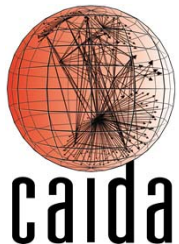# Revealing the load-balancing behavior of YouTube traffic of interdomain links

Ricky K. P. Mok[+]
Vaibhav Bajpai*, Amogh Dhamdhere[+], kc claffy[+]
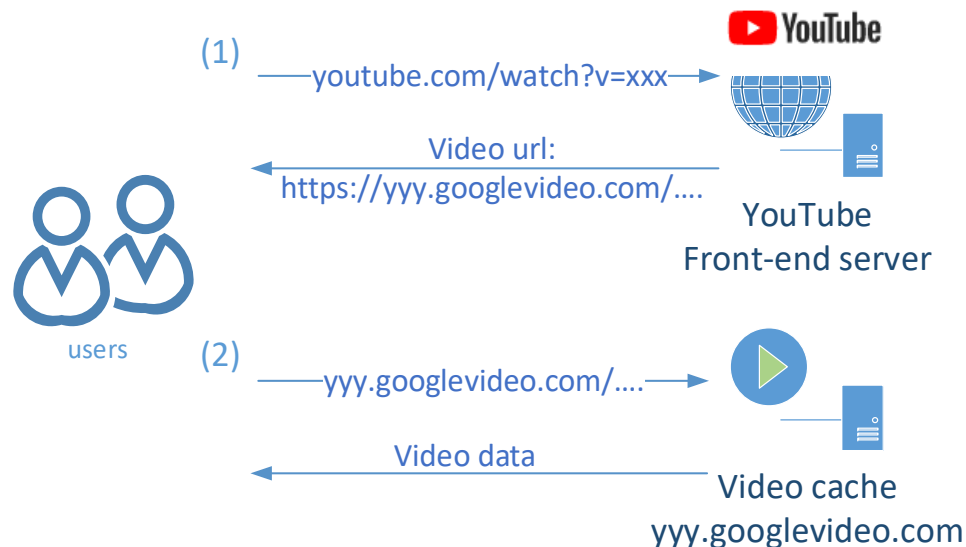[+]CAIDA/ University of California San Diego
* Technical University of Munich

1

YouTube Streaming Background

(1) Front-end server (youtube.com) dedicates users to a video cache after they selected a video

(2) Video caches are in the namespace of *.googlevideo.com

Connecting with YouTube

Three common approaches

1. Install Google Global Caches (GGCs) and peering with Google

2. Access via transit providers
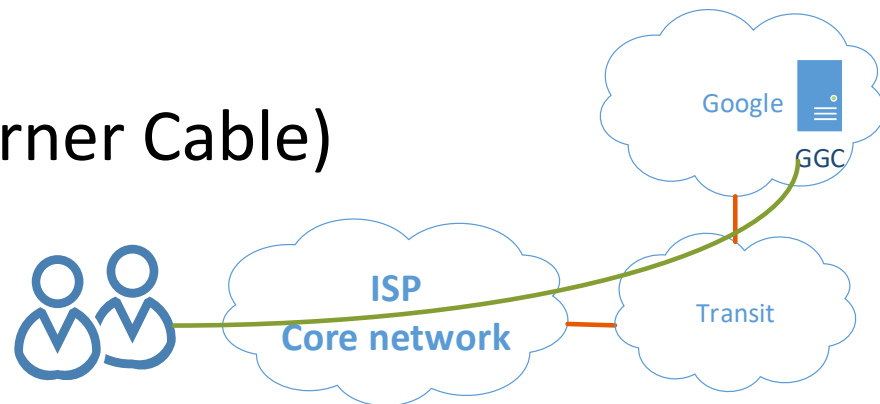
3. Peering with Google AS

GGC option

- Request Google to install the GGCs **inside** the ISP network

- Users mainly stream videos from GGCs
  - Still peering with Google for handling cache misses/loading new data from Google

