Network Security and Measurement

- BGP Measurements -

Prof. Dr. Thomas Schmidt

http://inet.haw-hamburg.de | t.schmidt@haw-hamburg.de

Agenda

Passive BGP measurements

Active BGP measurements

Inferring AS relations

Listen to the Change

PASSIVE BGP MEASUREMENTS

Observing the BGP Control Plane

BGP glues the Internet together in the default-free zone

- Each peer has a route to any prefix reachable on the Internet
- Large ISPs (Tier 1 ++) and IXPs operate default-free

Inspecting these ‘full tables’ opens a complete, location-specific view onto the Internet

- Paths to prefixes across ASs
- Per view point, ASs arrange in a tree

Observing the BGP Control Plane

Looking Glasses and Route Collectors are passive view points on the BGP control plane

BGP glues the Internet together in the default-free zone

- Each peer has a route to any prefix reachable on the Internet
- Large ISPs (Tier 1 ++) and IXPs operate default-free

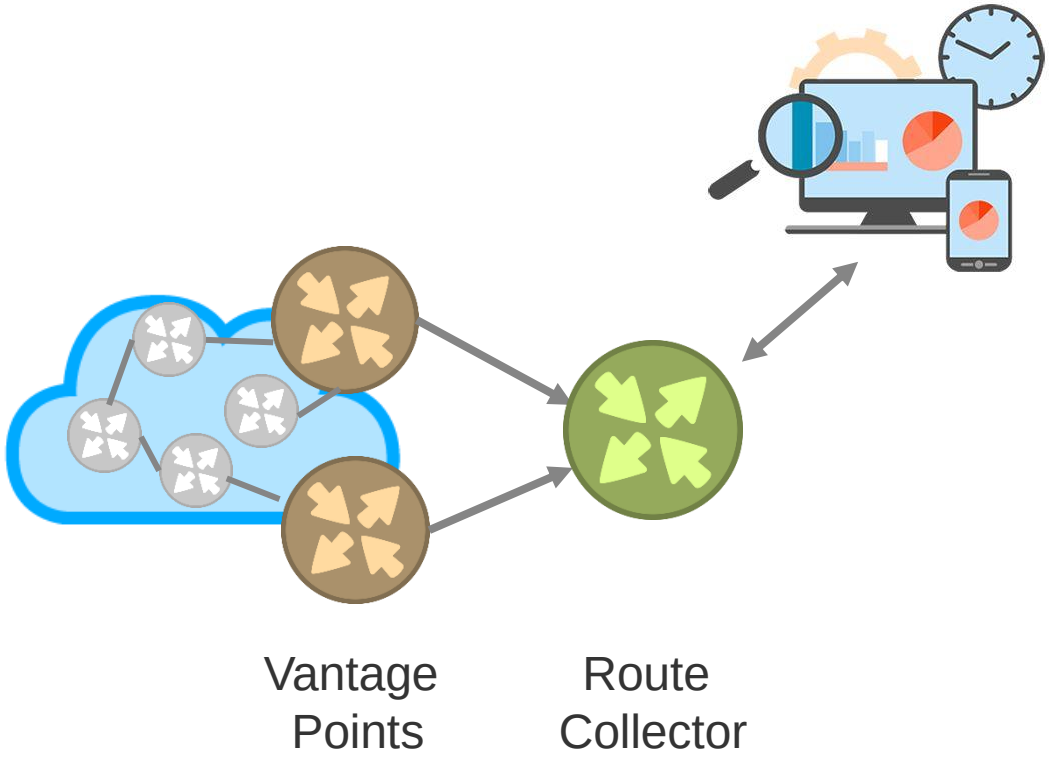
Inspecting these ‘full tables’ opens a complete, location-specific view onto the Internet

- Paths to prefixes across ASs
- Per view point, ASs arrange in a tree

How does a route collector work?

Route collector peers with other ASs: Vantage Points

- Takes role of a customer
- Should receive full BGP tables
- Route collector receives the view of how the Vantage Point sees the Internet
- Route collector grants access to its table



Two types of BGP data

Route table dumps

- From collectors or looking glasses
- Formats MRT (RFC 6396) or ASCII (Cisco)

Incremental BGP updates

- Life feeds of BGP speakers
- Console outputs in MRT format

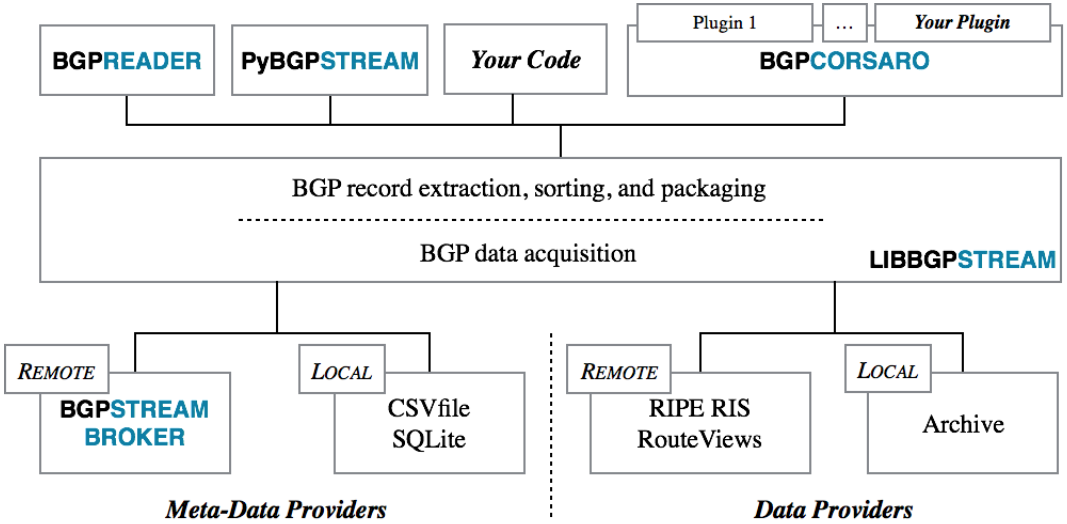
BGP Stream

An open source software framework for accessing BGP data

- Current collector dumps
- Real-time updates
- Historic dumps

Integrates various BGP sources and archives

Access via unified APIs



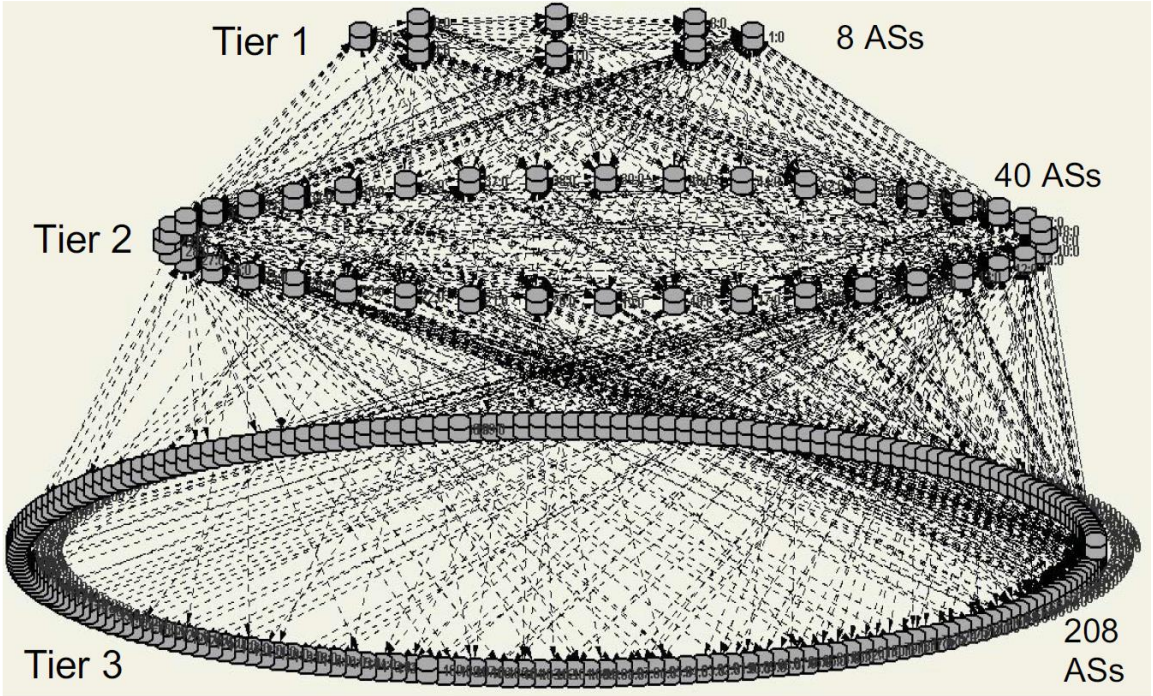
https://bgpstream.caida.org/

Reconstructing an AS-Level Topology

Collecting and combining all AS-AS links from BGP dumps allows reconstruction of an AS-level topology

Caveat:

- Many links remain hidden due to limited visibility
- Links are prefix-dependent and directed
- Links are not transitive



Picture downsampled. Souce: Sriram et al., JSAC 24 (10), Oct. 2006

Beacons

ACTIVE BGP MEASUREMENTS

Overall objective: Study BGP dynamics

What do we learn from monitoring BGP announcements and withdrawals?

Can we design a sound measurement methodology to answer questions about the behavior of BGP peers?

Motivation 1: Impact of implementations on BGP dynamics

Different implementations may behave differently (e.g., bugs, different default values)

Example: MinRouterAdvertTimer
Minimum amount of elapsed time between advertisement of routes to a destination

In an ideal world, different vendor implementations have no impact on the operation of BGP

Motivation 2: Route Flap Damping

Problem of frequently changing routes:
May introduce unnecessary load on routers.

Think about the BGP decision process, for example.

May have impact on upper layer

Different routes, different delays

Goal: Increase routing stability.

Approach: Delay BGP updates to allow updates in batches and reduce update traffic

Motivation 3: BGP convergence time

How long does it take until multiple (all) BGP routers know about the same routing changes?

Motivation 4: BGP Zombies

When an AS withdraws a prefix that it had originally announced, the prefix should disappear from all routing tables after some time.

Does this assumption hold in real-world?

We are interested in studying these examples.

What to do?

We are interested in studying these examples.

What to do?

We need active BGP measurements.

BGP Beacons [IMC'03]

Publicly documented prefixes having global visibility and a published schedule for announcements and withdrawals.

BGP Beacons [IMC’03]

Publicly documented prefixes having global visibility and a published schedule for announcements and withdrawals.

198.133.206.0/24

Announcement:
Withdrawal:

Some design challenges

Which frequency do you announce and withdraw?

Which timestamps do you use?

Which prefixes do you announce?

How do you identify the right signal?

At which frequency do you send a BGP Update?

The **Beacon period** is the time between each Beacon event.

The **Beacon event** is either an announcement of the Beacon prefix or a withdrawal.

At which frequency do you send a BGP Update?

The **Beacon period** is the time between each Beacon event.

The **Beacon event** is either an announcement of the Beacon prefix or a withdrawal.

Typical Beacon period is two hours, to allow route flap damping to expire.

Which timestamps do you use?

Problem

Local system delays

You want to know when the BGP update was actually sent.

