Network Metrology: Censorship

This class

- Definitions
- Motivation
  - Examples of censorship
- Internet censorship
  - Block IP address
  - Block hostnames
  - Measuring censorship at scale

DEFINITIONS

What is censorship?

- *Censorship, the changing, or the suppression, or prohibition of speech or writing that is deemed subversive of the common good –* Definition from Britannica Encyclopedia.

Network censor: Definition

- An entity that desires that some *identifiable* communication is *blocked* from being transmitted over the network
  - Without the authority to compel the content provider to remove the content
  - Without the authority to compel the client to install software of the censor's choosing
- Requires that the censor acts on network traffic

How to identify and block?

- Identification
  - The piece of information that allows the censor to identify content to be blocked is referred to as the censorship trigger
  - Example: IP address, hostname, URL, keywords etc.
- Blocking: The technical means used to restrict access to the content
  - Example: dropping packets, forging TCP RST packets or DNS responses

MOTIVATION

Internet Censorship is Widespread

- Practiced in **59+ countries around the world**
  - Many western countries
  - Several electoral democracies (e.g., S. Korea, Turkey) have significant censorship
    - YouTube blocked in Turkey for two years
    - Many North Korean sites blocked in South Korea
- Twelve countries have **centralized infrastructure for monitoring/blocking**

Source: Open Network Initiative

Examples of Recent Trends

- In 23 countries, a blogger or Internet user was **arrested** for content posted online
  - Chinese woman sent to labor camp for satirical Twitter message
  - Indonesian woman fined for an email complaining about a local hospital

- Twelve countries instituted **bans** on Twitter, YouTube or some other **online social media** service

Why do countries censor?

- **Political stability**

August 11, 2011, 12:21 PM

In British Riots, Social Media and Face Masks Are the Focus

Prime Minister David Cameron told Parliament on Thursday that if people are using social media to organize violence, as has been reported, than "we need to stop them." He asked the police to tell him if they need "new powers" to do so.

- **National security**

Internet 'Kill Switch' Legislation Back in Play

By David Kravets ✉ January 28, 2011 | 6:09 pm | Categories: Cyber Warfare, Cybersecurity

- **Social values**

NEWS - Written by Renai LeMay on Friday, June 24, 2011 14:34 - 28 Comments

Voluntary ISP filter attracts global attention

This week, Telstra and Optus reiterated that they were still planning to start filtering their customers' traffic for a list of internet addresses provided by the Australian Communications and Media Authority which it has deemed to contain child pornography. The initiative is a stop-gap measure agreed to by ISPs and the Federal Government in mid-2010 while a review is carried out into the Refused Classification category of content which the wider mandatory filter will block.

