IXPs: mapped?

Brice Augustin, Balachander Krishnamurthy, Walter Willinger

IXPs: Mapped?

Brice Augustin[†,*], Balachander Krishnamurthy[‡], Walter Willinger[‡]

† Université Pierre et Marie Curie, Paris ‡ AT&T Labs–Research, Florham Park

ABSTRACT

Internet exchange points (IXPs) are an important ingredient of the Internet AS-level ecosystem—a logical fabric of the Internet made up of about 30,000 ASes and their mutual business relationships whose primary purpose is to control and manage the flow of traffic. Despite the IXPs' critical role in this fabric, little is known about them in terms of their peering matrices (i.e., who peers with whom at which IXP) and corresponding traffic matrices (i.e., how much traffic do the different ASes that peer at an IXP exchange with one another). In this paper, we report on an Internet-wide traceroute study that was specifically designed to shed light on the unknown IXP-specific peering matrices and involves targeted traceroutes from publicly available and geographically dispersed vantage points. Based on our method, we were able to discover and validate the existence of about 44K IXP-specific peering links—nearly *18K more links* than were previously known. In the process, we also classified all known IXPs depending on the type of information required to detect them. Moreover, in view of the currently used inferred AS-level maps of the Internet that are known to miss a significant portion of the actual AS relationships of the peer-to-peer type, our study provides a new method for augmenting these maps with IXP-related peering links in a systematic and informed manner.

Categories and Subject Descriptors: C.2 [Computer Communication Networks]: Network Operations; Network Architecture and Design

General Terms: Measurement.

Keywords: IXP, peering, traceroute.

1. INTRODUCTION

The Internet AS-level ecosystem is a network of networks, where the individual networks or sovereign entities are Autonomous Systems (ASes), and two such ASes are connected if they have established a commercial relationship for fee-based (i.e., customer-provider) or settlement-free (i.e., peer-to-peer) traffic exchange.[1] This logical construct is largely a reflection of the prevailing economic conditions under which the key Internet players (e.g., service providers, content providers, business enterprises, Internet Exchange Points) have to operate. Understanding its structure and the main forces that shape this structure and its temporal evolution have been of great interest to networking researcher for some time (see for example [1, 2] and references therein). However, despite significant efforts, the Internet's AS-level ecosystem has remained an elusive object, mainly because of the various shortcomings of the available measurements that underlie most inference work to date. Among those measurements, the most commonly-used ones are either traceroute-based, or derived from BGP table dumps, or obtained from Internet Routing Registries or other publicly available data bases. While recent studies agree on the adequateness of a combination of these measurements for correctly inferring the number of ASes and adequately discovering and identifying the vast majority of customer-provider links, they all conclude that even the most carefully inferred AS maps currently in use still miss a substantial portion of AS connections of the peer-to-peer type, with estimates varying from 35% up to 95% [3, 4, 5, 6].

The key contribution of this paper is a flexible and efficient approach that sheds new light on the substrate of the AS-level ecosystem that consists of all known IXPs, their member ASes, and all the peerings among those members.[2] This IXP substrate is a critical component of the economic fabric of the Internet; IXPs are the physical infrastructures managed by third parties where member ASes can choose to peer with one another for the purpose of exchanging traffic directly and essentially for free rather than at a cost via some upstream service providers. With a few exceptions [7, 8, 4, 5], this substrate has been largely neglected in past AS-related studies. Our approach consists of launching targeted traceroutes from systematically selected sources to carefully chosen destination and checking the resulting paths for indications that they went through an IXP. It enables us to (i) classify known IXPs in terms of the information and efforts needed to discover them, (ii) check the accuracy of published IXP membership lists, and (iii) obtain for the IXPs we discover new insights into their unknown peering matrices; that

[*]This work was done as part of a summer internship at AT&T Labs-Research.

IMC'09, November 4–6, 2009, Chicago, Illinois, USA.

[1]Other less frequently used business relationships (e.g., sibling-sibling) exist but are not found at IXPs.

[2]The datasets, tools, and detailed results are available at http://www-rp.lip6.fr/~augustin/ixp/.

is, which member ASes peer with one another at these IXP. It is these peering links that have consistently eluded previous inference efforts and are the most difficult to detect using readily available traceroute datasets, BGP routing table dumps, or other publicly available information. At the same time, some recent papers [4, 5] have hypothesized that these very links may hold the key to solving the missing links problem for the AS-level Internet. Our work shows that this hypothesis is indeed true—we discover and validate the existence of about 44K IXP-related peering links or roughly 75% more than any previous study has reported. Moreover, our approach leaves little room for significantly improving our findings concerning the IXP substrate of the AS-level ecosystem. And where there is room for improvements, we provide details of the efforts required to achieve them.

A critical difference between our work and existing large-scale traceroute studies such as CAIDA's Skitter [9] (and its successor, Archipelago, or Ark) or the European Dimes project [10] is our exclusive focus on the IXP substrate of the AS-level Internet. Although an explicit goal of Skitter- or Dimes-like efforts has been the comprehensive mapping of the AS-level Internet as a whole, there has been an increasing awareness within the networking community of the limitations of such a pursuit. The main issue is the quality of the obtained measurements. In addition to well-known problems with traceroute (e.g., IP aliasing), the AS-level ecosystem has a rich set of policies by which individual ASes enforce the prevailing business agreements with their neighbors and hence impact what traffic crosses their networks. Given that most of the existing large-scale traceroute studies pay little attention to these issues, it should come as no surprise that the resulting measurements are more a reflection of what traceroute can measure than what these studies would like the tool to measure. This motivates our approach to stay clear of general-purpose traceroute experiments and target instead the IXP substrate where the economic conditions and routing policies are largely dominated by settlement-free peering agreements. This insight can be put to good use when trying to launch traceroutes between selected source-destination hosts for the main purpose of yielding useful information about the IXP substrate as a whole and the IXP-specific peering matrices in particular.

However, as is the case with most AS-related results, validation remains a serious challenge, and our work on the IXP substrate is not different. We use here some direct and indirect methods to tackle this problem. In the case where the ground truth is available (e.g., existence of IXPs), we use various types of available information (e.g., data bases, routing registries, web sites, search engines, private communications) to check for obvious inconsistencies between our findings and published reports or known facts. Despite our efforts, we find that 55 out of 278 detectable IXPs remain undetected by our method. We explain why they remain invisible and detail the effort and information needed to detect them. In the case where the ground truth is not available (e.g., IXP peering matrices[3], we device a mechanism to express our confidence in the validity of IXP-related peering link discovered by our method. In addition, for a few selected IXPs, we provide some absolute and relative comparisons of the number of validated links we found. We find that our method represents a significant improvement over past efforts, and while there may exist room for improvements, they are only significant (and require substantial more efforts) if typical peering matrices are not sparse.

The rest of the paper is structured as follows. Sec. 2 discusses related work, including two studies on which we build on. Sec. 3 describes the key ingredients of the IXP substrate of the AS-level ecosystem, reviews the use of traceroute, and lists the various sources of data we rely on. Sec. 4 describes the specifics of our methodology for mapping the IXP substrate. The experiments and results are described in Sec. 5, and the validation efforts are summarized in Sec. 6. We conclude in Sec. 8 with some lessons learned from our traceroute study and what they tell us about future AS-specific work.

2. RELATED WORK

The Internet research community's interest in the AS-level ecosystem started largely with the empirical observation reported in [11] that the node degree distribution of inferred AS graphs exhibit a power-law distribution. While this original claim was based on BGP-based data, later efforts such as Skitter [9] or Dimes [10] that relied on data obtained from large-scale traceroute experiments confirmed this finding. AS topology modeling has become a very active research area, largely dominated by novel graph-theoretic approaches aimed at developing mathematical graph models capable of reproducing the observed power-law node degree distributions and possibly other graph-based statistics (e.g., see [12] and references therein). However, at the same time, there has been increasing evidence that the available BGP- and traceroute-based measurements are of insufficient quality to support claims of inferred power-law distributions (or other commonly-used statistics) with any statistical significance. There has been a growing literature detailing the issues and problems associated with relying on BGP- and traceroute-based measurements for inferring AS-level connectivity [13, 14, 15, 16, 17]. These data hygiene efforts have led to an increasing awareness of the inadequacy or even futility of the dominant graph-theoretic treatment of the AS-level ecosystem and its almost exclusive focus on purely topology-related properties. But, they have also highlighted the need for alternative approaches to AS-level topology modeling; such as accounting for the critical forces at work in this economic fabric of the Internet and providing a deeper understanding of how these forces shape the structure and evolution of the AS-level ecosystem.

While Internet practitioners and network operators have advanced this more economic-based perspective for some time[18, 19, 20], the research community has been slow in adopting this view and making it the focal point of further AS-topology modeling. An early study[7] argued for abandoning the pure graph-theoretic treatment in favor of a more economic-oriented approach, and pointed towards the need for a more careful treatment of IXPs and their role in this economic fabric. IXPs are a focal point of the study by Xu et al. [8] and play again a significant role in the recent work by He et al. [21, 4]. In fact, [4] builds on the work by [8], but overcomes some of its limitations and also significantly extends its scope. By proposing to shed new light on the IXP substrate, our objective is similar to that of [8] and [4], and our basic approach is similar to theirs, with noticeable exceptions: our focus is on the IXP substrate and not on the

[3]IXPs rarely publish their peering matrices, and if they do, they are typically not obtained from IXP-internal databases but are inferred from observed traffic data (e.g. VIX `http://www.vix.at/vix_peeringmatrix.html`).