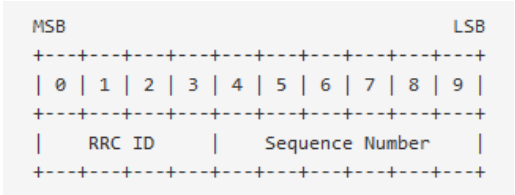
Encoding additional meta data

Overloading BGP Aggregator attribute

Set Aggregator IP attribute to `10.X.Y.Z`

Seconds since the start of the month (UTC)
and time of the announcement

Set Aggregator ASN attribute to

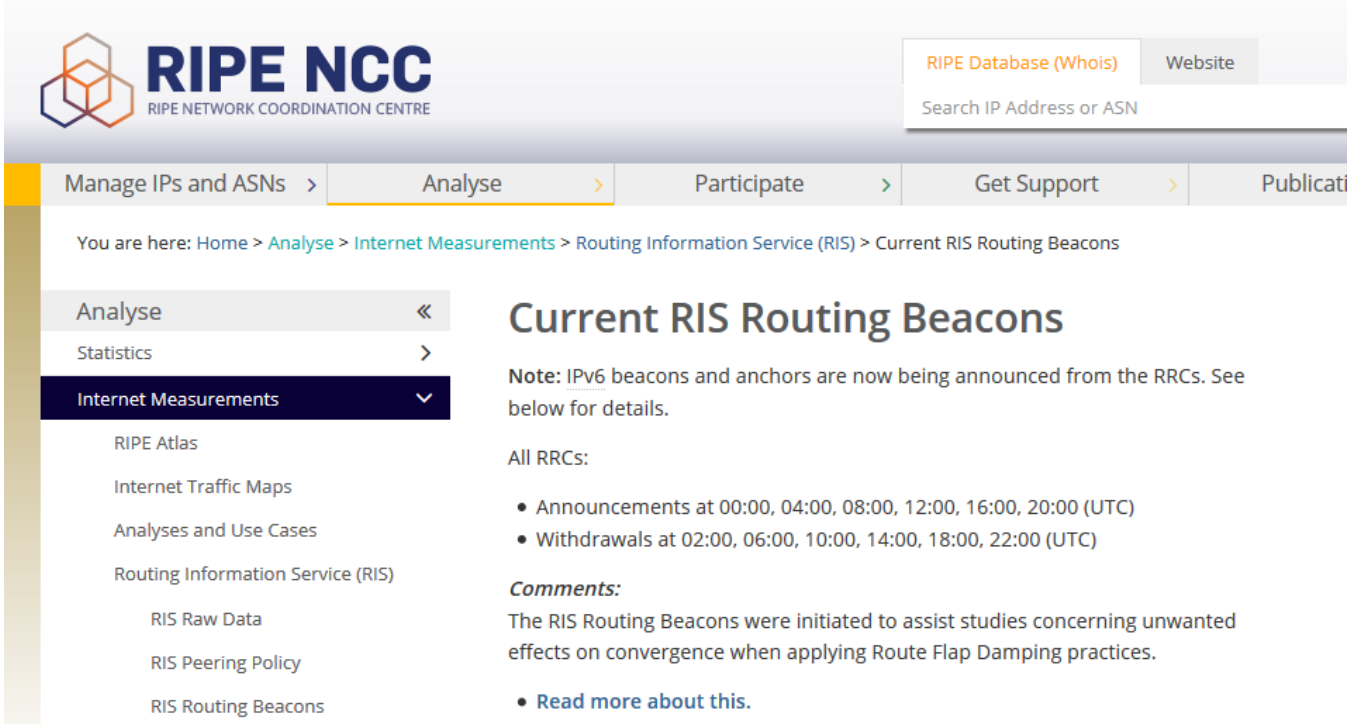


Which prefixes do you announce?

Prefixes should receive only a small amount or no traffic at all.

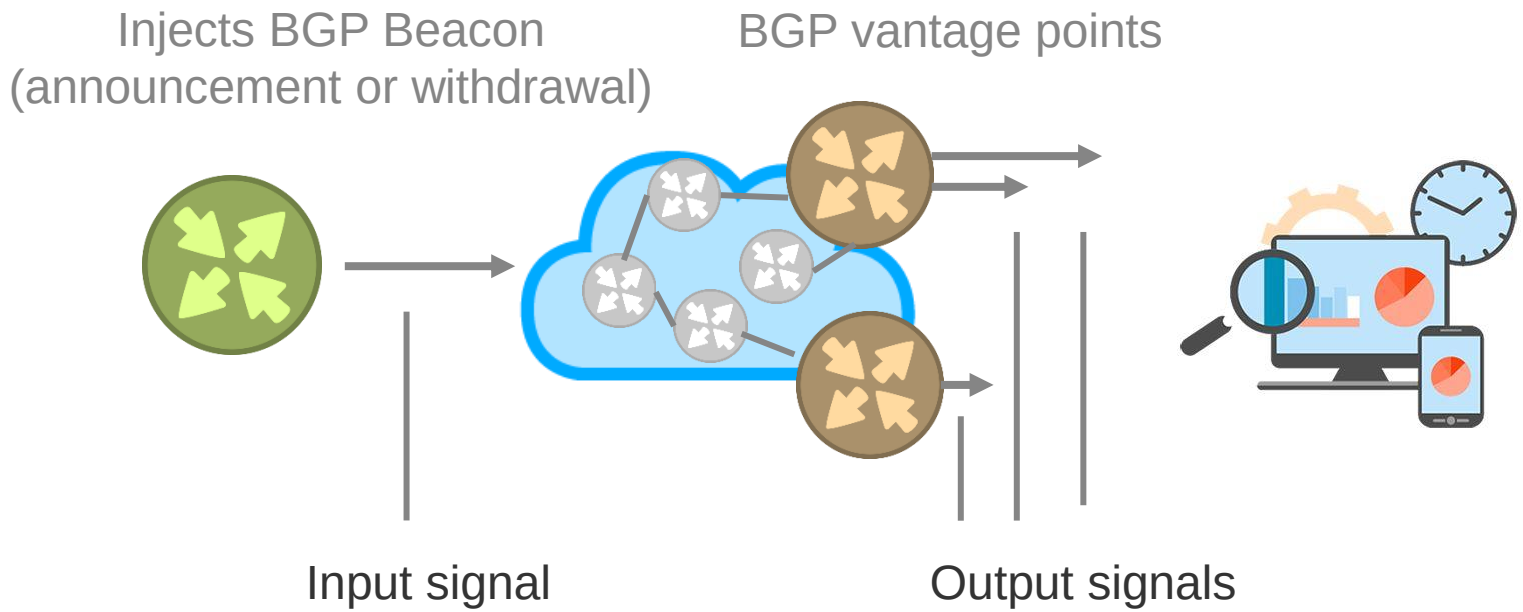
v4 Prefixes should be /24 (v6 /48) or less specific to prevent common filter rules.

Example: RIPE RIS BGP Beacons



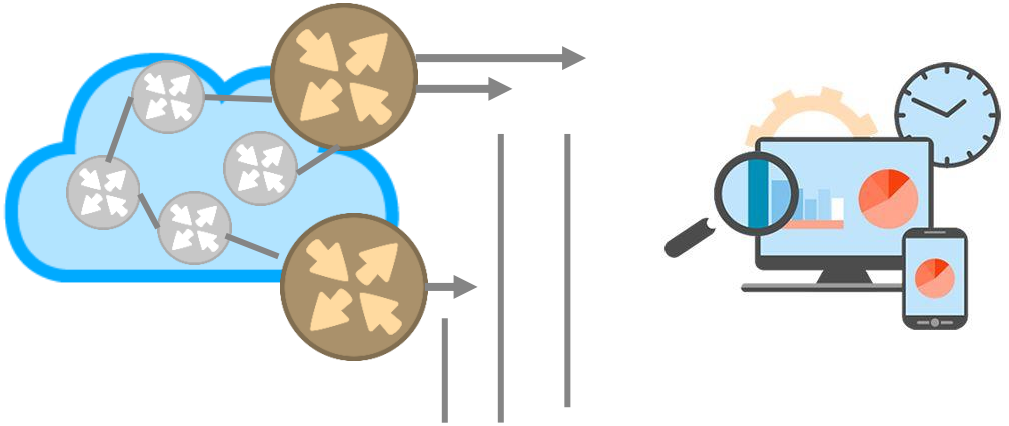
https://www.ripe.net/analyse/internet-measurements/routing-information-service-ris/current-ris-routing-beacons

A single BGP input signal can cause various BGP output signals



You may even see output signals for the Beacon prefix not triggered by the Beacon

Reasons for example:
Timeouts due to link breaks or congestion; routing changes of upstream providers.



Output signals

You need to

Identify signals

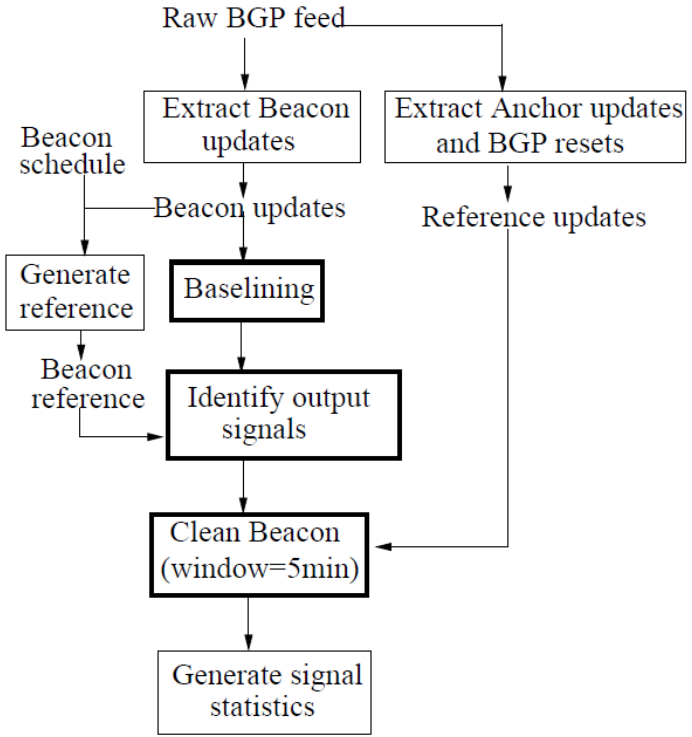
Associate output signals with a single input signal

Clean data

Remove updates that do not affect actual routing changes

Basic methodology: Overview

Objective: Filter events that are not caused by BGP Beacons



[Mao et al. “BGP Beacons,” ACM IMC 2003.]

(1) Baselining

Remove updates that are identical to previous updates or differ only in community or Multi-exit-discriminator (MED)

Makes comparison between peers more fair

(2) Signal identification:
Group updates together according to input signals

Case 1: Beacon AS peers directly with route collector

Beacon AS produces very clean output signals

Output signal will very likely receive first

Case 2: Beacon AS does not peer directly with route collector

Apply heuristics, e.g., look for large time gaps between updates

Use Beacon schedule as reference and sequence number in meta-data

(3) Filtering of noise

Idea: Compare to an anchor prefix

Anchor prefix is announced from the same origin AS as the Beacons but stable

Anchor serves as calibration point to identify non-Beacon routing changes

Delete signals that are also visible for the anchor prefix

Use cases [data from IMC'03 paper]

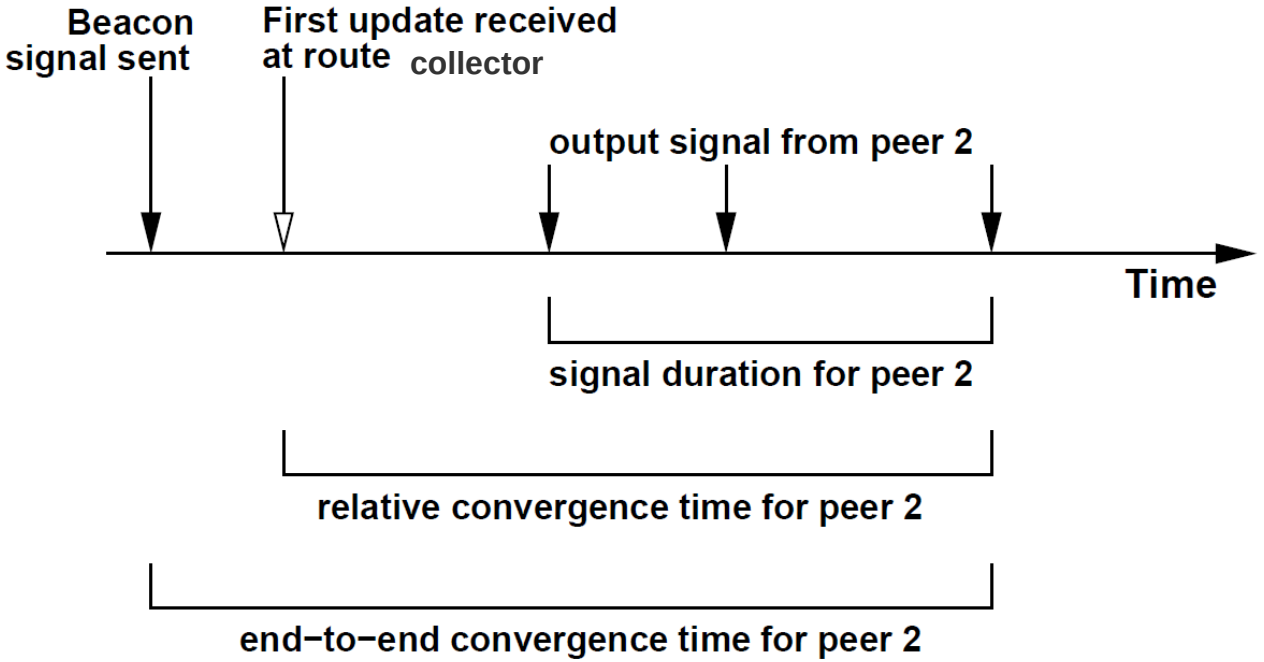
Implementation impact

Route flap damping

Convergence time

BGP Zombies

What to measure?