- E.g., Cox, Verizon

Google AS   Interdomain link

ISP
Core network

GGC
@NYC

GGC
@LA
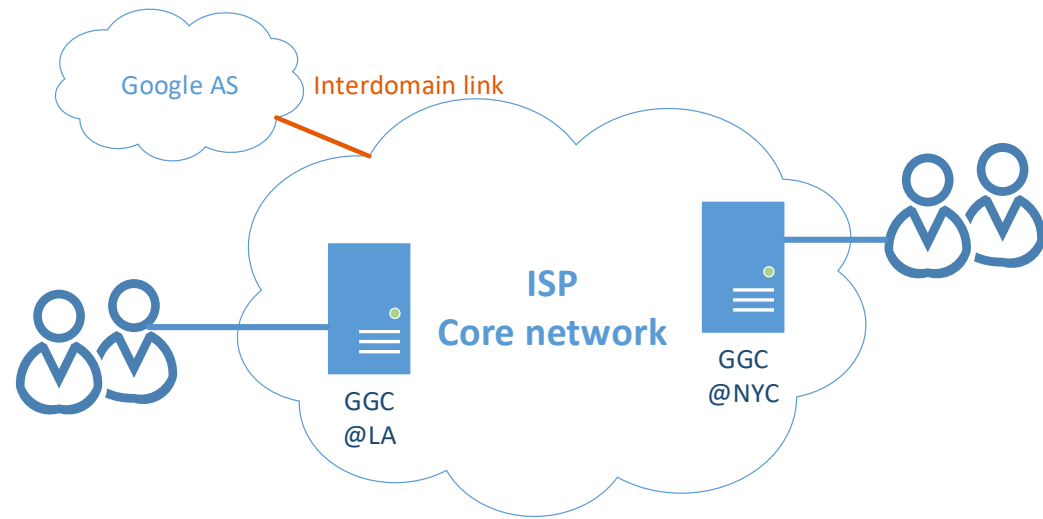
Transit provider

- Access ISP does not form direct peering with Google

- Relies on transit providers (e.g., TATA, Level 3) to connect with Google
  - Just like other public Internet destinations

- E.g., Spectrum
  (former Time Warner Cable)

Peering with Google

- Some ISPs do not prefer introducing GGCs into their network
    - IPv4 address space
    - Rack space
    - Physical security (24x7 access of the site)
- All traffic traverses the interdomain links peering with Google at IXPs
- E.g., AT&T, Comcast, CenturyLink

Google

GGC

Interdomain link

ISP
Core network
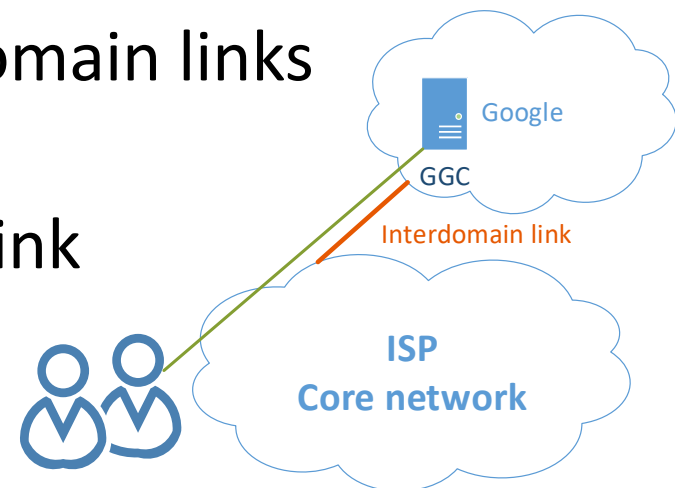
Our research questions

- Investigate how ISP distribute users into different interdomain links

- Did Google play a role in selecting interdomain links?

