Advanced Computer Networking (ACN)

IN2097 – WiSe 2023–2024

Prof. Dr.-Ing. Georg Carle

Sebastian Gallenmüller, Max Helm, Benedikt Jaeger,
Marcel Kempf, Patrick Sattler, Johannes Zirngibl

Chair of Network Architectures and Services
School of Computation, Information, and Technology
Technical University of Munich

Internet-wide Measurements

IPv6 Scans

Why do we need Internet measurements?

Security Measurements

- TLS
- QUIC Measurements
- BACnet

Passive Measurements

Impact of COVID-19 Pandemic on the Internet

Bibliography

Internet-wide Measurements

IPv6 Scans

Why do we need Internet measurements?

Security Measurements

Passive Measurements

Impact of COVID-19 Pandemic on the Internet

Bibliography

IPv6 Scans
ZMapv6

Original ZMap implementation supports only IPv4

- Extension of ZMap with IPv6 capabilities → ZMapv6
- https://github.com/tumi8/zmap
- Adaptation of scanning core to send and receive IPv6 packets
- Port probe modules for IPv6 scanning: ICMPv6, TCP over IPv6, UDP over IPv6

IPv6 Scans

Challenges

- Vast address space → “0/0” scan not possible
  - Scan rate 20k IP addr/s → 5.4 × 10^{26} years

→ Hitlists are required

IPv6 Scans

Challenges

- Vast address space → “0/0” scan not possible
  - Scan rate 20k IP addr/s → 5.4 × 10^{26} years

→ Hitlists are required

A list of **targets**, most likely **responsive**, of **feasible size**.

IPv6 Scans

Challenges

- Vast address space → “0/0” scan not possible
  - Scan rate 20k IP addr/s → 5.4×10^{26} years

→ Hitlists are required

A list of **targets**, most likely **responsive**, of **feasible size**.

Responsive:

- Responsive to at least one protocol (e.g., ICMP, HTTP, . . .)
- Different between addresses
- Changes over time

IPv6 Scans

Challenges

- Vast address space → “0/0” scan not possible
  - Scan rate 20k IP addr/s → 5.4×10^{26} years

→ Hitlists are required

A list of **targets**, most likely **responsive**, of **feasible size**.

Responsive:

- Responsive to at least one protocol (e.g., ICMP, HTTP, . . .)
- Different between addresses
- Changes over time

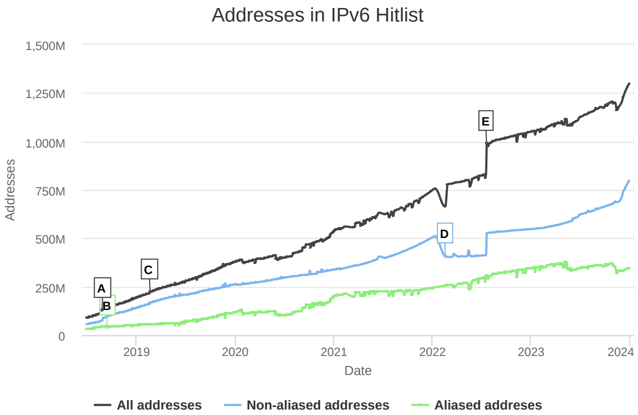
Feasible size:

- Scan duration
- Bandwidth limitations

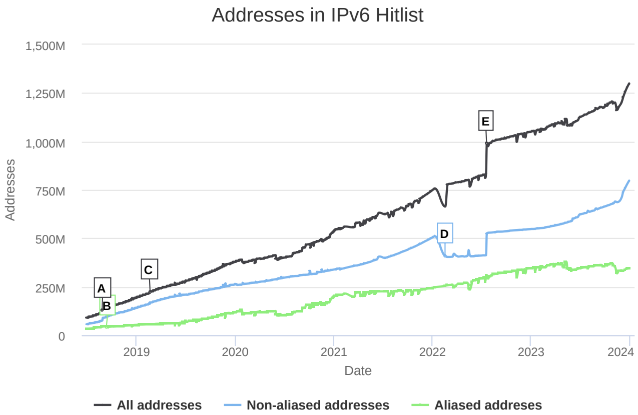
IPv6 Hitlist Service

- Research at this Chair
- Aggregates multiple inputs
- Filters aliased prefixes and applies blocklists
- Tests reachability daily
    - ICMPv6
    - TCP/80 (HTTP)
    - TCP/443 (HTTPS)
    - UDP/53 (DNS)
    - UDP/443 (QUIC)
- Uses ZMapv6

IPv6 Hitlist Service
Input

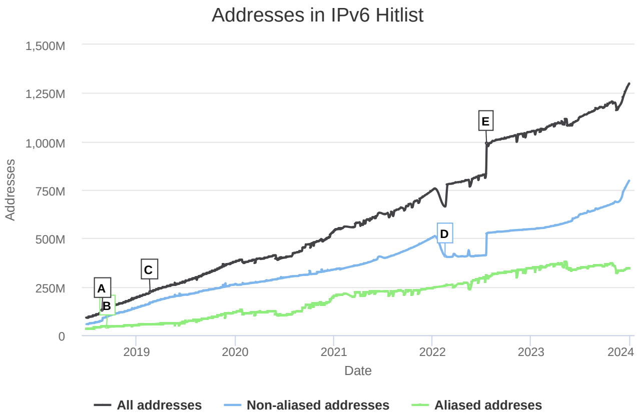


IPv6 Hitlist Service
Input



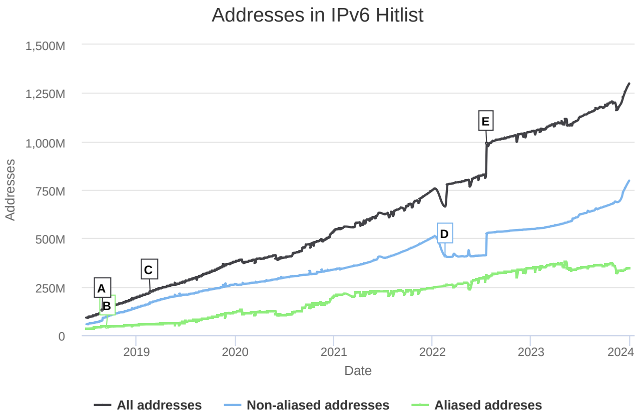
- A: Addition of IPv6 rDNS

IPv6 Hitlist Service
Input



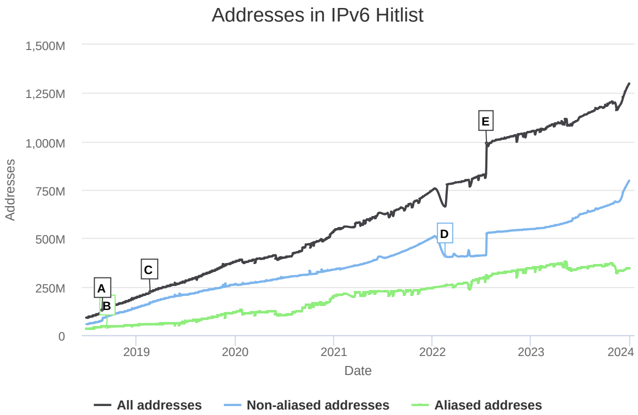
- A: Addition of IPv6 rDNS
- B: Withdrawal of two Amazon EC2 prefixes

IPv6 Hitlist Service
Input



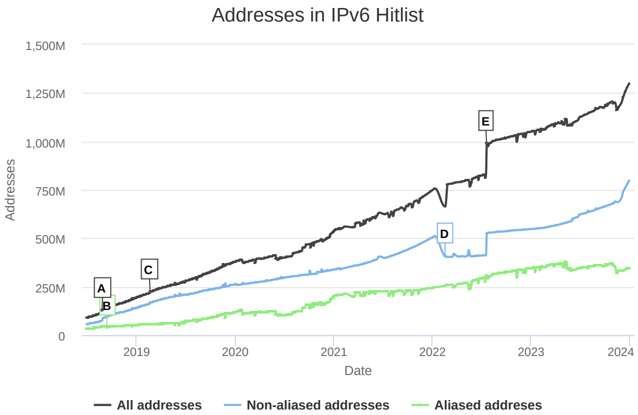
- A: Addition of IPv6 rDNS
- B: Withdrawal of two Amazon EC2 prefixes
- C: Additional IPv6 rDNS results

IPv6 Hitlist Service
Input



- A: Addition of IPv6 rDNS
- B: Withdrawal of two Amazon EC2 prefixes
- C: Additional IPv6 rDNS results
- D: Removal of addresses impacted by the GFW

IPv6 Hitlist Service
Input



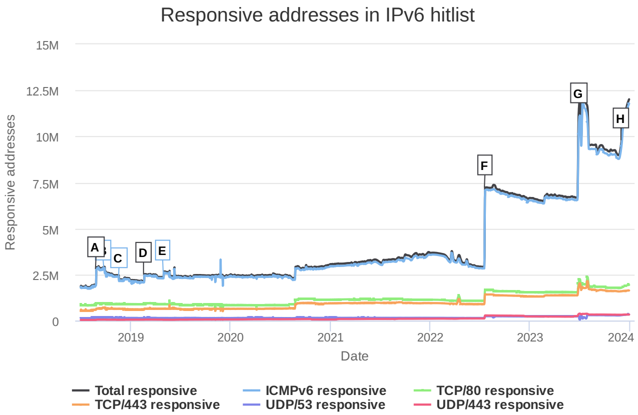
- A: Addition of IPv6 rDNS
- B: Withdrawal of two Amazon EC2 prefixes
- C: Additional IPv6 rDNS results
- D: Removal of addresses impacted by the GFW
- E: Addition of new addresses from passive sources and target generation methods

IPv6 Hitlist Service
Filter

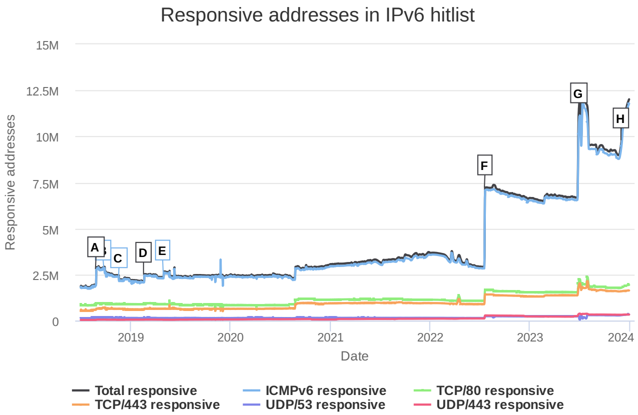
The input has to be filtered by several steps before the responsiveness can be tested.