AS-level ecosystem as a whole. Thus, any AS-related results we obtain are by-products of our IXP-centric work. More importantly, because of our exclusive focus on IXPs, we provide a more comprehensive and complete picture of the IXP substrate; in fact, our results provide a detailed account of the information and efforts needed to discover and map each and every IXP and illustrate the cost-benefit trade-offs associated with improving our IXP-related findings.

The IXP substrate of the AS-level Internet is an example where no central agency exists that contains all relevant information. However, since IXPs have in general economic incentives to attract business, IXP-related information is publicly available in various forms. Many IXPs have a website where they provide basic information about their location and facilities; basic architecture, fees, and services; list of AS members and total daily traffic. Two projects that systematically gather this information, augment it with knowledge obtained through personal communications, and make the resulting databases publicly available are PACKET CLEARING HOUSE (PCH) and PEERINGDB (PDB). While these efforts provide a great service to and are of enormous value for the Internet community, because all of the information is provided on a voluntary basis, the quality of these databases in terms of the accuracy and/or freshness of the data is unknown.[4] Nevertheless, some key players within the Internet's AS-level ecosystem require interested parties to first enroll in PEERINGDB before starting any discussion about potential peerings [22]. At the same time, since IXPs treat peering arrangements in general as proprietary information, the actual peering matrices of the IXPs are not part of these databases and have remained in general unknown. IXPs rarely publish their peering matrices, and if they do, they typically report inferred peerings; e.g., links between the IXP and its members that have seen non-zero traffic over some time interval in the recent past.

3. THE IXP SUBSTRATE

The IXP substrate of the Internet's AS-level ecosystem consists of all known IXPs, their member ASes, and all the peerings among those members. We now describe the basic features of this substrate and list the different data sources we rely on in our mapping effort.

3.1 A typical IXP architecture

Most of today's IXPs are composed of a layer-2 device, usually an Ethernet switch [5], where IXP members can plug in their access routers to interconnect directly with one another. IXPs typically deploy several redundancy mechanisms to ensure high resilience of their physical infrastructures, but these layer-2 mechanisms remain by and large invisible to IP.

Fig. 1 shows the typical architecture of an IXP. In this example, the IXP has six members, each represented by a router. When two members decide to peer, they just have to establish a BGP session between their routers. Since this requires that both routers have interfaces in the same IP subnet, the IXP will assign an IP address to the IXP-facing router interface of each of its members from the IP prefix(es) allocated to the IXP by the Internet Registry responsible for the IXP. In Fig. 1, these IXP-facing interface IP addresses for AS1 and AS4 are denoted by @IXP1 and @IXP4, respectively. This general practice is crucial for identifying IXPs in a traceroute path.

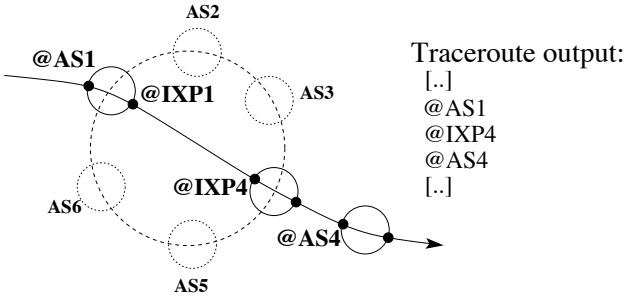

Figure 1: A typical IXP architecture with 6 AS members.

3.2 Identifying IXPs in traceroute

The basic method to identify an IXP in a traceroute path and infer properties such as peerings among IXP members is described in [8] and has been refined and significantly extended in [4]. Our approach relies and builds on these earlier efforts, and we briefly summarize them here for completeness. Key to the method's success is knowing the IXP prefixes. To illustrate, consider again Fig. 1 and assume we are interested in whether or not AS1 and AS4 peer at this IXP. Suppose that launching a traceroute probe from a source within AS1 to a destination in AS4 yields a sequence of IP addresses that contains the following contiguous subsequence: @AS1, @IXP4, and @AS4. If we know the IXP prefix and assume that each router on the path responds to traceroute with the IP address of the incoming interface, we can conclude that @IXP4 belongs to the prefix of this IXP; that is, the trace must have gone through the IXP.

3.3 Identifying members and peerings

By mapping @AS1 and @AS4 to their AS numbers, we obtain conclusive evidence that AS1 and AS4 are not only members of this IXP, but in fact peer with one another at this particular location. Unfortunately, there are a number of potential problems with this basic method. First, knowledge of the IXP prefixes precludes the discovery of IXPs that use private ASNs, and we provide details about the number of such IXPs in Sec. 6.1. Second, and more importantly, not all routers respond to traceroute probes with the incoming interface's IP address but use instead an alternate IP address, or do not respond at all. The alternate address might not belong to the member, but to one of its neighbors. Also the router could respond with a correct address, but the AS mapping might fail at finding the correct AS number. This could result in wrongly inferred peering arrangements, and we want to minimize the impact that such misclassifications can have. To map the second member (AS4), we rely not only on @AS4 but also on @IXP4. Indeed, @IXP4 was assigned by the IXP to AS4 and we have two additional ways to map IXP addresses to their corresponding members.

First, since the DNS names associated with each IXP address often provide information about the member identity, we can try and use the DNS name to derive the AS number

[4]For example, as of 5/5/09, the PDB entry for AMS-IX, a major IXP in Amsterdam, had 257 members, compared to the 312 members given on the AMS-IX website.

[5]A single active IXP (i.e., IXNM) supports ATM and Frame Relay in addition to Ethernet, and based on personal communication with PCH [23], all three planned efforts to build an MPLS-based IXP [24] have failed.

	DB	web	IRR	DNS	BGP	Tr	Ping
Prefixes	a	m	m	m	m	m	
Addr. map	a	n	m	a	a	a	
Existence					a	a	a
Members	a	a	a	a	a	a	
Peerings		m	a		a	a	

Table 1: Summary of our data sources and what we exploited them for. (a: automatically gathered and parsed; m: parsed manually (automated parsing was too hard); n: not gathered; blank: not relevant.

[4]. Second, a BGP table dump contains the IP addresses of the BGP peers, along with their AS numbers. At IXPs, routers use their IXP-assigned address to establish peering sessions. As a result, all BGP tables collected at routers located at IXPs contain these IXP-assigned addresses and their corresponding AS number. The next section gives more detail on how we extracted this data. We describe here how we use BGP tables to improve the inference of IXP peerings.

To map each IXP address to its AS number, we use three techniques, when available, where the ordering reflects the confidence (from high to low) we have in these techniques:

1. BGP dumps: By definition, the (IXP address to ASN) mapping is accurate, and thus we use it whenever it is available. However, since not all the IXP addresses appear in our BGP table dumps, we had to rely on subsequent, more error-prone techniques.

2. Majority-selection process: This heuristic was originally used in [4], but has been shown to work well in earlier studies [15, 25]. The idea of the majority-selection process is that in the majority of cases, routers will respond to traceroute probes with the incoming interface. If, in our traces, we find more than one IP address following @IXP4, we select the one that appears most frequently. We then map this address to its AS number and assign it to @IXP4.

3. DNS names: If the router after the IXP address never responds to our probes (i.e., the traceroute shows a "*" after the IXP address), we cannot use the majority-selection technique. However, we can still try and derive the member AS number based on the DNS name associated with the IXP address, assuming the DNS naming convention supports such an inference. We exclude from our study 1.5% of the traceroutes because they show a "*" after the IXP address and this DNS-based heuristic does not apply

Note that the three techniques just described cannot be applied to map @AS1. As a result, the mapping of this part of the IXP peering is less accurate. Later, in Section 6.3, we describe a method for evaluating the confidence in a discovered peering.

3.4 Data sources

As is typical for distributed and decentralized systems such as the AS-level ecosystem, there exists no central repository for IXP-specific data. Instead, network researchers and operators have access to multiple sources of IXP-related data of varying quality, and in this paper, we make extensive use of the following sources. Table 1 summarizes these sources and our use of them.

IXP databases: Two rich sources for IXP-specific data are Packet Clearing House [23] and PeeringDB [26]. The information they provide includes IXP names, geographic locations, IXP prefix(es) where available, list of members, and links to the IXP websites. Both databases are "best effort" and rely on voluntary contributions. It is important to note that PCH tends to never drop an exchange from the list; instead, as soon as sufficient evidence exists (e.g., through direct contacts), exchanges are marked as "defunct" or "down". PeeringDB encourages IXP operators to maintain their own information in their directory.

IXP websites: Another rich source of IXP-specific information is an IXP's website. Since most IXPs have an economic incentive to attract new members, a typical IXP website provides detailed information about location and facilities; basic architecture, fees, and services; list of AS members and, if available, overall traffic statistics. Using PCH, PeeringDB, and search engines, we collected URLs for over 200 IXPs and wrote one generic and 15 special-purpose parsers to extract the AS membership information available on those webpages.[6] By checking the **Last-Modified** HTTP header value of the corresponding webpages, we eliminated stale information to increase the overall quality of this valuable source of data.

Internet Routing Registries (IRR): Network operators are asked to use the IRR to share their BGP policies [28]. The *import* and *export* attributes may indicate the IP address of the BGP peering routers. When those attributes are provided, we search for addresses that belong to known IXP prefixes, resulting in an IXP-related peering. For example, the following extract reveals a peering between Linx Telecom (AS3327) and Google (AS15169) at the Amsterdam IXP (AMS-IX, allocating addresses in the range 195.69.144.0/23).

```
aut-num:    AS3327
import:     from AS15169 195.69.144.247 ..
```

As the information contained in the IRR is provided on a voluntary basis, its quality is unknown, but its freshness can be inferred by checking the date the entries were posted. Despite these limitations, we extracted the IRR-related peerings and used them to check for inconsistencies in our inferred IXP-specific peerings. We manually queried the IRR entries to discover additional IXP prefixes and collected IXP contacts.

