taxonomies that are frequently used within this document. Second, an overview of most common datasets, platforms, and tools which are used for topology discovery is given in Section III. Third, the review for recent studies on Internet topology discovery is presented in Section IV. Forth, Section V covers the recent studies which utilize Internet topologies to study the performance and resiliency of the Internet. Lastly, we explore a few open problems and possible venues for further research in Section VI.

II. Background

The Internet is a globally federated network composed of many networks each of which has complete autonomy over the structure and operation of its own network. These autonomous systems or networks (AS) can be considered as the building blocks of the Internet. Each AS represents a virtual entity and can be composed of a vast network infrastructure composed of networking equipment like routers and switches as well as transit mediums such as Ethernet and fiber optic cables. These ASes can serve various purposes such as providing transit or connectivity for other networks, generating or offering content such as video streams, or merely represent the network of an enterprise. Each of the connectivity provider ASes can be categorized into multiple tiers based on their size and how they are interconnected with other ASes. These tiers create a natural hierarchy of connectivity that is broadly composed of 3 tiers namely, (i) Tier-1: an AS that can reach all other networks without the need to pay for its traffic exchanges, (ii) Tier-2: an AS which can have some transit-free relations with other ASes while still needing to pay for transit for reachability to some portion of the Internet, and (iii) Tier-3: an AS that solely purchases transit for connectivity to the Internet. While each network has full control over its own internal network and can deliver data from one internal node to another, transmitting data from one AS to another requires awareness of a path that can reach the destination AS. This problem is solved by having each AS advertise its own address space to neighboring ASes through the border gateway protocol (BGP). Upon receiving a BGP announcement, each AS would prepend its own AS number (ASN) to the AS-path attribute of this announcement and advertise this message to its own neighbors. This procedure allows ASes to learn about other networks and the set of AS-paths or routes that they can be reached through. ASes can interconnect with each by linking their border routers at one or multiple physical locations. These border routers are responsible for advertising their prefixes in addition to performing the actual routing of traffic within the Internet. The border routers of ASes are placed within colocation facilities (colo) that offer space, power, security, and networking equipment to the tenants ASes. Each AS can have a physical presence in multiple metro areas. The collection of their routers within each of these metro areas are referred to as the points of presence (PoP) for these ASes. Figure 1 presents a high level abstraction of the aforementioned concepts. The figure consists of 3 ASes namely, AS_A, AS_B, and AS_C in red, blue, and green accordingly. The internal structure ASes is abstracted out presenting only the border routers of each AS. AS_A and AS_B have two PoPs one in LA and another in NY while AS_C is only present in NY. AS_A and AS_B establish a private interconnection with each other through their LA PoP within $colo_1$ while they peer with each other as well as AS_C in their NY PoP in $colo_2$ through an IXPs switching fabric.

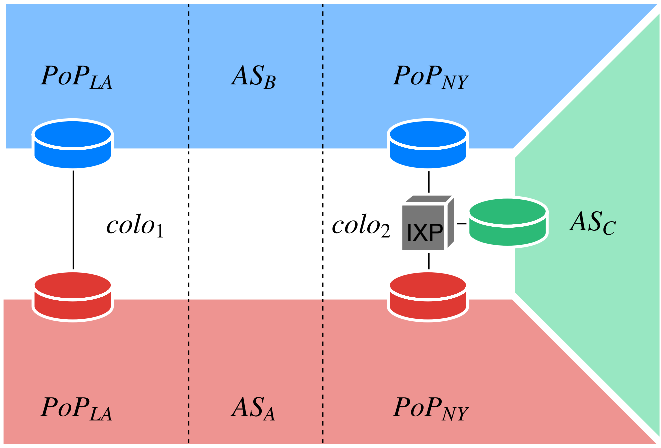

Fig. 1. Abstract representation for topology of AS_A, AS_B, and AS_C in red, blue, and green accordingly. AS_A and AS_B establish a private interconnection inside $colo_1$ at their LA PoP while peering with each other as well as AS_C inside $colo_2$ at their NY PoP facilitated by an IXP's switching fabric.

III. Tools & Datasets

This section provides an overview of various tools and datasets that have been commonly used by the measurement community for discovering Internet topology. We aim to familiarize the reader with these tools and datasets as they are continuously used within the literature by researchers. Researchers have utilized a wide range of tools for the discovery of topologies; they range from generic network troubleshooting tools such as traceroute or paris-traceroute to tools developed by the Internet measurement community such as Sibyl or MIDAR. Furthermore, researchers have benefited from many measurement platforms such as RIPE Atlas or PlanetLab which enable them to perform their measurements from a diverse set of ASes and geographic locations.

In addition to the aforementioned toolsets researchers have benefited from various datasets within their work. These datasets are collected by a few well-known projects in the Internet measurement community such as Routeviews [1], CAIDA's Ark [2], and CAIDA's AS relationships datasets or stem from other sources such as IP to geolocation datasets or information readily available on colocation facilities or IXP operators websites.

The remainder of this section is organized within two subsections. First, §III-A would provide an overview of the most commonly used tools and platforms for Internet topology discovery. Second, §III-B would give a brief overview of the datasets that appear in the literature presented within §IV and §V.

A. *Measurement Tools & Platforms*

Broadly speaking the tools used for Internet topology discovery can be categorized within three groups namely, (i) path discovery, (ii) alias resolution, and (iii) interface name decoding.

*1) **Path Discovery:*** Although originally developed for troubleshooting purposes, *traceroute* [3] has become one of the prominent tools used within the Internet measurement community. *traceroute* displays the set of intermediate router interfaces that are traversed towards a specific destination in the forward path. This is made possible by sending packets towards the destination with incremental TTL values, each router along the path would decrease the TTL value before forwarding the packet. If a router encounters a packet with a TTL value of 0 the packet would be dropped, and a notification message with its source address would be sent back to the originator of the packet. This, in turn, allows the originator of these packets to identify the source address of router interfaces along the forward path. Deployment of load-balancing mechanics by routers which rely on packet header fields can lead to inaccurate and incomplete paths to be reported by *traceroute*. Figure 2 illustrates an example of incorrect inferences by *traceroute* in the presence of load-balanced paths. Node a is a load-balancer and multiplexes packets between the top and bottom paths. In this example, the $TTL = 2$ probe originated from the source traverses the top path and expires at node b while the $TTL = 3$ probe goes through the bottom path and terminates at node e. These successive probes cause *traceroute* to incorrectly infer a non-existent link between nodes b and e. To address this problem, Augustin et al. [4] developed *paris-traceroute* which relies on packet header contents to enforce load-balancers to pick a single route for all probes of a single traceroute session. Furthermore, *paris-traceroute* uses a stochastic probing algorithm in order to enumerate all possible interfaces and links at each hop.

Given the scale of the Internet and its geographic span relying on a single vantage point (VP) to conduct topology discovery studies would likely lead to incomplete or inaccurate inferences. Researchers have relied on various active measurement platforms which either host a pre-defined set of tools, e.g. Dasu, Bismark, Dimes, Periscope, and RIPE Atlas [5], [6], [7], [8], [9] or provide full-access control, e.g. PlanetLab, CAIDA Archipelago, and GENI [10], [11], [12] to the user to conduct their measurements from a diverse set of networks and geographic locations. For example, RIPE Atlas [5] is composed of many small measurement devices (10k at the time of this survey) that are voluntarily hosted within many networks on a global scale. Hosting RIPE Atlas nodes would give credit to the hosting entity which later on could be used to conduct latency (*ping*) and reachability (*traceroute* and *paris-traceroute*) measurements. Periscope [7] is another platform that provides a unified interface for probing around 1.7k publicly available looking glasses (LGs) which provide a web interface to conduct basic network commands (*ping, traceroute,* and *bgp* on routers hosted in roughly 0.3k ASes. Periscope VPs are located at core ASes while RIPE Atlas probes are hosted in a mix of core and edge networks. Dasu [8] on the other hand mainly consists of VPs at edge networks and more specifically broadband users relying on ISPs to have Internet connectivity. Dasu consists of a plugin for the Vuze BitTorrent client that is able to conduct network measurement from the computers of users who have installed their plugin on their Vuze client. The authors of Dasu incentivize its adoption by reporting broadband network characteristics to its users. Cunha et al. [13] developed a route oracle platform named Sibyl which allowed users to define the path requirements for their measurement through an expressive input language based on symbolic regular-expressions after which Sibyl would select the source (LG) and destination pair that has the highest likelihood of satisfying the users path requirements based on its internal model.

Lastly, considering the large number of Internet hosts and networks, researchers have developed a series of tools that allow them to conduct large scale measurements in parallel. The methodology of *paris-traceroute* has been incorporated in *scamper* [14], an extensible packet prober that implements various common network measurement functionalities such as traceroute, ping, and alias resolution into a single tool. *scamper* is able to conduct measurements in parallel without exceeding a predefined probing rate. While *scamper* is able to run measurements in parallel, each measurement is conducted sequentially, this in turn could hinder its rate or induce overhead to the probing device in order to maintain the state of each measurement. *yarrp* [15], [16] is a high-rate IPv4 and IPv6 capable, Internet-scale probing tool inspired by the state-less design principles of *ZMap* [17] and *masscan* [18]. *yarrp* randomly permutates the IP and TTL space and encodes the state information of each probe within the IP and TCP header fields (which are included in the ICMP response) and is therefore able to conduct traceroute probes in parallel without incrementally increasing the TTL value.

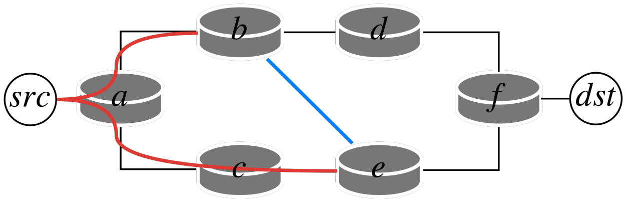

Fig. 2. Illustration of inferring and incorrect link ($b-e$) by *traceroute* due to load balanced paths. Physical links and traversed paths are shown with black and red lines accrodingly. The $TTL = 2$ probe traverses the top path and expires at node b while the $TTL = 3$ probe traverses the bottom path and expires at node e. This succession of probes causes *traceroute* to infer a non-existent link ($b-e$).

*2) **Alias Resolution**:* Paths which are obtained via the tools outlined in §III-A1 all specify the router interfaces that are encountered along the forward path. It is possible to observe multiple interfaces of a single router within different traceroute paths. The association of these interfaces to a single physical router is not clear from these outputs. Alias resolution tools have been developed to solve this issue. These tools would accept a set of interface addresses as an input and would provide a collection of interface sets, each of which corresponds to a single router. Alias resolution tools can broadly be categorized into two groups namely, (i) probing [19], [20], [21], [22], [23] and (ii) inference [24], [25], [26], [27] based techniques. The former would require a VP which would probe the interfaces in question to identify sets of interfaces which belong to the same router. Probe based techniques mostly rely on the IP ID field which is used for reassembling fragmented packets at the network layer. These techniques assume that routers rely on a single central incremental counter which assigns these ID values regardless of the interface. Given this assumption, Ally [19] probed IPs with UDP packets having high port numbers (most likely not in use) to induce an ICMP port unreachable response. Ally will infer IP addresses to be aliases if successive probes have incremental ID values within a short distance. Radargun [21] tries to address the probing complexity of Ally ($O(n^2)$) by iteratively probing IPs and inferring aliases based on the velocity of IP ID increments for each IP. MIDAR [23] presents a precise methodology for probing large scale pool of IP addresses by eliminating unlikely IP aliases using a velocity test. Furthermore, aliases are inferred by comparing the monotonicity of IP ID time series for multiple target IP addresses. MIDAR utilizes ICMP, TCP, and UDP probes to increase the likelihood of receiving responses from each router/interface. Palmtree [22] probes /30 or /31 mates of target IPs using a TTL value inferred to expire at the router in question to induce an ICMP_TTL_EXPIRED response from another interface of the router. Assuming no path changes have happened between measuring the routers hop distance and the time the ICMP_TTL_EXPIRED message has been generated, the source address of the ICMP_TTL_EXPIRED message should reside on the same router of the target IP and therefore are inferred to be aliases.

Inference based techniques accept a series of traceroute outputs and rely on a set of constraints and assumptions regarding the setting and environment which these routers are deployed to make inferences about interfaces that are most likely part of the same router. Spring et al. suggest a common successor heuristic to attribute IP addresses on the prior hop to the same router. This heuristic assumes that no layer-2 devices are present between the two routers in question. Gunes et al. Analytical Alias Resolution (AAR) [25] infers aliases using symmetric traceroute pairs by pairing interface addresses using the common address sharing convention of utilizing a /30 or /31 prefix for interfaces on both ends of a physical link. This method requires the routes between both end-pairs to be symmetrical. DisCarte [26] relies on the route record option to capture the forward and reverse interfaces for the first nine hops of a traceroute. Limited support and various route record implementations by routers in addition to the high complexity of the inference algorithm limits its applicability to wide/large scenarios.

*3) **Interface Name Decoding**:* Reverse DNS (RDNS) entries for observed interface addresses can be the source of information for Internet topology researchers. Port type, port speed, geolocation, interconnecting AS, and IXP name are examples of information which can be decoded from RDNS entries of router interfaces. These information sets are embedded by network operators within RDNS entries for ease of management in accordance to a (mostly) structured convention. For example, *ae-4.amazon.atlnga05.us.bb.gin.ntt.net* is an RNDS entry for a router interface residing on the border router of NTT (ntt.net) within Atlanta GA (atlnga) interconnecting with Amazon. Embedding this information is completely optional, and the structure of this information varies from one AS to another. Several tools have been developed to parse and extract the embedded information within RDNS entries [19], [28], [29], [30]. Spring et al. extracted DNS encoded information for the ISPs under study in their *Rocketfuel* project [19]. As part of this process, they relied on the city code names compiled in [31] to search for domain names which encode geoinformation in their name. *PathAudit* [28] is an extension to traceroute which report encoded information within observed router hops. In addition to geo information, *PathAudit* reports on interface type, port speed, and manufacturing vendor of the router. The authors of *PathAudit* extract common encodings (tags) from device configuration parameters, operator observations, and common naming conventions. Using this set of tags, RDNS entries from CAIDA's Ark project [2] are parsed to match against one or multiple of these tags. A clustering algorithm is employed to identify similar naming structures within domains of a common top level domain TLD. These common structures are translated into parsing rules which can match against other RDNS entries. *DDeC* [32] is a web service which decodes embedded information within RDNS entries by unifying the rulesets obtained by both *UNDNS* [19] and *DRoP* [29] projects.

B. Datasets

Internet topology studies have been made possible through various data sources regarding BGP routes, IXP information, colo facility listings, AS attributes, and IP to geolocation mapping. The following sub-section provides a short overview of data sources most commonly used by the Internet topology community.

*1) **BGP Feeds & Route Policies**:* University of Oregon's RouteViews and RIPE Routing Information Service (RIS) [1], [33] are projects originally conceived to provide real-time information about the global routing system from the standpoint of several route feed collectors. These route collectors periodically report the set of BGP feeds that they receive back to a server where the information is made publicly accessible. The data from these collectors have been utilized by researchers to map prefixes to their origin-AS or to infer AS relationships based on the set of observed AS-paths from all the route collectors. Routeviews and RIPE RIS provide a window into the global routing system from higher tier networks. Packet Clearing House (PCH) [34] maintains more than 100 route collectors which are placed within IXPs around the globe and provides a complementary view to the global routing system presented by Routeviews and RIPE RIS. Lastly, Regional Internet Registries (RIRs) maintain databases regarding route policies of ASes for each of the prefixes that are delegated to them using the Route Policy Specification Language (RPSL). Historically, RPSL entries are not well adopted and typically are not maintained/updated by ASes. The entries are heavily concentrated within RIPE and ARIN regions but nonetheless have been leveraged by researchers to infer or validate AS relationships [35], [36].

*2) **Colocation Facility Information**:* Colocation facilities (colo for short) are data-centers which provide space, power, cooling, security, and network equipment for other ASes to host their servers and also establish interconnections with other ASes that have a presence within the colo. PeeringDB and PCH [37], [38] maintain information regarding the list of colo facilities and their physical location as well as tenant ASes within each colo. Furthermore, some colo facility operators provide a list of tenant members as well as the list of transit networks that are available for peering within their facilities for marketing purposes on their website. This information has been mainly leveraged by researchers to define a set of constraints regarding the points of presence (PoP) for ASes.

