“Deep Dive into BGP Communities”

Georgios Smaragdakis

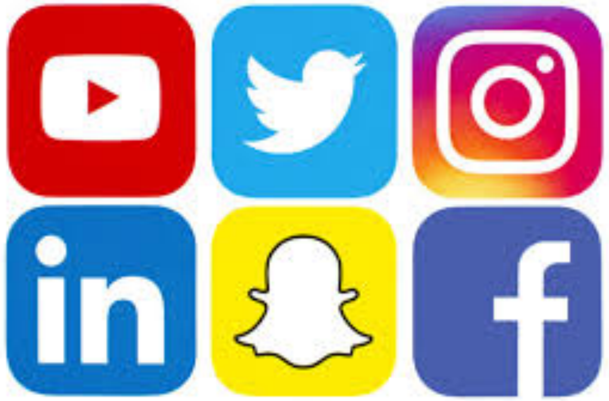
Joint work with Emile Aben, Arthur Berger, Robert Beverly, Randy Bush, Chris Dietzel, Anja Feldmann,
Vasileios Giotsas, Franziska Lichtblau, Cristel Pelsser, Philipp Richter, Florian Streibelt,
and many other colleagues!

MIT CSAIL

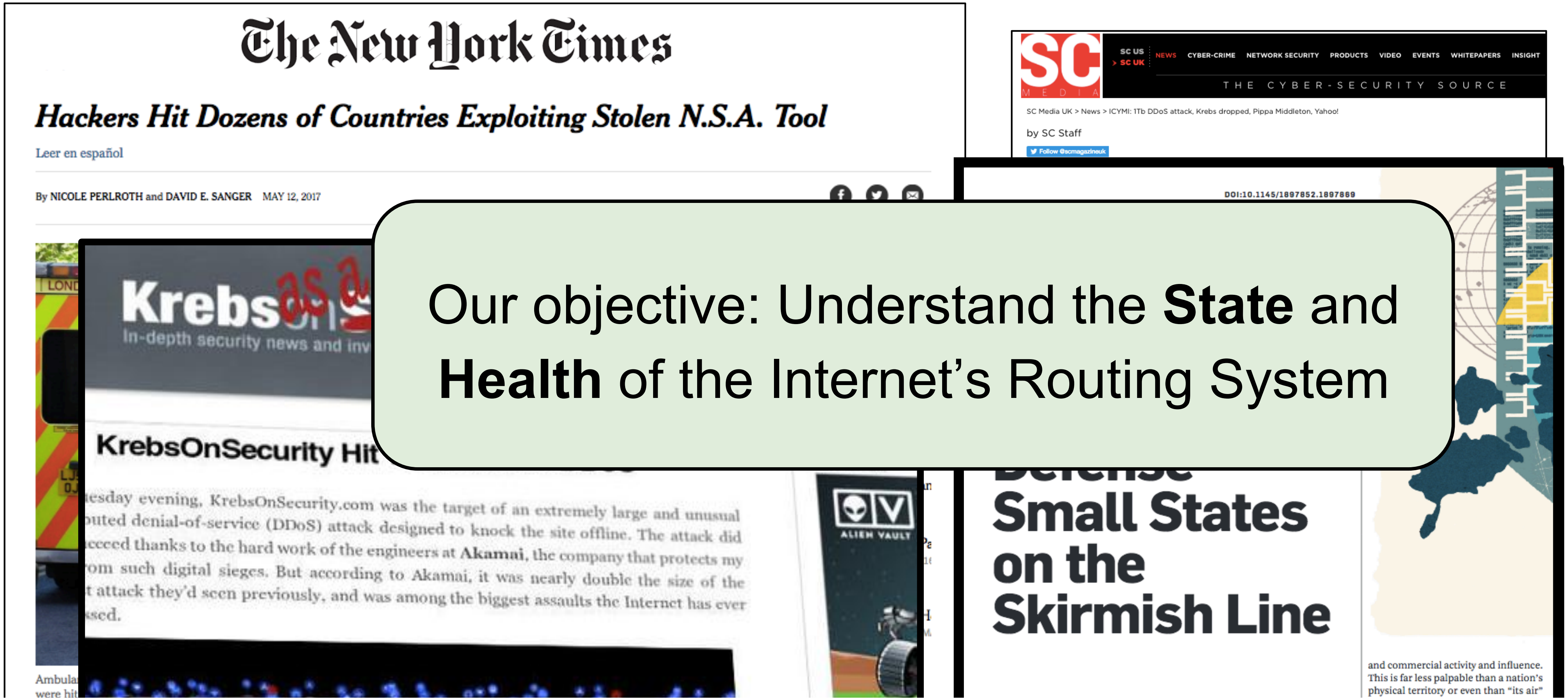
Technische Universität Berlin

European Research Council
Established by the European Commission
Supporting top researchers from anywhere in the world

The Internet is the Digital Backbone of our Civilization



Cyberattacks and Outages are Serious Threats



The New York Times

Hackers Hit Dozens of Countries Exploiting Stolen N.S.A. Tool

Leer en español

By NICOLE PERLROTH and DAVID E. SANGER MAY 12, 2017

Ambula
were hit

SC US
› SC UK
NEWS CYBER-CRIME NETWORK SECURITY PRODUCTS VIDEO EVENTS WHITEPAPERS INSIGHT

THE CYBER-SECURITY SOURCE

SC Media UK > News > ICYMI: 1Tb DDoS attack, Krebs dropped, Pippa Middleton, Yahoo!

by SC Staff

Follow @scmagazineuk

KrebsOnSecurity
In-depth security news and inv

KrebsOnSecurity Hit

esday evening, KrebsOnSecurity.com was the target of an extremely large and unusual
buted denial-of-service (DDoS) attack designed to knock the site offline. The attack did
cceed thanks to the hard work of the engineers at **Akamai**, the company that protects my
om such digital sieges. But according to Akamai, it was nearly double the size of the
t attack they'd seen previously, and was among the biggest assaults the Internet has ever
ssed.

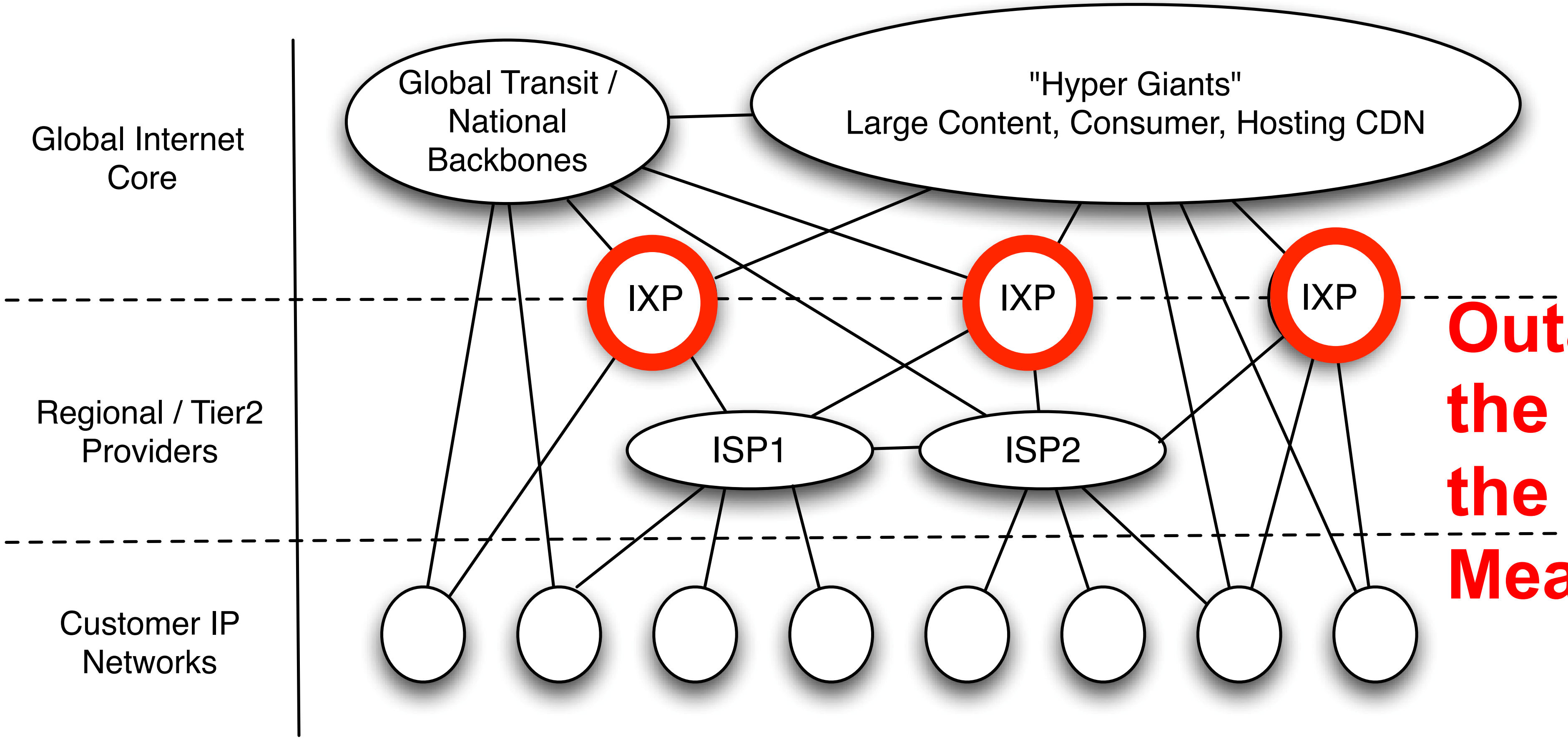
DOI:10.1145/1897852.1897869

Small States
on the
Skirmish Line

and commercial activity and influence.
This is far less palpable than a nation's
physical territory or even than "its air"

Our objective: Understand the **State** and **Health** of the Internet's Routing System

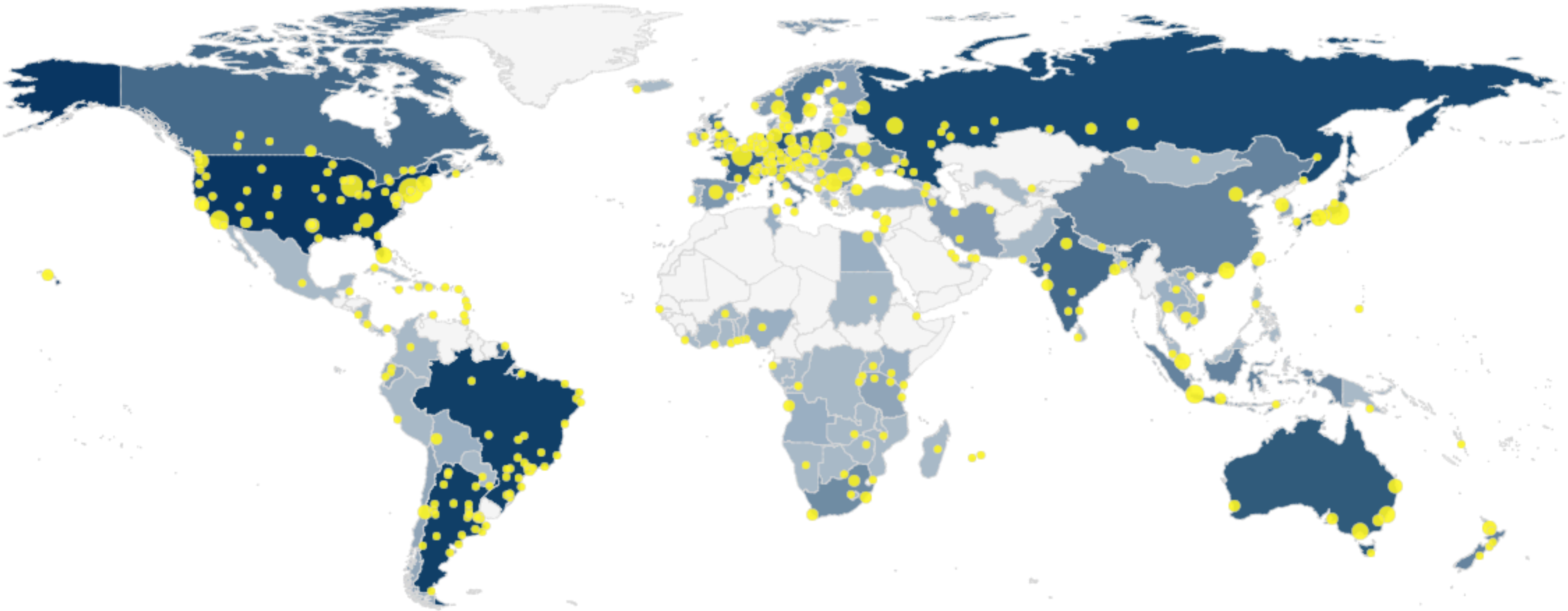
The New Internet


Global Internet Core
Regional / Tier2 Providers
Customer IP Networks
Global Transit / National Backbones
"Hyper Giants" Large Content, Consumer, Hosting CDN
IXP
IXP
IXP
ISP1
ISP2

Outages at the core of the Internet: Measured?

source: “Internet Interdomain Traffic”, Labovicz et al. SIGCOMM 2010

IXPs around the Globe



>300 active IXPs, **~125 Tbps** Traffic, **~2 Million** peerings

IXP is more than a Big Switch, it is an Ecosystem

LINX (London Internet Exchange) in Telehouse Colocation Facility (Telehouse North at Docklands)

1000s of cross-connects established in the datacenters

Peering Infrastructures are Critical Infrastructures

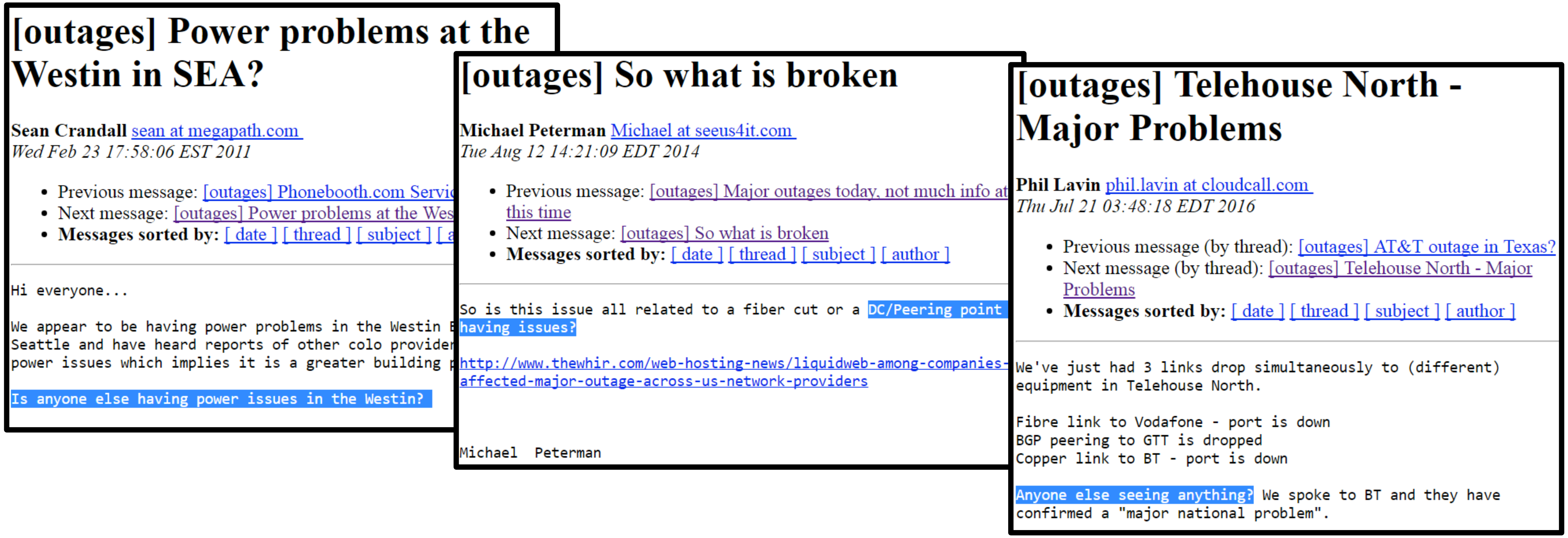
DHS and **ENISA** have characterized peering infrastructures as critical infrastructures – in the same category as nuclear reactors and power powerhouses. [An Annex to the National Infrastructure Protection Plan, 2010, 2015; Critical Infrastructures and Services, Internet Infrastructure: Internet Interconnections, 2010]

Internet Exchange Points: Typical SLA 99.99% (~52 min. downtime/year)[1]

Colocation facilities: Typical SLA 99.999% (~5 min. downtime/year)[2]

[1] https://ams-ix.net/services-pricing/service-level-agreement [2] http://www.telehouse.net/london-colocation/

Current practice: “Is anyone else having issues?”



[outages] Power problems at the Westin in SEA?

Sean Crandall sean at megapath.com
Wed Feb 23 17:58:06 EST 2011

- Previous message: [outages] Phonebooth.com Servi
- Next message: [outages] Power problems at the Wes
- **Messages sorted by:** [date] [thread] [subject] [a

```
Hi everyone...

We appear to be having power problems in the Westin B
Seattle and have heard reports of other colo provider
power issues which implies it is a greater building p

Is anyone else having power issues in the Westin?
```

[outages] So what is broken

Michael Peterman Michael at seeus4it.com
Tue Aug 12 14:21:09 EDT 2014

- Previous message: [outages] Major outages today, not much info at this time
- Next message: [outages] So what is broken
- **Messages sorted by:** [date] [thread] [subject] [author]

```
So is this issue all related to a fiber cut or a DC/Peering point
having issues?

http://www.thewhir.com/web-hosting-news/liquidweb-among-companies-
affected-major-outage-across-us-network-providers


Michael  Peterman
```

[outages] Telehouse North - Major Problems

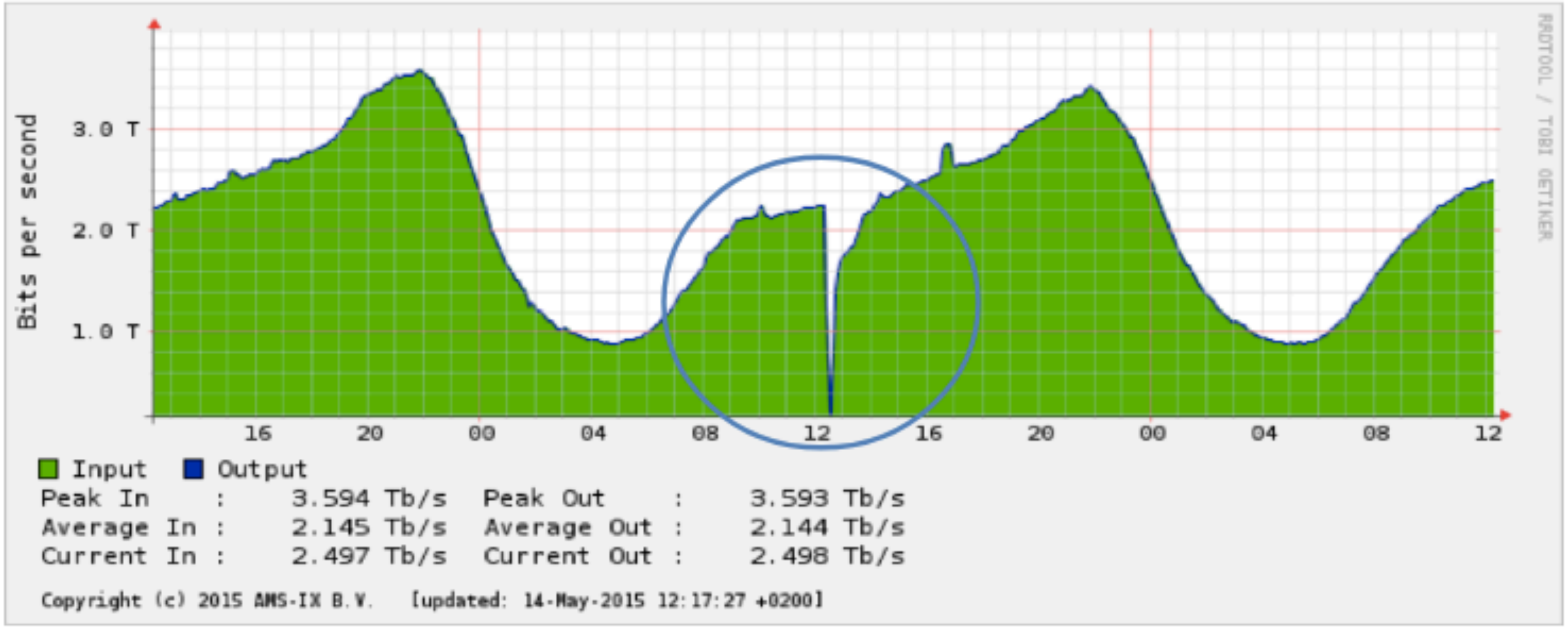
Phil Lavin phil.lavin at cloudcall.com
Thu Jul 21 03:48:18 EDT 2016

- Previous message (by thread): [outages] AT&T outage in Texas?
- Next message (by thread): [outages] Telehouse North - Major Problems
- **Messages sorted by:** [date] [thread] [subject] [author]

```
We've just had 3 links drop simultaneously to (different)
equipment in Telehouse North.

Fibre link to Vodafone - port is down
BGP peering to GTT is dropped
Copper link to BT - port is down

Anyone else seeing anything? We spoke to BT and they have
confirmed a "major national problem".
```

- ASes try to crowd-source the detection and localization of outages.
- Inadequate transparency/responsiveness from infrastructure operators.

The AMS-IX outage

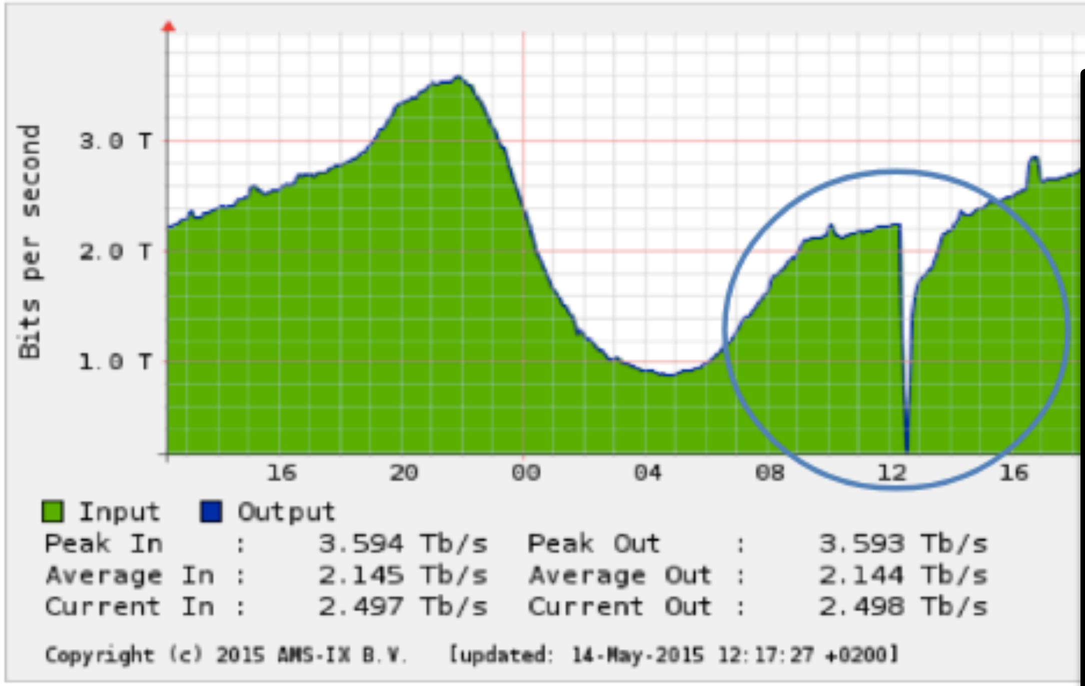


```
Peak In     :     3.594 Tb/s   Peak Out     :     3.593 Tb/s
Average In  :     2.145 Tb/s   Average Out  :     2.144 Tb/s
Current In  :     2.497 Tb/s   Current Out  :     2.498 Tb/s
```

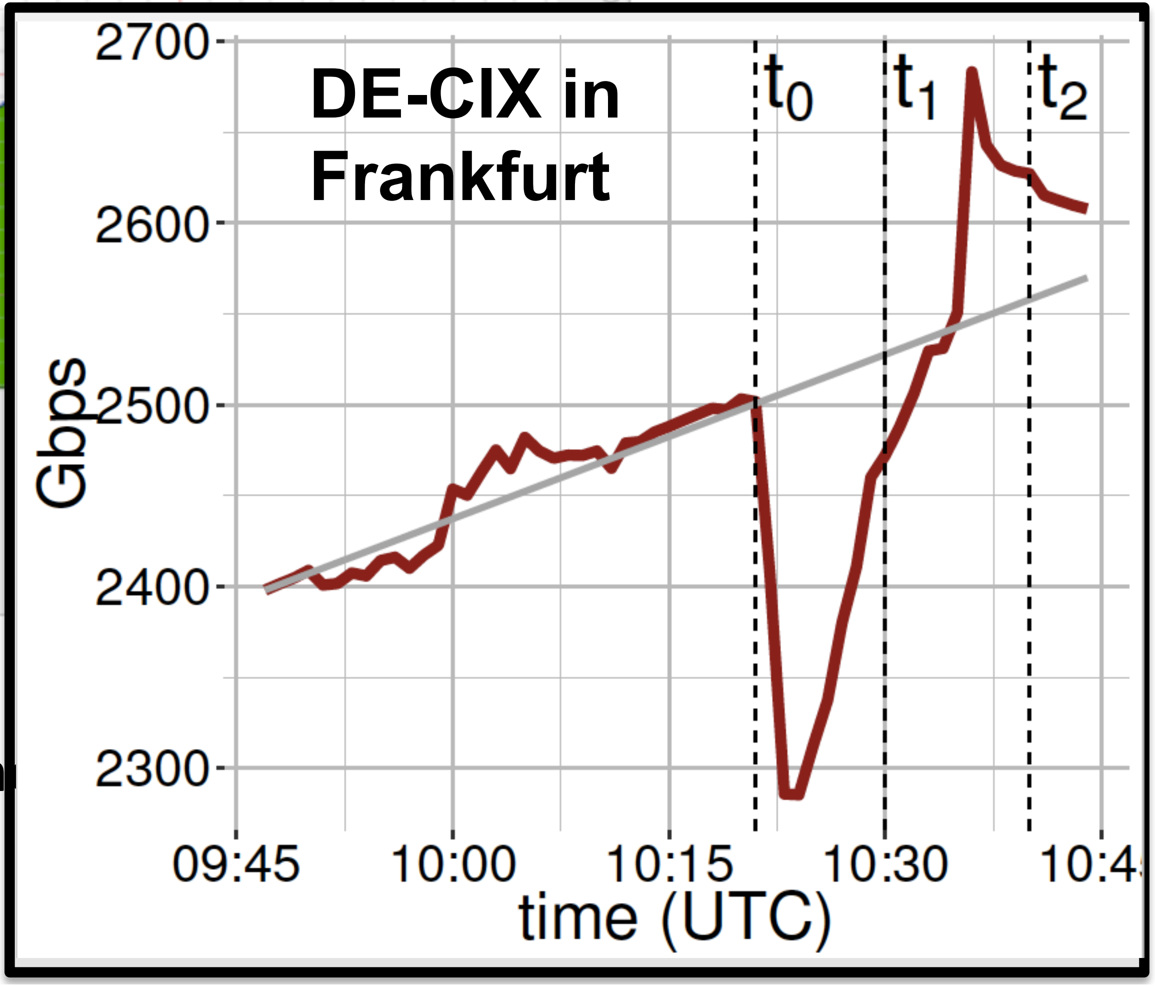
Copyright (c) 2015 AMS-IX B.V.   [updated: 14-May-2015 12:17:27 +0200]

