Internet Measurement
& Analysis (IMA)

Internet Measurement

Summer Semester 2019

Prof. Georgios Smaragdakis, Ph.D.

Instructor

Prof. Georgios Smaragdakis, Ph.D.

georgios@inet.tu-berlin.de

http://www.smaragdakis.net

Office: MAR 4.030
(please send me an email to arrange an appointment)

Requirements

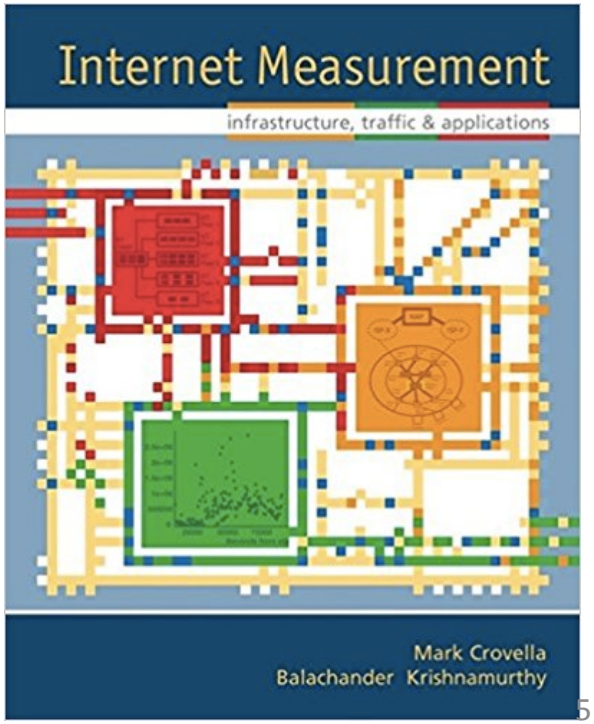
- BSc / MSc level
- Good understanding of basic Networking principles, e.g., Network Protocols and Architectures (prerequisite), Internet Control Plane.

Mission

- To understand how the Internet really works and how it can be hacked, protected, and improved using real network measurements, i.e., data you collect or analyze.

Readings

- All the course material will be available in ISIS
  - Papers from the Internet Measurement literature
  - "Internet Measurement: infrastructure, traffic, and applications" M. Crovella and B. Krishnamurthy



Lecture structure

- Lectures (please participate!)
- Practical exercises in class

  (how to use tools and analyze data)
- 2-3 invited lectures lectures by experts in Internet Measurements
- First invited lecture:
  - **Walter Willinger**, Niksun Inc.,

    May 16, 2pm – 4pm

Exam

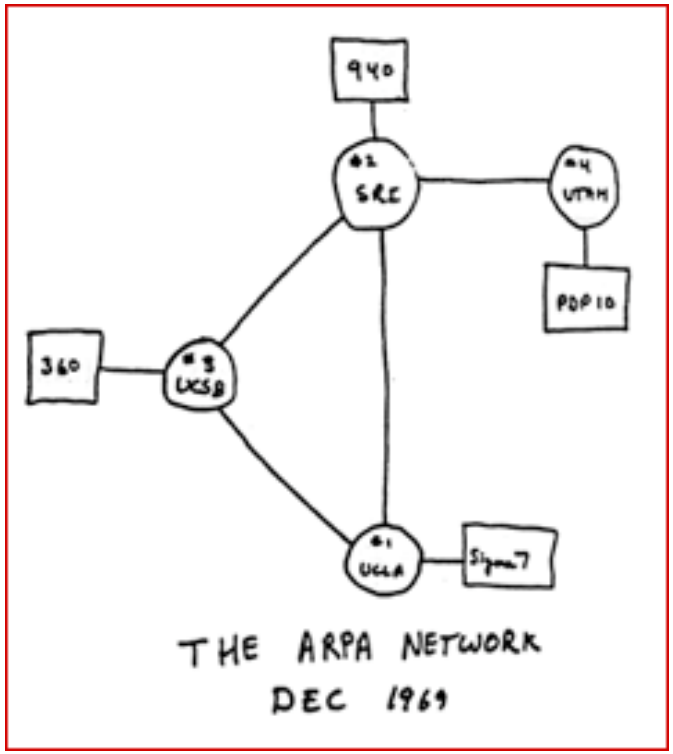
- Exam after the end of the semester
  (oral exam for the Small/Large module)

- If you are a visiting student, e.g., Erasmus student, please contact me!

Why Measure the Internet?

Why Measure the Internet?

- The Internet is not any more a small academic network!



The (Academic) Internet Growth

- Basic Principles

(end-to-end)

END-TO-END ARGUMENTS IN SYSTEM DESIGN

J.H. Saltzer, D.P. Reed and D.D. Clark*

M.I.T. Laboratory for Computer Science

- Inter-networking

(OSI, initial protocols)