*3) **IXP Information**:* IXPs are central hubs providing rich connectivity opportunities to the participating ASes. Their impact and importance regarding the topology of the Internet have been highlighted within many works [39], [40], [41], [42]. IXPs provide a switching fabric within one or many colo facilities where each participating AS connects their border router to this switch to establish bi-lateral peering with other member ASes or establishes a one to many (multi-lateral) peering with the route server that is maintained by the IXP operator. IXP members share a common subnet owned by the IXP operator. Information regarding the location, participating members, and prefixes of IXPs is readily available through PeeringDB, PCH, and the IXP operators website [37], [38].

*4) **IP Geolocation**:* The physical location of IP addresses isn't known. Additionally, IP addresses could correspond to mobile end-hosts or can be repurposed by the owner AS and therefore have a new geolocation. Several free and commercial databases have been made throughout the years that attempt to map IP addresses to physical locations. These datasets can vary in their coverage as well as the resolution of mapped addresses (country, state, city, and geo-coordinates). Maxmind's GeoIP2 [43], IP2Location databases [44], and NetAcuity [45] are among the most widely used IP geolocating datasets used by the Internet measurement community. Majority of these datasets have been designed to geolocate end-host IP addresses. Gharaibeh et al. [46] compare the accuracy of these datasets for geolocating router interfaces and while NetAcuity has relatively higher accuracy than Maxmind and IP2Location datasets, relying on RTT validated geocoding of RDNS entries is more reliable for geolocating router and core addresses.

IV. Capturing Network Topology

This section provides an overview of Internet measurement studies which attempt to capture the Internet's topology using various methodologies motivated by different end goals. Capturing Internet topology has been the focus of many pieces of research over the past decade, while each study has made strides of incremental improvements to present a more complete and accurate picture of Internet topology, the problem remains widely open and the subject of many recent studies.

Internet topology discovery has been motivated by a myriad of applications ranging from protocol design, performance measurement in terms of inter-AS congestion, estimating resiliency towards natural disasters and service or network interruptions, security implications of DDoS attacks and much more. A motivating example would be the Netflix Verizon dispute where the subpar performance of Netflix videos for Verizon customers lead to lengthy accusations from both parties [47]. The lack of proper methodologies to capture inter-AS congestion by independent entities at the time further elongated the dispute. Within Section V we provide a complete overview of works which rely on some aspect of Internet topology to drive their research and provide insight regarding the performance or resiliency of the Internet.

Capturing Internet topology is hard due to many contributing factors, the following is a summary of them:

- The Internet is by nature a decentralized entity composed of a network of networks, each of the constituent networks lacks any incentive to share their topology publicly and often can have financial gains by obscuring this information.
- Topology discovery studies are often based on "hackish" techniques that rely on toolsets which were designed for completely different purposes. The designers of the TCP/IP protocol stack did not envision the problem of topology discovery within their design most likely due to the centralized nature of the Internet in its inception. The *de facto* tool for topology discovery has been *traceroute* which is designed for troubleshooting and displaying paths between a host and a specific target address.
- Capturing inter-AS links within Internet topology becomes even more challenging due to lack of standardization for proper ways to establish these links. More specifically, the shared address between two border routers could originate from either of the participating networks. Although networks typically rely on common good practices such as *using addresses from the upstream provider*, the lack of any oversight or requirement within RFC standards does not guarantee its proper execution within the Internet.
- A certain set of RFCs regarding how routers should handle TTL expired messages has resulted in incorrect inferences of the networks which are establishing inter-AS interconnections. For example, responses generated by third-party interfaces on border routers could lead to the inference of an inter-AS link between networks which necessarily are not interconnected with each other.

Topology discovery studies can be organized according to many of their features; in particular, the granularity of the obtained topology seems to be the most natural fit. Each of the studies in this section based on the utilized dataset, or devised methodology results in topologies which capture the state of the Internet at different granularities, namely physical-level, router-level, PoP-level, and AS-level. The aforementioned resolutions of topology have a direct mapping to the abstract layers of the TCP/IP stack, e.g. physical-level corresponds to the first layer (physical), router-level can be mapped to the transport layer, and PoP-level as well as AS-level topologies are related to application layer at the top of the TCP/IP stack. These abstractions allow one to capture different features of interest without the need for dealing with the complexities of lower layers. For instance, the interplay of routing and the business relationships between different ASes can be captured through an AS-level topology without the need to understand how and where these inter-AS relationships are being established.

In the following subsections, we will provide an overview of the most recent as well as prominent works that have captured Internet topology at various granularities. We present all studies in accordance to their chronological order starting with works related to AS-level topologies as the most abstract representation of Internet topology within Section §IV-A, AS-level topologies are the oldest form of Internet topology but have retained their applicability for various forms of analyses throughout the years. Later we'll present router-level and physical-level topologies within Section §IV-B and §IV-D accordingly.

A. *AS-Level Topology*

The Internet is composed of various networks or ASes operating autonomously within their domain that interconnect with each other at various locations. This high-level abstraction of the Internet's structure is captured by graphs representing AS-level topologies where each node is an AS and edges present an interconnection between two ASes. These graphs lay-out virtual entities (ASes) that are interconnecting with each other and abstract out details such as the number and location where these inter-AS links are established. For example, two large Tier-1 networks such as Level3 and AT&T can establish many inter-AS links through their border routers at various metro areas. These details are abstracted out, and all of these inter-AS links are represented by a single edge within the AS-level topology. The majority of studies rely on control plane data that is obtained by active measurements of retrieving router dumps through available looking glasses or passive measurements that capture BGP feeds, RPSL entries and BGP community attributes. Path measurements captured through active or passive *traceroute* probes have been an additional source of information for obtaining AS-level topologies. The obtained *traceroute* paths have been mapped to their corresponding AS path by translating each hop's address to its corresponding AS. Capturing AS-level topology has been challenging mainly due to limited visibility into the global routing system, more specifically the limited set of BGP feeds that each route collector is able to observe. This limited visibility is known as the topology incompleteness problem within the community. Researchers have attempted to address this issue by either modeling Internet topology by combing the limited ground truth information with a set of constraints or by presenting novel methodologies that merge various data sources in order to obtain a comprehensive view of Internet topology. The later efforts lead to research's that highlighted the importance of IXPs as central hubs of rich connectivity. Within the remainder of this Section we organize works into the following three groups: (i) graph generative and modeling, (ii) topology incompleteness, and (iii) IXP's internal operation and peerings.

*1) **Graph Generation & Modeling**:* Graph generation techniques attempt to simulate network topologies by relying on a set of constraints such as the maximum number of physical ports on a router. These constraints coupled with the limited ground truth information regarding the structure of networks are used to model and generate topologies. The output of these models can be used in other studies which investigate the effects of topology on network performance and resiliency of networks towards attacks or failures caused by natural disasters.

Li et al. [48] argue that graph generating models rely on replicating too abstract measures such as degree distribution which are not able to express the complexities/realities of Internet topology. Authors aim to model ASes/ISPs as the building blocks of the Internet at the granularity of routers, where nodes represent routers and links are Layer2 physical links which connect them together. Furthermore, the authors argue that technological constraints on routers switching fabric dictate the amount of bandwidth-links we can have within this topology. Furthermore, due to economical reasons access providers aggregate their traffic over a few links as possible since the cost of laying physical links could surpass that of the switching/routing infrastructure. This, in turn, leads to lower degree core and high degree edge elements. The authors create five graphs with the same degree distribution but based on different heuristics/models and compare the performance of these models using a single router model. Interestingly graphs that are less likely to be produced using statistical measures have the highest performance.

Gregori et al. [49] conduct a structural interpretation of the Internet connectivity graph with an AS granularity. They report on the structural properties of this graph using k-core decomposition techniques. Furthermore, they report what effects IXPs have on the AS-level topology.

The data for this study is compiled from various datasets, namely CAIDA's Ark, DIMES, and Internet Topology Collection from IRL which is a combination of BGP updates from Routeviews, RIPE RIS, and Abilene. The first two datasets consist of traceroute data and are converted to AS-level topologies by mapping each hop to its corresponding ASN. A list of IXPs was obtained using from PCH, PeeringDB, Euro-IX, and bgp4.as. The list of IXP members was compiled either from the IXP websites or by utilizing the **show ip bgp summary command** from IXPs which host an LG.

Using the obtained AS-level graph resulted from combing various data sources the authors report on various characteristics of the graph namely: degree, average neighbor degree, clustering coefficient, betweenness centrality, and k-core decomposition. A k-core subgraph has a minimum degree of k for every node and is the largest subgraph which has this property. The authors present stats regarding the penetration of IXPs in different continents with Europe having the largest share (47%) and North America (19%) at second position. Furthermore using k-core decomposition, the authors identify a densely connected core and a loosely connected periphery which consists of the majority of nodes. The authors also look at the fraction of nodes in the core which are IXP participants and find that IXPs play a fundamental role in the formation of these cores.

*2) **Topology Incompleteness**:* Given the limited visibility of each of the prior works, researchers have relied on a diverse set of data sources and devised new methodologies for inferring additional peerings to address the incompleteness of Internet topology. These works have lead to highlighting the importance of IXPs as a means of providing the opportunity for establishing many interconnections with IXP members and a major source for identifying missing peering links. Peerings within IXPs and their rich connectivity fabric between many edge networks caused topological changes to the structure of the Internet deviating from the historical hierarchical structure and as a consequence creating a more flat Internet structure referred to as Internet flattening within the literature.

He et al. [50] address AS-level topology incompleteness by presenting tools and methodologies which identify and validate missing links. BGP snapshots from various (34 in total) Routeviews, RIPE RIS, and public route servers are collected to create a baseline AS-level topology graph. The business relationship of each AS edge is identified by using the PTE algorithm [51]. The authors find that the majority of AS links are of a c2p type, while most of the additional links which are found by additional collectors are p2p links. Furthermore, by parsing IRR datasets using Nemecis [52] to infer additional AS links. A list of IXP participants is compiled by gathering IXP prefixes from PCH and performing DNS lookups and parsing the resulting domain name to infer the participating ASN. Furthermore, the authors infer inter-AS links within IXPs by relying on traceroute measurements which cross IXP addresses and utilize a majority voting scheme to infer the participants ASN reliably. By Combing all these datasets and proposed methodologies, the authors find about 300% additional links compared to prior studies, most of which is found to be established through IXPs.

Augustin et al. [39] attempt to expand on prior works for discovering IXP peering relationships by providing a more comprehensive view of this ecosystem. They rely on various data sources to gather information on IXPs as much as possible, their data-sources are: (i) IXP databases such as PCH and PeeringDB, (ii) IXP websites which typically list their tenants as well as the prefixes which are employed by them, (iii) RIRs may include BGP policy entries specifically the *import* and *export* entries that expose peering relationships, (iv) DNS names of IXP addresses which include information about the peer, (v) BGP dumps from LGs, Routeviews, and RIPE's RIS can include next hop neighbors which are part of an IXP prefix. The authors conduct targeted traceroute measurements with the intention of revealing peering relationships between members of each IXP. To limit the number of conducted probes, the authors either select a vantage point within one of the member ASes or if not available they rely on the AS relationship datasets to discover a - at most 2 hops away - neighbor for each member which has a VP. Using the selected VPs, they conduct traceroutes towards alive addresses (or random address if such an address was not discovered) in the target network. Inference of peerings based on traceroutes is done using a majority voting scheme similar to [50]. The authors augment their collected dataset with the data plane measurements of CAIDA's Skitter, DIMES, and traceroutes measured from about 250 PlanetLab nodes. The resultant dataset is able to identify peerings within 223 (out of 278) IXPs which consisted of about 100% (40%) more IXPs (peerings) compared to the work of He et al. [50].

Ager et al. [53] rely on sFlow records from one of largest European/global IXPs as another source of information for inferring peering relationships between IXP tenants and provide insight on three fronts: (i) they outline the rich connectivity which is happening over the IXP fabric and contrast that with known private peerings which are exposed through general topology measurement studies, (ii) present the business dynamics between participants of the IXP and providing explanation for their incentives to establish peering relationships with others, and (iii) provide the traffic matrix between peers of the IXP as a microcosm of Internet traffic. Among the set of analyses that have been conducted within the paper one could point to: (i) comparison of peering visibility from Routeviews, RIPE, LGs, and the IXPs perspective, (ii) manual label for AS types as well as the number of established peerings per member, (iii) breakdown of traffic into various protocols based on port numbers as well as the share of each traffic type among various AS types, and (iv) traffic asymmetry, ratio of used/served prefixes and geo-distance between end-points.

Khan et al. [54] utilize LG servers to provide a complementary view to Routeviews and RIPE RIR of the AS-level Internet topology. A list of 1.2k LGs (420 were operational at the time of the study) has been built by considering various sources including PeeringDB, traceroute.org, traceroute.net.ru, bgp4.as, bgp4.net, and virusnet. AS-level topologies from IRL, CAIDA's Ark, iPlane, and IRR's are used to compare the completeness of the identified AS-links. For the duration of a month **show ip bgp summary** is issued twice a week and **BGP neighbor ip advertised** is issued once a week towards all LGs which support the command. The first command outputs each neighbor's address and its associated ASN while the second command outputs the routing table of the router, consisting of reachable prefixes, next hop IP as well as the AS path towards the given prefix. AS-level connectivity graph is constructed by parsing the output of the prior commands. Using this new data source enables the authors to identify an additional 11k AS-links and about 700 new ASes.

Kloti et al. [55] perform a cross-comparison of three public IXP datasets, namely PeeringDB [37], Euro-IX [56], and PCH [38] to study several attributes of IXPs such as location, facilities, and participants. Aside from the three aforementioned public IXP datasets, for validation purposes BGP feeds collected by PCH route collectors as well as data gathered from 40 IXP websites was used through the study. The three datasets lack common identifiers for IXPs across datasets, for this reason in a first pass IXPs are linked together through an automated process by relying on names and geo information, in the second pass linked IXPs are manually checked for correctness. The authors present one of the largest IXP information datasets at the time as a side effect of their study.

Geo coverage of each dataset is examined where the authors find relatively close coverage by each dataset except for North America region where PCH has the highest coverage. Facility location for IXPs is compared across datasets and is found that PCH lacks this information and in general facility information for IXPs is limited for other datasets. Complementarity of datasets is presented using both Jaccard and overlap index. It is found that PeeringDB and Euro-IX have the largest overlap within Europe and larger IXPs tend to have the greatest similarity across all pairs of datasets.

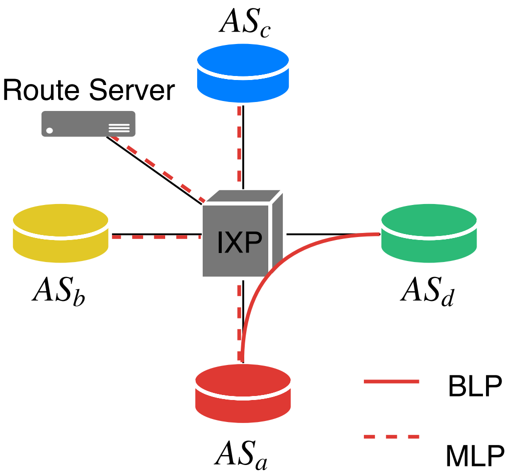

Fig. 3. Illustration of an IXP switch and route server along with 4 tenant networks AS_a, AS_b, AS_c, and AS_d. AS_a establishes a bi-lateral peering with AS_d (solid red line) as well as multi-lateral peerings with AS_b and AS_c (dashed red lines) facilitated by the route server within the IXP.