Outage in AMS-IX, Amsterdam, The Netherlands on May 14, 2015

The AMS-IX outage

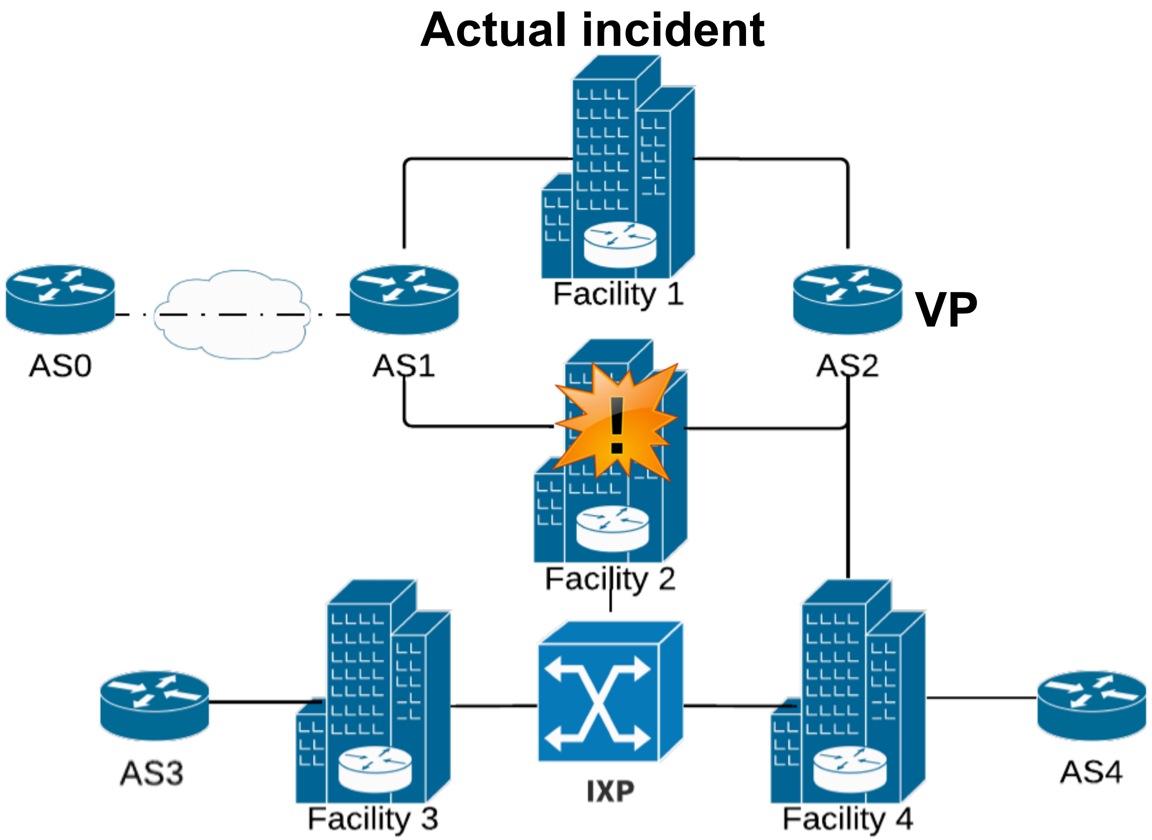


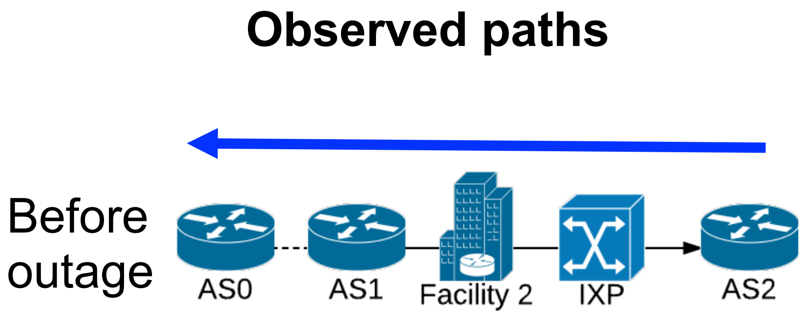
```
■ Input   ■ Output
Peak In      :    3.594 Tb/s   Peak Out     :    3.593 Tb/s
Average In   :    2.145 Tb/s   Average Out  :    2.144 Tb/s
Current In   :    2.497 Tb/s   Current Out  :    2.498 Tb/s

Copyright (c) 2015 AMS-IX B.V.   [updated: 14-May-2015 12:17:27 +0200]
```



Outage in AMS-IX, Amsterdam, Th

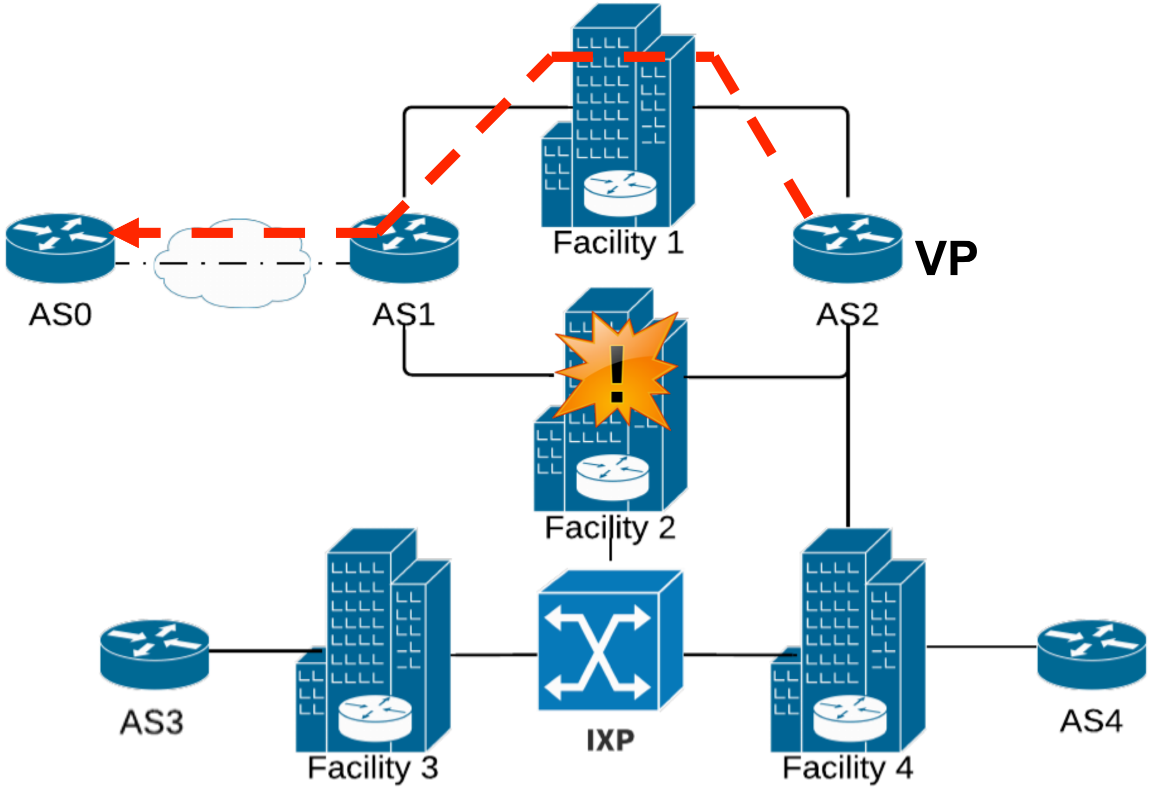
Challenges in detecting infrastructure outages



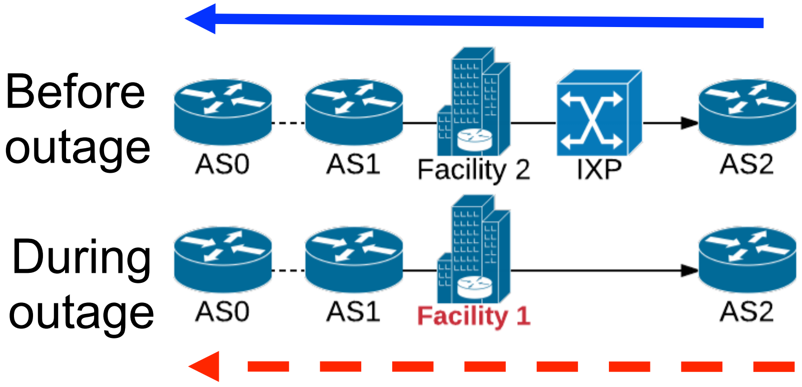


Challenges in detecting infrastructure outages

Actual incident



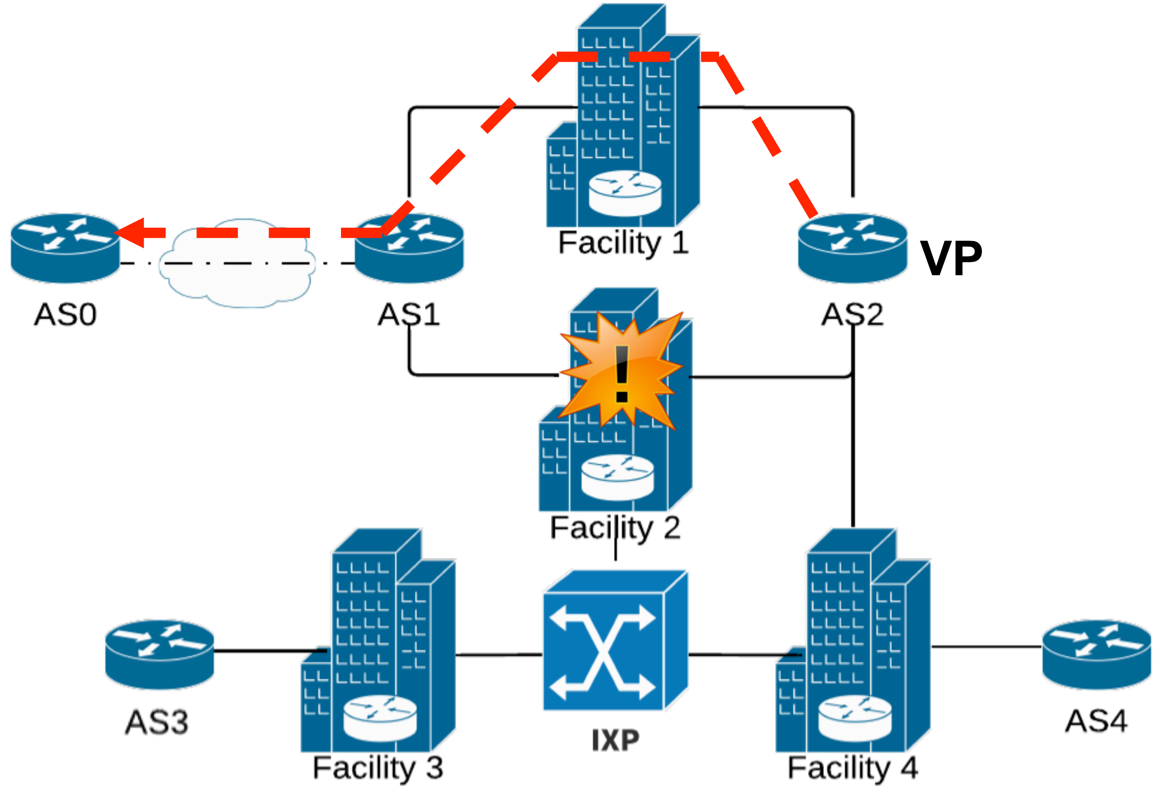
Observed paths



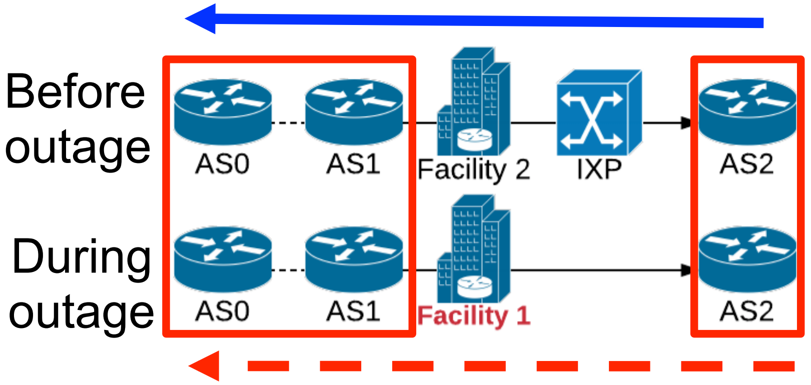
Challenges in detecting infrastructure outages

1. Capturing the infrastructure-level hops between ASes

Actual incident



Observed paths

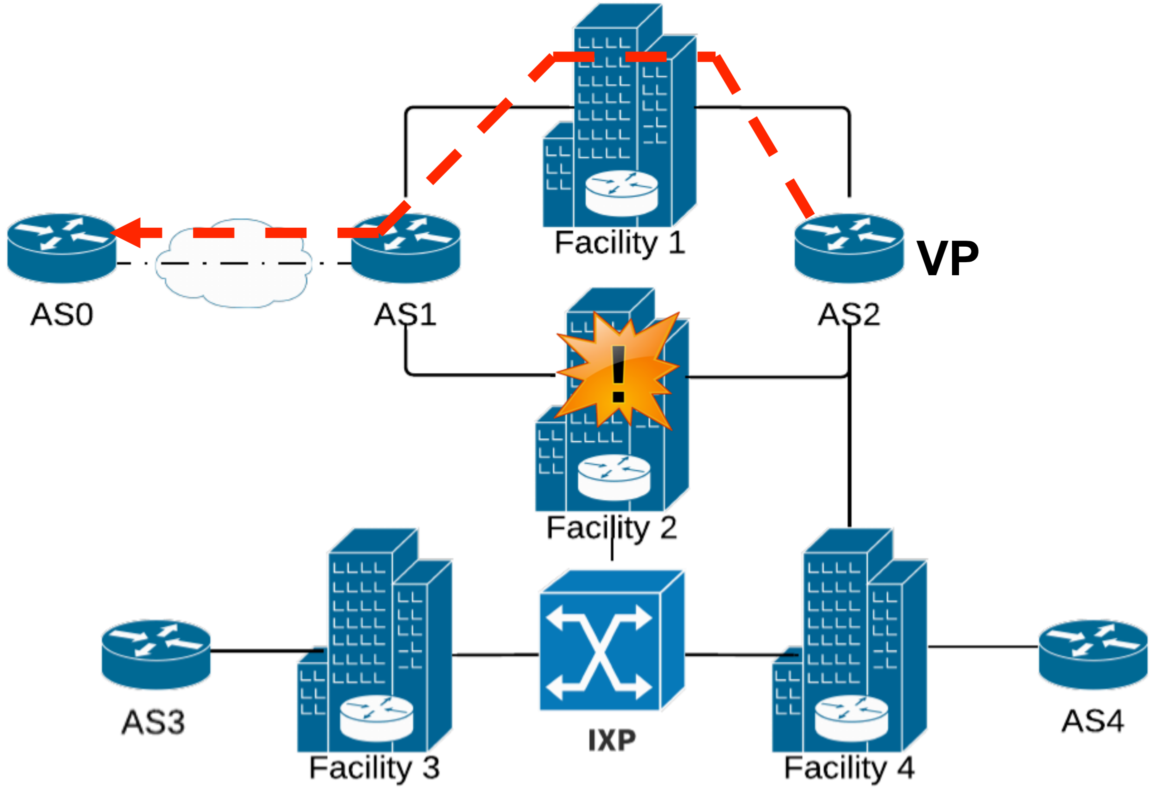


AS path does not change!

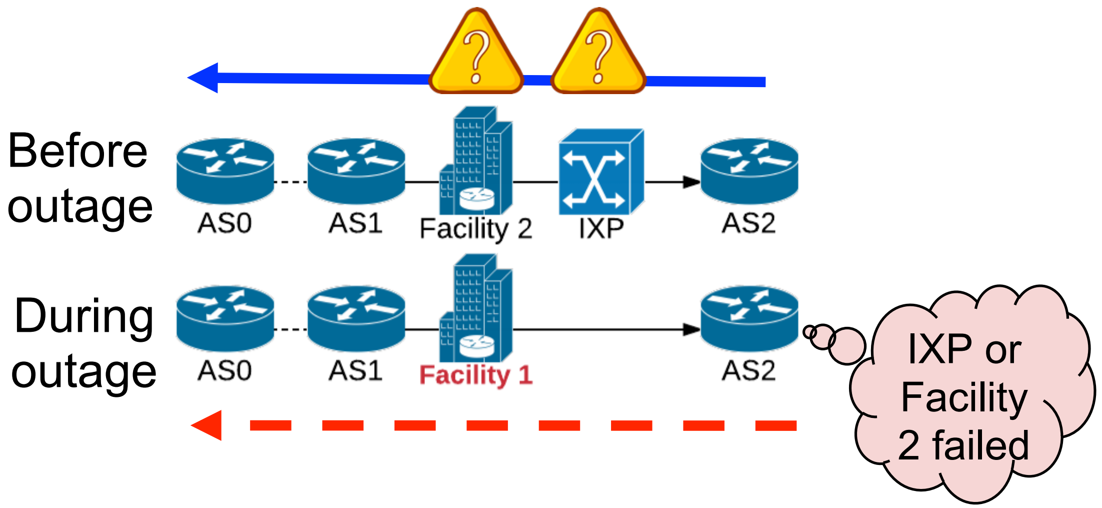
Challenges in detecting infrastructure outages

1. Capturing the infrastructure-level hops between ASes

Actual incident

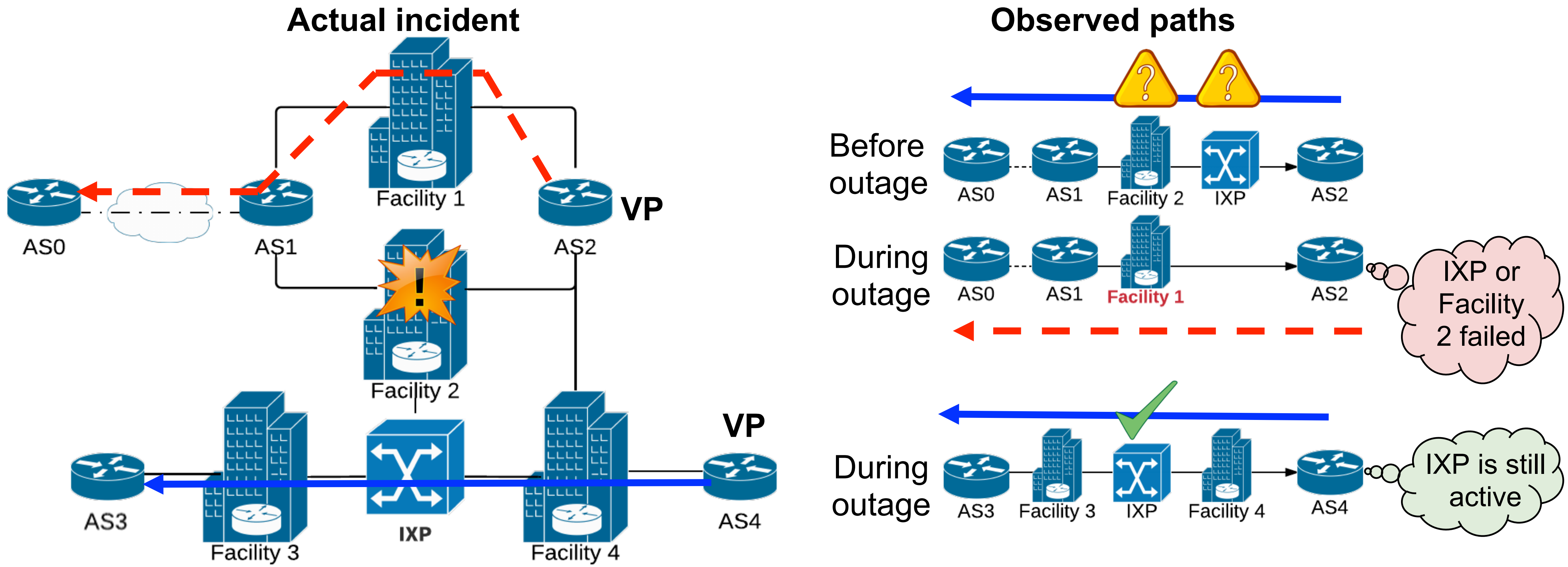


Observed paths



Challenges in detecting infrastructure outages

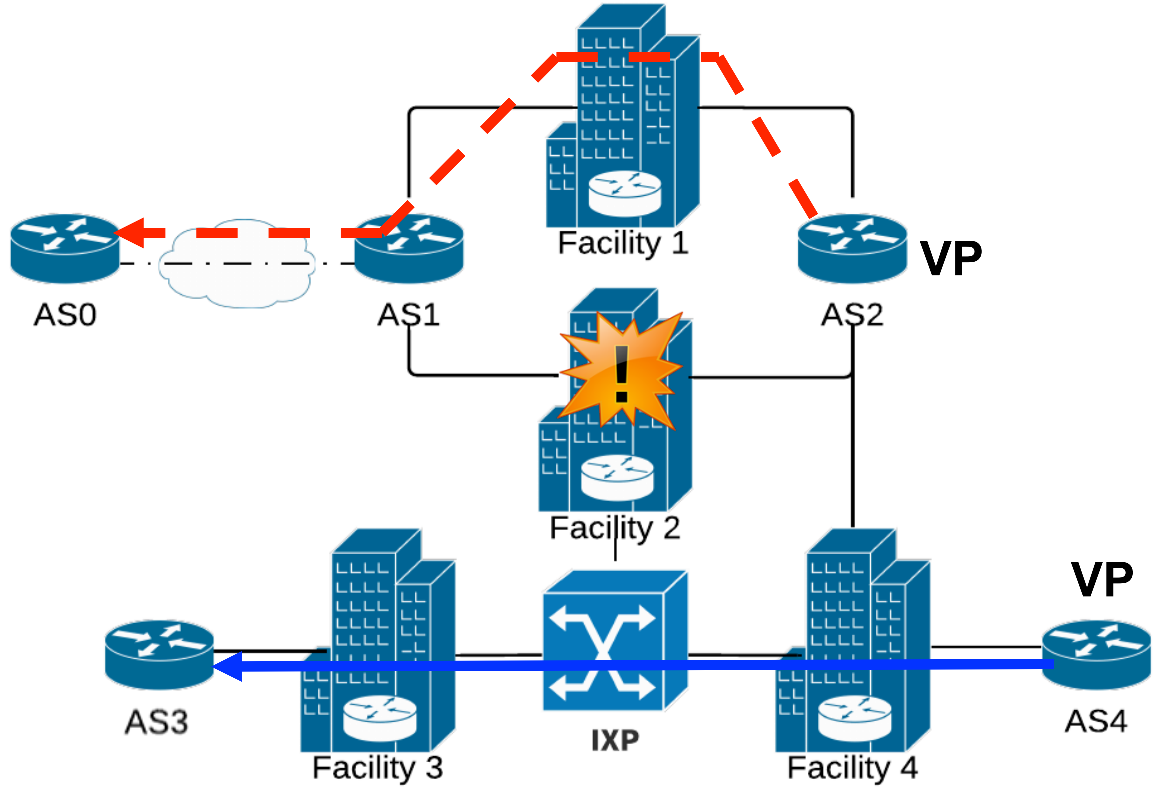
1. Capturing the infrastructure-level hops between ASes
2. Correlating the paths from multiple vantage points



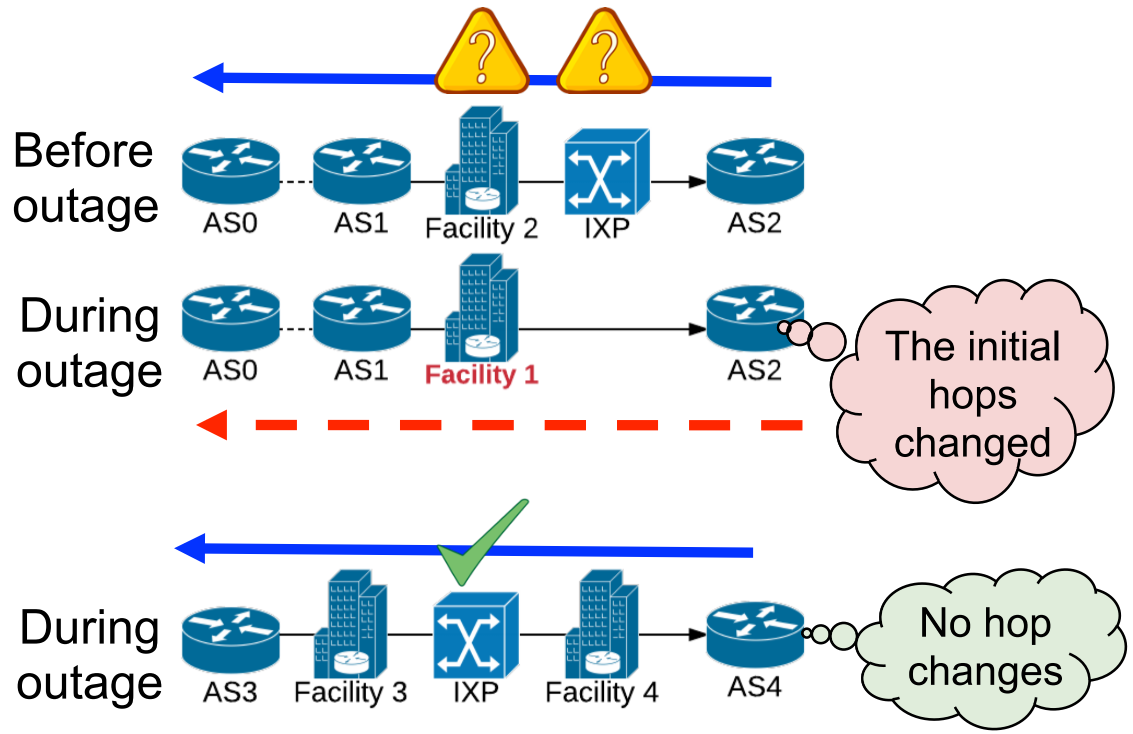
Challenges in detecting infrastructure outages

1. Capturing the infrastructure-level hops between ASes
2. Correlating the paths from multiple vantage points
3. Continuous monitoring of the routing system

Actual incident



Observed paths

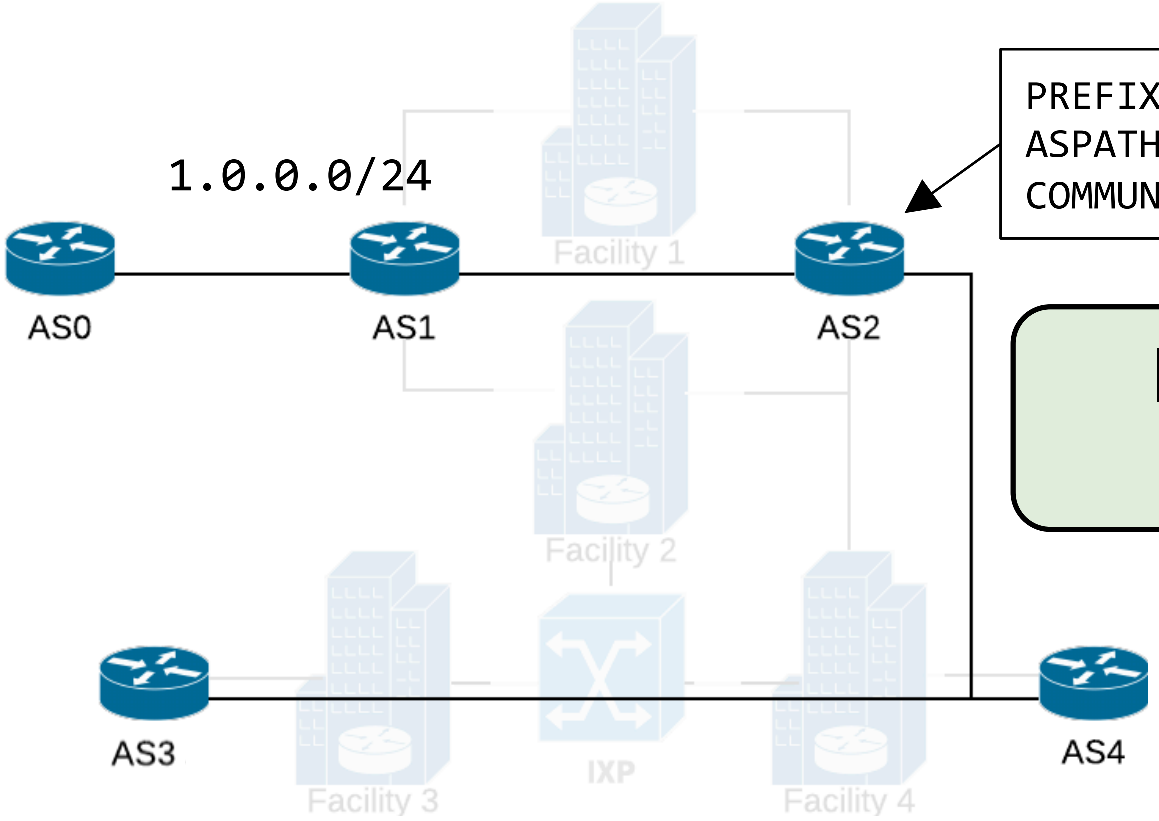


Challenges in detecting infrastructure outages

1. Capturing the infrastructure-level hops between ASes	✗ **BGP**	✓ **Traceroute**
2. Correlating the paths from multiple vantage points	✓ **BGP**	✗ **Traceroute**
3. Continuous monitoring of the routing system	✓ **BGP**	✗ **Traceroute**

Can we combine **BGP continuous passive** measurements with **fine-grained** topology discovery?