- Performance implication

Measurement Overview
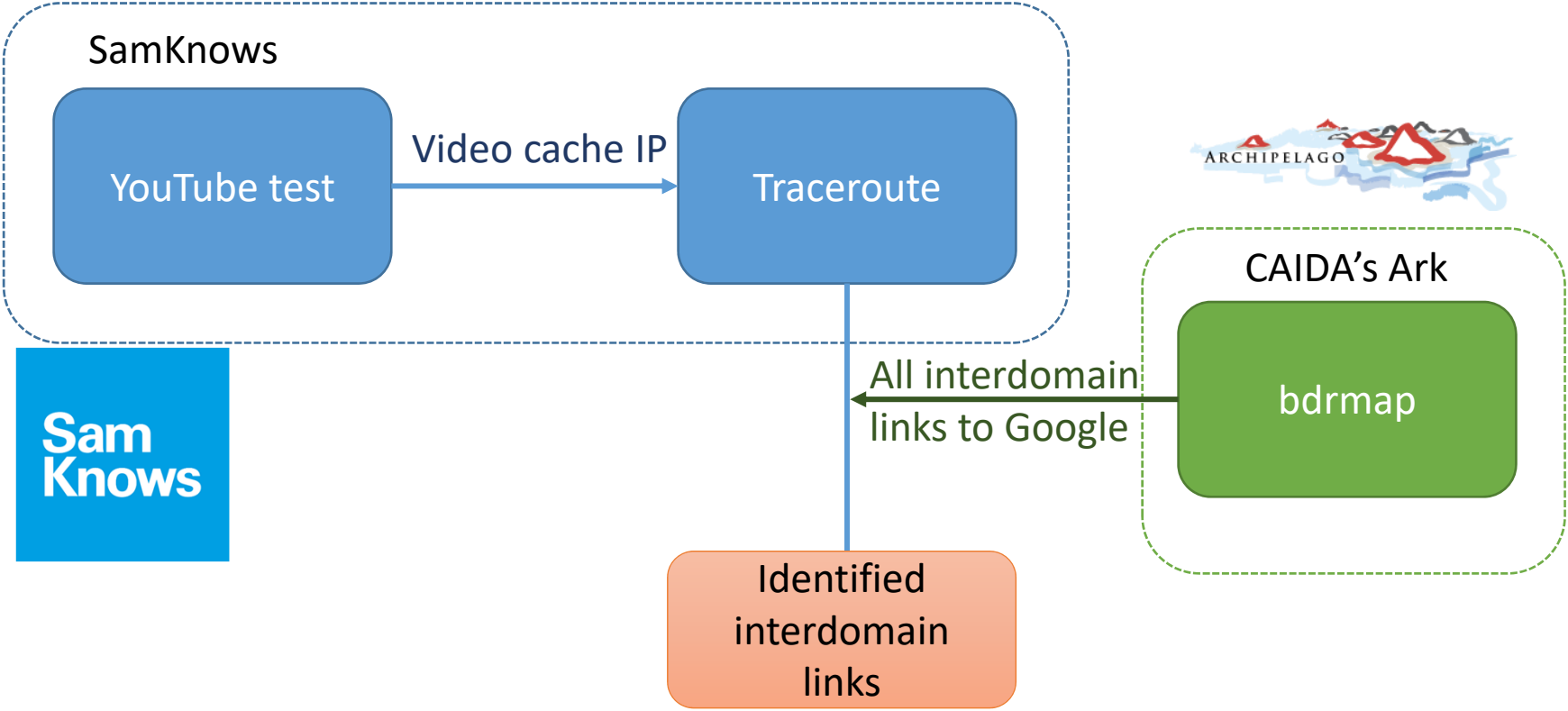
YouTube Measurement

- YouTube test
  - Streams a popular video as a normal YouTube viewer
  - Reports video cache hostname/IP, streaming performance

- Traceroute
  - toward the video cache IP (immediately after YouTube test)

Discovering interdomain links

- bdrmap [Luckie16IMC] is used to infer the links
  - Runs on CAIDA's Ark probes
- Each link is represented by two IPs (Near-side and Far-side)

Interdomain link

ISP network *Near* — *Far* Google
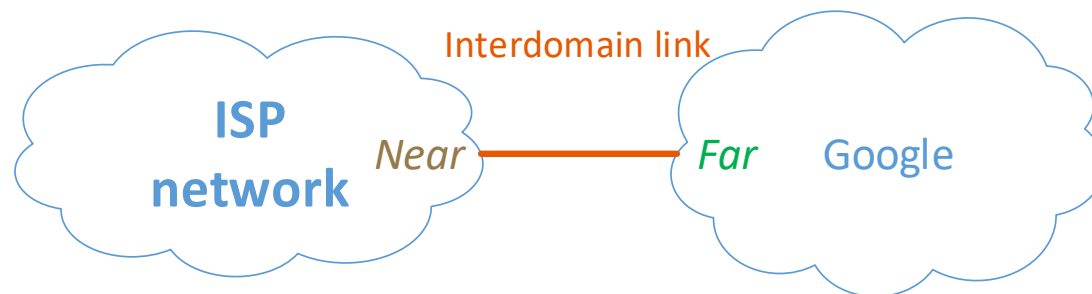
[Luckie16IMC] M. Luckie, A. Dhamdhere, B. Huffaker, D. Clark, and kc claffy.
bdrmap: Inference of borders between IP networks. In ACM IMC, 2016.

Identifying interdomain links

- Match the traceroutes with all interdomain links to Google (AS 15169/ AS 36040/ AS 43515)

```
Hop        IP
…
 5   137.164.11.24
 6   74.125.49.165
 7   108.170.247.225
 8   209.85.242.59
…
```