Use cases

Implementation impact

Route flap damping

Convergence time

BGP Zombies

Implementation impact

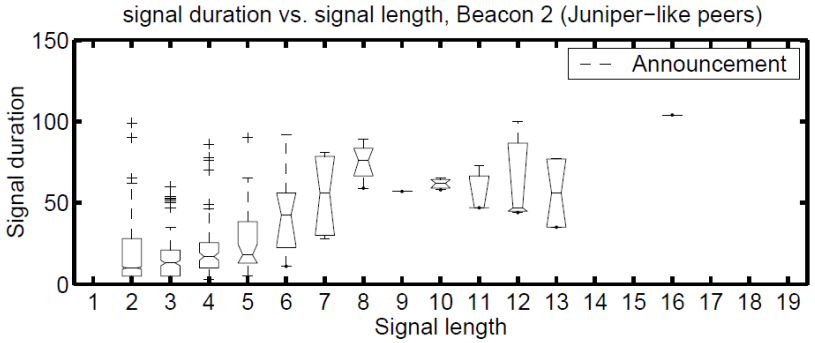
Analyze Cisco and Juniper routers?

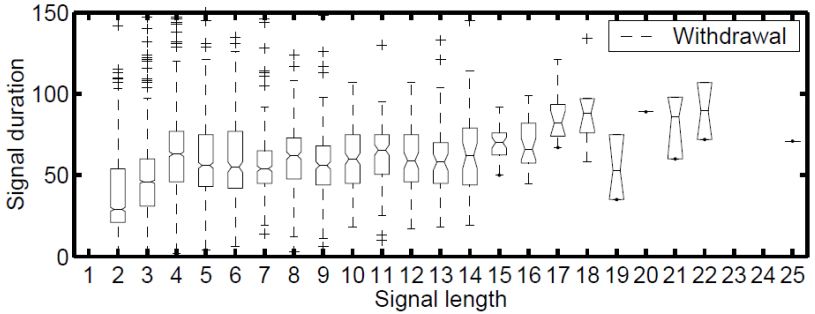
Ground truth regarding vantage points available.

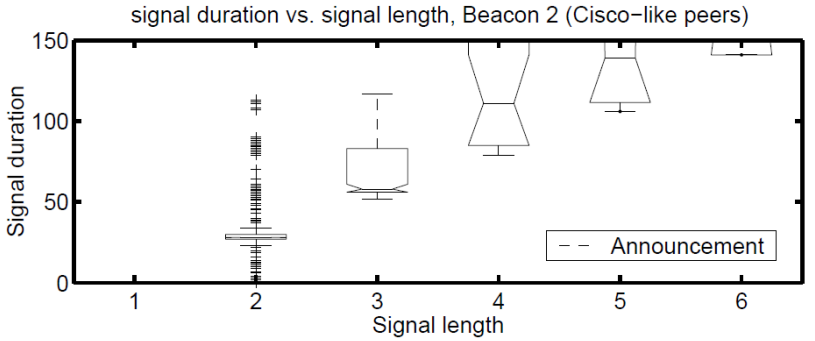
Average value

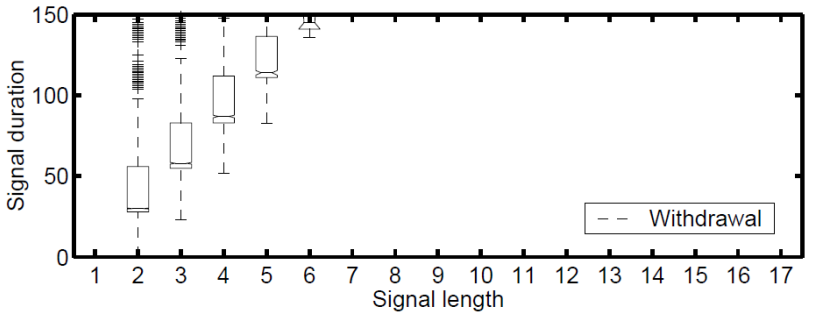
Peer	Type	signal length A	signal length W	duration A	duration W	inter-arrival A	inter-arrival W	% of short inter-arrivals A	% of short inter-arrivals W
147.28.255.1	Cisco	1.20	2.07	6.79	48.4	34.8	45.4	1.56	0.44
147.28.255.2	Juniper	1.50	2.49	7.13	44.3	14.2	29.6	12.76	4.37

Juniper did not use MinRouteAdvTimer by default.


signal duration vs. signal length, Beacon 2 (Juniper−like peers)
Announcement
Signal duration
Signal length
0
50
100
150
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19


Withdrawal
Signal duration
Signal length
0
50
100
150
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25


signal duration vs. signal length, Beacon 2 (Cisco−like peers)
Announcement
Signal duration
Signal length
0
50
100
150
1 2 3 4 5 6


Withdrawal
Signal duration
Signal length
0
50
100
150
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17

Use cases

Implementation impact

Route flap damping

Convergence time

BGP Zombies

Route Flap Damping: background

Router keeps track of penalty value, on a per route and per neighbor basis

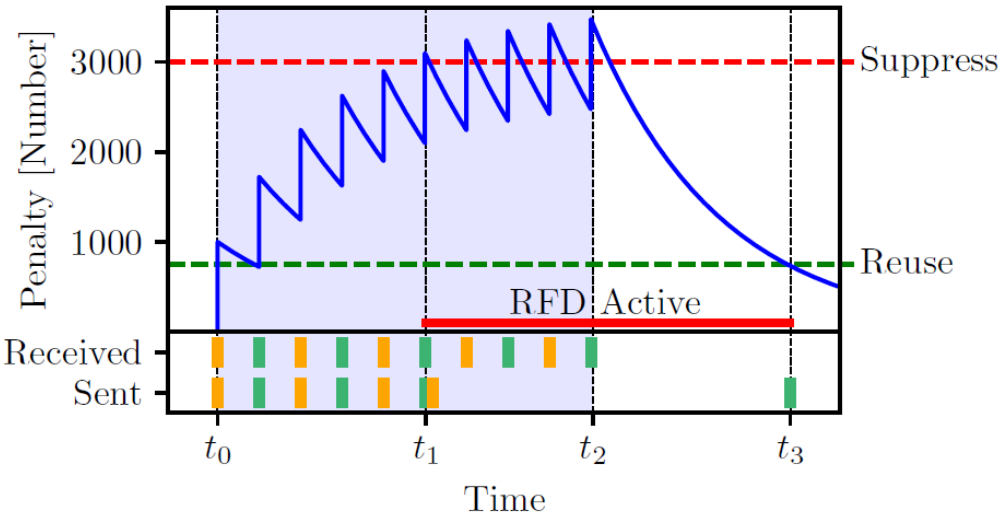
- Penalties increase additively, but decrease by a factor (AIMD)

Different penalty increments per implementation

When Suppress threshold is exceeded, route is not used anymore

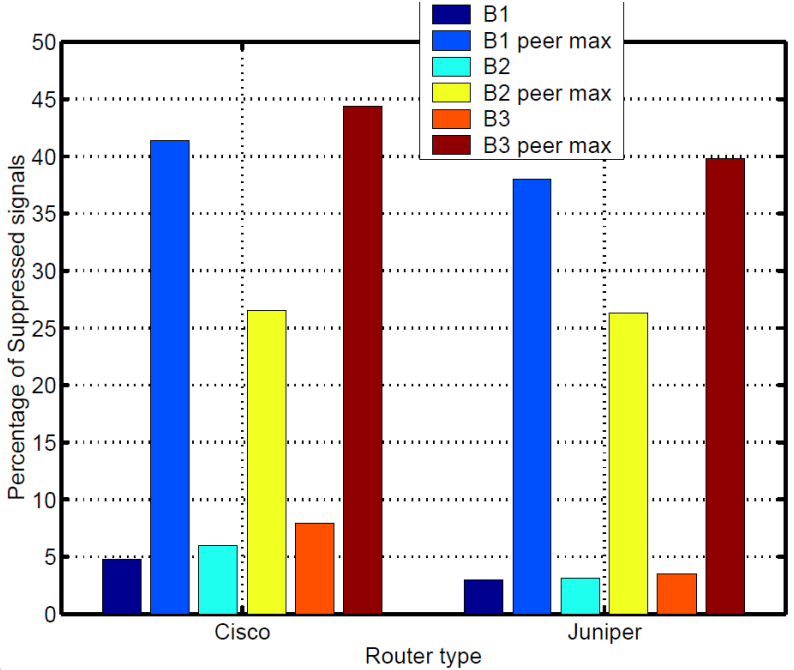
There is a limit how long a route is suppressed

If the suppressed route is the only route, prefix becomes unavailable.



To which extent does route flap damping suppress stable routes?

BGP Beacons are good infrastructure to answer this question



Use cases

Implementation impact

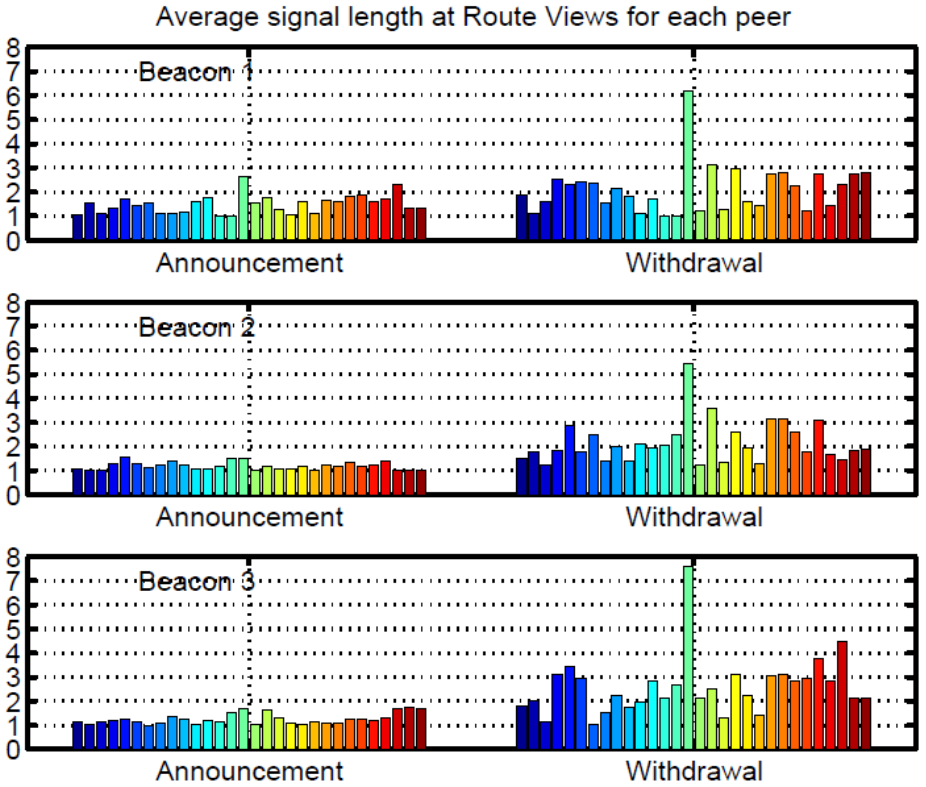
Route flap damping

Convergence time

BGP Zombies

Convergence time

Different peers see different numbers of announcements.