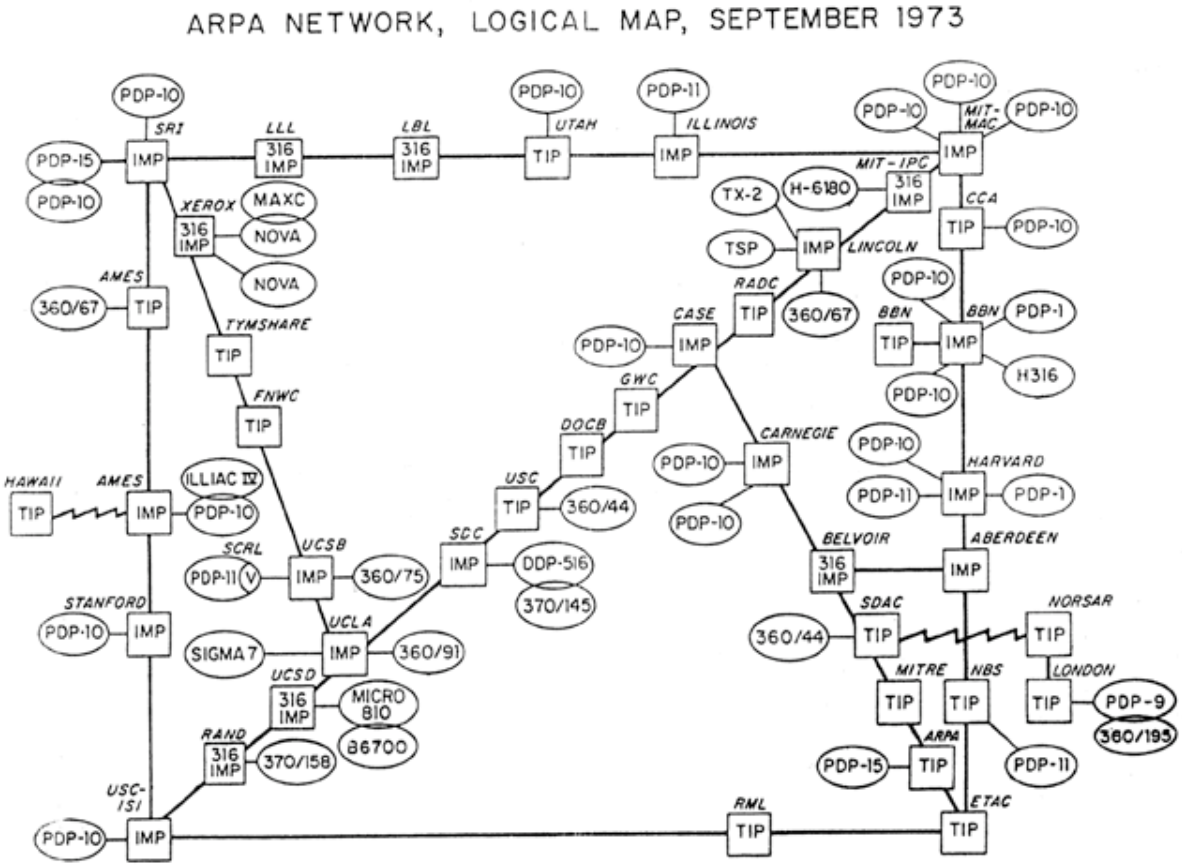
- First nodes outside

Continental US

- First scalability

and operational issues

(congestion)

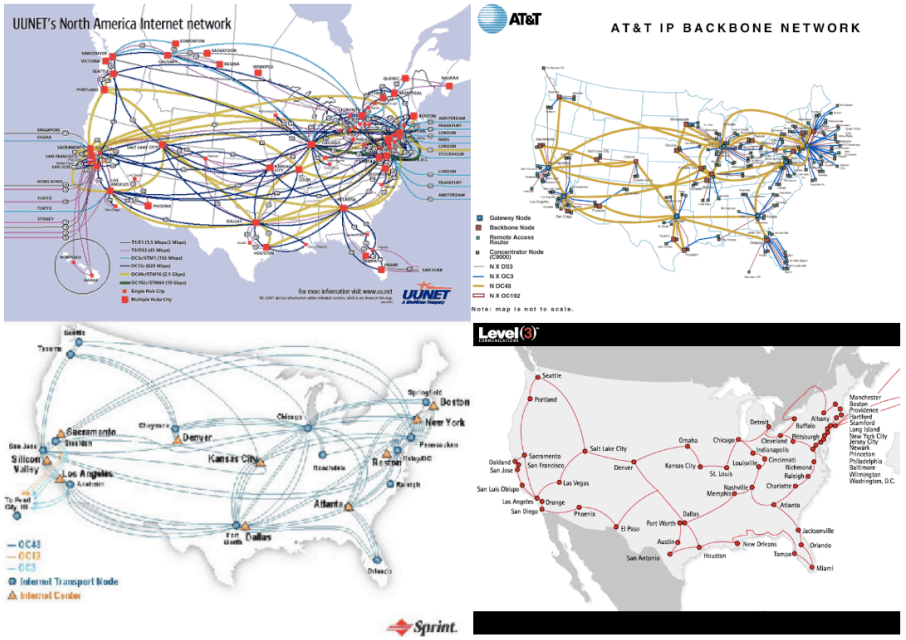


Commercialization of the Internet 1980s-1990s

- Congestion control (TCP/IP)
- Security issues (Morris worm)
- Network management and economics

