Incentivizing Censorship Measurements via Circumvention

Ihsan Ayyub Qazi

Aqib Nisar*

Aqsa Kashaf**

Zartash A. Uzmi

LUMS

* Now at USC

** Now at CMU

Internet censorship is pervasive!

- Over 70 countries restrict Internet access
  - Often due to political, social, or economic reasons

Censorship has a substantial impact

- … on different stakeholders in the Internet ecosystem

Users

ISPs

Advertisers

Content Providers

Government

It has led to the design of censorship…

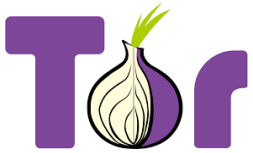
Measurement Systems

- *What is blocked?*
- *Where is it blocked?*
- *How is it blocked?*
- *When it is blocked?*

… CensMon, Iris, Augur, Encore

Circumvention Systems

How do we bypass censorship?



LANTERN

Current practice and limitations

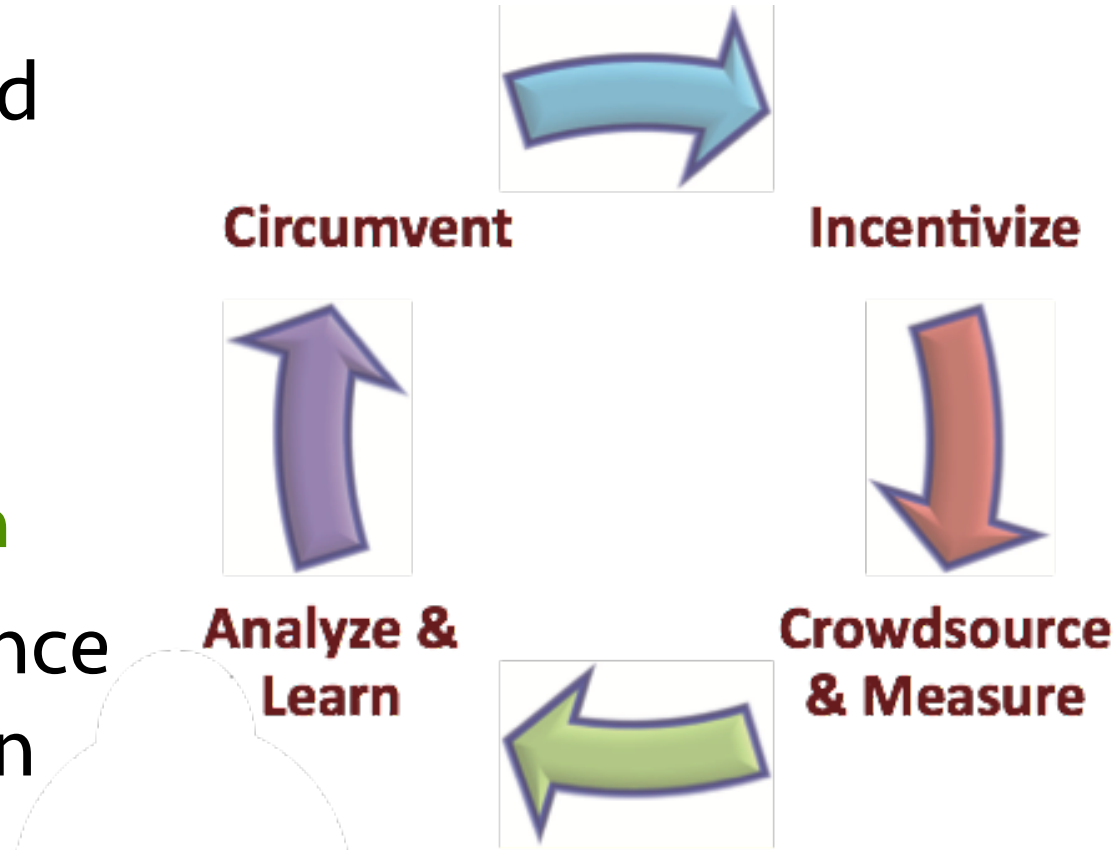
Existing measurement and circumvention systems are designed independently

- Circumvention systems are not data-driven
  - … leads to one-size-fits-all solutions!
- Censorship measurement systems lack incentives
  - … limits availability of geographically distributed probe points

In this work we ask, "*Can we address the limitations of individual systems by consolidating them in a single platform?*"

C-Saw in 1-slide

- Consolidates measurements and circumvention
  - Uses crowdsourcing to gather censorship measurements
  - Offers data-driven circumvention
- Better circumvention performance incentivizes more users to opt-in



Rest of the talk

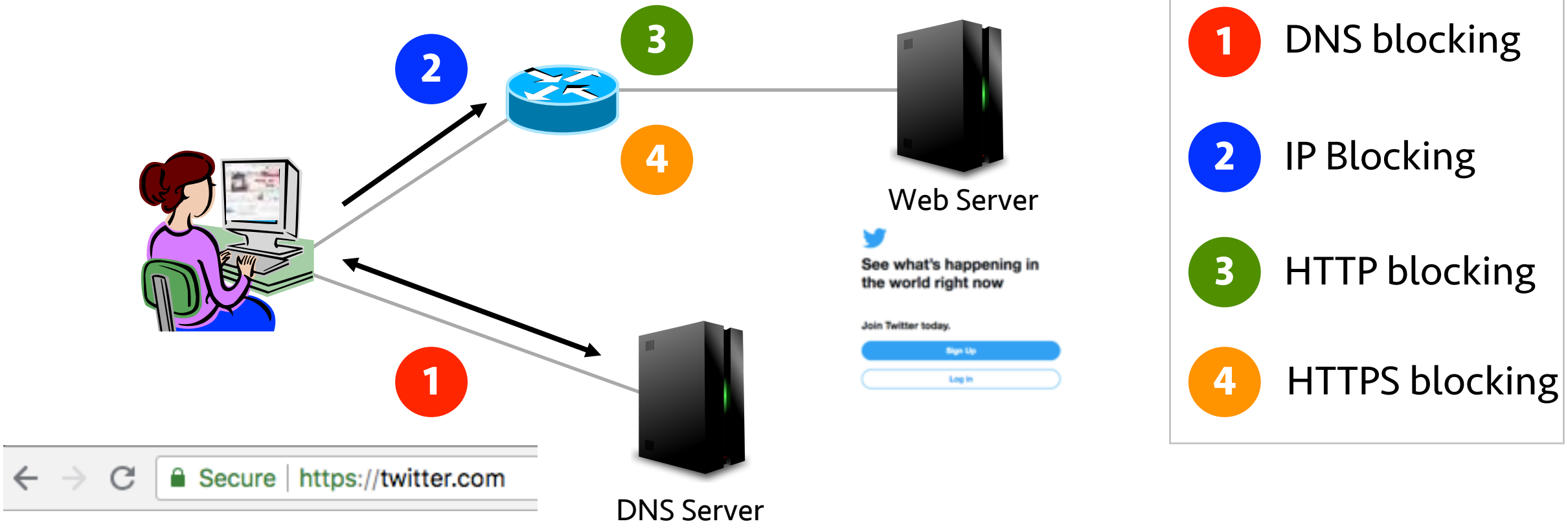
- Web Censorship & Circumvention
- C-Saw Design
- Evaluation
- Deployment