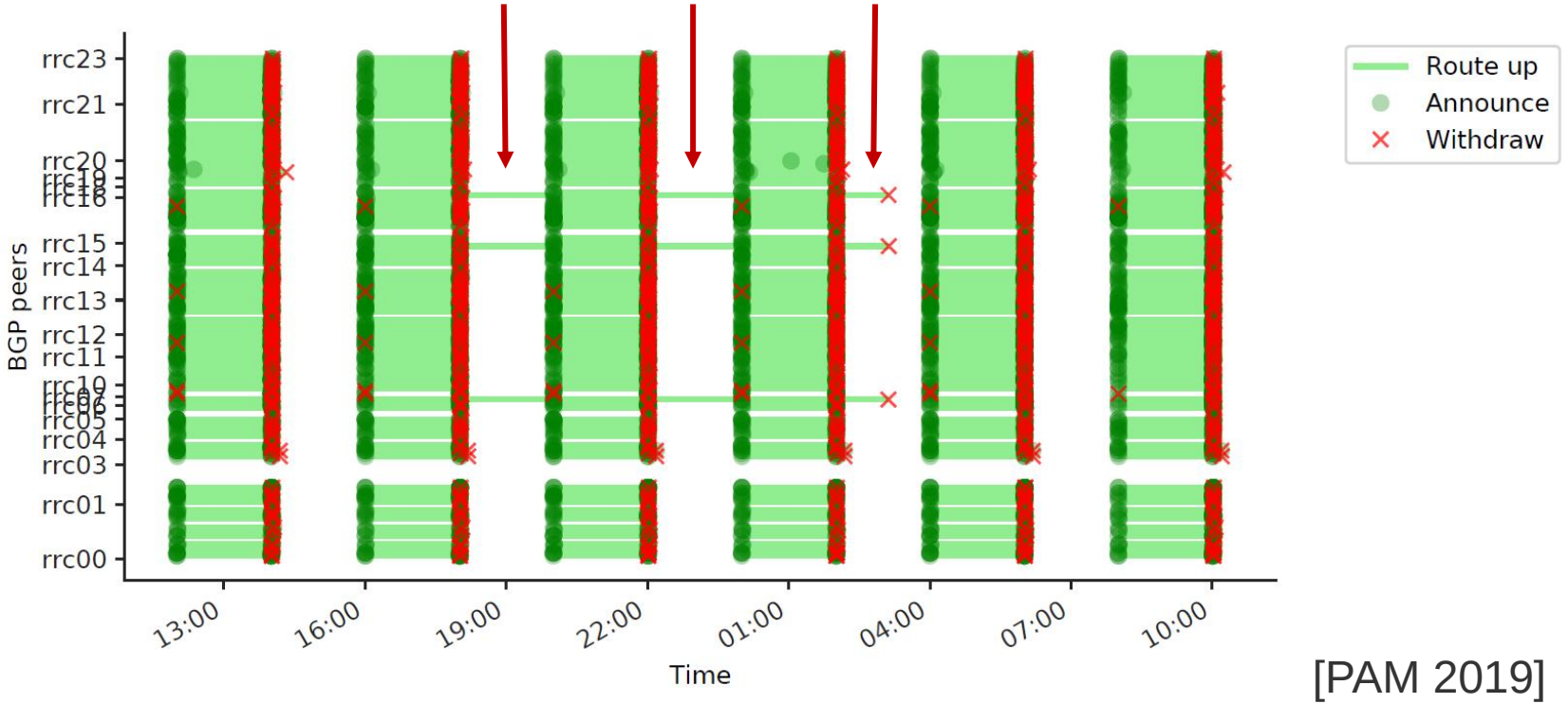


BGP Zombie

BGP Zombie refers to an active Routing Information Base (RIB) entry for a prefix that has been withdrawn by its origin network, and is hence not reachable anymore.

Also known as “ghosts” or “stuck routes”.

BGP Zombies: Analysis based on RIPE Beacons and RIS



[PAM 2019]

Literature: BGP Beacons

Z. Morley Mao, Randy Bush, Timothy G. Griffin, and Matthew Roughan. 2003. BGP Beacons. In *Proceedings of the 3rd ACM SIGCOMM conference on Internet measurement* (IMC '03). ACM, New York, NY, USA. https://doi.org/10.1145/948205.948207

BGP Beacons

Z. Morley Mao; Randy Bush†; Timothy G. Griffin‡; Matthew Roughan§

ABSTRACT

The desire to better understand global BGP dynamics has motivated several studies using active measurement techniques, which inject announcements and withdrawals of prefixes from the global routing domain. From these one can measure quantities such as the BGP convergence time. Previously, the route injection infrastructure of such experiments has either been temporary in nature, or its use has been restricted to the experimenters. The routing research community would benefit from a permanent and public infrastructure for such active probes. We use the term *BGP Beacon* to refer to a publicly documented prefix having global visibility and a published schedule for announcements and withdrawals. A BGP Beacon is to be used for the ongoing study of BGP dynamics, and so should be supported with a long-term commitment. We describe several BGP Beacons that have been set up at various points in the Internet. We then describe techniques for processing BGP updates when a BGP Beacon is observed from a BGP monitoring point such as Oregon's Route Views. Finally, we illustrate the use of BGP Beacons in the analysis of convergence delays, route flap damping, and update inter-arrival times.

Categories and Subject Descriptors

C.2.2 [**Computer-Communication Networks**]: Network Protocols—*Routing protocols*

General Terms

Measurement, Experimentation

Keywords

Network measurements, Border Gateway Protocol, convergence time

*University of California at Berkeley, email: zmao@eecs.berkeley.edu.
†Internet Initiative Japan, email: randy@psg.com.
‡Intel Research, email: tim.griffin@intel.com. This work was conducted while Tim was with AT&T Labs–Research.
§AT&T Labs–Research, email: roughan@research.att.com.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
IMC'03, October 27–29, 2003, Miami Beach, Florida, USA.
Copyright 2003 ACM 1-58113-773-7/03/0010 ...$5.00.

1. WHAT IS A BGP BEACON?

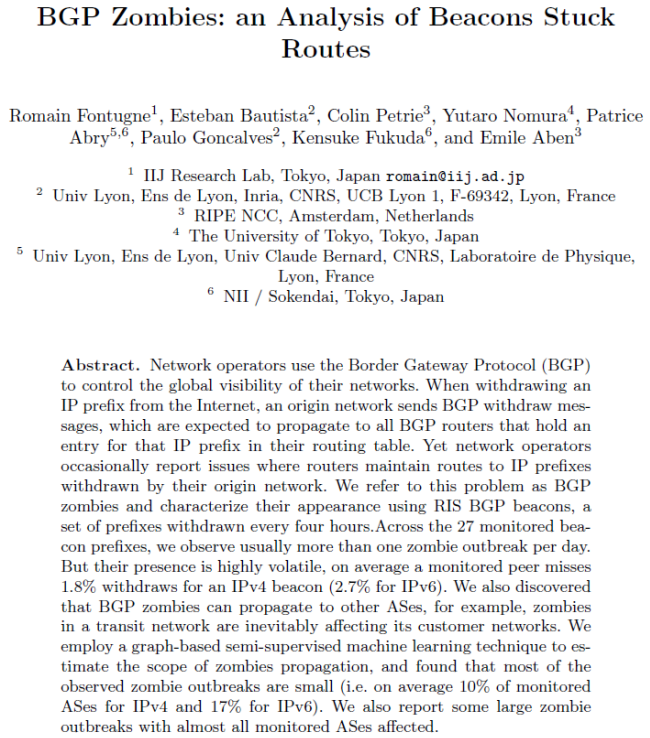
The Border Gateway Protocol (BGP) [1, 2, 3] is central to the stability and robustness of the Internet. Passive monitoring of BGP updates has resulted in important insights into the dynamics of BGP [4, 5, 6]. Several public sources, such as Oregon's Route Views [7] and the RIPE Routing Service [8], provide BGP updates collected from a large number of points in the Internet. Passive measurements are not sufficient for all purposes and so active techniques have also been employed in the analysis of BGP dynamics [9, 10]. With the active approach, prefixes are announced and withdrawn from the global routing domain while quantities such as convergence time are measured. The main advantage of the active approach is that the *input* to the routing system is known, which allows inferences to be made that would be difficult or impossible with purely passive measurements.

To date, the route injection infrastructure of such experiments has either been temporary in nature, or its use has been restricted to the experimenters. Mounting such an infrastructure is often beyond the means of many interested in this area of research. So we feel that the routing research community would benefit from a permanent and public infrastructure for such active routing probes. We use the term *BGP Beacon* to refer to a publicly documented prefix having global visibility and a published schedule for announcements and withdrawals. A BGP Beacon is to be used for the ongoing study of BGP dynamics, and should be supported with a long term commitment. We describe two collections of BGP Beacons that have been set up at various points in the Internet. We then describe techniques for processing BGP updates when a BGP Beacon is observed from a BGP monitoring point such as Route Views or RIPE. *Anyone* could get data from *any* public or private route monitor to study the Beacon dynamics, as the Beacon updates are globally visible.

We illustrate the use of BGP Beacons with four case studies. Each study relies on the fact that we are monitoring updates that have been generated by a Beacon event. First, we consider the impact of different implementations of BGP on the observations. Second, we investigate the potential that route flap damping [11] punishes "well behaved" routes. Simulation results in [12] have shown that it can punish "well behaved" as well as "misbehaving" routes. Here we use the BGP Beacons to validate those results in the global Internet. This is the first study of the impact of flap damping using real data from the Internet. Even though the BGP Beacons have a fairly long cycle (two hours between each announce or withdraw event), we see that *even announcements* can potentially trigger flap damping as much as 10 percent of the time at some locations on the Internet. For our third study, we present a novel analysis of the inter-arrival times of updates generated by BGP Beacons. Finally, we revisit the convergence-time issues studied in [9, 10].

Literature: BGP Zombies

R. Fontugne, E. Bautista, C. Petrie, Y. Nomura, P. Abry, P. Goncalves, K. Fukuda, E. Aben. "BGP Zombies: an Analysis of Beacons Stuck Routes", In *Proceedings of PAM'19*. LNCS vol. 11419, Springer. 2019. https://doi.org/10.1007/978-3-030-15986-3_13



BGP Zombies: an Analysis of Beacons Stuck Routes

Romain Fontugne[1], Esteban Bautista[2], Colin Petrie[3], Yutaro Nomura[4], Patrice Abry[5,6], Paulo Goncalves[2], Kensuke Fukuda[6], and Emile Aben[3]

[1] IIJ Research Lab, Tokyo, Japan romain@iij.ad.jp
[2] Univ Lyon, Ens de Lyon, Inria, CNRS, UCB Lyon 1, F-69342, Lyon, France
[3] RIPE NCC, Amsterdam, Netherlands
[4] The University of Tokyo, Tokyo, Japan
[5] Univ Lyon, Ens de Lyon, Univ Claude Bernard, CNRS, Laboratoire de Physique, Lyon, France
[6] NII / Sokendai, Tokyo, Japan

Abstract. Network operators use the Border Gateway Protocol (BGP) to control the global visibility of their networks. When withdrawing an IP prefix from the Internet, an origin network sends BGP withdraw messages, which are expected to propagate to all BGP routers that hold an entry for that IP prefix in their routing table. Yet network operators occasionally report issues where routers maintain routes to IP prefixes withdrawn by their origin network. We refer to this problem as BGP zombies and characterize their appearance using RIS BGP beacons, a set of prefixes withdrawn every four hours.Across the 27 monitored beacon prefixes, we observe usually more than one zombie outbreak per day. But their presence is highly volatile, on average a monitored peer misses 1.8% withdraws for an IPv4 beacon (2.7% for IPv6). We also discovered that BGP zombies can propagate to other ASes, for example, zombies in a transit network are inevitably affecting its customer networks. We employ a graph-based semi-supervised machine learning technique to estimate the scope of zombies propagation, and found that most of the observed zombie outbreaks are small (i.e. on average 10% of monitored ASes for IPv4 and 17% for IPv6). We also report some large zombie outbreaks with almost all monitored ASes affected.

