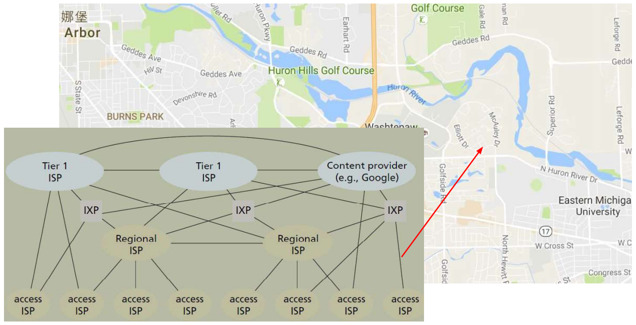
Mapping peering interconnections to a facility

Giotsas, V. et al.,
Proc. of ACM CoNEXT '15, Dec. 2015

Chun-Yu Chen
Ming zhi Yu

Why doing this research?

- Basic Idea
  - Mapping peering interconnections to physical facilities/locations
  - In other words: finding where different ASes interconnects
- Motivations
  - Network troubleshooting
  - Diagnosing attacks
  - Diagnosing congestions



Outline

- Introduction
- Preparation and data collection for traceroute campaign
- Constrained facility search
- Results and validations
- Our assessments

Terminologies

- Autonomous System (AS)
  - The classic definition of an Autonomous System is a set of routers under a single technical administration, using an interior gateway protocol and common metrics to route packets within the AS, and using an exterior gateway protocol to route packets to other ASes. (RFC 1930)
- Internet Exchange Point (IXP)
  - An IXP is a physical infrastructure composed of layer-2 Ethernet switches where participating networks can interconnect their routers
- Peering
  - Peering is a voluntary interconnection of administratively separate Internet networks for the purpose of exchanging traffic between the users of each network

Terminologies (Cont'd)

- Public Peering
  - Public peering, also referred to as public interconnect, is the establishment of peering connections between two members of an IXP via the IXP's switch fabric.
- Private Peering with cross-connect
  - Cross-connecting two interfaces of two networks with dedicated medium.

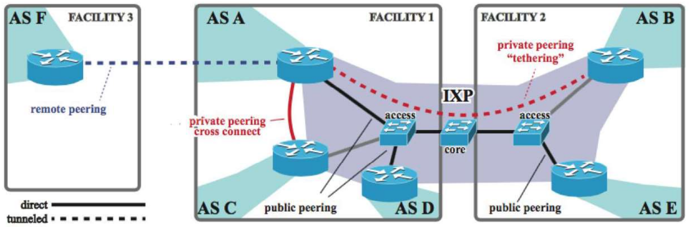
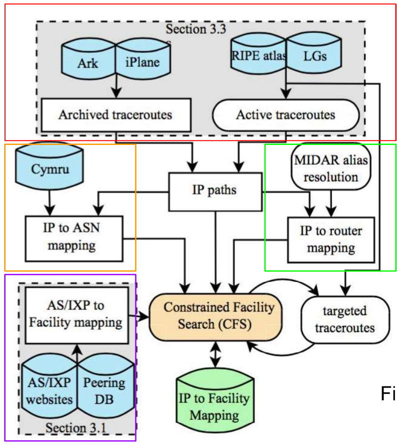


Figure 1

Methodology Overview



Section 3.3

Ark

iPlane

RIPE atlas

LGs

Archived traceroutes

Active traceroutes

Cymru

IP paths

MIDAR alias resolution

IP to ASN mapping

IP to router mapping

AS/IXP to Facility mapping

Constrained Facility Search (CFS)

targeted traceroutes

AS/IXP websites

Peering DB

IP to Facility Mapping

Section 3.1

Traceroute Data Gathering

Mapping IP to ASN

Resolve Alias

Traceroute => IP => ASN => Facility

Mapping AS/IXP to Facilities

Figure 4

Terminologies (Cont'd)

- Private peering over IXP
  - Private peering using VLAN on shared switching fabric
- Remote Peering

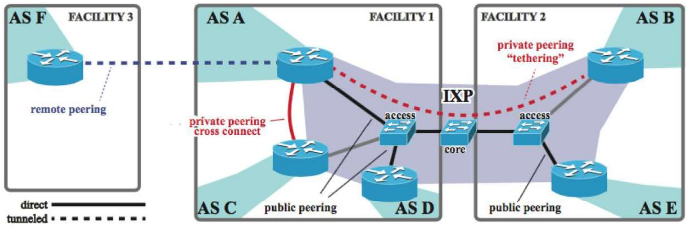


Figure 1

Preparation of Datasets

Two categories of data:

- Location-related information
  - In which facilities do an AS's routers or an IXP's switches locate? Without these, we will never know the exact location where networks interconnect.
- Network-related information
  - How to get IP address of a router interface (traceroute)
  - How to know when two networks interconnect? (IP router interface to ASN)
  - How to determine which router interface belongs to which AS?