- Not globally routed
  - Datasets might contain addresses that are not routed
  - Infrastructure changes might change the reachability of addresses
- Blocklisted
  - Aggregated list of blocklisting requests from all scans at our chair
- Aliased prefixes
- Not responsive for 30 consecutive days

IPv6 Hitlist Service
Results

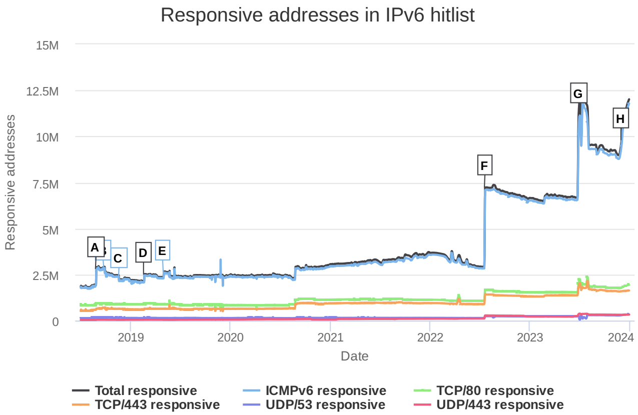


IPv6 Hitlist Service
Results



- A, D, E: Addition of IPv6 rDNS

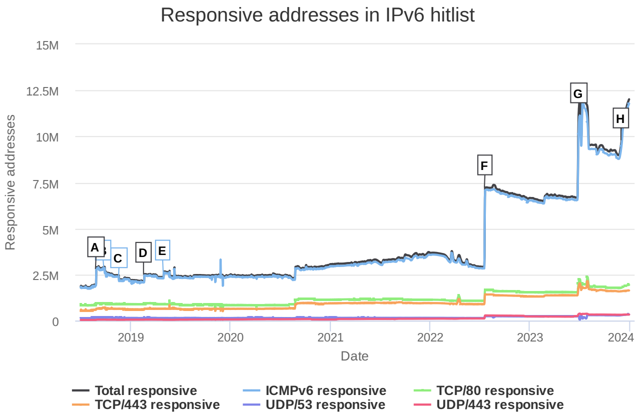
IPv6 Hitlist Service
Results



- A, D, E: Addition of IPv6 rDNS
- B: LATNET SERVISS stopped responding

IPv6 Hitlist Service

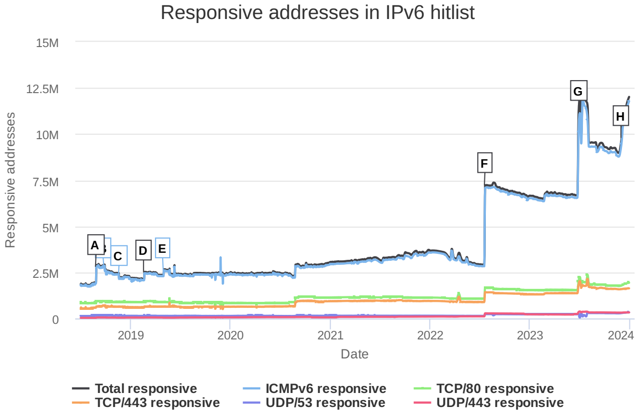
Results



- A, D, E: Addition of IPv6 rDNS
- B: LATNET SERVISS stopped responding
- C: Online S.A.S. stopped responding

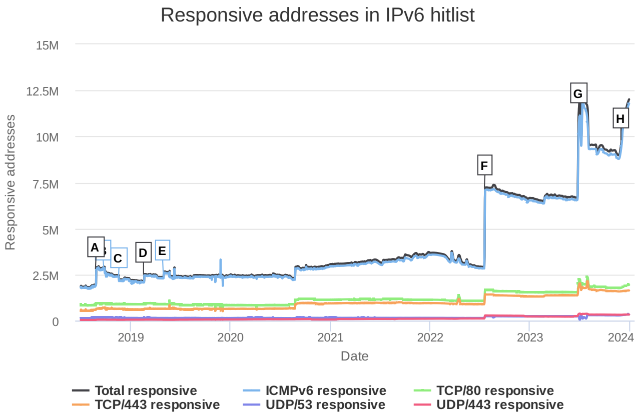
IPv6 Hitlist Service

Results



- A, D, E: Addition of IPv6 rDNS
- B: LATNET SERVISS stopped responding
- C: Online S.A.S. stopped responding
- F,G: Addition of new address candidates based on target generation algorithms

IPv6 Hitlist Service
Results



- A, D, E: Addition of IPv6 rDNS
- B: LATNET SERVISS stopped responding
- C: Online S.A.S. stopped responding
- F,G: Addition of new address candidates based on target generation algorithms
- H: Addition of ipinfo.io as new data collaborator

IPv6 Hitlist Service
New Sources

While the existing IPv6 Hitlist sources regularly update the input, new sources have not been added.

We evaluated different approaches to extend our hitlist [1]:

- Target Generation:
  - 6Tree, 6Graph, 6GAN, 6VecLM,
  - Distance Clustering (DC) (our custom algorithm)
- 30-day unresponsive addresses

		Responsive	
Method	Addr	Addr. ↓	ASes
6Graph	125.8 M		
6Tree	37.6 M		
DC	5.3 M		
6GAN	3.3 M		
6VecLM	70.3 k		
30-day Unresp.	405.0 M		

New Sources

While the existing IPv6 Hitlist sources regularly update the input, new sources have not been added.

We evaluated different approaches to extend our hitlist [1]:

- Target Generation:
  - 6Tree, 6Graph, 6GAN, 6VecLM,
  - Distance Clustering (DC) (our custom algorithm)
- 30-day unresponsive addresses

		Responsive	
Method	Addr	Addr. ↓	ASes
6Graph	125.8 M	3.8 M	10.7 k
6Tree	37.6 M	2.2 M	11.5 k
DC	5.3 M	651.9 k	5.5 k
6GAN	3.3 M	4.3 k	39
6VecLM	70.3 k	1.0 k	105
30-day Unresp.	405.0 M	1.3 M	9.0 k

→ All sources contribute additional responsive addresses.

→ We identified 5.6 M new responsive IPv6 addresses from 14.6 k ASes.

Summary

Hitlists:

- Lots of possible sources
- Knowledge about sources is important
- Number of IP addresses is not only metric → evaluate reachability and stability
- Optimal sources depend on type of measurement (end-user devices, servers, routers,. . .)
- Be aware of biases in your hitlist (address distribution, prefix/AS distribution, aliased prefixes)

Tracking Clients

Is client tracking impossible?

The privacy extension prevents tracking of clients by randomization of the interface ID.

- In general, most end devices implement the privacy extension
- 90% of IPv6 addresses seen by a large CDN are only seen once in long-running analyses [2]

Is client tracking impossible?

The privacy extension prevents tracking of clients by randomization of the interface ID.

- In general, most end devices implement the privacy extension
- 90% of IPv6 addresses seen by a large CDN are only seen once in long-running analyses [2]
- But how about Customer Premises Equipment (CPE)?
  - e.g., private networks use a single, fixed router (the CPE) as gateway to the Internet

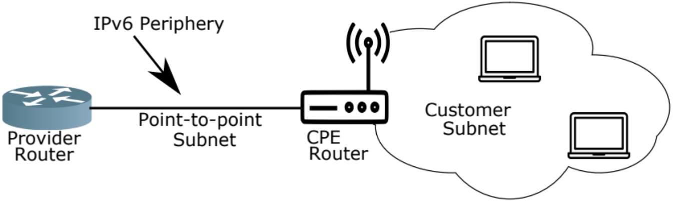


Figure 1: Common IPv6 architecture [3]

Is client tracking impossible?

The privacy extension prevents tracking of clients by randomization of the interface ID.