Many opportunities to censor

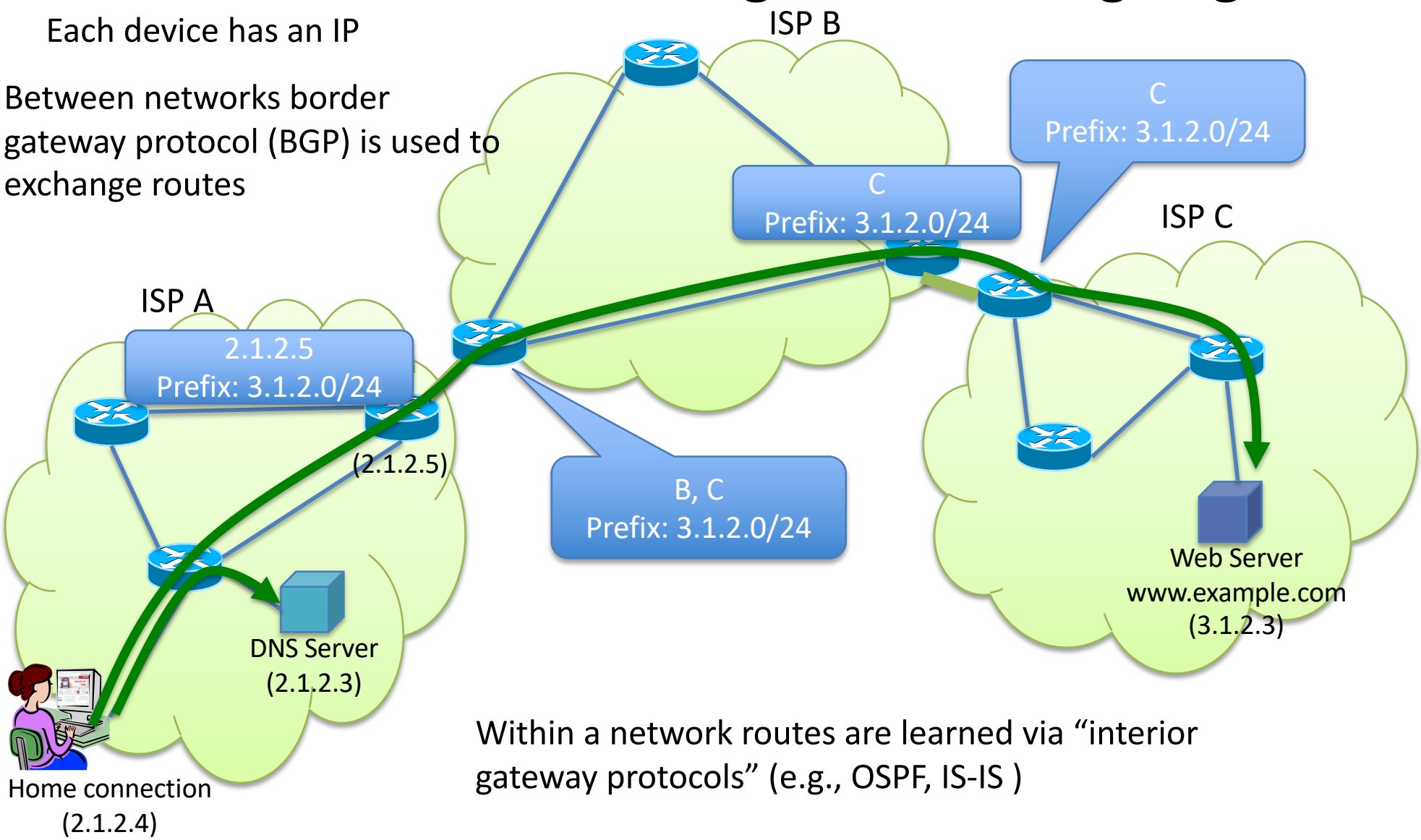
- Block IP addresses
  - IP layer
- Block hostnames
  - DNS (application layer)
- Disrupt TCP flows
  - TCP (transport layer)
  - Many possible triggers
- Disrupt HTTP transfers
  - HTTP (application layer)

Networking 101

- So how does our traffic get where its going?

Each device has an IP

Between networks border gateway protocol (BGP) is used to exchange routes



Within a network routes are learned via “interior gateway protocols” (e.g., OSPF, IS-IS)