DNS names: We resolved the DNS names for the IP addresses in all the IXP prefixes and inferred their corresponding members, when possible. [4] used this technique but did not evaluate its accuracy and completeness.

Looking glass (LG) servers: Many networks run public looking glass servers capable of issuing commands such as ping, traceroute, or `show ip bgp summary`. Based on [29] we collected a list of 2,329 working traceroute-capable LGs located in 66 countries and 406 ASes (column labeled "LG" in Table 2). [7] Of those 2,329 traceroute capable LGs, 1.1K were also capable of issuing the `show ip bgp summary command`.

[6]While IXPs typically list their members by ASN, 48 of them only publish their names/logos with corresponding hyperlinks. We convert the website DNS name to an IP address and then map the IP address to an AS number using the Team Cymru mapping service [27].

[7]Note that [29] has links to many more LGs, including many that are not maintained and thus not usable. By relying on the database in [29] and checking with the whois database, we were originally able the determine the country location for all but 33 of our LGs. Subsequently, we could also determine the country location of those 33 LGs. For 50% of our LGs, we coud even infer the city location, typically by extracting geographic information from the names of the LG servers themselves.

	CAIDA	PlanetLab	DIMES	LG
Sources	26	254	18K	2.3K
AS	26	223	n.a.	406
Countries	18	31	113	66

Table 2: Coverage of our datasets.

BGP tables: ROUTE VIEWS [30] and RIPE RIS [31] provide snapshots of BGP tables, and many of them are obtained from route collectors located at major IXPs. We also relied on the LG-based BGP data from the 1.1K LGs capable of issuing the `show ip bgp summary` command. This command lists the BGP sessions established with the router running the LG and indicates for each session the ASN and IP address of the peering router. If the routers peer at an IXP, then the IP address will be the one assigned by the IXP operator to the member. The following is an example extract of the command run on a BGP router operated by RUNNet (AS3267):

```
Neighbor                AS
193.232.244.232      15169
```

It shows that Google (AS15169) has a peering session with RUNNet at the Moscow Internet Exchange (MSK-IX, using prefix 193.232.244.0/23). Like the traceroute-based method, mining BGP tables reveals information on the existence of and memberships and peerings at IXPs, without the inaccuracies inherent in traceroute-based data. For instance, in our example, the address 193.232.244.232 will appear consistently in any traceroute to a machine in Google's AS15169 that traverses the MSK-IX. Knowing from the BGP table that this address is assigned to Google, we can directly map it to its correct ASN, and do not need to rely on the mapping of the next address in the path (which, for reasons discussed in Sec. 3.3, will not be necessarily mapped to Google's AS). Using BGP table dumps can be very efficient. While a single query to a BGP LG can yield dozens, or even hundreds of IXP-related peerings, a single traceroute yields at most a single such peering. Note however that although BGP tables provide accurate lists of members and peerings, these lists are in general not complete because we can only detect those members which have a peering relationship with our LGs.

Traceroute datasets: Many projects have generated large sets of traceroute data that have not been mined for IXP-specific information. For example, the SKITTER/SCAMPER measurement project [32] used 26 monitors to run pairwise traceroute probes and made the measurements publicly available. We downloaded and used a snapshot that was captured on April, 27th 2009. DIMES [10] employs about 18,000 agents scattered around the world that perform coordinated traceroute measurements. We downloaded and used the latest resulting dataset available (Feb. 2009). We also obtained access to a traceroute data collected as part of a project at NORTHWESTERN UNIVERSITY [33], but this dataset ended up contributing no new IXP-specific information. Similarly, the public data collected with the *mrinfo* tool [34] in July 2009 revealed no new IXPs and discovered only 200 new peerings at 12 IXPs. Finally, two recent and promising techniques remain to be investigated: traceroutes run from P2P users [17] and the *Reverse traceroute* tool [35].

Ping data: A positive response to pinging IP addresses in a known IXP prefix indicates that the IXP exists and is alive (i.e., responding to ping). On the other hand, a negative response can either mean that the IXP blocks ICMP packets or that the IXP prefix is not advertised in the global BGP tables, and thus not routable. We successively ran pings to each address in each IXP prefix until we got a positive response.

Miscellaneous: We used search engines and personal contacts to resolve inconsistencies in the available data and evaluate the relevance of 3rd-party information.

4. MAPPING THE IXP SUBSTRATE

4.1 List of IXPs and IXP prefixes

We first build our list of known IXPs using the PCH and PEERINGDB databases. While as of April 2009, PCH contained 332 IXPs marked as active, PEERINGDB listed only 253. Through the use of search engines and private communications with IXP operators, we discovered two additional IXPs not listed in either database. Merging these sources resulted in a list of 359 unique IXPs, each with its name and geographic location.

As mentioned earlier, key to our mapping efforts is an accurate and complete list of IXP prefixes. While IXPs with unknown prefixes remain necessarily invisible to our traceroute-based mapping technique, a wrong IXP prefix will lead to wrong inferences. As of April 2009, PCH and PEERINGDB had prefixes for 227 and 165 IXPs, respectively. By combining the two, we obtained 362 prefixes for 247 IXPs. We then augmented this list with prefixes of IXPs we obtained through ad-hoc methods (e.g., by checking for published IP address blocks on IXP websites, we found the family of Russian RIPN IXPs). After checking the validity of these newly obtained prefixes and IXPs (e.g., remove duplicates, check against information in the whois database, rely on naming convention whereby the IXP name generally appears either in the domain name of the host name[8]), we merged the resulting list of prefixes and IXPs with the combined PCH/PEERINGDB list and obtained our final list containing 393 prefixes for 278 IXPs.

4.2 Targeted traceroute

Building on previous efforts [4], the main feature of our approach is to infer peerings between members of an IXP by launching targeted traceroute probes. The main difference between our approach an these earlier efforts is the input data and the algorithms we use.

4.2.1 Input data

One critical component of the input data is our **list of 278 IXPs with a total of 393 known prefixes**, together with their geographic location and list of their members—the most complete list ever used in an IXP study. A second key component is our **list of 2.3K traceroute-enabled looking glass (LG) servers**, together with their geographic location and the ASN of the network they belong to. We built our list of traceroute-capable LGs from the traceroute.org database [29] and updated it with 486 additional LGs from PeeringDB and 20 LGs found through search engines. Our list is the most complete list of traceroute-enabled LGs ever used for providing more visibility into the Internet's IXP

[8]As an example, the AustinMAP has a prefix in PCH, but the DNS names of the corresponding IP addresses all end with "szixp.co.sz", suggesting that this prefix is in fact used by an IXP in Swaziland.

substrate. The third important component is our **inferred AS map** obtained by either merging BGP routing tables from the RIPE RIS and RouteViews projects or by downloading the latest map provided by the Cyclops project[36] and constantly updating it with the new peerings we discovered and validated in our experiments that we ran to fine-tune our methodology (see below). A typical Cyclops-based map contains about 110K AS links. Our augmented map contains 10-20% more links, all of them newly discovered IXP-related peering links, representing the most complete AS map currently available.

4.2.2 Our algorithm

For an IXP with N members, our algorithm has to examine at most $N * (N-1)$ potential peerings. To check a particular peering, say, between IXP members AS A and AS B, the algorithm proceeds in two phases. First, it requires the selection of a source from where to launch the traceroute; depending on whether or not AS A provides a LG server with traceroute capabilities, the traceroute probe will either start within AS A or start outside of AS A and traverse AS A. Second, it requires the selection of a destination target in AS B.

Phase 1 – Source selection:.

Ideally, we would like to launch a traceroute probe from the source network (i.e., AS A) itself; this is only possible if there is a LG server located within the source network. For 213 of the 278 IXPs for which we have prefixes, at least one of their members has a LG. We refer to this technique as the "basic targeted" technique.

If there is no LG in any member AS of an IXP, we rely on our inferred AS map, obtain the neighbors of AS A, and check if any of them has a LG. If neighbor AS C provides a LG server, and AS A and AS C have a customer-provider relationship, then launching a traceroute probe from AS C to a target destination in AS B has a good chance to go through AS A, increasing the likelihood of revealing an actual peer-to-peer link at the IXP. If this 1-hop exploration procedure finds no neighbor ASes of AS A with a LG server, we apply the same procedure to all the neighbor ASes of the neighbor ASes of AS A. If this 2-hop exploration yields no ASes with LG servers (happened in 9% of the cases), we give up and cannot discover the possible peering relationships of AS A. In total, for 253 IXPs, we managed to find at least one LG in at least one member of one of the IXPs' members. We refer to this technique as the "targeted+neighbors" technique (i.e., the "basic targeted" technique updated with information from neighboring LGs).

When the 1- or 2-hop exploration process reveals multiple candidate ASes with LGs, we order them according to the chances they have to reveal the peering. LGs found in the IXP members themselves have a bigger chance, and have priority over LGs located in neighbor ASes of IXP members. We give LGs that are 1 or 2 hops away from an IXP member the same priority and then order them according to the following criteria: (1) Success ratio: for each IXP, we keep track of the number of peerings discovered by each LG. If we have already seen the IXP from a particular LG, then we give that LG priority over those from which this particular IXP has never been seen. (2) Geographic location: we give LGs in the same city as the IXP priority over LGs that are in the same country, mainly because LGs in closer vicinity of the IXP have a higher likelihood of discovering the IXP than LGs further away.

Phase 2 – Target selection:.