Rest of the talk

- Web Censorship & Circumvention
- C-Saw Design
- Evaluation
- Deployment

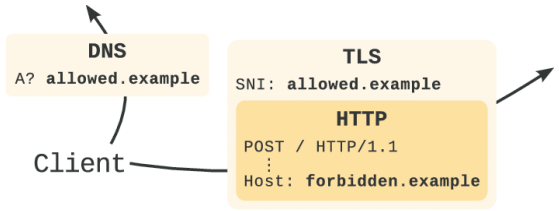
Web censorship techniques

- Web filtering can be performed by intercepting a user request at different levels of the protocol stack



Circumvention approaches

- Public DNS Servers
- Domain Fronting
- VPNs
- Tor
- Lantern
- ... others



ExpressVPN

Tor

LANTERN

Circumvention: local fix vs relay-based

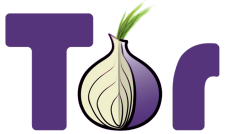
Local fix

2

Domain Fronting

1

Relay(s)

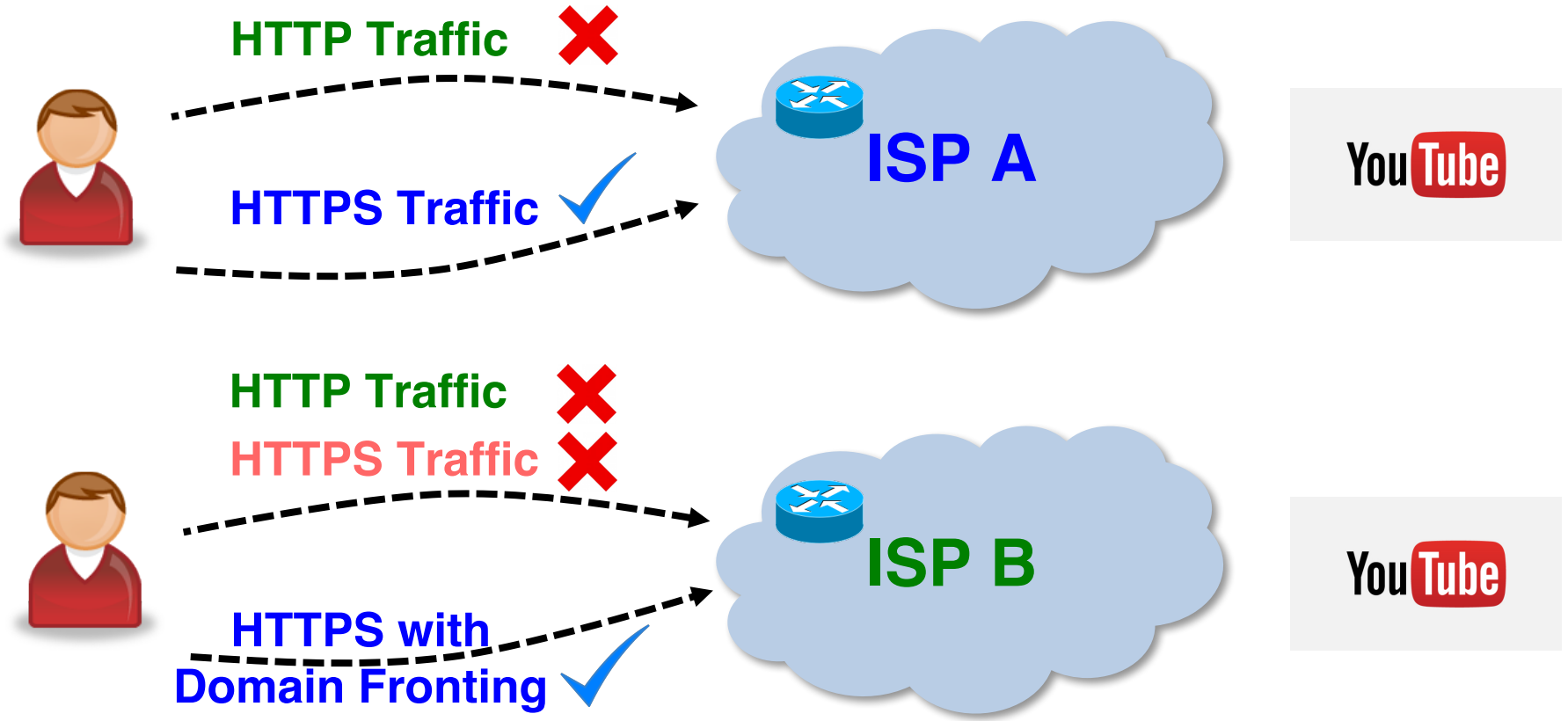


What are the opportunities for improving circumvention performance?