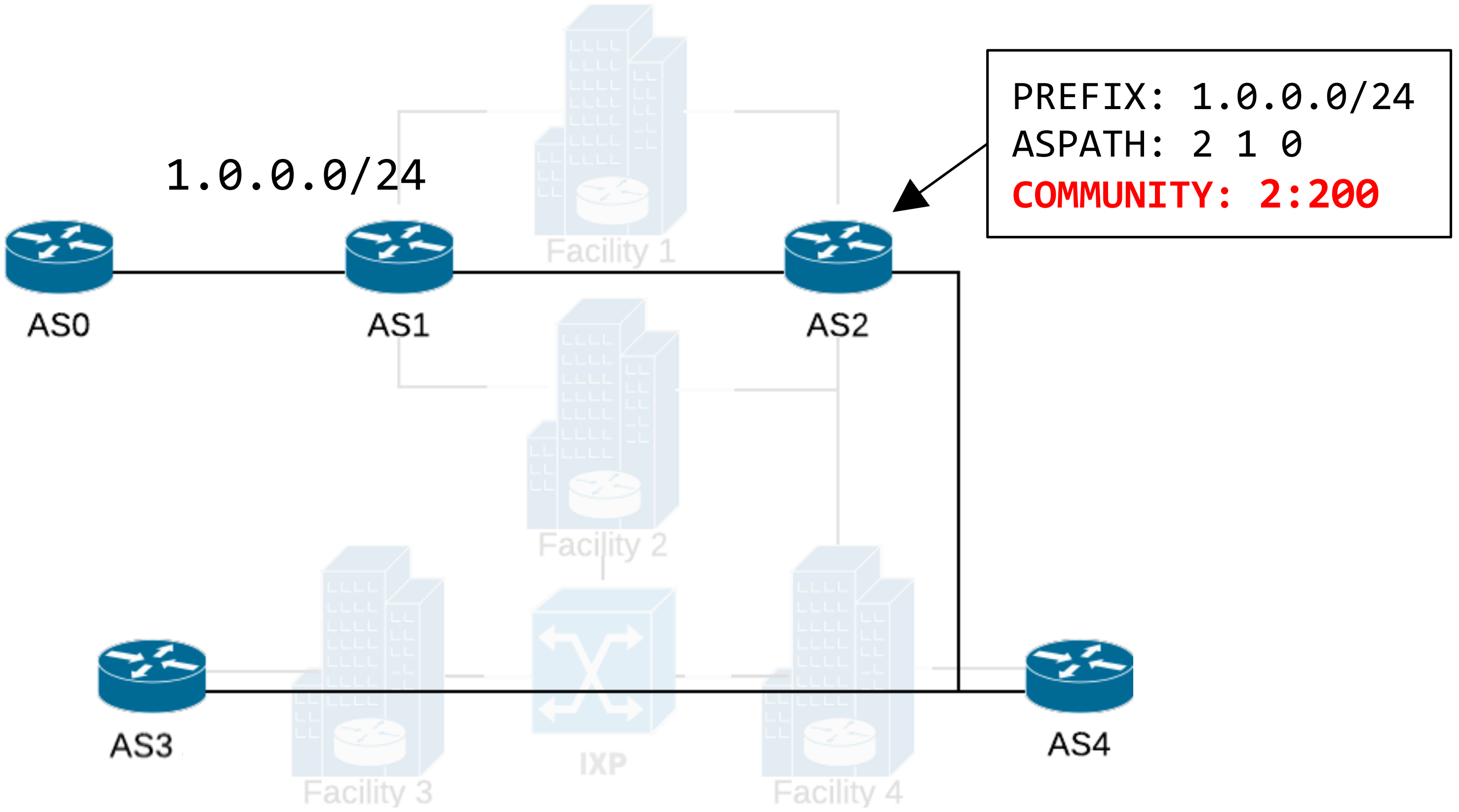
Deciphering location metadata in BGP



```
PREFIX: 1.0.0.0/24
ASPATH: 2 1 0
COMMUNITY: 2:200
```

Is BGP an information hiding protocol?

Deciphering location metadata in BGP

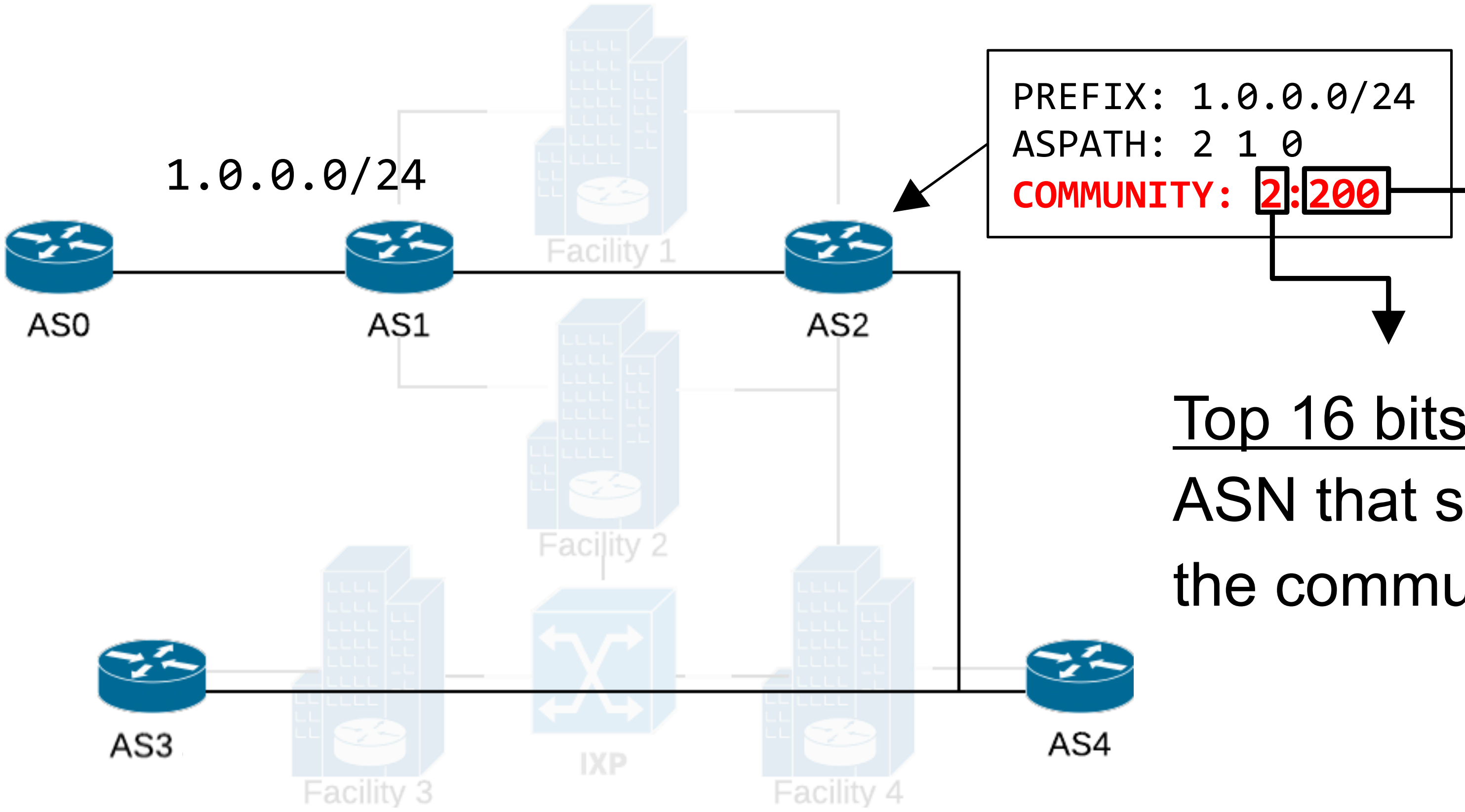


```
PREFIX: 1.0.0.0/24
ASPATH: 2 1 0
COMMUNITY: 2:200
```

BGP Communities:

- Optional attribute
- 32-bit numerical values
- Encodes **arbitrary** metadata

Deciphering location metadata in BGP

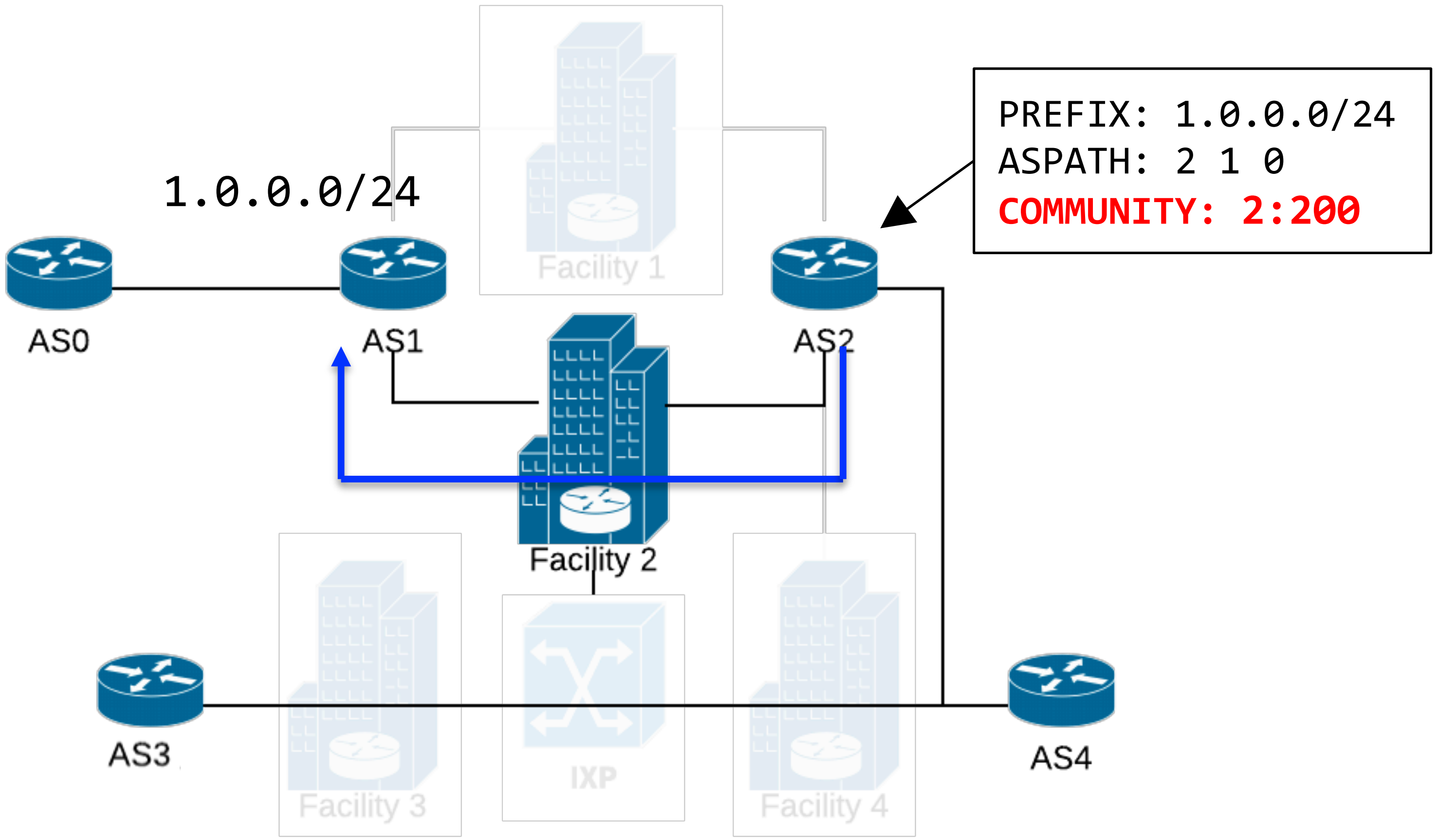


```
PREFIX: 1.0.0.0/24
ASPATH: 2 1 0
COMMUNITY: 2:200
```

Top 16 bits:
ASN that sets the community.

Bottom 16 bits:
Numerical value that encodes the actual meaning.

Deciphering location metadata in BGP

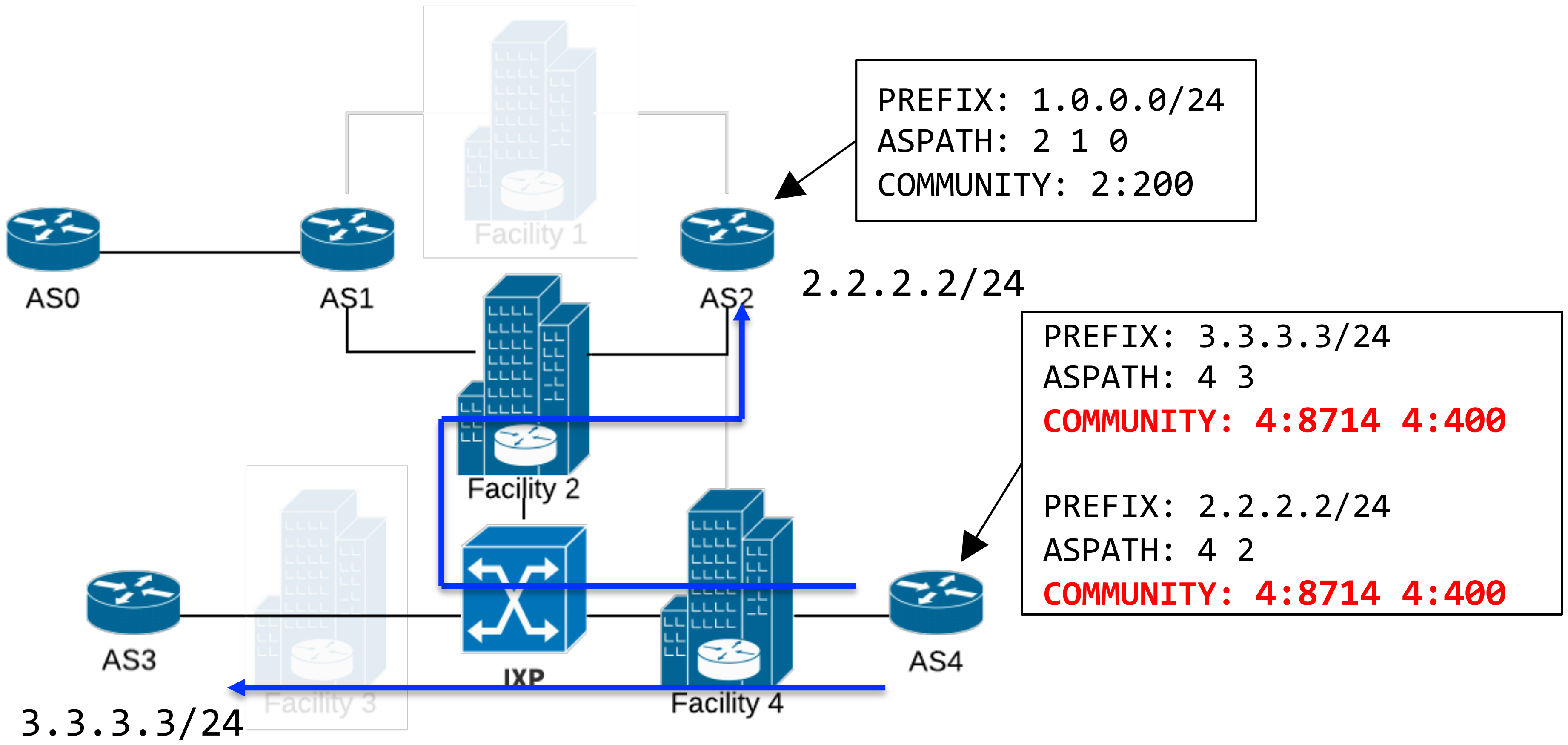


```
PREFIX: 1.0.0.0/24
ASPATH: 2 1 0
COMMUNITY: 2:200
```

The BGP Community 2:200 is used to tag routes received at Facility 2

i.e, Location Information!!

Deciphering location metadata in BGP

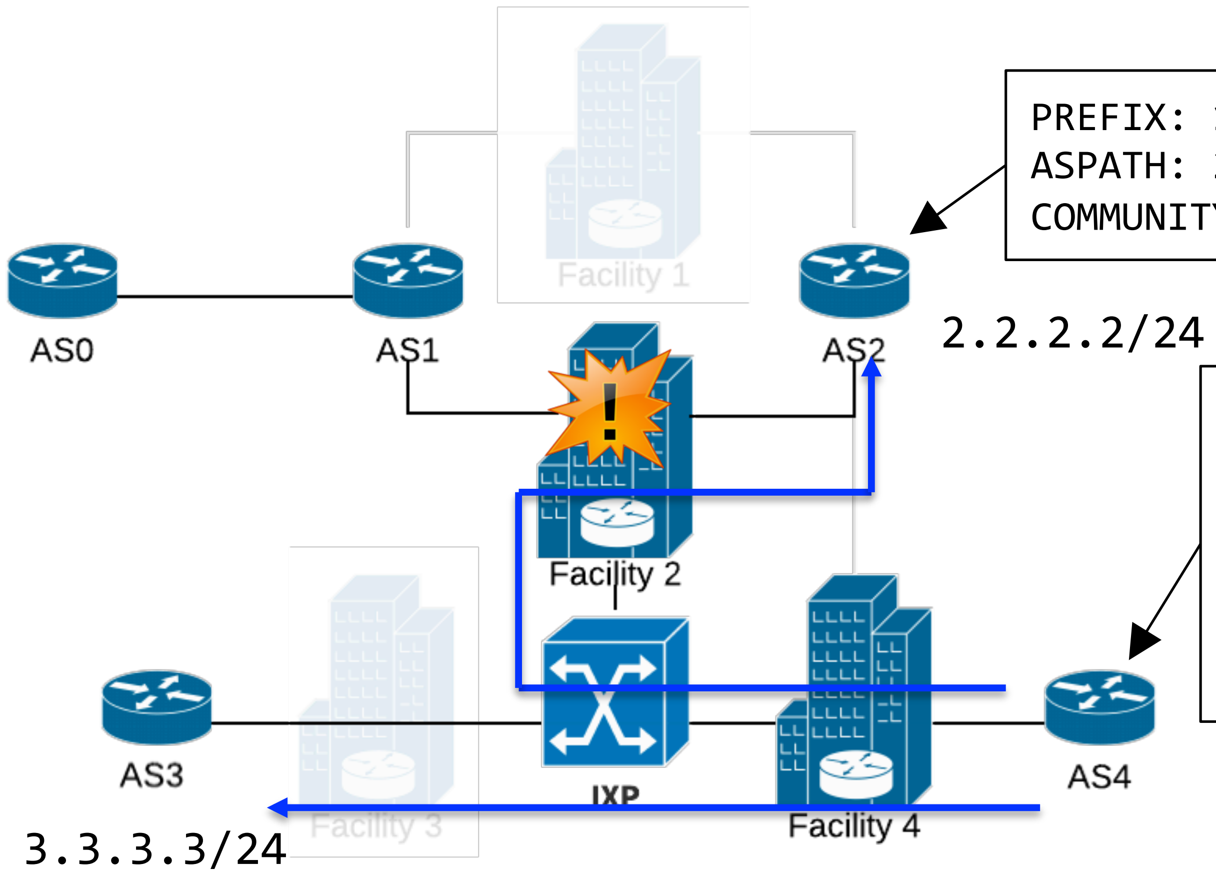


```
PREFIX: 1.0.0.0/24
ASPATH: 2 1 0
COMMUNITY: 2:200
```

```
PREFIX: 3.3.3.3/24
ASPATH: 4 3
COMMUNITY: 4:8714 4:400

PREFIX: 2.2.2.2/24
ASPATH: 4 2
COMMUNITY: 4:8714 4:400
```

The BGP Community 4:400 is used to tag routes received at Facility 4 and at the IXP

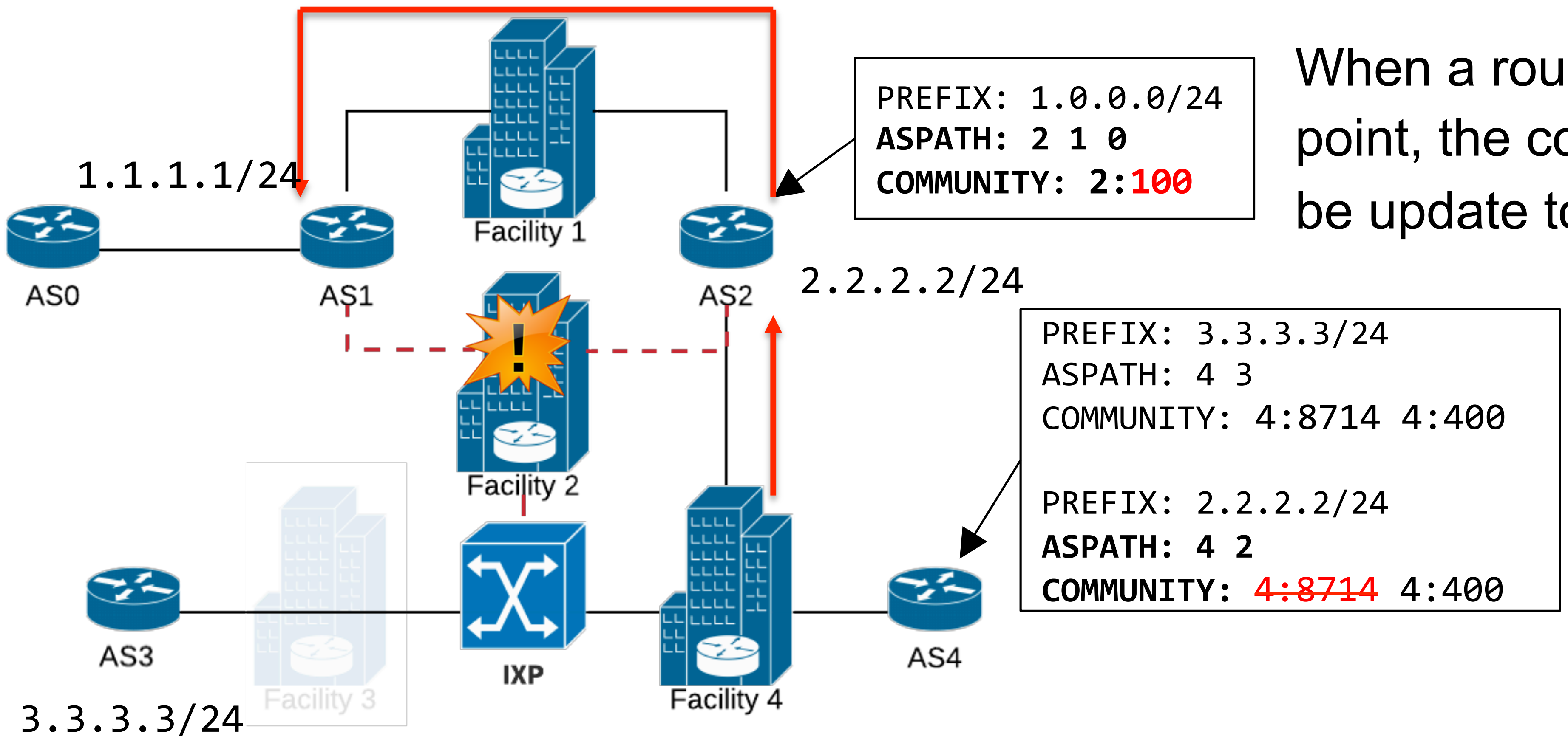
Deciphering location metadata in BGP



```
PREFIX: 1.0.0.0/24
ASPATH: 2 1 0
COMMUNITY: 2:200
```

```
PREFIX: 3.3.3.3/24
ASPATH: 4 3
COMMUNITY: 4:8714 4:400

PREFIX: 2.2.2.2/24
ASPATH: 4 2
COMMUNITY: 4:8714 4:400
```

Deciphering location metadata in BGP



PREFIX: 1.0.0.0/24
ASPATH: 2 1 0
COMMUNITY: 2:100

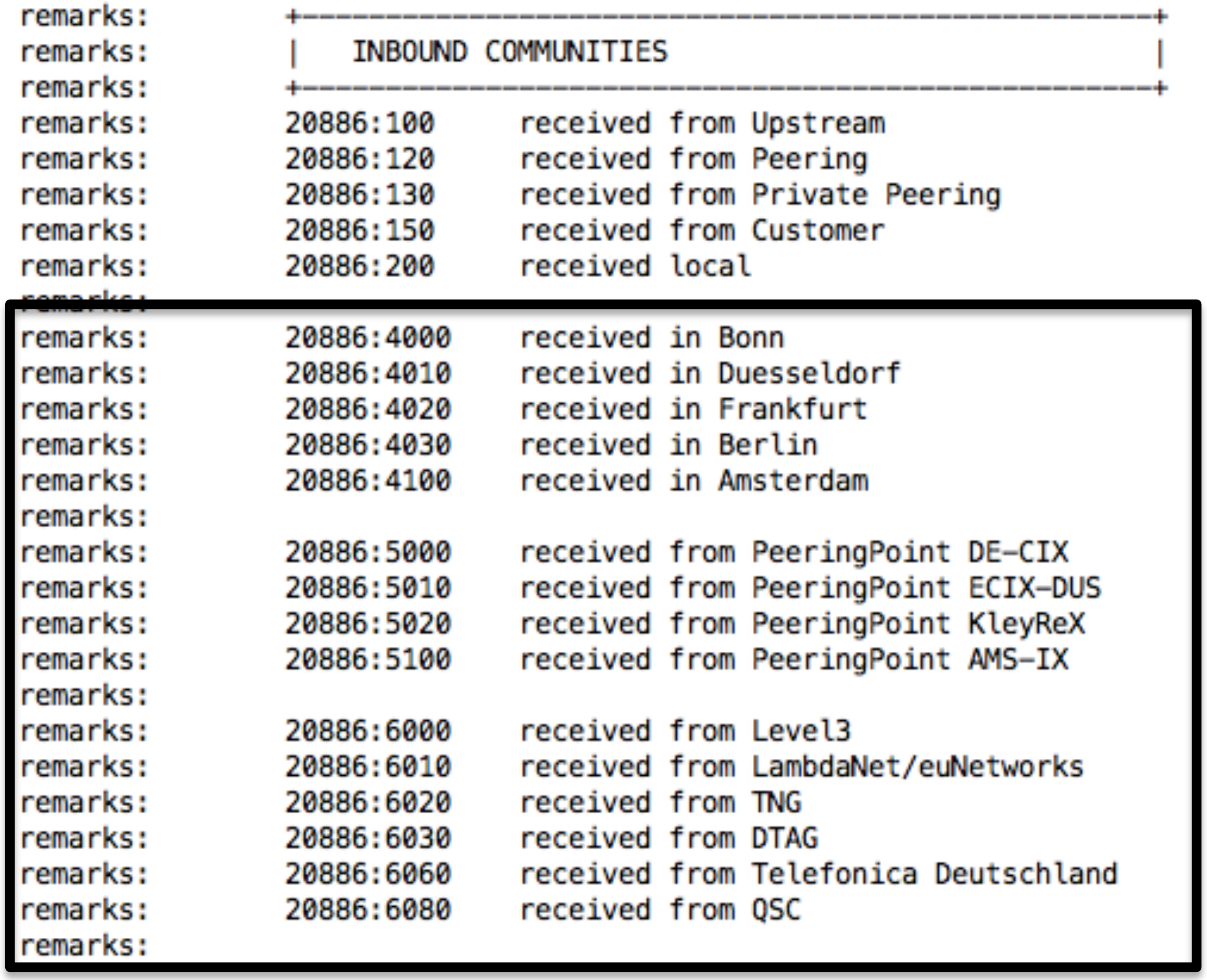
PREFIX: 3.3.3.3/24
ASPATH: 4 3
COMMUNITY: 4:8714 4:400

PREFIX: 2.2.2.2/24
ASPATH: 4 2
COMMUNITY: ~~4:8714~~ 4:400

When a route changes ingress point, the community values will be update to reflect the change.