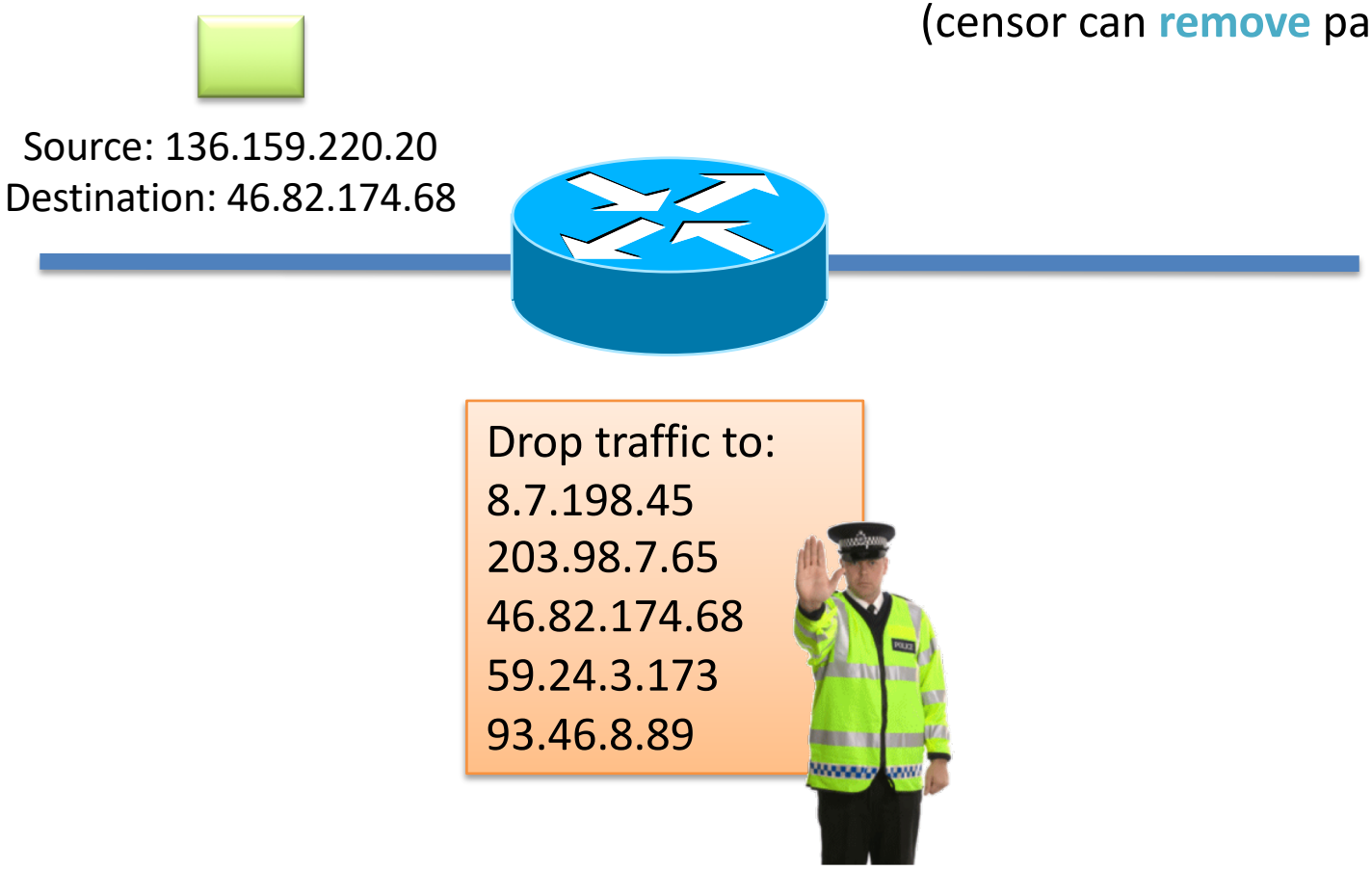
Many opportunities to censor

- Block IP addresses
  - IP layer
- Block hostnames
  - DNS (application layer)
- Disrupt TCP flows
  - TCP (transport layer)
  - Many possible triggers
- Disrupt HTTP transfers
  - HTTP (application layer)

IP-based blocking: ACL

- Option 1: Configure routers using an access control list (ACL) to drop traffic to a given IP address.

This is an example of in-path blocking (censor can **remove** packets)