A censorship case study in Pakistan

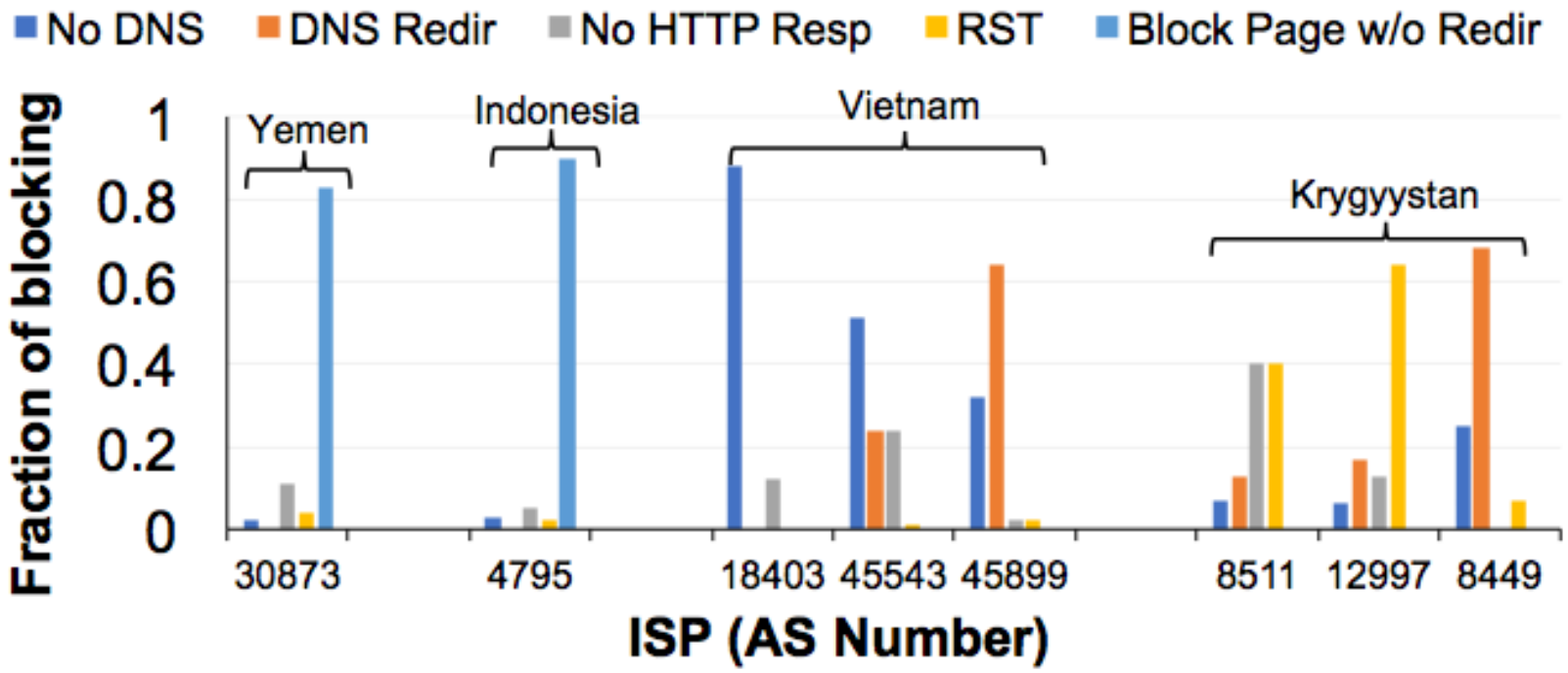
- Measurements taken from different vantage points
  - University campus (Lahore)
    - Served by ISP-A and ISP-B
  - Home users (Karachi)
    - Served by ISP-B only

A censorship case study in Pakistan



(1) Insights about censors

- Blocking mechanisms can differ *across* ISPs
- Blocking mechanisms can differ *across* URLs even within an ISP



Insights hold across several countries

(2) Circumvention insights - 1/2

Fetched:

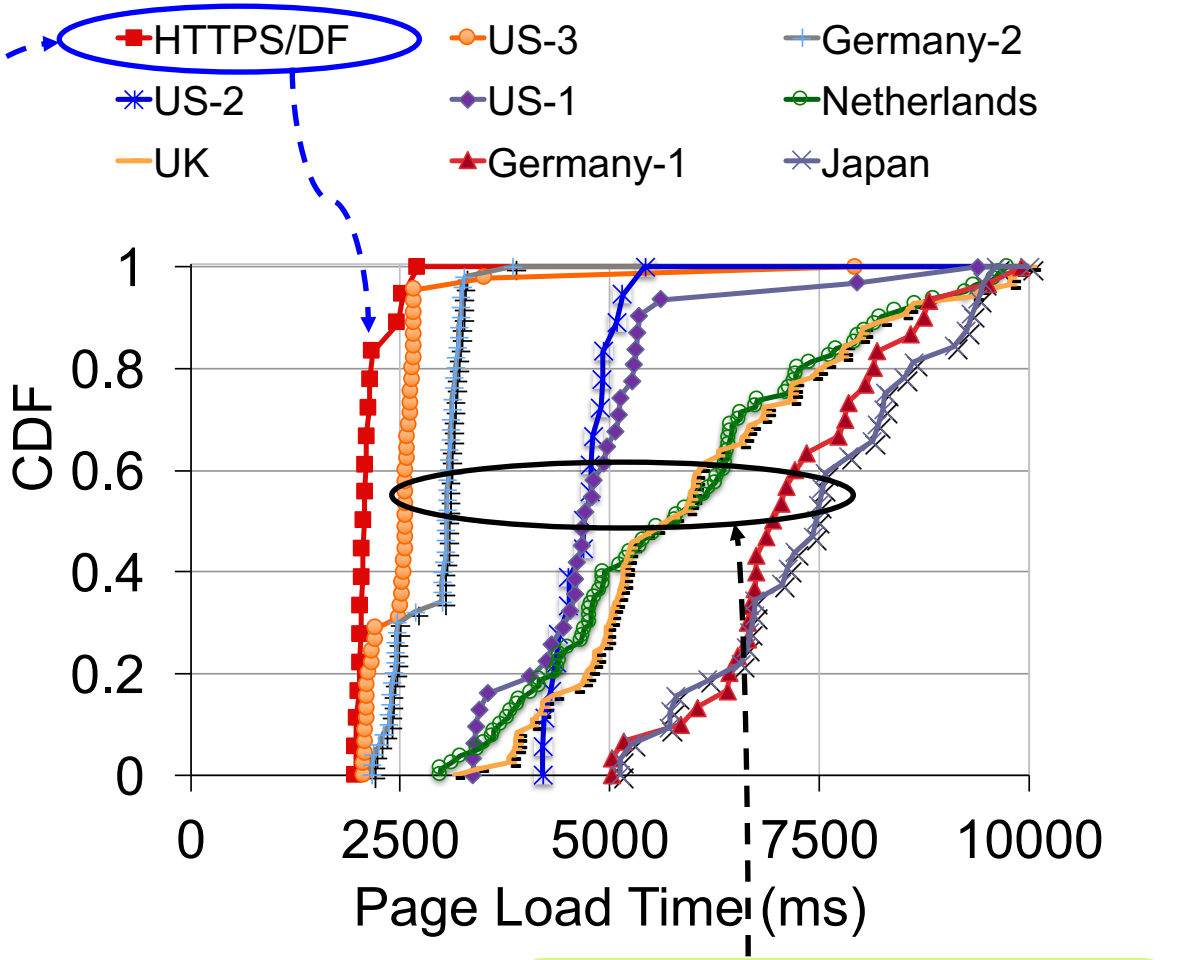
YouTube homepage
200 runs

ISP-B:

Blocking: HTTP & HTTPS
HTTPS/DF

Measurement point:

Campus network



All static proxies exhibited longer PLTs than the local fix

(2) Circumvention insights - 2/2

Fetched:

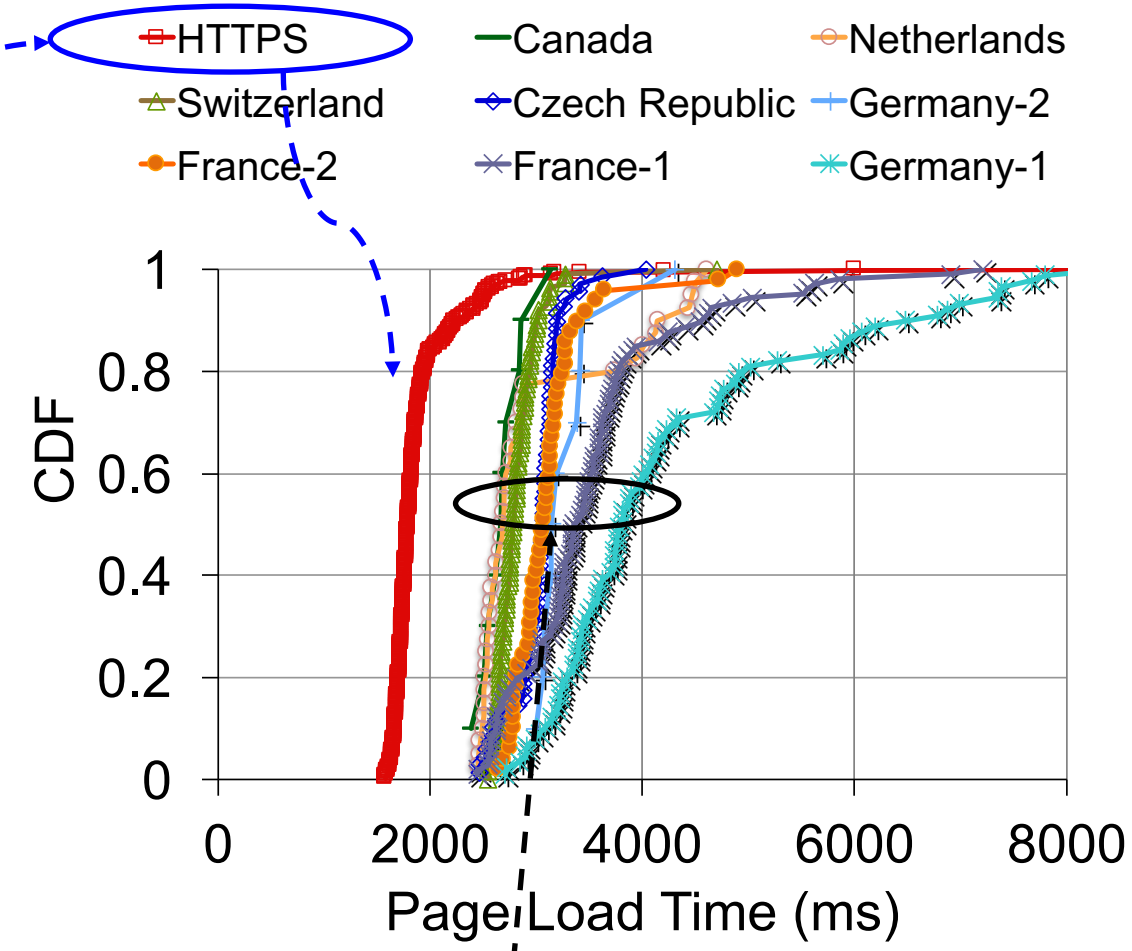
YouTube homepage
200 runs

ISP-A:

HTTP Blocking Only
HTTPS

Measurement point:

Campus network
Tor exit relay shown



All Tor results indicate longer PLTs

(2) Circumvention insights - 2/2

Fetched:

YouTube homepage

200 runs

Measurement point:

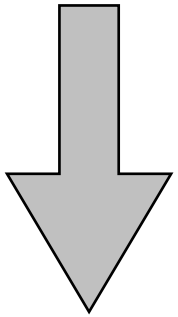
Campus network

Tor exit relay shown

Different circumvention strategies impose widely different overheads

Key implication for design

Measurements reveal differences in blocking mechanisms



Can pick the least overhead circumvention strategy

Rest of the talk

- Web Censorship & Circumvention
- C-Saw Design
- Evaluation
- Deployment

Rest of the talk

- Web Censorship & Circumvention
- C-Saw Design
- Evaluation
- Deployment

Design goals

1. Scalable measurements with user consent
2. Adaptive circumvention

- In addition, a practical and usable solution should
  - *require no target lists*
  - *preserve privacy of users contributing measurements*

How C-Saw meets these goals?