*3) **IXP Peerings**:* The studies within this section provide insight into the inner operation of IXPs and how tenants establish peerings with other ASes. Each tenant of an IXP can establish a one-to-one (bilateral) peering with other ASes of the IXP similar to how regular peerings are established. Given the large number of IXP members, a great number of peering sessions should be maintained over the IXP fabric. Route servers have been created to alleviate this issue where each member would establish a peering session with the route server and describe its peering preferences. This, in turn, has enabled one-to-many (multilateral) peering relationships between IXP tenants. Figure 3 illustrates an IXP with 4 tenant networks AS_a, AS_b, AS_c, and AS_d. AS_a established a bi-lateral peering with AS_d (solid red line) as well as multi-lateral peerings with AS_b and AS_c (dashed red lines) that are facilitated by the route server within the IXP. Studies within this section propose methodologies for differentiating these forms of peering relationships from each other and emphasize the importance of route servers in the operation of IXPs.

Giotsas et al. [57] present a methodology to discover multi-lateral peerings within IXPs using the BGP communities attributes and route server data. The BGP communities attribute which is 32bits follows specific encoding to indicate either of the following policies by each member of an IXP: (i) *ALL* routes are announced to all IXP members. (ii) *EXCLUDE* block an announcement towards a specific member, this policy is usually used in conjunction with the ALL policy. (iii) *NONE* block an announcement towards all members, and (iv) *INCLUDE* allow an announcement towards a specific member, this policy is used with the NONE policy. Using a combination of prior policies a member AS can control which IXP members receive its BGP announcements. By leveraging available LGs at IXPs and issuing router dump commands, the authors obtain the set of participating ASes and the BGP communities values for their advertised prefixes which in turn allows them to infer the connectivity among IXP participants. Furthermore, additional BGP communities values are obtained by parsing BGP feeds from Routeviews and RIPE RIS archives. Giotsas et al. infer the IXP by either parsing the first 16bits of the BGP communities attribute or by cross-checking the list of excluded ASes against IXP participants.

By combining the passive and active measurements, the authors identify 207k multilateral peering (MLP) links between 1.3k ASes. They validate their findings by finding LGs which are relevant to the identified links from PeeringDB, by testing 26k different peerings they are able to confirm 98.4% of them. Furthermore Giotsas et al. parse the peering policies of IXP members either from PeeringDB or from IXP websites which provide this information and find that 72%, 24%, and 4% of members have an open, selective, and restrictive peering policy accordingly. Participation in a route server seems to be positively correlated to a networks openness in peering. The authors present the existence of a binary pattern in terms of the number of allowed/blocked ASes where ASes either allow or block the majority of ASes from receiving their announcements. Peering density as a representation of the percentage of established links against the number of possible links is found to be between 80%-95%.

Giotsas et al. [58] expand their prior work [57] by inferring multi-lateral peering (MLP) links between IXP tenants by merely relying on passive BGP measurements. BGP feeds are collected from both Routeviews and RIPE RIS collectors. Additionally, the list of IXP looking glasses, as well as their tenants, are gathered from PeeringDB and PCH. The authors compile a list of IXP tenants, using which the setter of each BGP announcement containing the communities attribute is determined by matching the AS path against the list of IXP tenants. If less than two ASes match against the path, no MLP link can be identified. From the two matching ASes, the AS which is closest to the prefix would be the setter, if more than two ASes match, only two ASes which have a p2p relationship according to CAIDA's AS relationship dataset are selected and the one closer to the prefix is identified as the setter. Depending on a blacklist or whitelist policy that the setter AS has chosen a list of multi-lateral peers for each setter AS is compiled.

The methodology is applied to 11 large IXP route servers; the authors find about 73% additional peering links out of which only 3% of the links are identified within CAIDA's Ark and DIMES datasets. For validation, the authors rely on IXP LGs and issue a *show ip bgp* command for each prefix. About 3k links where tested for validation and 94% of them were found to be correct.

Richter et al. [59] outline the role and importance of route servers within IXPs. For their data, weekly snapshots of peer and master RIBs from two IXPs which exposes the multi-lateral peerings that have been happening at the IXP are used. Furthermore, the authors have access to sFlow records which are sampled from the IXP's switching infrastructure. This dataset allows the authors to identify peerings between IXP members which have been established without the help of route servers. Using peer RIB snapshots peering relationships between IXP members as well as the symmetrical nature of it is identified. For the master RIB, Richter et al. assume peering with all members unless they find members using BGP community values to control their peering. The data plane sFlow measurements would correspond to a peering relationship if BGP traffic is exchanged between two members of the IXP. The proclivity of multi-lateral peering over bi-lateral peering is measured and found that ASes favor multi-lateral peerings with a ratio of 4:1 and 8:1 in the large and medium IXPs accordingly. Furthermore, traffic volumes transmitted over multi-lateral and bi-lateral peerings are measured and found that ASes tend to send more traffic over bi-lateral links with a ratio of 2:1 and 1:1 for the large and medium IXPs accordingly. It is found that ASes have binary behavior of either advertising all or none of their prefixes through the route server. Additionally, when ASes establish hybrid (multi and bi-lateral) peerings, they do not advertise further prefixes over their bi-lateral links. Majority of additional peerings happen over multi-lateral fabric while traffic ratios between multi(bi)-lateral peerings remain fairly consistent over the period of study.

Summary: This subsection provided an overview of researches concerned with AS-level topology. The majority of studies were concerned with the incompleteness of Internet topology graphs. These efforts lead to highlighting the importance of IXPs as central hubs of connectivity. Furthermore, various sources of information such as looking glasses, router collectors within IXPs, targeted traceroutes, RPSL entries, and traffic traces of IXPs were gleaned together to provide a more comprehensive view of inter-AS relationships within the Internet. Lastly the importance of route servers to the inner operation of IXPs and how they enable multi-lateral peering relationships was brought into attention.

B. *Router-Level Topology*

Although AS-level topologies provide a preliminary view into the structure and peering relations of ASes, they merely represent virtual relationships and do not reflect details such as the number and location where these peerings are established. ASes establish interconnections with each other by placing their border routers within colos where other ASes are also present. Within these colos ASes can establish one to one peerings through private interconnections or rely on an IXPs switching fabric to establish public peerings with the IXP participants. Furthermore, some ASes extend their presence into remote colos to establish additional peerings with other ASes by relying on layer2 connectivity providers. Capturing these details can become important for accurately attributing inter-AS congestion to specific links/routers or for pin-pointing links/routers that are responsible for causing outages or disruptions within the connectivity of a physical region or network. Studies within this section aim to present methodologies to infer router-level topologies using data plane measurements in the form of traceroute. These methods would address the aforementioned shortcomings of AS-level topologies by mapping the physical entities (border routers) which are used to establish peering relations and therefore can account for multiple peering links between each AS. Furthermore, given that routers are physical entities, researchers are able to pinpoint these border routers to geo locations using various data sources and newly devised methodologies. Creating router-level topologies of the Internet can be challenging due to many reasons. First, given the span of the Internet as well as the interplay of business relationships and routing dynamics, *traceroute* as the de-facto tool for capturing router-level topologies is only capable of recording a minute fraction of all possible paths. Routing dynamics caused by changes in each ASes route preference as well as the existence of load-balancers further complicate this task. Second, correctly inferring which set of ASes have established an inter-AS link through traceroute is not trivial due to non-standardized practices for establishing interconnections between border routers as well as several RFCs regarding the operation of routers that cause *traceroute* to depict paths that do not correspond to the forward path. Lastly, given the disassociation of the physical layer from the transport layer establishing the geolocation for the set of identified routers is not trivial. Within Section III we presented a series of platforms which try to address the first problem. The following studies summarize recent works which try to address the latter two problems.

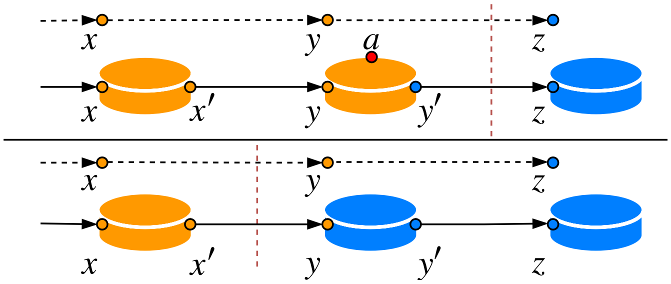

Fig. 4. Illustration of address sharing for establishing an inter-AS link between border routers. Although the traceroute paths (dashed lines) are identical the inferred ownership of router interfaces and the placement of the inter-AS link differs for these two possibilities.

*1) **Peering Inference**:* As briefly mentioned earlier, inferring inter-AS peering relationships using traceroute paths is not trivial. To highlight this issue, consider the sample topology within Figure 4 presenting the border routers of AS_1 and AS_2 color coded as orange and blue accordingly. This figure shows the two possibilities for address sharing on the inter-AS link. The observed traceroute path traversing these border routers is also presented at the top of each figure with dashed lines. Within the top figure AS_2 is providing the address space for the inter-AS link $(y' - z)$ while AS_1 provides the address space for the inter-AS link $(x' - y)$ for the bottom figure. As we can see both of the traceroute paths are identical to each other while the ownership of router interfaces and the placement of the inter-AS link differs for these two possibilities. To further complicate the matter, a border router can respond with an interface (a in the top figure using address space owned by AS_3 color coded with red), not on the forward path of the traceroute leading to incorrect inference of an inter-AS link between AS_1 and AS_3. Lastly, the border routers of some ASes are configured to not respond to traceroute probes which restrict the chances of inferring inter-AS peerings with those ASes. The studies within this section try to address these difficulties by using a set of heuristics which are applied to a set of traceroutes that allow them to account for these difficulties.

Spring et al. [19] done the seminal work of mapping networks of large ISPs and inferring their interconnections through traceroute probes. They make three contributions namely, (i) conducting selective traceroute probes to reduce the overall overhead of running measurements, (ii) provide an alias resolution technique to group IP address into their corresponding router, and (iii) parse DNS information to extract PoP/GEO information. Their selective probing method is composed of two main heuristics: (i) directed probing, which utilizes Routeviews data and the advertised paths to probe prefixes which are likely to cross the target network, (ii) path reduction, that avoids conducting traceroutes which would lead to redundant paths, i.e., similar ingress or egress points. Additionally, an alias resolution technique named *Ally* is devised to group interfaces from a single network into routers. Lastly, a series of DNS parsing rules are crafted to extract geoinformation from router interface RDNS entries. The extracted geo information allows the authors to identify the PoPs of each AS. Looking glasses listed on *traceroute.org* are used to run *Rocketfuel*'s methodology to map the network of 10 ISPs including AT&T, Sprint, and Verio. The obtained maps were validated through private correspondence with network operators and by comparing the set of identified BGP neighbors with those obtainable through BGP feeds.

Nomikos et al. [60] develop an augmented version of traceroute (*traIXroute*) which annotates the output path and reports whether (and at which exact hop) an IXP has been crossed along the path. The tool can operate with either traceroute or scamper as a backend. As input, *traIXroute* requires IXP membership and a list of their corresponding prefixes from PeeringDB and PCH as well as Routeviews' prefix to origin-AS mapping datasets. *traIXroute* annotates the hops of the observed path with the origin AS and tags hops which are part of an IXP prefix and also provides the mapping between an IXP address and the members ASN if such a mapping exists. Using a sliding window of size three the hops of the path are examined to find (i) hops which are part of an IXP prefix, (ii) hops which have an IXP to ASN mapping, and (iii) whether the adjacent ASes are IXP members or not. The authors account for a total of 16 possible combinations and present their assessment regarding the location of the IXP link for 8 cases that were most frequent. About 75% of observed paths matched rules which rely on IXP to ASN mapping data. The validity of this data source is looked into by using BGP dumps from routers that PCH operates within multiple IXPs. A list of IXP address to ASN mappings was compiled by using the next hop address and first AS within the AS path from these router dumps. The authors find that 92% (93%) of the IXP to ASN mappings reported by PeeringDB (PCH) are accurate according to the BGP dumps. Finally, the prevalence of IXPs along Internet paths are measured by parsing a CAIDA Ark snapshot. About 20% of paths are reported to cross IXPs, the IXP hop on average is located on the 6th hop at the middle of the path, and only a single IXP is observed along each route which is in accordance with valley-free routing.

Luckie et al. [61] develop *bdrmap*, a method to identify inter-domain links of a target network at the granularity of individual routers by conducting targeted traceroutes. As an input to their method, they utilize originated prefixes from Routeviews and RIPE RIS, RIR delegation files, list of IXP prefixes from PeeringDB and PCH, and CAIDA's AS-to-ORG mapping dataset. Target prefixes are constructed from the BGP datasets by splitting overlapping prefixes into disjoint subnets, the first address within each prefix is targeted using *paris-traceroute*, neighbors border addresses are added to a stop list to avoid further probing within the customer's network. IP addresses are grouped together to form a router topology by performing alias resolution using Ally and Mercator. By utilizing the prefixscan tool, they try to eliminate third-party responses for cases where interfaces are responsive to alias resolution. Inferences to identify inter-AS links are done by iteratively going through a set of 8 heuristics which are designed to minimize inference errors caused by address sharing, third-party response, and networks blocking traceroute probes. Luckie et al. deploy their tool within 10 networks and receive ground truth results from 4 network operators; their method is able to identify 96-99% of inter-AS links for these networks correctly. Furthermore, the authors compare their findings against BGP inferred relationships and find that they are able to observe between 92% - 97% of BGP links. Using a large US access network as an example, the authors study the resiliency of prefix reachability in terms of the number of exit routers and find that only 2% of prefixes exit through the same router while a great majority of prefixes had about 5-15 exit routers. Finally, the authors look at the marginal utility of using additional VPs for identifying all inter-AS links and find that results could vary depending on the target network and the geographic distribution of the VPs.

Marder et al. [62] devise a tool named *MAP-IT* for identifying inter-AS links by utilizing data-plane measurements in the form of traceroutes. The algorithm developed in this method requires as input the set of traceroute measurements which were conducted in addition to prefix origin-AS from BGP data as well as a list of IXP prefixes and CAIDA's AS to ORG mapping dataset. For each interface a neighbor set (N_s) composed of addresses appearing on prior (N_b) and next (N_f) hops of traceroute is created. Each interface is split into two halves, the forward and backward halves. Direct inferences are made regarding the ownership of each interface half by counting the majority ASN based on the current IP-to-AS mapping dataset. At the end of each round, if a direct inference has been made for an interface half, the other side will be updated with an indirect inference. Furthermore, within each iteration of the algorithm using the current IP-to-AS mapping, *MAP-IT* visits interface halves with direct inferences to check whether the connected AS still holds the majority, if not the inference is reduced to indirect, after visiting all interface halves any indirect inference without an associated direct inference is removed. *MAP-IT* would update the IP to AS mapping dataset based on the current inferences and would continue this process until no further inferences are made. For verification Marder et al. use Internet2's network topology as well as a manually compiled dataset composed of DNS names for Level3 and TeleSonera interfaces. The authors investigate the effect of the hyper parameter f which controls the majority voting outcome for direct inferences and empirically find that a value of 0.5 yields the best result. Using *f=0.5 MAP-IT* has a recall of 82% - 100% and a precision of 85% - 100% for each network. The authors also look into the incremental utility of each iteration of *MAP-IT*, interestingly the majority (80%) of inferences can be made in the first round which is equivalent to making inferences based on a simple IP2AS mapping. The algorithm converges quickly after its 2nd and 3rd iterations.

Marder et al. [63] combine the best practices of *bdrmap* [61] and *MAP-IT* [62] into *bdrmapIT*, a tool for identifying the border routers that improves *MAP-IT*'s coverage without loosing *bdrmap*'s accuracy at identifying border routers of a single ASN. The two techniques are mainly made compatible with the introduction of "Origin AS Sets" which annotates each link between routers with the set of origin ASes from the prior hop. *bdrmapIT* relies on a two-step iterative process. During the first step, the owner of routers are inferred by counting the routers majority subsequent interfaces votes. Exceptions in terms of the casted vote for IXP interfaces, reallocated prefixes, and multi-homed routers are made to account for these cases correctly. During the second step, interfaces are annotated with an ASN using either the origin AS (if router annotation matches that of the interface) or the majority vote of prior connected routers (if router annotation differs from the interface). The iterative process is repeated until no further changes are made to the connectivity graph. The methodology is evaluated using *bdrmap*'s ground truth dataset, as well as the ITDK dataset by removing the probes from a ground truth VP. The authors find that *bdrmapIT* improves the coverage of *MAP-IT* by up to 30% while maintaining the accuracy of *bdrmap*.