IP-Based blocking with ACL: pros and cons

Advantages

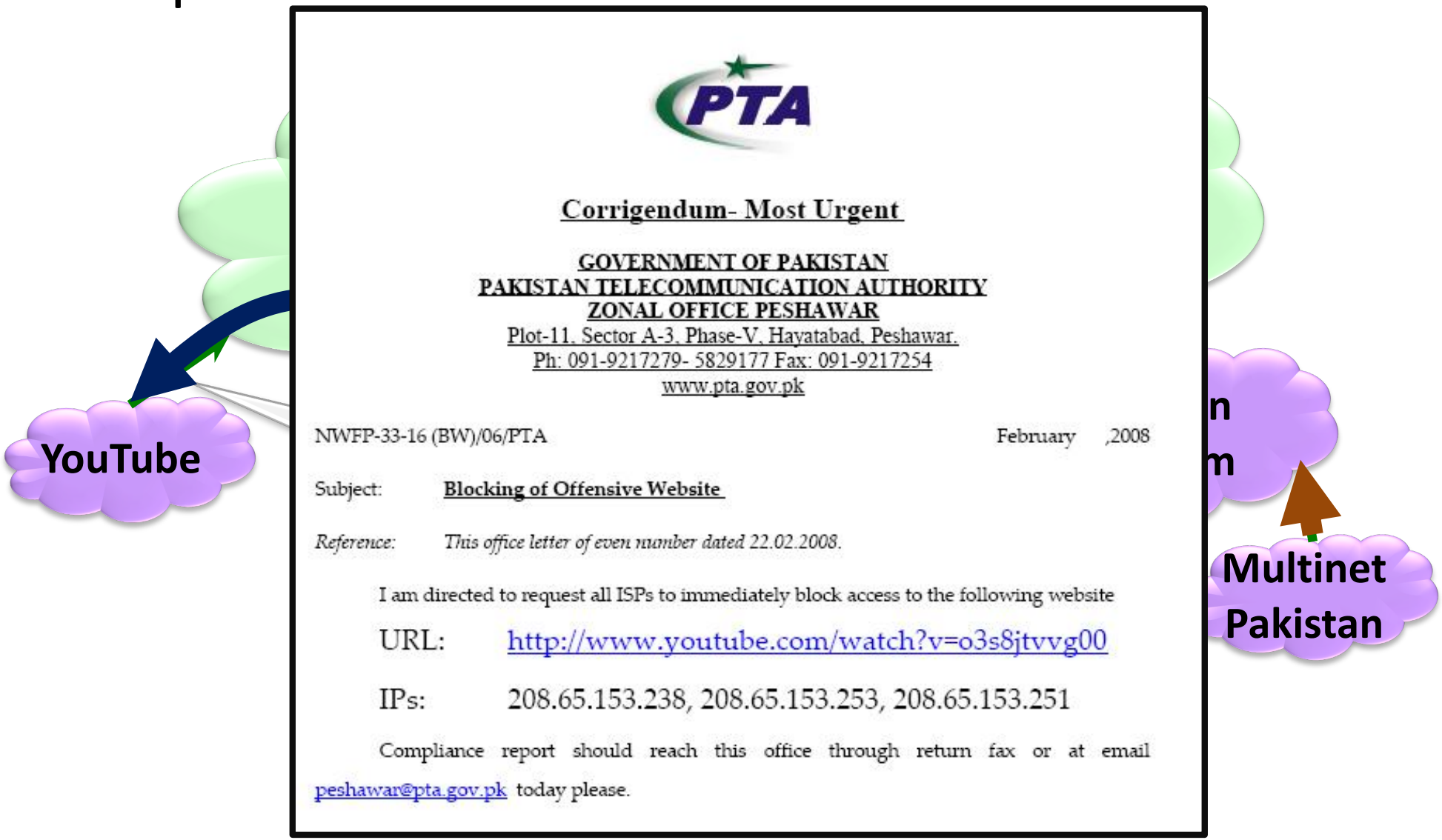
- Quick and easy to configure
- Routers have efficient techniques for IP matching

Disadvantages

- Need to know the IP
  - Easily evadable!
- High collateral damage: IP != Web host
  - Noticeable if high profile site is hosted on the same system
- Location of the censor can be determined from within the censored network and censored hosts