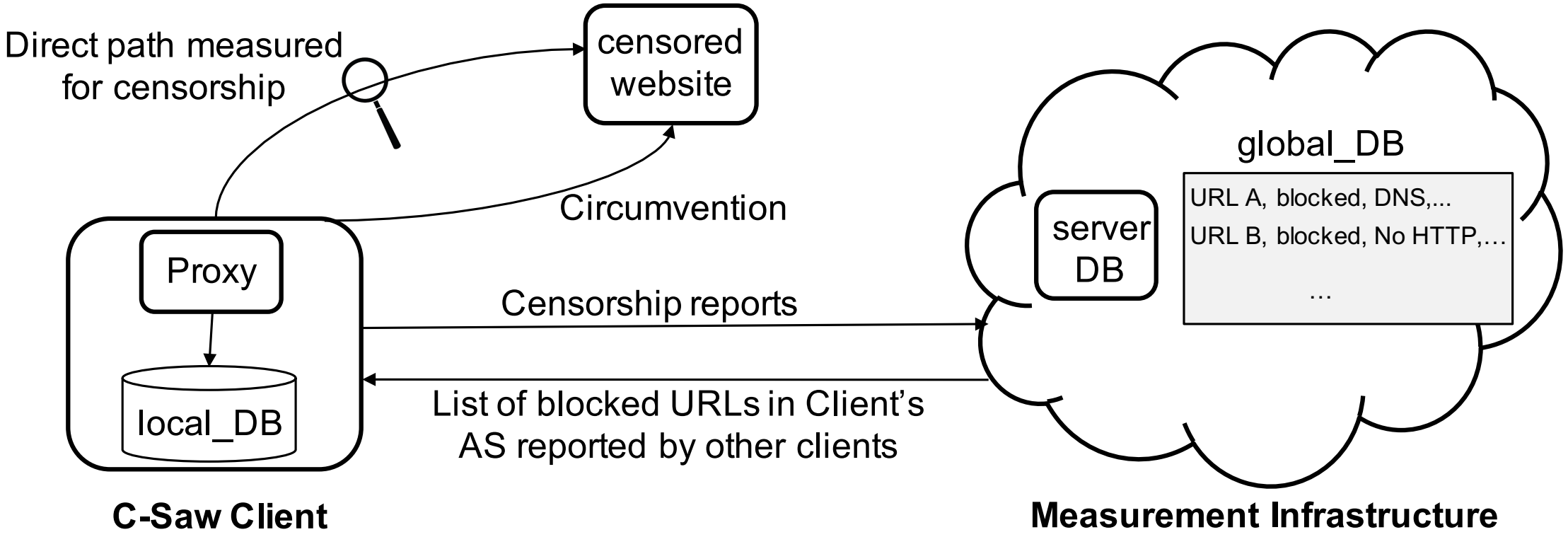
1 Scalable measurements with user consent

- C-Saw offers small PLTs as an incentive
  - It only measures those URLs that a user actually visits
- As a result, it requires no target lists!

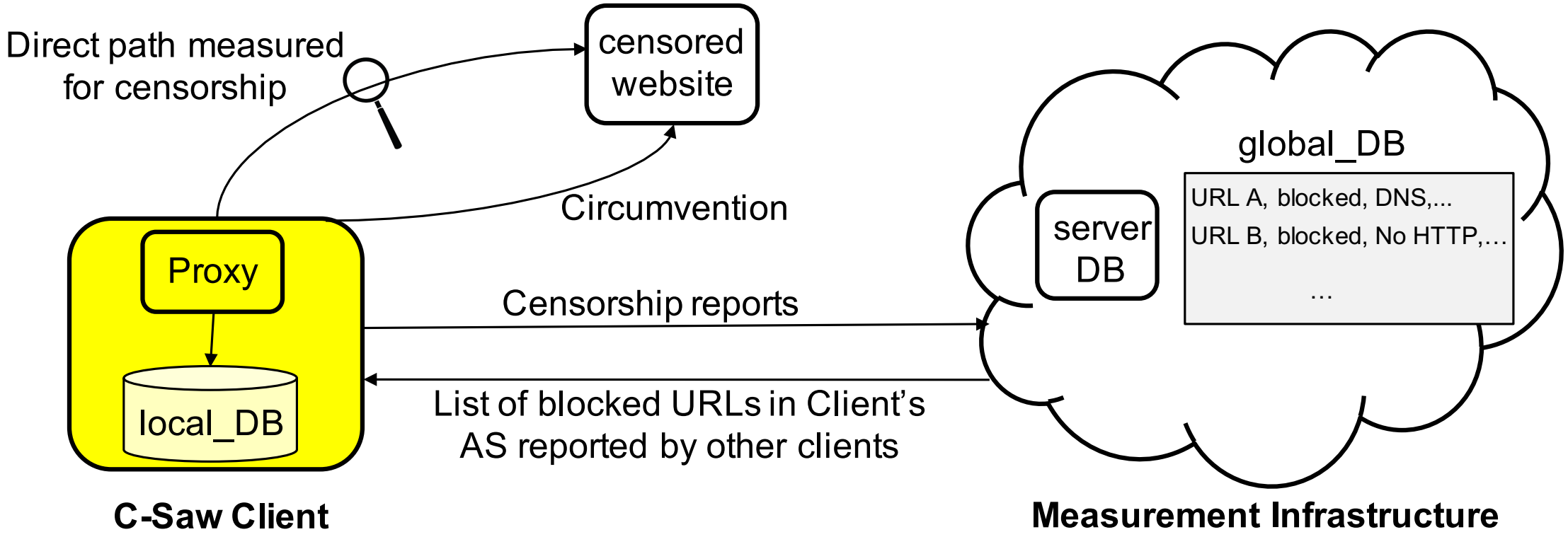
2 Adaptive circumvention

- C-Saw measures the blocking mechanism used by a censor
- Selects the least overhead circumvention strategy