https://www.vox.com/a/internet-maps

- Researchers start measuring the Internet



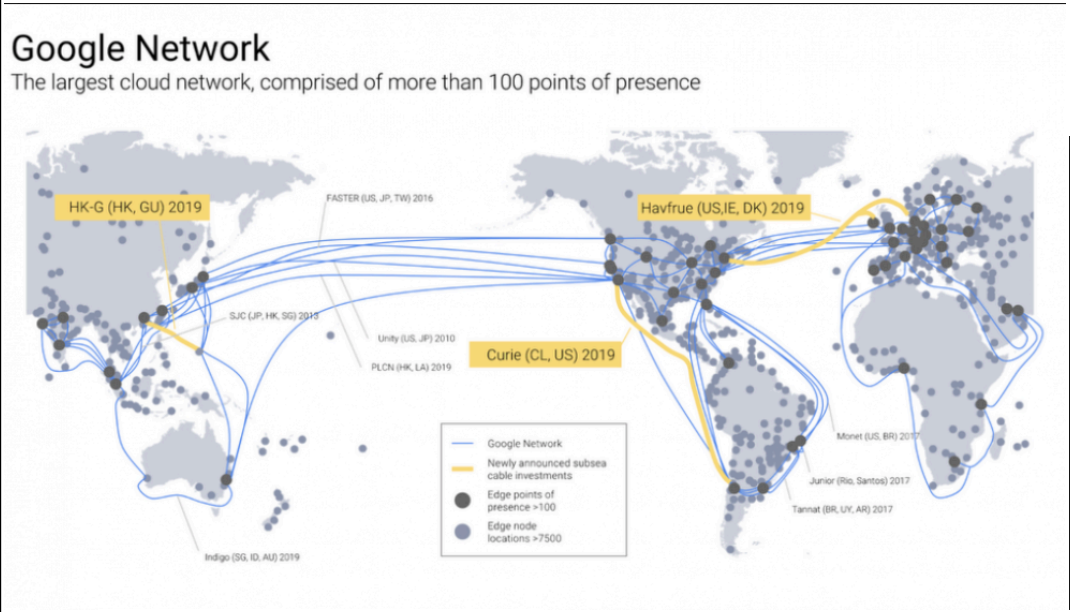
End-to-End Internet Packet Dynamics

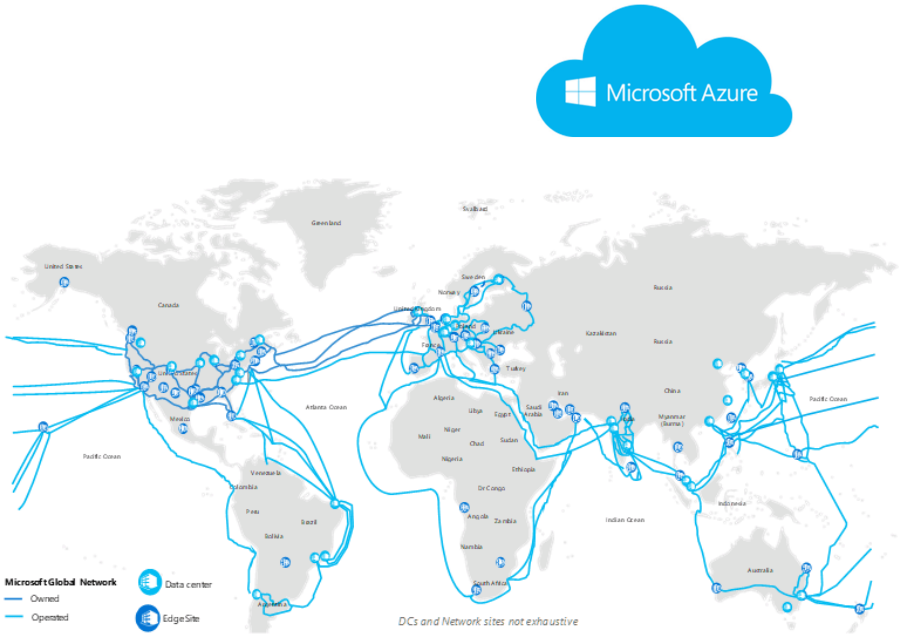
Vern Paxson
Network Research Group
Lawrence Berkeley National Laboratory*
University of California, Berkeley
vern@ee.lbl.gov

LBNL-40488

June 23, 1997

The Internet is the Substrate of the Web



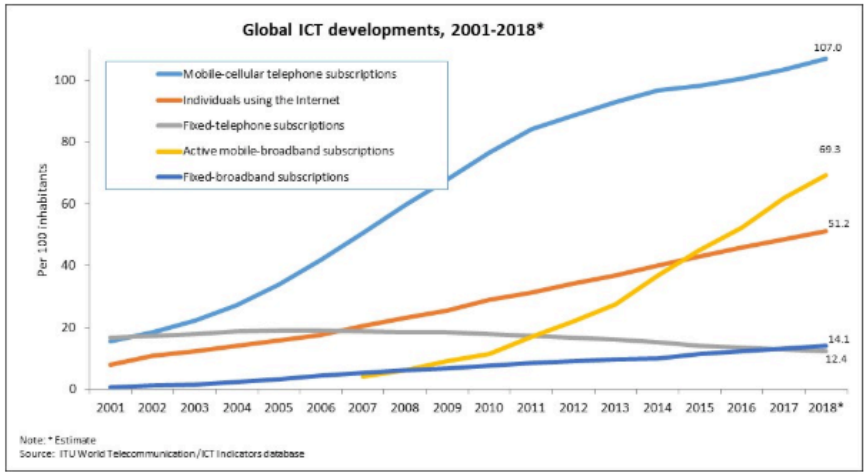


Akamai

Network Deployment: Akamai has deployed the most pervasive, highly-distributed content delivery network (CDN) with approximately 239,000 servers in 139 countries and nearly 1,600 networks around the world.

Why Measure the Internet?

- Internet and increasingly Mobile Internet is the becoming the de-facto means for communication, commerce, information, and entertainment
- How many servers/IoTs/Smartphones are connected to the Internet?



A Multi-perspective Analysis of Carrier-Grade NAT Deployment

Philipp Richter[1], Florian Wohlfart[2], Narseo Vallina-Rodriguez[3], Mark Allman[3], Randy Bush[5], Anja Feldmann[1], Christian Kreibich[3,6], Nicholas Weaver[3], Vern Paxson[3,4]

[1]TU Berlin, [2]TU München, [3]ICSI, [4]UC Berkeley, [5]Internet Initiative Japan, [6]Lastline

ZMap: Fast Internet-Wide Scanning and its Security Applications

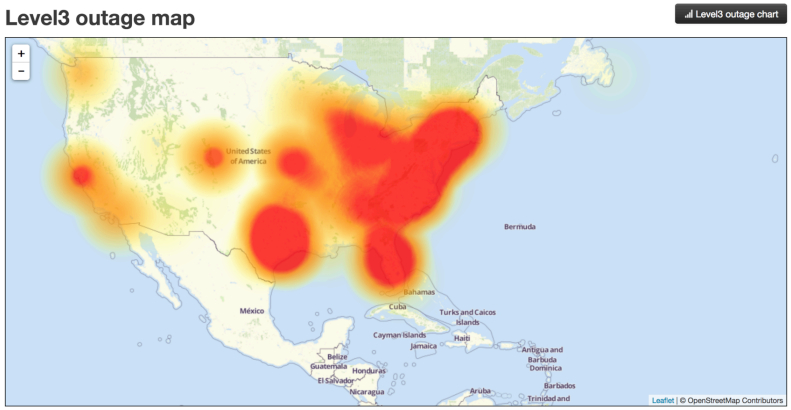
Zakir Durumeric
University of Michigan
zakir@umich.edu

Eric Wustrow
University of Michigan
ewust@umich.edu

J. Alex Halderman
University of Michigan
jhalderm@umich.edu

Why Measure the Internet?

- The Internet is a "critical infrastructure". Failures, attacks, and outages in the Internet will effect many other activities.



Level3 Communications offers telecommunications services to business customers. Level 3 services include internet connectivite and managed services such as VPN, collaboration, voice and video.