When looking for traceroute targets, our goal is to locate destinations that respond to ping, implying that they belong to a routable prefix. This approach prevents artifacts like routing loops and has additional benefits. It typically speeds up any traceroute experiments that involve public LGs and also makes the experiments in general more efficient. Many LGs are configured to buffer the entire traceroute output before responding, and if traceroute has not completed after a timeout, they simply send an empty response. Probes to a non-responding address causes long delays and yields non-informative results. Rather than using trial-and-error scanning to find pingable addresses, we rely on heuristics (e.g., try the first address in each prefix, then the second, etc., until we get a response) that mimic how IP addresses are often allocated inside a prefix. If this incremental search heuristic cannot find any pingable IP address within an AS, we simply select an address randomly in one of its prefixes. For about 4% of all IXP members, we were unable to find a single pingable address. Also, for about 6% of the ASes we did not find any prefix. This can happen when ASes are part of some bigger network that advertises their prefixes. For these ASes, our method fails and we have no way to check the peerings of interest.

4.2.3 Output data and implementation

Our algorithm outputs a set of traceroute probes that can be mined to shed new light on the IXP substrate. We can add a third phase for launching special-purpose traceroute probes using extra information that was unavailable when running the generic version of our algorithm—e.g., a specific source-destination pair for forcing a traceroute probe to go through a particular IXP that remained undetected by the generic algorithm.

Looking glass servers are intended to be queried manually via a browser. To automate, we built a parser that builds the appropriate HTTP queries[9] The parser outputs a list of URLs associated with a LG-specific string with all the information required to query the LG. This data is used as input to our targeted traceroutes algorithm. Our scheduler issues 64 queries in parallel, with the constraints that a LG can only issue a single query at a given time and waits 10s between two queries to the same LG to limit the effects of rate-limiting LGs. Next we wait up to 30s for the HTML response with formatted traceroute output which is then parsed via a generic parser recognizing the dozen different trace formats we encountered.

4.3 Targeted Source Routing

The performance of our basic methodology is highly dependent on the number and location of the traceroute-capable LGs. To reduce this dependence, we developed an extension of targeted traceroutes to exploit the IPv4 *Loose Source Record Route* (LSRR) option and increase the coverage of our method without increasing the number of LGs. The LSRR option has previously been employed in Internet mapping projects [37, 38], and we use it here to force our targeted

[9]Of the 2.3K LGs, only about 40 required manual formatting (e.g., when part of a Javascript/AJAX script, when a cookie is required, or when a session ID is carried in a PHP script).

traceroutes to traverse a particular pair of IXP members, thus allowing us to check a particular IXP-related peering.

LSRR targeted traceroutes differ from our basic targeted traceroute method only in its first phase. To check a peering between ASes A and B, instead of selecting a LG in or near AS A, we pick a LSRR-capable router in A and force our traceroute probes to go through this router before reaching its final destination (in B, selected as in Phase 2). Ideally, if we can find such a router in each IXP member, we could systematically check each IXP peering from a single source in the Internet. However, this extension has limitations because packets with LSRR options can cause many problems.

We first use the basic technique described in [37] to check if a router is LSRR-capable. We send a UDP probe with a high TTL value to an IP address d which we know is responsive to UDP packets and insert a LSRR option to force the probe to traverse r before reaching the destination. Receiving a response from d means that r forwarded the packet and thus is LSRR-capable. We had to extend this technique because routers can block packets with LSRR options, either silently or by sending an ICMP *Source routing failed* error.[10] Specifically, for each member's ASN, we first build a list of candidate IP addresses belonging to this ASN, extracted from the CAIDA, PlanetLab, and DIMES traceroute datasets. We then test each candidate to see whether it is LSRR-capable or not, using the test described above. For each member's ASN, we keep trying candidate addresses until the list is empty or we find a successful candidate. The result of this phase is a list of LSRR-capable routers associated with the ASNs of the IXP members they belong to. We found LSRR-capable routers in 847 IXP members.

Another known problem that packets with LSRR options often encounter is that they are blocked. We initially ran the experiment from 250 PlanetLab nodes, and restricted the subsequent experiments to the 30 nodes which revealed at least one LSRR-capable router (i.e., those nodes allow the injection of packets with LSRR options in the network). Performing the experiments from just those 30 nodes did not prevent our probes from being widely filtered by routers spread across many different ASes. Note however that if the probes are dropped after they go through an IXP, we still have an opportunity to check the peerings.

Lastly, routers often do not respond to traceroute probes with LSRR options. In a regular traceroute path, it can happen that routers do not response to probes, thus they remain anonymous [39]. While this phenomenon is relatively rare in the case of regular traceroute, it happens much more frequently for traceroute probes with LSRR options. This empirical finding limits the usefulness of source routing and is one of the main reasons why it is used so infrequently.

5. RESULTS

5.1 Experiments

Using as input (i) our list of 278 IXPs with 393 known prefixes, (ii) a list of 2,329 traceroute-capable looking glass servers located in 66 countries and 406 ASes, and (iii) an AS map that we obtained either from merging BGP routing table information from RIPE RIS and RouteViews ourselves or from downloading an inferred topology from the Cyclope website and augmenting it with our most recently discovered IXP peerings, we run our targeted traceroute experiments (without targeted source routing) in July 2008, December 2008, and April 2009. We report results from our April 2009 experiment as it uses the most up-to-date list of LGs and inferred AS map and subsumes the two earlier experiments.

[10]Even if we receive such an error message, we still consider the router as potentially LSRR-capable.

	CAIDA	P-lab	DIMES	LG	All
Direct	218	117	n.a.	83	65
Neighbors	141	50	n.a.	35	32

Table 3: IXPs with no coverage.

Region	# LGs	Region	# LGs
Europe	1,361	South America	84
North America	718	Australia & New Z.	58
Asia	104	Africa	4

Table 4: Geographic distribution of the 2.3K LGs.

We compare the results from mining our dataset with those obtained from mining CAIDA's Skitter dataset and the Dimes dataset of traceroute measurements for IXP-specific information. While [8, 4] also used Skitter-based data in their searches for IXPs, to best of our knowledge, the Dimes dataset has not been analyzed for IXP-specific information. To have yet another point for comparison, we also run an experiment using PlanetLab [40], where we selected 254 alive nodes, one in each site. We then compiled a list of IXP member ASes; i.e., ASes that are known (or believed) to be a member of any of the IXPs with known prefixes. From each alive node and for each AS on our list, we launched a traceroute to a single IP address responding to ping. We ran this experiment twice, on October 24 and December 9, 2008 and report here the results from the December run. Finally, we ran our source-routed traceroute experiments three times in March 2009 (we report the results from our March 30 experiment) and mined the BGP tables in February 2009.

To help calibrate the results obtained from these various traceroute-based studies, Table 2 shows for each study the number of different ASes, countries, and regions where the traceroute sources are situated. In addition Table 3 lists for each measurement study the number of IXPs for which there is no traceroute source available in any of its (direct) members or in any of its members' neighbors. Both tables show that detecting IXPs is largely a visibility problem; that is, detecting IXPs is less about the total number of available traceroute sources and more about where these sources are located with respect to the IXPs.

Fig. 2 shows the distribution of our 2.3K traceroute servers among the 406 ASes. Some ASes provide a large number of LGs, and their locations are in general spread across each network. Most ASes provide only a single vantage point. To illustrate the geographic distribution of the LGs, Table 4 breaks them down by continent. The vast majority of LGs are found in Europe and North America, and only four LGs are located in Africa (Egypt and South Africa).

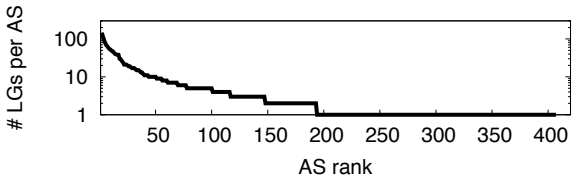

Figure 2: Distribution of the number of LG per AS.

5.2 Existence of IXPs

To recall, there are 278 IXPs with known prefixes and another 81 for which we have no prefix information. In theory, all 278 IXPs with known prefixes should be detectable via traceroute, but some of them will be harder to discover than others, depending on how they are designed (e.g., using private addresses) and, more importantly, where they are located within the IXP substrate with respect to the sources available for launching traceroute probes. The column "IXP" in Table 5 summarizes the results of mining the different datasets for IXPs, and the same column in Table 6 breaks down the results of mining our dataset[11] by the particular technique we used. The three sub-columns give the number of detected IXPs, the percentage, and the number of unique IXPs found (i.e., IXPs that none of the other datasets or techniques detected).

Combining the results from all datasets, we established the existence of 223 of 278 IXPs with known prefixes. The bulk of the discovered IXPs (i.e., 214) is found in our dataset, and of the techniques we used, the most successful one was the targeted traceroutes method with 176 "hits". In contrast, the May 2005 experiment reported in He et al. [4] found 110 IXPs.

Note that with less than 2.4K sources, the targeted traceroutes method does better than DIMES which has over 18K sources at its disposal. This observation illustrates that the success rate for discovering IXPs depends critically on finding LG sites with good visibility into the IXP substrate. The number of sources is less important than their relative location with respect to the IXPs and their members. Also note that our technique of traceroutes with source-routing detected a total of 118 IXPs which shows its potential. The reason why the pinging technique revealed only 74 IXPs is that most IXP prefixes are not advertised in the global BGP tables. All in all, relying on our various techniques to mine our dataset produced all but five of the IXPs present in the CAIDA, DIMES or PlanetLab datasets; all our attempts to detect them (using brute-force LSRR experiments) failed.

With 223 of the 278 IXPs with known prefixes discovered, we classify the 55 unaccounted IXPs in Sec.6, giving detailed reasons as to why they remained undetected and providing specifics about the information needed to detect them.