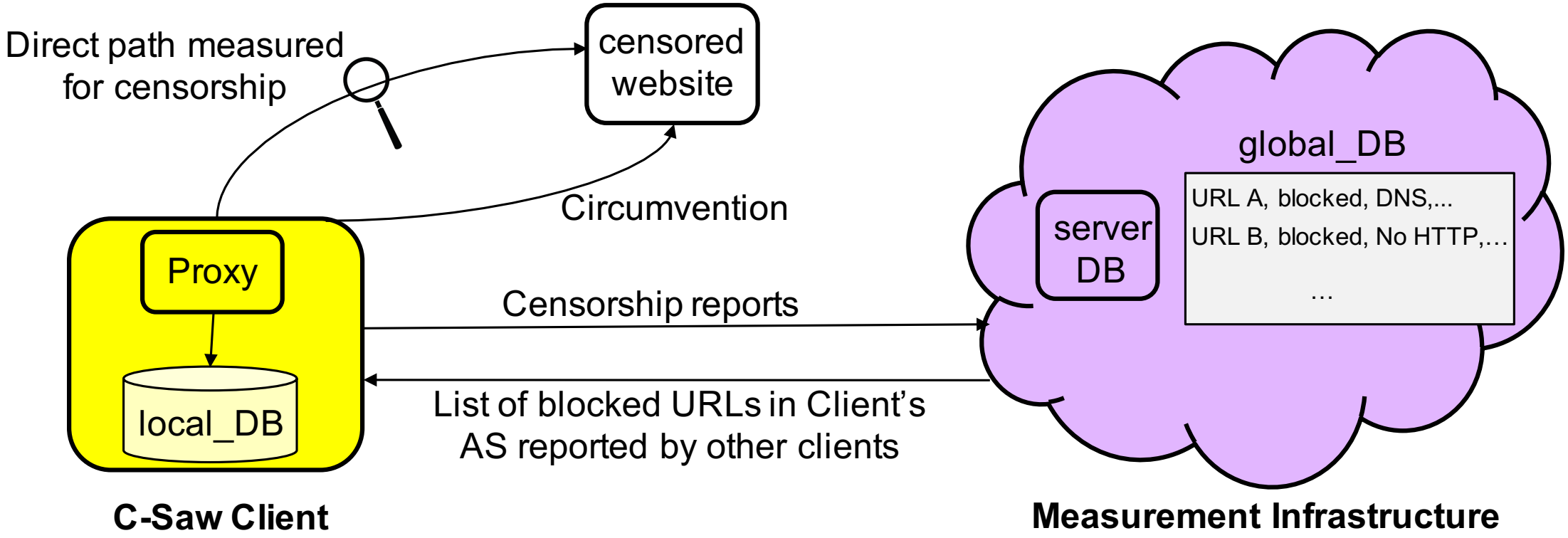
C-Saw components



C-Saw components

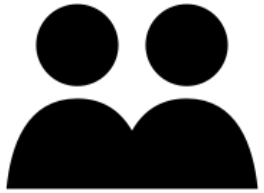


C-Saw components



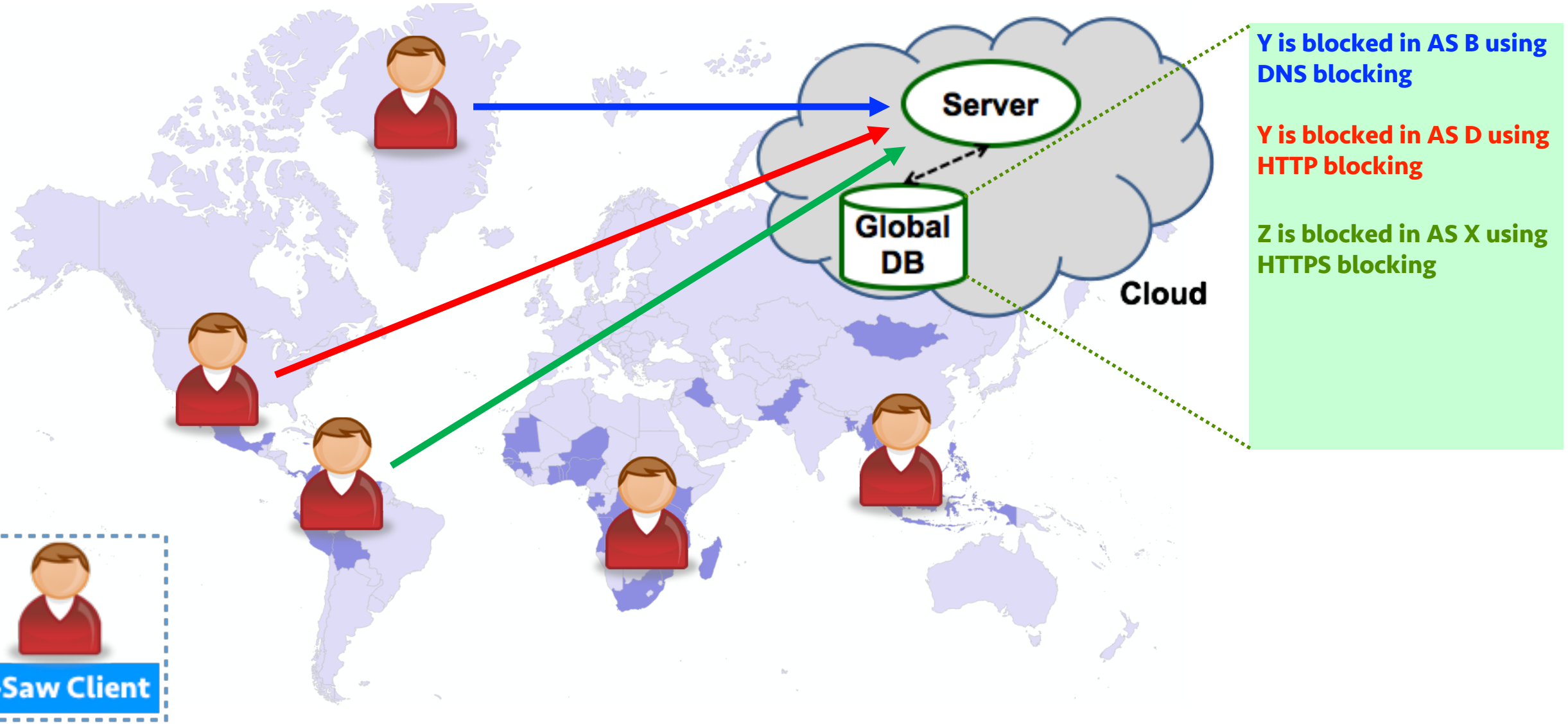
C-Saw proxy

- Measurement module
  - Runs a censorship detection algorithm
  - Issues redundant requests
  - Achieves resilience to false reports



- Circumvention module
  - Selects a circumvention approach (e.g., Public DNS, Domain Fronting, or Tor)

C-Saw Big Picture



Security and privacy considerations

- Interference with C-Saw measurements
  - Rate limits creation of fake IDs and uses a voting mechanism
- Blocking access to the measurement infrastructure
  - One can use Tor hidden services
- User privacy and resilience to detection