2) ***Geo Locating Routers & Remote Peering***: Historically ASes would have established their peering relations with other ASes local to their PoPs and would have relied on their upstream providers for connectivity to the remainder of the Internet. IXPs enabled ASes to establish peerings that both improved their performance due to shorter paths and reduced their overall transit costs by offload upstream traffic on p2p links instead of c2p links. With the proliferation of IXPs and their aforementioned benefits, ASes began to expand their presence not only within local IXPs but also remote ones as well. ASes would rely on layer2 connectivity providers to expand their virtual PoPs within remote physical areas. Layer3 measurements are agnostic to these dynamics and are not able to distinguish local vs. remote peering relations from each other. Researchers have tried to solve this issue by pinpointing border routers of ASes to physical locations. The association of routers to geolocations is not trivial, researchers have relied on a collection of complementary information such as geocoded embeddings within reverse DNS names or by constraining the set of possible locations through colo listings offered by PeeringDB and similar datasets. In the following, we present a series of recent studies which tackle this unique issue.

Castro et al. [41] present a methodology for identifying remote peerings, where two networks interconnect with each other via a layer-2 connectivity provider. Furthermore, they derive analytical conditions for the economic viability of remote peering versus relying on transit providers. Levering PeeringDB, PCH, and information available on IXP websites a list of IXP's as well as their tenants, prefixes and interface to member mapping is obtained. For this study, IXPs which have at least one LG or RIPE NCC probe (amounting to a total of 22) are selected. By issuing temporally spaced probes towards all of the identified interfaces within IXP prefixes and filtering interfaces which either do not respond frequently or do not match an expected maximum TTL value of 255 or 64 a minimum RTT value for each interface is obtained. By examining the distribution of minimum RTT for each interface, a conservative threshold of 10ms is selected to consider an interface as remote. A total of 4.5k interfaces corresponding to 1.9k ASes in 22 IXPs are probed in the study. The authors find that 91% of IXPs have remote peering while 285 ASes have a remote interface. Findings including RTT measures as well as remote labels for IXP members were confirmed for TorIX by the staff. One month of Netflow data captured at the border routers of RedIRIS (Spain's research and education network) is used to examine the amount of inbound and outbound traffic between RedIRIS and its transit providers, using which an upper bound for traffic which can be offloaded is estimated. Furthermore, the authors create a list of potential peers (2.2k) which are reachable through Euro-IX, these potential peers are also categorized into different groups based on their peering policy which is listed on PeeringDB. Considering all of the 2.2k networks RedIRIS can offload 27% (33%) of its inbound (outbound) traffic by remotely peering with these ASes. Through their analytical modeling, the authors find that remote peering is viable for networks with global traffic as well as networks which have higher ratios of traffic-independent cost for direct peering compared to remote peering such as networks within Africa.

Giotsas et al. [64] attempt to obtain a peering interconnection map at the granularity of colo facilities. Authors gather AS to facility mapping information from PeeringDB as well as manually parsing this information for a subset of networks from their websites. IXP lists and members were compiled by combining data from PeeringDB, PCH, and IXP websites. For data-plane measurements, the authors utilize traceroute data from RIPE Atlas, iPlane, CAIDA's Ark, and a series of targeted traceroutes conducted from looking glasses. The authors annotate traceroute hops with their corresponding ASN and consider the segment which has a change in ASN as the inter-AS link. Using the colo-facility listing obtained in the prior step the authors produce a list of candid facilities for each inter-AS link which can result in three cases: (i) a single facility is found, (ii) multiple facilities match the criteria, or (iii) no candid facility is found. For the latter two cases, the author's further constraint the search space by either benefiting from alias resolution results (two alias interfaces should reside in the same facility) or by conducting further targeted probes which are aimed at ASNs that have a common facility with the owner AS of the interface in question. The methodology is applied to five content providers (Google, Yahoo, Akamai, Limelight, and Cloudflare) and five transit networks (NTT, Cogent, DT, Level3, and Telia). The authors present the effect of each round of their constrained facility search (CFS) algorithm's iteration (max iteration count of 100), the majority of pinned interfaces are identified up to the 40th iteration with RIPE probes providing a better opportunity for resolving new interfaces. The authors find that DNS-based pinning methods are able to identify only 32% of their findings. The authors also cross-validate their findings using direct feedback from network admins, BGP communities attribute, DNS records, and IXP websites with 90% of the interfaces being pinned correctly and for the remainder, the pinning accuracy was correct at a metro granularity.

Nomikos et al. [42] present a methodology for identifying remote peers within IXPs, furthermore they apply their methodology to 30 large IXPs and characterize different aspects of the remote peering ecosystem. They define an IXP member as a remote peer if it is not physically connected to the IXPs fabric or reaches the IXP through a reseller. The development of the methodology and the heuristics used by the authors are motivated by a validation dataset which they obtain through directly contacting several IXP operators. A collection of 5 heuristics are used in order to infer whether an IXP member is peering locally or remotely these heuristics in order of importance are: (i) the port capacity of a customer, (ii) latency measurements from VPs within IXPs towards customer interfaces, (iii) colocation locations within an RTT radius, (iv) multi-IXP router inferences by parsing traceroutes from publicly available datasets and corroborating the location of these IXPs and whether the AS in question is local to any of them, and (v) identifying private peerings (by parsing public traceroute measurements) between the target AS and one or more local IXP members is used as a last resort to infer whether a network is local or remote to a given IXP. The methodology is applied to 30 large IXPs, and the authors find that a combination of RTT and colo listings to be the most effective heuristics in inferring remote peers. Overall 28% of interfaces are inferred to be peering remotely and for 90% of IXPs. The size of local and remote ASes in terms of customer cone is observed to be similar while hybrid ASes tend to have larger network sizes. The growth of remote peering is investigated over a 14 month period, and the authors find that the number of remote peers grew twice as fast as the number of local peers.

Motamedi et al. [65] propose a methodology for inferring and geolocating interconnections at a colo level. The authors obtain a list of colo facility members from PeeringDB and colo provider webpages. A series of traceroutes towards the address space of prior steps ASes are conducted using available measurement platforms such as looking glasses and RIPE Atlas nodes in the geo proximity of the targeted colo. *tracerotue* paths are translated to a router-level connectivity graph using alias resolution and a set of heuristics based on topology constraints. The authors argue that a router-level topology coupled with the prevalence of observations allows them to account for *traceroute* anomalies and they are able to infer the correct ASes involved in each peering. To geolocate routers, an initial set of *anchor* interfaces with a known location is created by parsing reverse DNS entries for the observed router interfaces. This information is propagated/expanded through the router-level graph by a Belief Propagation algorithm that uses a set of co-presence rules based-on membership in the same alias set and latency difference between neighboring interfaces.

Summary: while traceroutes have been historically utilized as a source of information to infer inter-AS links, the methodologies did not correctly account for the complexities of inferring BGP peerings from layer-3 probes. The common practice of simply mapping interface addresses along the path to their origin-AS based on BGP data does not account for the visibility of BGP collectors, address sharing for establishing inter-AS links, third-party responses of TTL expired messages by routers, and unresponsive routers or firewalled networks along the traceroute path. The presented methodologies within this section attempt to account for these difficulties by corroborating domain knowledge for common networking practices and relying on a collection of traceroute paths and their corresponding router view (obtained by using alias resolution techniques) to make accurate inferences of the entities which are establishing inter-AS links. Furthermore, pin-pointing routers to physical locations was the key enabler for highlighting remote peerings that are simply not visible from an AS-level topology.

C. *PoP-Level Topology*

PoP-level topologies present a middle ground between AS-level and router-level topologies. A PoP-level graph presents the points of presence for one or many networks. These topologies inherently have geo information at the granularity of metro areas embedded within. They have been historically at the center of focus as many ASes disclose their topologies at a PoP level granularity and do not require detailed information regarding each individual router and merely represent a bundle of routers within each PoP as a single node. They have lost their traction to router-level topologies that are able to capture the dynamics of these topologies in addition to providing finer details of information. Regardless of this, due to the importance of some ASes and their centrality in the operation of today's Internet, several studies [66], [67], [68] outlining the internal operation of these ASes within each PoP have emerged. These studies offer insight into the challenges these ASes face for peering and serving the vast majority of the Internet as well as the solutions that they have devised.

Cunha et al. [13] develop *Sibyl*, a system which provides an expressive interface that allows the user to specify the requirements for the path of a traceroute, given the set of requirements *Sibyl* would utilize all available vantage points and rely on historical data to conduct a traceroute from a given vantage point towards a specific destination that is most likely to satisfy the users constraints. Furthermore, given that each vantage point has limited probing resources and that concurrent requests can be made, *Sibyl* would pick source-destination pairs which optimize for resource utilization. *Sibyl* combines PlanetLab, RIPE Atlas, traceroute servers accessible through looking glasses, DIMES, and Dasu measurement platforms to maximize its coverage. Symbolic regular expressions are used for the query interface where the user can express path properties such as the set of traversed ASes, cities, and PoPs. The likelihood of each source-destination pair matching the required path properties is calculated using a supervised machine learning technique (RuleFit) which is trained based on prior measurements and is continuously updated based on new measurements. Resource utilization optimization is addressed by using a greedy algorithm, *Sibyl* chooses to issue traceroutes that fit the required budget and that have the largest marginal expected utility based on the output of the trained model.

Schlinker et al. [67] outline Facebook's edge fabric within their PoPs by utilizing an SDN based system that alters BGP local-pref attributes to utilize alternative paths towards specific prefixes better. The work is motivated by BGP's shortcomings namely, lack of awareness of link capacities and incapability to optimize path selections based on various performance metrics. More specifically BGP makes its forwarding decisions using a combination of AS-path length and the local-perf metric. Facebook establishes BGP connections with other ASes through various means namely, private interconnections, public peerings through IXPs, and peerings through router servers within IXPs. The authors report that the majority of their interconnections are established through public peerings while the bulk of traffic is transmitted over the private links. The later reflects Facebook's preference to select private peerings over public peerings while peerings established through route servers have the lowest priority. Furthermore, the authors observe that for all PoPs except one, all prefixes have at least two routes towards each destination prefix. The proposed solution isolates the traffic engineering per PoP to simplify the design, the centralized SDN controller within each PoP gathers router RIB tables through a BMP collector. Furthermore, traffic statistics are gathered through sampled sFlow or IPFIX records. Finally, interface information is periodically pulled by SNMP. The collector emulates BGP's best path selection and projects interface utilization. For overloaded interfaces prefixes with alternative routes are selected, an alternative route is selected based on a set of preferences. The output of this step generates a set of route overrides which are enforced by setting a high local-pref value for them. The authors report that their deployed system detours traffic from 18% of interfaces. The median of detour time is 22 minutes and about 10% of detours last as long as 6 hours. The detoured routes resulted in 45% of the prefixes achieving a median latency improvement of 20ms while 2% of prefixes improved their latency by 100ms.

Yap et al. [66] discuss the details of Espresso, an application-aware routing system for Google's peering edge routing infrastructure. Similar to the work of Schlinker et at. [67] Espresso is motivated by the need for a more efficient (both technically and economically) edge peering fabric that can account for traffic engineering constraints. Unlike the work of Schlinker et al. [67] Espresso maintains two layers of control plane one which is localized to each PoP while the other is a global centralized controller that allows Google to perform further traffic optimizations. Espresso relies on commodity MPLS switches for peering purposes, traffic between the switches and servers are encapsulated in IP-GRE and MPLS headers. IP-GRE header encodes the correct switch, and the MPLS header determines the peering port. The global controller (GC) maintains an egress map that associates each client prefix and PoP tuple to an edge router/switch and egress port. User traffic characteristics such as throughput, RTT, and re-transmits are reported at a /24 granularity to the global controller. Link utilization, drops, and port speeds are also reported back to the global controller. A greedy algorithm is used by the GC to assign traffic to a candid router port combination. The greedy algorithm starts by making its decisions using traffic priority metrics and orders its available options based on BGP policies, user traffic metrics, and the cost of serving on a specific link. Espresso has been incrementally deployed within Google and at the time of the study was responsible for serving about 22% of traffic. Espresso is able to maintain higher link utilization while maintaining low packet drop rates even for fully utilized links (95% less than 2.5%). The authors report that the congestion reaction feature of the GC results in higher goodput and mean time between re-buffers for video traffic.

Wohlfart et al. [68] present an in-depth study of the connectivity fabric of Akamai at its edge towards its peers. The authors account 3.3k end-user facing (EUF) server deployments with varying size and capabilities which are categorized into four main groups. Two of these groups have Akamai border routers and therefore establish explicit peerings with peers and deliver content directly to them while the other two groups are hosted within another ASes network and are responsible for delivering content implicitly to other peers. Customers are redirected to the correct EUF server through DNS, the mapping is established by considering various inputs including BGP feeds collected by Akamai routers, user performance metrics, and link cost information. To analyze Akamai's peering fabric, the authors rely on proprietary BGP snapshots obtained from Akamai routers and consist of 3.65M AS paths and about 1.85M IPv4 and IPv6 prefixes within 61k ASes (ViewA). As a point of comparison, a combination of daily BGP feeds from Routeviews, RIPE RIS, and PCH consisting of 21.1M AS paths and 900k prefixes within 59k ASes is used (ViewP). While at an AS level both datasets seem to have a relatively similar view, ViewA (ViewP) observes 1M (0.1M) prefixes the majority of which are prefixes longer than /25. Only 15% of AS paths within ViewP are observed by ViewA which suggests that a large number of AS paths within ViewP are irrelevant for the operation of Akamai. Wohlfart et al. report 6.1k unique explicit peerings between Akamai and its neighbors by counting the unique number of next-hop ASN from the Akamai BGP router dumps. About 6k of these peerings happen through IXPs while the remainder are established through PNIs. In comparison, only 450 peerings between Akamai and other ASes are observed through ViewP. Using AS paths within ViewP the authors report about 28k implicit peers which are within one AS hop from Akamai's network. Lastly, the performance of users sessions are looked into by utilizing EUF server logs containing the clients IP address, throughput, and a smoothed RTT value. The performance statistics are presented for two case studies (i) serving a single ISP and (ii) serving customers within 6 distinct metros. Overall 90% of traffic is coming from about 1% of paths and PNIs are responsible for delivering the bulk of traffic and PNIs and cache servers within eyeball ASes achieve the best performance regarding RTT.

Nur et al. [69] study the Internet AS-level topology using a multigraph representation where AS pairs can have multiple edges between each other. Traceroute measurements from CAIDA's Ark and iPlane projects are collected for this study. For IP to AS mapping Routeviews' BGP feed is utilized. Next hop addresses for BGP announcements are extracted from Routeviews as well as RIPE RIS. For mapping IP addresses to their corresponding geo-location various data sources have been employed namely, (i) UNDNS for DNS parsing, (ii) DB-IP, (iii) Maxmind GeoLite2 City, and (iv) IP2Location DB5 Lite.

Each ASes border interface is identified by tracking ASN changes along the hops of each traceroute. Each cross border interface X-BI is geolocated to the city in which it resides by applying one of the following methods in order of precedence: (i) relying on UNDNS for extracting geoinformation from reverse DNS names, (ii) majority vote along three (DB-IP, Maxmind, and IP2Location) IP to GEO location datasets, (iii) sandwich method where an unresolved IP between two IPs in the same geolocation is mapped to the same location, (iv) RTT based geo locating which relies on the geolocation of prior or next hops of an unresolved address that have a RTT difference smaller than 3 ms for mapping them to the same location, and (v) if all of the prior methods fail Maxmind's output is used for mapping the geolocation of the X-BI. The set of inter-AS links resulting from parsing traceroutes is augmented by benefiting from BGP data. If an AS relationship exists between two ASes but is missing from the current AS-level graph and all identified X-BIs corresponding to these ASes are geolocated to a single city, a link will be added to the AS-level topology graph under the assumption that this is the only possible location for establishing an interconnection between these two ASes.