5.3 Membership of IXPs

The columns "Members" in Tables 5 and 6 give the number of discovered IXP members for each dataset and technique, respectively. In summary, combining the results obtained from mining all the datasets, we obtain a total of 3.5K IXP members. The vast majority of them (i.e., 3.3K) resulted from mining our dataset, and among the techniques we used to detect IXP members, the two most successful ones were targeted traceroutes and BGP LGs. Using BGP LGs to detect members is more efficient, though; it also has the additional benefit of producing very accurate results. The fact that LSRR probes frequently get blocked along their routes explains why the source-routing technique systematically yields a lower number of IXP members. When comparing our results to the 2.3K members detected by He et al. [4] in their May 2005 experiment, it is important to note that their members include IXP members found in their traceroutes as well as IXP members collected from 66 IXP websites and inferred from their IXP addresses' DNS names.

[11] When referring in the rest of the paper to "our dataset", we mean the dataset consisting of (i) targeted traceroutes, (ii) targeted LSRR traceroutes, (iii) BGP LGs, and (iv) pings.

5.4 Peerings at IXPs

The columns "Peerings" in Tables 5 and 6 show the total number of IXP-related peering links as well as the number of unique peering links (i.e., links only seen in one dataset or discovered by one technique) discovered using the various datasets and techniques. In summary, when combining all datasets and techniques, we discover a total of 58K peerings at IXPs. Our dataset is responsible for the bulk of it (i.e., 44K), and the targeted traceroute method yields roughly twice as many IXP peerings than either the BGP LG-based method or the targeted source-routing method. In their May 2005 study, He et al. [4] reported in their paper a total of 7.7K IXP-related peerings. However, while their goal was to discover additional links in the overall AS topology (i.e. links that were not present in current AS maps), our goal is more specific and aims at detecting all peerings at IXPs.

Focusing first on the publicly available datasets, we observe that despite its large size, the CAIDA-provided traceroute data is a relatively poor source, producing only 2.6K peerings. As already noticed, the methodology to collect these traces suffers from a poor coverage in terms of vantage points. DIMES clearly has a better coverage than CAIDA and yields 17K peerings. Note however that the DIMES study has to be viewed as a general-purposes traceroute collection effort which is not optimized in any way for the purpose of discovering peerings at IXPs. While such datasets are a good starting point for detecting peerings at IXPs, the large number of DIMES agents suggests the design of an IXP-specific experiment that has the potential to discover IXPs and peerings that remained invisible to our approach. For example, if there are DIMES clients in areas where we have no or poor coverage in terms of LG sites, traceroutes to and from those clients are likely to reveal some of the IXPs, their members, and peerings among those members that we were unable to detect.

Next, relying on our dataset and techniques yields 44K peerings, and outperforms the use of any of the publicly available data sources. Our regular traceroutes from the 254 PlanetLab nodes produced only 8K peerings due to the location of those nodes (generally universities, typically not connected to IXPs) in spite of their geographical dispersion. We discovered slightly more peerings (10.4K) in the traces collected with LSRR traceroutes from our 30 nodes, even though they were run from PlanetLab nodes. Since source-routing forces probes to take a certain path before reaching an IXP, the sources of the traceroutes have little impact on the results. The results involving LSRR traceroutes may be in part due to the problems of source-routing discussed earlier. Out of 183K paths traversing IXPs, we dropped 57K because they were incomplete in the sense that either the router before or after the IXP address was not responding. Compared to our regular traceroute traces where we had to dismiss around 1.5% of the paths, for our LSRR traceroute traces, this percentage was 31% of the interesting paths.

Fig.3 shows how using our dataset and mining it with our techniques clearly outperforms the other methods that rely on datasets from CAIDA, DIMES and PlanetLab. Our method performs worse (on the right-hand side of the figure) on IXPs with less than 20 peerings. Table 6 shows that our targeted traceroute technique performs well, generating

	IXP (total of 278)			Members		Peerings				Cost	
Dataset	#	%	unique	#	%	#	%	validated	unique	Time	Queries
Our dataset	214	77%	50	3.3K	94%	44K	76%	36K	29.6K	14d	16M
DIMES	155	56%	3	1.9K	53%	17.5K	30%	10.5K	5.5K	n.a.	n.a.
PlanetLab	122	43%	0	1.8K	51%	8.3K	14%	5.6K	1.6K	1h	1.1M
CAIDA	102	37%	0	1K	28%	2.6K	4%	1.6K	0.3K	3d	2.9M
Personal	3	1.3%	3	7	0.2%	6	0	6	6	n.a.	10
Total	223	80%		3.5K		58K		44K			
He et al. [4]	110	n.a.	n.a.	2.4K	n.a.	7.7K	n.a.	n.a.	n.a.	n.a.	23K

Table 5: Contributions of the datasets we used.

	IXP (total of 214)			Members		Peerings				Cost	
Technique	#	%	unique	#	%	#	%	validated	unique	Time	Queries
Targeted (basic)	170	79%	28	2.1K	63%	25.3K	57%	19.8K	11.2K	2d	150K
Targeted+neighbors	176	82%	34	2.3K	70%	28.8K	65%	23K	13K	14d	1M
BGP sum	119	55%	13	2.5K	76%	10.7K	25%	10.7K	6.8K	1h	3.5K
LSRR	118	55%	3	1K	30%	10.4K	24%	7.5K	5.2K	10d	15M
Ping	74	34%	11	0	0	0	0	0	0	1h	3K

Table 6: Contributions of our techniques.

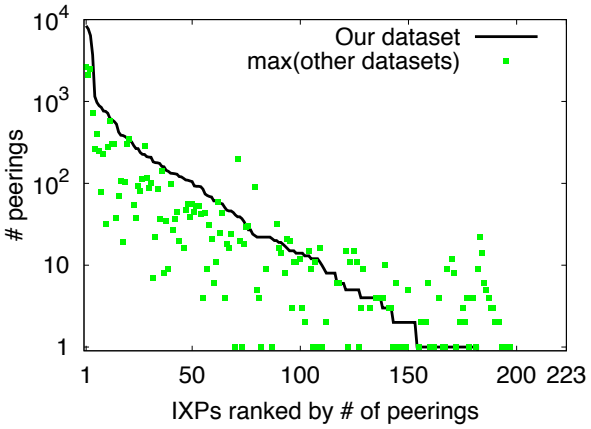

Figure 3: Comparison of the number of peerings found with our technique and the best of the other datasets (CAIDA, PlanetLab and DIMES).

at least twice as many peerings than the other techniques, due to the great source diversity and the ability to launch specificly targeted traceroutes from them. In contrast, mining BGP LGs produced relatively poor results with 11K peerings—IXPs are hard to find in a BGP table and require the BGP LG to be installed at a router located at the IXP. In such a case, the router will peer with other IXP members, and their IXP-allocated address will appear in the BGP dump. As a result we will be able to detect that the peering is actually done at the IXP. Unfortunately, most LGs are not located at an IXP. Being located outside an IXP, they are of no use for us as source of informative BGP data. On the one hand, the targeted traceroute method is clearly superior because it depends to a lesser degree on the LG location. On the other hand, it is slower, creates more load on the network, and is not as accurate as the use of BGP LGs.

5.5 Cost

The "Cost" columns in Tables 5 and 6 also give the cost associated with the various datasets used and techniques applied. Here, we measured cost in terms of duration of the experiments and number of queries issued during the experiments. We note that as far as our techniques are concerned, LSRR traceroute is clearly the most expensive technique, while the use of BGP LGs yields a high number of peerings for a very low overall cost. Although our targeted traceroute technique is expensive in terms of time, it seems possible to reduce this cost significantly by more effectively incorporating the results learned from previous runs of the experiment. At the same time, it is important to keep in mind the high failure rate of this technique. Indeed, around 27% of the issued queries did not complete, either because of a timeout (4% of the queries) or because of temporarily down and rate-limiting LGs.

Database	IXP	Members	Peerings
DNS	84	2K	0
PeeringDB	253	1K	0
PCH	332	2.2K	0
IXP websites	166	4K	3.1K
IRR	80	2K	18K

Table 7: Contributions of IXP databases.

5.6 Comparison with public IXP databases

Table 7 is a summary of what we could extract from the various IXP-related databases. Recall that the quality of these databases is largely unknown, implying that the information extracted from them may be far from the ground truth and can at best be used to support some rough qualitative statements. We present this information here for completeness. Specifically, we could extract membership information for 84 IXPs based to their DNS naming conventions, resulting in nearly 2K members. Unfortunately, the remaining IXPs either do not name their IP addresses or do it in a way which is not easily parsable. We have seen earlier that the data quality of the two main IXP repositories differ, with PCH showing twice as many members as PeeringDB. In contrast, IRR yields 2K members at only 80 IXPs. Still, the richest source for membership information are the IXP websites themselves which provide a total of 4K members at 166 IXPs. One of the possible reasons for the difference between this number and the 3.5K members given in Table 5 is

that networks serving content tend to peer at several IXPs. Since we often have only a single LG in a given member, the traceroutes run from such a LG are likely to go through the closest IXP, thus preventing the detection of members at other IXPs.

6. VALIDATION

6.1 Undetected IXPs

We detected 223 of the 278 IXPs with known prefixes. We went through the 55 IXPs that remained invisible to the various traceroute experiments and explain below why they were not discovered. We collected IXP contact information from IXP databases and websites, IRR databases, and contacted operators, network administrators, and teams at PCH and EP.net. We summarize our main findings and refer to the accompanying webpage (http://www-rp.lip6.fr/~augustin/ixp/) for more details.