  - All measurement reports are carried over the Tor network

Rest of the talk

- Web Censorship & Circumvention
- C-Saw Design
- Evaluation
- Deployment

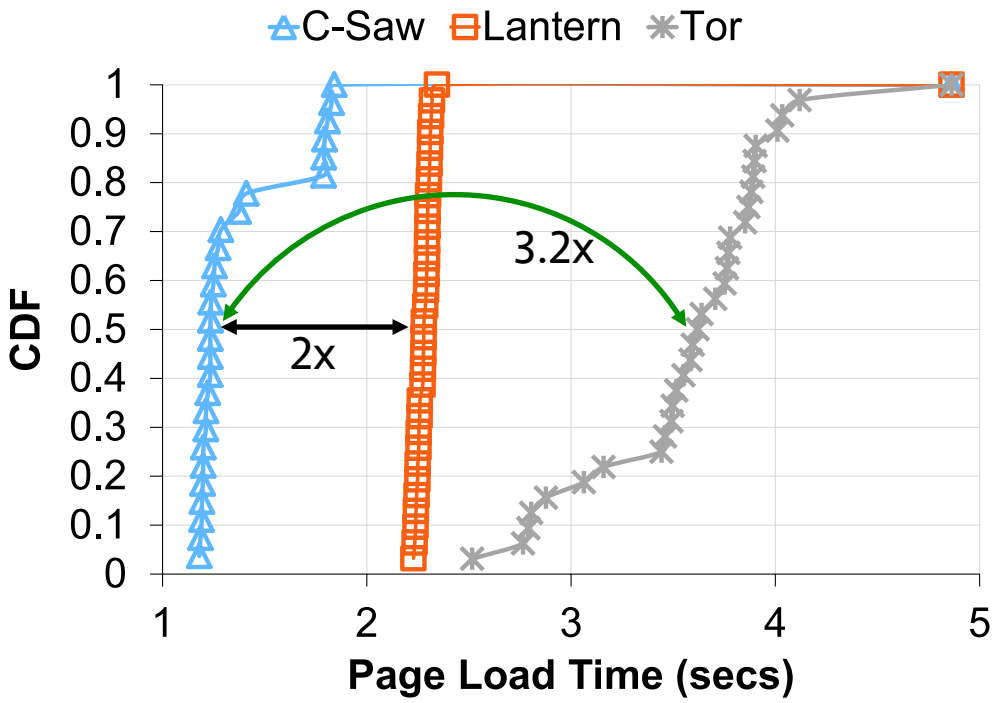
Rest of the talk

- Web Censorship & Circumvention
- C-Saw Design
- Evaluation
- Deployment

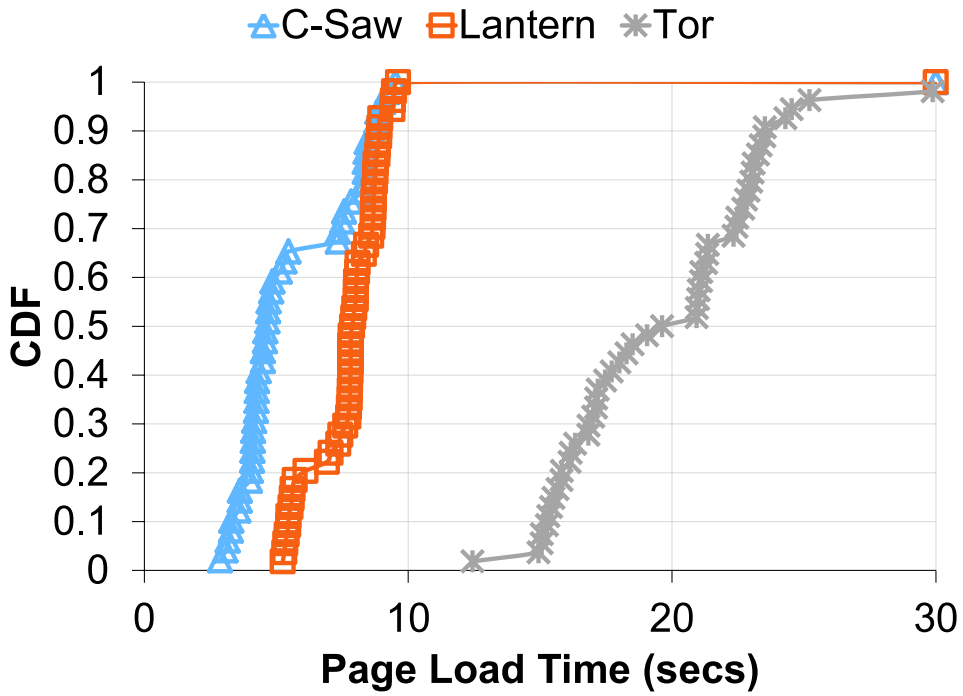
Evaluation

- We implemented C-Saw using GitHub's electron framework
  - Measures common forms of censorship
  - Implements several local fixes and optimizations
  - Supports Tor and Lantern as relay-based circumvention approaches
- Evaluation
  - Macro-benchmarks: C-Saw with Tor and Lantern
  - Micro-benchmarks: Impact of redundant requests, URL aggregation