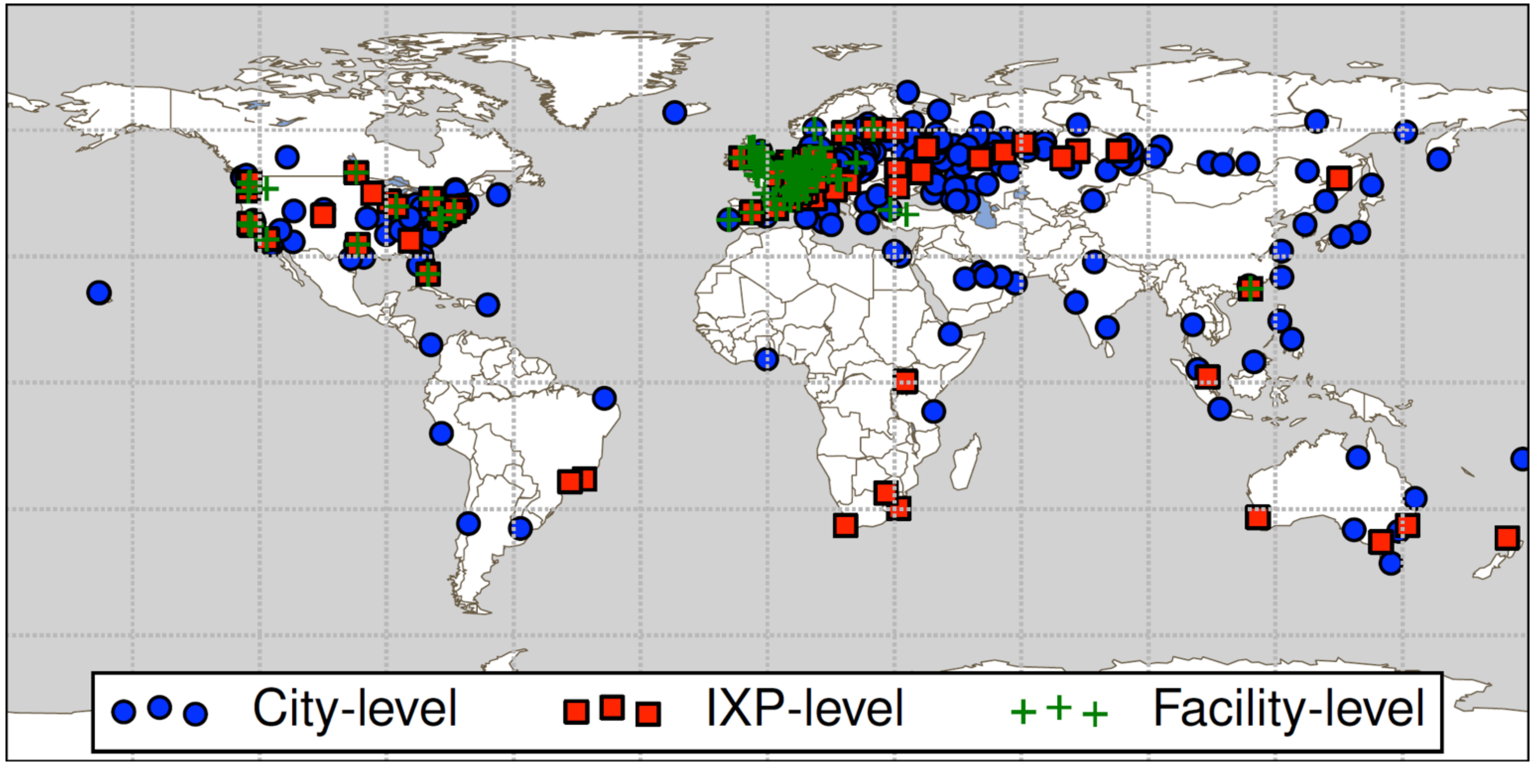
Building a BGP Communities Dictionary

- Community values **not** standardized
- Natural Language Tools
- Documentation in public data sources: Internet Routing Registries (IRRs), NOCs websites



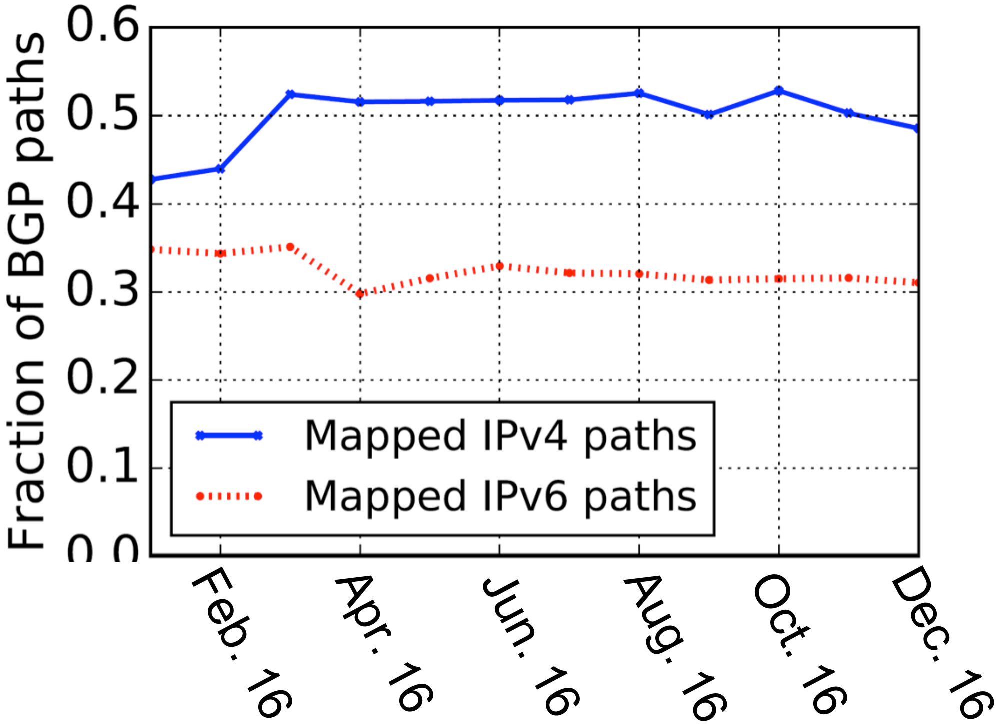
```
remarks:        +-----------------------------------------------------+
remarks:        |   INBOUND COMMUNITIES                               |
remarks:        +-----------------------------------------------------+
remarks:        20886:100     received from Upstream
remarks:        20886:120     received from Peering
remarks:        20886:130     received from Private Peering
remarks:        20886:150     received from Customer
remarks:        20886:200     received local
remarks:
remarks:        20886:4000    received in Bonn
remarks:        20886:4010    received in Duesseldorf
remarks:        20886:4020    received in Frankfurt
remarks:        20886:4030    received in Berlin
remarks:        20886:4100    received in Amsterdam
remarks:
remarks:        20886:5000    received from PeeringPoint DE-CIX
remarks:        20886:5010    received from PeeringPoint ECIX-DUS
remarks:        20886:5020    received from PeeringPoint KleyReX
remarks:        20886:5100    received from PeeringPoint AMS-IX
remarks:
remarks:        20886:6000    received from Level3
remarks:        20886:6010    received from LambdaNet/euNetworks
remarks:        20886:6020    received from TNG
remarks:        20886:6030    received from DTAG
remarks:        20886:6060    received from Telefonica Deutschland
remarks:        20886:6080    received from QSC
remarks:
```

Building a BGP Communities Dictionary



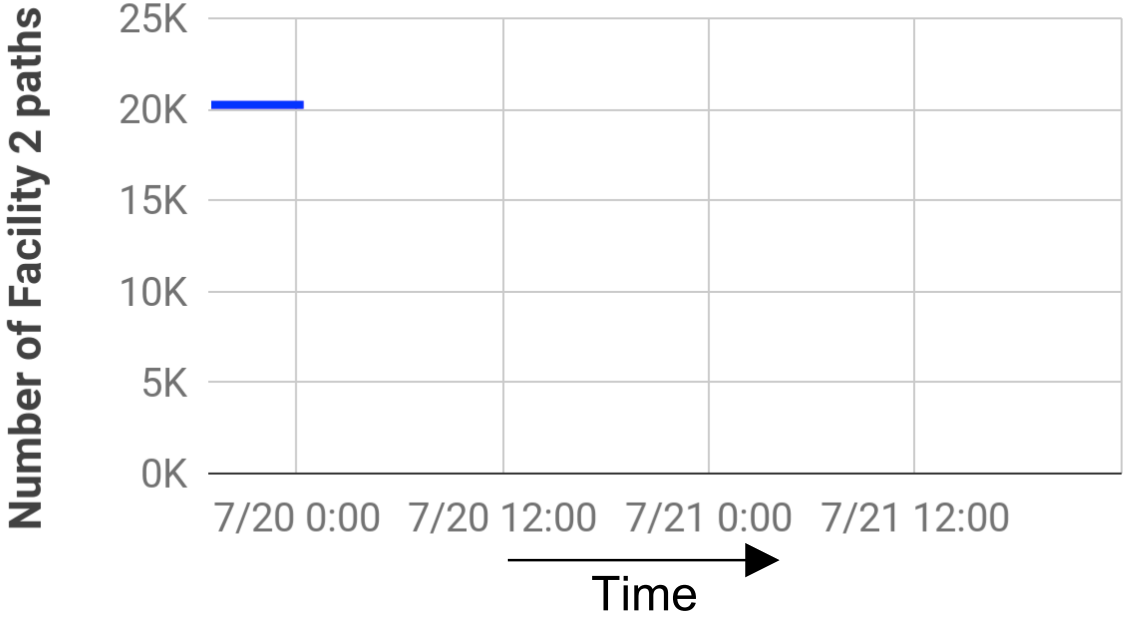
3,049 communities for **locations** used by 468 Ases

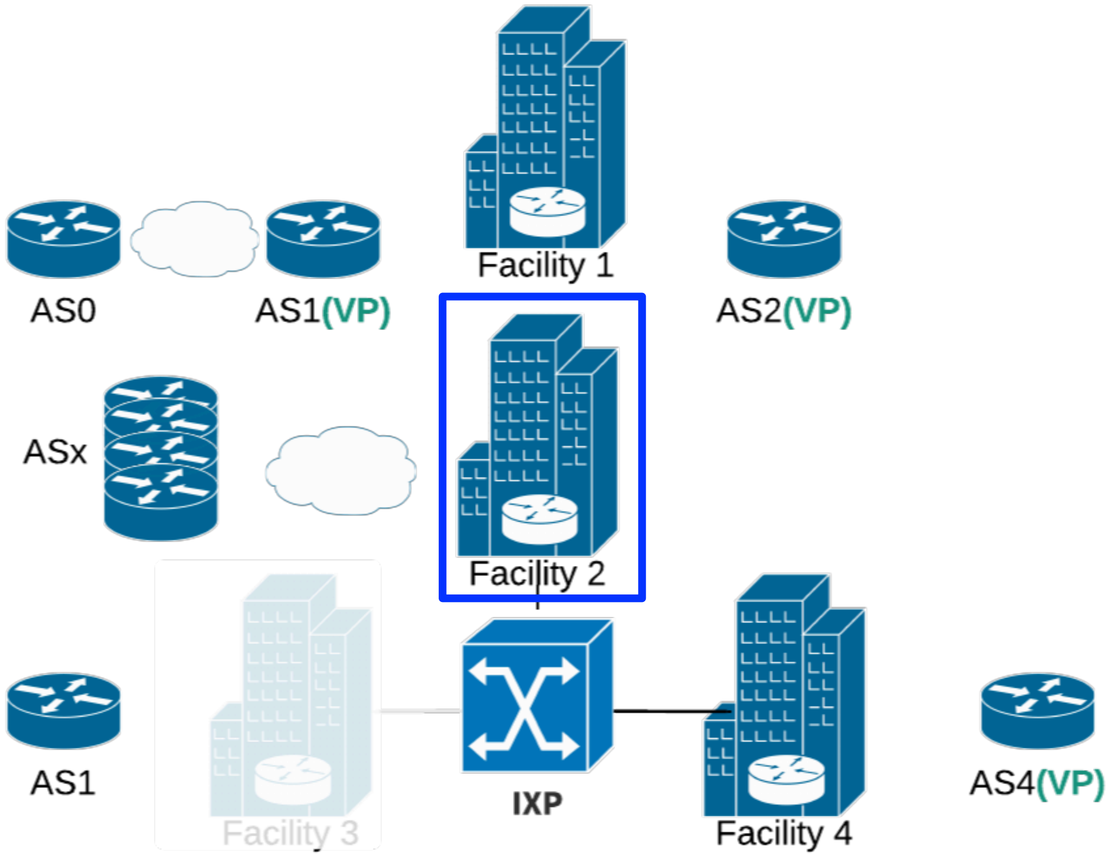
Topological coverage



- **~50%** of IPv4 and **~30%** of IPv6 paths annotated with at least one Community in our dictionary.
- **24%** of the facilities in PeeringDB, **98%** of the facilities with at least 20 members.

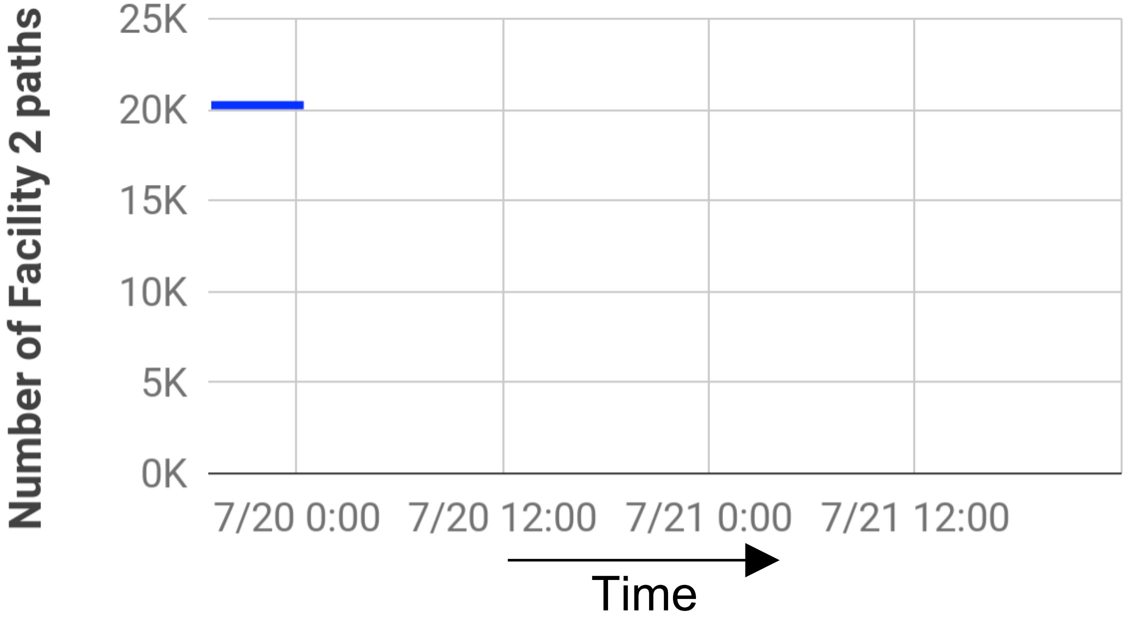
Passive outage detection: Initialization

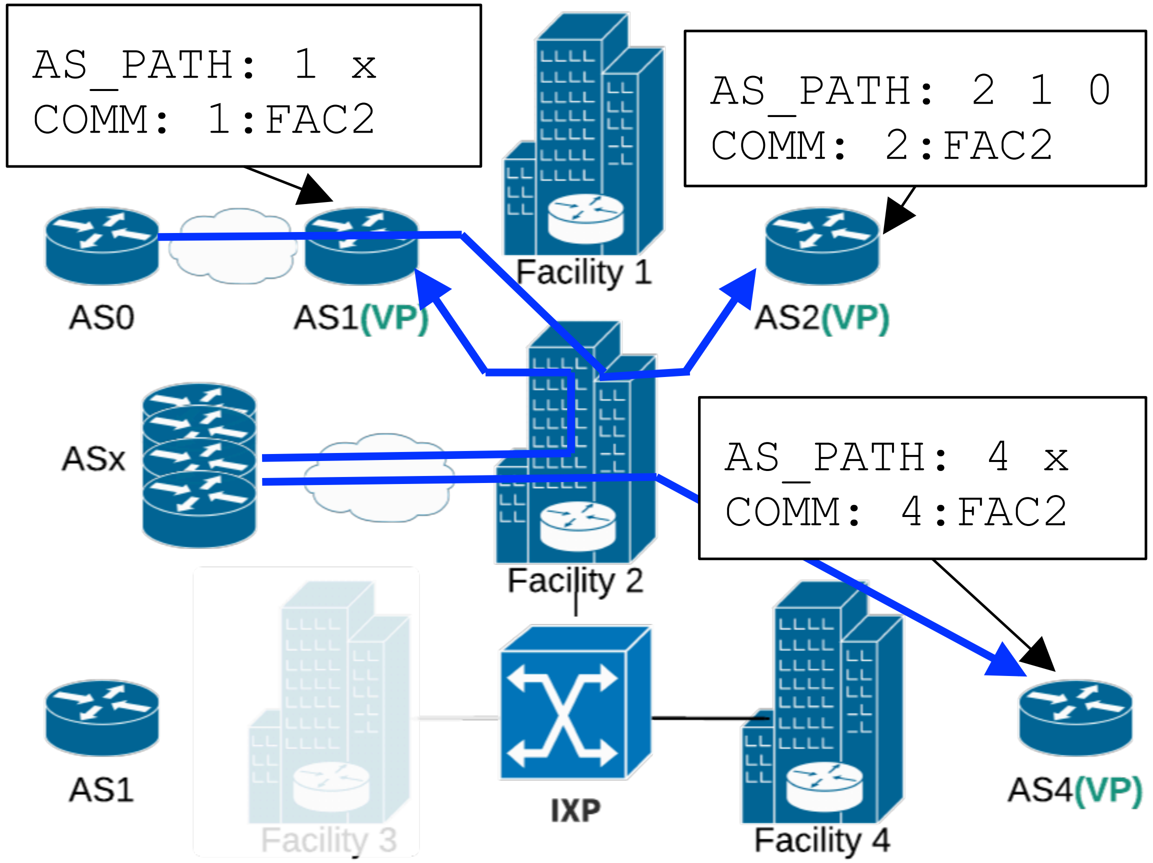




For each vantage point **(VP)** collect all the **stable** BGP routes tagged with the communities of the target facility (Facility 2)

Passive outage detection: Initialization



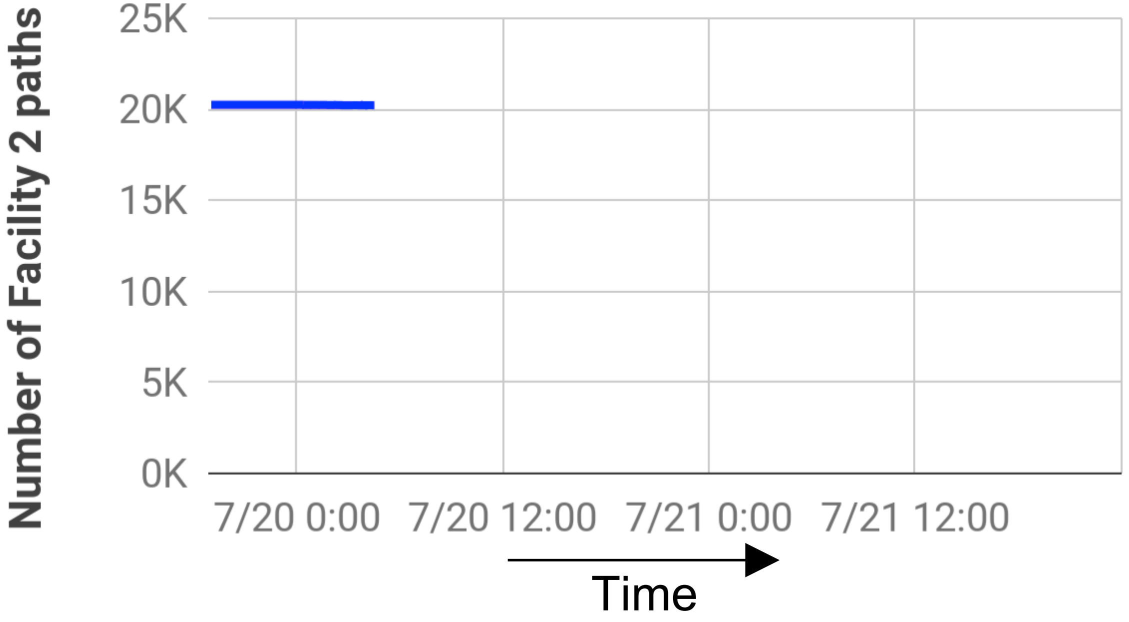

```
AS_PATH: 1 x
COMM: 1:FAC2
```

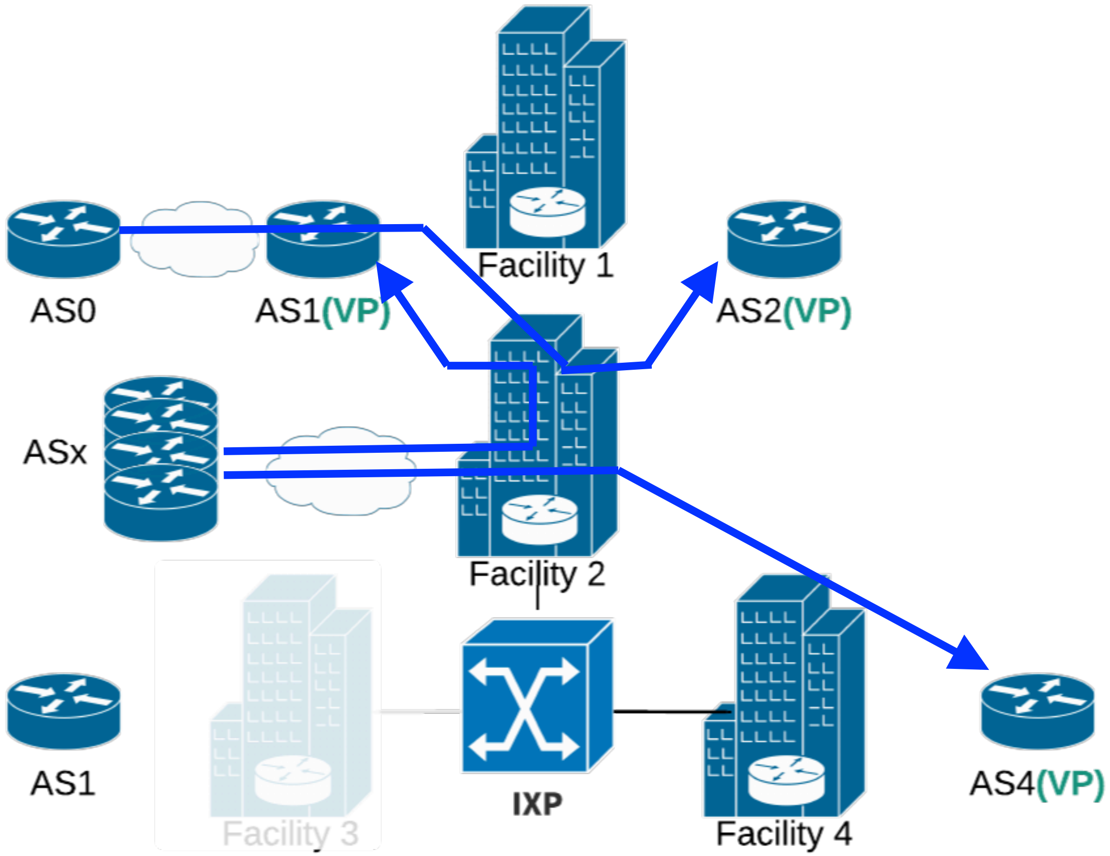
```
AS_PATH: 2 1 0
COMM: 2:FAC2
```

```
AS_PATH: 4 x
COMM: 4:FAC2
```

For each vantage point **(VP)** collect all the **stable** BGP routes tagged with the communities of the target facility (Facility 2)

