



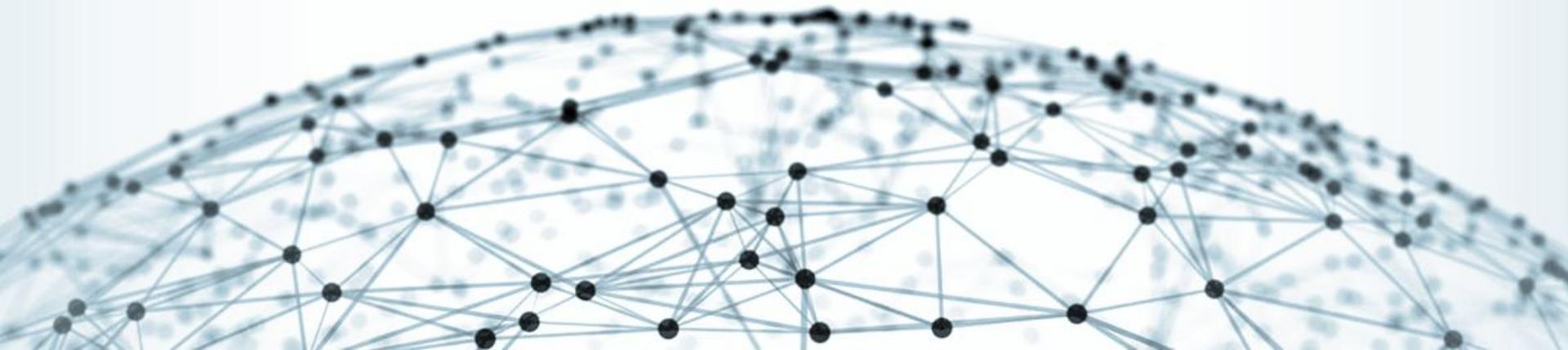
Censored Planet

Censored Planet

An Internet-wide, Longitudinal Censorship Observatory

Ram Sundara Raman, Prerana Shenoy, Katharina Kohls*, Roya Ensafi

University of Michigan, *Ruhr University Bochum



Internet censorship is pervasive and evolving



**Measuring Internet
censorship globally is
a complex problem**

Measuring Internet censorship is complex



Censorship
method variance

Measuring Internet censorship is complex



Censorship method variance

DNS manipulation

TCP/IP blocking

Application layer (HTTP)
connection

Measuring Internet censorship is complex

Censorship
method variance

Geographical and
Network variance



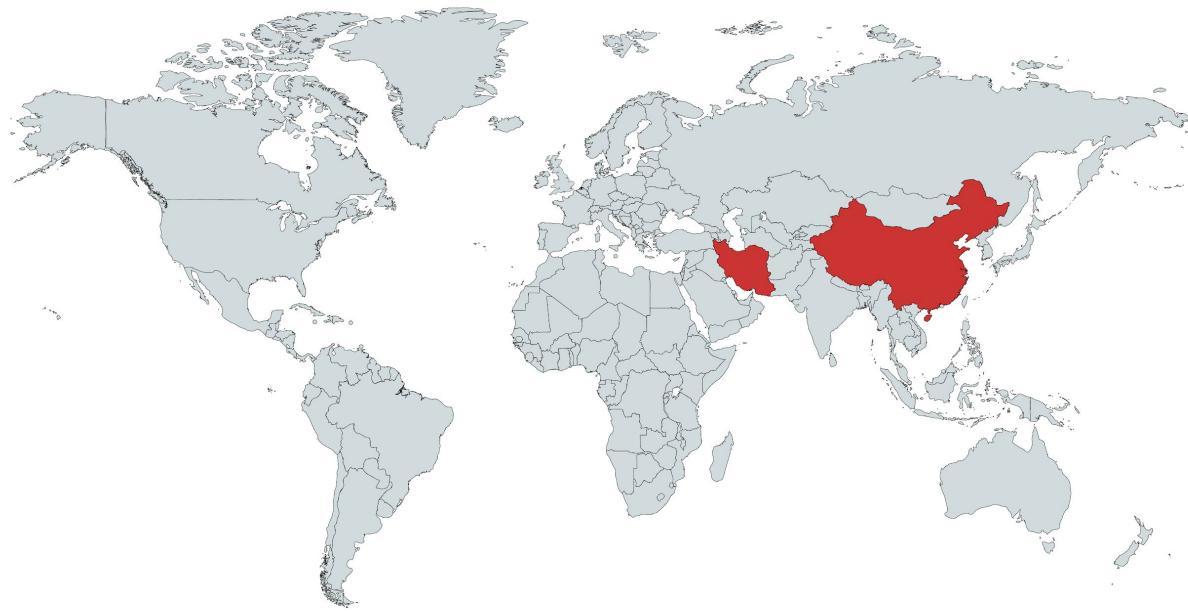
Measuring Internet censorship is complex



Censorship
method variance

Geographical and
Network variance

Longitudinal
variance

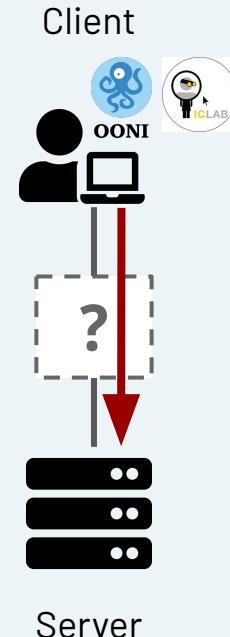


Previous studies: Few countries and limited snapshots

Direct Censorship

Measurement Platforms

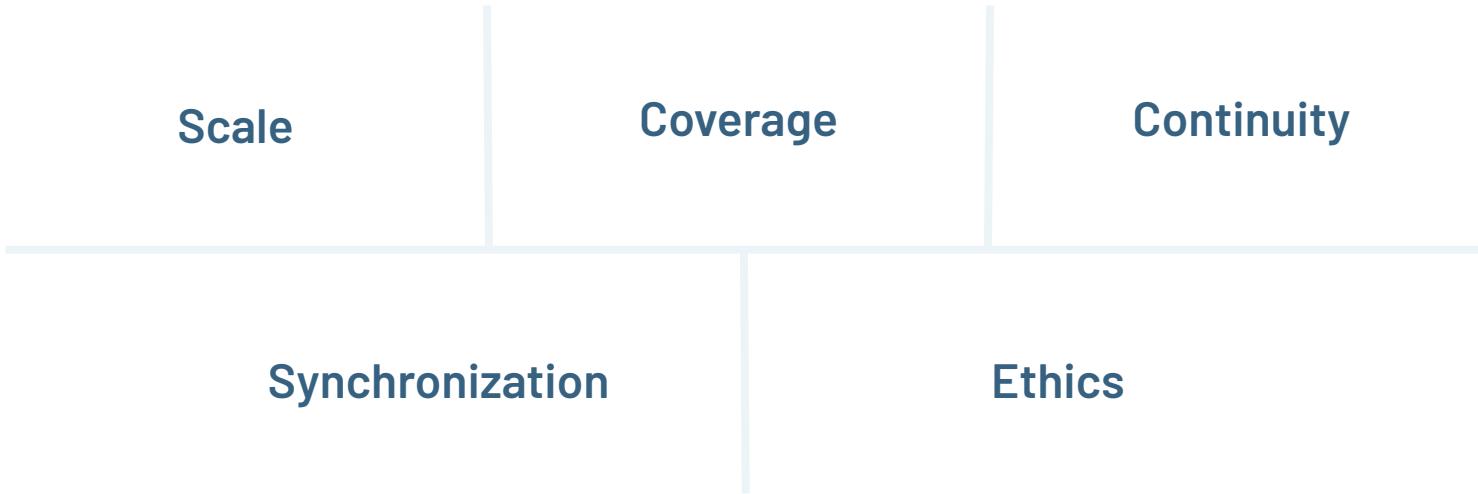
- Ask **volunteers** on the ground, or deploy software or hardware in censored region (e.g. OONI probe^[1])
- Use **VPNs**, or **research networks** (e.g. ICLab^[2])



[1] Open Observatory of Network Interference, <https://ooni.org>

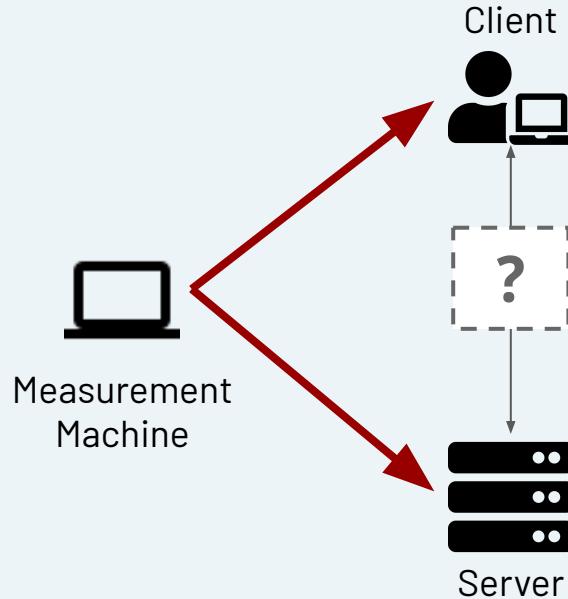
[2] A. Akhavan Niaki, S. Cho, Z. Weinberg, N. P. Hoang, A. Razaghpanah, N. Christin, and P. Gill. ICLab: A Global, Longitudinal Internet Censorship Measurement Platform. In IEEE Symposium on Security and Privacy (SP), 2020.