Level(3)

Detecting Peering Infrastructure Outages in the Wild

Vasileios Giotsas
CAIDA/TU Berlin
vgiotsas@ucsd.edu

Christoph Dietzel
TU Berlin/DE-CIX
christoph@inet.tu-berlin.de

Georgios Smaragdakis
MIT/TU Berlin
gsmaragd@csail.mit.edu

Anja Feldmann
TU Berlin
anja@inet.tu-berlin.de

Arthur Berger
MIT/Akamai
awberger@csail.mit.edu

Emile Aben
RIPE NCC
emile.aben@ripe.net

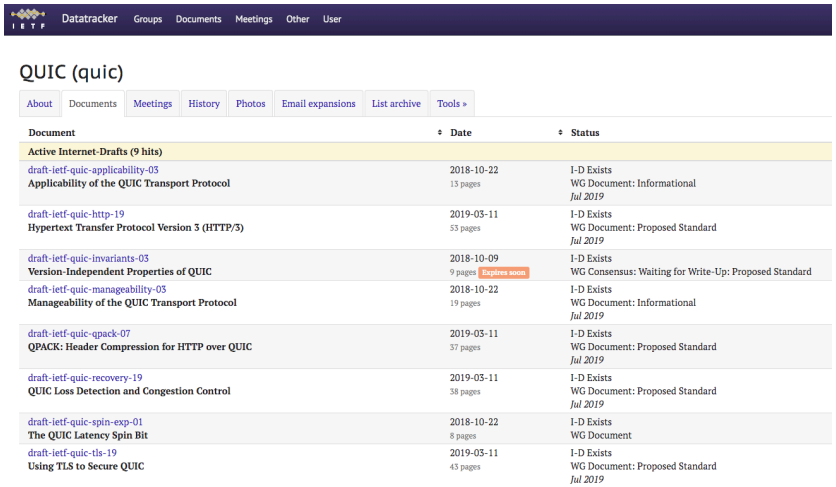
Understanding the Mirai Botnet

Manos Antonakakis[◇] Tim April[‡] Michael Bailey[†] Matthew Bernhard[◁] Elie Bursztein[○]
Jaime Cochran[▷] Zakir Durumeric[◁] J. Alex Halderman[◁] Luca Invernizzi[○]
Michalis Kallitsis[§] Deepak Kumar[†] Chaz Lever[◇] Zane Ma[†*] Joshua Mason[†]
Damian Menscher[○] Chad Seaman[‡] Nick Sullivan[▷] Kurt Thomas[○] Yi Zhou[†]

[‡]*Akamai Technologies* [▷]*Cloudflare* [◇]*Georgia Institute of Technology* [○]*Google*
[§]*Merit Network* [†]*University of Illinois Urbana-Champaign* [◁]*University of Michigan*

Why Measure the Internet?

- There is a big gap between the design and implementation/adoption of protocols!

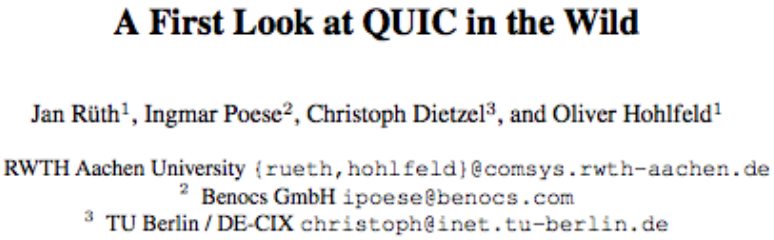


Datatracker Groups Documents Meetings Other User

QUIC (quic)

About | Documents | Meetings | History | Photos | Email expansions | List archive | Tools »

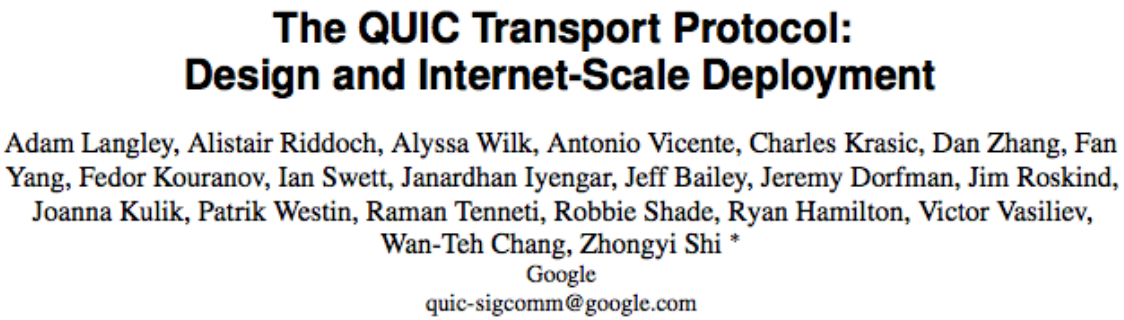
Document	Date	Status
Active Internet-Drafts (9 hits)		
draft-ietf-quic-applicability-03 **Applicability of the QUIC Transport Protocol**	2018-10-22 13 pages	I-D Exists WG Document: Informational *Jul 2019*
draft-ietf-quic-http-19 **Hypertext Transfer Protocol Version 3 (HTTP/3)**	2019-03-11 53 pages	I-D Exists WG Document: Proposed Standard *Jul 2019*
draft-ietf-quic-invariants-03 **Version-Independent Properties of QUIC**	2018-10-09 9 pages Expires soon	I-D Exists WG Consensus: Waiting for Write-Up: Proposed Standard
draft-ietf-quic-manageability-03 **Manageability of the QUIC Transport Protocol**	2018-10-22 19 pages	I-D Exists WG Document: Informational *Jul 2019*
draft-ietf-quic-qpack-07 **QPACK: Header Compression for HTTP over QUIC**	2019-03-11 37 pages	I-D Exists WG Document: Proposed Standard *Jul 2019*
draft-ietf-quic-recovery-19 **QUIC Loss Detection and Congestion Control**	2019-03-11 38 pages	I-D Exists WG Document: Proposed Standard *Jul 2019*
draft-ietf-quic-spin-exp-01 **The QUIC Latency Spin Bit**	2018-10-22 8 pages	I-D Exists WG Document
draft-ietf-quic-tls-19 **Using TLS to Secure QUIC**	2019-03-11 45 pages	I-D Exists WG Document: Proposed Standard *Jul 2019*



A First Look at QUIC in the Wild

Jan Rüth[1], Ingmar Poese[2], Christoph Dietzel[3], and Oliver Hohlfeld[1]

[1] RWTH Aachen University {rueth,hohlfeld}@comsys.rwth-aachen.de
[2] Benocs GmbH ipoese@benocs.com
[3] TU Berlin / DE-CIX christoph@inet.tu-berlin.de