- In general, most end devices implement the privacy extension
- 90% of IPv6 addresses seen by a large CDN are only seen once in long-running analyses [2]
- But how about Customer Premises Equipment (CPE)?
  - e.g., private networks use a single, fixed router (the CPE) as gateway to the Internet
  - → In 2020, an approach to discover the IPv6 periphery was presented [3]
  - → A measurement revealed 64.8M router addresses
  - → 30M addresses were found using EUI-64

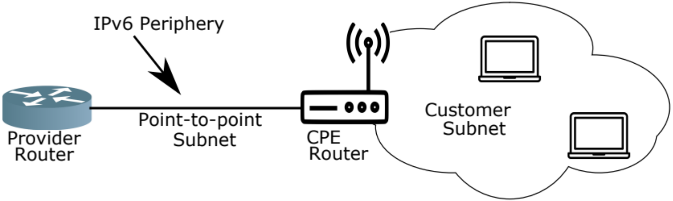


Figure 1: Common IPv6 architecture [3]

Tracking Clients

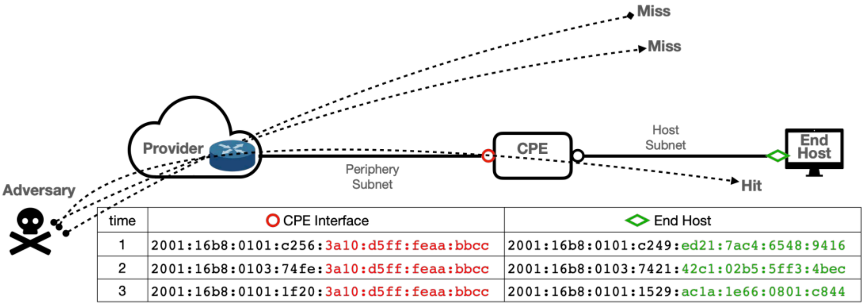
Is client tracking impossible?

- To prevent tracking based on the assigned prefix, providers often rotate their assignments

Tracking Clients

Is client tracking impossible?

- To prevent tracking based on the assigned prefix, providers often rotate their assignments
- But behavioral analysis of providers and CPE's using EUI-64 can be used to track prefixes [4]
- While clients can not be tracked, CPEs using EUI-64 identifiers can be actively found.



time	○ CPE Interface	◇ End Host
1	2001:16b8:0101:c256:3a10:d5ff:feaa:bbcc	2001:16b8:0101:c249:ed21:7ac4:6548:9416
2	2001:16b8:0103:74fe:3a10:d5ff:feaa:bbcc	2001:16b8:0103:7421:42c1:02b5:5ff3:4bec
3	2001:16b8:0101:1f20:3a10:d5ff:feaa:bbcc	2001:16b8:0101:1529:ac1a:1e66:0801:c844

Figure 2: CPE with missing privacy extension [4]

Tracking Clients

Is client tracking impossible?

- In theory, testing all targets in a /48 is infeasible
- How can we reduce the search space and effectively track EUI-64 identifier?
  - Customers receive at least /64 prefixes and often larger
  - Providers often use only parts of their owned prefixes
  - Prefixes are often assigned at nibbles (e.g., /56, /60, /64)

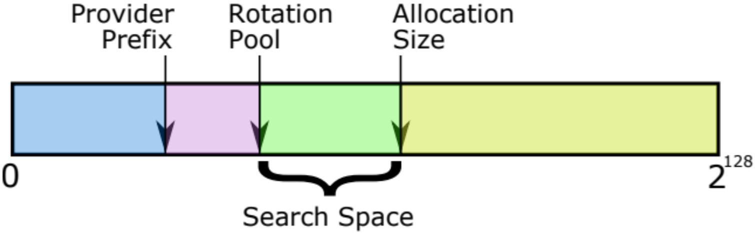
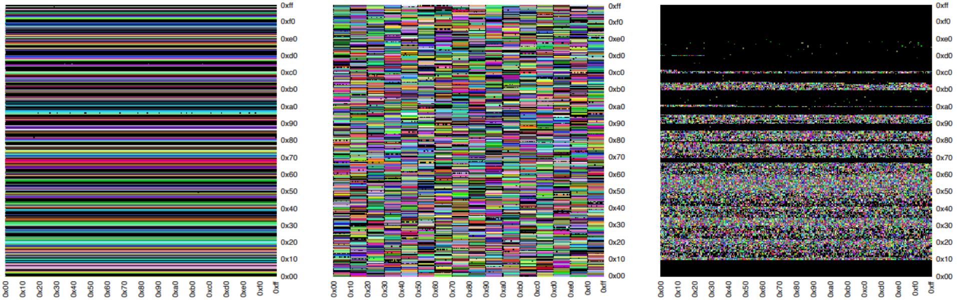


Figure 3: Limiting the search space to track IPv6 hosts: for a provider, infer the : i) size of the allocations to customers; and ii) range of prefixes used for rotation. [4]

Tracking Clients

Is client tracking impossible?

- A recent study reveals that high speed probing, behavioral analysis of providers and CPE’s using EUI-64 can be used to track prefixes [4]
- Different allocation schemes can be found, e.g., /56 or /64 assignments



(a) Entel (Bolivia): /56 allocations

(b) BH Telecom (Bosnia): /60 allocations

(c) Starcat (Japan): /64 allocations

Figure 4: The y-axis of plots represents the 7^{th} byte of a probed address, while the x-axis denotes the 8^{th} byte; each pixel represents a probed /64 network. Each color represents a different responsive source address, while black indicates no response was received when probing to an address in that /64 network. [4]

Is client tracking impossible?

- Active scans can be used to identify specific CPEs even if providers rotate their assignments
- Based on a reduced search space, an adversary can effectively implement tracking

→ The tracking relies on CPEs using EUI-64

→ CPE manufacturers should change the device behavior

- The given paper resulted in a change of behavior within a large manufacturer
- For more details see the original work [4]

Internet-wide Measurements

IPv6 Scans

Why do we need Internet measurements?

Security Measurements

Passive Measurements

Impact of COVID-19 Pandemic on the Internet

Bibliography

Why do we need Internet measurements?

Do we really have to?

- The network is well engineered
- Well documented protocols, mechanisms, . . .
- Everything built by humans
  - → No unknowns (compare this to physics)
- In theory, we can know everything that is going on

→ No need for measurements?!

Why do we need Internet measurements?

Do we really have to?

- The network is well engineered
- Well documented protocols, mechanisms, . . .
- Everything built by humans
  - → No unknowns (compare this to physics)
- In theory, we can know everything that is going on

→ No need for measurements?!

But:

- Distributed multi-domain network
  - → Information only partially available
- Moving target
  - Requirements change
  - Growth, usage, structure changes
- Highly interactive system
- Heterogeneity in all directions
- The total is more than the sum of its pieces
- Built, driven, and used by humans
  - → Errors, misconfigurations, flaws, failures, misuse, . . .

Why do we need Internet measurements?

Do we really have to?

- The network is well engineered
- Well documented protocols, mechanisms, . . .
- Everything built by humans
  - → No unknowns (compare this to physics)
- In theory, we can know everything that is going on

→ No need for measurements?!

But:

- Distributed multi-domain network
  - → Information only partially available
- Moving target
  - Requirements change
  - Growth, usage, structure changes
- Highly interactive system
- Heterogeneity in all directions
- The total is more than the sum of its pieces
- Built, driven, and used by humans
  - → Errors, misconfigurations, flaws, failures, misuse, . . .

Active network measurements are an important research area to understand the Internet and interactions between all its components.

Why do we need Internet measurements?
Why do we measure the network?

Network provider view

- Manage traffic
  - Model reality
  - Predict future
  - Plan network
  - Avoid bottlenecks in advance
- Reduce cost
- Accounting

Why do we need Internet measurements?
Why do we measure the network?

Network provider view

- Manage traffic
  - Model reality
  - Predict future
  - Plan network
  - Avoid bottlenecks in advance
- Reduce cost
- Accounting

Service provider view

- Get information about clients
- Adjust service to demands
- Reduce load on servers
- Accounting

Why do we need Internet measurements?
Why do we measure the network?

Network provider view

- Manage traffic
  - Model reality
  - Predict future
  - Plan network
  - Avoid bottlenecks in advance
- Reduce cost
- Accounting

Service provider view

- Get information about clients
- Adjust service to demands
- Reduce load on servers
- Accounting

Client view

- Get the best possible service
- *Do I get what I paid for?*

Why do we need Internet measurements?
Why do we measure the network?

Network provider view

- Manage traffic
  - Model reality
  - Predict future
  - Plan network
  - Avoid bottlenecks in advance
- Reduce cost
- Accounting

Service provider view

- Get information about clients
- Adjust service to demands
- Reduce load on servers
- Accounting

Client view

- Get the best possible service
- *Do I get what I paid for?*

Security view

- Detect malicious traffic
- Detect malicious hosts
- Detect malicious networks

Why do we need Internet measurements?
Why do we measure the network?

Network provider view

- Manage traffic
  - Model reality
  - Predict future
  - Plan network
  - Avoid bottlenecks in advance
- Reduce cost
- Accounting

Service provider view

- Get information about clients
- Adjust service to demands
- Reduce load on servers
- Accounting

Client view

- Get the best possible service
- *Do I get what I paid for?*

Security view

- Detect malicious traffic
- Detect malicious hosts
- Detect malicious networks

Researcher view

- Understand the Internet better
- *Could our new routing algorithm handle all this real-world traffic?*
- . . .

Internet-wide Measurements

IPv6 Scans

Why do we need Internet measurements?