Inferring the Hidden

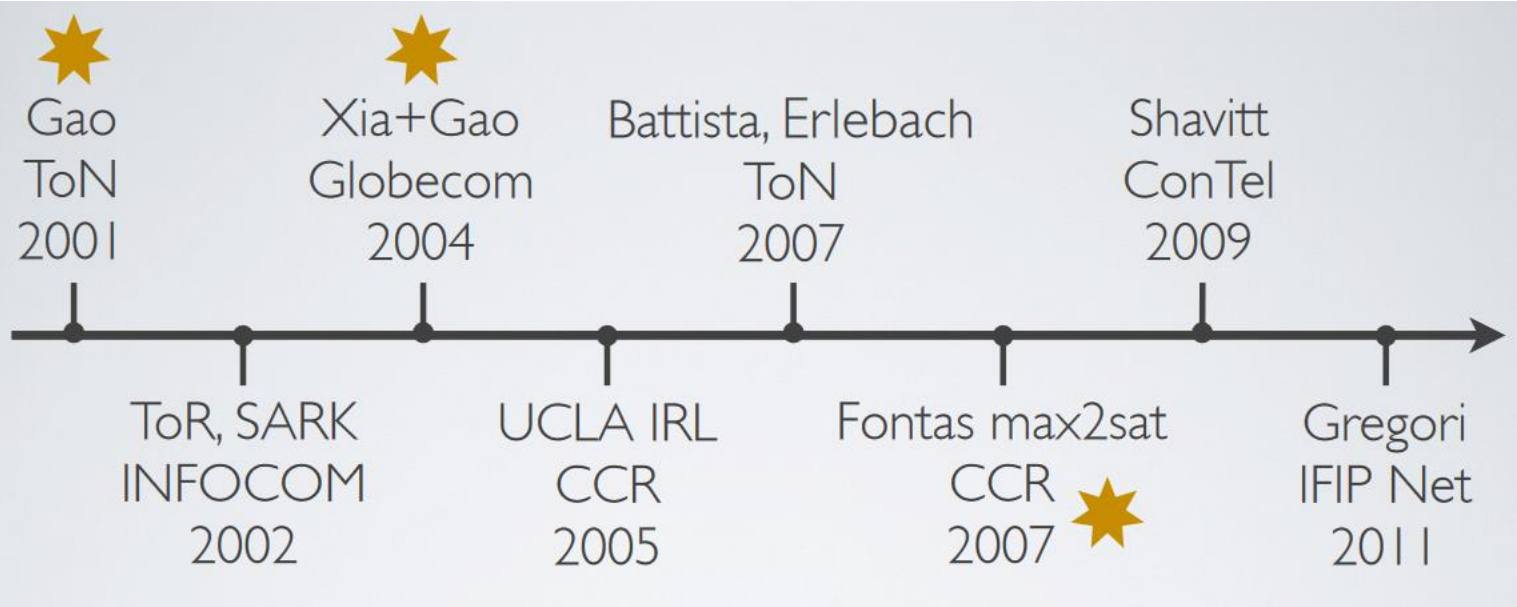
AS RELATIONSHIPS AND CUSTOMER CONES

Business Relations and BGP

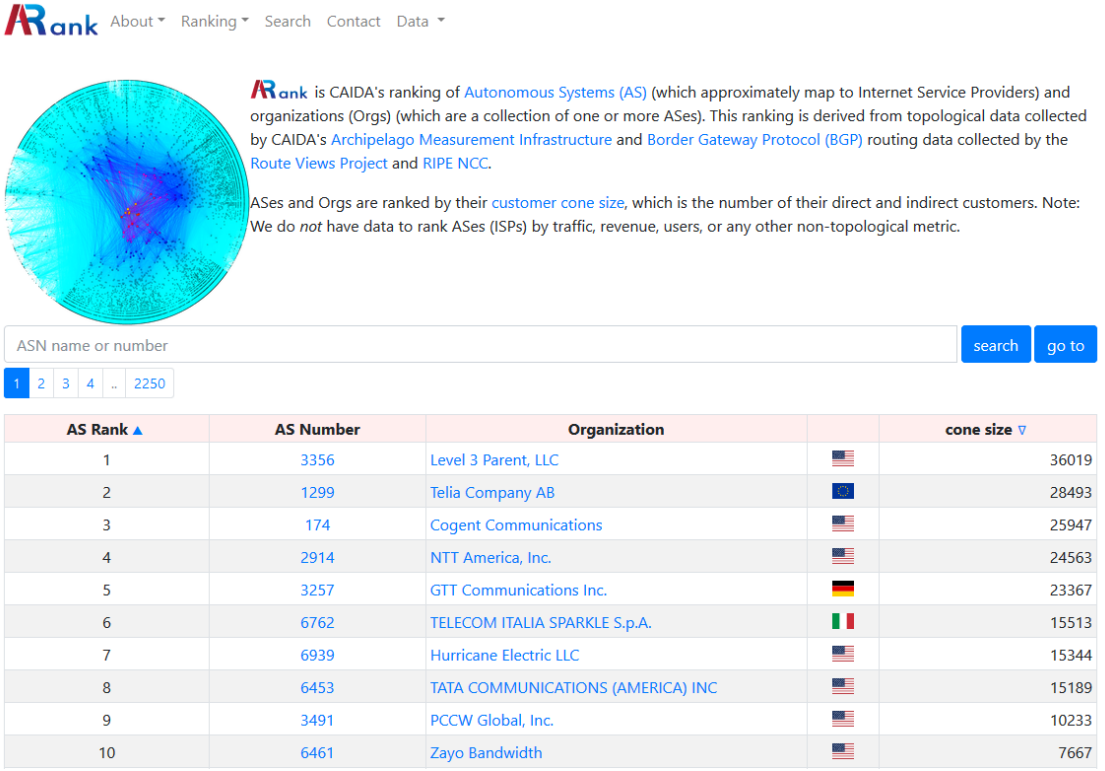
How to rank an autonomous systems?

How to infer hidden AS relationships?

Long-term research



CAIDA AS rank



AS Rank ▲	AS Number	Organization		cone size ▽
1	3356	Level 3 Parent, LLC		36019
2	1299	Telia Company AB		28493
3	174	Cogent Communications		25947
4	2914	NTT America, Inc.		24563
5	3257	GTT Communications Inc.		23367
6	6762	TELECOM ITALIA SPARKLE S.p.A.		15513
7	6939	Hurricane Electric LLC		15344
8	6453	TATA COMMUNICATIONS (AMERICA) INC		15189
9	3491	PCCW Global, Inc.		10233
10	6461	Zayo Bandwidth		7667

http://as-rank.caida.org

What can we extract from RIB dumps?

Data sources to infer AS relationships

IANA List of alloc. ASNs

RPSL

BGP communities

Route Collectors

Directly reported

Data sources to infer AS relationships

IANA List of alloc. ASNs

RPSL

BGP communities

Route Collectors

Directly reported

IANA list of allocated ASNs

https://www.iana.org/assignments/as-numbers/as-numbers.xhtml

Allows to identify valid AS numbers assigned to organizations and RIR

Preferably, you still know about sub-assignments

Routing Policy Specification Language (RPSL)

List what to import from peers

List what to export to peers

Availability of data depends on the region

RIPE Whois data is largest source

e.g., because European IXPs require operators to register policies

Routing Policy Specification Language (RPSL)

```
aut-num: AS39063
import:  from AS3320 accept ANY
import:  from AS174 accept ANY
import:  from AS1299 accept ANY
import:  from AS9002 accept ANY
export:  to AS3320 announce AS39063
export:  to AS174 announce AS39063
export:  to AS1299 announce AS39063
export:  to AS9002 announce AS39063
```

List what to import from peers

List what to export to peers

Availability of data depends on the region

RIPE Whois data is largest source

e.g., because European IXPs require operators to register policies

You need to resolve to route objects (prefixes)

Derive C2P relationships based on RPSL

`ANY` in import and export rules indicates customer/provider relationships

If X has rule that imports ANY from Y then

If Y exports ANY to X then

Y = provider, X = customer

BGP communities

Communities can be tagged to BGP routes

Common convention

Tagging AS places its ASN (or neighbor) in the first 16 bits, remaining 16 bits not well-defined but usually published on AS website

Idea

Build a dictionary of community attributes and policy meanings

BGP communities: Example

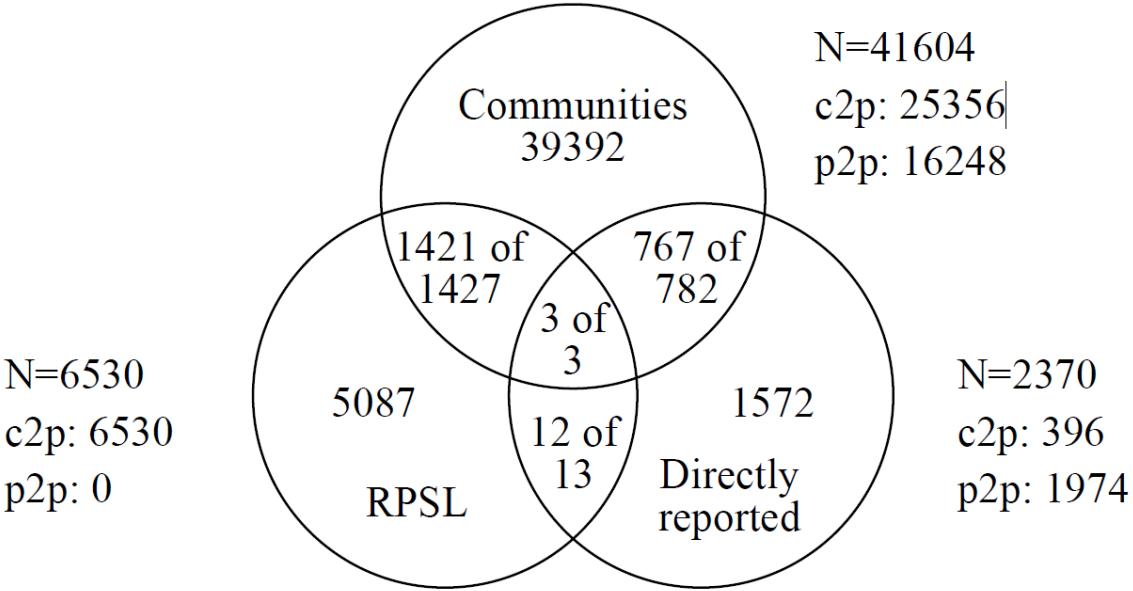
T · · Systems · · ·

IP Transit

AS3320 BGP Communities

Route Classification by Neighbor Type	
Community Value	Description
3320 : 9010	Imported from a customer
3320 : 9020	Imported from a peer
3320 : 9030	Imported from an upstream provider

What do we gain?