We found that 17 of the 55 undetected IXPs are active. Of those 17 IXPs, 11 were confirmed to be active, but our attempts to detect them failed. Regional IXPs typically allow traffic between members but forbid any transit traffic, which forces members to be multi-homed. If we do not have a LG in one of the members, we have no chance to find a traceroute going through the IXP. This is the case for many of the IXPs in Africa where we have only four LGs. We selected two African IXPs (in Swaziland and Uganda) where we don't have any LG in either their members or the neighbors of their members and designed a brute-force experiment using LSRR targeted traceroutes. Checking more than *100K* IP addresses, we found only *four* LSRR routers in two members of the Uganda IX and none in the members of the other IXP. Yet, none of our source-routed traceroutes revealed the two IXPs: they were either blocked by an intermediate router or experienced non-responses. We thus lack sufficient information to infer the presence of an IXP that is known to exist and be active. IXPs that fall in this category are typically small and isolated.

For the remaining 6 undetected but active IXPs, we were unable to find target addresses and run traceroutes to their networks. While IXPs typically tend to disclose their member list, we found four IXPs that didn't disclose them. Interestingly, DNS names or other techniques can often be used to reverse-engineer the peering participants, and for two of the above IXPs, we succeeded in identifying their members. We also encountered two IXPs that assigned private ASNs to their members. This policy prevents us from finding addresses to launch our traceroutes.

Lastly, based on 3rd-party information, we were able to classify the remaining 38 undetected IXPs into defunct (22), planned (7), not an IXP (3), temporarily down (1), and unknown (5). Here, unknown means that the evidence we have is either too weak or contradictory.

6.2 Membership and Mapping methods

Sec. 3.3 describes three techniques for mapping IXP addresses to the ASN of their corresponding members. The DNS mapping and the majority selection are error-prone as they rely on IP to AS mapping which is inaccurate. The mapping extracted from BGP tables at IXPs is accurate and has additional benefits, and hence we make extensive use of this technique. One benefit of relying on the mapping extracted from BGP tables at IXPs is that it helps discovering new peerings that we would have ignored otherwise. For example, we typically ignore traceroutes that contains a non-responsive router ("*") before or after the IXP address. However, if we can map the IXP address directly to its corresponding member, we do not need the presence of an IP address after the IXP address. This technique is particularly efficient when the traceroute contains many "*"s as is typically the case for traceroutes with source-routing (see Sec. 4.3). Using this property, we discovered 20% additional peerings in our LSRR traces. The difference is less significant in regular traceroutes where "*"s occur less frequently.

A second benefit of BGP mapping is for assessing the accuracy of the majority selection and DNS-based techniques. Our targeted traceroutes traversed a total of 4,114 IXP addresses, and 65% of them were present in BGP tables and could thus be checked. 94% of those verifiable addresses were confirmed. The DNS mapping is less accurate: 38% of the 7,019 addresses could be checked, but only 77% of them were correct. Note, however, that there can be cases where the majority selection gives a result that is different than expected, but not necessarily wrong. For example, we checked the members of the Amsterdam Exchange Point (AMS-IX) for which the majority selection and the BGP method disagreed. Several ISPs own different ASNs, and so the different mapping techniques yielded different ASNs. E.g., for EUnet, the BGP method gave AS6667, while the majority selection technique produced AS790. EUnet uses the former ASN in its international backbone, and the latter for domestic operations.

Regarding completeness, Sec. 5.6 showed a gap between the member lists published by IXPs on their websites and the member lists inferred from our dataset. More specifically, we have a coverage greater than 90% for only 17% of the 112 IXPs that we checked. Among them, we find some of the bigger IXPs like AMS-IX, DE-CIX and LINX. Half of the checked IXPs have a coverage of lower than 60%.

6.3 Peerings

As mentioned earlier, 166 IXPs publish a list of their members. While the lists for big IXPs are clearly dynamically updated (e.g., they are read directly from the IXP route server), others are manually maintained and tend to be out-of-date (e.g., checking last modification date, we found pages that are up to seven years old).

Given that the peering matrices published by just a few IXPs are of unknown quality, we lack ground truth. Our measurement method outputs a list of peerings extracted from various datasets, but these peerings may be incorrect (see Sec. 3.3). We next describe a mechanism for assigning a level of confidence to our detected IXP peerings by combining several sources of information to "rate" the validity of each peering. Our method is based on the following properties of a given peering `AS1 IXP AS4` (see Fig. 1):

rev: We also observed the reverse peering `AS4 IXP AS1`. Note however that this only applies if there is a LG in AS4, and routing between AS1 and AS4 through the IXP is symmetric. Note that in many cases, we do not have an appropriate LG and so these two conditions may not be satisfied.

p1bgp: AS1 is a member of the IXP, as per our BGP LGs. Therefore, even if AS1 was obtained by mapping an IP address, we know that this mapping was correct.

p2bgp: AS4 was obtained directly by mapping the IXP address to its member's ASN, according to our BGP LGs.

Thus we did not rely on the majority selection at all, and can consider this mapping to be correct.

p1right: We found AS1 in the "right" part of another peering (e.g. `AS3 IXP AS1`). The fact that we observe a member on both sides of a peering reduces the chance that the IP-to-AS mapping is incorrect.

p2left: We found AS4 in the "left" part of another peering (e.g. `AS4 IXP AS3`); same reasoning as **p1right**.

p1maddr: We found multiple IP addresses on the "left" part of the peering, all of which were mapped to AS1. If we find a single address that maps to AS1, it is possible that the mapping is incorrect because the paths go through a single router which systematically responds with an incorrect interface. On the contrary, if the paths go through multiple addresses, it is less likely that all these routers respond with an incorrect address.

mjuasn: We applied the majority selection rule over at least two addresses, all of which were mapped to AS4. If we see multiple addresses, and all addresses map to the same AS, the mapping is likely correct.

After determining these properties for each peering, we define the following combinations of properties to classify the peerings; i.e., to rate our level of confidence in the correctness of the discovered peerings:

High confidence: Assigned to peerings with the properties `rev or (p1bgp and p2bgp)`; i.e., peerings which we have observed in both directions, or for which both ASNs are known to belong to the IXP.

Medium confidence: Assigned to peerings with the properties `(p1bgp or p1right or p1maddr) and (p2bgp or p2left or mjuasn)`; i.e., peerings for which only one member is known to belong to the IXP and which rely only on a (seemingly correct) majority selection process.

Low confidence: Assigned to the remaining peerings. This designation does not mean that the peering is incorrect; we just do not have enough evidence to assert its correctness.

The "validated" columns in Tables 5 and 6 show the number of discovered peerings for which we have a "high confidence" and which we view as being validated. Of the 28.8K peerings found by targeted traceroutes, we were able to assigned a "high confidence" to 23K of them (75%). Of those 23K links, 2.7K were classified this way because they satisfied **rev**, i.e., we observed the peering in both directions. The remaining ones were assigned "high confidence" because both members are known to belong to the IXP. While 3K peerings were assign a "medium confidence", only 2.8K peering ended up being classifies to have "low confidence". To illustrate the "best effort" nature of the IRR, we found that about 15% of the "high confidence" peerings are present in the IRR; for peerings with a "medium" and "low" confidence, the numbers are 7% and 2%, respectively.

6.4 Peering matrices

We examined 111 IXPs, and only four of them publish their peering matrices. Even for those four, we do not know if it represents the ground truth. For example, the matrix published on the LONAP website is obtained by extracting peering information about each member from the *whois* database, which is known to have stale entries and does not reflect the peerings actually made at the IXP. In the case of VIX, the published peering matrix is inferred from measured traffic that traversed the links over some time interval in the past. Thus, any comparisons of published peering matrices with those obtained by our method have to be interpreted with care. Table 8 shows the number of peerings found at 8 selected IXPs. To build this table, we only considered peerings discovered in the various datasets that fell in the "high confidence" category and ignored the "medium" and "low" confidence peerings. The last column gives the total number of ("high confidence") peerings detected in all of the datasets. In the absence of the ground truth, to provide some calibration of the number of discovered IXP peerings, we computed the maximum number of peerings for each of the 8 IXPs (i.e., $n*(n-1)/2$ where n is the number of members/participants at the IXP) and consider the two cases where 60% and 30% of the entries of the peering matrices are populated. While the 30% case reflects a relatively sparse peering matrix, the 60% case represents a possibly unrealistically high degree of peering at IXPs. We note that our method clearly outperforms the other methods, even though there are peerings discovered in the other datasets that we miss, mainly because of the constraints imposed by the locations and number of our LGs.

Examining Table 7, we note that even when assuming, for example, that real-world peering matrices tend to be sparse and have only about 30% of their entries populated, the number of peerings we find with our method at these selected IXPs is still off this 30% target. Assuming less sparser peering matrices (e.g., 60% of all possible peerings have been established), this difference becomes even larger. This suggests that despite the dominance of our method over the other methods, there is room for improvements. Other indications that improvements may be possible are are seen in Table 4 where we observe that the different datasets yield a substantial number of unique peerings. The typically low level of overlap between the different methods suggests that more peerings exists and that the total number of IXP peerings among all the datasets exceeds the 58K that we found to date or the 44K that we have validated. This observation is further supported by Table 7 that shows the number of IXP-related peerings discovered by mining the IXP databases discussed earlier. While PeeringDB and PCH do not provide any peering information, the IRR seems to be a rich source of information since it yields some 18K peerings at only 80 IXPs. However, in the absence of any ground truth for peerings at IXPs, all that Tables 4, 5, and 7 say is that while our proposed method clearly outperforms the currently available methods as far as detecting peerings at IXPs is concerned, there may still be room for improvements. How much room remains an open problem, though, mainly because the reasons why peering matrices may not be full can vary. We discuss some of these reasons in the next section.