The inferred PoP nodes in the AS graph are validated for major research networks as well as several commercial ISPs. The overlap of identified PoPs is measured for networks which have publicly available PoP-level maps. The maps align with the set of identified cities by X-AS with deviations in terms of number of PoPs per city. This is a limitation of X-AS as it is only able to identify one PoP per city. Identified AS-links are compared against CAIDA's AS relationships dataset, the percentage of discrepancy for AS links of each AS is measured. For 78% of ASes, the maps agree with each other completely, and the average link agreement is about 85% for all ASes. Various properties of the resulting graph are analyzed in the paper, the authors find that the number of X-BI nodes per AS, X-BI nodes degree, and AS degree all follow a power law distribution.

Summary: PoP-level topologies can offer a middle ground between router-level and AS-level topologies offering an understanding of inter-AS peering relationships while also being able to distinguish instances of these peerings happening at various geo-locations/PoPs. Additionally, we reviewed studies that elaborate on the faced challenges as well as the devised solutions for content provider (Google, Facebook) and CDN (Akamai) networks which are central to the operation of today's Internet.

*D. **Physical-Level Topology***

This subsection is motivated by the works of Knight et al. [70] and Durairajan et al. [71], [72], [73] which presented the groundwork for having a comprehensive physical map of the Internet consisting of edges corresponding to fiber optic cables providing connectivity between metro areas and PoPs as nodes within these topologies. A sample of this topology for CenturyLink's fiber-optic backbone network within the continental US is presented in Figure 5. Physical maps were mostly neglected by the Internet topology community mainly due to two reasons: (i) the scarcity of well-formatted information and (ii) the complete disassociation of physical layers from probes conducted within higher layers of the TCP/IP stack. The following set of papers try to address the former issue by gathering various sources of information and compiling them into a unified format.

Knight et al. [70] present the Internet topology Zoo which is a collection of physical maps of various networks within the Internet. The authors rely on ground truth data publicly provided by the network operators on their websites. These maps are presented in various formats such as static images or flash objects. The authors transcribe all maps using yEd (a graph editor and diagraming program) into a unified graph specification format (GML) and annotate nodes and links with any additional information such as link speed, link type, longitude, and latitudes that is provided by these maps. Each map and its corresponding network is classified as a backbone, testbed, customer, transit, access or internet exchange based on the properties of their network. For example, backbone networks should connect at least two cities together while access networks should provide edge access to individuals. A total of 232 networks are transcribed by the authors. About 50% of networks are found to have more than 21 PoPs and each of these PoPs have an average degree of about 3. Lastly similar to [49] the core density of networks is examined by measuring the 2-core size of networks. A wide degree of 2-core sizes ranging from 0 (tree-like networks) to 1 (densely connected core with hanging edges) are found within the dataset.

Durairajan et al. [71] create a map of the physical Internet consisting of nodes representing colocation facilities and datacenters, links representing conduits between these nodes and additional metadata related to these entities. The authors rely on publicly available network maps (images, Flash objects, Google Maps overlays) provided by ASes. The methodology for transcribing images consists of 5 steps: (i) capturing high-resolution sub-images, (ii) patching sub-images into a composite image, (iii) extracting a link image using color masking techniques, (iv) importing link image into ArcGIS using geographic reference points, and (v) using link vectorization in ArcGIS to convert links into vectors. Given that each map has a different geo resolution, different scores are attributed to nodes with lat/lon or street level, city, and state having a corresponding score of 1.0, 0.75, 0.5. All maps have at least city level resolution with about 20% of nodes having lat/lon or street level accuracy.

Durairajan et al. [72] work is motivated by two research

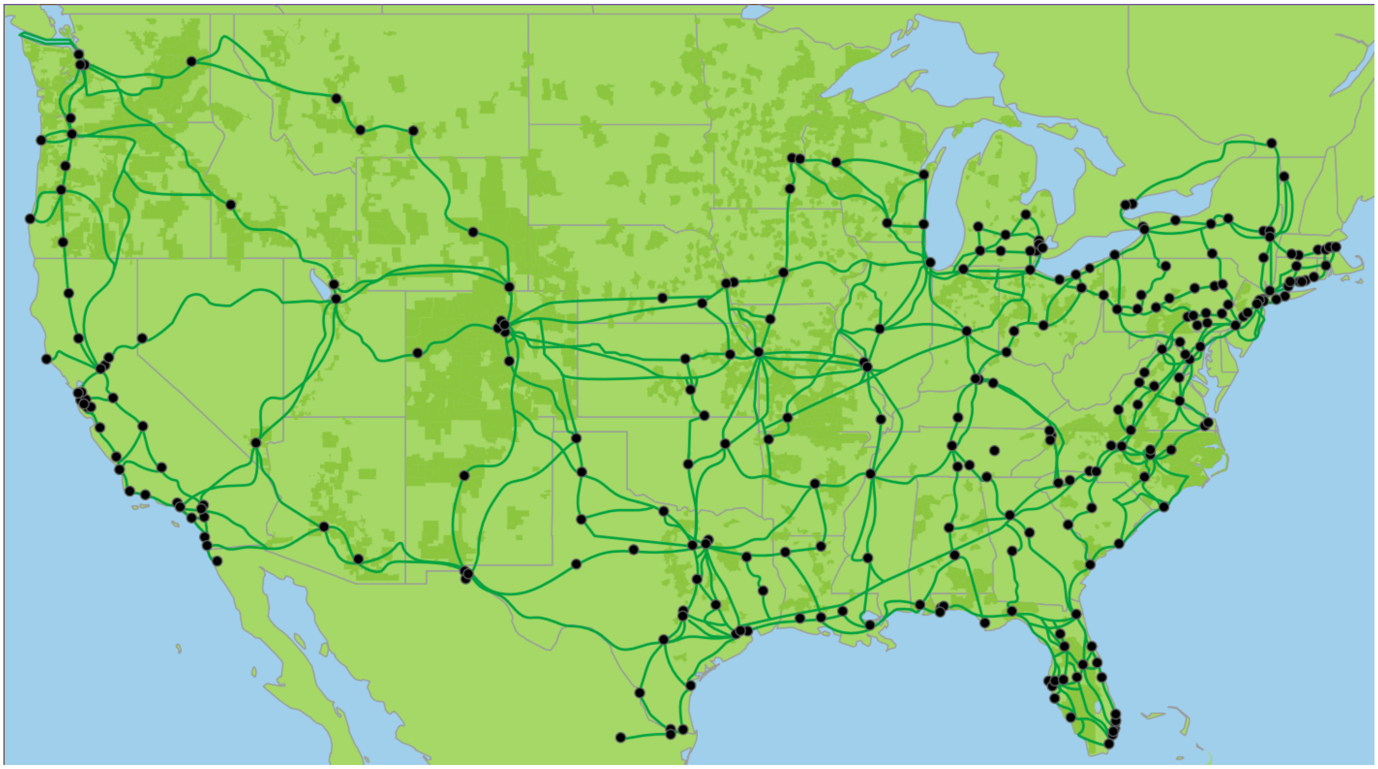

Fig. 5. Fiber optic backbone map for CenturyLink's network in continental US. Each node represents a PoP for CenturyLink while links between these PoPs are representative of the fiber optic conduits connecting these PoPs together. Image courtesy of CenturyLink.

questions: (i) how do physical layer and network layer maps compare with each other? and (ii) how can probing techniques be improved to reveal a larger portion of physical infrastructure? For physical topologies, the authors rely on maps which are available from the Internet Atlas project. From this repository the maps for 7 Tier-1 networks and 71 non-Tier-1 networks which are present in North America are gathered, these ASes collectively consist of 2.6k PoPs and 3.6k links. For network layer topologies, traceroutes from the CAIDA Ark project during the September 2011 to Match 2013 period are used. Additionally DNS names for router interfaces are gathered from the IPv4 Routed /24 DNS Names Dataset which includes the domain names for IP addresses observed in the CAIDA Ark traceroutes. Traceroute hops are annotated with their corresponding geo information (extracted with DDeC) as well as the AS number which is collected from TeamCymru's service. Effects of vantage point selection on node identification are studied by employing public traceroute servers. Different modalities depending on the AS ownership of the traceroute server and the target address are considered ([VP_in, t_in], [VP_in, t_out], [VP_out, t_in]). Their methodology (POPsicle) chooses VPs based on geo proximity towards the selected targets and along the pool of destinations, those which have a square VP to destination distance greater than the sum of squares of the distance between target VP and destination are selected to create a measurement cone. For this study 50 networks that have a comprehensive set of geo-information for their physical map are considered. Out of these 50 networks, 21 of them do not have any geo information embedded in their DNS names. Furthermore, 16 ASes were not observed in the Ark traces. This results in 13 ASes out of the original 50 which have both traces and geo-information in the network layer map. POPsicle was deployed in an IXP (Equinix Chicago) to identify the PoPs of 10 tenants. Except for two networks, POPsicle was able to identify all known PoPs of these networks. Furthermore, POPsicle was evaluated by targeting 13 ISPs through Atlas probes which were deployed in IXPs, for all of these ISPs POPsicle was able to match or outperform Ark and Rocketfuel. Furthermore for 8 of these ISPs POPsicle found all or the majority of PoPs present in Atlas maps.

Durairajan et al. [73] obtain the long-haul fiber network within the US and study its characteristics and limitations. For the construction of the long-haul fiber map, Durairajan et al. rely on the Internet Atlas project [71] as a starting point and confirm the geo-location or sharing of conduits through legal documents which outline laying/utilization of infrastructure. The methodology consists of four steps: (i) using Internet Atlas maps for tier-1 ASes that have geo-coded information, a basic map is constructed, (ii) the geolocation of nodes and links for the map is confirmed through any form of legal document which can be obtained, (iii) the map is augmented with additional maps from large transit ASes which lack any geo-coded information, (iv) the augmented map is once again confirmed through any legal document that would either confirm the geolocation of a node/link or would indicate conduit sharing with links that have geo-coded information. The long-haul fiber map seems to be physically aligned with roadway and railway routes, the authors use the polygon overlap feature of ArcGIS to compare the overlap of these maps and find that most often long-hauls run along roadways. The authors also assess shared conduit risks, for this purpose they construct a conduit sharing matrix were rows are ASes and columns are conduits the value within each row indicates the number of ASes which are utilizing that conduit. Out of 542 identified conduits about 90% of them are shared by at least one other AS. Using the risk matrix the hamming distance for each AS pair is measured to identify ASes which have similar risk profiles. Using traceroute data from Edgescape and parsing geoinformation in domain names the authors infer which conduits were utilized by each traceroute and utilize the frequency of traceroutes as a proxy measure of traffic volume. Finally a series of risk mitigation analysis are conducted namely: (i) the possibility of increasing network robustness by utilizing available conduits or by peering with other networks is investigated for each AS (ii) increasing network robustness through the addition of additional k links is measured for each network, and lastly (iii) possibility for improving latency is investigated by comparing avg latencies against right of way (ROW), line of sight (LOS), and best path delays.

Summary: the papers within this sub-section provided an overview of groundbreaking works that reveal physical-level topologies of the Internet. The researchers gathered various publicly available maps of ASes as well as legal documents pertaining to the physical location of these networks to create a unified, well-formatted repository for all these maps. Furthermore, the applicability of these maps towards the improvement of targeted probing methodologies and the possibility of improving and provisioning the infrastructure of each network is investigated. Although the interplay of routing on top of these physical topologies is unknown and remains as an open problem, these physical topologies provide complementary insight into the operation of the Internet and allow researchers to provision or design physical infrastructure supporting lower latency Internet access or to measure the resiliency of networks towards natural disasters.

V. Implications & Applications of Network Topology

This section will provide an overview of the studies which rely on Internet topology to provide additional insight regarding the performance, resiliency, and various characteristics of the Internet. The studies which are outlined in this section look into various properties of the Internet including but not limited to: path length both in terms of router and AS hops, latency, throughput, packet loss, redundancy, and content proximity. In a more broad sense, we can categories these studies into three main groups: *(i) studying performance characteristics of the Internet*, *(ii) studying resiliency of the Internet*, and *(iii) classifying the type of inter-AS relationships between ASes*. Depending on the objective of the study one or more of the aforementioned properties of the Internet could be the subject which these studies focus on. Each of these studies would require different resolutions of Internet topology. As outlined in Section IV obtaining a one to one mapping between different resolutions is not always possible. For example, each AS link can correspond to multiple router level links while each router level link can correspond to multiple physical links. For this reason, each study would rely on a topology map which better captures the problems objectives. As an example, studying the resiliency of a transit ASes backbone to natural disasters should rely on a physical map while performing the same analyses using an AS-level topology could lead to erroneous conclusions given the disassociation of ASes to physical locations. While on the other hand studying the reachability and visibility of an AS through the Internet would require an AS-level topology and conducting the same study using a fiber map would be inappropriate as the interplay of the global routing system on top of this physical map is not known. The remainder of this section would be organized into three sub-sections presenting the set of studies which focus on the *(i) Internet performance*, *(ii) Internet resiliency*, and *(iii) AS relationship classification*. Furthermore, each sub-section would further divide the studies based on the granularity of the topology which is employed.

A. ***Performance***

Raw performance metrics such as latency and throughput can be conducted using end-to-end measurements without any attention to the underlying topology. While these measurements can be insightful on their own, gaining a further understanding of the root cause of subpar performance often requires knowledge of the underlying topology. For example, high latency values reported through end-to-end measurements can be a side effect of many factors including but not limited to congestion, a non-optimal route, an overloaded server, and application level latencies. Many of these underlying causes can only be identified by a correct understanding of the underlying topology. Congestion can happen on various links along the forward and reverse path, identifying the faulty congested link or more specifically the inter-AS link requires a correct mapping for the traversed topology. Expanding infrastructure to address congestion or subpar latency detected through end-to-end measurements is possible through an understanding of the correct topology as well as the interplay of routing on top of this topology. In the following Section, we will present studies that have relied on router, AS, and physical level topologies to provide insight into various network performance related issues.

1) ***AS-Level Topology****:* Studies in this section rely on BGP feeds as well as traceroute probes that have been translated to AS paths to study performance characteristics such as increased latency and path lengths due to insufficient network infrastructure within Africa [74], [75], path stability and the latency penalties due to AS path changes [76], IXPs centrality in Internet connectivity as a means for reducing path distances towards popular content [77], and estimating traffic load on inter-AS links through the popularity of traversed paths [78]. Chatzis et al. [77] demonstrate the centrality of a large European IXP in the Internet's traffic by relying on sampled sFlow traces captured by the IXP operator. Peering relationships are identified by observing BGP as well as regular traffic being exchanged between tenant members. The authors limit their focus to web traffic as it constitutes the bulk of traffic which is observed over the IXP's fabric. Endhost IP addresses are mapped to the country which they reside in by using Maxmind's IP to GEO dataset. The authors observe traffic from nearly every country (242 out of 250). While tenant ASes generated the bulk of traffic, about 33% of traffic originated from ASes which were one or more hops away from the IXP. The authors find that recurrent IP addresses generate about 60% of server traffic. Finally, the authors highlight the heterogeneity of AS traffic by identifying servers from other ASes which are hosted within another AS. Heterogeneous servers are identified by applying a clustering algorithm on top of the SOA records of all observed IP addresses. Lastly, the share of heterogeneous traffic on inter-AS links is presented for Akamai and Cloudflare. It is found that about 11% (54%) of traffic (servers) are originated (located) within 3rd-party networks.

Sanchez et al. [78] attempt to characterize and measure inter-domain traffic by utilizing traceroutes as a proxy measure. Traceroute probes towards random IP addresses from the Ono BitTorrent extension are gathered over two separate months. Ground truth data regarding traffic volume is obtained from two sources: (i) sampled sFlows from a large European IXP and (ii) link utilization for the customers of a large ISP presenting the 95th percentile of utilization using SNMP.

AS-link traversing paths (ALTP) are constructed by mapping each hop of traceroutes to their corresponding ASN. For each ALTP-set a relative measure of link frequency is defined which represents the cardinality of the link to the sum of cardinalities of all links in that set. This measure is used as a proxy for traffic volume. The authors measure different network syntax metrics namely: connectivity, control value, global choice, and integration for the ALTP-sets which have common links with their ground truth traffic data. r^2 is measured for regression analysis of the correlation between network syntax metrics and traffic volume. ALTP-frequency shows the strongest correlation with r^2 values between 0.71 - 0.97 while the remainder of metrics also show strong and very-strong correlations. The authors utilize the regression model to predict traffic volume using ALTP-frequency as a proxy measure. Furthermore Sanchez et al. demonstrate that the same inferences cannot be made from a simple AS-level connectivity graph which is derived from BGP streams. Finally, the authors apply the same methodology to CAIDA's Ark dataset and find similar results regarding the correlation of network syntax metrics and traffic volume.