Interdomain Links
Near: 74.125.49.165
Far: 108.170.247.225

- Aggregate with Far-side IP
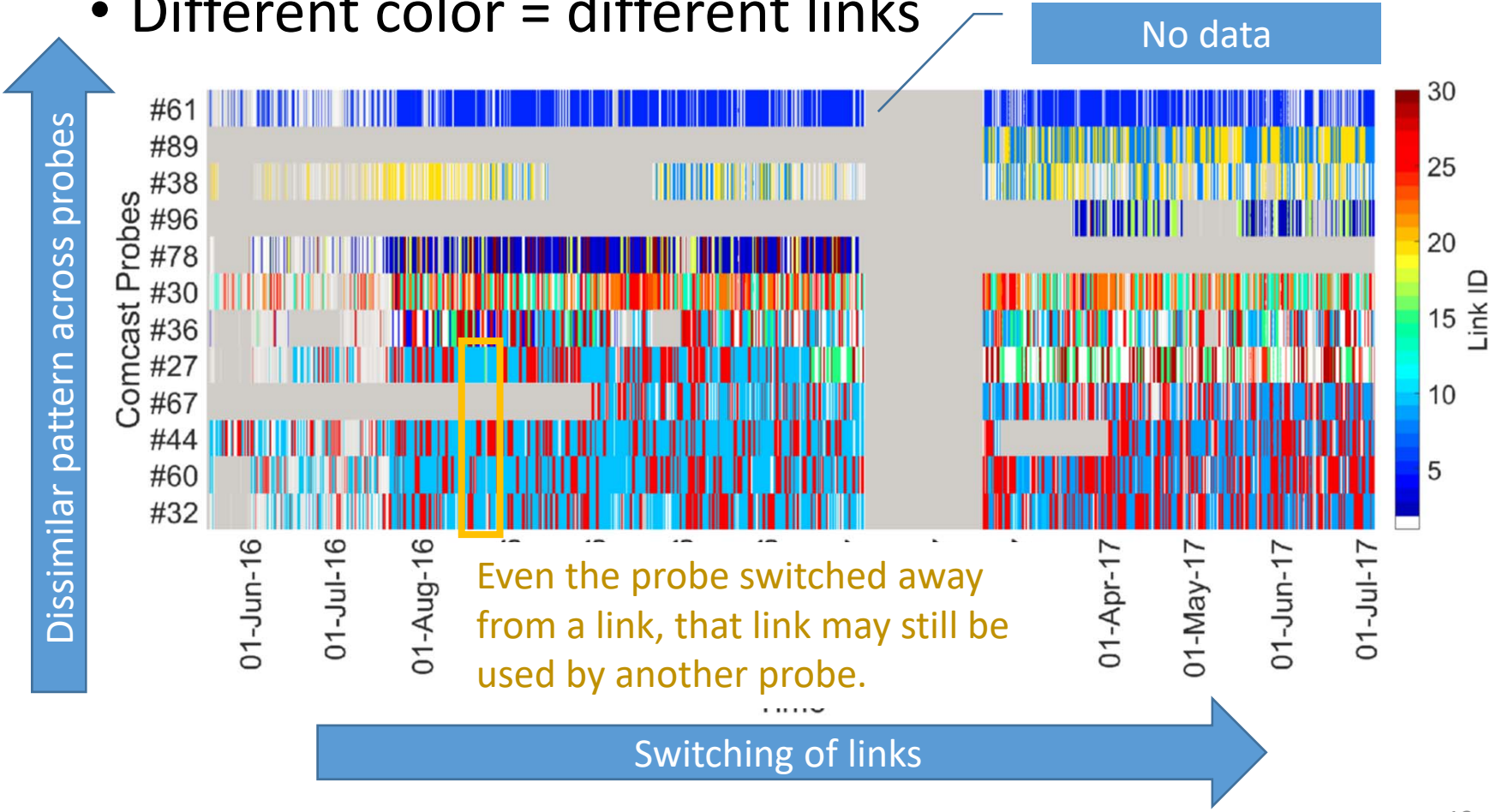 - IP alias on the near-side router

Data

- ~ 1 year of YouTube data/traceroute
  - May 2016- July 2017
- 74,000 YouTube measurement/traceroute
  - ~100 SamKnows boxes (home router)
- Identified 45 interdomain links
  - 15 SamKnows boxes
  - 4 ISPs (1 US, 1 DE, 1 IT, 1 FR)

ISP	# of Google links
Comcast (US)	26
Kabel (DE)	5
Italia (IT)	10
Free (FR)	4

Links - Comcast

- Different color = different links

No data

Even the probe switched away from a link, that link may still be used by another probe.