Obtaining AS and IXP-related Information

- AS-to-facilities mapping
  - ○ a list of interconnection facilities where ASes' routers are present
  - ○ Data sources: PeeringDB and Network Operating Centers (NOCs)
- IXP related information
  - ○ Compiled a list of IXPs.
  - ○ For each IXP, obtained its prefix, associated interconnection facilities, and members
  - ○ Data sources: IXP websites, PeeringDB, and Packet Clearing House
  - ○ Example: https://github.com/euro-ix/json-schemas

Tools to Perform Traceroute

Several publicly available traceroute servers were used.

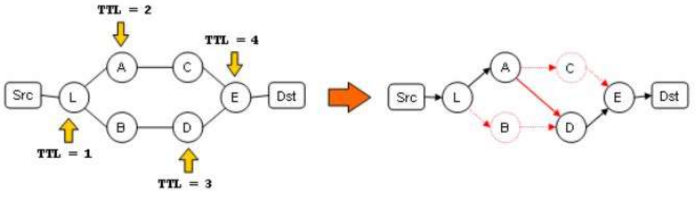
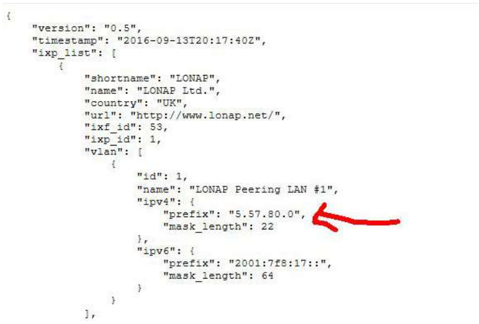


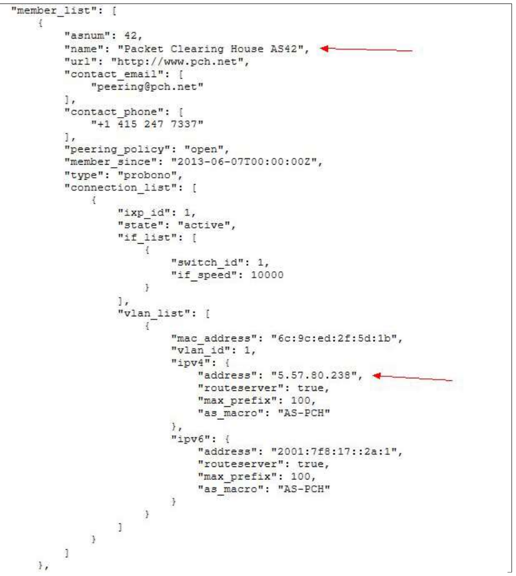
Figure 1 *Traceroute's deficiencies under load balancing*

- RIPE /
  - ○ A (
  - ○ Pa
- Lookin
  - ○ Pr ı-privileged de
- iPlane
  - ○ Performs daily IPv4 traceroute campaigns. Two archives were used.
- CAIDA Archipelago (Ark)
  - ○ A globally distributed measurement platform.

IXP Info Examples



```
{
    "version": "0.5",
    "timestamp": "2016-09-13T20:17:40Z",
    "ixp_list": [
        {
            "shortname": "LONAP",
            "name": "LONAP Ltd.",
            "country": "UK",
            "url": "http://www.lonap.net/",
            "ixf_id": 53,
            "ixp_id": 1,
            "vlan": [
                {
                    "id": 1,
                    "name": "LONAP Peering LAN #1",
                    "ipv4": {
                        "prefix": "5.57.80.0",
                        "mask_length": 22
                    },
                    "ipv6": {
                        "prefix": "2001:7f8:17::",
                        "mask_length": 64
                    }
                }
            ],
```