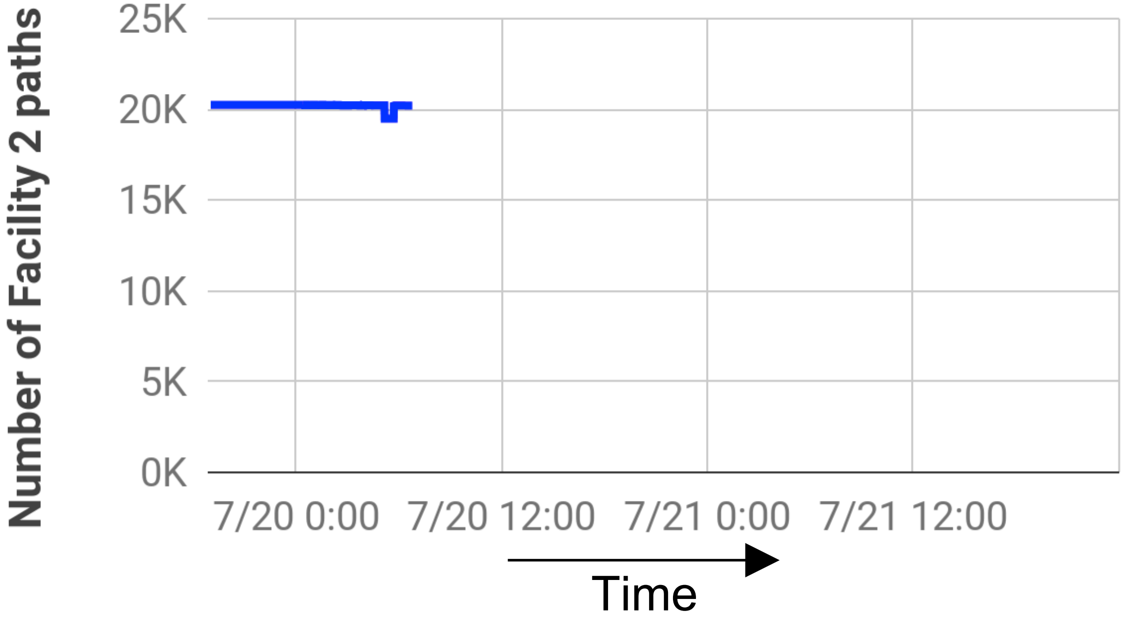
Passive outage detection: Monitoring

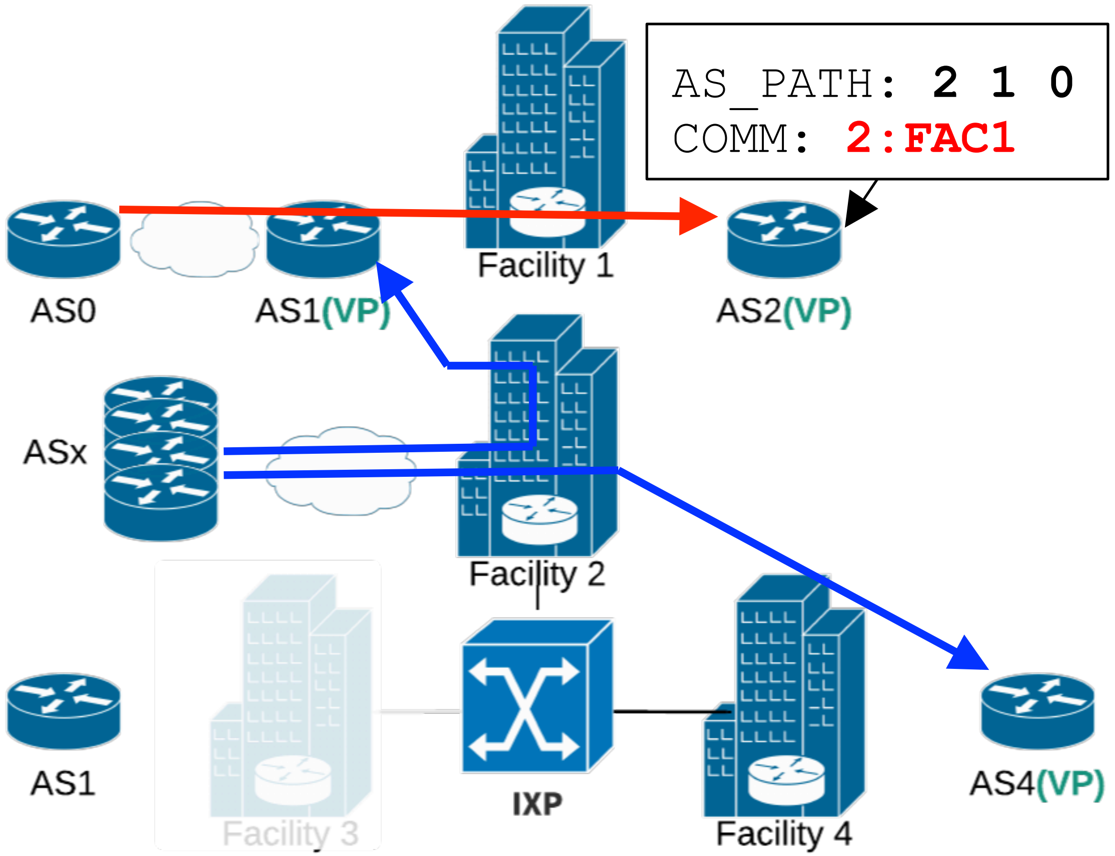




Track the BGP updates of the stable paths for changes in the communities values that indicate ingress point change.

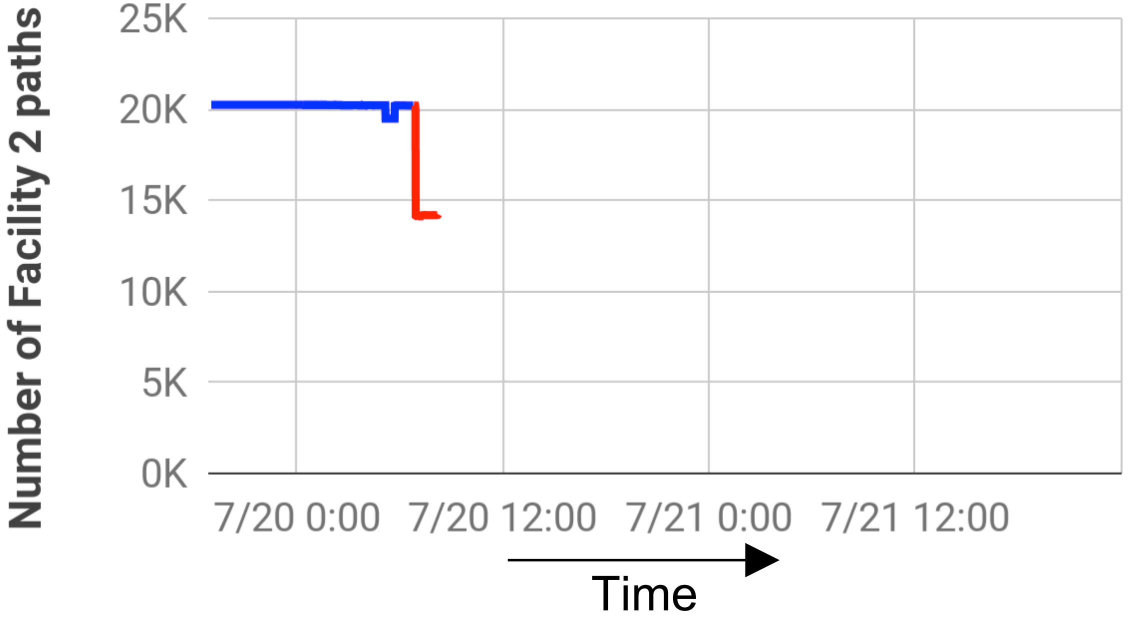
Passive outage detection: Monitoring

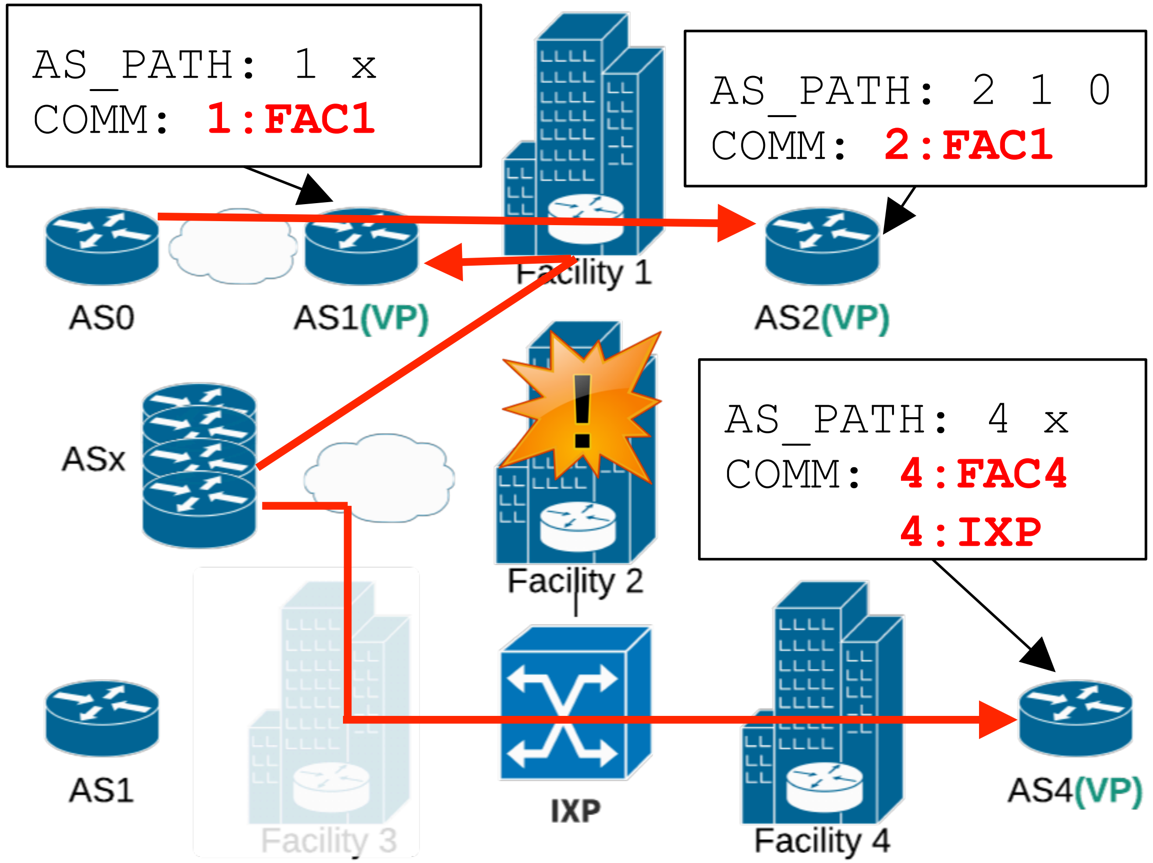




We ignore about single router-level/
AS-level path changes if the ingress-tagging
communities remain the same.

Passive outage detection: Outage signal



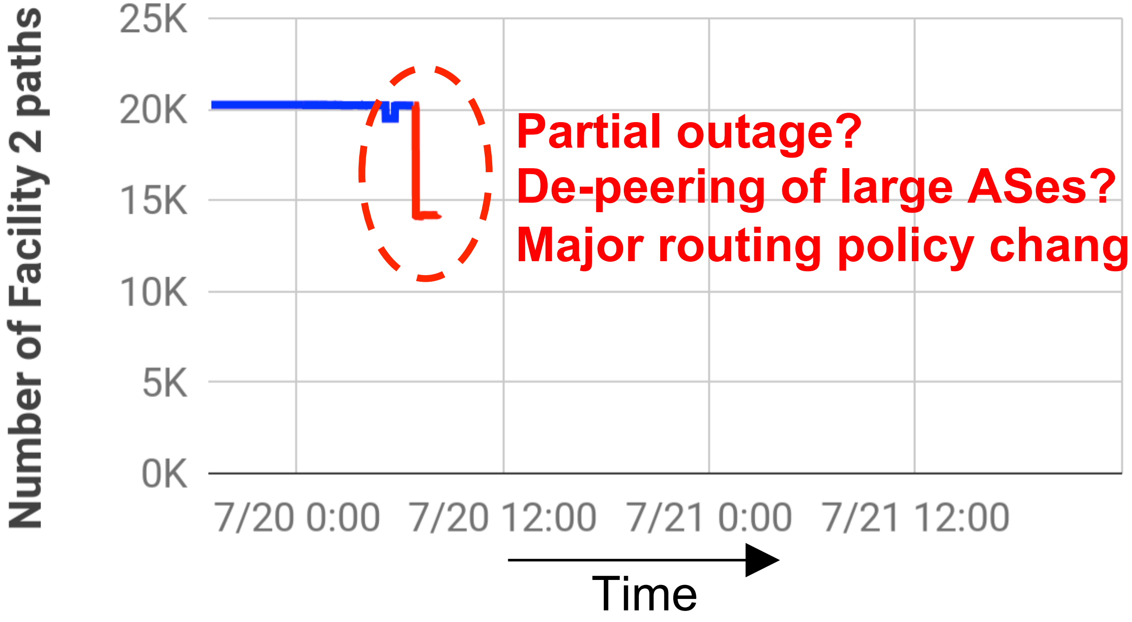

AS_PATH: 1 x
COMM: 1:FAC1

AS_PATH: 2 1 0
COMM: 2:FAC1

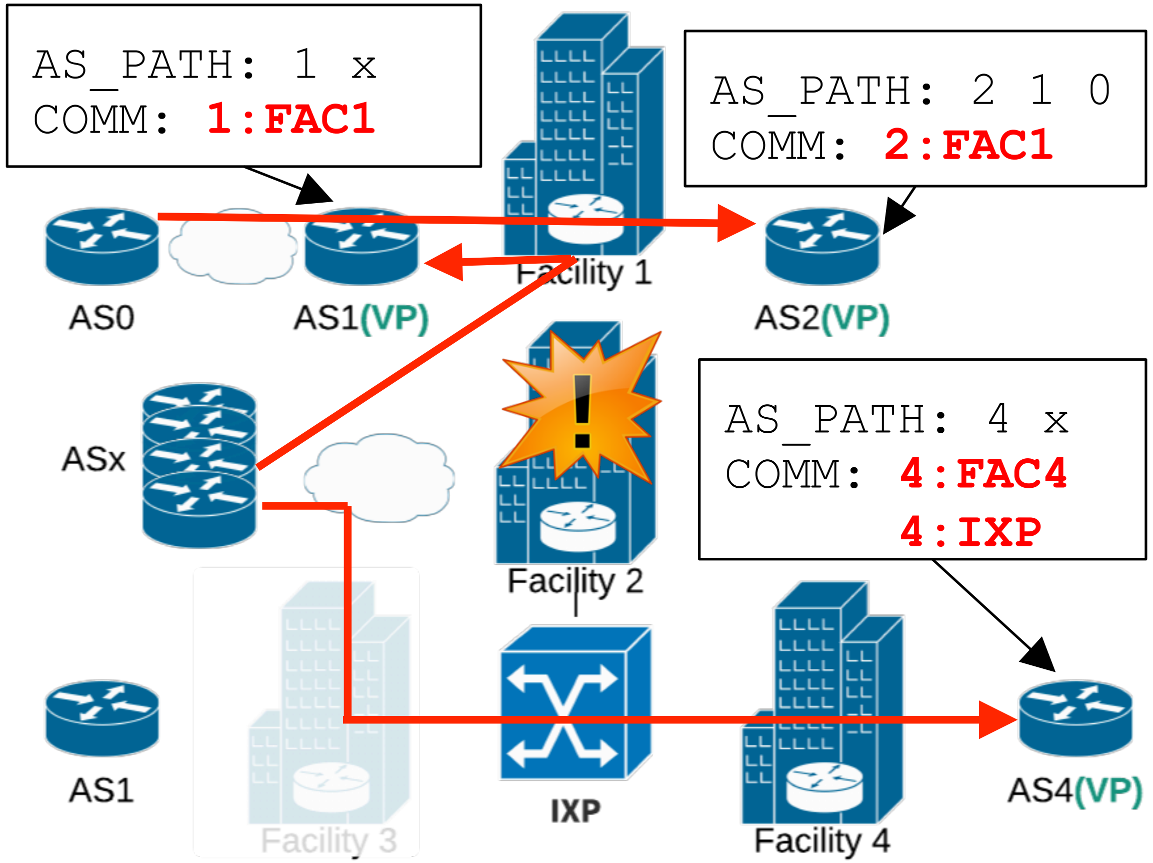
AS_PATH: 4 x
COMM: 4:FAC4
4:IXP

Crowdsourcing mechanism: Concurrent changes of communities values for multiple networks for the same facility is an indication of outage.

Passive outage detection: Outage signal



Partial outage?
De-peering of large ASes?
Major routing policy change?



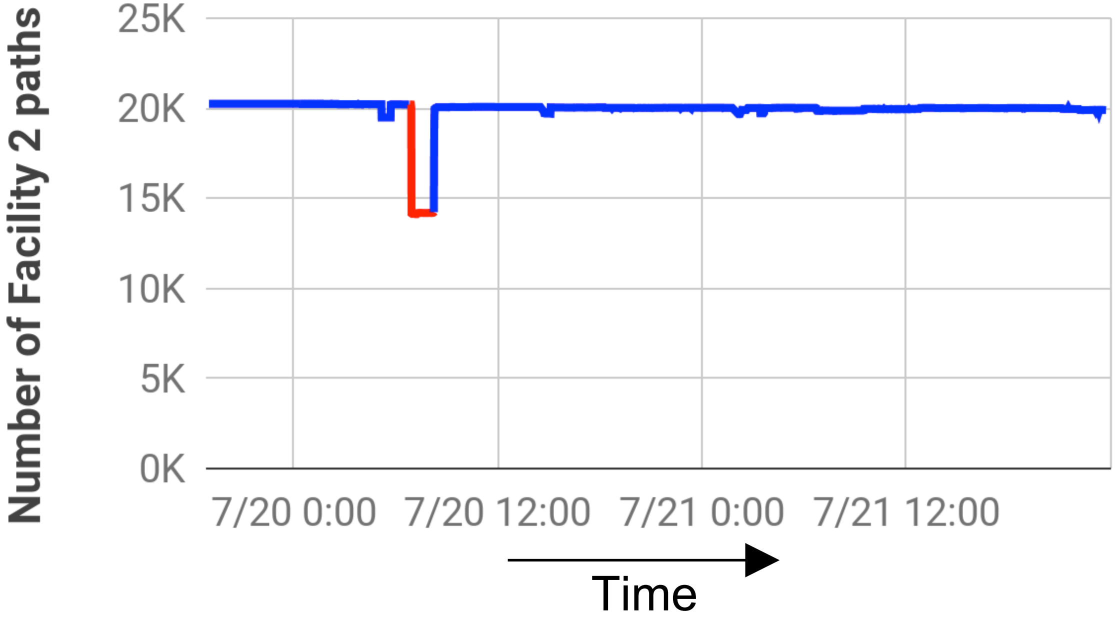
```
AS_PATH: 1 x
COMM: 1:FAC1
```

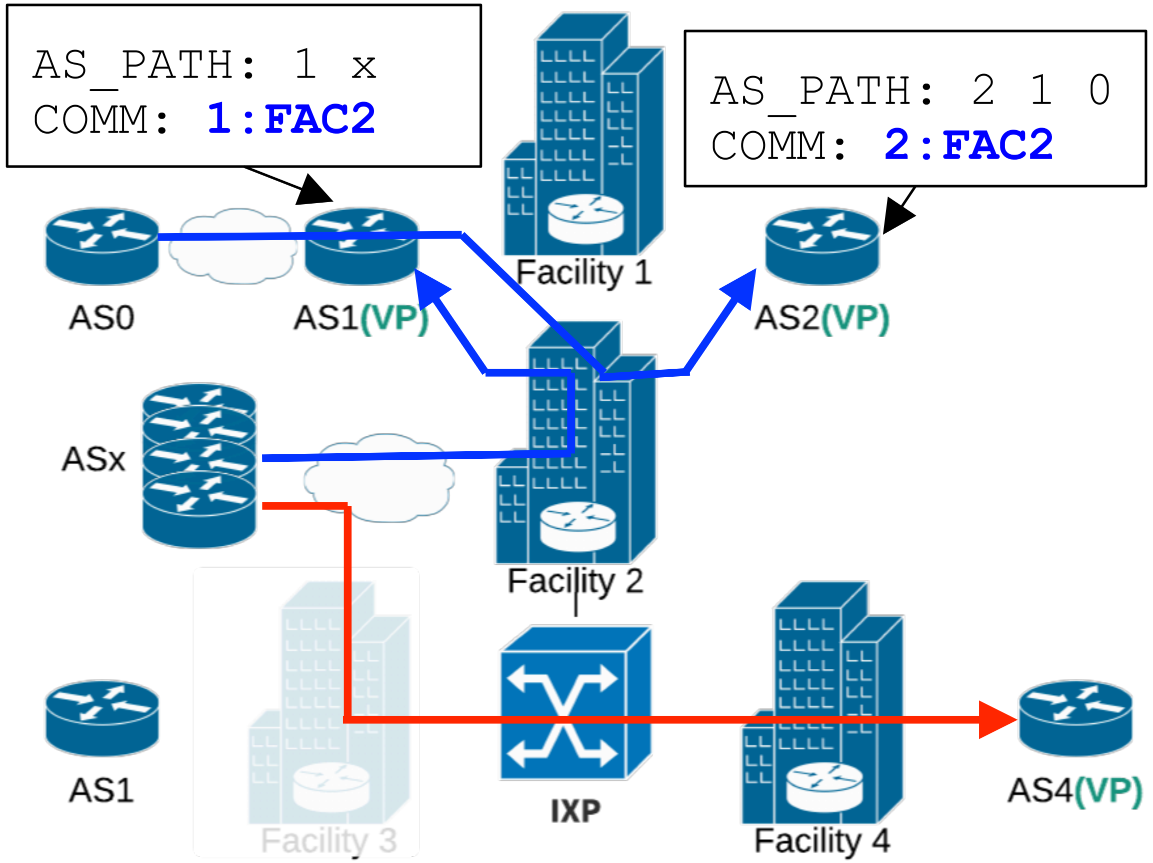
```
AS_PATH: 2 1 0
COMM: 2:FAC1
```

```
AS_PATH: 4 x
COMM: 4:FAC4
      4:IXP
```

Crowdsourcing mechanism: Concurrent changes of communities values for multiple networks for the same facility is an indication of outage.

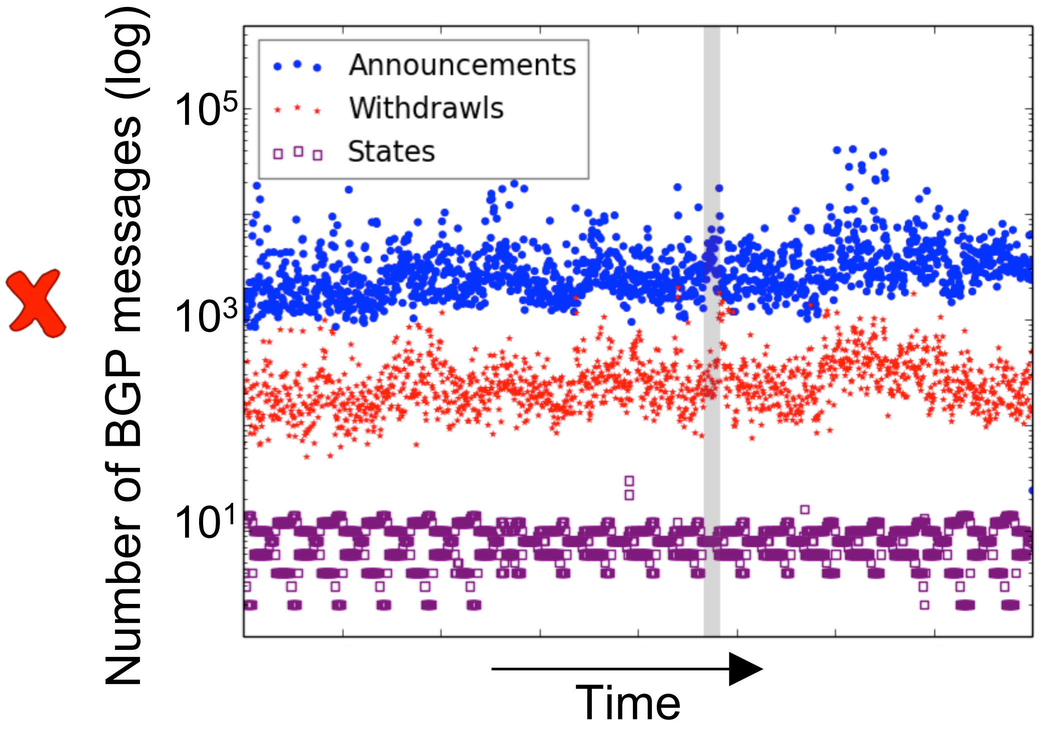
Passive outage detection: Outage tracking





End of outage inferred when the majority of paths return to the original facility.

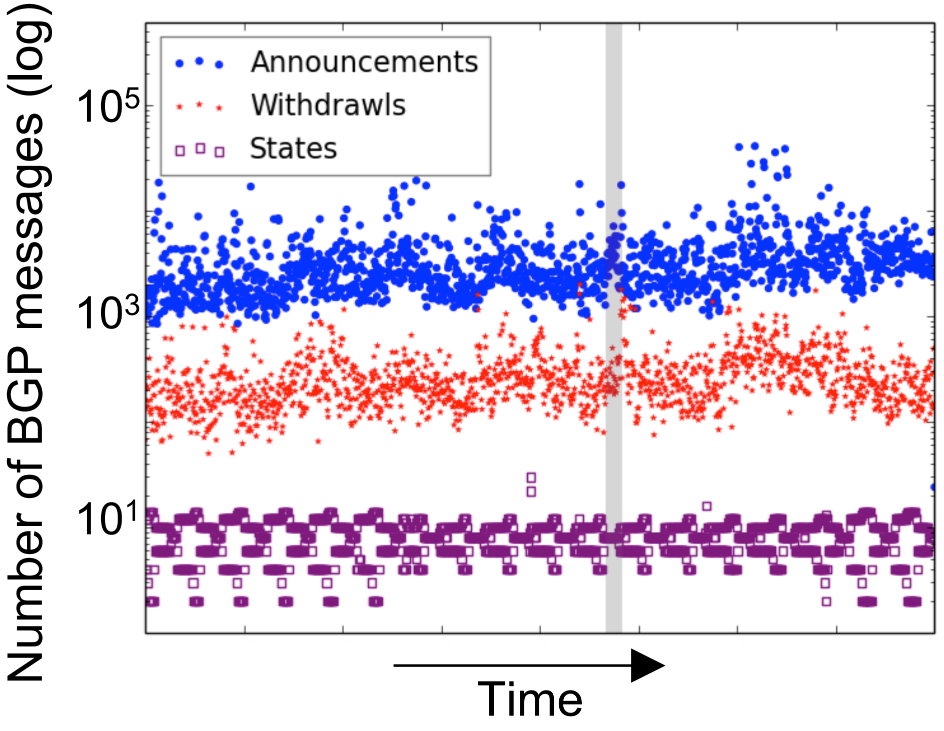
De-noising BGP routing activity



The aggregated activity of BGP messages (announcements, withdrawals, states) provides no outage indication.