Dissimilar pattern across probes

Switching of links

Links – European ISPs

- No cross-ISP links

Non-identified link

Host/ISP switched
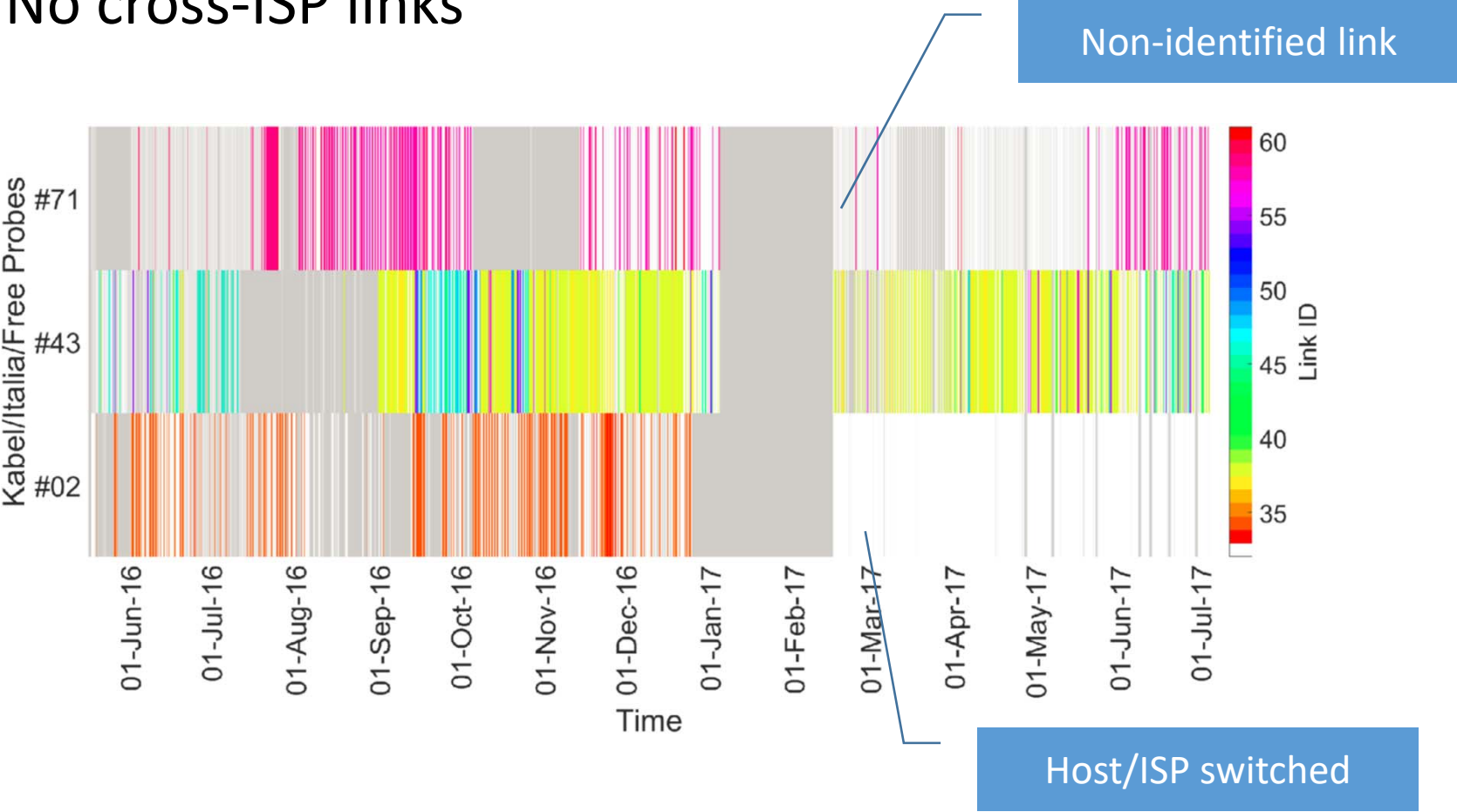
Non-identified links

- bdrmap and the YouTube traceroute ran from different vantage points (VPs)

- Even though in the same ISP, some links may not be observed in some VPs.

- Depends on coverage of Ark

- Comcast (83.4%) *vs.* Free (40.2%)