The QUIC Transport Protocol: Design and Internet-Scale Deployment

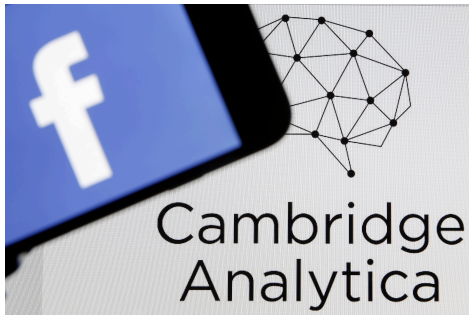
Adam Langley, Alistair Riddoch, Alyssa Wilk, Antonio Vicente, Charles Krasic, Dan Zhang, Fan Yang, Fedor Kouranov, Ian Swett, Janardhan Iyengar, Jeff Bailey, Jeremy Dorfman, Jim Roskind, Joanna Kulik, Patrik Westin, Raman Tenneti, Robbie Shade, Ryan Hamilton, Victor Vasiliev, Wan-Teh Chang, Zhongyi Shi *

Google

quic-sigcomm@google.com

Why Measure the Internet?

- Understanding how the Internet works can inform policymaking and improve the life and protect privacy of Billion of people.



On the Origins of Memes by Means of Fringe Web Communities

Savvas Zannettou*, Tristan Caulfield‡, Jeremy Blackburn†, Emiliano De Cristofaro‡,
Michael Sirivianos*, Gianluca Stringhini°, and Guillermo Suarez-Tangil+

*Cyprus University of Technology, ‡University College London,
†University of Alabama at Birmingham, °Boston University, +King's College London
sa.zannettou@edu.cut.ac.cy, {t.caulfield,e.decristofaro}@ucl.ac.uk, blackburn@uab.edu,
michael.sirivianos@cut.ac.cy, gian@bu.edu, guillermo.suarez-tangil@kcl.ac.uk

Tracing Cross Border Web Tracking

Costas Iordanou
TU Berlin / UC3M

Georgios Smaragdakis
TU Berlin

Ingmar Poese
BENOCS

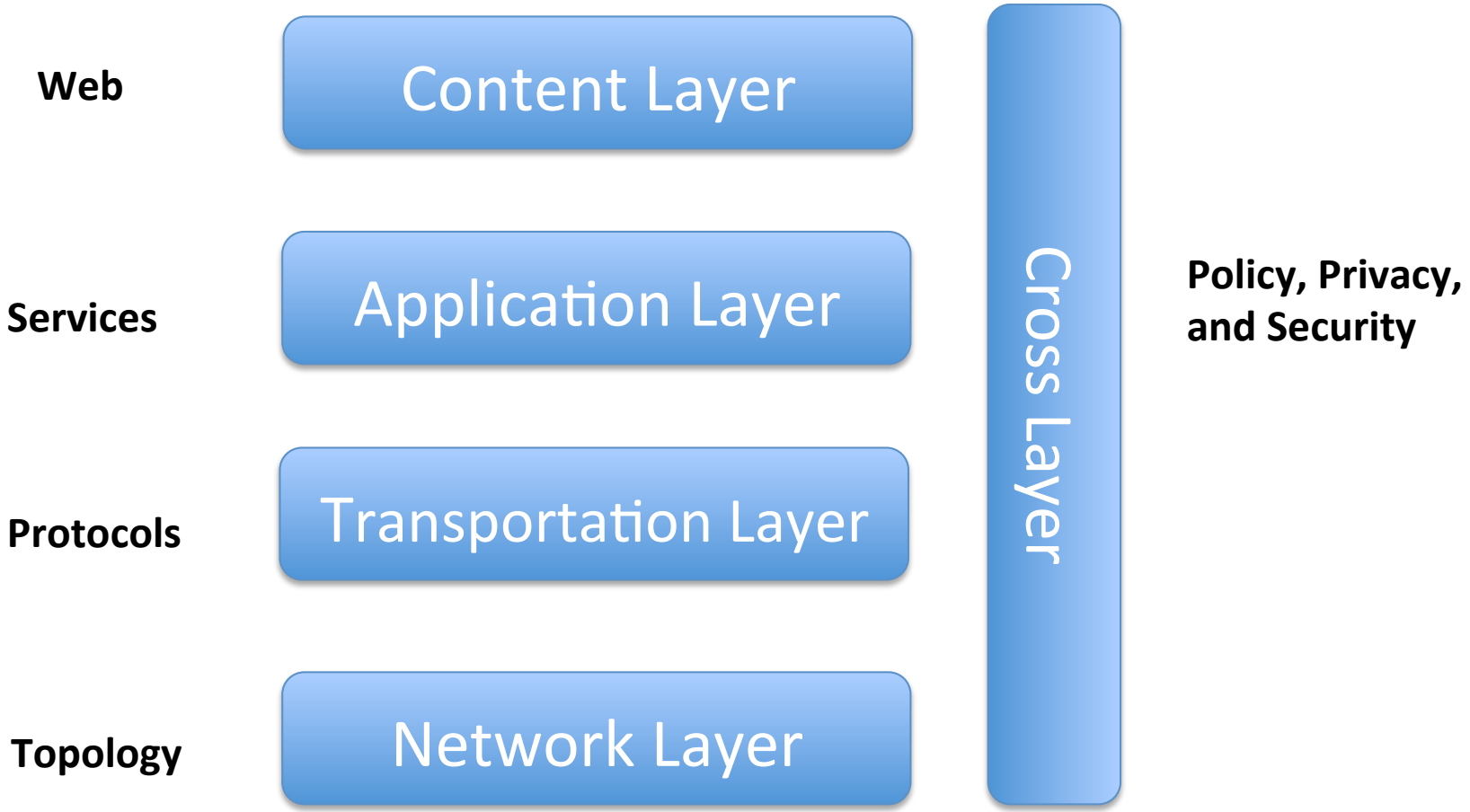
Nikolaos Laoutaris
Data Transparency Lab / Eurecat

Why Measure the Internet?

- Despite the importance of the Internet, our ability to answer even simple questions about the Internet is quite **limited**!

- No central authority
- Lack of "ground truth"
- Fast evolving Internet and Web!