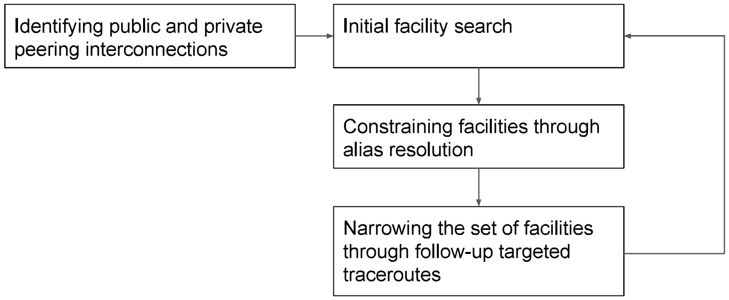
```
"member_list": [
    {
        "asnum": 42,
        "name": "Packet Clearing House AS42",
        "url": "http://www.pch.net",
        "contact_email": [
            "peering@pch.net"
        ],
        "contact_phone": [
            "+1 415 247 7337"
        ],
        "peering_policy": "open",
        "member_since": "2013-06-07T00:00:00Z",
        "type": "probono",
        "connection_list": [
            {
                "ixp_id": 1,
                "state": "active",
                "if_list": [
                    {
                        "switch_id": 1,
                        "if_speed": 10000
                    }
                ],
                "vlan_list": [
                    {
                        "mac_address": "6c:9c:ed:2f:5d:1b",
                        "vlan_id": 1,
                        "ipv4": {
                            "address": "5.57.80.238",
                            "routeserver": true,
                            "max_prefix": 100,
                            "as_macro": "AS-PCH"
                        },
                        "ipv6": {
                            "address": "2001:7f8:17::2a:1",
                            "routeserver": true,
                            "max_prefix": 100,
                            "as_macro": "AS-PCH"
                        }
                    }
                ]
            }
        ]
    },
```

Linking Route Information with Location Information

Traceroute provides IP addresses along a route, a mapping from each IP router interface to an ASN is needed.

- Team Cymru's IP-to-ASN service was used to construct such mapping
- Mapping based on longest prefix matching is prone to errors
  - ○ Same prefix could be shared between siblings or neighboring AS
  - ○ Alias resolution is performed to group IP interfaces into routers (i.e. IP^1_x and IP^2_x are two interfaces of the same router)
- MIDAR system was used to group IP interfaces into routers

Constrained Facility Search Overview


Identifying public and private peering interconnections

Initial facility search

Constraining facilities through alias resolution

Narrowing the set of facilities through follow-up targeted traceroutes

Step 1 - Identifying public and private peering interconnections

- Ex. If we have traceroute data (IP_A, IP_e, IP_B)
  - If IP_e is in an IXP address space => IP_A and IP_B is connected through public peering

- Ex. If we have traceroute data (IP_A, IP_B)
  - IP_A and IP_B is connected through private peering

Step 2 - Initial facility search

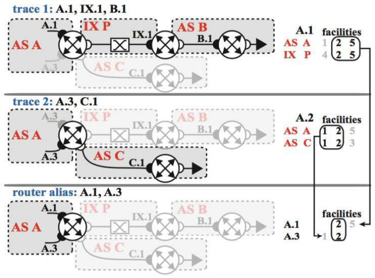
- For public peering traceroute (IP_A, IP_e, IP_B)
  - If there is only one common facility:
    AS A {f_1, f_2, f_3} and the IXP {f_3, f_4, f_5} => IP_A and IP_e are in the same facility f_3
  - If there are multiple common facilities
    AS A {f_1, f_2, f_3} and the IXP {f_2, f_3, f_4} => IP_A is in {f_2, f_3}
  - If there is no common facility
    AS A {f_1, f_2, f_3} and the IXP {f_4, f_5, f_6}
    => Either IP_A and IP_B are connected through remote private peering or the data of facilities is incomplete
- For private peering (IP_A, IP_B)
  - Also compare {F_A} and {F_B} like discussed process above

Step 3 - Constraining facilities through alias resolution

- For unresolved interface try to find the interface to a facility by comparing the possible locations to their aliases
  - Ex. IP_A^1 is in {f_1, f_2} and IP_A^2 is in {f_2, f_3}.
    IP_A^1 and IP_A^2 are aliases. => IP_A^1 and IP_A^2 are in f_2