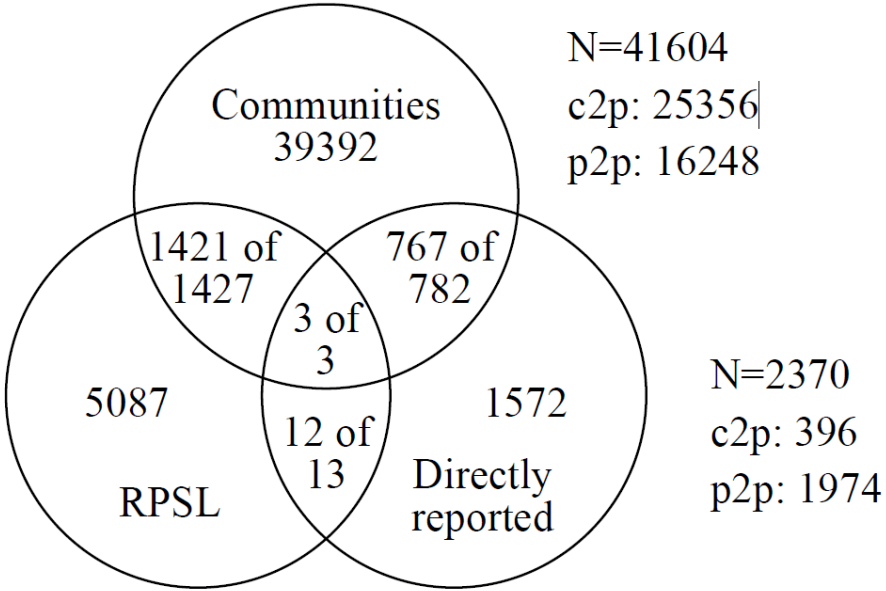
[Luckie et al., IMC 2013.]

What do we gain?

Reasons for discrepancies:

Some providers mistakenly import all routes from their customers; some customers mistakenly export all routes to their provider

Incorrect community tagging



[Luckie et al., IMC 2013.]

Deriving AS relationships from public BGP dumps (routing table dumps)

Assumptions (based on ISP discussions)

Multiple large transit providers form a peering mesh (Clique)

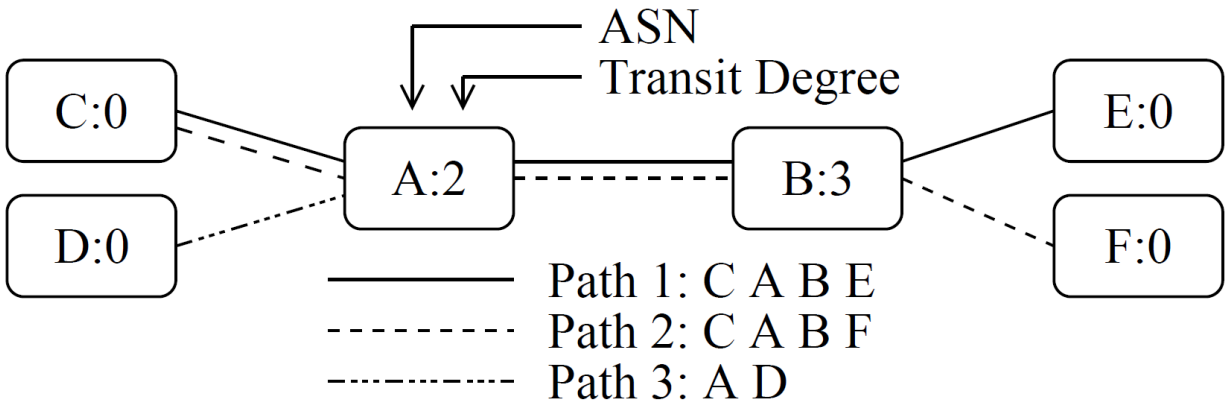
A provider will announce customer routes to its provider (except those that are in the clique)

AS topology is a directed acyclic graph (no cycles of p2c links)

Difference between transit degree and node degree

node degree is the number of neighbors an AS has

transit degree is the number of ASes that appear on either side of an AS in adjacent links



The following inference algorithm benefits from transit degree.

Initially sorting by transit degree reduces ordering errors of stub ASes with large peering visibility, i.e., stubs that provide a VP or peer with many VPs

High-level idea

Filter and sanitize AS paths

Initially sorting by transit degree reduces ordering errors of stub ASes with large peering visibility, i.e., stubs that provide a VP or peer with many VPs

Infer clique and resulting p2p mesh

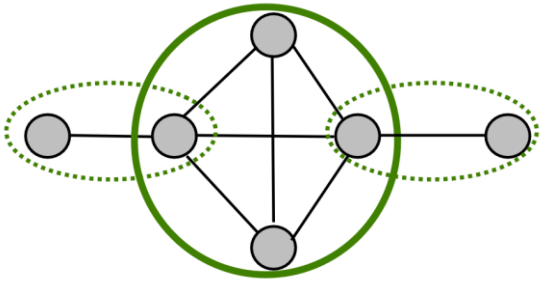
Infer providers, customers, and peers

Sanitizing

Remove paths that include loops (path poisoning) or unassigned ASes

Remove ASes that are used to operate IXP route servers

Inferring Clique: Background



Complete subgraph of a graph: part of a graph in which all nodes are connected to each other

Cliques: maximal complete subgraphs (not subsumed by any other complete subgraph)

Bron and Kerbosch (1973) algorithm allows to compute all cliques in linear time (relative to the number of cliques)

Inferring Clique

(1) Find max. clique (C1) involving the largest ten ASes by transit degree (start small)

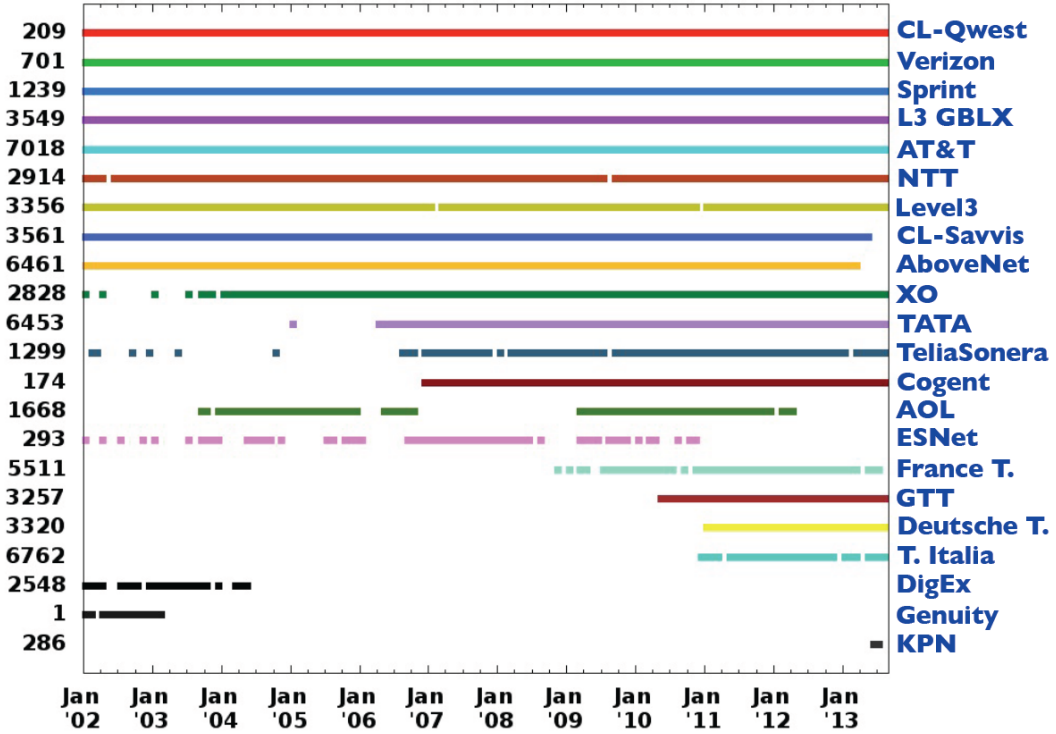
(2) Test every other AS to complete the clique

(3) If an AS would be admitted to the clique except for a single missing link, add to backup clique (C2)

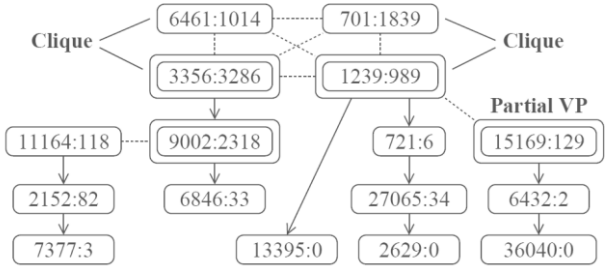
(4) Reapply Bron/Kerbosch to find largest clique (transit degree sum) from AS links involving both cliques C1 and C2

Clique members over time IPv4

Peering disputes and mergers of ASes can disrupt inference



Inferring providers, customers, and peers



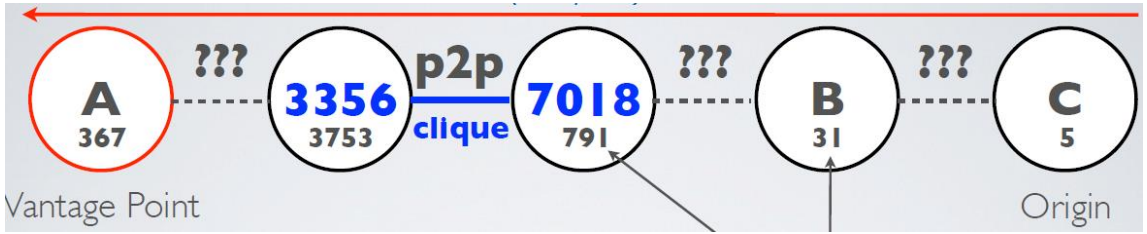
Objective

Be efficient, consider constrains necessary to infer c2p relationships, ignore non-hierarchical segments (p2p)