Limitations of Direct Measurements

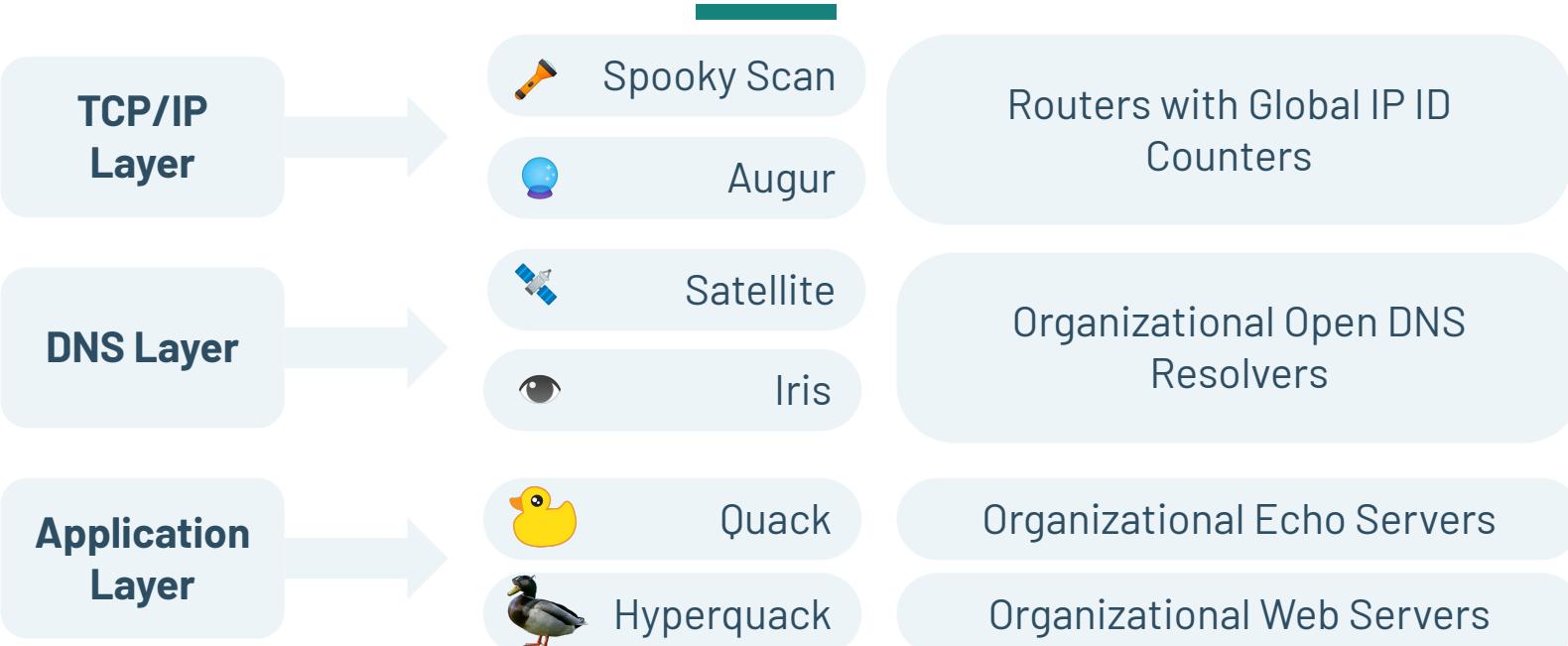


Remote Censorship Measurements

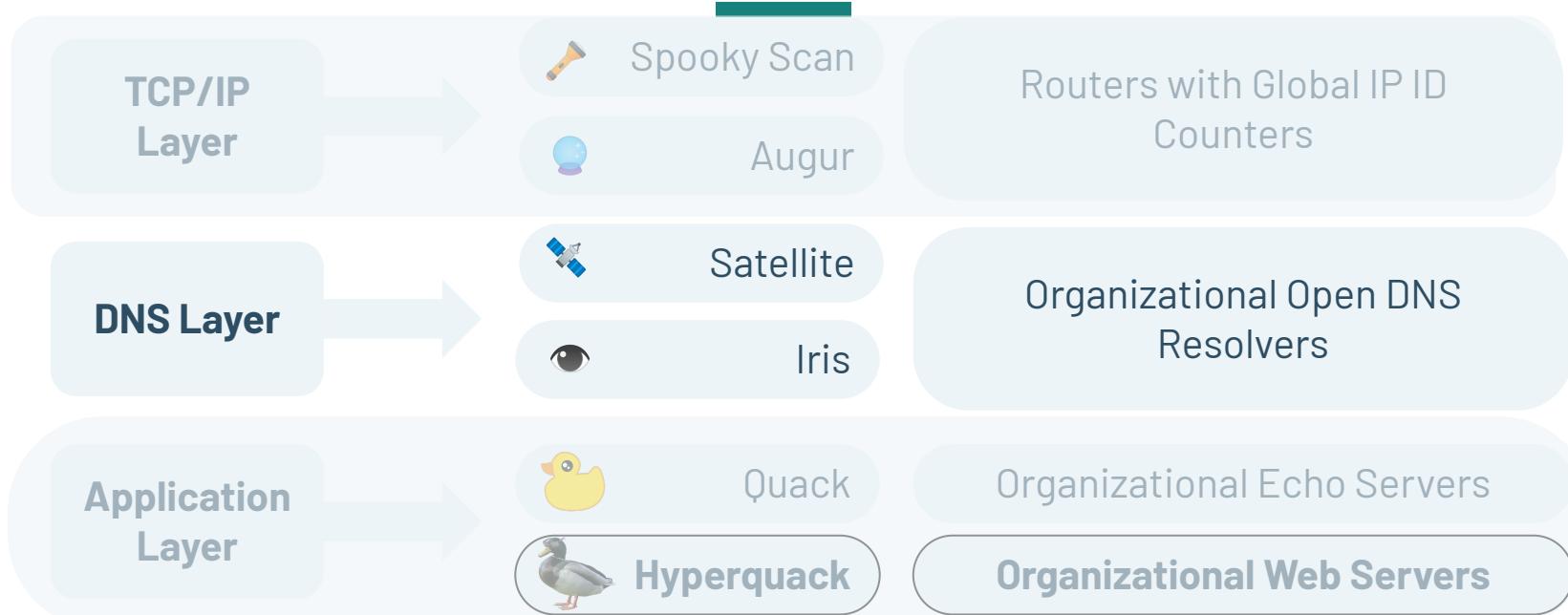
Detect whether pairs of hosts around the world can talk to each other without controlling either endpoint.