Page Load Times with C-Saw



DNS Blocked Webpage



Unblocked Webpage

Rest of the talk

- Web Censorship & Circumvention
- C-Saw Design
- Evaluation
- Deployment

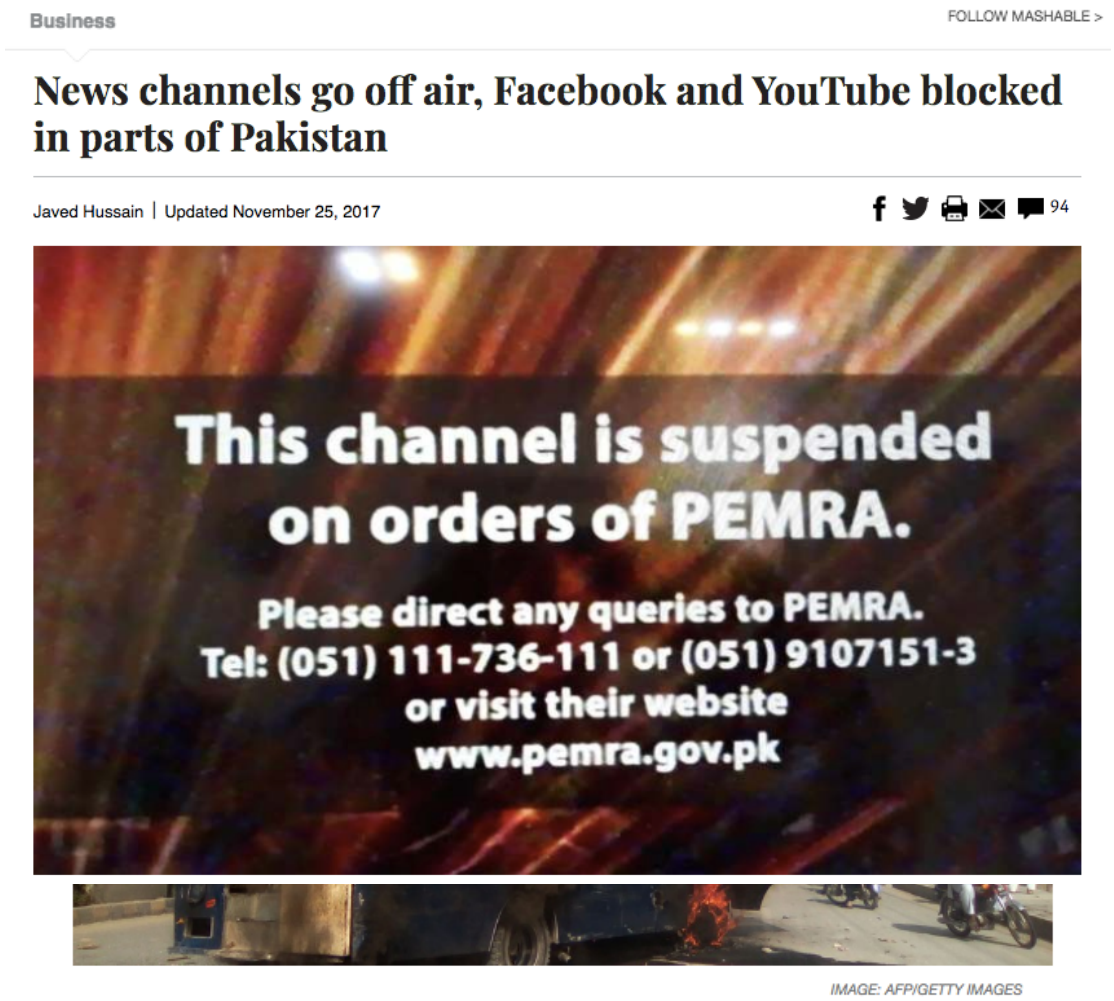
Rest of the talk

- Web Censorship & Circumvention
- C-Saw Design
- Evaluation
- Deployment

Deployment study

- We released C-Saw to 123 consenting users (3-month measurements)
  - Residential, Enterprise, and University network users in Pakistan
  - Users were carefully informed about C-Saw
    - … but were not given any list of blocked websites they needed to visit
- Insights
  - Users visited 420 blocked domains accessed through 16 different ASes
  - For majority of URLs, a block page was returned followed by DNS blocking
  - We found blocking of CDN servers

C-Saw in the wild



- Twitter was found blocked at 13:32 on Nov 25, 2017 from AS 17557 (Response: HTTP_GET_BLOCKPAGE)
- Instagram was found blocked at 4:51 on Nov 26, 2017 from AS 38193 (Response: DNS blocking)
- Instagram was found blocked at 9:06 on Nov 26, 2017 from AS 59257 (Response: DNS blocking)
- Instagram was found blocked at 9:31 on Nov 26, 2017 from AS 45773 (Response: DNS blocking)

Limitations and discussion

- Scope of measurements
  - Difficult to measure unpopular websites or censorship at specific times
- Robustness of C-Saw
  - Relies on Tor as one possible circumvention strategy
  - Arms race between Tor and some censors (e.g., China)
  - New circumvention approaches can be easily incorporated in C-Saw
- Non-Web filtering

Summary

- Censorship Measurements
  - C-Saw uses crowdsourcing to collect measurements
- Circumvention Performance
  - Censorship measurements enable adaptive circumvention
  - Small PLTs incentivize users to opt-in