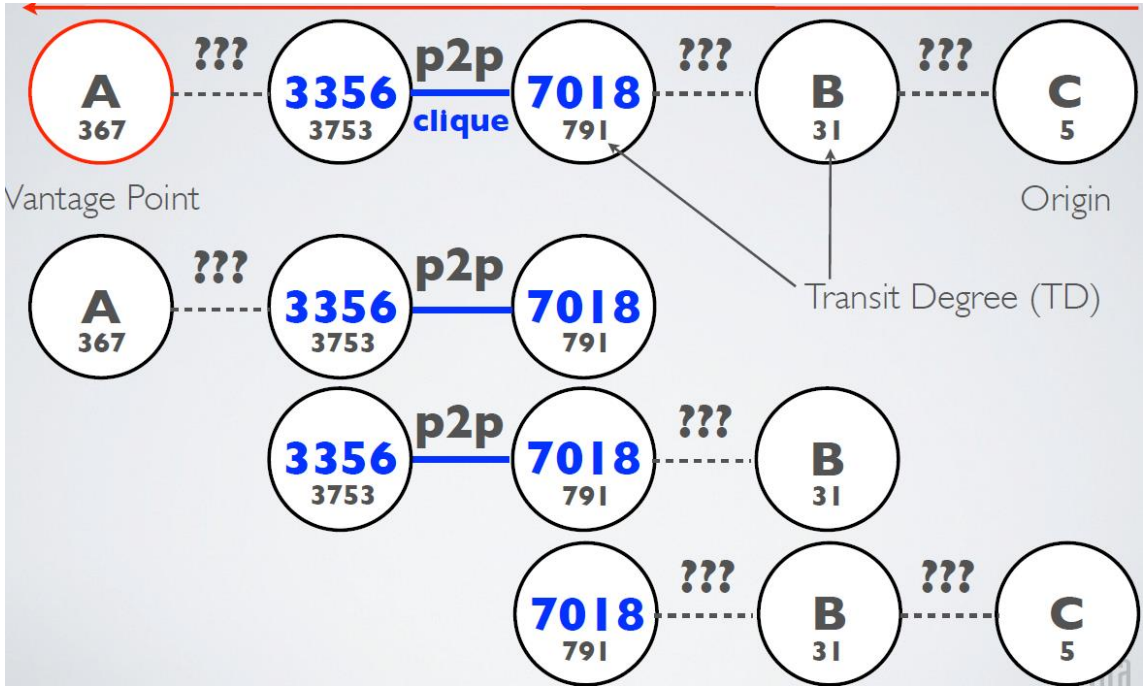
Solution

Process triplets instead of full paths

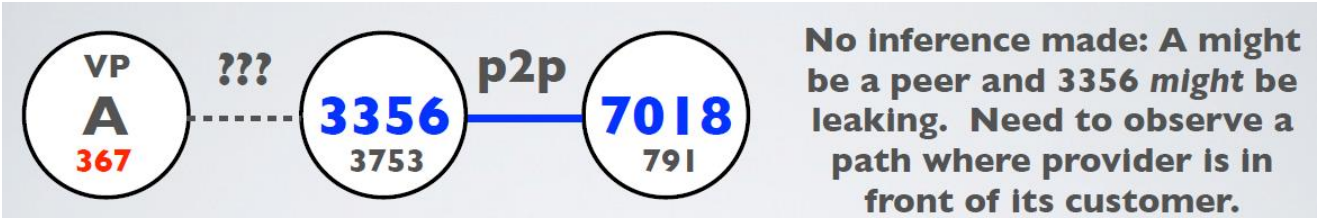
From paths to triplets



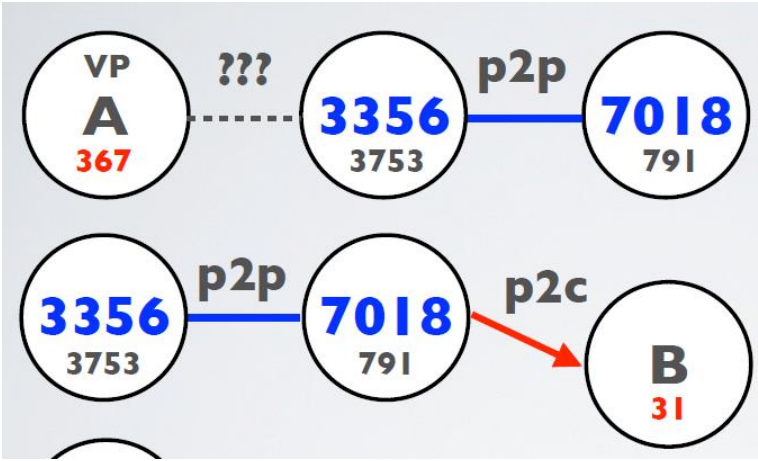
From paths to triplets



C2P inference, top-down



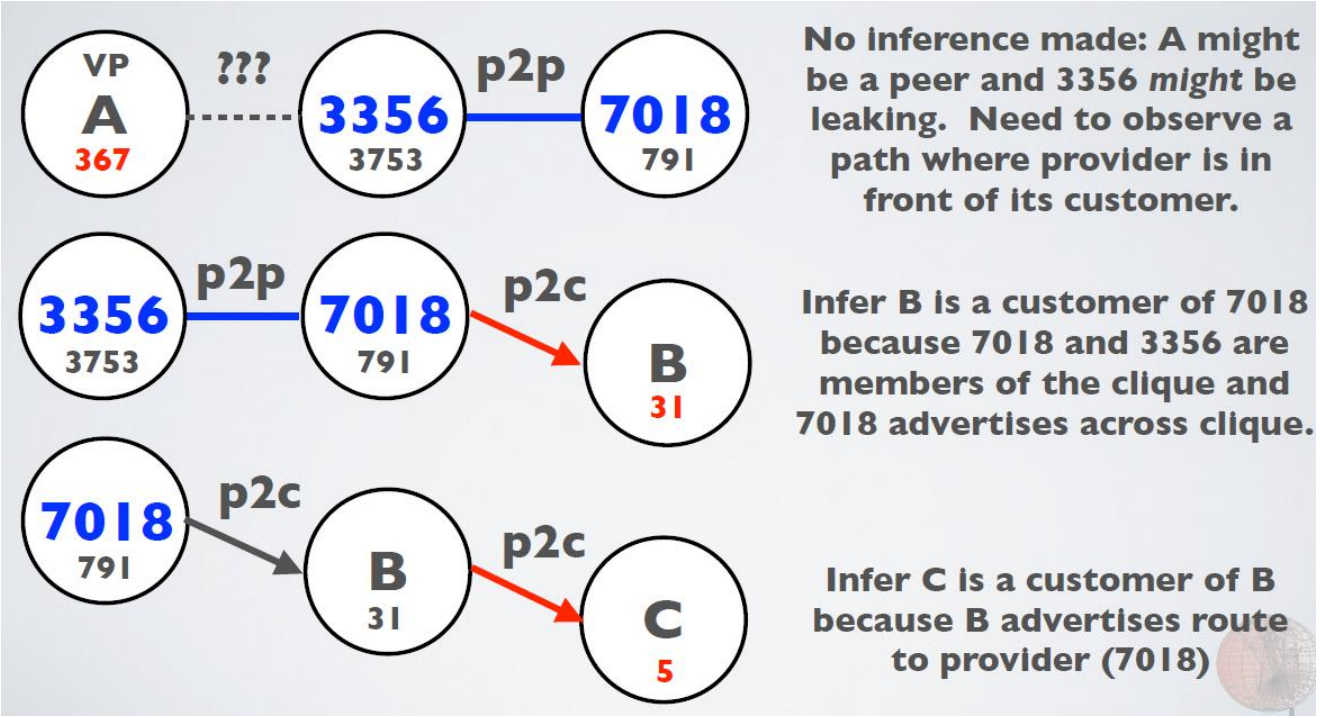
C2P inference, top-down



No inference made: A might be a peer and 3356 *might* be leaking. Need to observe a path where provider is in front of its customer.

Infer B is a customer of 7018 because 7018 and 3356 are members of the clique and 7018 advertises across clique.

C2P inference, top-down



Special cases need to be considered separately

Vantage points send only p2p routes to route collector

ASes with no providers

Stub clique

Adjacent links with no relationships

Complex relationships

Sibling Relationships and Mutual Transit
Indistinguishable from each other, poisoning, leaking.
No solution currently; as2org unreliable

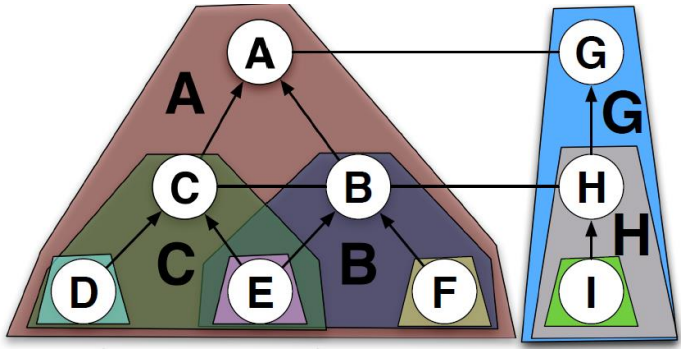
Partial Transit and Traffic Engineering
Handle in “customer cone”

Paid Peering
Unable to observe financial flows

Backup Transit
Rare in public BGP data. Mostly inferred as p2p.

Based on the relationship inference, we can create customer cones.

Customer cones



Set of ASes that an AS can reach through its customers.

Three methods to infer customer cones

Recursively inferred

BGP observed

Provider/peer observed

Recursively inferred

Visit recursively each AS reachable from p2c links, all customers would be part of the cone

Recursively inferred

Visit recursively each AS reachable from p2c links, all customers would be part of the cone

Problem: Assumes (unrealistically) that a provider receives all customer routes.

The error may affect the size of the customer cone.

BGP observed

C is included in A’s customer cone if we observe a BGP path where C is reached following a sequence of p2c links from A

BGP observed

C is included in A's customer cone if we observe a BGP path where C is reached following a sequence of p2c links from A

Advantage

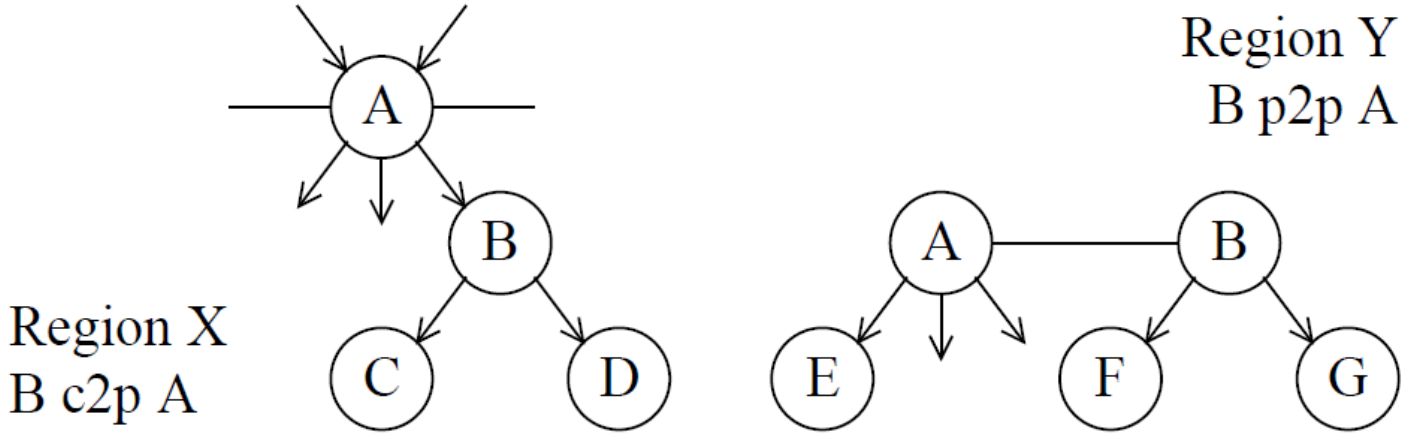
Doesn't include recursively sub-customers, of which prefixes have never been announced to the provider

Problems

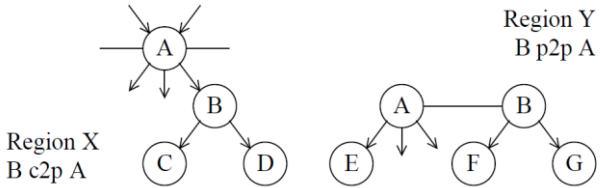
(1) Customer cones of ASes w/ hybrid relations will still include customers of peers

(2) Customer cones of ASes that provide a VP are more likely to be complete and therefore appear larger (measurement artifact)

Example of hybrid relationships



Provider/peer observed



Compute customer cone of A using routes observed from providers and peers of A

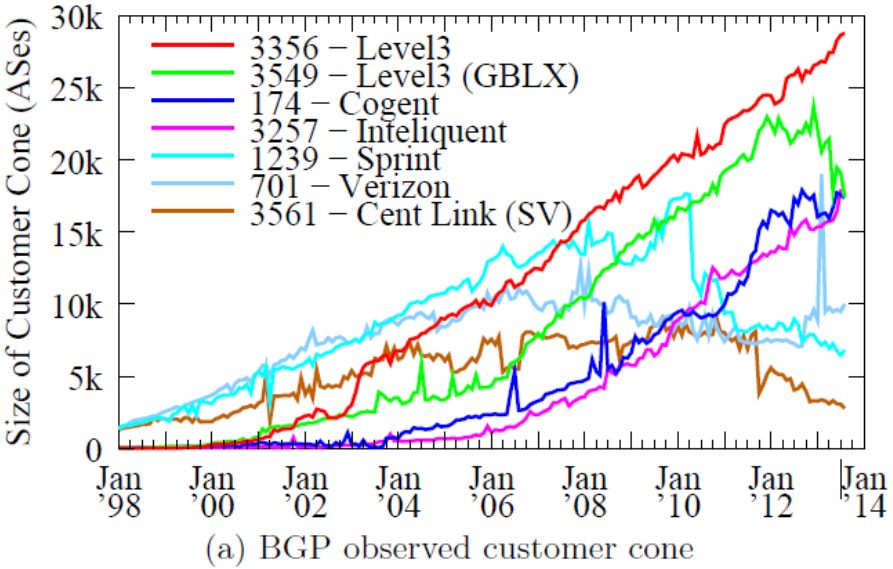
Advantage

(1) Does not include customers of B observed from the p2p portion in the customer cone of A.