6.5 Weakness of our methodology

The main weakness of our methodology is that we entirely depend on the available LGs (and LSRR-capable routers). As a result, we will never be able to check the peering between two members if we do not have a LG in one of them or in one of their 1- or 2-hop neighbors. Given that we rely on publicly available LGs, some of the actual peerings will remain uncovered by our method. Furthermore, having a LG in a member might not be enough, as its geographic location also matters. Consider for example Limelight (AS22822). It offers a LG at multiple routers in different locations which enables us to check its peerings at many IXPs. However, many other networks only provide a LG at a single router.

IXP	max	60%	30%	CAIDA	PlanetLab	DIMES	Our dataset	All datasets
VIX	5.6K	3.4K	1.7K	63	182	186	945	1081
SIX	1K	0.6K	0.3K	1	32	22	88	102
MANAP	0.7K	0.4K	0.2K	4	0	6	26	39
AMS-IX	48K	29K	14.5K	352	1.3K	2.6K	7.1K	8.6K
DE-CIX	36K	22K	11K	307	1.2K	2.3K	7.3K	8.9K
LINX	45K	27K	13.5K	355	1.2K	2.5K	5.2K	7.3K
LAIIX	1K	0.6K	0.3K	11	26	29	73	99
FreeIX	5.4K	3.2K	1.6K	7	10	28	241	309

Table 8: Peerings found at selected IXPs.

This router's particular location is likely to preclude the discovery of the network's other peerings at other locations. One way trying to circumvent this problem is to make more and better use of source-routing, but this comes at a cost as shown in Table 5.

Second, for several reasons, two members at an IXP might simply not peer at that IXP. For example, they might have a direct peering. For instance, BT (AS5400) and Google (AS15169) are both members of the AMS-IX, but we did not detect a peering at this IXP in spite of having 3 LGs in BT. Instead, we observed a direct path from BT to Google. Another possibility is that the members have established a *private* peering at the IXP. In this case, they directly plugged a cable between their routers, and the path does not go through the public fabric, which prevents the detection via our method. Finally, there are the possibilities of configuration errors, or members are simply not interested in peering with one another at the IXP because they exchange only a very small amount of traffic.

7. ANALYSIS

Using the information we obtained for the 223 IXPs we detected, we present below the results of our analysis of IXP-specific data. The data consists of discovered peerings (with associated confidence attributes) and IXP membership information, and the metrics of interest are IXP size, member presence, member connectivity, and member multi-connectivity.

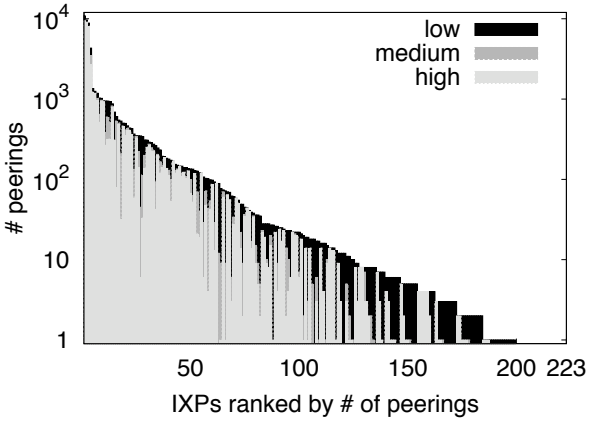

Figure 4: Peerings per IXP.

After associating each of the 58K discovered peerings (irrespective of the assigned confidence) with the corresponding IXP, we rank the IXPs in decreasing order of the total number of associated discovered peerings and show in Fig. 4 the ranked IXPs (x-axis) and their sizes (y-axis), broken down by the portions of peerings to which we assigned "high", "medium", and "low" confidence, resp. The plot shows that the "high confidence" or "validated" peerings dominate the left-hand-side of the figure where the number of discovered peerings at an IXP is high. At the same time, there are many IXPs on the right-hand-side of the figure that have a majority of "low confidence" peerings, consistent with our earlier observation that the success or failure of discovering the peering matrices of small IXPs is highly dependent on the location of our LGs with respect to those IXPs and their members. Also note that the absence of the ground truth as far as the total number of peerings at an IXP is concerned prevents us from indicating in Fig. 4 how far off we are compared to the total number of actual peerings that exist at each IXP.

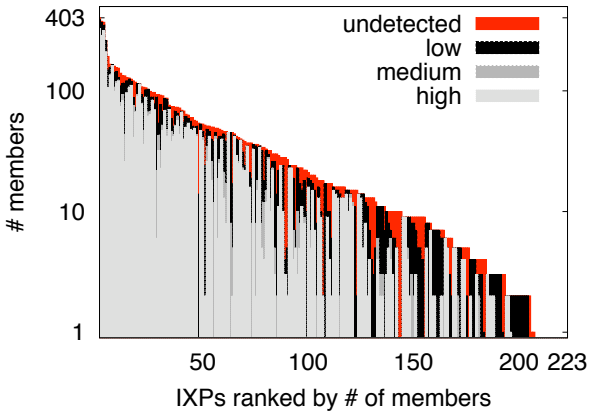

Figure 5: Members per IXP.

A method for measuring the size of an IXP and for which a comparison with the ground truth is possible is in terms of number of members. On the one hand, we take as ground truth the member lists obtained from the IXP websites (when available) or from PCH or PeeringDB, yielding a total of 4K member ASes. On the other hand, we infer the number of members of an IXP via the discovered peerings associated with that IXP. Note that the latter results in a natural classification of IXP members into "high", "medium", and "low" confidence members, depending on the confidence we assigned to the discovered peering links between the ASes involved in the discovered peerings at that IXP. For example, if IXP X is known to have three members (i.e., AS1, AS2, and AS3), and if we discover the peerings AS1-AS2 and AS2-AS3 with any of the three confidence levels, we conclude that AS1, AS2, and AS3 all are members at IXP X, irrespective of what we know or don't know about the relationship between AS2 and AS3. Moreover, AS1, AS2, and AS3 will be classified as "high", "medium", or "low" confidence mem-

bers based on the highest confidence attribute assigned to the corresponding peerings. Fig. 5 shows the the 223 IXPs ranked in decreasing order of the total number of members (based on the ground truth) [12]. The plot depicts the portions of members that were classified as "high", "medium", and "low" confidence members and also shows the fraction of members of each IXP that was not discovered by our method. The good news is that our method performs well for the large IXPs where the difference between the ground truth and the discovered members is typically small and the confidence assigned to the discovered members is in general high. An exception is Interlan IX (at x=48), for which we could not validate any member.

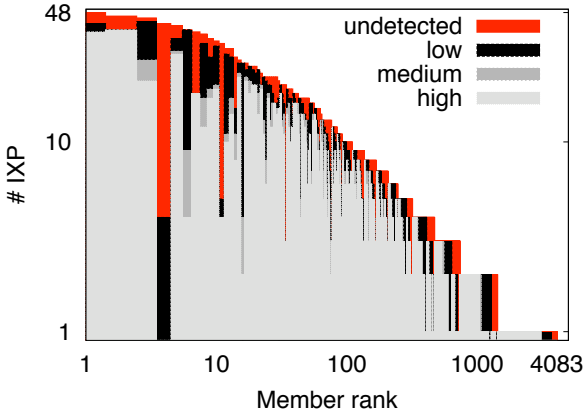

Figure 6: Member presence (log-log).

Sizing IXPs as shown in Fig. 5 only counts the number of IXP members, but does not distinguish them in terms of size or type. In particular, there are several reasons why one and the same AS can show up in the membership list at a number of different IXPs; e.g., increased resilience, connectivity with a particular member, connectivity in a particular location served by the IXP. Counting for each of the 4K member ASes the number of IXPs where its presence was discovered by our method, we then rank the ASes in decreasing order of this *member presence* and plot in Fig.6 for each AS the number of IXPs where it is present, broken down by the level of confidence we have in the discovered membership. Among the top networks we find two content providers: Google (present in 35 IXPs) and Limelight Networks (present in 29 IXPs). For obvious reasons, PCH tops all networks with its presence in 39 IXPs. Note that if an organization uses different ASNs to peer at IXPs, then each ASN will be plotted separately on the x axis. For example, according to the IXP databases, ISC is present as AS3557 in 42 different IXPs, and this number is plotted at x=4 in Fig. 6 as representing the ground truth. However, in our datasets, ISC appears with ASN 3557 at only 4 IXPs and with 18 other ASNs at other IXPs (showing up at a different x-value in the Fig. 6). This deficiency of our analysis could be addressed by grouping ASes that are part of one and the same organization or company, but we have currently no principled method for performing such a grouping.

When an AS becomes a member of an IXP, it will typically peer with several other members at that IXP. For each of the 4K member ASes, we count the number of discovered IXP-related peerings established by this member AS, rank them in decreasing order of this *member connectivity*, and show in Fig. 7 a rank-ordered plot, broken down by the level of confidence we assigned to the discovered peerings. Among the networks that peer most aggressively at the IXPs, we encountered PCH (again for obvious reasons), several Tier-2 ISPs, and Limelight Networks with 500 peerings. Google appears in this analysis with 151 peerings (at a total of 35 IXPs). Similar to ISC, EUnet (x=25 in Fig. 7) owns two different AS numbers. Our validation method could not recognize that they belong to the same organization, and could assign a high confidence to only 6 out of 359 discovered peerings.

[12]Note that the IXP ranking in Fig. 5 is not the same as in Fig. 4, and so a direct comparison between equally ranked IXPs is not meaningful.