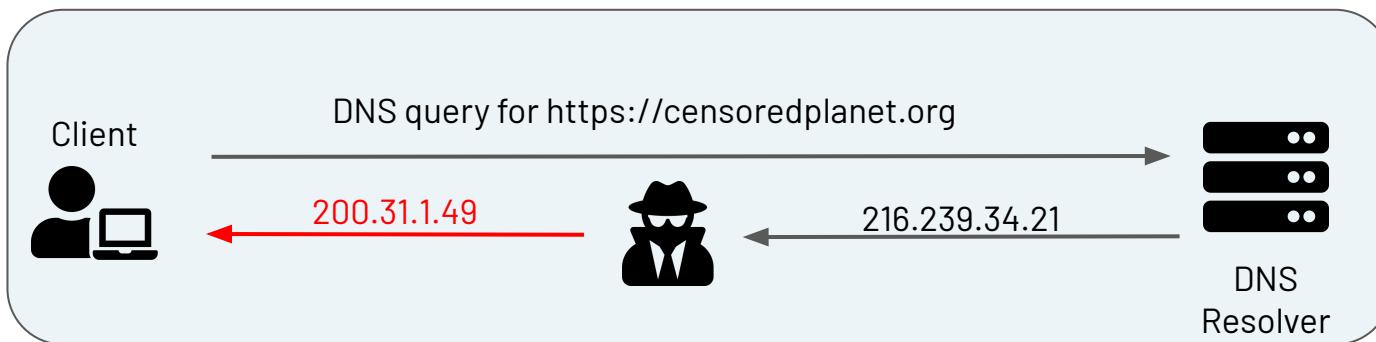
Remote Measurement Techniques

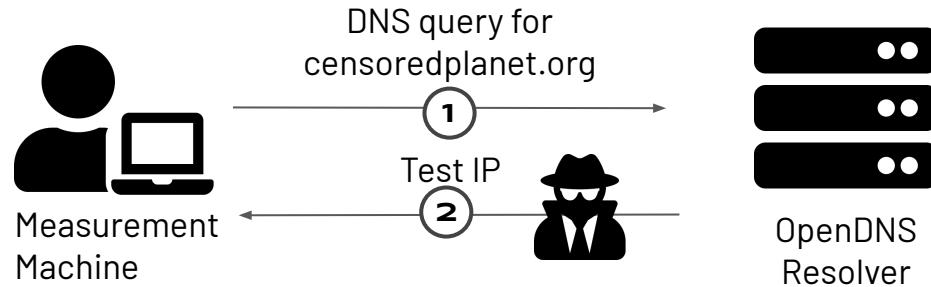


Remote Measurement Techniques



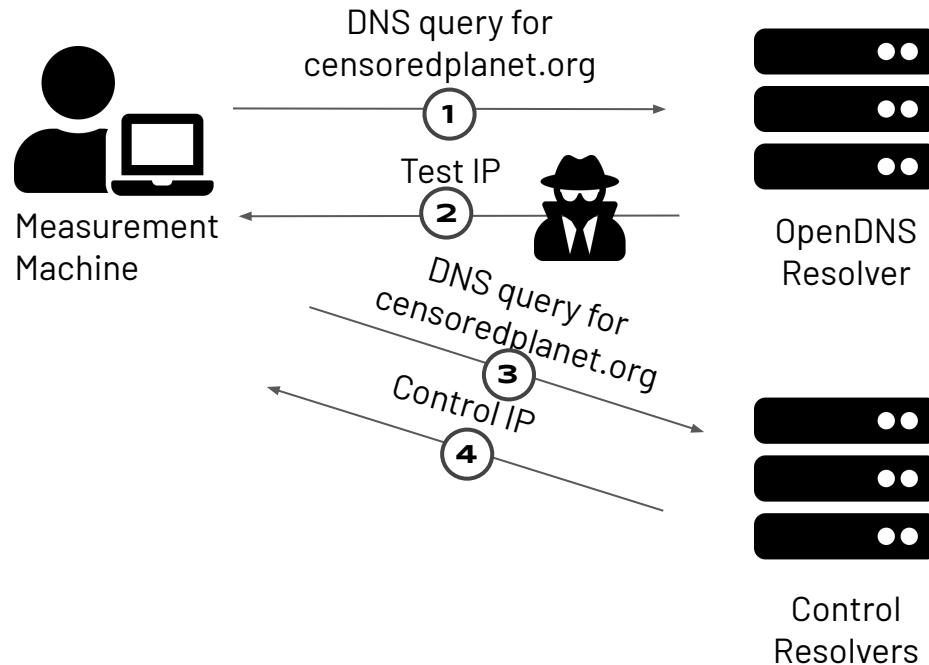
DNS Manipulation



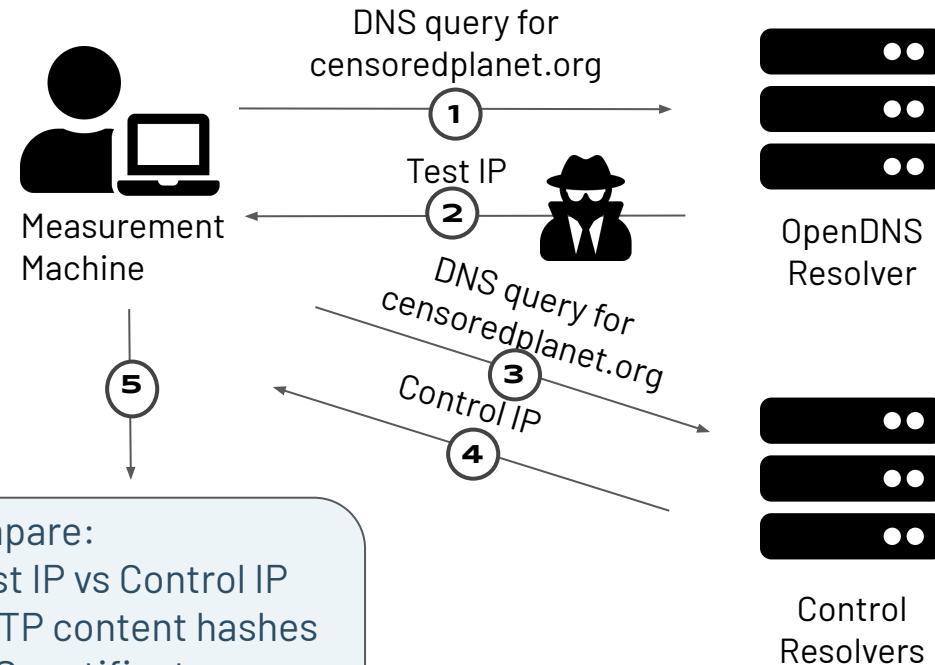


Satellite & Iris

Satellite & Iris



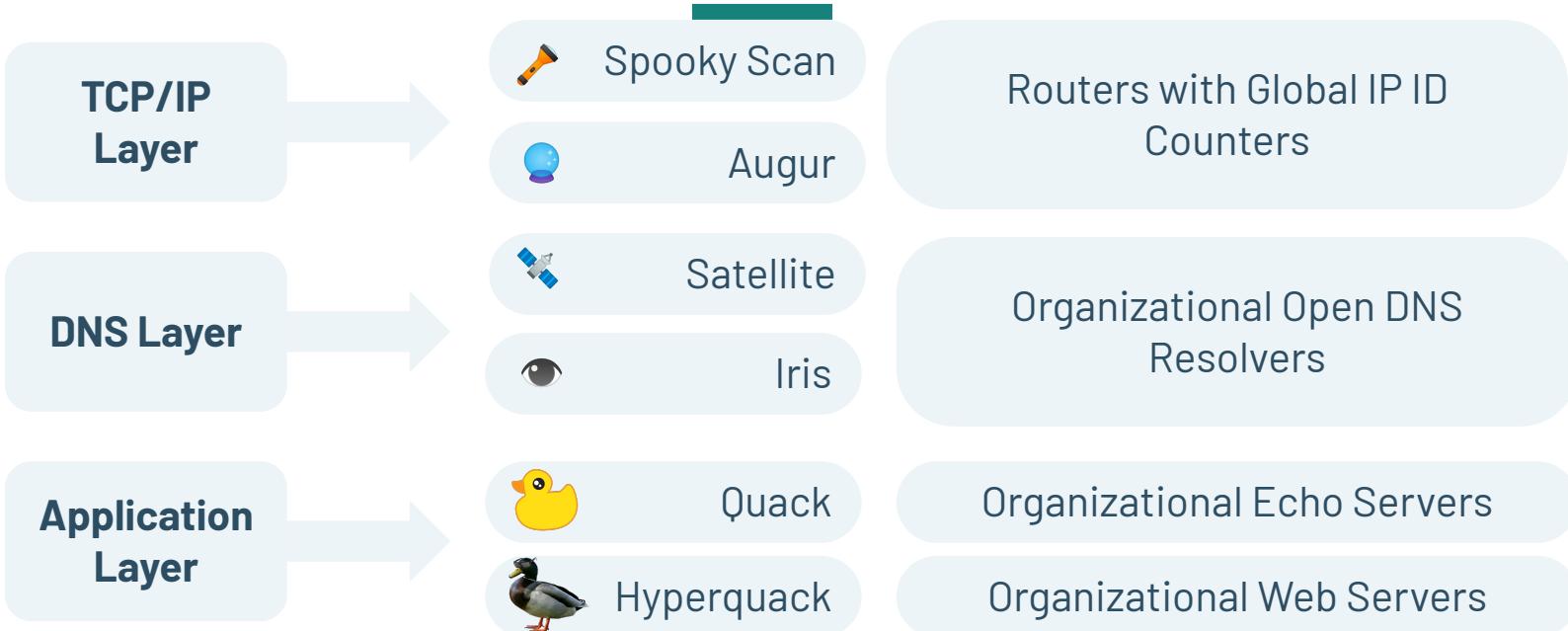
Satellite & Iris



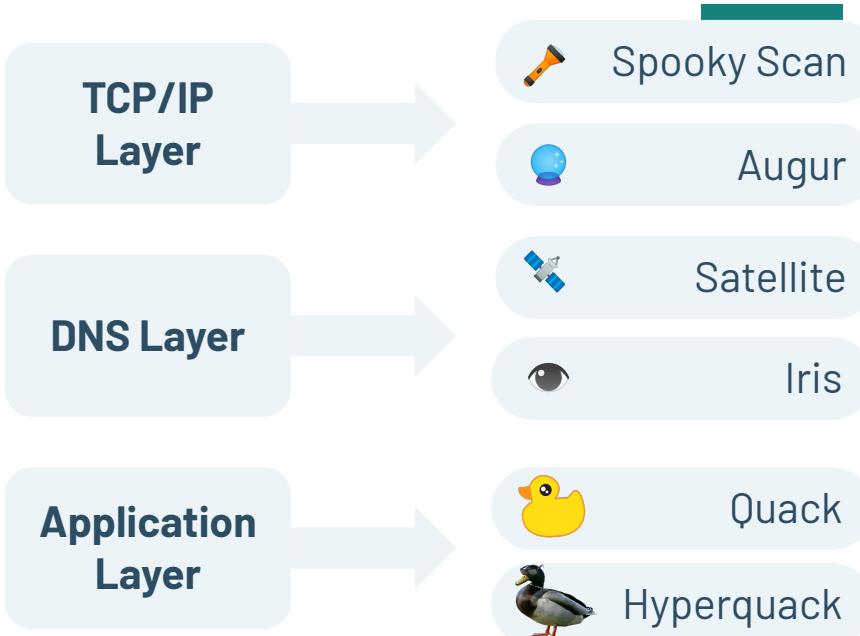
Satellite Scale, Coverage and Ethics

- More than 8.2 million OpenDNS resolvers in 232 countries
- To reduce risk, we want to choose infrastructural resolvers
- We use resolvers with a valid PTR record beginning with the subdomain ns[0-9]* or nameserver[0-9]* → Likely to be part of big organizations
- 30k resolvers in ~4,500 ASes in 175 countries
- Stable DNS resolvers allow us to repetitively run measurements over time

Remote Measurement Techniques



Remote Measurement Techniques



Limitations:

- Specialized techniques
- Limited snapshots
- Labor intensive
- Accuracy

Censored Planet Observatory

- Collect global censorship measurement data continuously using remote measurement techniques (Augur, Satellite, Quack, Hyperquack)
- Analyze the data to create a more complete and accurate view of global Internet censorship
- Custom rapid focus measurements to analyze censorship events quickly



Censored Planet: An Internet-wide, Longitudinal Censorship Observatory. ACM CCS 2020. R. Sundara Raman, P. Shenoy, K. Kohls, R. Ensafi.



Censored Planet

Censored Planet Observatory

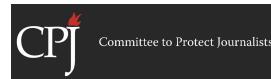
- Collect global censorship measurement data continuously using remote measurement techniques (Augur/Spooky, Satellite/Iris, Quack, Hyperquack)
- Analyze the data to create a complete and accurate view of global censorship
- Custom rapid focus measurements to analyze censorship events quickly



Research team investigating Internet censorship with tracking system

MIT Technology Review

Online censorship in Saudi Arabia soared after Jamal Khashoggi's murder



Laws, cheap web filters arm Russia to block news, says Censored Planet

FINANCIAL TIMES
US blocks Hong Kong users from some government websites

Censored Planet Observatory

- Started in August 2018 and continuously collecting censorship data on 6 Internet protocols (TCP, DNS, Echo, Discard, HTTP, HTTPS)
- Continuous baseline of reachability data for **2000 sensitive domains and IP addresses (From Alexa and Citizen Lab) each week**
- More than **95,000 vantage points in 221 countries and territories** (updated every week)
- Rapid focus capabilities to analyze censorship events in detail