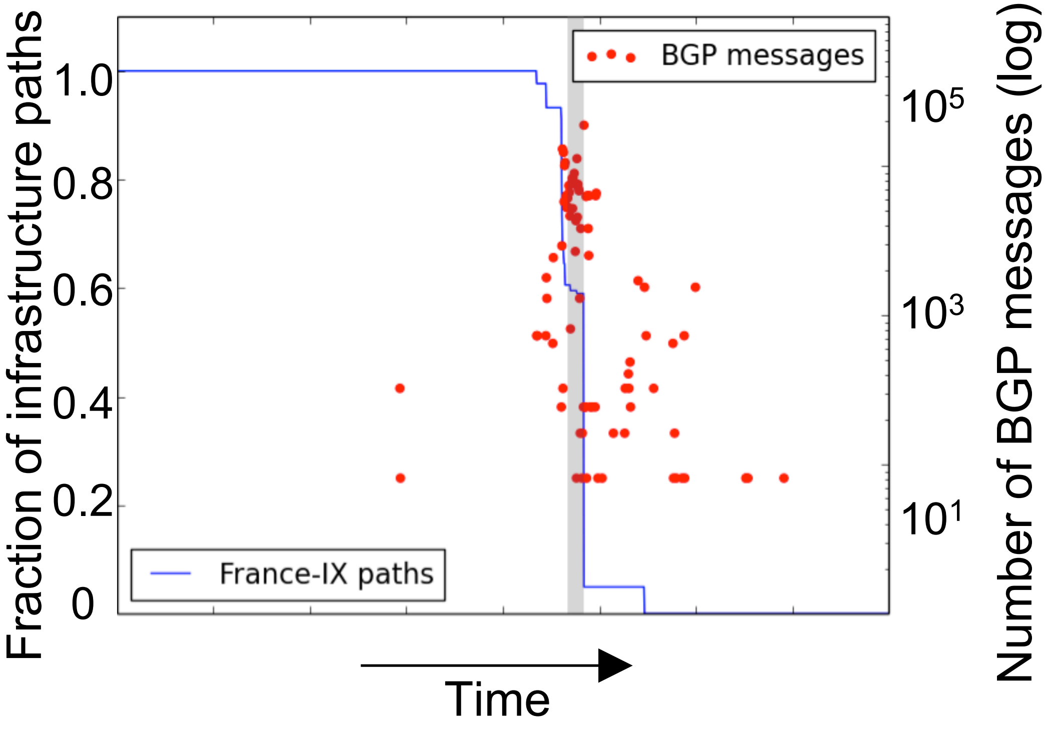
De-noising BGP routing activity

✗



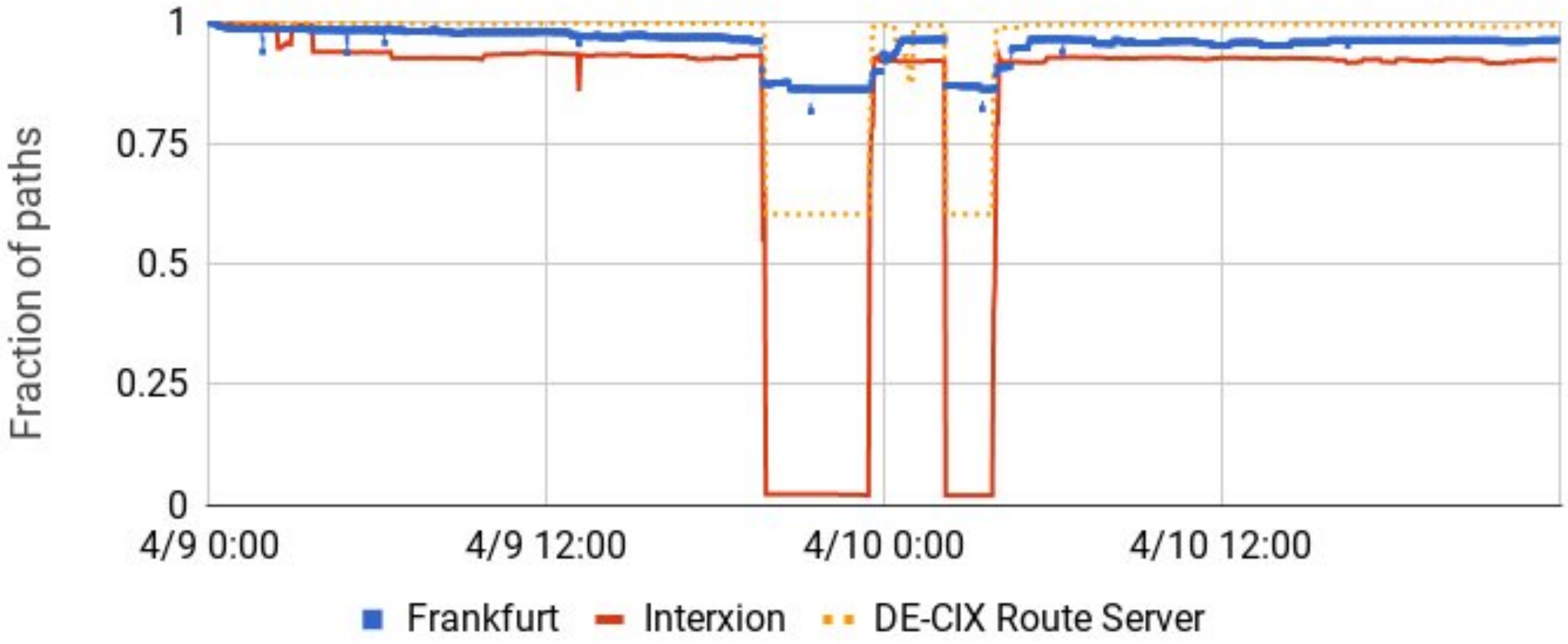
The aggregated activity of BGP messages (announcements, withdrawals, states) provides no outage indication.

✓



The BGP activity filtered using communities provides **strong outage signal**.

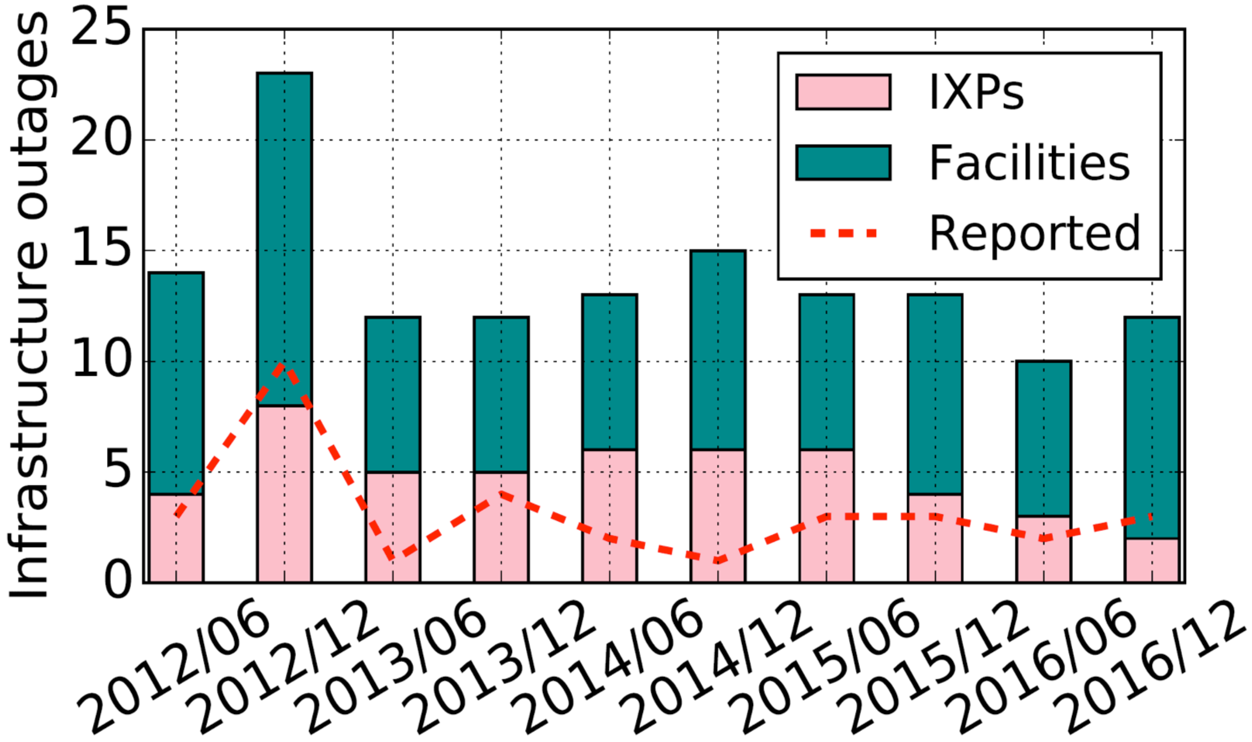
Providing Hard Evidence: DE-CIX? Outage



Interxion Frankfurt Outage (2018/04/09)

Changes in BGP paths annotated with communities that tag the location of inter-domain connections

Observed outages



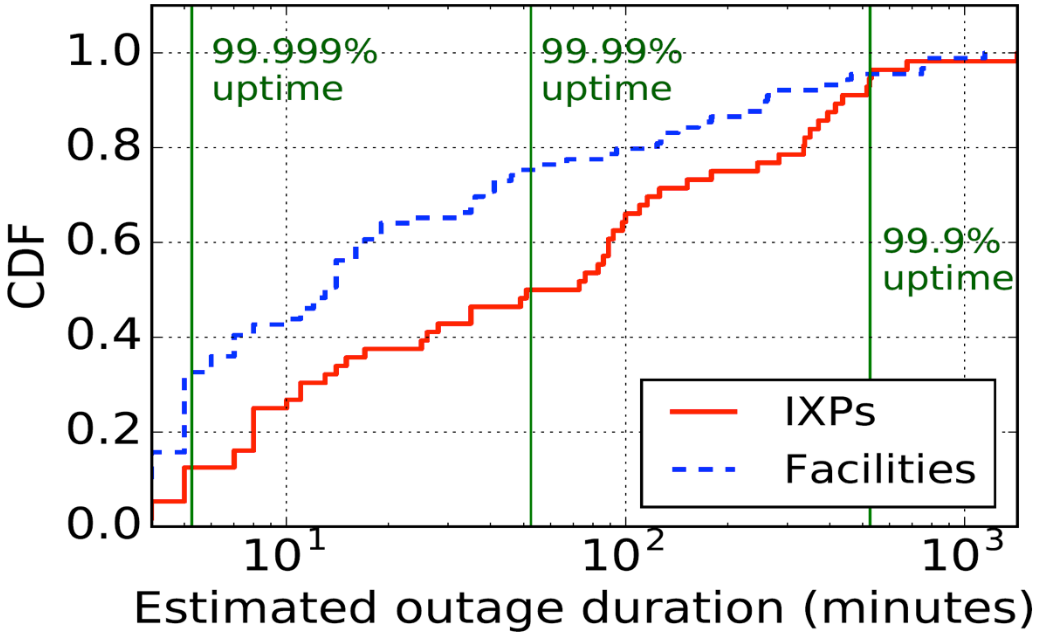
- **159** outages in 5 years of BGP data

  76% of the outages not reported in popular mailing lists/websites

- Validation through status reports, direct feedback, social media

  90% accuracy, **93%** precision (for trackable PoPs)

Effect of outages on Service Level Agreements

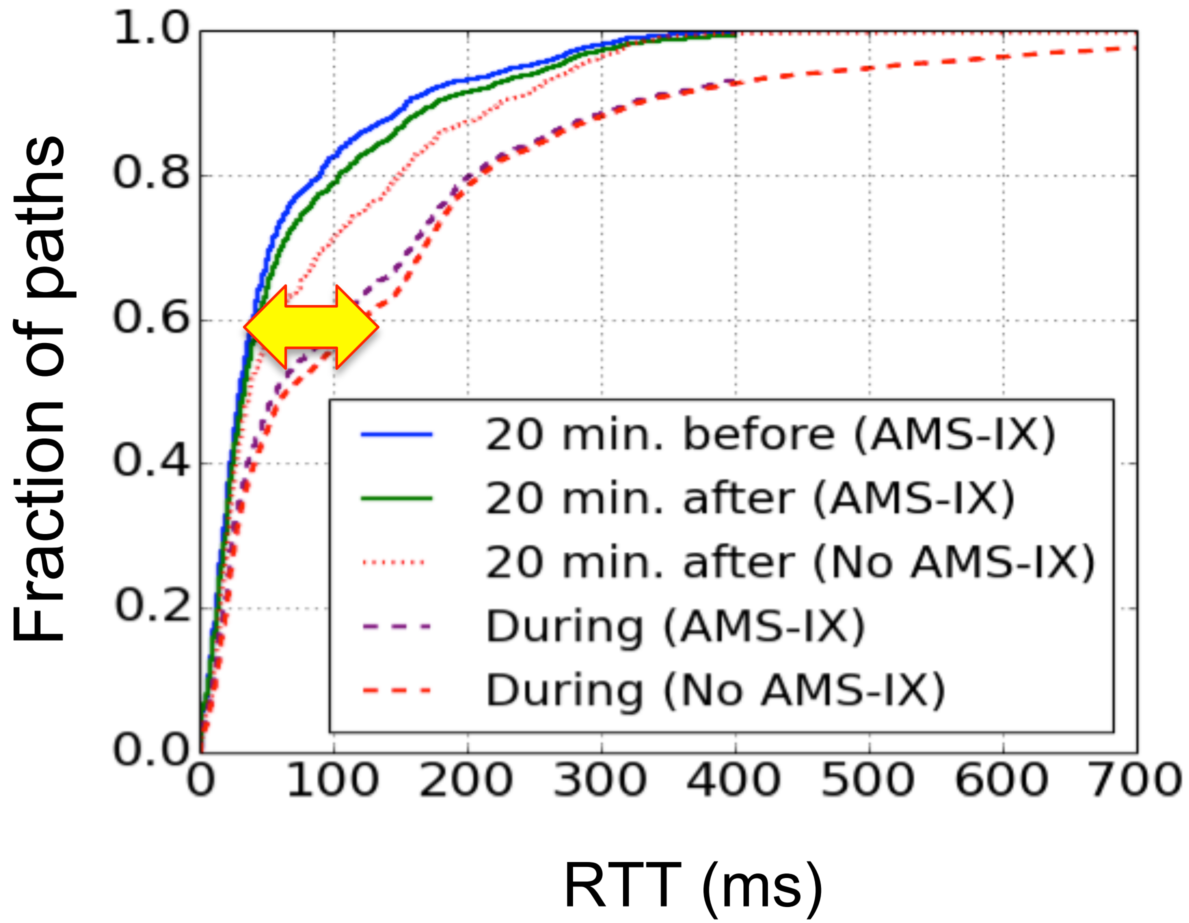


~70% of failed facilities worse than 99.999% uptime

~50% of failed IXPs worse than 99.99% uptime

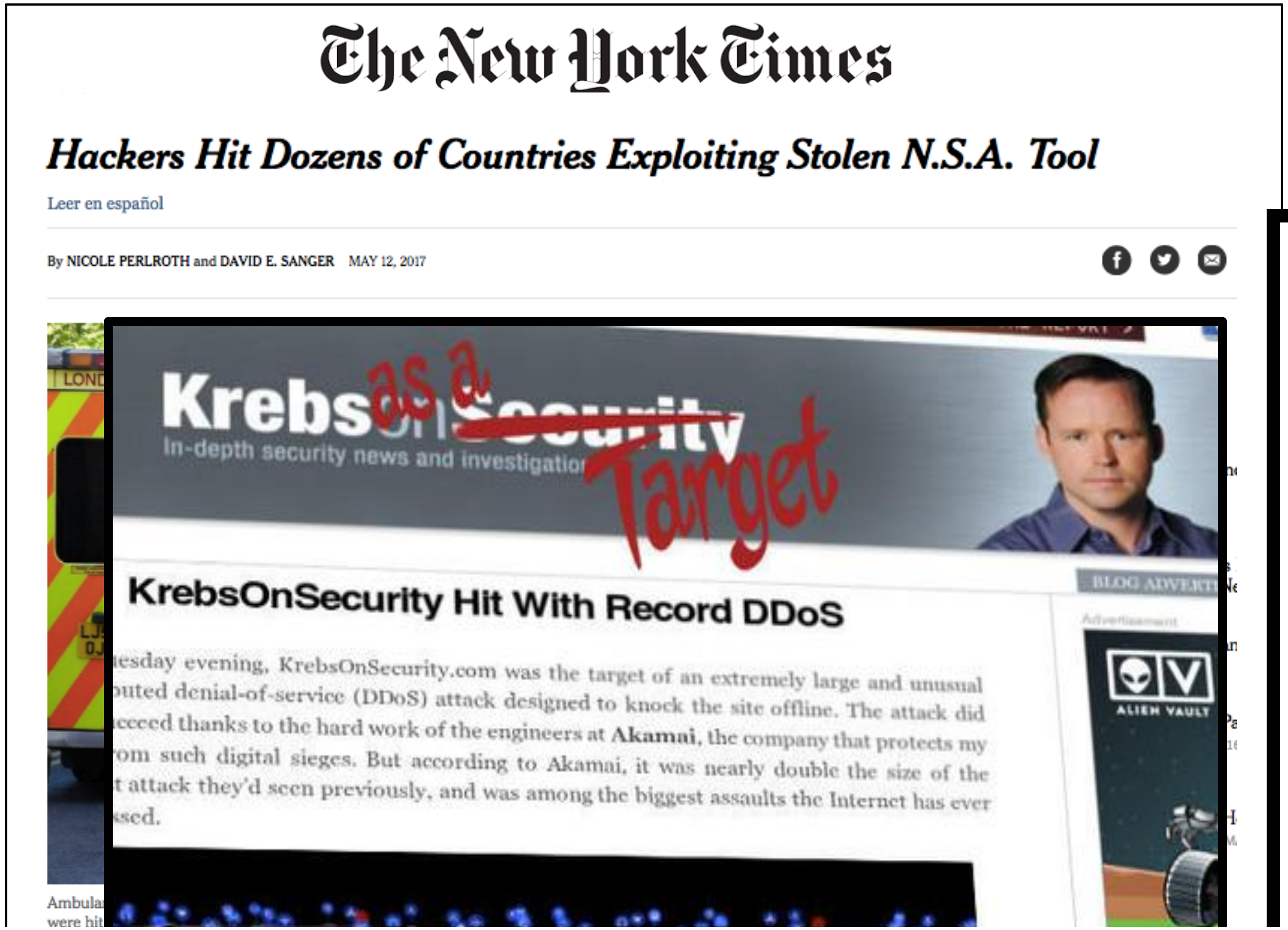
5% of failed infrastructures worse than 99.9% uptime!

Measuring the performance impact of outages



Median RTT rises by **> 100 ms** for rerouted paths during AMS-IX outage.

Cyberattacks ~~and Outages~~ are Serious Threats



The New York Times

Hackers Hit Dozens of Countries Exploiting Stolen N.S.A. Tool

Leer en español

By NICOLE PERLROTH and DAVID E. SANGER MAY 12, 2017

Ambula
were hit

Krebs ~~OnSecurity~~ as a Target

In-depth security news and investigation

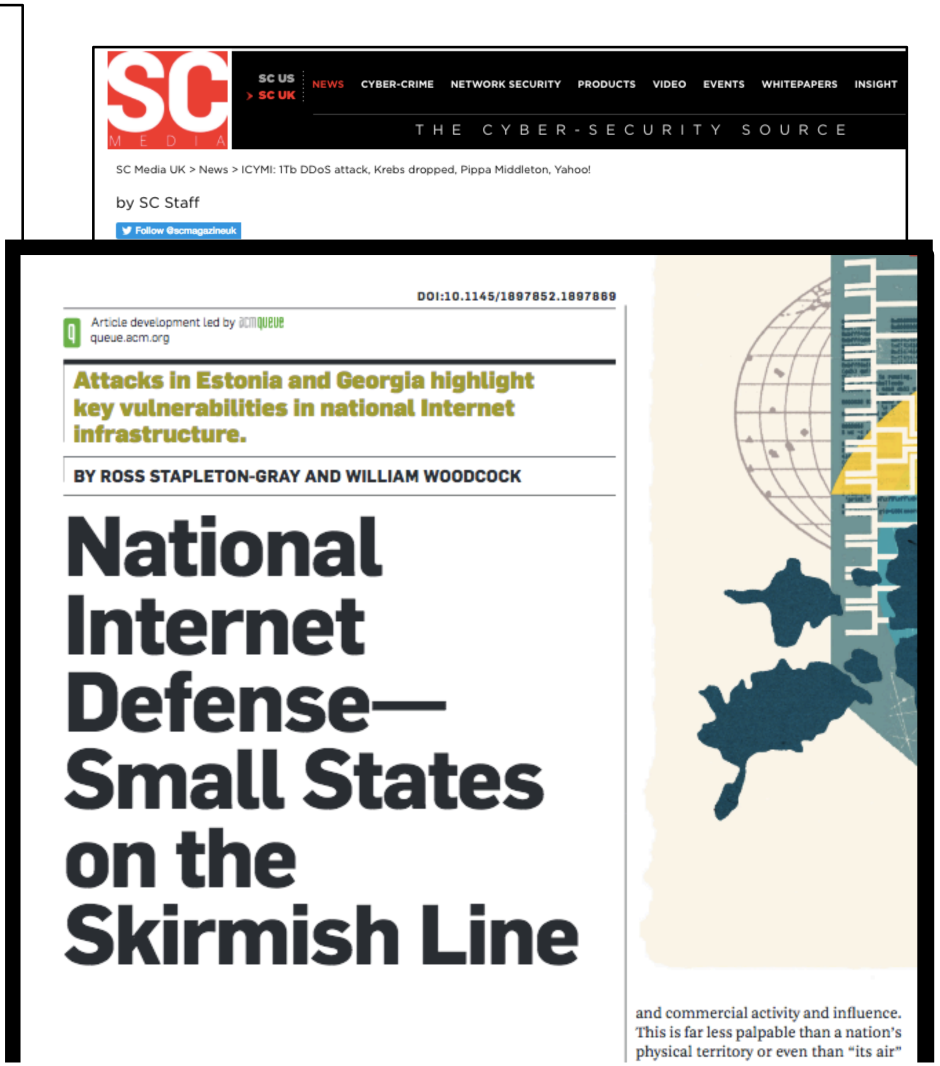
BLOG ADVERTI

Advertisement

ALIEN VAULT

KrebsOnSecurity Hit With Record DDoS

esday evening, KrebsOnSecurity.com was the target of an extremely large and unusual
buted denial-of-service (DDoS) attack designed to knock the site offline. The attack did
cceed thanks to the hard work of the engineers at **Akamai**, the company that protects my
rom such digital sieges. But according to Akamai, it was nearly double the size of the
t attack they'd seen previously, and was among the biggest assaults the Internet has ever
ssed.



SC MEDIA

SC US
› SC UK

NEWS CYBER-CRIME NETWORK SECURITY PRODUCTS VIDEO EVENTS WHITEPAPERS INSIGHT

THE CYBER-SECURITY SOURCE

SC Media UK > News > ICYMI: 1Tb DDoS attack, Krebs dropped, Pippa Middleton, Yahoo!

by SC Staff

Follow @scmagazineuk

DOI:10.1145/1897852.1897869

Article development led by acmqueue
queue.acm.org

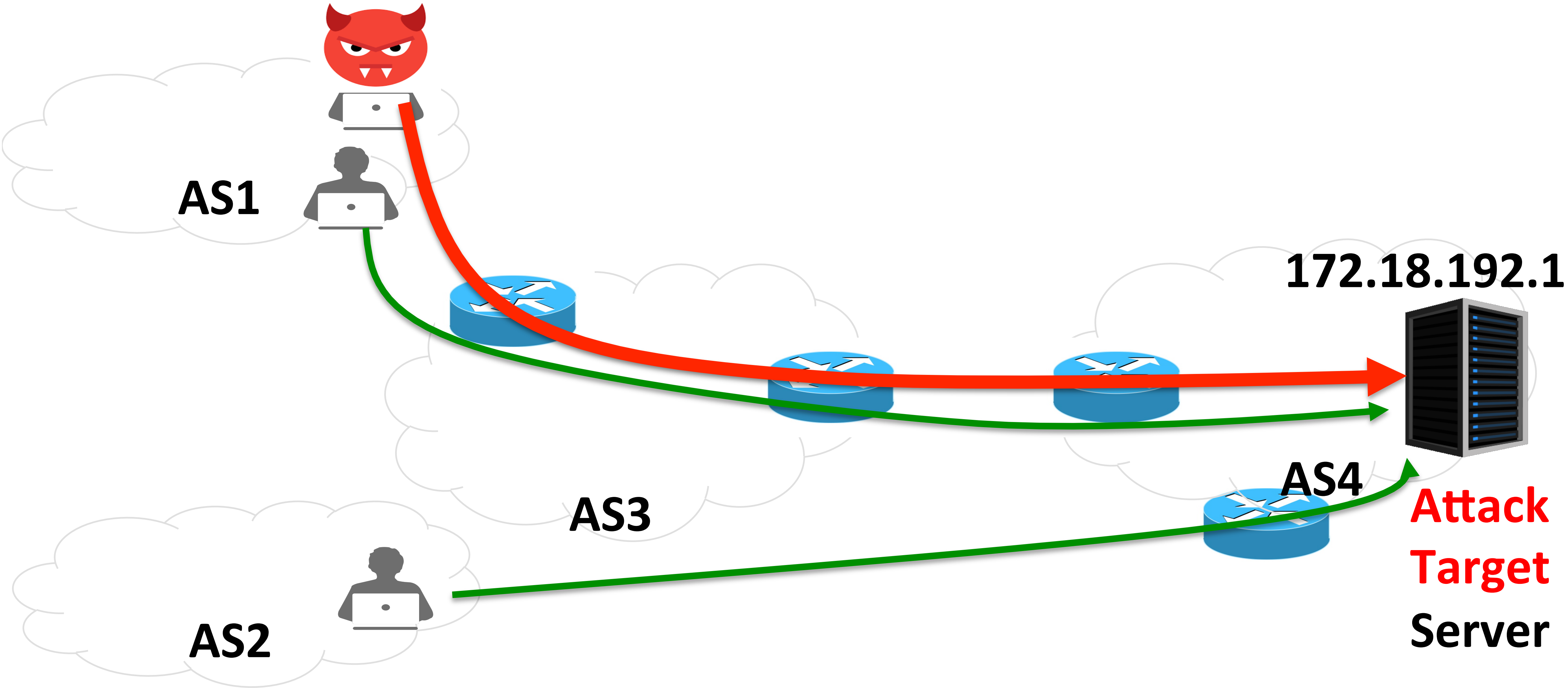
Attacks in Estonia and Georgia highlight key vulnerabilities in national Internet infrastructure.