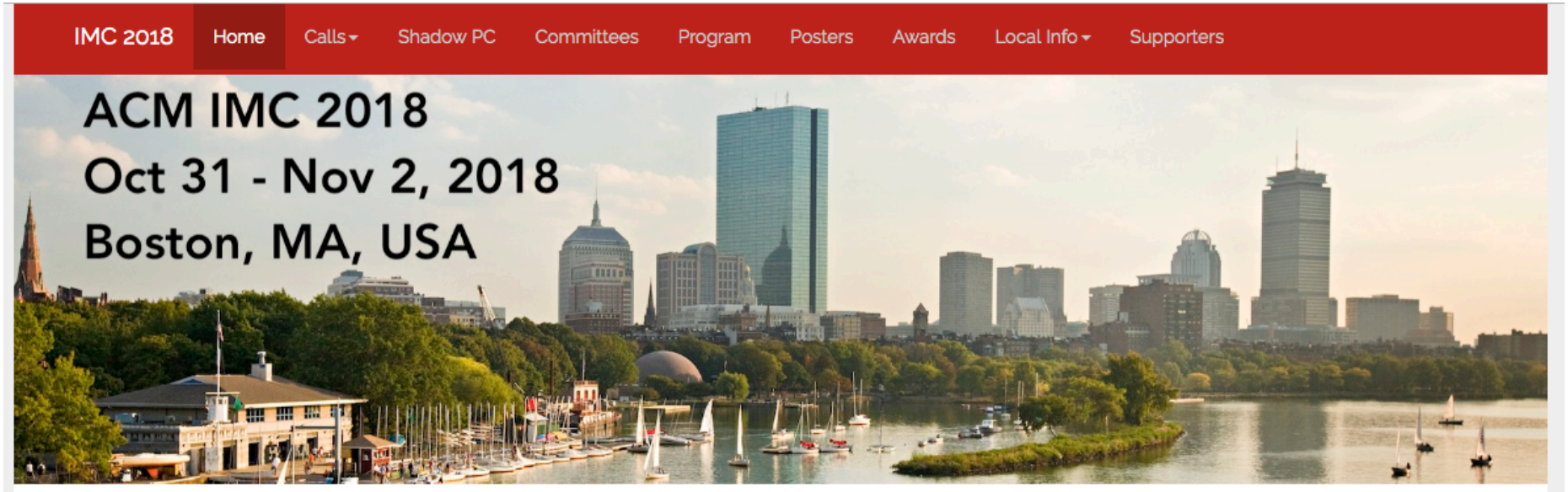
Measuring the Different Internet Layers

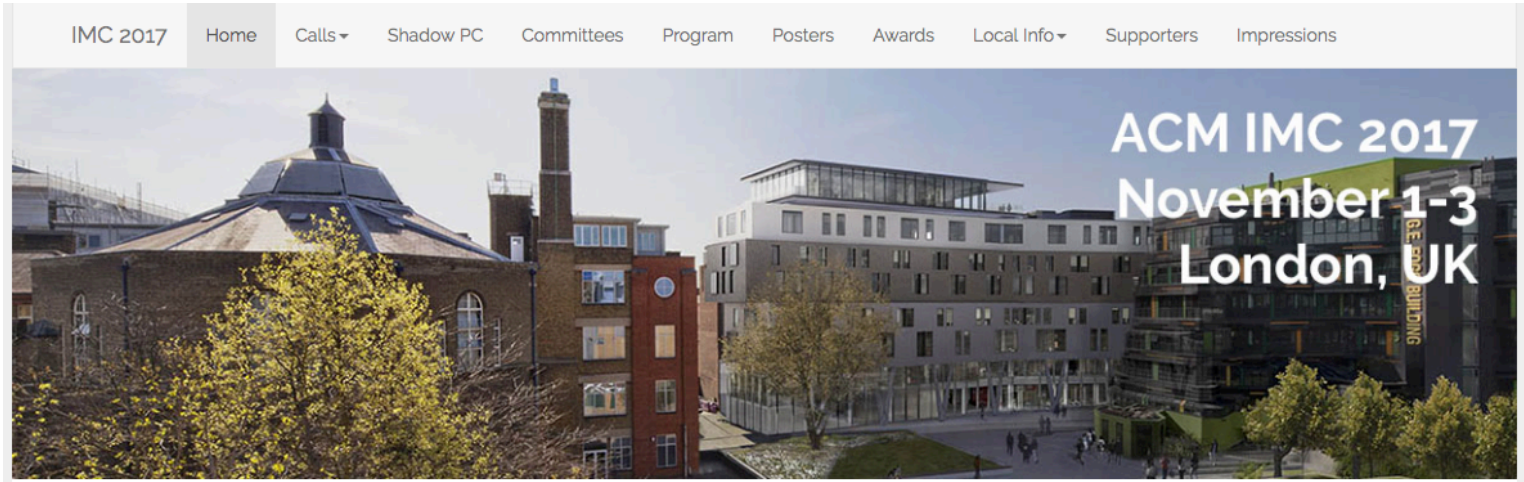


Topics that we will cover in the Lectures:

- Internet Topology Discovery
  - Using passive measurements (BGP)
  - Using active measurements (Traceroute)
- Measuring the New Internet (ISPs, IXPs)
- Measuring the Web Infrastructure (CDNs, Cloud, DNS)
- Measuring Web Performance and Privacy
- Modeling of Network Traffic (Traffic flows)
- Assessing the Internet's Resilience and Security
- Topics: IP Geolocation, Network Attack Mitigation, etc.

Internet Measurement Conference





18 years conference https://www.sigcomm.org/events/imc-conference

Ethical Consideration

- Be a good Internet citizen!
  - Do not try create harm
- If you find a vulnerability, report it!
- Be very careful with private data!!

Example: IPv4 activity

- How many IPv4 addresses exist?

Example: IPv4 activity

- How many IPv4 addresses exist?
  - from your Networking (NPA) course:

    32 bit addresses, i.e.,

    2^{32} (around 4.3 Billion addresses)

Example: IPv4 activity

- How many IPv4 addresses exist?
  - from your Networking (NPA) course:

    2^{32} (around 4.3 Billion addresses)

  - How many of them have been allocated to networks?
  - Are there differences around the world?
  - Are there differences across networks?
  - How many of them are really used?
  - How about IPv6?

Are there public data we can use?