Step 4 - Narrowing the set of facilities through follow-up targeted traceroutes

- Conduct more traceroute targeting those unresolved interfaces to obtain more data for those unresolved interfaces and then repeat step 2 to step 4.
  - Traceroute (IP_A^1, IP_e, IP_B)
    AS A {f_1, f_2, f_5}, IXP {f_2, f_4, f_5} => IP_A^1 is in f_2 or f_5
  - If we already know that IP_A^1 and IP_A^3 are aliases
  - We target both IP_A^1 and IP_A^3 for more traceroute data
  - New traceroute (IP_A^3, IP_C^1)
    AS C {f_1, f_2, f_3} => IP_A^3 is in f_1 or f_2
  - IP_A^1 and IP_A^3 are in f_2



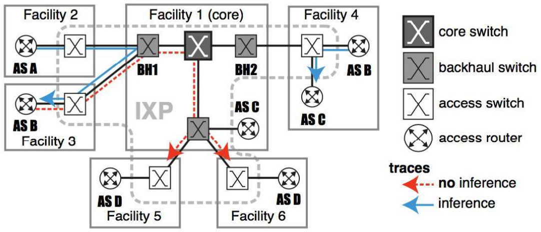
Evaluation Method and Results

The algorithm was evaluated on a number of important groups of interconnections. The following ASes were selected by the author as traceroute targets.

- Content providers: Google, Yahoo!, Akamai, Limelight and Cloudfare
  - BGP announcements, whitelists, URLs served by theses networks→ IP
- Tier 1 ISP: NTT, Cogent, and Deutsche Telekom, Level 3 and Telia
  - Constructed a list of all their peers and selected one active IP for each peer.
- Results: after 100 iterations, 9812 router interfaces (70.65%) were mapped to a single interconnection facility.
  - About 9% of unresolved interfaces were constrained to the level of city
  - For unresolved interfaces, about 33% of those interfaces have no facility data

Proximity Heuristic

- Networks connected to the same switch, or connected to switches attached with the same back-haul switch, exchange traffic locally and not via the core switch.
  - We have traceroute (IP_A, IP_e, IP_B) and we already found that IP_A is at f_2
    => IP_B is more likely to be in f_3 than f_4



Validation

Inferences about interconnection facility was validated by four different sources of information:

- Direct feedback
  - Obtained validation from two CDN operators. 88% (474/540) were correct at the facility level.
- BGP communities
  - Entry point of a route is often tagged by an operator using BGP communities. 92% (76/83) of public peering interfaces and 89% (94/106) of cross-connect interfaces were correct.
- DNS records
  - Hostnames of router interfaces sometimes contain facility information. 91% (91/100) of public peering interfaces and 89% (191/213) of cross-connect interfaces were correct.
- IXP websites
  - A few IXPs' websites list the exact facilities where and IP interfaces with which their members are connected. 99.1% (322/325) public and 44/48 (91.7%) private were correct.

Related work

One advantage of CFS compared to methods proposed in related works is that CFS is able to locate interconnection facility at a much finer granularity.

- Methods that provide mapping at city level
  - Augustin et al.[1],
  - Dasu[2],
  - Giotsas et al.[3],
  - Calder et al[4].
- Methods that provide mapping at country or state level
  - IP geolocation

Interesting Facts (Aha points)

- Proximity heuristic
  - Networks connected to the same switch, or connected to switches attached with the same back-haul switch, exchange traffic locally and not via the core switch
- Multiple roles that routers play at interconnection facilities
  - Private interconnections and public peering
- EU IXP is more public than their counterparts in the US

What we like about the paper and novel points

- The proposed method systematically extracts hidden information (i.e., where do two ASes interconnect, etc.) from data available to the public (i.e., mapping info)
- CFS can be generalized to map interfaces in different granularity (i.e., building, city, state, etc.)

What can be improved or extended?

- What is overlooked?
  - The total number of peering facilities that had been validated are unmentioned.
  - CSF has limited usage while dealing with remote peering and tethering.
- How can it be improved?
  - The result is highly dependent on the data available to the public. The true questions is that how do we detect if there is an error during mapping?
  - What can we do after obtaining the result?
- Overlooked Advantage
  - This method can be applied to mapping problems that map a certain set of elements into another set of elements? (e.g. locating PoP facilities)

Thank you for your attention!

References

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

[2] M. A. Sanchez, J. S. Otto, Z. S. Bischof, D. R.Chones, F. E. Bustamante, B. Krishnamurthy, and W. Willinger. Dasu: Pushing Experiments to the Internet's Edge. In NSDI, 2013.

[3] V. Giotsas, M. Luckie, B. Huaker, and kc clay.Inferring Complex AS Relationships. In ACM IMC,2014.

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