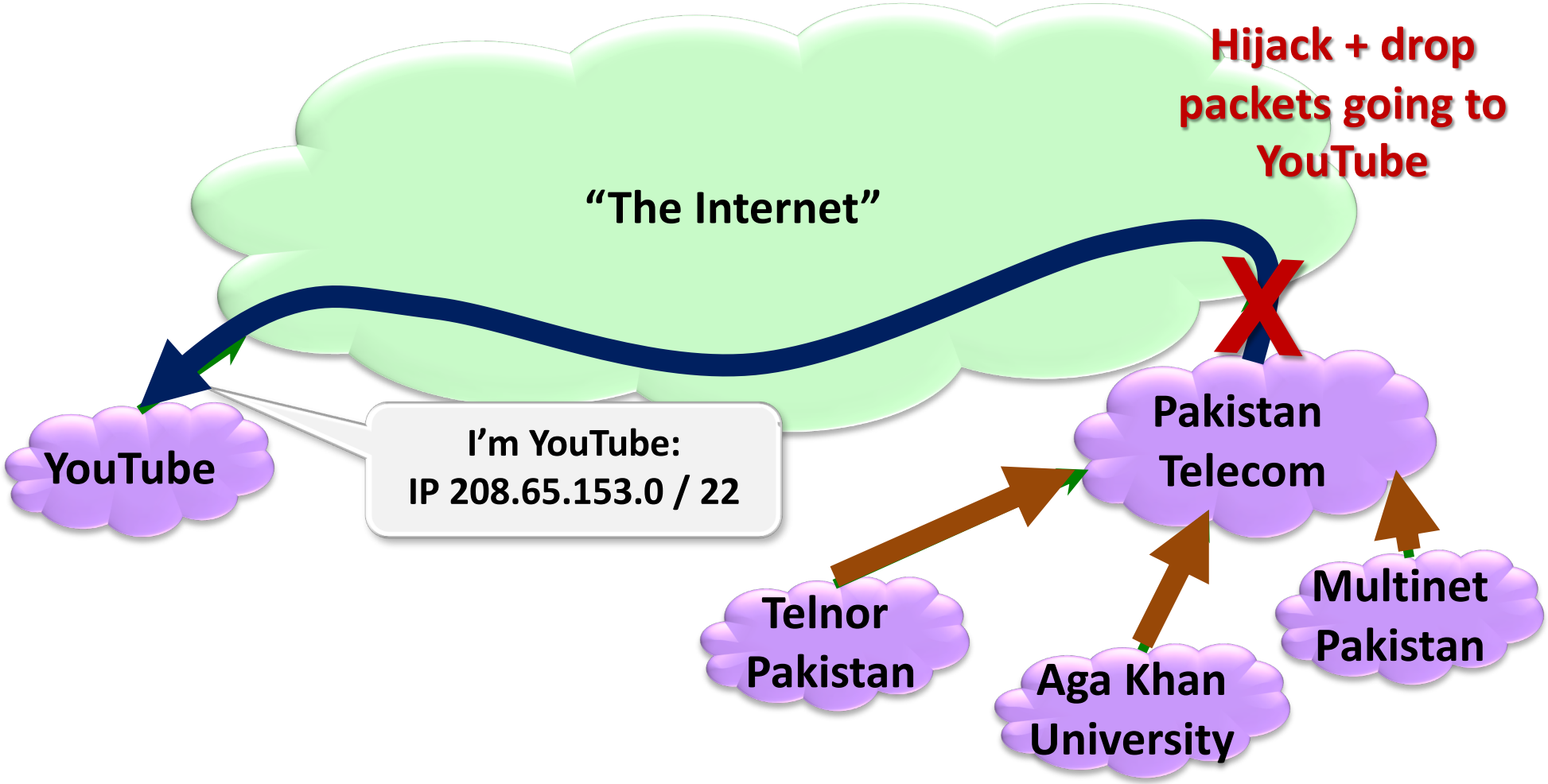
IP-BASED BLOCKING

- Option 2: Use BGP to block IPs



IP-BASED BLOCKING

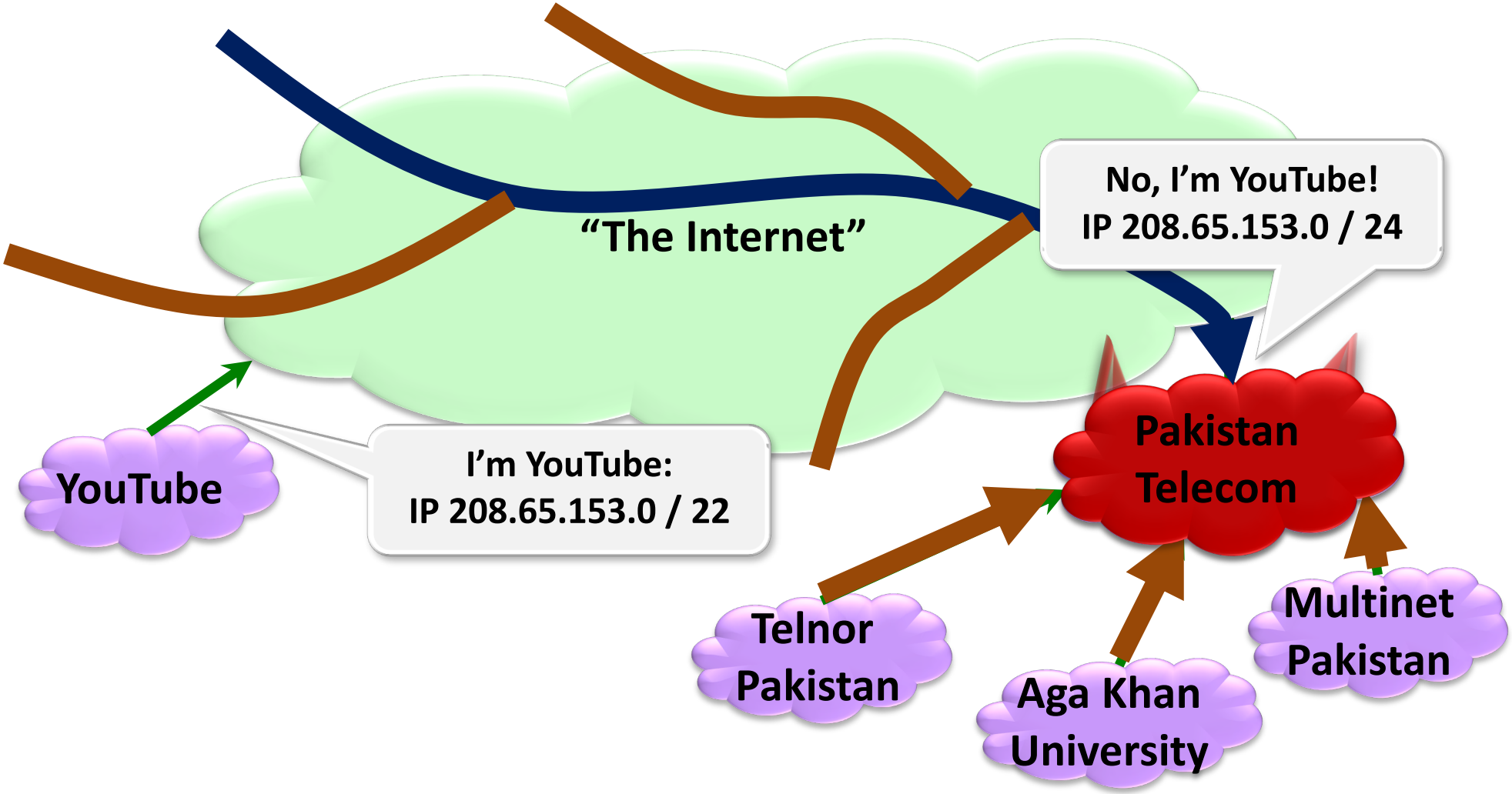
Here's what should have happened….



Block your own customers.

IP-BASED BLOCKING

But here's what Pakistan ended up doing…



IP-based blocking: BGP poisoning

- Option 2: BGP route poisoning
  - Instead of configuring router ACLs, just advertise a bogus route
    - Causes routers close to the censor to route traffic to the censor, which just drops the traffic
  - How to detect this type of censorship?
    - BGP looking glass servers in the impacted region
    - Sometimes global monitors as well …
  - Challenges
    - Can cause international collateral damage!
    - Will block all content on a given prefix
      - Could announce a /32 to get a single address but most ISPs will not propagate beyond a /24

Known users of IP-based blocking

- Pakistan using IP-based blocking for YouTube address ranges
  - Can interfere with other Google services
- China
  - Some reports of IP blocking
  - Many URLs redirected to small set of IP-addresses, possibly this is the set used for ACLs
- UK
  - Uses IP blocking of the Pirate Bay's IP address
- Australia
  - IP blocking for Melbourne Free University IPs (precise motivation unclear…)
  - https://www.eff.org/deeplinks/2013/04/australian-networks-censor-community-education-site
- In general, too much collateral damage of IP-based blocking.