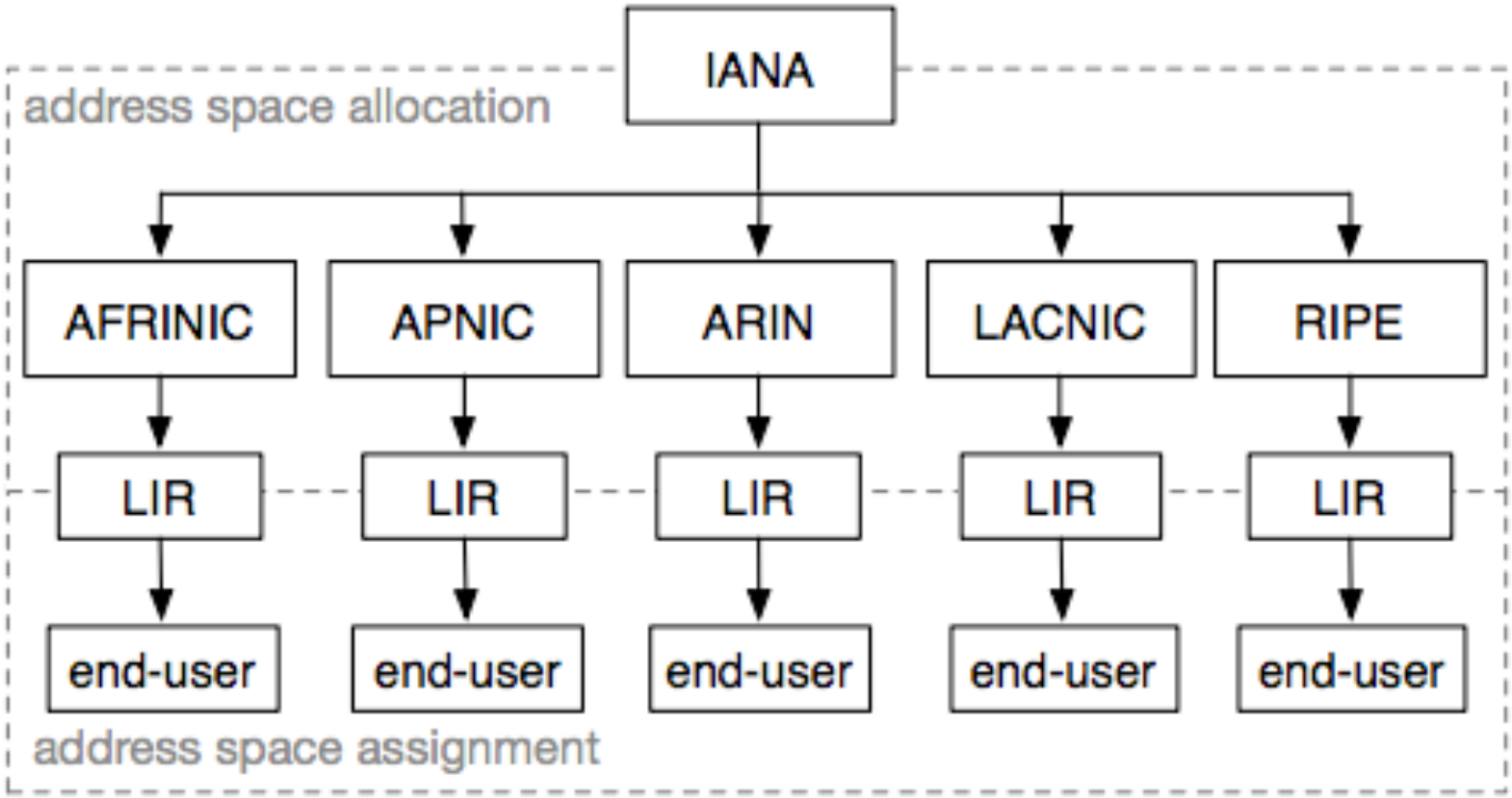
Are there public data we can use? (Address Allocation)

Internet Registries:

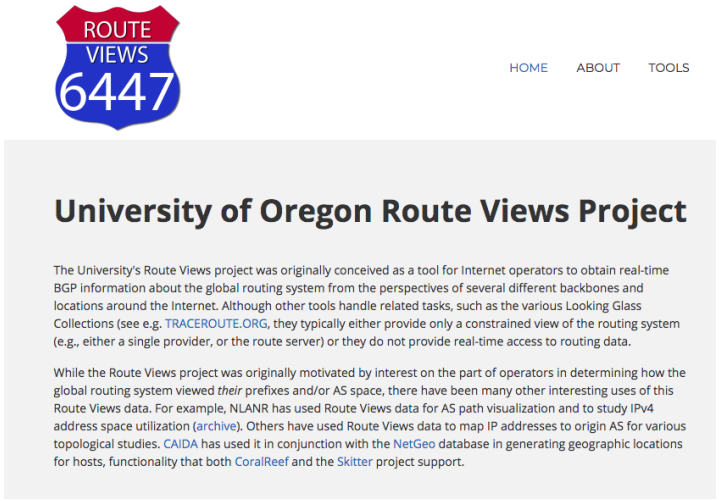
In recent years, address space that is allocated to networks is reported in databases (for accounting, to claim ownership of the address space, for filtering etc.).

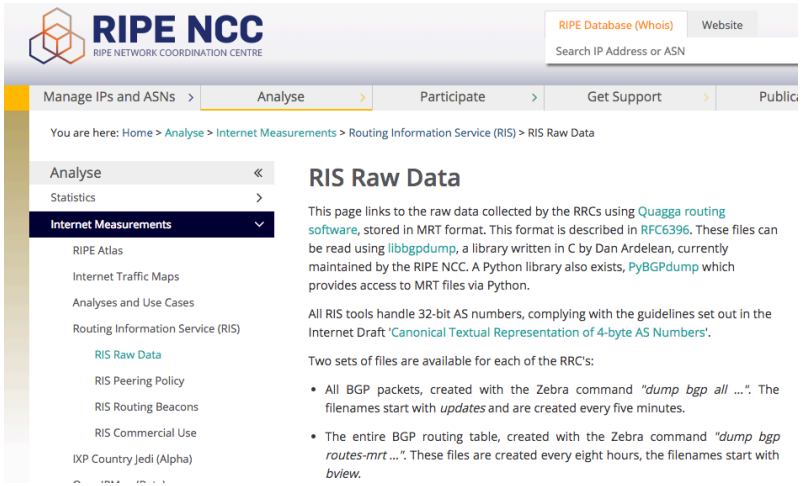
Archives of Internet Registries are available.