Security Measurements

- TLS
- QUIC Measurements
- BACnet

Passive Measurements

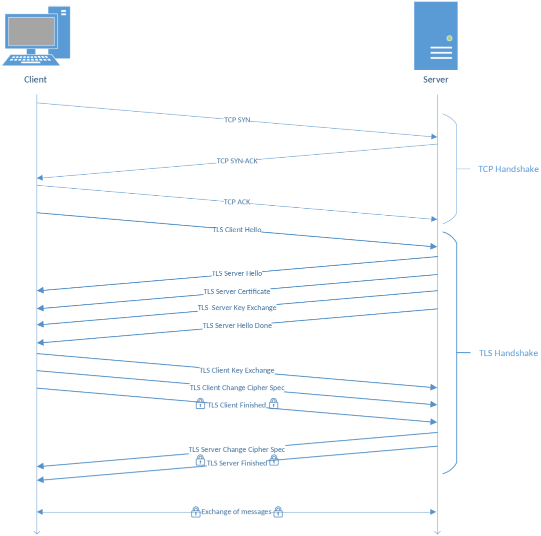
Impact of COVID-19 Pandemic on the Internet

Bibliography

TLS

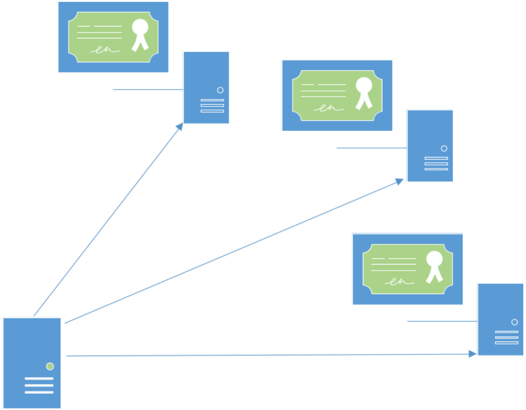
- TLS: Transport Layer Security
  - SSL 3.0
  - TLS 1.0
  - TLS 1.1
  - TLS 1.2
  - TLS 1.3
- Security foundation for HTTPS, IMAPS, SMPTS, DoT, DoH, . . .

→ Evaluate TLS Deployment



Certificate Scanning

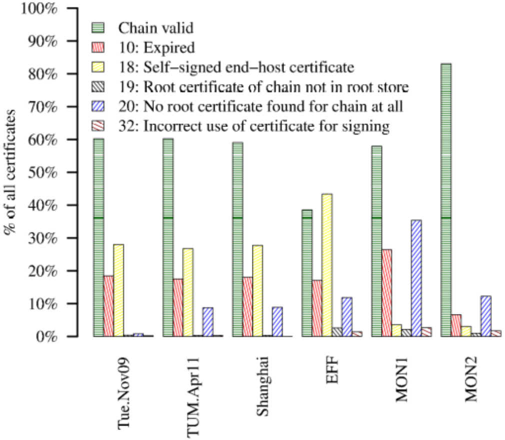
- Methodology
  1. Identify hosts offering TLS service (HTTPS, IMAPS,. . .)
  2. Download certificate chains
  3. Analyze and validate chains
- Challenges
  - Targets (0/0?)
  - Performance
  - Evaluation metrics



Certificate Scanning

Analysis of the TLS landscape [5]

- Active and passive measurements
  1. Analyses of certificate chains
  2. Expiry
  3. Algorithms
- Conclusion:
  - TLS landscape in sorry state (expired, no root cert, . . .)
  - But: situation improves over time [6]



Evolution of TLS Scanning

	Holz et al. (2011) [5]	**Now**
Targets	• Alexa Top 1M	• Full IPv4 & IPv6 hitlist

Evolution of TLS Scanning

	Holz et al. (2011) [5]	**Now**
Targets	• Alexa Top 1M	• Full IPv4 & IPv6 hitlist
Server Name Indication (SNI)	• Not used	• Alexa Top 1M • $>$ 1000 TLD Zone files • Reverse DNS

TLS

Evolution of TLS Scanning

	Holz et al. (2011) [5]	**Now**
Targets	• Alexa Top 1M	• Full IPv4 & IPv6 hitlist
Server Name Indication (SNI)	• Not used	• Alexa Top 1M • $>$ 1000 TLD Zone files • Reverse DNS
Software stack	• Nmap • OpenSSL	• ZMap • Custom-built scanner for TLS and HTTPS

TLS

Evolution of TLS Scanning

	Holz et al. (2011) [5]	**Now**
Targets	• Alexa Top 1M	• Full IPv4 & IPv6 hitlist
Server Name Indication (SNI)	• Not used	• Alexa Top 1M • $>$ 1000 TLD Zone files • Reverse DNS
Software stack	• Nmap • OpenSSL	• ZMap • Custom-built scanner for TLS and HTTPS
Performance	• Weeks for 1M hosts	• Day(s) for complete Internet (several hundert millions of hosts)

TLS

Evolution of TLS Scanning

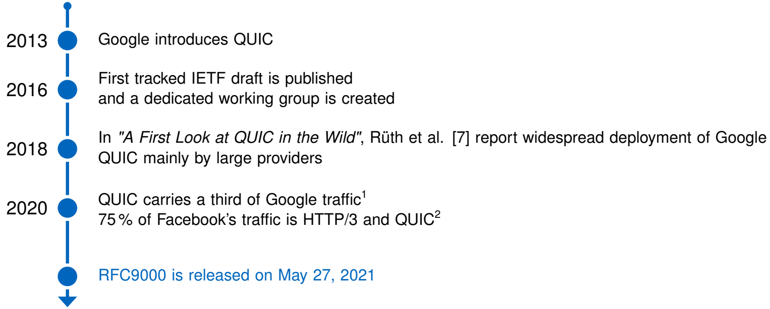
	Holz et al. (2011) [5]	**Now**
Targets	• Alexa Top 1M	• Full IPv4 & IPv6 hitlist
Server Name Indication (SNI)	• Not used	• Alexa Top 1M • $>$ 1000 TLD Zone files • Reverse DNS
Software stack	• Nmap • OpenSSL	• ZMap • Custom-built scanner for TLS and HTTPS
Performance	• Weeks for 1M hosts	• Day(s) for complete Internet (several hundert millions of hosts)
Frequency	• Single measurements	• Continuously running measurement service

TLS 1.3

New features in TLS 1.3

- 1-RTT handshakes by default
    - Use presumed cipher suite selection
- 0-RTT handshake with resumption possible
    - PSK for early data
    - Forward secrecy after early data
- Privacy
    - Client certificates are encrypted
    - SNI not encrypted (RFC Draft for encrypted SNI in TLS 1.3)
- Grease mechanism
    - Send random version data to increase robustness

QUIC Measurements



2013 Google introduces QUIC

2016 First tracked IETF draft is published
and a dedicated working group is created

2018 In *"A First Look at QUIC in the Wild"*, Rüth et al. [7] report widespread deployment of Google QUIC mainly by large providers

2020 QUIC carries a third of Google traffic[1]
75 % of Facebook's traffic is HTTP/3 and QUIC[2]

RFC9000 is released on May 27, 2021

[1] https://blog.chromium.org/2020/10/chrome-is-deploying-http3-and-ietf-quic.html

[2] https://engineering.fb.com/2020/10/21/networking-traffic/how-facebook-is-bringing-quic-to-billions/

[3] https://hacks.mozilla.org/2021/04/quic-and-http-3-support-now-in-firefox-nightly-and-beta/

QUIC Measurements

As a new fundamental network protocol with widespread early adoption, QUIC requires early analysis and researchers tools to analyze QUIC deployments.

→ We provided an Internet-wide measurement study shortly before the final RFC release [8]

Research Questions:

1. How can we detect QUIC deployments?
   - → IPv4 + IPv6 ZMap modules
   - → HTTPS DNS RR
   - → HTTP ALT-SVC header
2. Who deploys QUIC?
3. Which QUIC versions are deployed?
4. Can we successfully connect to QUIC servers and analyze deployments?
   - → We developed and published the QScanner, a highly parallelized stateful QUIC scanner

How can we detect QUIC deployments?

ZMap module:

- QUIC relies on UDP
  → ZMap needs to send valid QUIC packets
- Relies on the QUIC version negotiation
  - Server responses should contain all supported versions
  - No state is created at the server
  - No computational expensive cryptography is necessary
- Requires no input (at least for IPv4)
- ZMap reports most addresses supporting the QUIC version negotiation
  - Domains can be mapped to only 10 % of addresses

		Scanned		Results	
		Targets	Addresses	ASes	Domains
ZMap	IPv4	3 023 298 514	2 134 964	4736	30 970 316
	IPv6	24 434 296	210 997	1704	17 972 799

QUIC Measurements
How can we detect QUIC deployments?

HTTPS DNS Resource Records

- Based on a new IETF draft [9]
- Specifies DNS resource records to provide service information
  - Can include ALPN values indicating QUIC support
  - `simple.example 7200 IN HTTPS 1 . alpn=h3`
- Requires domains to resolve

→ HTTPS DNS RRs results in the fewest amount of deployments

		Scanned		Results	
		Targets	Addresses	ASes	Domains
HTTPS	IPv4	213 689 057	85 092	1287	2 962 708
	IPv6		69 684	112	2 736 040

QUIC Measurements

How can we detect QUIC deployments?

HTTP ALTSVC Headers

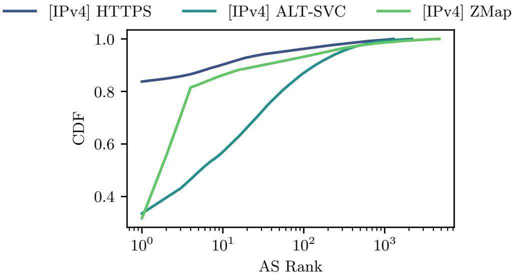
- HTTP header containing alternative service information
  - Can include ALPN values indicating QUIC support
  - `alt-svc: h3=":443"; ma=86400, h3-29=":443"; ma=86400, h3-28=":443"; ma=86400, h3-27=":443"; ma=86400`
- Requires HTTP(s) capable targets and scans
- ALT-SVC reveals the most domains with QUIC support

		Scanned		Results	
		Targets	Addresses	ASes	Domains
ALT-SVC	IPv4	375 338 772	232 585	2174	36 907 770
	IPv6	69 458 318	283 169	292	16 979 759

Who deploys QUIC?