21.8 billion

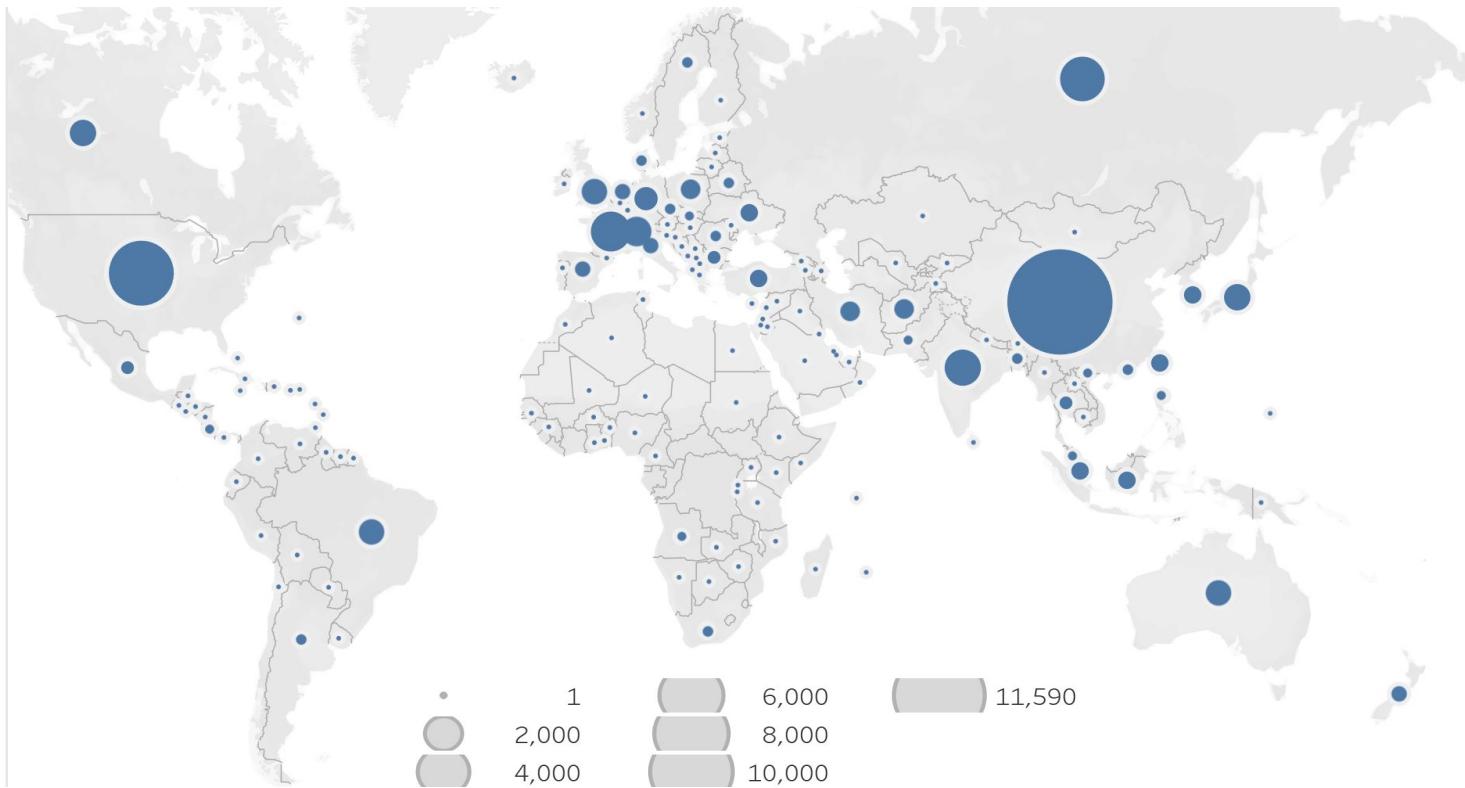
Measurements over 20 Months

221 countries

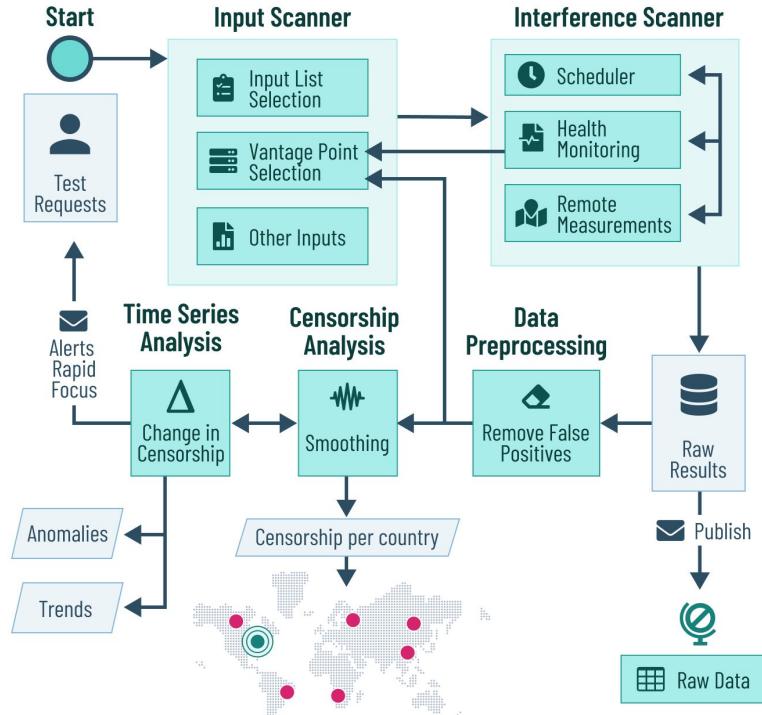
42%-360% increase compared to OONI, ICLab

8 ASes (median)/country

Median increase of 4-7 ASes per country



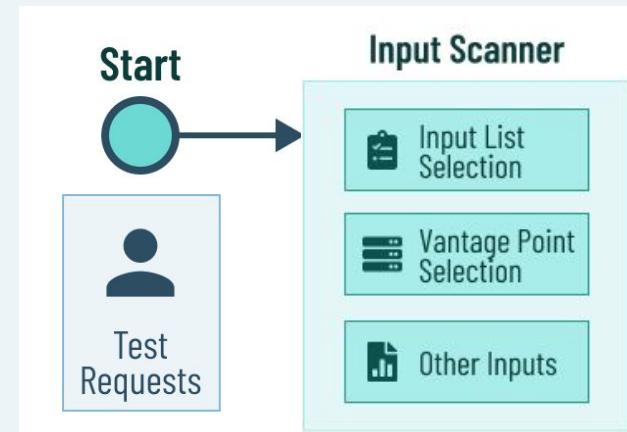
Open DNS resolvers used in January 2021



Modular Design

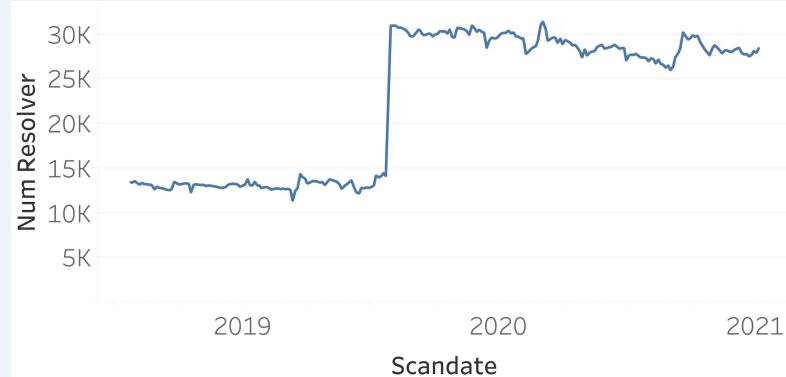
Input Scanner

- **Vantage Point Selection**
 - Internet-wide scans for infrastructural machines
 - Consistency and diversity
- **Test List Selection**
 - Popular domains (Alexa)
 - Sensitive domains (Citizen Lab)
- Updated every week



Input Scanner

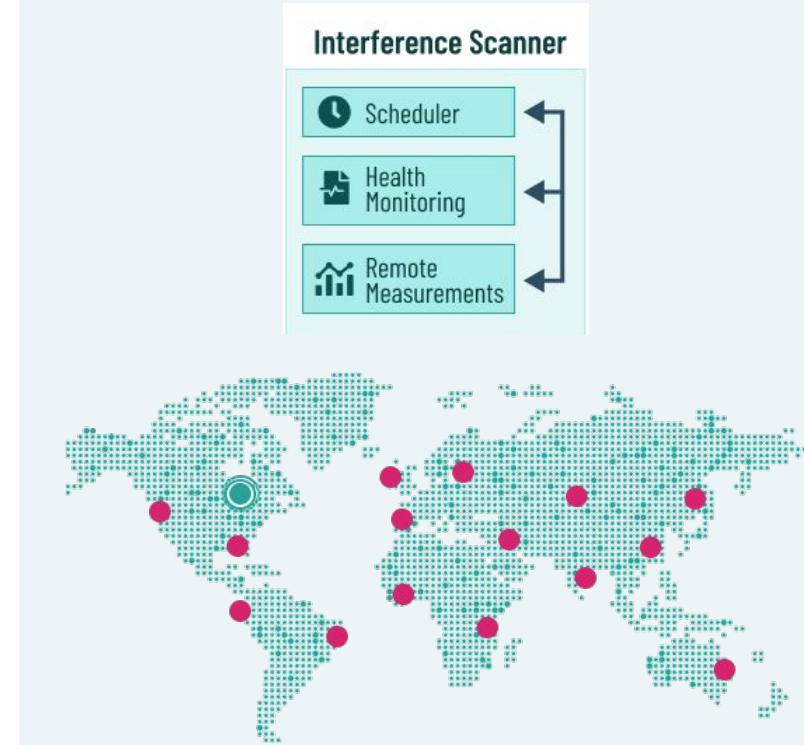
- **Vantage Point Selection**
 - Internet-wide scans for infrastructural machines
 - Consistency and diversity
- **Test List Selection**
 - Popular domains (Alexa)
 - Sensitive domains (Citizen Lab)
- Updated every week



Number of DNS resolvers used
for measurements over time

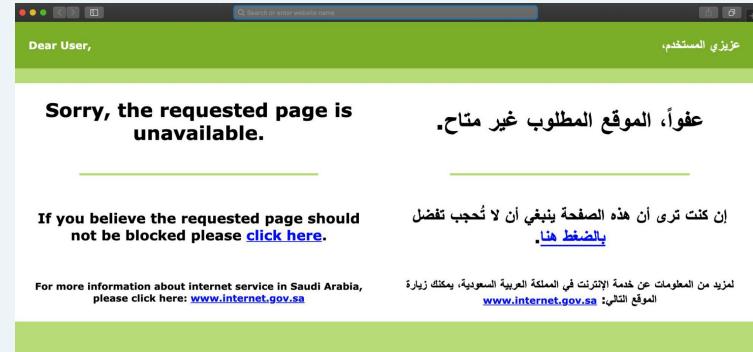
Interference Scanner

- **Scheduler**
 - Manages vantage points and synchronizes measurements
- **Health Monitoring**
 - Vantage point health
 - Measurement errors
- Remote Measurements performed every week - Augur, Satellite, Quack, Hyperquack
- Raw data published



Data Pre-processing

- Aggregating to common data schema
- **Confirming censorship**
 - Use clustering techniques in previous work^[1] to find and group blockpages.
 - Consider clear signs of censorship