Many opportunities to censor

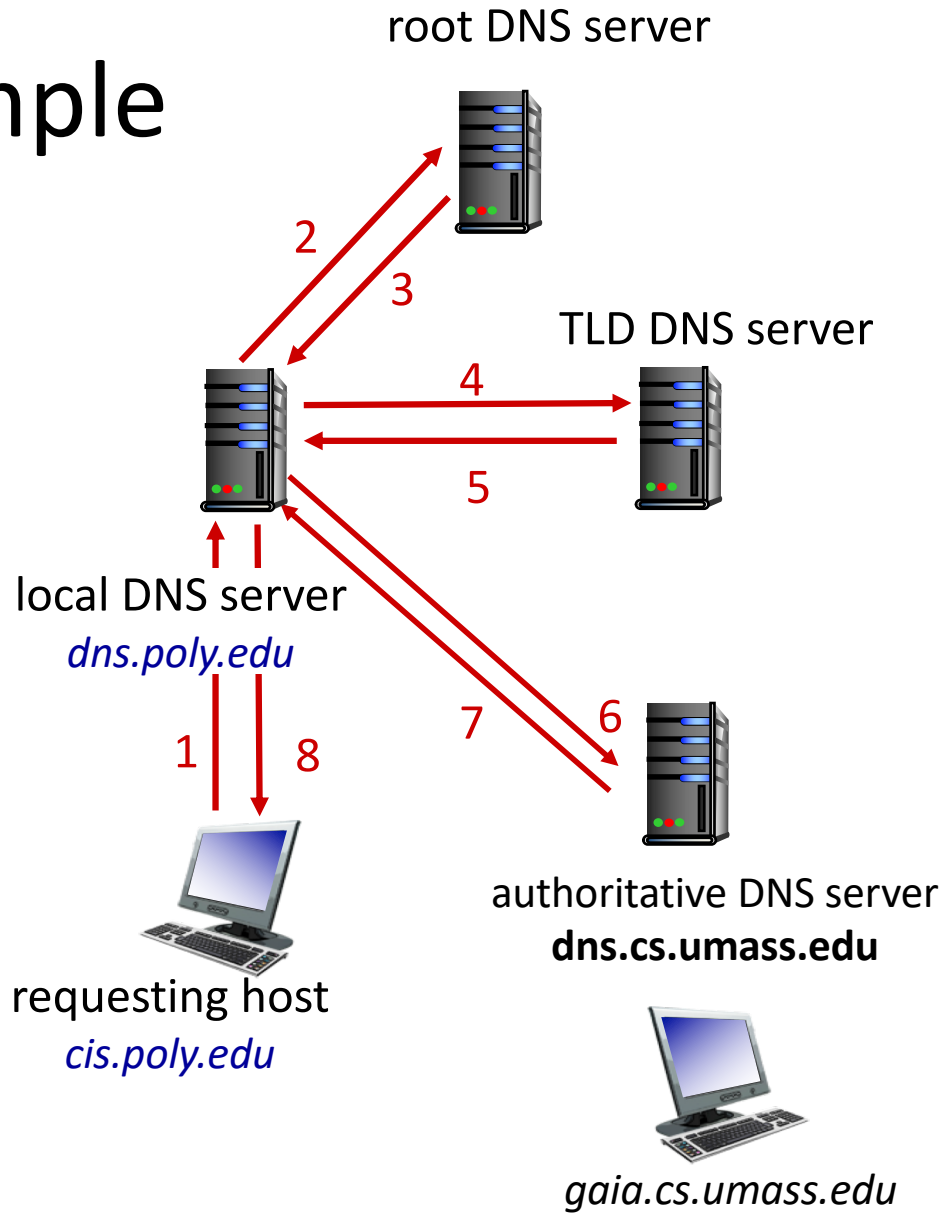
- Block IP addresses
  - IP layer
- Block hostnames
  - DNS (application layer)
- Disrupt TCP flows
  - TCP (transport layer)
  - Many possible triggers
- Disrupt HTTP transfers
  - HTTP (application layer)