To analyze who is involved in the deployment of QUIC, we analyzed originating ASes:

- Deployments are dominated by large providers
- ZMap results in addresses located in more than 4.7 k ASes
- HTTPS DNS Resource Records are strongly biased towards Cloudflare

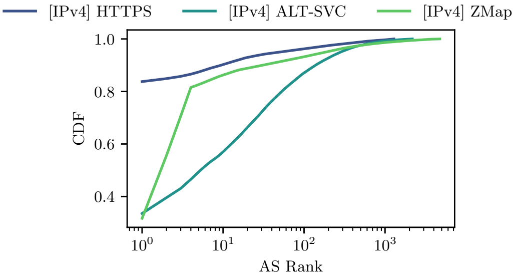


Rank	Provider	ZMap #IPv4 Addr.	#Domains
1	Cloudflare	676 483	23 843 989
2	Google	510 450	6 006 547
3	Akamai	320 646	23 206
4	Fastly	232 776	938 649
5	Cloudflare London	23 489	61 979

Who deploys QUIC?

To analyze who is involved in the deployment of QUIC, we analyzed originating ASes:

- Deployments are dominated by large providers
- ZMap results in addresses located in more than 4.7 k ASes
- HTTPS DNS Resource Records are strongly biased towards Cloudflare

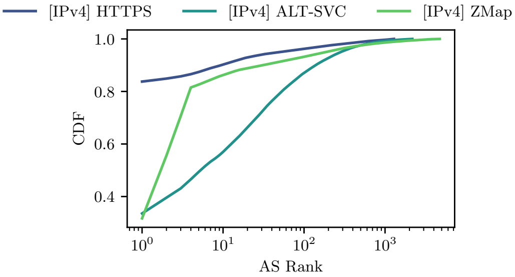


Rank	Provider	HTTPS DNS RR #IPv4 Addr.	#Domains
1	Cloudflare	71 278	2 887 327
2	DigitalOcean	969	1256
3	Google	719	1235
4	Amazon	709	814
5	OVH	708	1034

Who deploys QUIC?

To analyze who is involved in the deployment of QUIC, we analyzed originating ASes:

- Deployments are dominated by large providers
- ZMap results in addresses located in more than 4.7 k ASes
- HTTPS DNS Resource Records are strongly biased towards Cloudflare



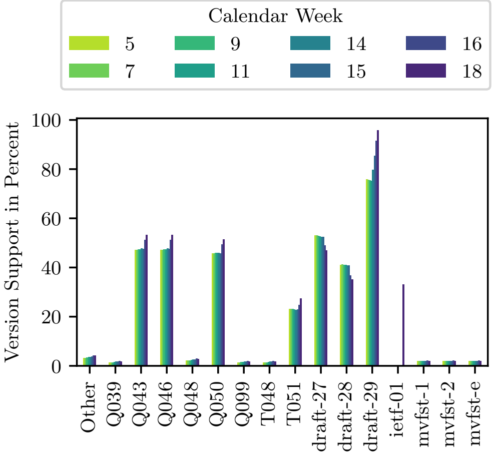
Rank	Provider	ALT-SVC #IPv4 Addr.	#Domains
1	Cloudflare	78 033	19 286 420
2	OVH	14 011	1 691 721
3	GTS Telecom	8160	234 149
4	A2 Hosting	8068	858 932
5	DigitalOcean	6556	135 910

QUIC Measurements

Which QUIC versions are deployed?

We regularly scanned with ZMap between February and May 2021:

- 50 % of found targets still supported Google QUIC versions
- More than 90 % supported the latest draft that should be deployed (Draft-29)
- First deployments announced Version 1 even before the final RFC release



QUIC Measurements
Can we successfully connect to QUIC servers?

QScanner (`https://github.com/tumi8/QScanner`)

- Stateful scanner based on quic-go that conducts full handshakes
- Supports the latest drafts and Version 1
- Allows HTTP requests after successful handshakes
- Extracts widespread information:
  - connection information
  - TLS properties
  - X.509 certificates
  - HTTP headers

→ We are able to successfully complete handshakes with more than 26 M targets

QUIC Measurements

	IPv4 (%)	
	no SNI	SNI
Total Targets	2 M	17M
Success	7.25	76.06
Version Mismatch	8.83	5.77
Timeout	34.50	11.09
Crypto Error (0x128)	48.26	5.73
Other	1.16	1.35

- Low success rate without a server name identifier
- Version mismatches were mainly due to an iterative roll-out of IETF QUIC at Google
  - They do not occur in current scans
- Including the server name identifier drastically increases the success rate
  - Addresses from ZMap without domains have to be treated carefully

Can we identify different QUIC deployments based on configurations?

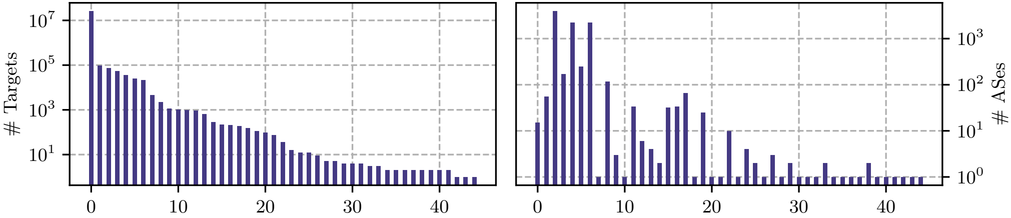
Servers share a set of QUIC Transport Parameters during the handshake:

- 17 different parameters exist, e.g.,
  - initial size of the flow control window
  - the maximum number of allowed streams
- A new TLS extension was defined to send transport parameters (see RFC9001)

→ The QScanner extracts server values

→ Can we identify different QUIC deployments based on configurations?

Can we identify different QUIC deployments based on configurations?



Transport parameters differ within order of magnitudes

- We find 45 different parameter sets
- The most common set is used by Cloudflare and 15 additional ASes
- Three parameter sets are seen in more than 1000 ASes
- Two out of these are seen in combination with a single HTTP Server header value:
  - *proxygen-bolt*

→ These targets are edge PoPs from Facebook and not set up by individuals

Conclusion

→ Different means to detect QUIC deployments exist, each offering unique targets

→ Widespread deployment of QUIC can be found

- more than 2M addresses in 4700 ASes

→ The overall state was solid and ready for the RFC release

- 26 M targets result in successful handshakes
- More than 90 % of targets support the latest draft or version 1

→ Mainly driven by large providers

- We identified deployments in many ASes as edge PoPs of large providers

BACnet

BACnet: Building Automation and Control Networks

- Used to control heating, solar panels, ventilation and other building automation aspects
- Unsolicited access can have real-world consequences
  - Presence detection

  → Break into home
  - Manipulate heating, water flow, . . .
- Security & safety critical protocol

→ evaluate BACnet deployment

BACnet

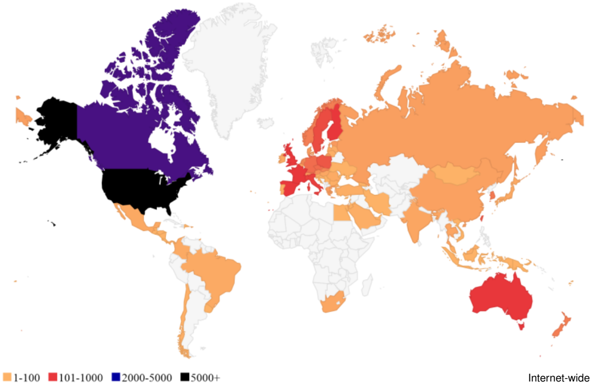
BACnet Protocol:

- Simple UDP-based request-response protocol
- Default port: UDP/47808
- BACnet devices have properties (e.g. device name, temperature, heating level) which can be set and retrieved
  - SingleProperty message
  - MultiProperty message
- No security built in

BACnet

Internet-wide BACnet scans [10]

- Conducted two Internet-wide scans (SingleProperty, MultiProperty)
  - Found 13 k devices
- Evaluated deployment
  - Vendors: Top 5 → ~65%
  - ASes: Top 5 → 30%
  - Countries → see figure



Internet-wide BACnet scans [10]

- Amplification attack vulnerability characteristics
  - Stateless → UDP
  - No authentication
  - Larger response → client can choose returned property
- Amplification
  - Factor of 10-30x possible
  - Extreme example: Hwy 57; Located in the silver box on the electrical pole in front of Grove Primary Care Clinic. Pole 688

BACnet
Summary

Active security measurements can help to improve the Internet‘s security

- Find insecure device and network configurations and notify affected parties
- Analyze deployment over time to observe remediation
- Find weaknesses in protocols
- Identify protocols vulnerable to amplification attacks before they are being exploited

Internet-wide Measurements

IPv6 Scans

Why do we need Internet measurements?