Gupta et al. [74] study circuitous routes in Africa and their degrading effect on latency. Circuitous routes are between two endpoints within Africa that traverse a path outside of Africa, i.e. the traversed route should have ideally remained within Africa but due to sub-par connectivity has detoured to a country outside of Africa. Two major datasets are used for the study, (i) BGP routing tables from Routeviews, PCH, and Hurricane Electric, and (ii) periodic (every 30 minutes) traceroute measurements from BISmark home routers towards MLab servers, IXP participants, and Google cache servers deployed across Africa. Traceroute hops are annotated with their AS owner and inter-AS links are identified with the observation of ASN changes along the path. Circuitous routes are identified by relying on high latency values for the given path. Latency penalty is measured as the ratio of path latency to the best case latency between the source node and a node in the same destination city. The authors find two main reasons for paths with high latency penalty values namely, (i) ASes along the path are not physically present at a local IXP, or (ii) the ASes are present at a geographically closeby IXP but do not peer with each other due to business preferences.

Fanou et al. [75] study Internet topology and its characteristics within Africa. By expanding RIPE's Atlas infrastructure within African countries, the authors leverage this platform to conduct traceroute campaigns with the intention of uncovering as many as possible AS paths. To this end, periodic traceroutes were ran between all Atlas nodes within Africa. These probes would target both IPv4 and IPv6 addresses if available. Traceroute hops were mapped to their corresponding country by leveraging six public datasets, namely OpenIPMap, MaxMind, Team Cymru, AFRINIC DB, Whois, and reverse DNS lookup. Upon disagreement between datasets, RIPE probes within the returned countries were employed to measure latency towards the IP addresses in question, the country with the lowest latency was selected as the host country. Interface addresses are mapped to their corresponding ASN by utilizing Team Cymru's IP to AS service [79], using the augmented traceroute path the AS path between the source and destination is inferred. Using temporal data the preference of AS pairs to utilize the same path is studied, 73% (82%) of IPv4 (IPv6) paths utilize a path with a frequency higher than 90%. Path length for AS pairs within west and south Africa are studied, with southern countries having a slightly shorter average path of 4 compare to 5. AS path for pairs of addresses which reside within the same country in each region is also measured where it's found that southern countries have a much shorter path compared to pairs of addresses which are in the same western Africa countries (average of 3 compared to 5). AS-centrality (percentage of paths which AS appears in and is not the source or destination) is measured to study transit roles of ASes. Impact of intercontinental transit on end-to-end delay is measured by identifying the IP path which has the minimum RTT. It is found that intercontinental paths typically exhibit higher RTT values while a small fraction of these routes still have relatively low RTT values ($< 100ms$) and are attributed to inaccuracies in IP to geolocation mapping datasets.

Green et al. [76] leverage inter-AS path stability as a measure for conducting Internet tomography and anomaly detection. Path stability is analyzed by the stability of a *primary path*. The primary path of router *r* towards prefix *p* is defined as the most prevalent preferred path by *r* during the window time-frame of *W*. Relying on 3 months of BGP feeds from RIPE RIS' LINX collector it is demonstrated that 85% (90%) of IPv4 (IPv6) primary paths are in use for at least half of the time. Any deviation from the primary path are defined as pseudo-events which are further categorized into two groups: (i) transient events where a router explores additional paths before reconverging to the primary path, and (ii) structural events where a router consistently switches to a new primary path. For each pseudo-event, the duration and set of new paths that were explored are recorded. About 13% of transient pseudo-events are found to be longer than an hour while 12% of structural pseudo-events last less than 7 days. The number of explored paths and the recurrence of each path is measured for pseudo-events. It is found that MRAI timers and route flap damping are efficient at regulating BGP dynamics. However, these transient events could be recurrent and require more complex mechanisms in order to be accounted for. For anomaly detection about 2.3k AS-level outages and hijack events reported by BGPmon during the same period of the study are used as ground truth. About 84% of outages are detected as pseudo-events in the same time window while about 14% of events the detection time was about one hour earlier than what BGPmon reported. For hijacks, the announced prefix is looked-up amongst pseudo-events if no match is found less specific prefixes are used as a point of comparison with BGPmon. For about 82% of hijacking events, a matching pseudo-event was found, and the remainder of events are tagged as explicit disagreements.

*2) **Router-Level Topology**:* With the rise of peering disputes highlighted by claims of throttling for Netflix's traffic access to unbiased measurements reflecting the underlying cause of subpar performance seems necessary more than before. Doing so would require a topology map which captures inter-AS links. The granularity of these links should be at the router level since two ASes could establish many interconnections with each other, each of which could exhibit different characteristics in terms of congestion. As outlined in Section IV various methodologies have been presented that enable researchers to infer the placement of inter-AS links from data plane measurements in the form of traceroutes. A correct assessment of the placement of inter-AS links is necessary to avoid attributing intra-AS congestion to inter-AS congestion, furthermore incorrectly identifying the ASes which are part of the inter-AS link could lead to attributing congestion to incorrect entities.

Dhamdhere et al. [80] rely on prior techniques [61] to infer both ends of an interconnect link and by conducting time series latency probes (TSLP) try to detect windows of time where the latency time series deviates from its usual profile. Observing asymmetric congestion for both ends of a link is attributed to inter-AS congestion. The authors deploy 86 vantage points within 47 ASes on a global scale. By conducting similar TSLP measurements towards the set of identified inter-AS links over the span of 21 months starting at March 2016, the authors study congestion patterns between various networks and their upstream transit providers as well the interconnections they establish with content providers. Additionally, the authors conduct throughput measurements using the Network Diagnostic Tool (NDT) [81] as well as SamKnows [82] throughput measurements of Youtube servers and investigate the correlation of inter-AS congestion and throughput.

Chandrasekaran et al. [83] utilize a large content delivery networks infrastructure to assess the performance of the Internet's core. The authors rely on about 600 servers spanning 70 countries and conduct pairwise path measurements in both forward and reverse directions between the servers. Furthermore, AS paths are measured by translating router hop interfaces to their corresponding AS owner, additionally inter-AS segments are inferred by relying on a series of heuristics developed by the authors based on domain knowledge and common networking practices. Latency characteristics of the observed paths are measured by conducting periodic ping probes between the server pairs. Consistency and prevalence of AS paths for each server pair are measured for a 16 month period. It is found that about 80% of paths are dominant for at least half of the measurement period. Furthermore, about 80% of paths experience 20 or fewer route changes during the 16 month measurement period. The authors measure RTT inflation in comparison to optimal AS paths and find that sub-optimal paths are often short-lived although a small number (10%) of paths experience RTT inflation for about 30% of the measurement period. Effects of congestion on RTT inflation are measured by initially selecting the set of server pairs which experience RTT inflation using ping probe measurements while the first segment that experiences congestion is pinpointed by relying on traceroute measurements which are temporally aligned with the ping measurements. The authors report that most inter and intra-AS links experience about 20 to 30 ms of added RTT due to congestion.

Chiu et al. [84] assess path lengths and other properties for paths between popular content providers and their clients. A collection of 4 datasets were used throughout the study namely: (i) iPlane traceroutes from PlanetLab nodes towards 154k BGP prefixes, (ii) aggregated query counts per /24 prefix (3.8M) towards a large CDN, (iii) traceroute measurements towards 3.8M + 154k prefixes from Google's Compute Engine (GCE), Amazon Elastic Cloud, and IBM's Softlayer VMs, and (iv) traceroutes from RIPE Atlas probes towards cloud VMs and a number of popular websites. Using traceroute measurements from various platforms and converting the obtained IP hop path to its corresponding AS-level path the authors assess the network distance between popular content providers and client prefixes. iPlane traceroutes are used as a baseline for comparison, only 2% of these paths are one hop away from their destination this value increases to 40% (60%) for paths between GCE and iPlane (end user prefixes). This indicates that Google peers directly with the majority of networks which host its clients. Using the CDN logs as a proxy measure for traffic volume the authors find that Google peers with the majority of ASes which carry large volumes of traffic. Furthermore Chiu et al. find that the path from clients towards *google.com* due to off-net hosted cache servers is much shorter where 73% of queries come from ASes that either peer with Google or have an off-net server in their network or their upstream provider. A similar analysis for Amazon's EC2 and IBM's Softlayer was performed each having 30% and 40% one hop paths accordingly.

Kotronis et al. [85] study the possibility of improving latency performance through the employment of relay nodes within colocation facilities. This work tries to (i) identify the best locations/colos to place relay nodes and (ii) quantify the latency improvements that are attainable for end pairs. The authors select a set of ASes per each country which covers at least 10% of the countries population by using APNIC's IPv6 measurement campaign dataset [86]. RIPE Atlas nodes within these AS country pairs are selected which are running the latest firmware, are connected and pingable, and have had stable connectivity during the last 30 days. Colo relays are selected by relying on the set of pinned router interfaces from Giotsas et al. [64] work. Due to the age of the dataset, a series of validity tests including conformity with PeeringDB data, pingability, consistent ASN owner, and RTT-based geolocation test with Periscope LGs have been conducted over the dataset to filter out stale information. A set of PlanetLab relays and RIPE Atlas relays are also considered as reference points in addition to the set of colo relays. The measurement framework consists of 30 minute rounds between April 20th - May 17th 2017. Within each round, ping probes are sent between the selected end pairs to measure direct latency. Furthermore, the relay paths latency is estimated by measuring the latency between the <src, relay> and <dst, relay> pairs. The authors observed improve latency for 83% of cases with a median of 12-14ms between different relay types. Colo relays having the largest improvement. The number of required relays for improved latency is measured, the authors find that colo relays have the highest efficiency where 10 relays account for 58% of improved cases while the same number of improved cases for RIPE relays would require more than 100 relays. Lastly, the authors list the top 10 colo facilities which host the 20 most effective colo relays, 4 of these color are in the top 10 PeeringDB colos in terms of the number of colocated ASes and all host at least 2 or more IXPs within them.

Fontugne et al. [87] introduce a statistical model for measuring and pin-pointing delay and forwarding anomalies from traceroute measurements. Given the prevalence of route asymmetry on the Internet, measuring the delay of two adjacent hops is not trivial. This issue is tackled by the key insight that differential delay between two adjacent hops is composed of two independent components. Changes in link latency can be detected by having a diverse set of traceroute paths that traverse the under study link and observing latency values disrupting the normal distribution for latency median. Forwarding patterns for each hop are established by measuring a vector accounting for the number of times a next hop address has been observed. Pearson product-moment correlation coefficient is used as a measure to detect deviations or anomalies within the forwarding pattern of a hop. RIPE Atlas' *built-in* and *anchoring* traceroute probes for an eight-month period in 2015 are used for the study. The authors highlight the applicability of their proposed methodology by providing insight into three historical events namely, DDoS attacks on DNS root servers, Telekom Malaysia's BGP route leak, and Amsterdam IXP outage.

*3) **Physical-Level Topology**:* Measuring characteristics of physical infrastructure using data plan measurements is very challenging due to the disassociation of routing from the physical layer. Despite these challenges, we overview two studies within this section that investigate the effects of sub-optimal fiber infrastructure on latency between two end-points [88] and attempt to measure and pinpoint the causes of observing subpar latency within fiber optic cables [89].

Singala et al. [88] outline the underlying causes of sub-par latency within the Internet. The authors rely on about 400 Planet Lab nodes to periodically fetch the front page of popular websites, geolocate the webserver's location and measure the optimal latency based on speed of light (*c-latency*) constraints. Interestingly the authors find that the median of latency inflation is about 34 times greater than *c-latency*. Furthermore, the authors breakdown the webpage fetch time into its constituent components namely, DNS resolution, TCP handshake, and TCP transfer. Router path latency is calculated by conducting traceroutes towards the servers and lastly, minimum latency towards the web server is measured by conducting periodic ping probe. It is found that the median of router paths experience about 2.3x latency inflation. The authors hypothesize that latencies within the physical layer are due to sub-optimal fiber paths between routers. The validity of this hypothesis is demonstrated by measuring the pairwise distance between all nodes of Internet2 and GEANT network topologies and also computing road distance using Google Maps API. It is found that fiber links are typically 1.5-2x longer than road distances. While this inflation is smaller in comparison to webpage fetching component's latency the effects of fiber link inflation are evident within higher layers due to the stacked nature of networking layers.

Bozkurt et al. [89] present a detailed analysis of the causes for sub-par latency within fiber networks. The authors rely on Durairajan et al. [72] InterTubes dataset to estimate fiber lengths based on their conduits in the dataset. Using the infrastructure of a CDN, server clusters which are within a 25km radius of conduit endpoints were selected, and latency probes between pairs of servers at both ends of the conduit were conducted every 3 hours for the length of 2 days. The conduit length is estimated using the speed of light within fiber optic cables (f-latency), and the authors find that only 11% of the links have RTTs within 25% of the f-latency for their corresponding conduit. Bozkurt et al. enumerate various factors which can contribute to the inflated latency that they observed within their measurements namely, (i) refraction index for different fiber optic cables varies, (ii) slack loops within conduits to account for fiber cuts, (iii) latency within optoelectrical and optical amplifier equipment, (iv) extra fiber spools to compensate for chromatic dispersion, (v) publication of mock routes by network operators to hide competitive details, and (vi) added fiber to increase latency for price differentiation. Using published latency measurements from AT&T and CenturyLink RTT inflation in comparison to f-latency from InterTubes dataset is measured to have a median of 1.5x (2x) for AT&T and CenturyLink's networks. The accuracy of InterTubes dataset is verified for Zayo's network. Zayo published detailed fiber routes on their website. The authors find great conformity for the majority of fiber conduit lengths while for 12% of links the length difference is more than 100km.

*B. **Resiliency***

Studying the resiliency of Internet infrastructure has been the subject of many types of research over the past decade.

While many of these studies have reported postmortems regarding natural disasters and their effects on Internet connectivity, others have focused on simulating what-if scenarios to examine the resiliency of the Internet towards various types of disruptions. Within these studies, researchers have utilized Internet topologies which were contemporary to their time. The resolution of these topologies would vary in accordance with the stated problem. For example, the resiliency of long haul fiber infrastructure to rising sea levels due to global warming is measured by relying on physical topology maps [90] while the effects of router outages on BGP paths and AS reachability is studied using a combination of router and AS level topologies within Luckie et al. work [91]. The remainder of this section is organized according to the resolution of the underlying topology which is used by these studies.

*1) **AS-Level Topology**:* Katz-Bassett et al. [92] propose LIFEGUARD as a system for recovering from outages by rerouting traffic. Outages are categorized into two groups of forward and reverse path outages. Outages are detected and pinpointed by conducting periodic ping and traceroute measurements towards the routers along the path. A historical list of responsive routers for each destination is maintained. Prolonged unresponsive ping probes are attributed to outages. For forward path outages, the authors suggest the use of alternative upstream providers which traverse AS-paths that do not overlap with the unresponsive router. For reverse path outages, the authors propose a BGP poisoning solution where the origin AS would announce a path towards its own prefix which includes the faulty AS within the advertised path. This, in turn, causes the faulty AS to withdraw the advertisement (to avoid a loop) of the prefix and therefore cause alternative routes to be explored in the reverse path. A less-specific sentinel prefix is advertised by LIFEGUARD to detect the recovery of the previous path.

Luckie et al. [91] correlate BGP outage events to inferred router outages by relying on time-series of IPID values obtained through active measurements. This work is motivated by the fact that certain routers rely on central incremental counters for the generated IPID values, given this assumption one would expect to observe increasing IPID values for a single router. Any disruption in this pattern can be linked to a router reboot. IPID values for IPv4 packets are susceptible to counter rollover since they are only 16 bits wide. The authors rely on IPID values obtained by inducing fragmentation within IPv6 packets. The authors rely on a hit list of IPv6 router addresses which is obtained from intermediate hops of CAIDA's Ark traceroute measurements. By analyzing the time series of IPID values, an outage window is defined for each router. Router outages are correlated with their corresponding BGP control plane events by looking at BGP feeds and finding withdrawal and announcement messages occurring during the same time frame. It is found that for about 50% of router/prefix pairs at least 1-2 peers withdrew the prefix and nearly all peers withdrew their prefix announcement for about 10% of the router/prefix pairs. Luckie et al. find that about half of the ASes which had outages were completely unrouted during the outage period and had single points of failure.