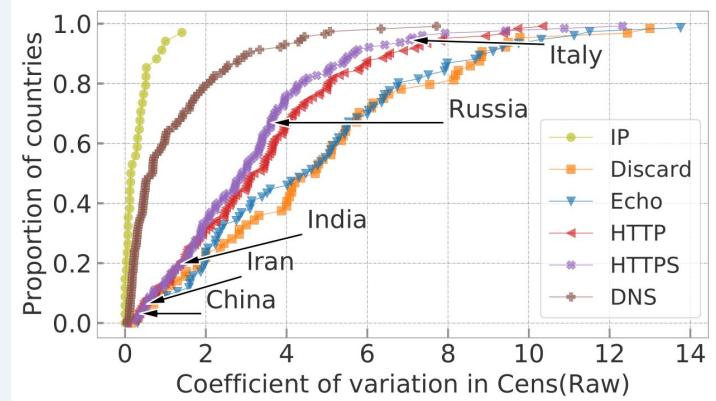


Example blockpage in Saudi Arabia

[1] Measuring the Deployment of Network Censorship Filters at Global Scale, NDSS 2020

Censorship Analysis

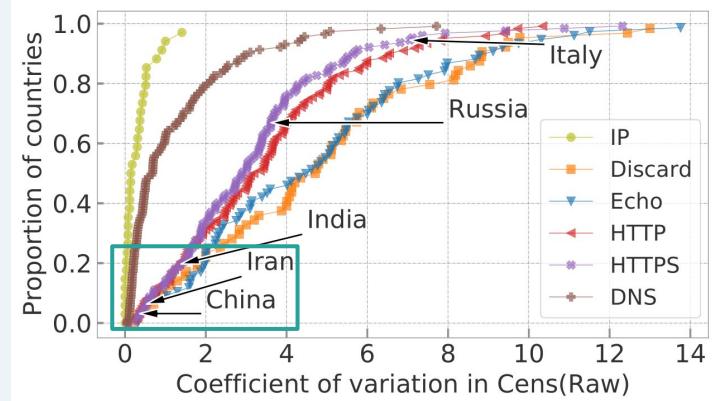
- Censorship values also vary *within* countries
- Countries with heterogeneous censorship policies have high variance in raw censorship values



Coefficient of variation in raw censorship metric

Censorship Analysis

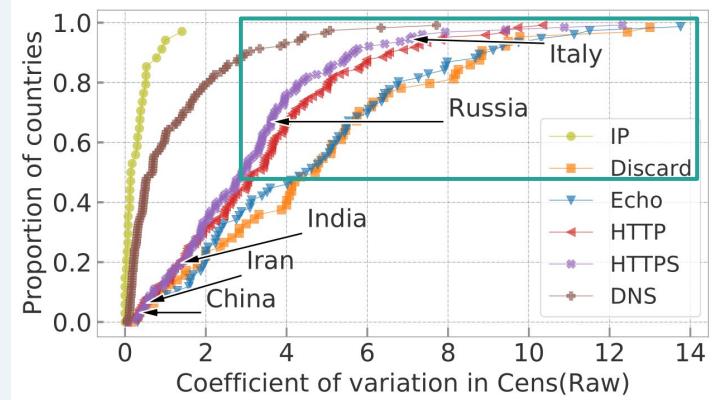
- Censorship values also vary *within* countries
- Countries with heterogeneous censorship policies have high variance in raw censorship values



Coefficient of variation in raw
censorship metric

Censorship Analysis

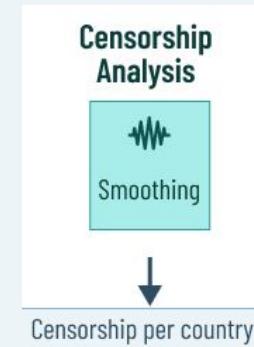
- Censorship values also vary *within* countries
- Countries with heterogeneous censorship policies have high variance in raw censorship values



Coefficient of variation in raw
censorship metric

Censorship Analysis

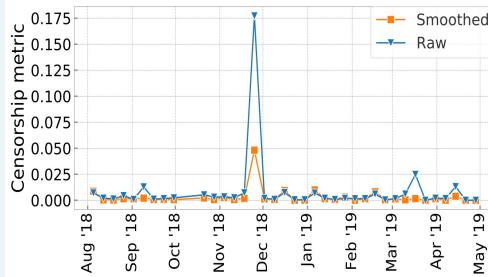
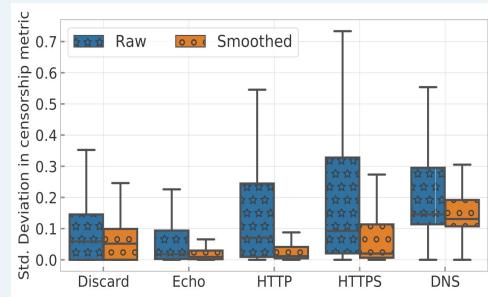
- Objective: Obtain a representative metric of censorship within a country that is not affected by outlier vantage points
- Apply an **optimization model** (Nelder-Mead) to obtain a weight for each Autonomous System that smooths the metric.



Censorship Analysis

- Objective: Obtain a representative metric of censorship within a country that is not affected by outlier vantage points
- Apply an **optimization model** (Nelder-Mead) to obtain a weight for each Autonomous System that smooths the metric.

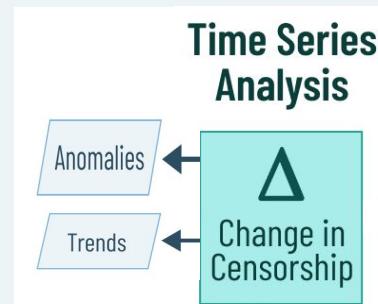
Standard deviation in censorship metrics



Censorship over time (Pakistan)

Time Series Analysis

- **Anomaly Detection** - Bitmap-based detection
- **Trend Analysis** - Mann-Kendall test

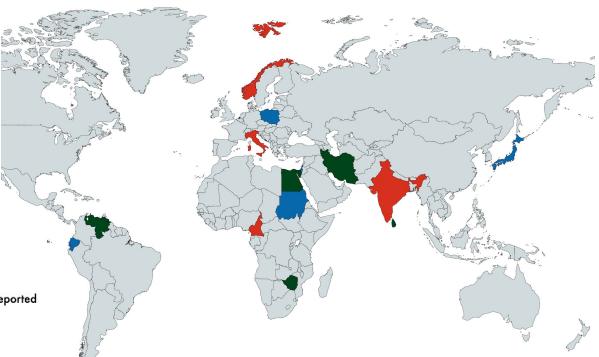


Findings

- Censorship Events
- Censorship Trends
- Case Studies

Censorship Events

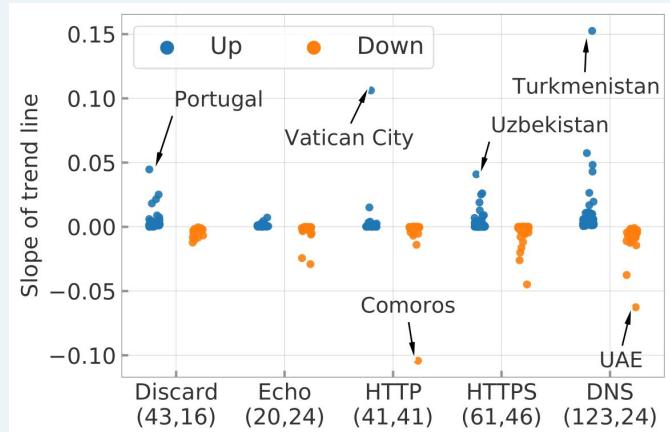
- Identified 15 key censorship events
 - 5 previously reported
 - 10 unreported



Country	Period	Method	Category or Domain blocked	Event
Egypt	26 Sep 2019	HTTP, HTTPS	News Media	Protests
Iran	Mar 2020	HTTP, Echo	wikimedia.com, wikia.com	Policy
Sri Lanka	21 Apr–12 May 2019	HTTP, HTTPS	Social Networking	Terrorism
Venezuela	12–29 Jan 2019	HTTP, HTTPS	Social Networking, wikipedia.org	Unrest
Zimbabwe	20 Jan 2019	HTTP, HTTPS	Social Networking	Protests
Ecuador	8 Oct 2019	DNS	Social Networking	Protests
India	6 Sep 2018	DNS	Online Dating	Law
Israel	May 2019–Jun 2019	DNS	Foreign Relations and Military	Conflict
Japan	28 Jun 2019	DNS, Echo	News Media	Summit
Poland	22 Jul 2019	DNS, HTTP, HTTPS	Govt., News Media, Human Rights	Unrest
Sudan	11 Apr 2019	HTTP, HTTPS	Social Networking	Unrest
Cameroon	25 Nov 2018	HTTP	Gambling	Unknown
India	Feb–Mar 2020	Echo, HTTPS	Illegal	Unknown
Italy	22 Dec 2019	Discard	Human Rights	Unknown
Norway	Dec 2019–Mar 2020	DNS	Multiple	Unknown