BY ROSS STAPLETON-GRAY AND WILLIAM WOODCOCK

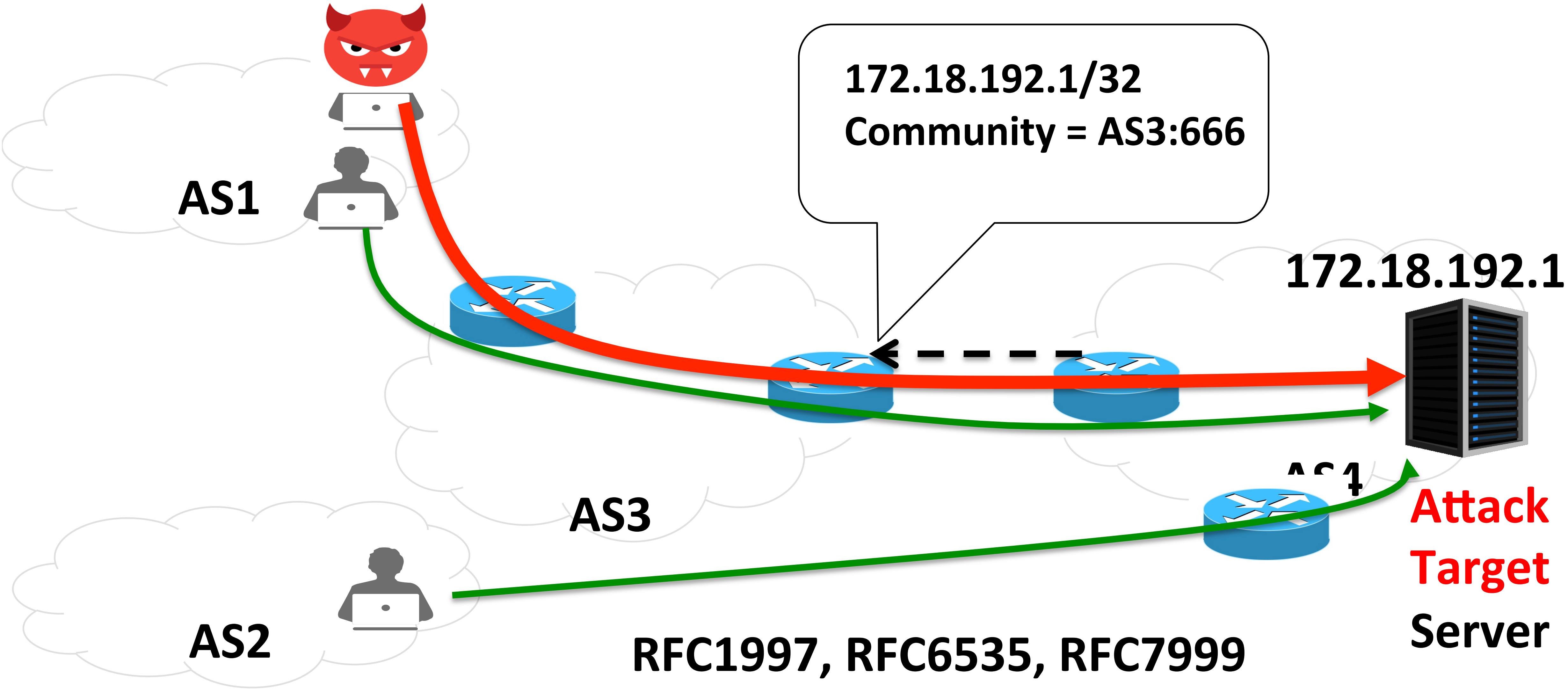
National Internet Defense—Small States on the Skirmish Line

and commercial activity and influence. This is far less palpable than a nation's physical territory or even than "its air"

Networks under Attack



BGP Blackholing in the Internet



172.18.192.1/32
Community = AS3:666

AS1

AS2

AS3

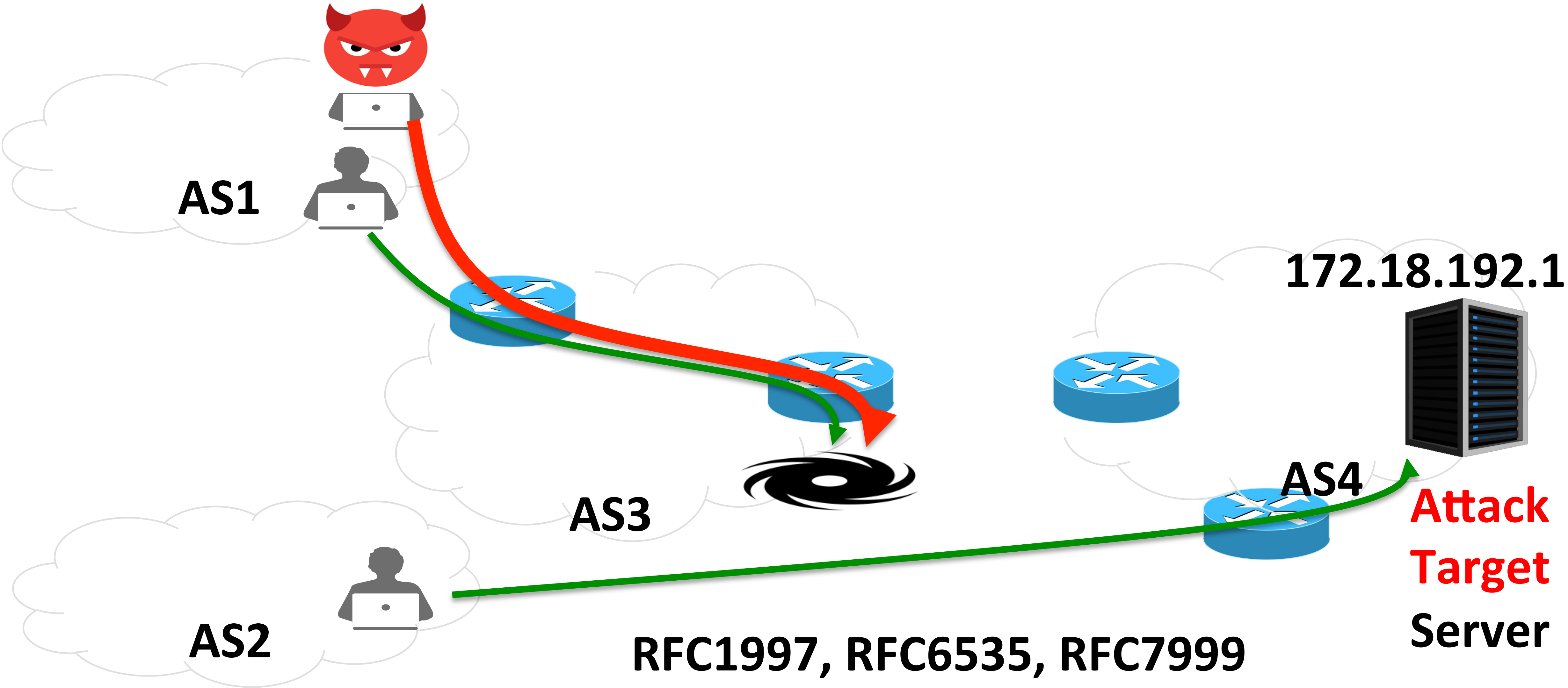
AS4

172.18.192.1

Attack
Target
Server

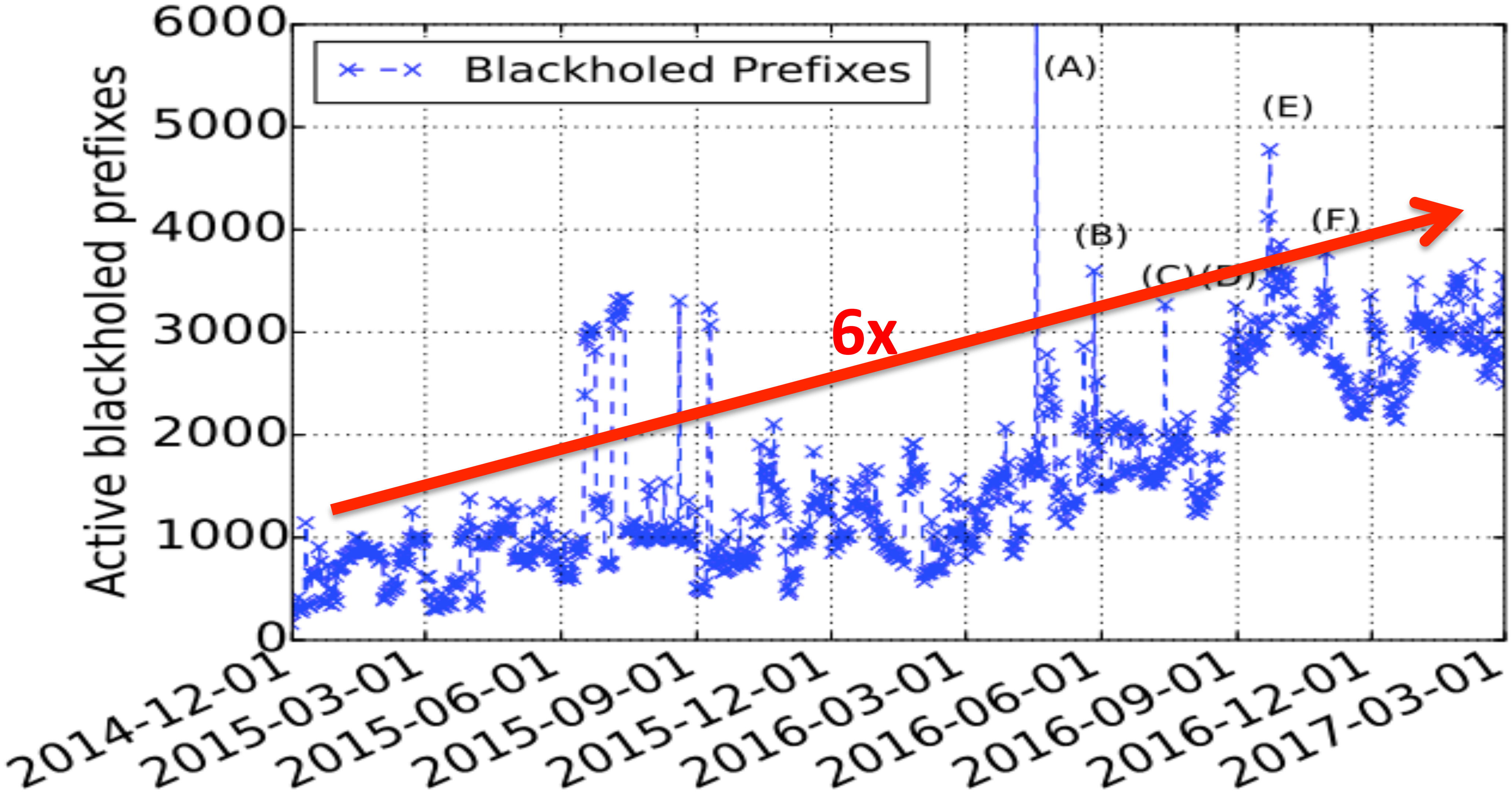
RFC1997, RFC6535, RFC7999

BGP Blackholing in the Internet

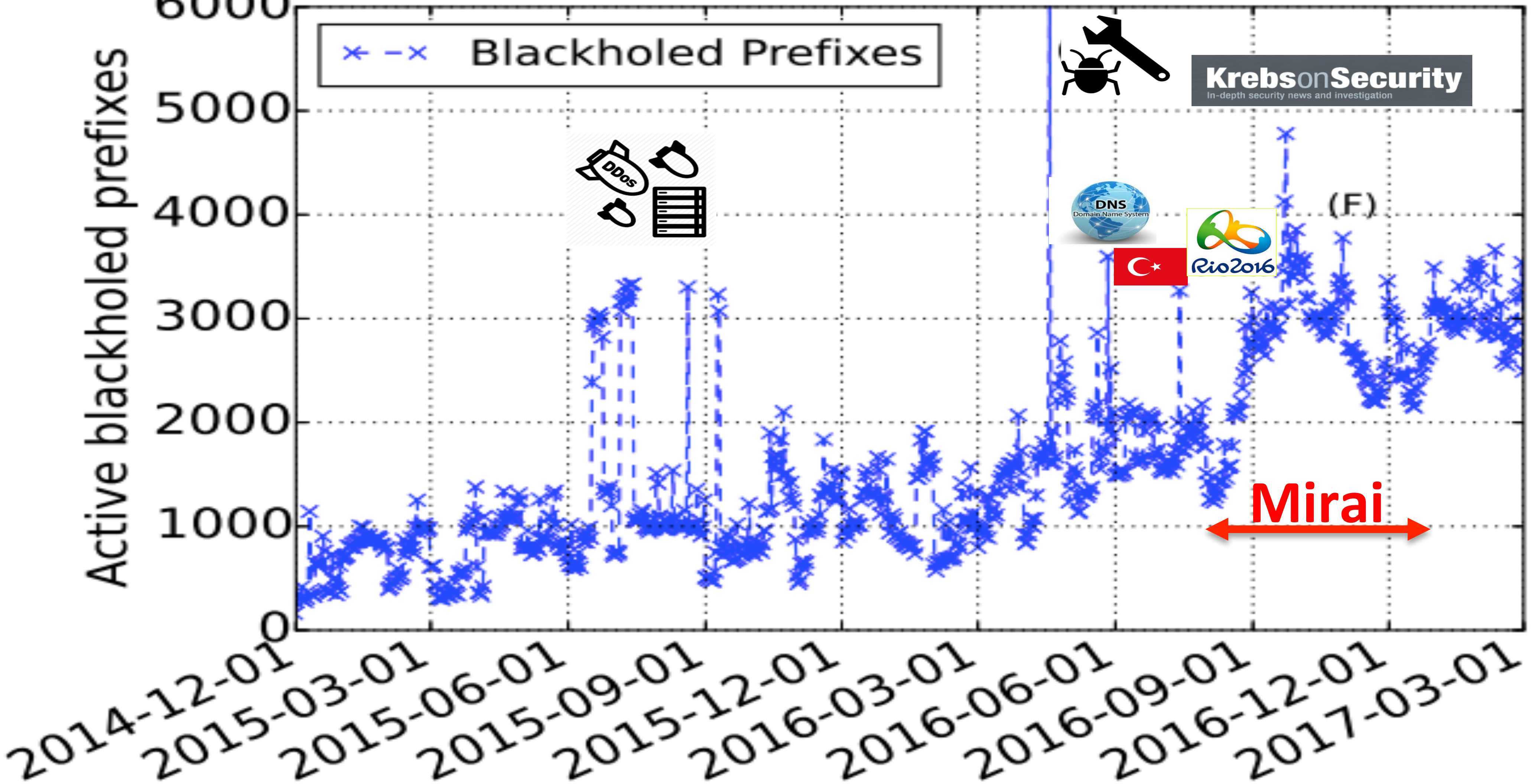


RFC1997, RFC6535, RFC7999

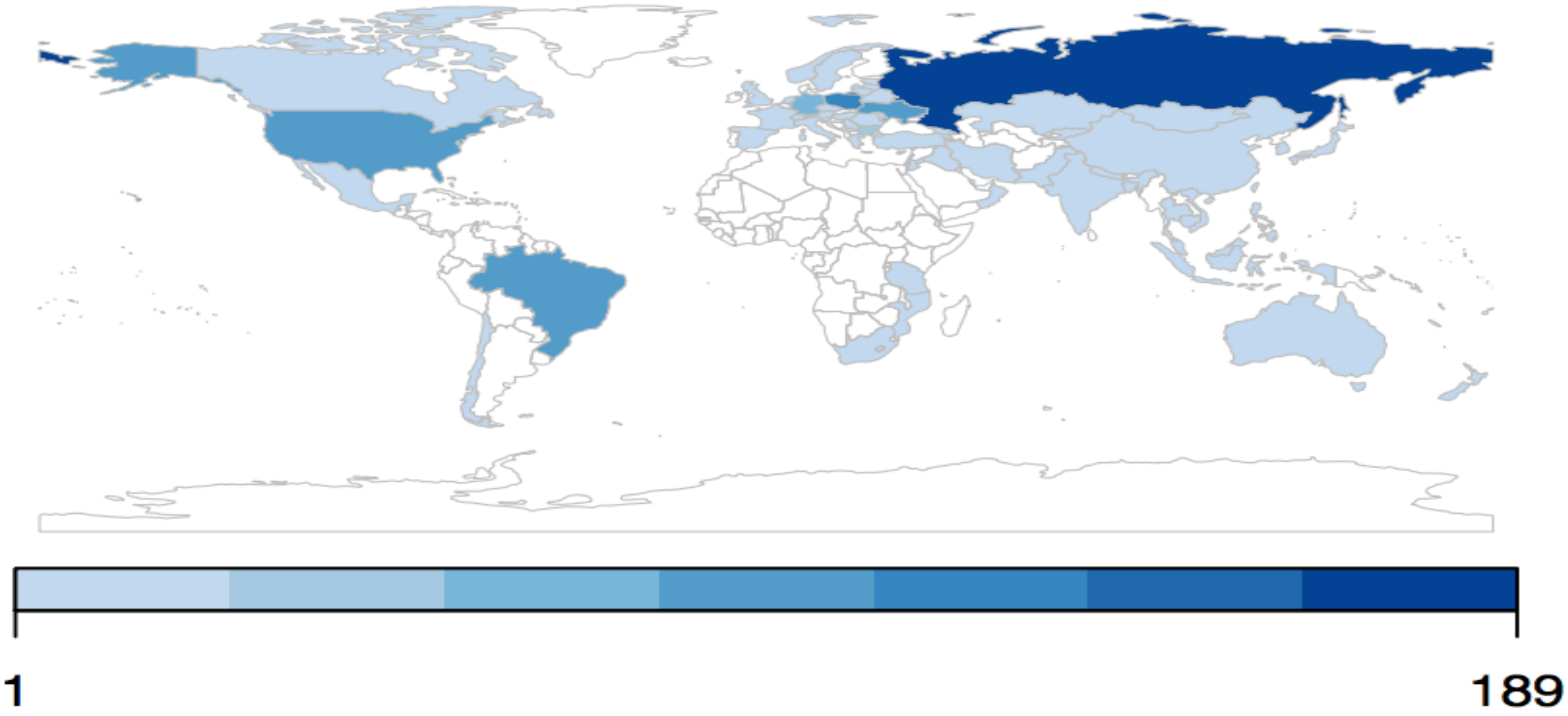
The Rise of BGP Blackholing



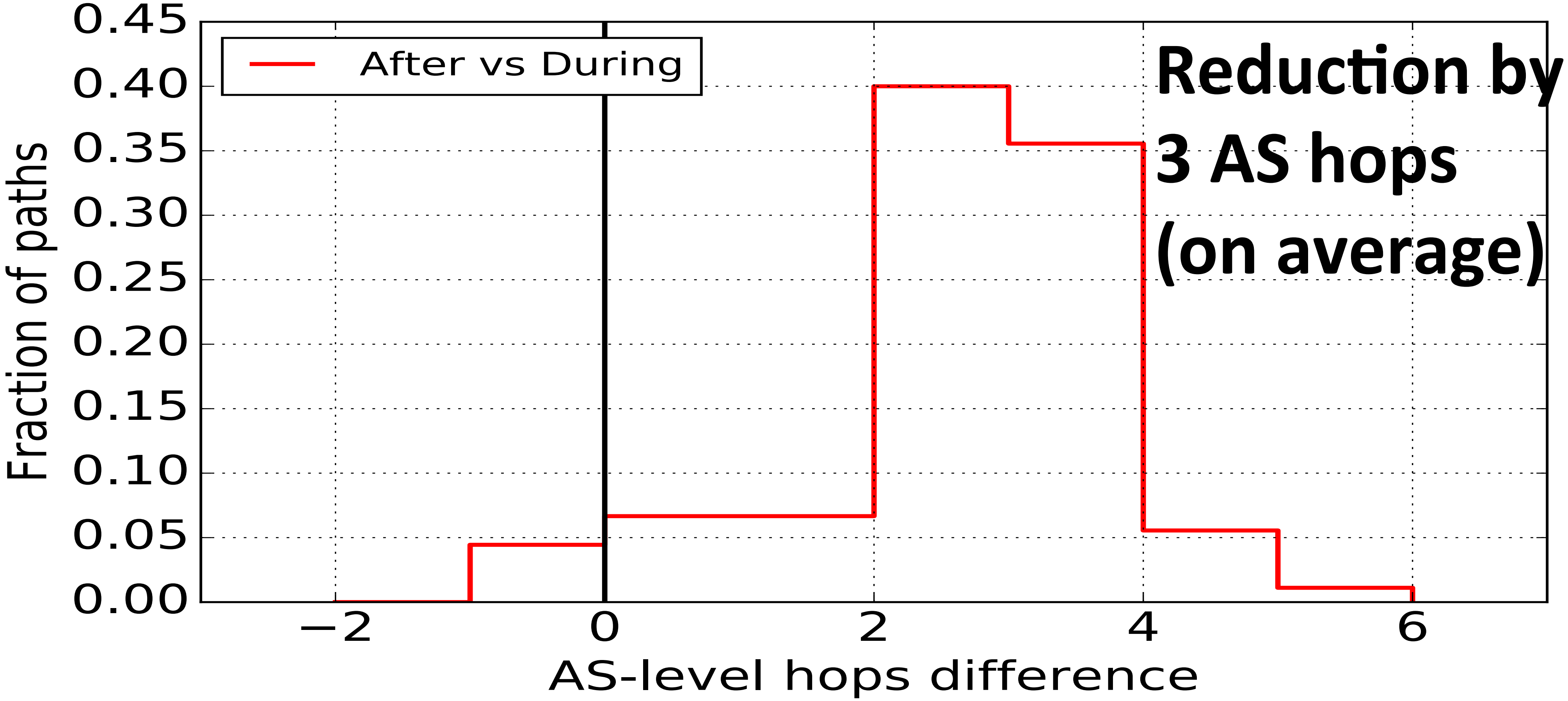
The Rise of BGP Blackholing



Popularity of Blackholing Users



BGP Blackholing Efficacy: Active Measurements



Reduction by 3 AS hops (on average)

~~Cyberattacks and Outages are Serious Threats~~

The New York Times

Hackers Hit Dozens of Countries Exploiting Stolen N.S.A. Tool

Leer en español

By NICOLE PERLROTH and DAVID E. SANGER MAY 12, 2017

KrebsOnSecurity Hit With Record DDoS

esday evening, KrebsOnSecurity.com was the target of an extremely large and unusual buted denial-of-service (DDoS) attack designed to knock the site offline. The attack did cceed thanks to the hard work of the engineers at Akamai, the company that protects my om such digital sieges. But according to Akamai, it was nearly double the size of the t attack they'd seen previously, and was among the biggest assaults the Internet has ever ssed.

SC

SC US · SC UK · NEWS · CYBER-CRIME · NETWORK SECURITY · PRODUCTS · VIDEO · EVENTS · WHITEPAPERS · INSIGHT

THE CYBER-SECURITY SOURCE

SC Media UK > News > ICYMI: 1Tb DDoS attack, Krebs dropped, Pippa Middleton, Yahoo!

by SC Staff

DOI:10.1145/1897852.1897869

Article development led by queue.acm.org

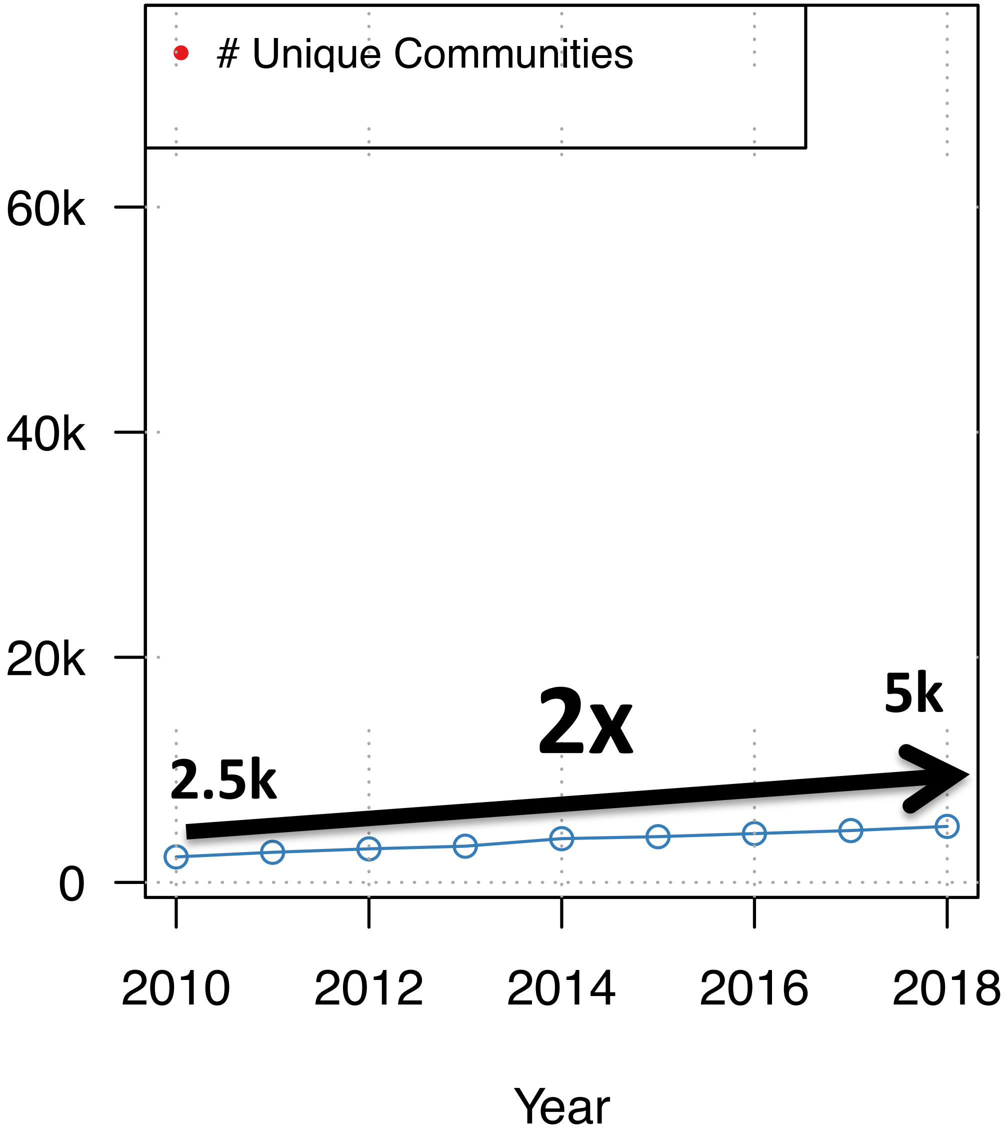
Attacks in Estonia and Georgia highlight key vulnerabilities in national Internet infrastructure.

National Internet Defense— Small States on the Skirmish Line

and commercial activity and influence. This is far less palpable than a nation's physical territory or even than "its air"

Can BGP Communities be Abused?