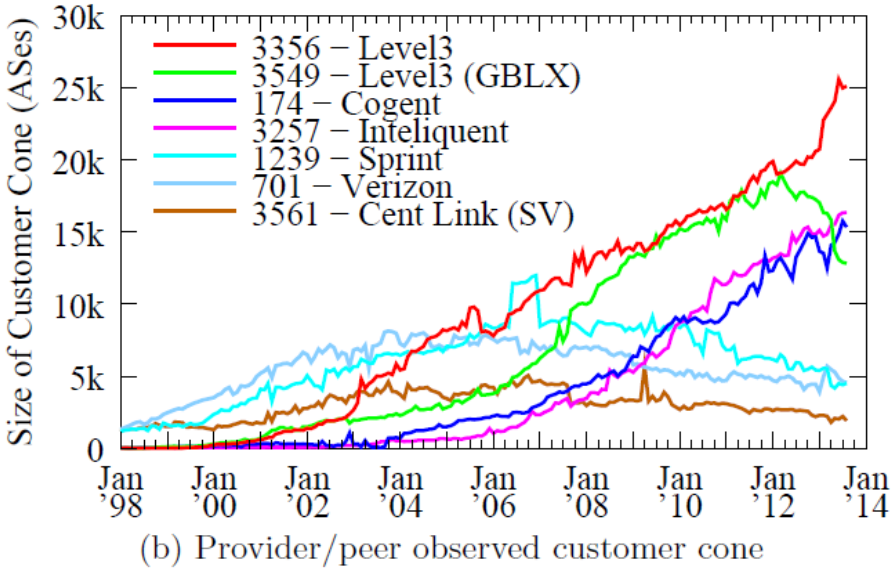
(2) Presence of VP set will not inflate A's customer cone.

Evaluation, April 2012

ASN	VP	PP Obs.	BGP Obs.	Recursive
3356	⋆	46.8 (1)	59.1 (1)	78.0 (1)
3549	⋆	45.2 (2)	54.2 (2)	72.3 (2)
3257	⋆	32.6 (3)	33.8 (5)	59.3 (5)
174		31.1 (4)	39.9 (4)	65.1 (3)
1299	⋆	29.3 (5)	40.0 (3)	64.6 (4)
2914	⋆	24.6 (6)	29.8 (6)	57.4 (6)
6453	⋆	18.9 (7)	28.1 (7)	55.8 (7)
6762	⋆	16.9 (8)	18.5 (9)	44.5 (11)
1239	⋆	15.2 (9)	21.0 (8)	51.0 (8)
3491		13.8 (10)	13.9 (12)	32.1 (13)
701	⋆	12.0 (11)	18.2 (10)	47.4 (9)
2828		11.3 (12)	11.4 (13)	45.7 (10)
7018	⋆	10.2 (13)	15.3 (11)	43.7 (12)
1273		8.4 (14)	8.4 (14)	26.7 (14)
6939	⋆	8.1 (15)	8.3 (15)	18.6 (15)

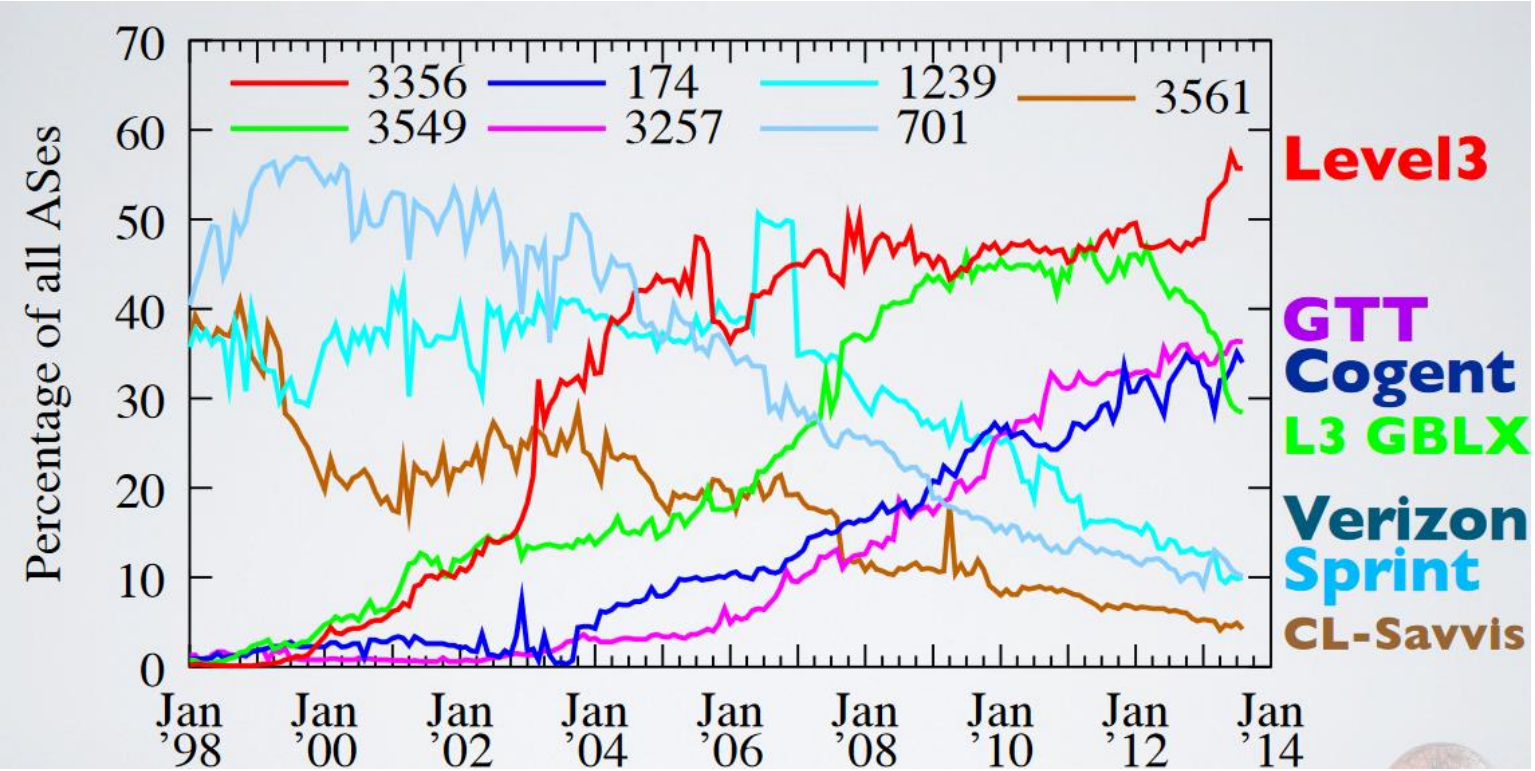


(a) BGP observed customer cone



(b) Provider/peer observed customer cone

Customer cones over time (Provide/peer method)



Literature

Matthew Luckie, Bradley Huffaker, Amogh Dhamdhere, Vasileios Giotsas, and kc claffy. 2013. AS relationships, customer cones, and validation. In *Proceedings of the 2013 conference on Internet measurement conference* (IMC '13). ACM, New York, NY, USA, 243-256. DOI: https://doi.org/10.1145/2504730.2504735

AS Relationships, Customer Cones, and Validation

Matthew Luckie
CAIDA / UC San Diego
mjl@caida.org

Bradley Huffaker
CAIDA / UC San Diego
bradley@caida.org

Amogh Dhamdhere
CAIDA / UC San Diego
amogh@caida.org

Vasileios Giotsas
University College London
V.Giotsas@cs.ucl.ac.uk

kc claffy
CAIDA / UC San Diego
kc@caida.org

ABSTRACT

Business relationships between ASes in the Internet are typically confidential, yet knowledge of them is essential to understand many aspects of Internet structure, performance, dynamics, and evolution. We present a new algorithm to infer these relationships using BGP paths. Unlike previous approaches, our algorithm does not assume the presence (or seek to maximize the number) of valley-free paths, instead relying on three assumptions about the Internet's inter-domain structure: (1) an AS enters into a provider relationship to become globally reachable; and (2) there exists a peering *clique* of ASes at the top of the hierarchy, and (3) there is no cycle of p2c links. We assemble the largest source of validation data for AS-relationship inferences to date, validating 34.6% of our 126,082 c2p and p2p inferences to be 99.6% and 98.7% accurate, respectively. Using these inferred relationships, we evaluate three algorithms for inferring each AS's *customer cone*, defined as the set of ASes an AS can reach using customer links. We demonstrate the utility of our algorithms for studying the rise and fall of large transit providers over the last fifteen years, including recent claims about the flattening of the AS-level topology and the decreasing influence of "tier-1" ASes on the global Internet.

Categories and Subject Descriptors

C.2.5 [**Local and Wide-Area Networks**]: Internet; C.2.1 [**Network Architecture and Design**]: Network topology

Keywords

AS relationships; routing policies; customer cones

1. INTRODUCTION

The Internet consists of thousands of independent interconnected organizations, each driven by their own business model and needs. The interplay of these needs influences, and sometimes determines, topology and traffic patterns, i.e., connectivity between networked organizations and routing across the resulting mesh. Understanding the underlying business relationships between networked organizations provides the strongest foundation for understanding many other aspects of Internet structure, dynamics, and evolution.

Business relationships between ASes, which are typically congruent with their routing relationships, can be broadly classified into two types: customer-to-provider (c2p) and peer-to-peer (p2p). In a c2p relationship, the customer pays the provider for traffic sent between the two ASes. In return, the customer gains access to the ASes the provider can reach, including those which the provider reaches through its own providers. In a p2p relationship, the peering ASes gain access to each others' customers, typically without either AS paying the other. Peering ASes have a financial incentive to engage in a *settlement-free* peering relationship if they would otherwise pay a provider to carry their traffic, and neither AS could convince the other to become a customer. Relationships are typically confidential so must be *inferred* from data that is available publicly. This paper presents a new approach to inferring relationships between ASes using publicly available BGP data.

Measurement and analysis of Internet AS topologies has been an active area of research for over a decade. While yielding insights into the structure and evolution of the topology, this line of research is constrained by systematic measurement and inference challenges [32], many of which our approach proactively addresses. First, the BGP-based collection infrastructure used to obtain AS-level topology data suffers from artifacts induced by misconfigurations, poisoned paths, and route leaks, all of which impede AS-relationship inference. Our algorithm incorporates steps to remove such artifacts. Second, AS topologies constructed from BGP data miss many peering links [6]. We show this lack of visibility does not hinder the accuracy of inferences on the links we do observe. Third, most AS-relationship algorithms rely on "valley-free" AS paths, an embedded assumption about the rationality of routing decisions that is not always valid [32], and which our algorithm does not make. Fourth, import and export filters can be complicated; some operators caveat their c2p links as being region or prefix specific. However, they still describe themselves as customers even if they do not receive full transit. Therefore, we make a c2p inference when any transit is observed between two ASes. We argue that relationship inferences can still be c2p or p2p with the caveat that the c2p relationship may be partial. We develop techniques to mitigate the effects of such *hybrid relationships* when computing an AS's customer cone, described later in

IMC'13, October 23–25, 2013, Barcelona, Spain.

http://dx.doi.org/10.1145/2504730.2504735.