Quantifying the use of links

- Compute the probability of a probe using each link

$$P_l^b = \frac{\text{\# of observations on link } l \text{ from probe } b}{\text{\# of traceroute from probe } b \text{ with identified links}}$$
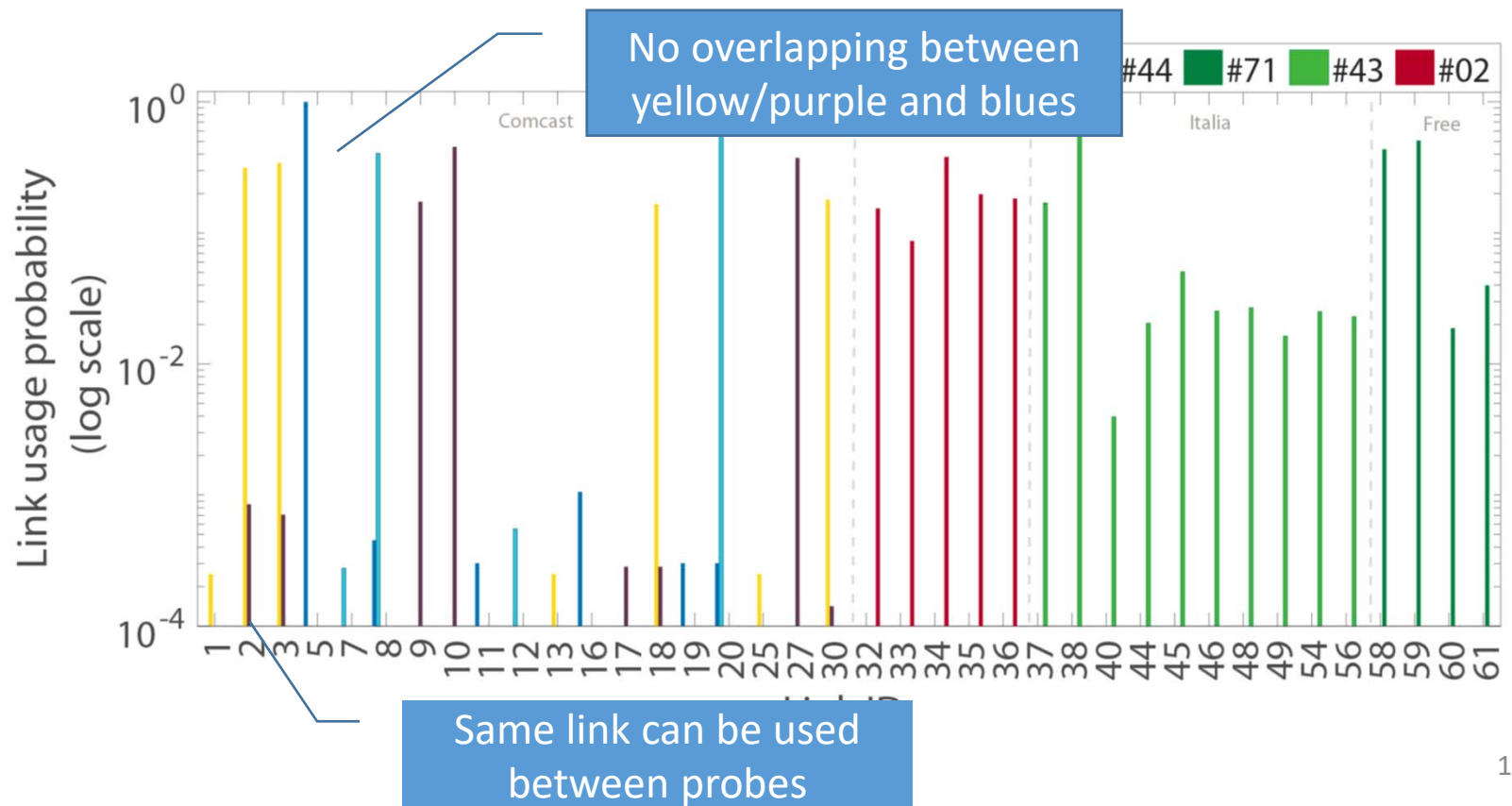
- We can obtain a feature vector for each probe

$$\vec{P^b} = <P_1^b, P_2^b, ...>$$

Overall use of links

- At least two links were used.
  - 10 links for the probe connected to Italia

Same link can be used between probes

No overlapping between yellow/purple and blues

Questions

- Why in the same ISPs, some probes show "share use" of links, some are not?

- For the same probe, how ISP diverts the traffic to different links?