BGP Communities Usage is on the Rise



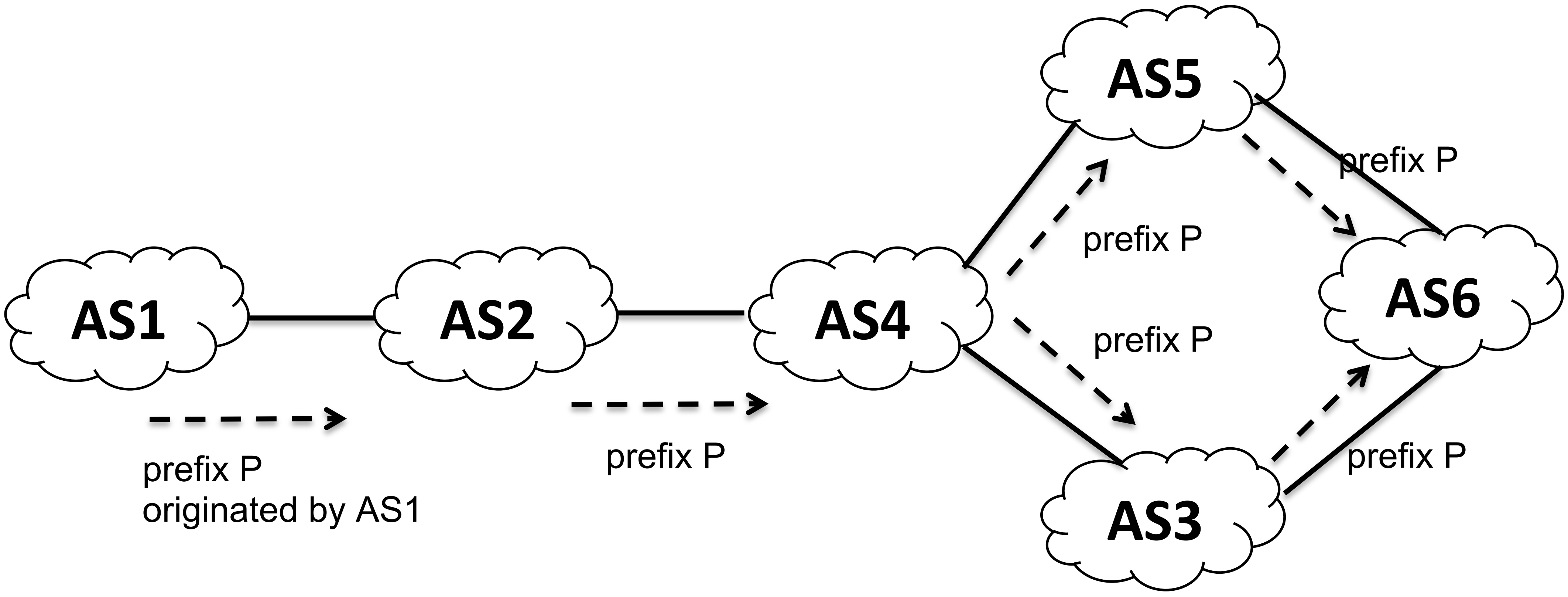
Communities is the Swiss Knife of operators:

- 75% of the BGP announcement have >1 community

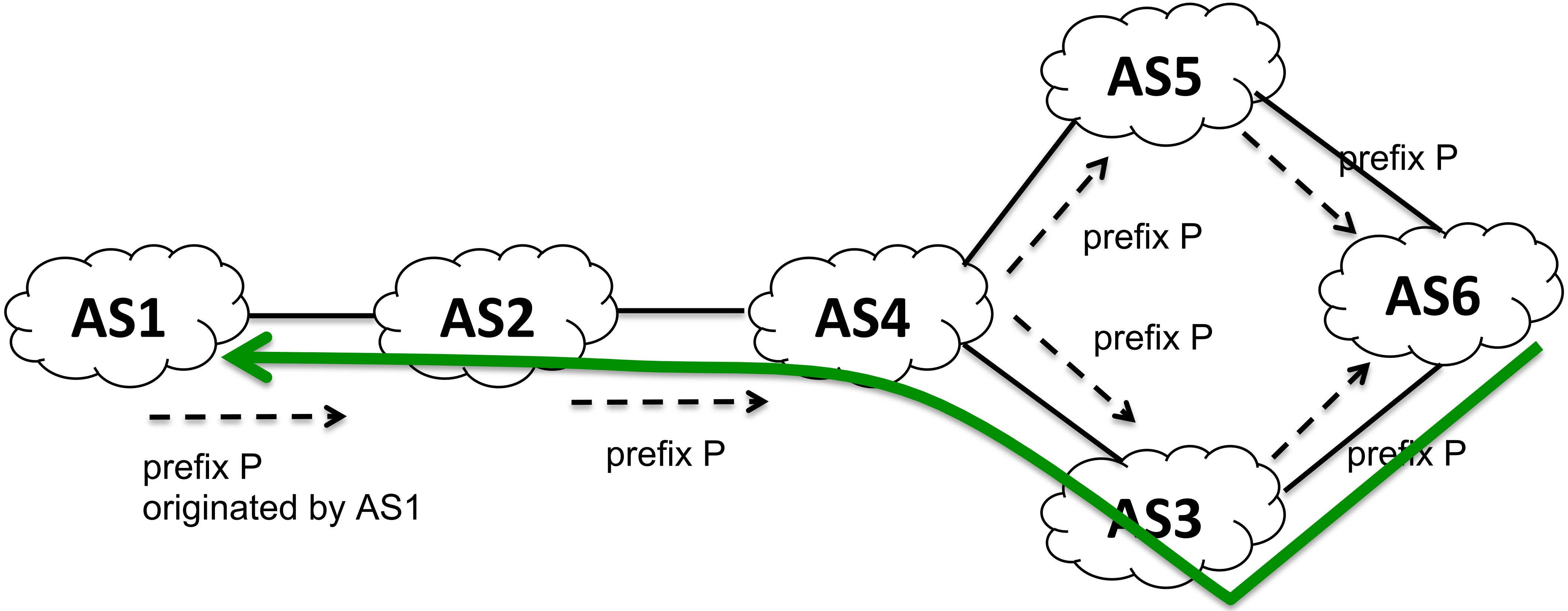
Usage:

- location
- blackholing
- Traffic Engineering: path prepending,
  local preference, selective announcements
- RTT delays

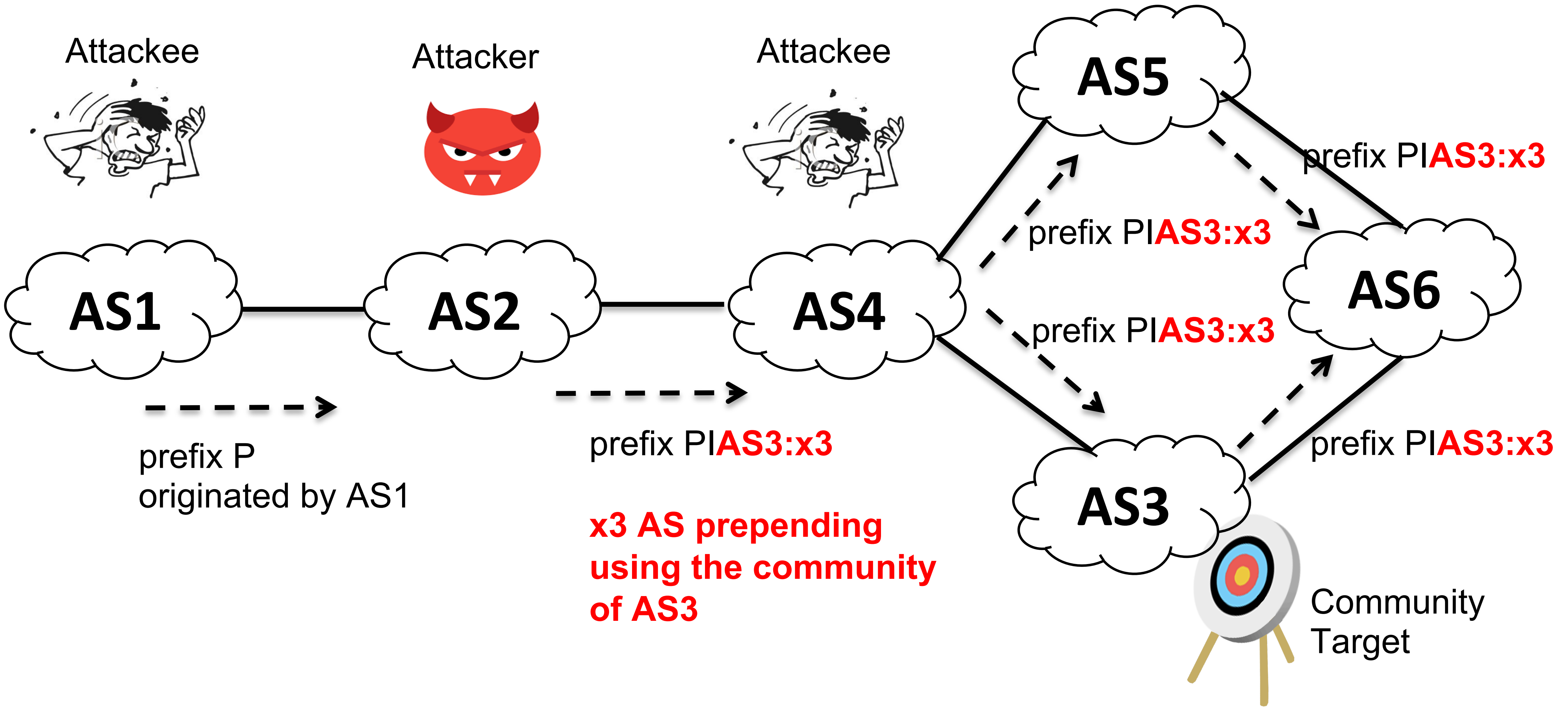
Teaser Example of BGP Communities Attacks



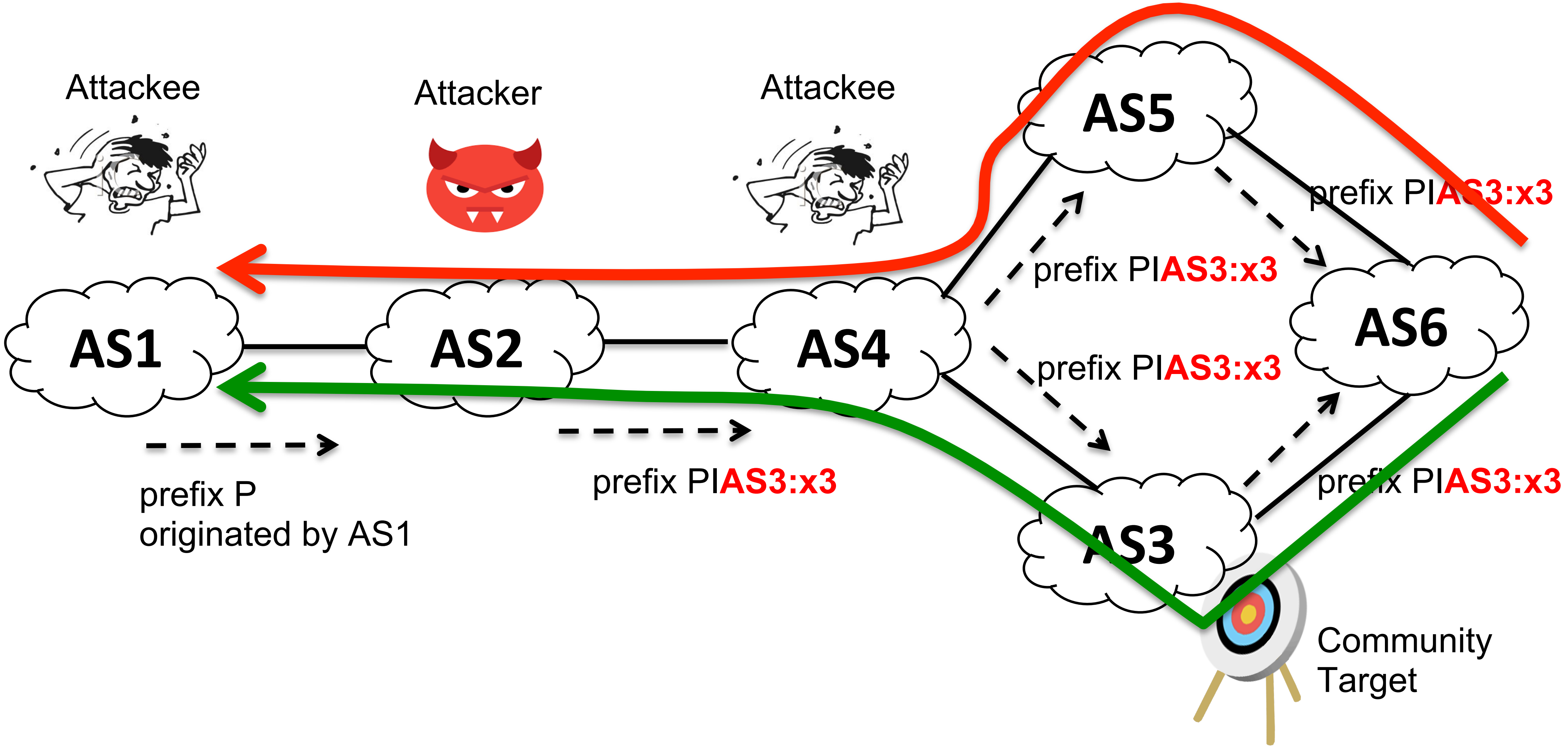
Teaser Example of BGP Communities Attacks



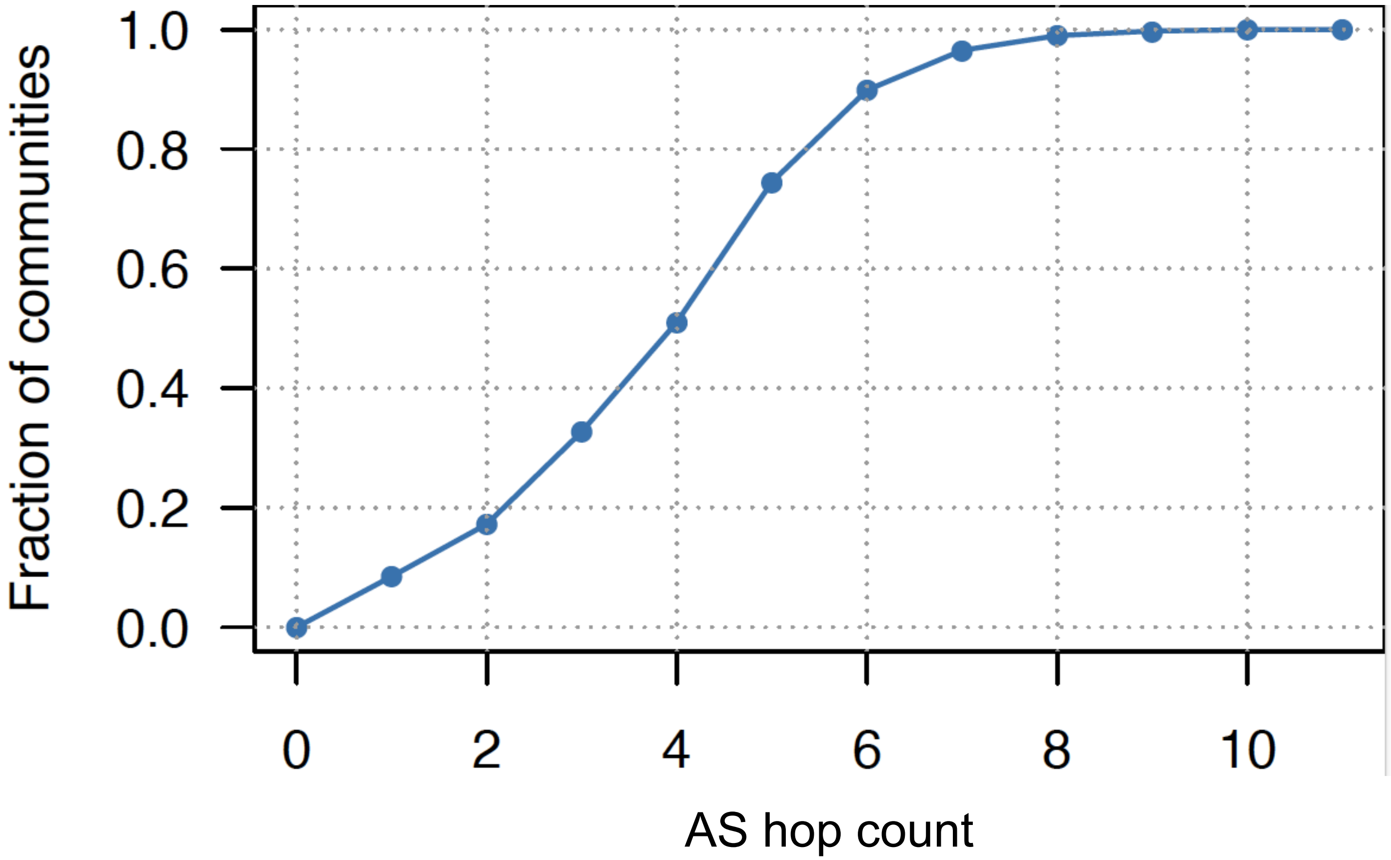
Teaser Example of BGP Communities Attacks



Teaser Example of BGP Communities Attacks

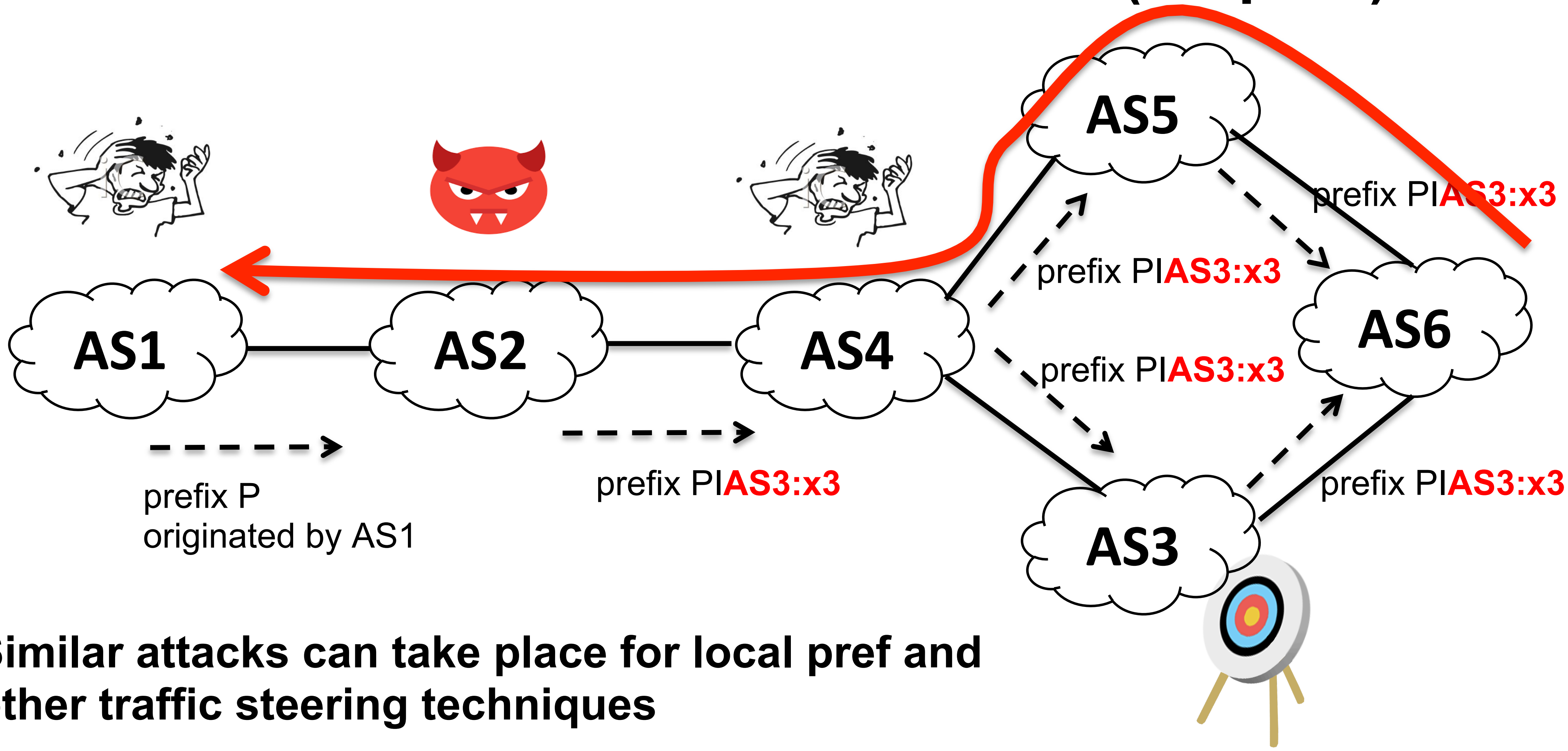


Propagation of Communities (necessary condition)



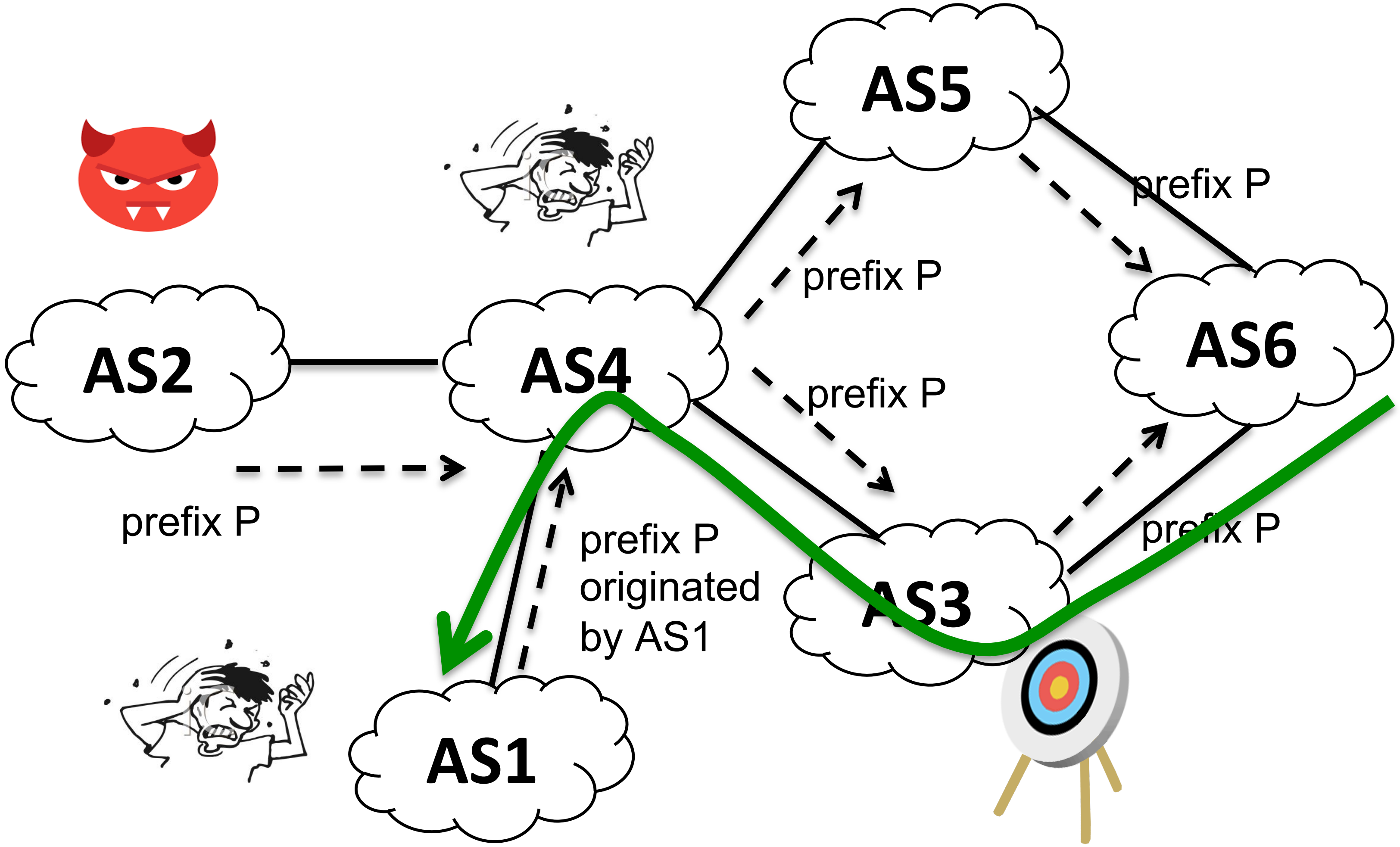
BGP communities is an optional and transitive attribute:
14% of transit provider (2.2K our of 15.5K) propagate communities

AS path prepending Attack without Hijack even if route is authenticated (on-path)

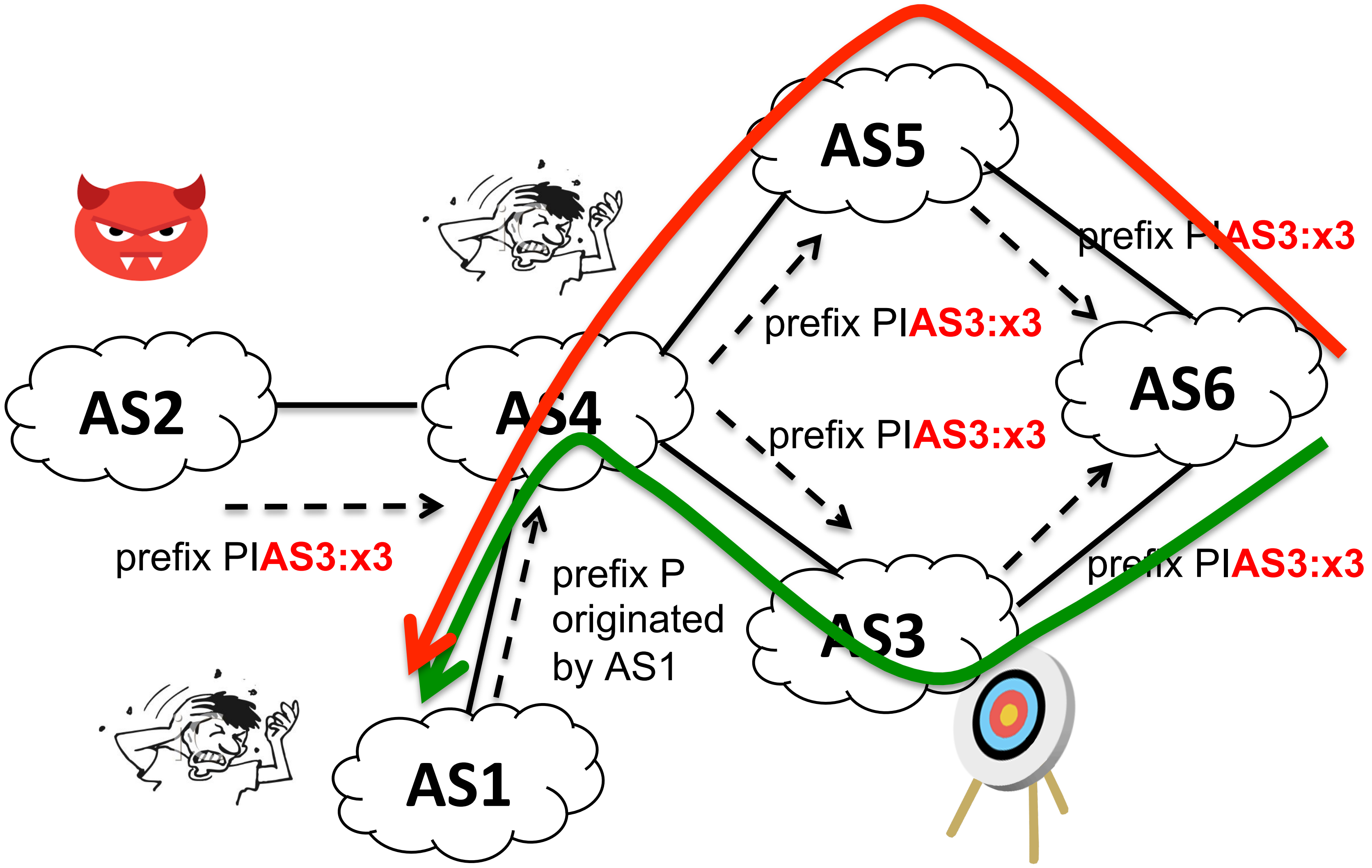


Similar attacks can take place for local pref and other traffic steering techniques

AS path prepending Attack with Hijack (off-path)



AS path prepending Attack with Hijack (off-path)



Experimentation

PEERING
The BGP Testbed

With Ethical Considerations!

Traffic Steering	✓		✓	AS relationship plays a role, IRR is checked (difficult)
Blackholing	✓		✓	Accepted independent of AS relationship, high evaluation order (easy)
Route Manipulation	✓		✓	May have to modify IRR (involved)

cisco

Does not propagate communities by default

JUNIPer NETWORKS

Propagates Communities by default

Order of rules in configuration plays an important role!

Discussion

- Have we gone too far with BGP communities? Propagate **only** communities to the peer, o.w. there is a risk of a global effect
- Need for BGP communities **authentication**
- Be aware of **standardized** BGP communities
- Need for proper BGP communities **documentation**
- **Monitor** the hygiene and propagation of BGP communities usage

Conclusion

- BGP communities is on the rise and provide a unique, yet **unexplored** source of information about the **State** and **Health** of the Internet

- BGP communities are increasingly **popular** to cope with **complex** operational taks

- We showcase:

- How to use BGP communities to detect **peering infrastructure outages** and assess their impact

- How to use BGP communities as a proxy to infer **attacks** and **mitigation strategies**

- Assess **vulnerabilities** due to the abuse of BGP communities abuse

Thank you!

Published papers supported by ERC StG ResolutioNet:

“Detecting Peering Infrastructure Outages in the Wild”, ACM SIGCOMM 2017

“Inferring BGP Blackholing Activity in the Internet”, ACM IMC 2017

“BGP Communities: Even More Worms in the Routing Can”, ACM IMC 2018