Security Measurements

Passive Measurements

Impact of COVID-19 Pandemic on the Internet

Bibliography

Passive Measurements

Methodology

- Observation of existing traffic using monitoring probes in the network
- Measurement of traffic volume, traffic composition, packet inter-arrival times
- Different levels of granularity
  - Packet-level
  - Flow-level
  - Link-level

Passive Measurements

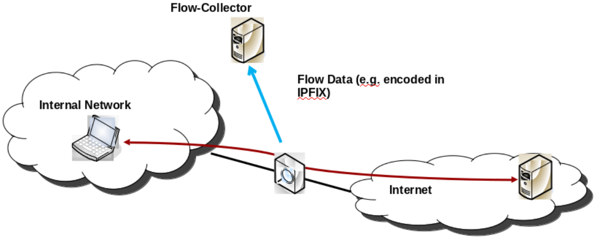
Applications

- Traffic analysis
  - Traffic engineering
  - Anomaly detection
- Accounting
  - Resource utilization
  - Accounting and charging
- Security
  - Intrusion detection
  - Detection of prohibited data transfers (e.g., P2P applications)
- Research

Issues

- Protection of measurement data against illegitimate use (encryption, . . .)
- Applicable law (“lawful interception”, privacy laws, . . .)

Passive Measurements
Flow-Level

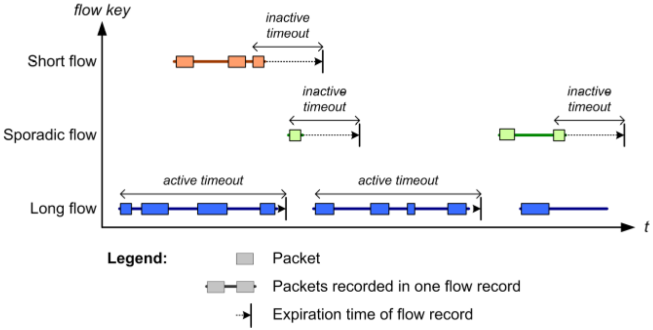


- Network devices create flow data
- Flow data exported to a central collector
- Evaluate communication patterns

Passive Measurements
Flow-Level

Export timeouts to trigger flow expiration

- Inactive timeout
  - → export at the end of flow
- Active timeout
  - → export periodically for long-lived flows
- Timeouts can be configured



Passive Measurements
Flow Data

Flows describe packets which belong together

- E.g. all packets in a TCP connection, i.e. with same 5-tuple:
  - Source IP Address
  - Destination IP Address
  - Transport Protocol
  - Source Port
  - Destination Port
- Various flow metrics can be generated
  - Number of Packets
  - Number of Bytes
  - Duration

Passive Measurements
IPFIX

IPFIX (IP Flow Information eXport) is a protocol to export flow data

- Open: defined by the IETF in RFCs (3917, 3955, 5103, 5153, 5470, 7011, 7012, 7014, 7015)
- Standard track protocol based on Cisco Netflow v5 - v9
- Extensible: Companies can add their own flow definitions and metrics

IPFIX format differentiates between

- Template Records
- Data Records

Design approach: separate flow metric definition from actual data

→ compact data format

Passive Measurements
IPFIX Approach

- Flow definition
  - NetFlow: Flows are always represented by IP 5-tuple
  - IPFIX & Flexible NetFlow: Flows can have arbitrary flow keys
- Update statistic counters of appropriate flow for each arriving packet
- Whenever a flow is terminated its record is exported
  - E.g. TCP FIN, TCP RST, timeout
- Sampling algorithms can reduce the number of flows to be analyzed
  - E.g. update flow cache only for every 10,000th packet
- Transport protocol:
  - SCTP must be implemented, TCP and UDP may be implemented
  - SCTP should be used
  - TCP may be used
  - UDP may be used (with restrictions – congestion control!)

Passive Measurements
IPFIX - Terminology

IP Traffic Flow

- A flow is defined as a set of IP packets passing an observation point in the network during a certain time interval.
- All packets belonging to a particular flow have a set of common properties.

Observation Point

- The observation point is a location in the network where IP packets can be observed.
- One observation point can be a superset of several other observation points.

Metering Process

- The metering process generates flow records.
- It consists of a set of functions that includes
  - packet header capturing
  - timestamping
  - sampling
  - classifying
  - and maintaining flow records.

Passive Measurements
IPFIX - Terminology

Flow Record

- A flow record contains information about a specific flow that was metered at an observation point.
- A flow record contains measured properties of the flow (e.g. the total number of bytes of all packets of the flow) and usually also characteristic properties of the flow (e.g. the source IP address).

Exporting Process

- The exporting process sends flow records to one or more collecting processes.
- The flow records are generated by one or more metering processes.

Collecting Process

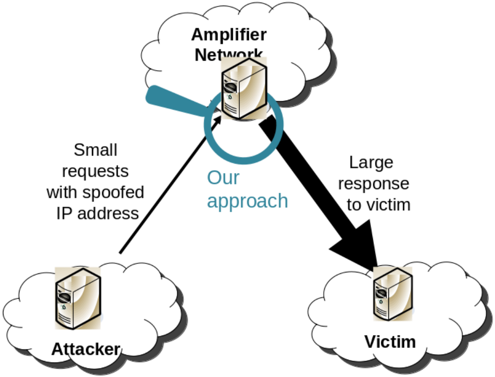
- The collecting process receives flow records from one or more exporting processes for further processing.

Passive Measurements
Amplification Attack Detection

- Example for amplification attack: short UDP packet with DNS request and spoofed IP packet resulting in large response
- Amplification attacks can have drastic effect on network availability
- Goal: Detect amplification attacks at the amplifier [11]
- Use traffic characteristics to discern benign from amplification traffic
- Many protocols can be abused for this type of attack [12]
  - Network services (NTP, SNMP, SSDP and NetBios)
  - Legacy services (CharGen and QOTD)
  - P2P networks (BitTorrent and Kademlia)
  - Game servers (Quake 3 and Steam)
  - P2P-based botnets

Passive Measurements
Amplification Attack Detection

Detect amplification attacks at the amplifier [11]



Passive Measurements
Amplification Attack Detection

Detect amplification attacks at the amplifier [11]

Detection methodology

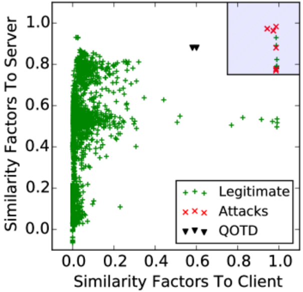
- Amplification factor
    - Attacker sends packets that generate larger response than request
    - → Asymmetric traffic can be indicator for amplification attack
- Packet size similarity
    - Attacker sends few variations of packets that are sure to create large amplification factor → similar length
    - → Similar packet sizes can be indicator for amplification attack
- Payload similarity
    - Attacker sends few variations of packets that are sure to create large amplification factor → similar payload content
    - → Similar payload can be indicator for amplification attack
- Unsolicited ICMP messages
    - Victim does not expect amplification traffic
    - → Backscatter ICMP can be indicator for amplification attack
- TTL measurements
    - Path from attacker to amplifier $\neq$ path from amplifier to victim
    - → Different path length can be indicator for amplification attack

Passive Measurements
Amplification Attack Detection

Detect amplification attacks at the amplifier [11]

How can we compare payload similarity of packets within one flow?

- Similar data has low entropy
- Compression determines entropy as a side product
  - Repetitive data
  - → highly compressible
  - Different data
  - → bad compression factor



Passive Measurements
Amplification Attack Detection

Summary:

- Amplification Attack: Small request of spoofed traffic → large response sent to victim (DoS)
- Detection at amplifier allows to see request and response
- Flow data can help to tackle (performance & encryption) challenges
- Characteristics of flow data well suited to detect amplification traffic

Internet-wide Measurements

IPv6 Scans

Why do we need Internet measurements?

Security Measurements

Passive Measurements

Impact of COVID-19 Pandemic on the Internet

Bibliography

Impact of COVID-19 Pandemic on the Internet
Introduction

- Pandemic is a rare and special event
- Work from home and Stay at home orders posed challenges to the Internet
- Fundamental importance of the Internet and digitalization in general to these measures

Impact of COVID-19 Pandemic on the Internet
Introduction

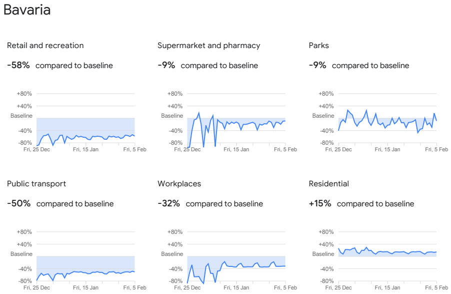
- Pandemic is a rare and special event
- Work from home and Stay at home orders posed challenges to the Internet
- Fundamental importance of the Internet and digitalization in general to these measures
- Expectation
    - Increased load with abnormal patterns and access points
    - Higher load on residential networks
    - General higher load due to higher media consumption and video conferencing

Impact of COVID-19 Pandemic on the Internet
Introduction

- Pandemic is a rare and special event
- Work from home and Stay at home orders posed challenges to the Internet
- Fundamental importance of the Internet and digitalization in general to these measures
- Expectation
    - Increased load with abnormal patterns and access points
    - Higher load on residential networks
    - General higher load due to higher media consumption and video conferencing
- Overall the Internet managed to handle the traffic increase

Impact of COVID-19 Pandemic on the Internet
Motivation

- Google and Apple provided mobility reports based on their data
- What is the effect on the Internet?