- Spatial
  - Geographic location
- Temporal
  - Time-of-the-day

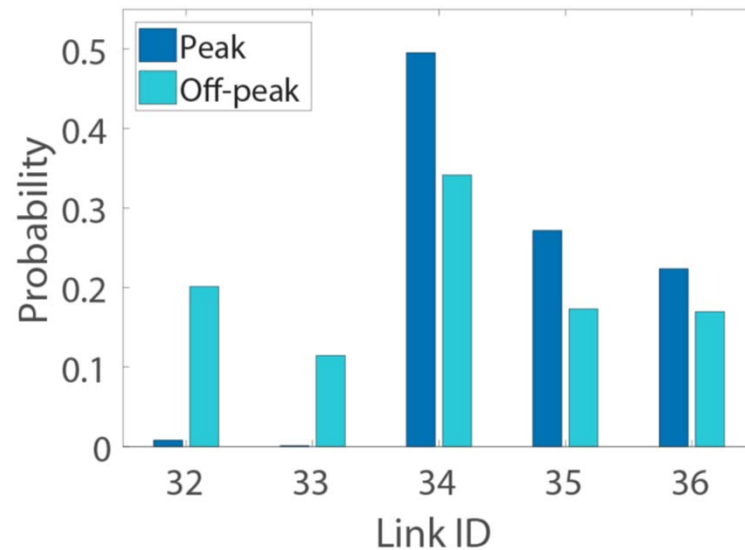# Recall the case of Comcast
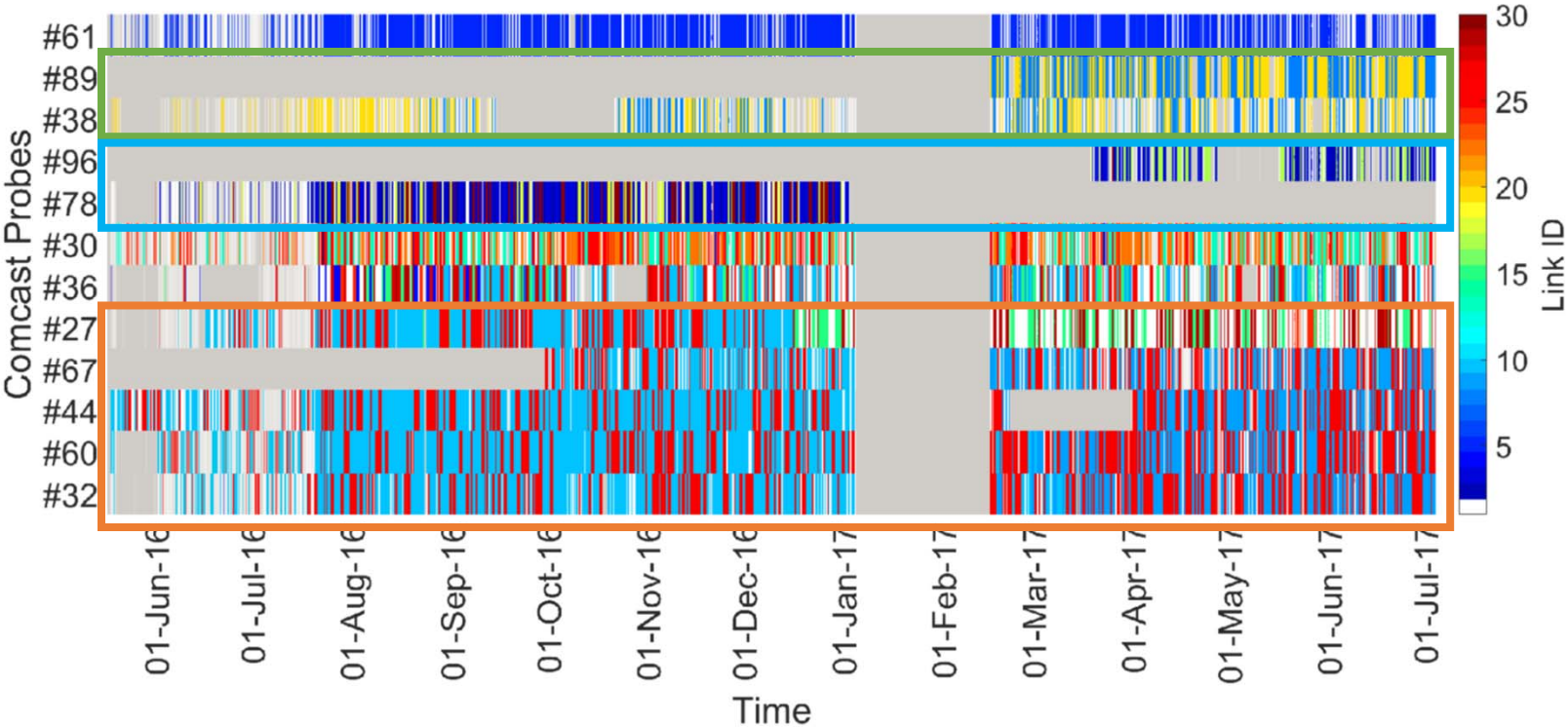
Geographic location