Censorship Trends

- Increasing levels of DNS censorship in more than 100 countries.
- Anonymization and circumvention tools most blocked using DNS.
- 11 categories of domains increasingly blocked (News Media, Provocative Attire)



Upward and downward trends
in censorship methods

Case Study

To be presented at FC 2021

Lost in Transmission: Investigating Filtering of COVID-19 Websites

Anjali Vyas, Ram Sundara Raman, Nick Ceccio,
Philipp M. Lutscher*, Roya Ensafi

Select 81 COVID-19 search terms



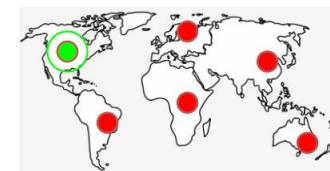
Search engine crawl in 9 countries



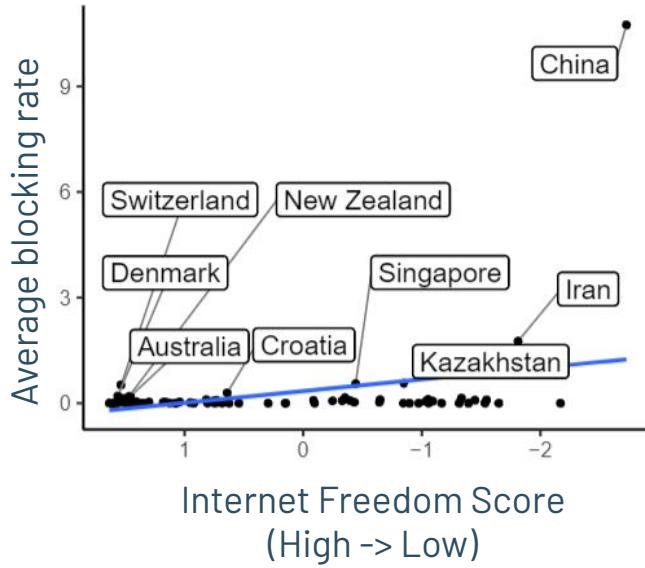
Extract top 1,291 domains for testing

bbc.com
covid-19.uk.com
covid-19-stats.info

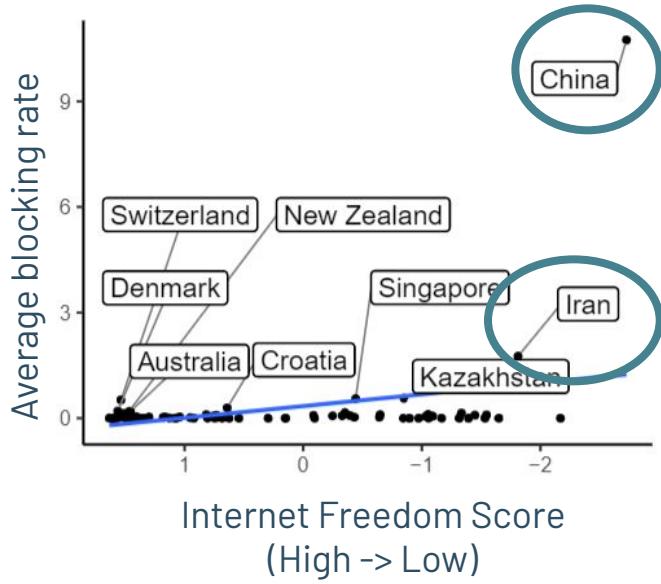
Perform Satellite measurements to 29,113 vantage points



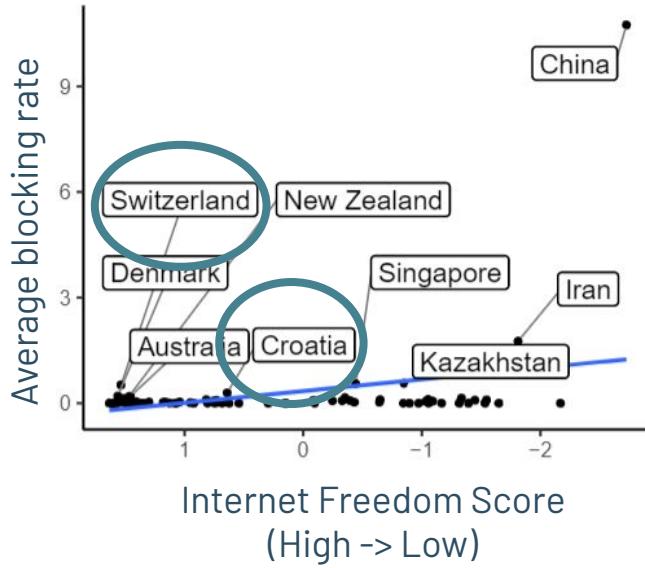
Test lists and measurement



DNS filtering of COVID-19 related websites



DNS filtering of COVID-19 related websites



DNS filtering of COVID-19 related websites

Zscaler Message:

Restricted Website Ahead!

We have noticed that you are trying to access [REDACTED] This website has been restricted by your organization's Internet filtering and security policy which is why we have prevented your computer from connecting to it.

For further assistance, please contact your IT support organization.

[Report an Inaccuracy](#)

Swisscom Internet Guard Message:

Dangerous website blocked

Swisscom Internet Guard warns you before you access this website. It might try to transfer dangerous software to your device (viruses, malware).

[Back](#)

[Access blocked website despite risk.](#)

DNS filtering of COVID-19 related websites in Switzerland

Domains	% of VPs	Category (Fortiguard Web Filter)
www.covid-19.uk.com	66.44	Phishing
coronavirus-realtime.com	66.40	Malicious
covid19graph.work	66.36	Phishing
www.covid19ireland.com	66.22	Phishing
www.covid19maps.info	65.83	Phishing

DNS filtering of COVID-19 related websites in Switzerland

Are these websites dangerous?

- Classified 46 websites most blocked manually and with 3 independent website categorization tools.
- Only 0-36.67% websites marked as security risks.

Large-scale, continuous censorship measurement is essential!

Our study finds:

- **Censorship varies over time** → Continuous censorship measurement is crucial
- **DNS Censorship in ‘free’ countries** → Censorship measurements should cover all countries
- **Censorship increase in encrypted communication**

Website

<https://censoredplanet.org>

 Censored Planet
University of Michigan Lab

Team Projects Reports Publications Observatory Log In

Percentage of resolvers facing interference by country



0 20%

Top disrupted domains by country

Brazil

Country	Domain	Disrupted percentage
Brazil	www.date.com	76.65
Brazil	www.agentprovocateur.com	76.42
Brazil	www.hrw.org	75.94
Brazil	www.163.com	71.46
Brazil	creditkarma.com	65.68

Date and Time of Scan	File Name	Scan Tool	Scan Type	Size of File in MB
2020-06-24T06:01:03	CP_Quack-echo-2020-06-24-06-01-03.tar.gz	Quack - echo	Application Layer	621.177
2020-06-23T00:08:31	CP_Quack-https-2020-06-23-00-08-31.tar.gz	Quack - https	Application Layer	3940.94
2020-06-22T14:45:38	CP_Quack-https-2020-06-22-14-45-38.tar.gz	Quack - https	Application Layer	3340.128
2020-06-22T01:02:10	CP_Quack-http-2020-06-22-01-02-10.tar.gz	Quack - http	Application Layer	1580.374
2020-06-21T12:00:01	CP_Satellite-2020-06-21-12-00-01.tar.gz	Satellite	DNS Layer	7137.384

 Censored Planet

49



Censored Planet

Thank you!

<https://censoredplanet.org>

A nighttime satellite view of Earth from space, showing city lights and auroras.