Unlike Luckie et al. approach which relied on empirical data to assess the resiliency of Internet, Lad et al. [93] investigate both the impact and resiliency of various ASes to prefix hijacking attacks by simulating different attacks using AS-level topologies obtained through BGP streams. Impact of prefix hijacking is measured as the fraction of ASes which believe the false advertisement by a malicious AS. Similarly, the resiliency of an AS against prefix hijacks is measured as the number of ASes which believe the true prefix origin announcement. Surprisingly it is found that 50% of stub and transit ASes are more resilient than Tier-1 ASes this is mainly attributed to valley-free route preferences.

Fontugne et al. [94] look into structural properties, more specifically AS centrality, of AS-level IPv4 and IPv6 topology graphs. AS-level topologies are constructed using BGP feeds of Routeviews, RIPE RIS, and BGPmon monitors. The authors illustrate the sampling bias of betweenness centrality (BC) measure by sub-sampling the set of available monitors and measuring the variation of BC for each sample. AS hegemony is used as an alternative metric for measuring the centrality of ASes which accounts for monitor biases by eliminating monitors too close or far from the AS in question and averaging the BC score across all valid monitors. Additionally, BC is normalized to account for the size of advertised prefixes. The AS hegemony score is measured for the AS-level graphs starting from 2004 till 2017. The authors find a great decrease in the hegemony score throughout the years supporting Internet flattening reports. Despite these observations, the hegemony score for ASes with the largest scores have remained consistent throughout the years pointing to the importance of large transit ASes in the operation of the Internet. AS hegemony for Akamai and Google is measured, the authors report little to no dependence for these content providers to any specific upstream provider.

*2) **Router-Level Topology**:* Palmer et al. [95] rely on topology graphs gathered by *SCAN* and *Lucent* projects consisting of 285k (430k) nodes (links) to simulate the effects of link and node failures within the Internet connectivity graph. The number of reachable pairs is used as a proxy measure to assess the impact of link or node failures. It is found that the number of reachable nodes does not vary significantly up to the removal of 50k links failures while this value drops to about 10k for node removals.

Kang et al. [96], [97] propose the *Crossfire* denial of service attack that targets links which are critical for Internet connectivity of ASes, cities, regions, or countries. The authors rely on a series of traceroute measurements towards addresses within the target entity and construct topological maps from various VPs towards these targets. The attacker would choose links that are "close" to the target (3-4 router hops) and appear with a high frequency within all paths. The attacker could cut these entities from the Internet by utilizing a botnet to launch coordinated low rate requests towards various destinations in the target entity. Furthermore, the attacker can avoid detection by the target by targeting addresses which are in close proximity of the target entity, e.g. sending probes towards addresses within the same city where an AS resides within. The pervasiveness and applicability of the *Crossfire* attack is investigated by relying on 250 PlanetLab [10] nodes to conduct traceroutes towards 1k web servers located within 15 target countries and cities. Links are ranked according to their occurrence within traceroutes and for all target cities and countries, the authors observe a very skewed power-law distribution. This observation is attributed to cost minimization within Internet routing (shortest path for intra-domain and hot-potato for inter-domain routing). Bottleneck links are measured to be on average about 7.9 (1.84) router (AS) hops away from the target.

Giotsas et al. [98] develop *Kepler* a system that is able to detect peering outages. *Kepler* relies on BGP communities values that have geocoded embeddings. Although BGP community values are not standardized, they have been utilized by ASes for traffic engineering, traffic blackholing, and network troubleshooting. Certain ASes use the lower 16bits of the BGP communities attribute as a unique identifier for each of their border routers. These encodings are typically documented on RIR webpages. The authors compile a dictionary of BGP community values and their corresponding physical location (colo or IXP) by parsing RIR entries. Furthermore, a baseline of stable BGP paths is established by monitoring BGP feeds and removing transient announcements. Lastly, the tenants of colo facilities and available IXPs and their members is compiled from PeeringDB, DataCenterMap, and individual ASes websites. Deviations in stable BGP paths such as explicit withdrawal or change in BGP community values are considered as outage signals.

*3) **Physical-Level Topology**:* Schulman et al. [99] investigate outages within the last mile of Internet connectivity which are caused by severe weather conditions. The authors design a tool called *ThunderPing* which relies on weather alerts from the US National Weather Service to conduct connectivity probes prior, during, and after a severe weather condition towards the residential users of the affected regions. A list of residential IP addresses is compiled by parsing the reverse DNS entry for 3 IP addresses within each /24 prefix. If any of the addresses have a known residential ISP such as Comcast or Verizon within their name the remainder of addresses within that block are analyzed as well. IP addresses are mapped to their corresponding geolocation by relying on Maxmind's IP to GEO dataset. Upon the emergence of a weather alert *ThunderPing* would ping residential IP addresses within the affected region for 6 hours before, during, and after the forecasted event using 10 geographically distributed PlanetLab nodes. A sliding window containing 3 pings is used to determine the state of a host. A host responding with more than half of the pings is considered to be *UP*, not responding to any pings is considered to be *DOWN*, and host responding to less than half of the pings is in a *HOSED* state. The authors find that failure rates are more than double during thunderstorms compared to other weather conditions. Furthermore, the median for the duration of *DOWN* times is almost an order of magnitude larger (10^4 seconds) during thunderstorms compared to clear weather conditions.

Erikson et al. [100] present a framework (*RiskRoute*) for measuring the risks associated with various Internet routes. *RiskRoute* has two main objectives namely, (i) computing backup routes and (ii) to measure new paths for network provisioning. The authors introduce the *bit-risk miles* measure which quantifies the geographic distance that is traveled by traffic in addition to the outage risk along the path both in historical and immediate terms. Furthermore *bit-risk miles* is scaled to account for the impact of an outage by considering the population that is in the proximity of an outage. The likelihood of historical outage for a specific location is estimated using a Gaussian kernel which relies on observed disaster events at all locations. For two PoPs, RiskRoute aims to calculate the path which minimizes the *bit-risk mile* measure. For intra-domain routes, this is simply calculated as the path which minimizes the *bit-risk mile* measure among all possible paths which connect the two PoPs. For inter-domain routing the authors estimate BGP decisions using geographic proximity and rely on shortest path routes. Using the RiskRoute framework, improvements in the robustness of networks is analyzed by finding an edge which would result in the largest increase in *bit-risk* measure among all possible paths. It is found that Sprint and Teliasonera networks observe the greatest improvement in robustness while Level3's robustness remains fairly consistent mostly due to rich connectivity within its network.

Durairajan et al. [90] assess the impact of rising sea levels on the Internet infrastructure within the US. The authors align the data from the sea level rise inundation dataset from the National Oceanic and Atmospheric Administration (NOAA) with long-haul fiber maps from the Internet Atlas project [71] using the *overlap* feature of ArcGIS. The amount of affected fibers as well as the number of PoPs, colos, and IXPs that will be at risk due to the rising sea levels is measured. The authors find that New York, Seattle, and Miami are among the cities with the highest amount of vulnerable infrastructure.

C. AS Relationship Inference

ASes form inter-AS connections motivated by different business relationships. These relationships can be in the form of a transit AS providing connectivity to a smaller network as a customer (c2p) by charging them based on the provided bandwidth or as a settlement-free connection between both peers (p2p) where both peers exchange equal amounts of traffic through their inter-AS link. These inter-AS connections are identical from topologies obtained from control or data plan measurements. The studies within this section overview a series of methodologies developed based on these business relationships in conjunction with the valley free routing principle to distinguish these peering relationships from each other.

*1) **AS-Level**:* Luckie et al. [101] develop an algorithm for inferring the business relationships between ASes by solely relying on BGP data. Relationships are categorized as a customer to provider (c2p) relationship were a customer AS pays a provider AS for its connectivity to the Internet or a peer to peer (p2p) relationship were two ASes provide connectivity to each other and often transmit equal amounts of traffic through their inter-AS link(s). Inference of these relationships are based on BGP data using three assumptions: (i) there is a clique of large transit providers at the top of the Internet hierarchy, (ii) customers enter a transit agreement to be globally reachable, and (iii) we shouldn't have a cycle in customer to provider (c2p) relationships. The authors validate a subset (43k) of their inferences, which is the largest by the time of publication, and finally they provide a new solution for inferring customer cones of ASes. For their analyses, the authors rely on various data sources namely BGP paths from Routeviews and RIPE's RIS, any path containing origin ASes which do not contain valid ASNs (based on RIRs) is excluded from the dataset. For validation Luckie et al. use three data sources: validation data reported by network operators to their website, routing policies reported to RIRs in *export* and *import* fields, and finally they use the communities attribute of BGP announcements based on the work of Giotsas et al. [102]. The authors define two metrics node degree and transit degree which can be measured from the AS relationship graph.

Giotsas et al. [36] modify CAIDA's IPv4 relationship inference algorithm [101] and adapt it to IPv6 networks with the intention of addressing the lack of a fully-connected transit-free clique within IPv6 networks. BGP dumps from Routeviews and RIPE RIS which announce reachability towards IPv6 prefixes are used throughout this study. For validation of inferred relationships three sources are used: BGP communities, RPSL which is a route policy specification language that is available in WHOIS datasets and is mandated for IXPs within EU by RIPE, and local preference (LocPref) which is used to indicate route preference by an AS where ASes assign higher values to customers and lower values to providers to minimize transit cost. Data is sanitized by removing paths with artifacts such as loops or invalid ASNs. The remainder of the algorithm is identical to [101] with modifications to two steps: i) inferring the IPv6 clique and ii) removing c2p inferences made between stub and clique ASes. In addition to considering the transit degree and reachability, peering policy of ASes is also taken into account for identifying cliques. Peering policy is extracted from PeeringDB, a restrictive policy is assumed for ASes who do not report this value. ASes with selective or restrictive policies are selected as seeds to the clique algorithm. For an AS to be part of the clique, it should provide BGP feeds to Routeviews or RIPE RIS and announce routes to at least 90% of IPv6 prefixes available in BGP. The accuracy of inferences is validated using the three validation sources which where described, a consistent accuracy of at least 96% was observed for p2c and p2p relationships for the duration of the study. The fraction of congruent relationships where the relationship type is identical for IPv4 and IPv6 networks is measured. The authors find that this fraction increases from 85% in 2006 to 95% in 2014.

2) ***PoP-Level:*** Giotsas et al. [35] provide a methodology for extending traditional AS relationship models to include two complex relationships namely: hybrid and partial transit relationships. Hybrid relations indicate different peering relations at different locations. Partial transit relations restrict the scope of a customers relation by not exporting all provider paths to the customer. AS path, prefixes, and communities strings are gathered from Routeviews and RIPE RIS datasets. CAIDA's Ark traceroutes in addition to a series of targeted traceroutes launched from various looking glasses are employed to confirm the existence of various AS relationships. Finally, geoinformation for AS-links are gathered from BGP community information, PeeringDB's reverse DNS scan of IXP prefixes, DNS parsing of hostnames by CAIDA's DRoP service, and NetAcuity's IP geolocation dataset is used as a fallback when other methods do not return a result. Each AS relationship is labeled into one of the following export policies: i) full transit (FT) where the provider exports prefixes from its provider, ii) partial transit (PT) where prefixes of peers and customers are only exported, and iii) peering (P) where prefixes of customers are only exported. Each identified relationship defaults to peering unless counter facts are found through traceroute measurements which indicate PT or FT relationships. Out of 90k p2c relationships 4k of them are classified as complex with 1k and 3k being hybrid and partial-transit accordingly. For validation (i) direct feedback from network operators, (ii) parsed BGP community values, and (iii) RPSL objects are used. Overall 19% (7%) of hybrid (partial-transit) relationships were confirmed.

VI. Open Problems

We presented an overview of recent and seminal works regarding the topology of the Internet. Capturing the topology has many inherent difficulties added to that it is under a constant evolution which makes the problem even more challenging. Throughout the year's researchers have proposed solutions to capture a comprehensive and high-resolution representation of the Internet's topology enabling us to have a fuller and better understanding of the Internet's structure. The Internet has come a long way since its inception days where it was mainly used for sharing scientific research, exchanging email, or hosting static web pages. The ever-increasing demand for live content such as streaming videos and rendering video games on clusters of servers instead of a local console has driven content providers to create structural shifts within Internet's topology where they make best efforts to decrease their distance towards their customers. These efforts have lead to a topological change referred to as the *flattening of the Internet* where the edge of the Internet is experiencing rapid growth in connectivity consequently leading to less traffic traversing the classical hierarchical structure of the Internet. Furthermore, the advent of cloud providers and the virtualization of hardware has enabled many small companies to host their services and deliver their content to their users without the overheads of maintaining a global infrastructure. Fueled by increasing demands for higher and more consistent performance, many cloud providers are offering direct peering opportunities intended solely for connecting customer networks to their cloud infrastructure. Conventionally a network would rely on the Internet to route its traffic towards any cloud provider in contrast these cloud interconnections bypass any intermediate network and enable the customer to exchange traffic directly with the cloud provider. This in turn alleviates the uncertainties of routing and enables cloud providers to offer higher QoS guarantees. Given the competitive offerings within the cloud market enterprises can be tempted to rely on different services offered by each cloud provider or in extreme cases completely migrate their service from one provider to another. These demands have lead to colo facility operators and many other 3rd party cloud connectivity partners to offer connectivity fabrics (SDN enabled switches) which are connected to multiple cloud providers. A customer can obtain connectivity with one or multiple of these partnered cloud providers through a single switch port. Furthermore, due to the programmability of these switches, the customers can modify link speed and whom they are interconnected with through an exposed interface. Due to their inherent nature, these interconnections remain invisible from any Internet topology measurement platforms. Additionally, the existence of layer-2 devices between two peering routers invalidates presumptions of recent Internet topology capturing tools [62], [63]. Given the popularity of cloud providers and their increasing demand within today's Internet, presenting new methodologies which can accurately capture these interconnections and quantify their share in offloading traffic from the conventional Internet infrastructure would enable researchers to have a better understanding of Internet's topology.

VII. Conclusion

The Internet is a critical component of our everyday lives from performing trivial tasks such as sending instant messages to friends, watching and streaming videos to banking and operation of power grids and defense operations. The Internet is an interplay of many elements such as physical infrastructure, topology, routing, and applications with each other. Dissecting each elements effect is essential for having a correct understanding of the effects for each of these components. Topology is central to the Internet and its effect is reflected in every aspect of it. Measuring entirety of topology is a difficult task given the scope and scale of the problem at hand coarser and smaller topologies have been employed. Throughout the years many hurdles in topology discovery have been addressed and the community has shifted its attention to inferring inter-AS links between border routers as an abstraction of Internet topology that is feasible to obtain and features sufficient information to address common issues within this domain. Although great strides have been the discovered topologies are limited to the visible portion of the Internet and the interplay of routing on top of physical topologies remains as open and interesting problems that need to be addressed in the future.

References

[1] University of Oregon, "University of oregon route views project," http://www.routeviews.org/routeviews/, 2018.

[2] CAIDA, "Archipelago (Ark) measurement infrastructure," http://www.caida.org/projects/ark/, 2018.

[3] V. Jacobson, "traceroute," ftp://ftp.ee.lbl.gov/traceroute.tar.gz, 1989.

[4] B. Augustin, T. Friedman, and R. Teixeira, "Multipath tracing with paris traceroute," in *End-to-End Monitoring Techniques and Services*. IEEE, 2007.

[5] RIPE NCC, "RIPE Atlas," 2016.

[6] Y. Shavitt and E. Shir, "Dimes: Let the internet measure itself," *SIGCOMM CCR*, 2005.

[7] V. Giotsas, A. Dhamdhere, and K. C. Claffy, "Periscope: Unifying looking glass querying," in *PAM*. Springer, 2016.