- Perform agglomerative hierarchical clustering on the feature vectors $\overrightarrow{P^b}$ of all Comcast probes
- The results reflect the physical location of the probes
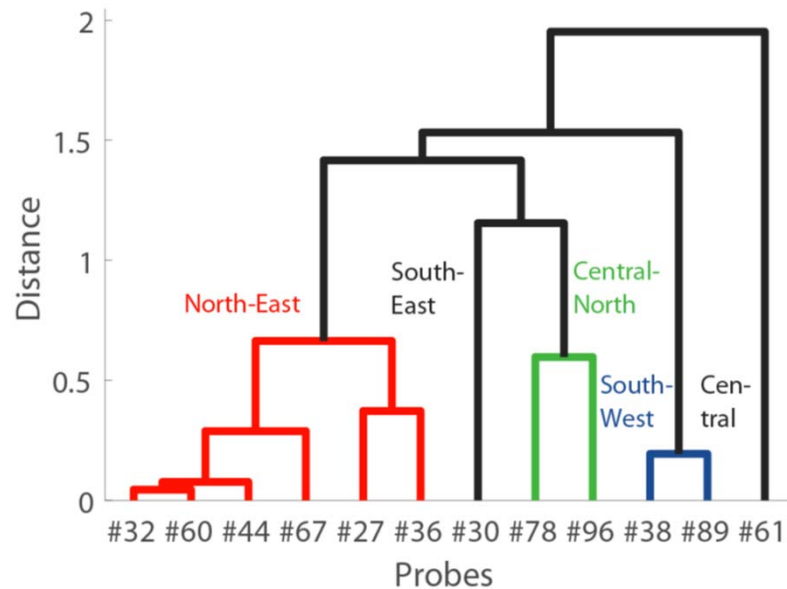
Time of the day

- Subset the traceroutes according to the time of the day
  - Peak hours (7pm – 11pm local time)
- Kabel uses two more links during the non-peak hours.

Destination Google ASes

- Currently, ISP can peer with Google using [google]
  - AS 15169 (All Google/YouTube content, primary option)
  - AS 36040 (Google's most popular content)
  - AS 43515 (Internal use, not available for public peering)

- European ISPs were still heavily rely on AS 36040 for YouTube content
  - Historical reason

Percentage of YouTube test sessions streaming from 3 Google ASes

Google AS	Comcast	Kabel (DE)	Italia (IT)	Free (FR)
15169	99%	76%	16%	4%
36040	1%	24%	84%	94%
43515	0%	0%	0%	2%

[google] https://peering.google.com/#/options/peering

Destination Google ASes

- Why ISP Free can still connect to AS 43515?
  - The prefix of the video cache's IP was announced by AS 43515
  - Streaming was also success
  - Now, the hostnames are no longer resolvable, and the IPs do not respond to ICMP ping.

→ Google recently changed the function of this AS

Google AS	Free (FR)
15169	4%
36040	94%
43515	2%

Destination Google ASes

- The use of destination AS also related to time of the day.

AS 36040   AS 15169
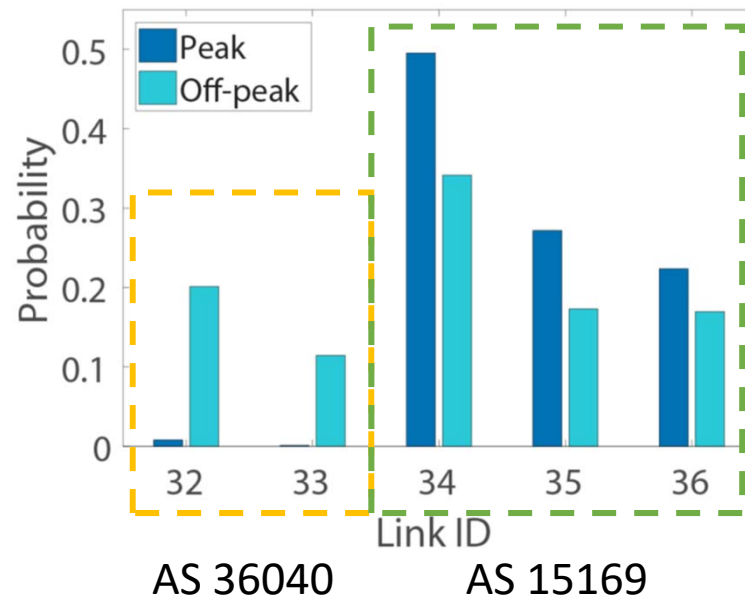
Conclusion

- ISPs distributed traffic to multiple interdomain links to Google
  - Geographic location
  - Time of day
- Different Google ASes are used to stream videos

In the paper and beyond

- YouTube video cache selection mechanism has limited influence on which interdomain links to be used
  - Overlapping cache hostname/IPs
  - Multiple interdomain links can access to the same cache
- Streaming performance implications
- Data will be available on CAIDA website (`caida.org`)

Sign up for CAIDA's Ark probe!
✓YouTube

Thanks!

cskpmok@caida.org

standard micro-USB phone charger

Raspberry Pi + SD card

HDMI to DVI-D display adapter