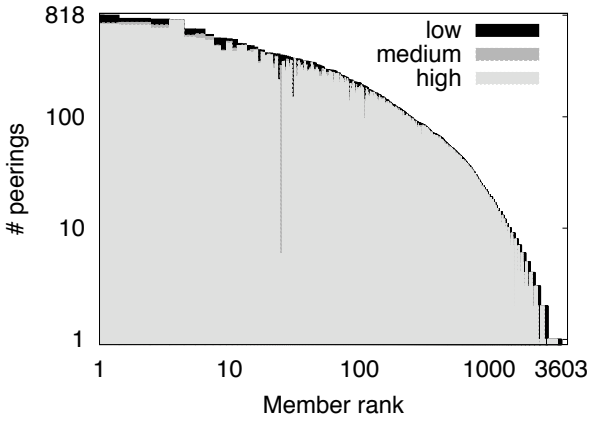

Figure 7: Member connectivity (log-log).

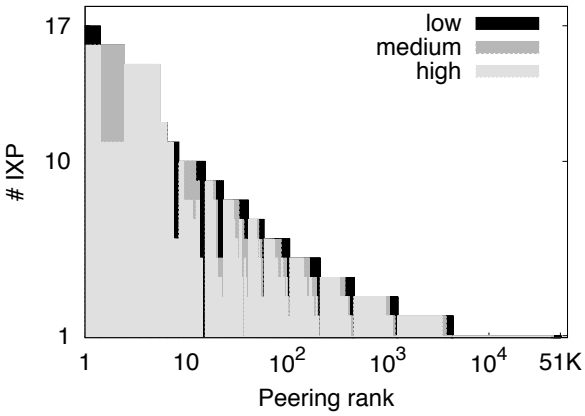

Figure 8: Member multi-connectivity (log-log).

Lastly, a network may not only be present at multiple IXPs, but may in fact have several peering sessions with the same member AS at those multiple IXPs. For example, to eliminate long transit traffic and keep the traffic local, two geographically distributed networks will establish a peering session at each IXP location where they are both present. For each AS pair, we count the total the number of discovered peerings, rank the pairs by their *member multi-connectivity*, and show in Fig. 8 a plot broken down by the level of confidence that we assigned to the discovered peerings. Among the top AS pairs are Google and Limelight who peer with one another at 16 different IXPs. The peering at x=15 for which we have only low confidence information, corresponds to links between Cogent (AS174) and AT&T (AS7018). Again, the use of multiple AS numbers by the same entity, explains the problem. For instance, at Espanix and AMS-IX, AT&T appears under the AS number 2686.

8. CONCLUSION

We have attempted to map all IXPs using the most complete input data, various databases (IXP databases, websites, IRR), and by looking for IXPs in all the known publicly available datasets produced and used by Internet topology researchers (CAIDA, DIMES, PlanetLab). We propose new methods to build additional datasets (targeted traceroutes, source routing, BGP tables). We detect 223 IXPs out of 278 and show that most of the remaining undetected IXPs are actually inactive or not visible to tracerouting. We also discover significantly more IXP-related peerings than previous studies and show that these peerings are not present in currently-used AS maps of the Internet.

As for future work, running our tools regularly will help us understand the evolution of IXPs. New members are added regularly[13], and large IXPs claim to have witnessed an exponential growth during the last few years. Fine-tuning our techniques to eliminate the number of low and medium confidence peerings and focusing on the remaining "islands" in the IXP substrate of the AS-level Internet that have remained by and large invisible to our method will bring our IXP mapping effort to a successful conclusion. Such fine-tuning will also require the development of a principled approach for identifying and dealing with ASes that use multiple ASNs. However, deriving the traffic matrices of IXPs on top of the peering matrices (the main focus of this paper) looms as an important but challenging open problem. Large IXPs typically report a total volume of traffic on the order of Gbps. The use of state-of-the-art tools for bandwidth measurement might help us shed light on this important part of the Internet traffic.

Acknowledgments

We are grateful to Bill Woodcock (PCH) and Bill Manning (EP.net) for sharing their knowledge on IXPs and for patiently answering our questions. We also thank the numerous IXP and network operators who responded to our inquiries and provided invaluable information.

9. REFERENCES

[1] R. Oliveira, B. Zhang, and L. Zhang, "Observing the Evolution of Internet AS Topology," in *SIGCOMM*, 2007.

[2] A. Dhamdhere and C. Dovrolis, "Ten Years in the Evolution of the Internet Ecosystem," in *IMC*, 2008.

[3] Chang et al., "Towards Capturing Representative AS-level Internet Topologies," in *Computer Networks, 44(6):737–755*, 2004.

[4] He et al., "Lord of the Links: A Framework for Discovering Missing Links in the Internet Topology," *IEEE/ACM Trans. Networking*, vol. 17, no. 2, 2009.

[5] Oliveira et al., "In Search of the Elusive Ground Truth: The Internet's AS-level Connectivity Structure," in *SIGMETRICS*, 2008.

[6] M. Roughan, J. Tuke, , and O. Maennel, "Bigfoot, Sasquatch, the Yeti and Other Missing Links: What We Don't Know About The AS Graph," in *IMC*, 2008.

[7] H. Chang, "An Economic-Based Empirical Approach to Modeling the Internet's Inter-Domain Topology and Traffic Matrix," Ph.D. Thesis, University of Michigan, 2006.

[8] Xu et al., "On Properties of Internet Exchange Points and Their Impact on AS Topology and Relationship," in *NETWORKING*, 2004.

[13]For example, between April 2009 and Sept. 2009, we noticed a 7% increase (from 253 to 271) of the number of IXPs in PeeringDB, but only one new IXP in PCH.

[9] B. Huffaker, D. Plummer, D. Moore, and k. claffy, "Topology discovery by active probing," in *Proc. Symposium on Applications and the Internet*, Jan. 2002.

[10] Y. Shavitt and E. Shir, "DIMES: Let the Internet Measure Itself," *CCR*, vol. 35, no. 5, pp. 71 – 74, October 2005.

[11] M. Faloutsos, P. Faloutsos, and C. Faloutsos, "On Power-law Relationships of the Internet Topology," in *SIGCOMM*, 1999, pp. 251–262.

[12] Mahadevan et al., "Orbis: Rescaling Degree Correlations to Generate Annotated Internet Topologies," in *CCR*, 2007.

[13] H. Chang and W. Willinger, "Difficulties Measuring the Internet's AS-level Ecosystem," in *40th Conf. Information Sciences and Systems*, 2006.

[14] Teixeira et al., "In Search of Path Diversity in ISP Networks," in *IMC*, October 2003.

[15] Mao et al., "Towards an Accurate AS-level Traceroute Tool," in *SIGCOMM*, 2003.

[16] Viger et al., "Detection, Understanding, and Prevention of Traceroute Measurement Artifacts," in *Computer Networks, 52(5): 998–1018*, 2008.

[17] Chen et al., "Where the Sidewalk Ends: Extending the Internet AS Graph Using Traceroutes From P2P Users," in *Proc. ACM CoNEXT*, 2009.

[18] W. Norton, "The Evolution of the U.S. Internet Peering Ecosystem," www.nanog.org/mtg-0405/pdf/norton.pdf.

[19] ——, "Internet service providers and peering," www.nanog.org/papers/isp.peering.doc.

[20] B. Woodcock, "Introduction to Exchange Point Economics," www.pch.net/documents/papers/intro-economics, 2006.

[21] He et al., "A Systematic Framework for Unearthing the Missing Links: Measurements and Impact," in *NSDI*, 2009.

[22] Google, "Brief Introduction to Peering," LACNIC meeting, July 2008.

[23] Packet Clearing House, "Internet Exchange Directory," https://prefix.pch.net/applications/ixpdir/.

[24] S. Tomic and A. Jukan, "GMPLS-based Exchange Points: Architecture and Functionality," in *Proc. Workshop on High Performance Switching and Routing*, June 2003.

[25] Amimi et al., "Issues with Inferring Internet Topological Attributes," in *Proc. SPIE*, 2002.

[26] PeeringDB, "Exchange Points List," https://www.peeringdb.com/private/exchange_list.php.

[27] Cymru, "IP to BGP ASN Lookup and Prefix Mapping Services," http://www.cymru.com/BGP/asnlookup.html.

[28] G. Siganos and M. Faloutsos, "Analyzing BGP Policies: Methodology and Tool," in *INFOCOM*, 2004.

[29] T. Kernen, "traceroute.org," www.traceroute.org, 2008.

[30] University of Oregon, "Route Views," www.routeviews.org.

[31] RIPE, "Ris raw data," 2008.

[32] Cooperative Association for Internet Data Analysis, http://www.caida.org/tools/measurement/skitter/.

[33] D. Choffnes and F. Bustamante, "Taming the Torrent: A Practical Approach to Reducing Cross-ISP Traffic in P2P Systems," in *SIGCOMM*, 2008.

[34] Mérindol et al., "Quantifying ASes Multiconnectivity using Multicast Information," in *Proc. ACM SIGCOMM Internet Measurement Conference*, 2009.

[35] Katz-Bassett et al., "Reverse Traceroute," Technical report, 2009.

[36] Y. Chi, R. Oliveira, and L. Zhang, "Cyclops: The Internet AS-level Observatory," *CCR*, October 2008.

[37] R. Govindan and H. Tangmunarunkit, "Heuristics for Internet Map Discovery," in *INFOCOM*, March 2000.

[38] Chang et al., "Inferring AS-level Internet Topology from Router-level Path Traces," in *SPIE ITCom 2001*, 2001.

[39] Yao et al., "Topology Inference in the Presence of Anonymous Routers," in *INFOCOM*, April 2003.

[40] "PlanetLab," www.planetlab.org.