Backup Slides

Ethics

- Censorship research frequently raises ethical considerations e.g., under what conditions is it safe enough to use remote vantage points?
- IRBs are often not positioned to help. We turn to authorities such as the **Belmont and Menlo Reports** to guide ethical thinking.
- Frequently consult with colleagues to check our reasoning and conclusions.
- Ensure suitable protections in place, including technical practices to minimize risk to individuals
- Use hosts that are a part of Internet infrastructure
- Follow the ethical scanning guidelines, coordinate with our network administrators and our upstream ISP, host web pages indicating purpose of probes, set rate limits and more

OONI's approach

- “Provide as much informed choice to the user as possible”
 - Choose websites to test
 - Choose whether to upload measurement
 - Choose what type of data to submit
- Establish relationships with local & civil society
- Keep the community of volunteer involved

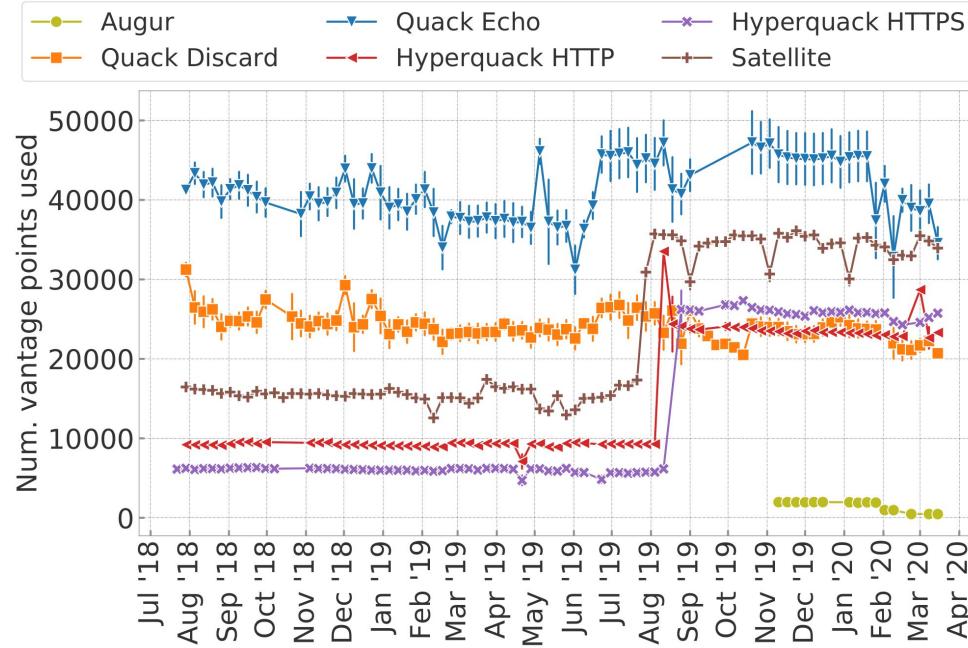
Coverage



Platforms	# AS	# Country	Median ASes / country
ICLab	56	48	1
OONI	1,915	155	4
Censored Planet Observatory	9,014	221	8



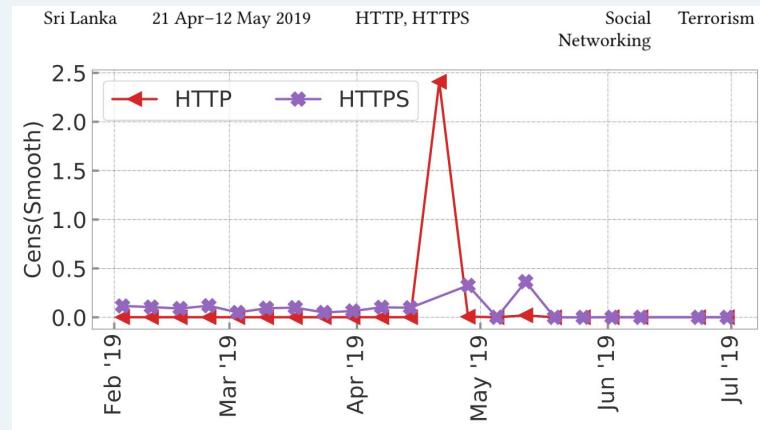
IPv4 hosts - Internet infrastructure is everywhere



Vantage Points over time

Censorship Events - Sri Lanka

- Bomb blasts on 21 April, 2019, triggered Internet shutdowns and Social media censorship.
- 22 domains blocked - Twitter, Whatsapp, Facebook, Instagram



Raw Data



Censored Planet
University of Michigan Lab

Team

Projects

Reports

Publications

Observatory ▾

Log In

DATA

Raw Data

Raw data files from Censored Planet tools

Please note: We are changing the formats of the output of our tools in 2019. If you need any information or help interpreting our data or the format, please reach out to us.

Note: We recently upgraded our measurement infrastructure, and are in the process of fixing some issues with the data collection. The raw data posted here needs to be processed to avoid false inferences. Thank you for your patience while we fix these issues and enhance our tools.

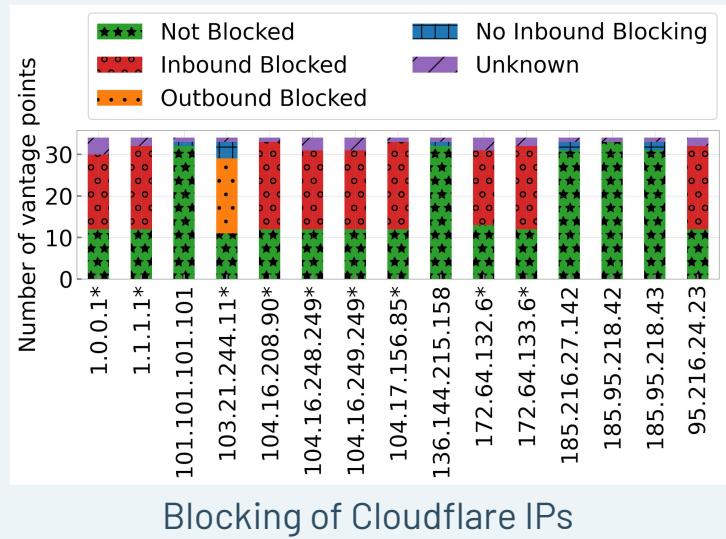
Date and Time of Scan	File Name	Scan Tool	Scan Type	Size of File in MB
2020-06-25T06:01:03	CP_Quack-echo-2020-06-25-06-01-03.tar.gz	Quack - echo	Application Layer	625.728
2020-06-25T01:02:10	CP_Quack-http-2020-06-25-01-02-10.tar.gz	Quack - http	Application Layer	1583.897
2020-06-24T06:01:03	CP_Quack-echo-2020-06-24-06-01-03.tar.gz	Quack - echo	Application Layer	621.177
2020-06-24T01:02:10	CP_Quack-http-2020-06-24-01-02-10.tar.gz	Quack - http	Application Layer	2203.408
2020-06-23T00:08:31	CP_Quack-https-2020-06-23-00-08-31.tar.gz	Quack - https	Application Layer	3940.94

Data Formats - General

- Scan Date
- Vantage Point
- Test keyword/domain/URL
- Blocked
- Country

Rapid Focus Study - Turkmenistan

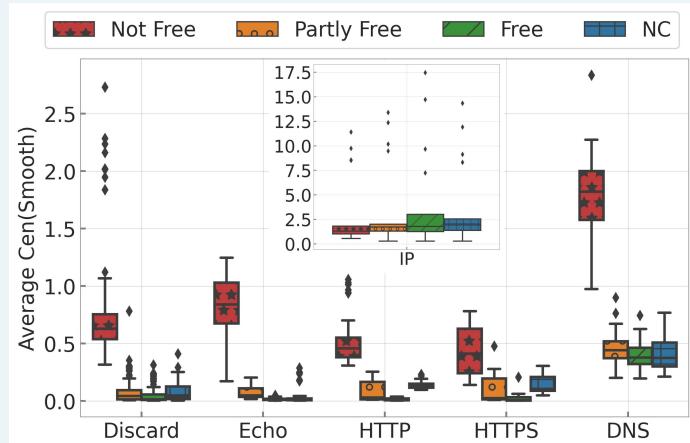
- Very strict censorship policies
- Conducted rapid focus Augur measurements to DoH server IPs and Cloudflare IPs in April 2020.
- 52.9% of vantage points in Turkmenistan block all Cloudflare IPs, restricting access to thousands of services.
- Other DoH servers (eg. Snopyta) also blocked.



Freedom on the Net Report

- Observe significant censorship in countries labelled as “Free” or Not Considered by qualitative reports like Freedom on the Net^[1].
- Data-driven insights from Censored Plane can significantly complement qualitative reports.

[1] <https://freedomhouse.org/countries/freedom-world/scores>, 2019.



Censorship in Freedom on the Net categories