Bavaria

Retail and recreation

-58% compared to baseline

Supermarket and pharmacy

-9% compared to baseline

Parks

-9% compared to baseline

Public transport

-50% compared to baseline

Workplaces

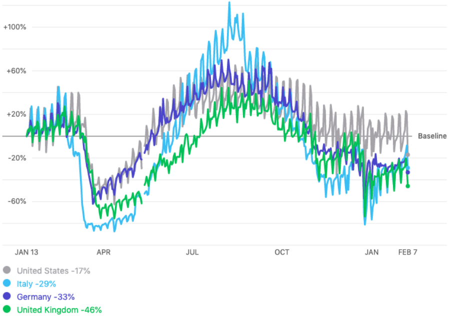
-32% compared to baseline

Residential

+15% compared to baseline

Google Mobility Report

`https://www.google.com/covid19/mobility/`



Apple Mobility Report

`https://covid19.apple.com/mobility`

Impact of COVID-19 Pandemic on the Internet
IMC 2020

- Early research from IMC 2020[3]
- Submission deadline was in begin of June 2020
- Presentations were in October 2020
- Four interesting papers on the topic:
  - Feldmann et al., The Lockdown Effect: Implications of the COVID-19 Pandemic on Internet Traffic [13]
  - Lutu et al., A Characterization of the COVID-19 Pandemic Impact on a Mobile Network Operator Traffic [14]
  - Fontugne et al., Persistent Last-mile Congestion: Not so Uncommon [15]
  - Böttger et al., How the Internet reacted to Covid-19 – A perspective from Facebook's Edge Network [16]

[3] https://conferences.sigcomm.org/imc/2020/

The Lockdown Effect [13]

Weekend effect

Approach by Feldmann et al. [13]

- Compared traffic volume throughout the day on a Wednesday and a Saturday, pre and during lockdown

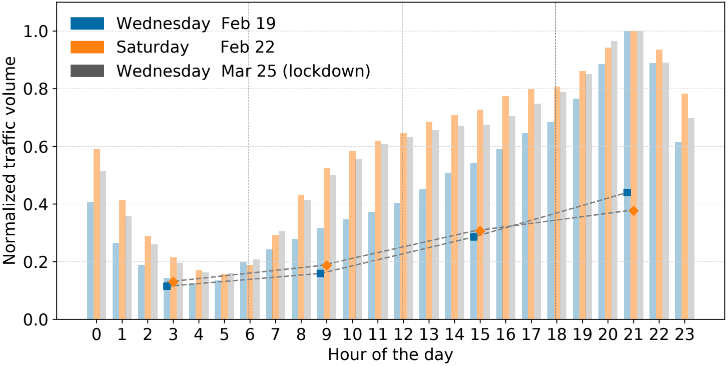


Figure 2a by Feldmann et al. [13]

The Lockdown Effect [13]

ISP Day Patterns

- They used the learned pattern and assigned each day a label
- Blue if the day matches the usual pattern (e.g. Sunday with weekend pattern)
- Orange if it does not match (Wednesday with weekend pattern)
- Data from a Central European ISP

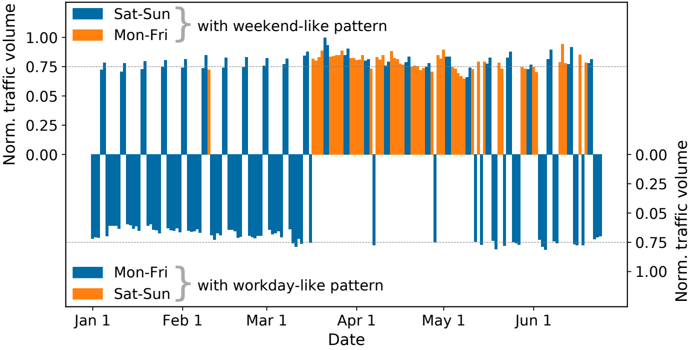


Figure 2b by Feldmann et al. [13]

The Lockdown Effect [13]

IXP Day Patterns

- Same approach as before from a Central European IXP

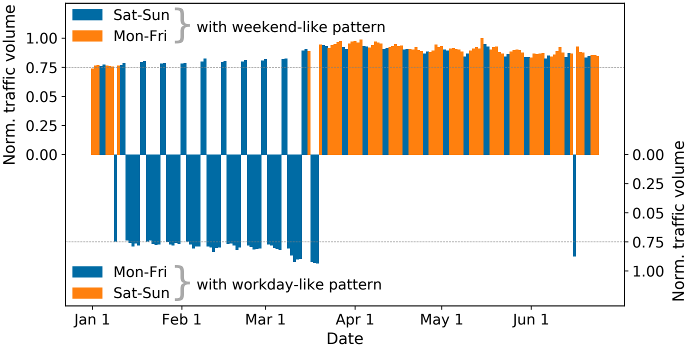


Figure 2c by Feldmann et al. [13]

The Lockdown Effect [13]
Hypergiants

Definition

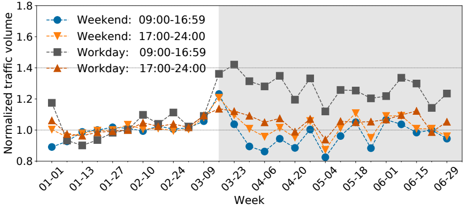
- Originally called so by Arbor networks
- First defined by Labovitz et al. [17]
- Describes companies which generate a disproportionate share of the traffic (high outbound traffic ratios)
- E.g. Google, Netflix, Cloudflare, Akamai

The Lockdown Effect [13]

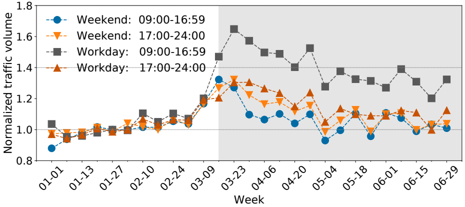
Hypergiants traffic

Analysis by Feldmann et al. [13]

- Used NetFlow and IPFIX data to analyze traffic of hypergiants
- No difference between the four categories until lockdown
- Increase of hypergiants by 40 %
- Other ASes increase by about 60 %



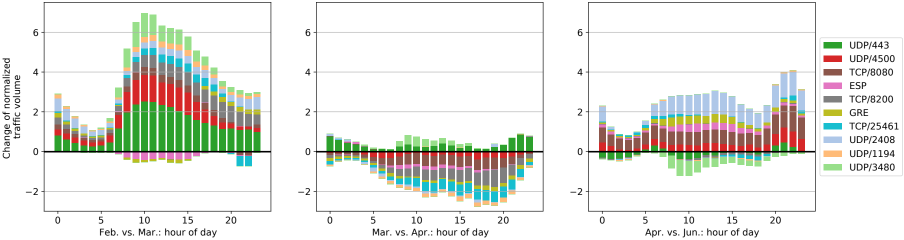
Hypergiants traffic. Figure 4a by Feldmann et al. [13]



Other ASes traffic. Figure 4b by Feldmann et al. [13]

The Lockdown Effect [13]
Transport Layer Analysis

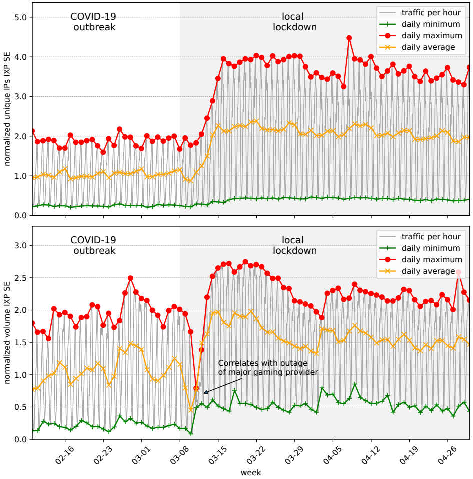
- By analyzing the used destination ports Feldmann et al. [13] inferred service usage
- UDP/443 is QUIC and mainly used by Google and Akamai
- UDP/4500 is for IPSec NAT traversal
- GRE and ESP transport the real IPSec traffic
  - Usually mainly used between companies
- TCP/8200 and TCP/25461 are used by TV streaming services



IXP in Central Europe. Figure 7 by Feldmann et al. [13]

The Lockdown Effect [13]
Gaming Category

- Filters for 5 ASNs and 57 known gaming related ports
- Used number of IP addresses as an abstraction for households
- Data shown is from an IXP in Southern Europe



Correlates with outage of major gaming provider

Figure 9 by Feldmann et al. [13]

The Lockdown Effect [13]
All Categories

- All labeled categeries
- Paper also contains the graphs for the Central European IXP and Southern European IXP

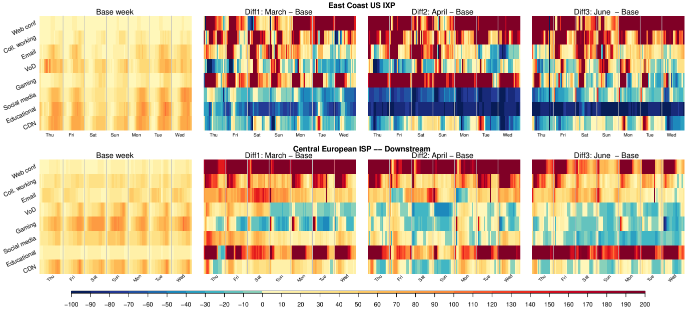


Figure 10 by Feldmann et al. [13]

Characterization of the Pandemic [14]

Analysis of the Mobile Network

Analysis by Lutu et al. [14]:

- Investigated effect on UK Mobile Network Operator (Telefonica)
- E.g.: Used the cell data to quantify mobility
- Can provide local data for cities and city districts
- Expecially analyzed mobility of inner London residents (see figure below)