Internet Registries (Allocation of IP space)



Are there public data to use? (Route Advertisements)

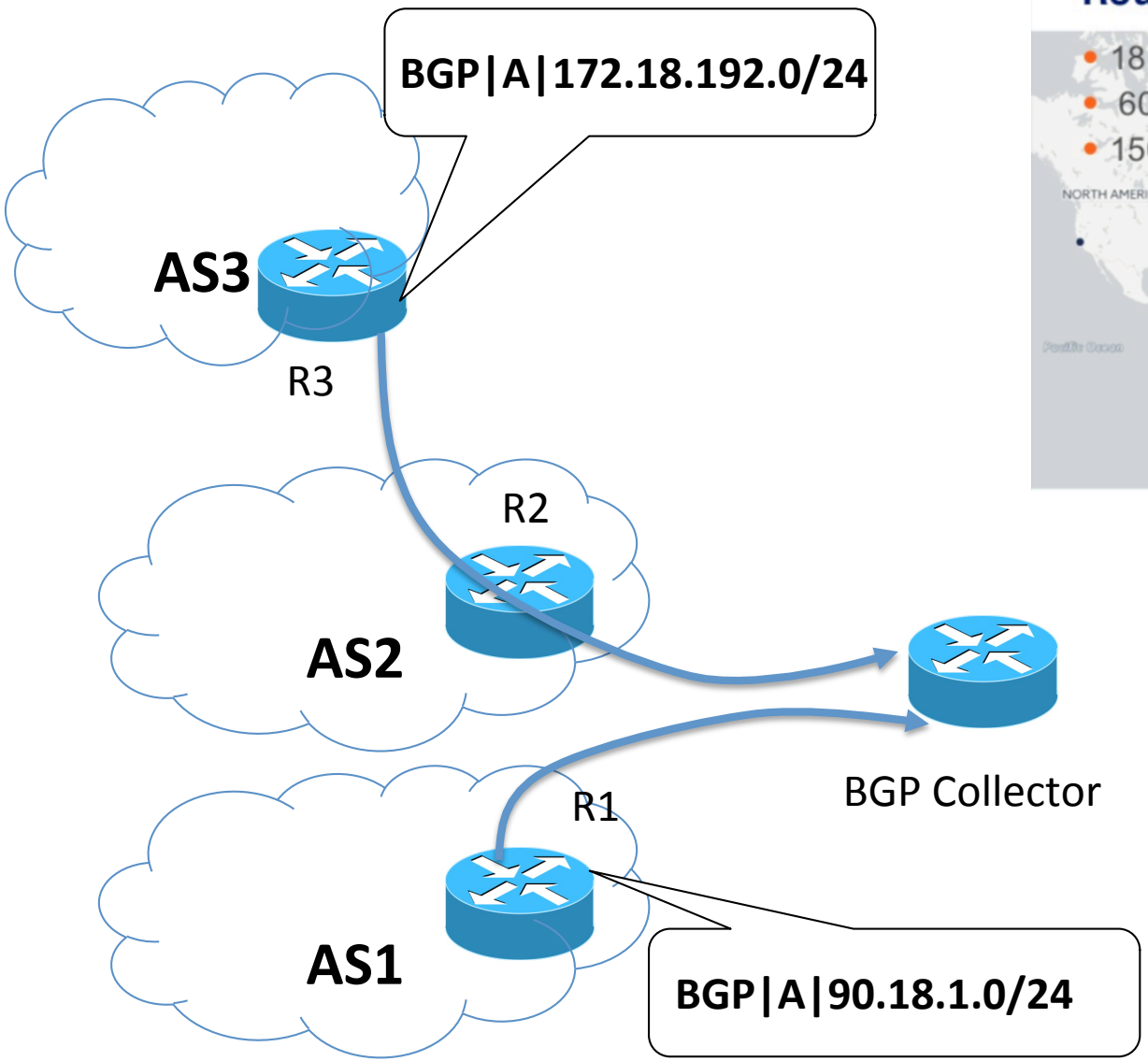


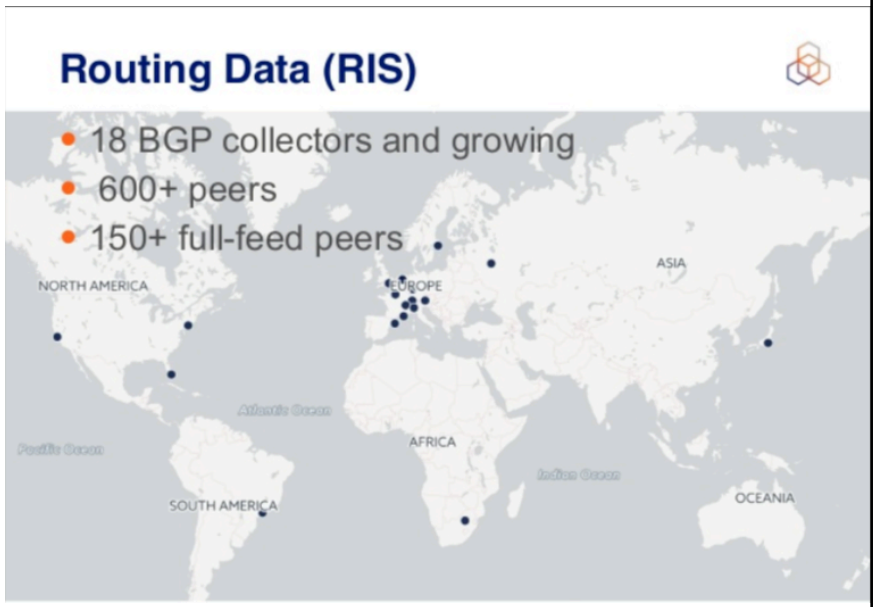


Route Collectors:

Projects like RouteViews and RIPE RIS collect raw BGP data (announcements, withdrawals) at multiple locations around the globe that go back to 1999!

How it works





...

t1|R2|A|172.18.192.0/24|AS2 AS3
t2|R3|A|90.18.1.0/24|AS1

...

Reading

"A Prime on IPv4 Scarcity"

P. Richter, M. Allman, R. Bush, and V. Paxson

Computer Communication Review, April 2015

(awarded best of CCR)

- Paper:

https://dl.acm.org/citation.cfm?id=2766335

- Presentation:

https://people.csail.mit.edu/richterp/ipv4_primer_slides_sigcomm.pdf