[8] M. A. Sánchez, J. S. Otto, Z. S. Bischof, D. R. Choffnes, F. E. Bustamante, B. Krishnamurthy, and W. Willinger, "Dasu: Pushing experiments to the internet's edge." in *NSDI*, 2013.

[9] S. Sundaresan, S. Burnett, N. Feamster, and W. De Donato, "Bismark: A testbed for deploying measurements and applications in broadband access networks." in *USENIX Annual Technical Conference*, 2014.

[10] B. Chun, D. Culler, T. Roscoe, A. Bavier, L. Peterson, M. Wawrzoniak, and M. Bowman, "Planetlab: an overlay testbed for broad-coverage services," *SIGCOMM CCR*, 2003.

[11] M. Berman, J. S. Chase, L. Landweber, A. Nakao, M. Ott, D. Raychaudhuri, R. Ricci, and I. Seskar, "Geni: A federated testbed for innovative network experiments," *Computer Networks*, 2014.

[12] Y. Hyun, "Archipelago measurement infrastructure," in *CAIDA-WIDE Workshop*, 2006.

[13] Í. Cunha, P. Marchetta, M. Calder, Y.-C. Chiu, B. V. Machado, A. Pescapè, V. Giotsas, H. V. Madhyastha, and E. Katz-Bassett, "Sibyl: a practical internet route oracle," *NSDI*, 2016.

[14] M. Luckie, "Scamper: a scalable and extensible packet prober for active measurement of the internet," in *IMC*. ACM, 2010.

[15] R. Beverly, "Yarrp'ing the internet: Randomized high-speed active topology discovery," in *IMC*. ACM, 2016.

[16] R. Beverly, R. Durairajan, D. Plonka, and J. P. Rohrer, "In the ip of the beholder: Strategies for active ipv6 topology discovery," in *IMC*. ACM, 2018.

[17] Z. Durumeric, E. Wustrow, and J. A. Halderman, "Zmap: Fast internet-wide scanning and its security applications." in *USENIX Security Symposium*, 2013.

[18] R. Graham, P. Mcmillan, and D. Tentler, "Mass scanning the internet: Tips, tricks, results," in *Def Con 22*, 2014.

[19] N. Spring, R. Mahajan, and D. Wetherall, "Measuring isp topologies with rocketfuel," *SIGCOMM CCR*, 2002.

[20] R. Govindan and H. Tangmunarunkit, "Heuristics for internet map discovery," in *INFOCOM*. IEEE, 2000.

[21] A. Bender, R. Sherwood, and N. Spring, "Fixing ally's growing pains with velocity modeling," in *IMC*. ACM, 2008.

[22] M. E. Tozal and K. Sarac, "Palmtree: An ip alias resolution algorithm with linear probing complexity," *Computer Communications*, 2011.

[23] K. Keys, Y. Hyun, M. Luckie, and K. Claffy, "Internet-scale ipv4 alias resolution with midar," *TON*, 2013.

[24] N. Spring, M. Dontcheva, M. Rodrig, and D. Wetherall, "How to resolve ip aliases," Citeseer, Tech. Rep., 2004.

[25] M. H. Gunes and K. Sarac, "Analytical ip alias resolution," in *International Conference on Communications*. IEEE, 2006.

[26] R. Sherwood, A. Bender, and N. Spring, "Discarte: a disjunctive internet cartographer," *SIGCOMM CCR*, 2008.

[27] M. H. Gunes and K. Sarac, "Resolving ip aliases in building traceroute-based internet maps," *TON*, 2009.

[28] J. Chabarek and P. Barford, "What's in a name?: decoding router interface names," in *HotPlanet*. ACM, 2013.

[29] B. Huffaker, M. Fomenkov *et al.*, "Drop: Dns-based router positioning," *SIGCOMM CCR*, 2014.

[30] Q. Scheitle, O. Gasser, P. Sattler, and G. Carle, "Hloc: Hints-based geolocation leveraging multiple measurement frameworks," in *Network Traffic Measurement and Analysis Conference*. IEEE, 2017.

[31] V. N. Padmanabhan and L. Subramanian, "An investigation of geographic mapping techniques for internet hosts," in *SIGCOMM CCR*. ACM, 2001.

[32] CAIDA, "Ddec - dns decoder," http://ddec.caida.org/, 2018.

[33] RIPE, "Routing information service (ris)," https://www.ripe.net/analyse/internet-measurements/routing-information-service-ris, 2018.

[34] Packet Clearing House, "Mrt routing updates," https://www.pch.net/resources/Raw_Routing_Data/, 2018.

[35] V. Giotsas, M. Luckie, and B. Huffaker, "Inferring Complex AS Relationships," *IMC*, 2014.

[36] V. Giotsas, M. Luckie, B. Huffaker, and K. Claffy, "Ipv6 as relationships, cliques, and congruence," in *PAM*. Springer, 2015.

[37] "Peeringdb," https://peeringdb.com/.

[38] "Packet clearing house (pch) - data," https://www.pch.net/resources/data.php.

[39] B. Augustin, B. Krishnamurthy, and W. Willinger, "Ixps: mapped?" in *SIGCOMM*. ACM, 2009.

[40] G. Comarela, E. Terzi, and M. Crovella, "Detecting unusually-routed ases: Methods and applications," in *IMC*. ACM, 2016.

[41] I. Castro, J. C. Cardona, S. Gorinsky, and P. Francois, "Remote Peering: More Peering without Internet Flattening," *CoNEXT*, 2014.

[42] G. Nomikos, V. Kotronis, P. Sermpezis, P. Gigis, L. Manassakis, C. Dietzel, S. Konstantaras, X. Dimitropoulos, and V. Giotsas, "O peer, where art thou?: Uncovering remote peering interconnections at ixps," in *IMC*. ACM, 2018.

[43] MaxMind, "GeoIP2 databases," https://www.maxmind.com/en/geoip2-databases, 2018.

[44] IP2Location, "IP address geolocaiton," https://www.ip2location.com/database/ip2location, 2018.

[45] NetAcuity, "Industry-standard geolocation," https://www.digitalelement.com/solutions/, 2018.

[46] M. Gharaibeh, A. Shah, B. Huffaker, H. Zhang, R. Ensafi, and C. Papadopoulos, "A look at router geolocation in public and commercial databases," in *IMC*. ACM, 2017.

[47] J. Engebretson, "Verizon-netflix dispute: Is netflix using direct connections or not?" https://www.telecompetitor.com/verizon-netflix-dispute-netflix-using-direct-connections/, 2014.

[48] L. Li, D. Alderson, W. Willinger, and J. Doyle, "A first-principles approach to understanding the internet's router-level topology," in *SIGCOMM CCR*. ACM, 2004.

[49] E. Gregori, A. Improta, L. Lenzini, and C. Orsini, "The impact of ixps on the as-level topology structure of the internet," *Computer Communications*, 2011.

[50] Y. He, G. Siganos, M. Faloutsos, and S. Krishnamurthy, "Lord of the links: a framework for discovering missing links in the internet topology," *TON*, 2009.

[51] J. Xia and L. Gao, "On the evaluation of as relationship inferences [internet reachability/traffic flow applications]," in *GLOBECOM*. IEEE, 2004.

[52] G. Siganos and M. Faloutsos, "Analyzing BGP policies: Methodology and tool," in *INFOCOM*. IEEE, 2004.

[53] B. Ager, N. Chatzis, A. Feldmann, N. Sarrar, S. Uhlig, and W. Willinger, "Anatomy of a large european ixp," *SIGCOMM CCR*, 2012.

[54] A. Khan, H.-c. Kim, and Y. Choi, "AS-level Topology Collection through Looking Glass Servers," *IMC*, 2013.

[55] R. Klöti, B. Ager, V. Kotronis, G. Nomikos, and X. Dimitropoulos, "A comparative look into public ixp datasets," *SIGCOMM CCR*, 2016.

[56] "European internet exchange association," https://www.euro-ix.net/.

[57] V. Giotsas, S. Zhou, and M. Luckie, "Inferring Multilateral Peering," *CoNEXT*, 2013.

[58] V. Giotsas and S. Zhou, "Improving the discovery of ixp peering links through passive bgp measurements," in *INFOCOM*. IEEE, 2013.

[59] P. Richter, G. Smaragdakis, A. Feldmann, N. Chatzis, J. Boettger, and W. Willinger, "Peering at peerings: On the role of ixp route servers," in *IMC*. ACM, 2014.

[60] G. Nomikos and X. Dimitropoulos, "traixroute: Detecting ixps in traceroute paths," in *PAM*. Springer, 2016.

[61] M. Luckie, A. Dhamdhere, B. Huffaker, D. Clark *et al.*, "Bdrmap: Inference of borders between ip networks," in *IMC*. ACM, 2016.

[62] A. Marder and J. M. Smith, "Map-it: Multipass accurate passive inferences from traceroute," in *IMC*. ACM, 2016.

[63] A. Marder, M. Luckie, A. Dhamdhere, B. Huffaker, J. M. Smith *et al.*, "Pushing the boundaries with bdrmapit: Mapping router ownership at internet scale," in *IMC*. ACM, 2018.

[64] V. Giotsas, G. Smaragdakis, B. Huffaker, and M. Luckie, "Mapping Peering Interconnections to a Facility," *CoNEXT*, 2015.

[65] R. Motamedi, B. Yeganeh, B. Chandrasekaran, R. Rejaie, B. Maggs, and W. Willinger, "On Mapping the Interconnections in Today's Internet," *under review for TON*, 2019.

[66] K.-K. Yap, M. Motiwala, J. Rahe, S. Padgett, M. Holliman, G. Baldus, M. Hines, T. Kim, A. Narayanan, A. Jain *et al.*, "Taking the edge off with espresso: Scale, reliability and programmability for global internet peering," in *SIGCOMM*. ACM, 2017.

[67] 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 *SIGCOMM*. ACM, 2017.

[68] F. Wohlfart, N. Chatzis, C. Dabanoglu, G. Carle, and W. Willinger, "Leveraging interconnections for performance: the serving infrastructure of a large cdn," in *SIGCOM*. ACM, 2018.

[69] A. Y. Nur and M. E. Tozal, "Cross-as (x-as) internet topology mapping," *Computer Networks*, 2018.

[70] S. Knight, H. X. Nguyen, N. Falkner, R. Bowden, and M. Roughan, "The internet topology zoo," *IEEE Journal on Selected Areas in Communications*, 2011.

[71] R. Durairajan, S. Ghosh, X. Tang, P. Barford, and B. Eriksson, "Internet atlas: a geographic database of the internet," in *HotPlanet*. ACM, 2013.

[72] R. Durairajan, J. Sommers, and P. Barford, "Layer 1-informed internet topology measurement," in *IMC*. ACM, 2014.

[73] R. Durairajan, P. Barford, J. Sommers, and W. Willinger, "InterTubes: A Study of the US Long-haul Fiber-optic Infrastructure," *SIGCOMM*, 2015.

[74] A. Gupta, M. Calder, N. Feamster, M. Chetty, E. Calandro, and E. Katz-Bassett, "Peering at the Internet's Frontier: A First Look at ISP Interconnectivity in Africa," *PAM*, 2014.

[75] R. Fanou, P. Francois, and E. Aben, "On the diversity of interdomain routing in africa," in *PAM*. Springer, 2015.

[76] T. Green, A. Lambert, C. Pelsser, and D. Rossi, "Leveraging interdomain stability for bgp dynamics analysis," in *PAM*. Springer, 2018.

[77] N. Chatzis, G. Smaragdakis, J. Böttger, T. Krenc, and A. Feldmann, "On the benefits of using a large ixp as an internet vantage point," in *IMC*. ACM, 2013.

[78] M. A. Sanchez, F. E. Bustamante, B. Krishnamurthy, W. Willinger, G. Smaragdakis, and J. Erman, "Inter-domain traffic estimation for the outsider," in *IMC*. ACM, 2014.

[79] TeamCymru, "IP to ASN mapping," https://www.team-cymru.com/IP-ASN-mapping.html, 2008.

[80] A. Dhamdhere, D. D. Clark, A. Gamero-Garrido, M. Luckie, R. K. Mok, G. Akiwate, K. Gogia, V. Bajpai, A. C. Snoeren, and K. Claffy, "Inferring persistent interdomain congestion," in *SIGCOMM*. ACM, 2018.

[81] M-Lab, "NDT (Network Diagnostic Tool)," https://www.measurementlab.net/tests/ndt/, 2018.

[82] SamKnows, "The internet measurement standard," https://www.samknows.com/, 2018.

[83] B. Chandrasekaran, G. Smaragdakis, A. W. Berger, M. J. Luckie, and K.-C. Ng, "A server-to-server view of the internet," in *CoNEXT*, 2015.

[84] Y.-C. Chiu, B. Schlinker, A. B. Radhakrishnan, E. Katz-Bassett, and R. Govindan, "Are we one hop away from a better internet?" in *IMC*. ACM, 2015.

[85] V. Kotronis, G. Nomikos, L. Manassakis, D. Mavrommatis, and X. Dimitropoulos, "Shortcuts through colocation facilities," in *IMC*. ACM, 2017.

[86] APNIC, "Measuring IPv6," https://labs.apnic.net/measureipv6/, 2018.

[87] R. Fontugne, C. Pelsser, E. Aben, and R. Bush, "Pinpointing delay and forwarding anomalies using large-scale traceroute measurements," in *IMC*. ACM, 2017.

[88] A. Singla, B. Chandrasekaran, P. Godfrey, and B. Maggs, "The internet at the speed of light," in *HotNet*. ACM, 2014.

[89] I. N. Bozkurt, W. Aqeel, D. Bhattacherjee, B. Chandrasekaran, P. B. Godfrey, G. Laughlin, B. M. Maggs, and A. Singla, "Dissecting latency in the internet's fiber infrastructure," *arXiv preprint arXiv:1811.10737*, 2018.

[90] R. Durairajan, C. Barford, and P. Barford, "Lights out: Climate change risk to internet infrastructure," in *Proceedings of the Applied Networking Research Workshop*. ACM, 2018.

[91] M. Luckie and R. Beverly, "The impact of router outages on the as-level internet," in *SIGCOMM*. ACM, 2017.

[92] E. Katz-Bassett, C. Scott, D. R. Choffnes, Í. Cunha, V. Valancius, N. Feamster, H. V. Madhyastha, T. Anderson, and A. Krishnamurthy, "Lifeguard: Practical repair of persistent route failures," in *SIGCOMM*. ACM, 2012.

[93] M. Lad, R. Oliveira, B. Zhang, and L. Zhang, "Understanding resiliency of internet topology against prefix hijack attacks," in *International Conference on Dependable Systems and Networks*. IEEE, 2007.

[94] R. Fontugne, A. Shah, and E. Aben, "The (thin) bridges of as connectivity: Measuring dependency using as hegemony," in *PAM*. Springer, 2018.

[95] C. R. Palmer, G. Siganos, M. Faloutsos, C. Faloutsos, and P. Gibbons, "The connectivity and fault-tolerance of the internet topology," in *Proceedings of the Workshop on Network-Related Data Management (NRDM)*, 2001.

[96] M. S. Kang, S. B. Lee, and V. D. Gligor, "The crossfire attack," *Symposium on Security and Privacy*, 2013.

[97] M. S. Kang and V. D. Gligor, "Routing bottlenecks in the internet: Causes, exploits, and countermeasures," in *Computer and Communications Security*. ACM, 2014.

[98] V. Giotsas, C. Dietzel, G. Smaragdakis, A. Feldmann, A. Berger, and E. Aben, "Detecting peering infrastructure outages in the wild," in *SIGCOMM*. ACM, 2017.

[99] A. Schulman and N. Spring, "Pingin'in the rain," in *IMC*. ACM, 2011.

[100] B. Eriksson, R. Durairajan, and P. Barford, "RiskRoute: A Framework for Mitigating Network Outage Threats," *CoNEXT*, 2013.

[101] M. Luckie, B. Huffaker, A. Dhamdhere, and V. Giotsas, “AS Relationships, Customer Cones, and Validation,” *IMC*, 2013.

[102] V. Giotsas and S. Zhou, “Valley-free violation in internet routing—analysis based on bgp community data,” in *International Conference on Communications*. IEEE, 2012.