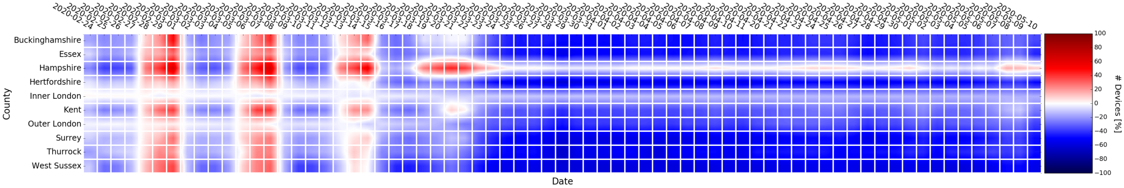


Figure 7 by Lutu et al. [14]

Last-mile Congestion [15]
Inferring Congestion from Traceroutes

Approach by Fontugne et al. [15] to analyze last-mile congestion

- Uses data from RIPE Atlas
- Subtracted latency of last non public routed address from latency of first public routed address
- Apply medians on 30 minute buckets to reduce noise
- Compute queuing delay by observing deviation from minimum median RTT value

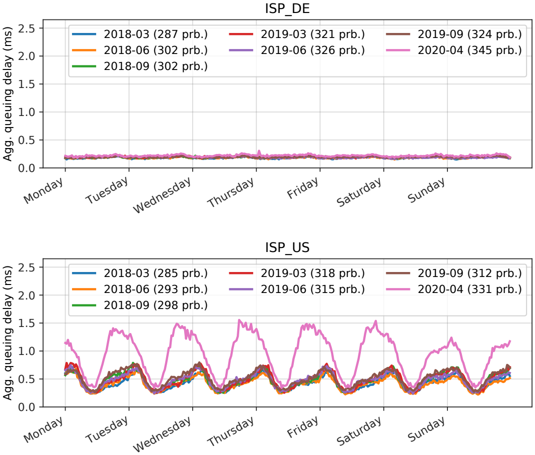


Figure 1 by Fontugne et al. [15]

Last-mile Congestion [15]

Inferring Congestion from Traceroutes

- Uses frequency analysis to find last-mile congestions
- Finds persistent last mile congestion for the US ISP
- Number of congested ASes increseas from 10% to 55% during in April 2020

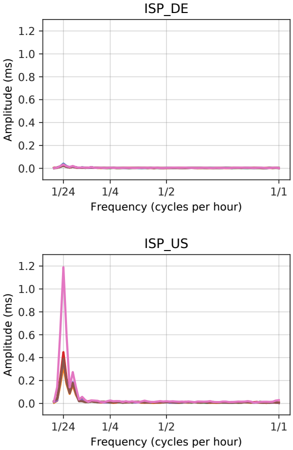
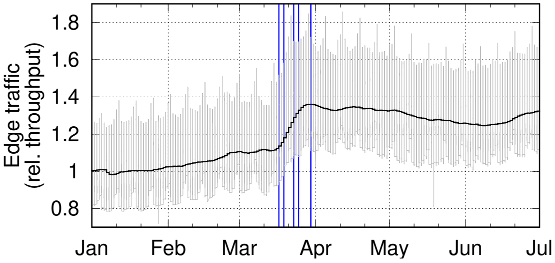


Figure 2 by Fontugne et al. [15]

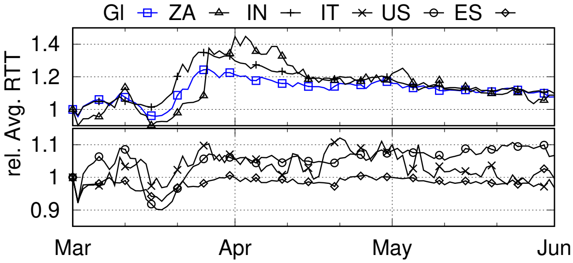
A Perspective from FBs edge [16]
Main Contributions

Approach by Böttger et al. [16]

- Used data collected at Facebooks edge to infer changes



Total traffic growth. Figure 1 by Böttger et al. [16]



Change in latency of selected countries. Figure 11 by Böttger et al. [16]

Internet-wide Measurements

IPv6 Scans

Why do we need Internet measurements?

Security Measurements

Passive Measurements

Impact of COVID-19 Pandemic on the Internet

Bibliography

Internet-wide Measurements

[1] J. Zirngibl, L. Steger, P. Sattler, O. Gasser, and G. Carle, “Rusty Clusters? Dusting an IPv6 Research Foundation,”, Nice, France, 2022.

[2] D. Plonka and A. Berger, “Temporal and spatial classification of active ipv6 addresses,” in Proceedings of the 2015 Internet Measurement Conference, ser. IMC ’15, Tokyo, Japan: Association for Computing Machinery, 2015, 509–522, ISBN: 9781450338486.

[3] E. C. Rye and R. Beverly, “Discovering the ipv6 network periphery,” in Passive and Active Measurement, A. Sperotto, A. Dainotti, and B. Stiller, Eds., Cham: Springer International Publishing, 2020, pp. 3–18.

[4] E. Rye, R. Beverly, and K. C. Claffy, “Follow the scent: Defeating ipv6 prefix rotation privacy,” in Proceedings of the 21st ACM Internet Measurement Conference, ser. IMC ’21, Virtual Event: Association for Computing Machinery, 2021, 739–752, ISBN: 9781450391290.

[5] R. Holz, L. Braun, N. Kammenhuber, and G. Carle, “The ssl landscape: A thorough analysis of the x.509 pki using active and passive measurements,” in Proceedings of the 2011 ACM SIGCOMM Conference on Internet Measurement Conference, ser. IMC ’11, New York, NY, USA: Association for Computing Machinery, 2011, 427–444, ISBN: 9781450310130. [Online]. Available: https://doi.org/10.1145/2068816.2068856.

[6] P. Kotzias, A. Razaghpanah, J. Amann, K. G. Paterson, N. Vallina-Rodriguez, and J. Caballero, “Coming of age: A longitudinal study of tls deployment,” in Proceedings of the Internet Measurement Conference 2018, ser. IMC ’18, New York, NY, USA: Association for Computing Machinery, 2018, 415–428, ISBN: 9781450356190. [Online]. Available: https://doi.org/10.1145/3278532.3278568.

[7] J. Rüth, I. Poese, C. Dietzel, and O. Hohlfeld, “A First Look at QUIC in the Wild,” in Passive and Active Measurement, Springer International Publishing, 2018, pp. 255–268, ISBN: 978-3-319-76481-8.

Internet-wide Measurements

[8] J. Zirngibl, P. Buschmann, P. Sattler, B. Jaeger, J. Aulbach, and G. Carle, "It's over 9000: Analyzing early quic deployments with the standardization on the horizon," in Proceedings of the 2021 Internet Measurement Conference, Virtual Event, USA: ACM, Nov. 2021. DOI: 10.1145/3487552.3487826.

[9] B. M. Schwartz, M. Bishop, and E. Nygren, "Service binding and parameter specification via the DNS (DNS SVCB and HTTPS RRs)," Internet Engineering Task Force, Internet-Draft draft-ietf-dnsop-svcb-https-08, Oct. 2021, Work in Progress, 60 pp. [Online]. Available: https://datatracker.ietf.org/doc/html/draft-ietf-dnsop-svcb-https-08.

[10] O. Gasser, Q. Scheitle, B. Rudolph, C. Denis, N. Schricker, and G. Carle, "The amplification threat posed by publicly reachable bacnet devices," Journal of Cyber Security and Mobility, 2017.

[11] T. Böttger, L. Braun, O. Gasser, F. von Eye, H. Reiser, and G. Carle, "Dos amplification attacks – protocol-agnostic detection of service abuse in amplifier networks," in Traffic Monitoring and Analysis, M. Steiner, P. Barlet-Ros, and O. Bonaventure, Eds., Cham: Springer International Publishing, 2015, pp. 205–218, ISBN: 978-3-319-17172-2.

[12] C. Rossow, "Amplification hell: Revisiting network protocols for ddos abuse," in In Proceedings of the 2014 Network and Distributed System Security Symposium, NDSS, Citeseer, 2014.

[13] A. Feldmann, O. Gasser, F. Lichtblau, et al., "The lockdown effect: Implications of the covid-19 pandemic on internet traffic," in Proceedings of the ACM Internet Measurement Conference, ser. IMC '20, New York, NY, USA: Association for Computing Machinery, 2020, 1–18.

[14] A. Lutu, D. Perino, M. Bagnulo, E. Frias-Martinez, and J. Khangosstar, "A characterization of the covid-19 pandemic impact on a mobile network operator traffic," in Proceedings of the ACM Internet Measurement Conference, ser. IMC '20, New York, NY, USA: Association for Computing Machinery, 2020, 19–33, ISBN: 9781450381383.

Internet-wide Measurements

[15] R. Fontugne, A. Shah, and K. Cho, “Persistent last-mile congestion: Not so uncommon,” in Proceedings of the ACM Internet Measurement Conference, ser. IMC ’20, New York, NY, USA: Association for Computing Machinery, 2020, 420–427, ISBN: 9781450381383.

[16] T. Böttger, G. Ibrahim, and B. Vallis, “How the internet reacted to covid-19: A perspective from facebook’s edge network,” in Proceedings of the ACM Internet Measurement Conference, ser. IMC ’20, New York, NY, USA: Association for Computing Machinery, 2020, 34–41, ISBN: 9781450381383.

[17] C. Labovitz, S. Iekel-Johnson, D. McPherson, J. Oberheide, and F. Jahanian, “Internet inter-domain traffic,” in Proceedings of the ACM SIGCOMM 2010 Conference, ser. SIGCOMM ’10, New York, NY, USA: Association for Computing Machinery, 2010, 75–86, ISBN: 9781450302012.