DNS name resolution example

- host at cis.poly.edu wants IP address for gaia.cs.umass.edu

iterated query:

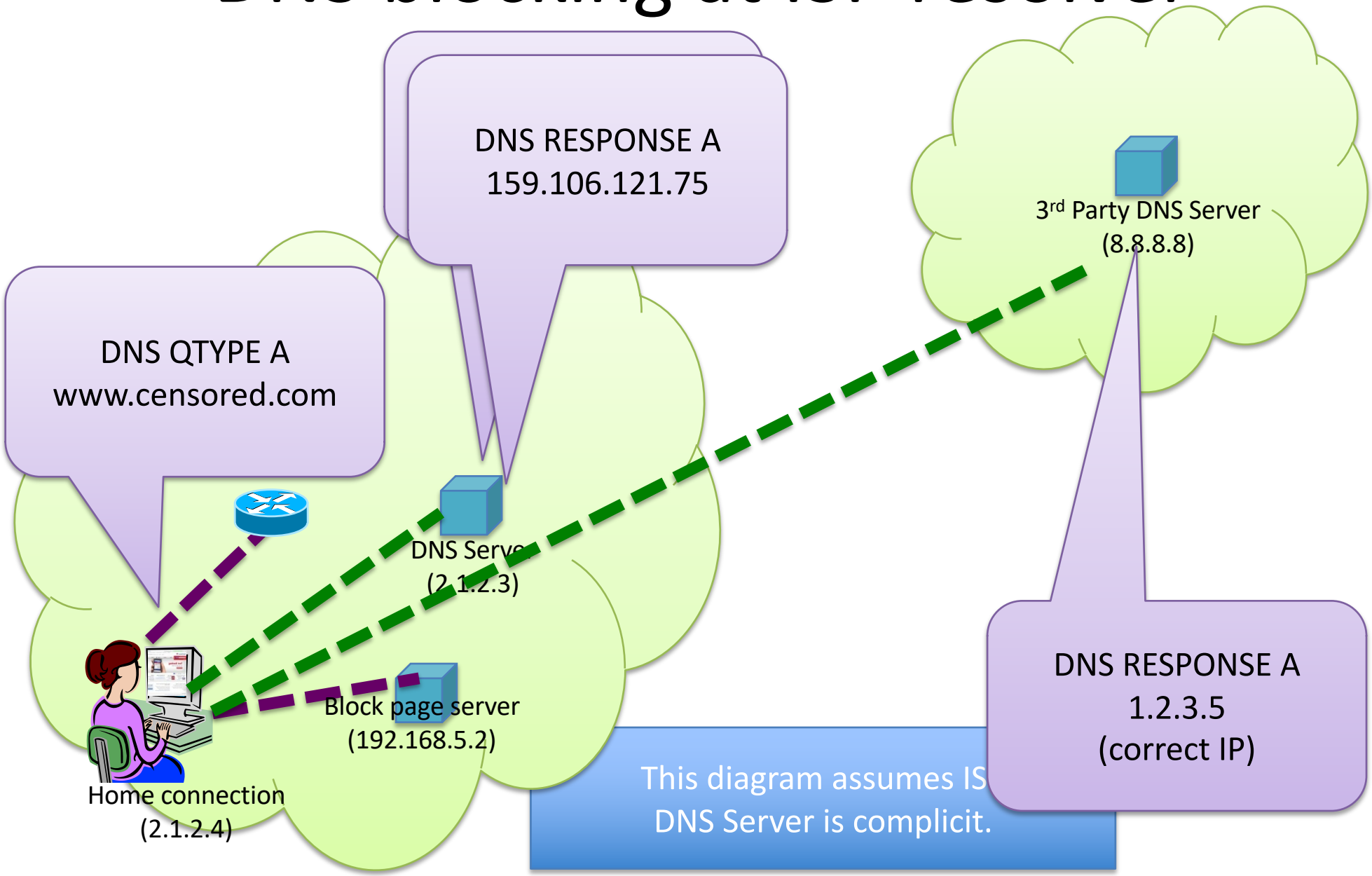
- contacted server replies with name of server to contact
- "I don't know this name, but ask this server"



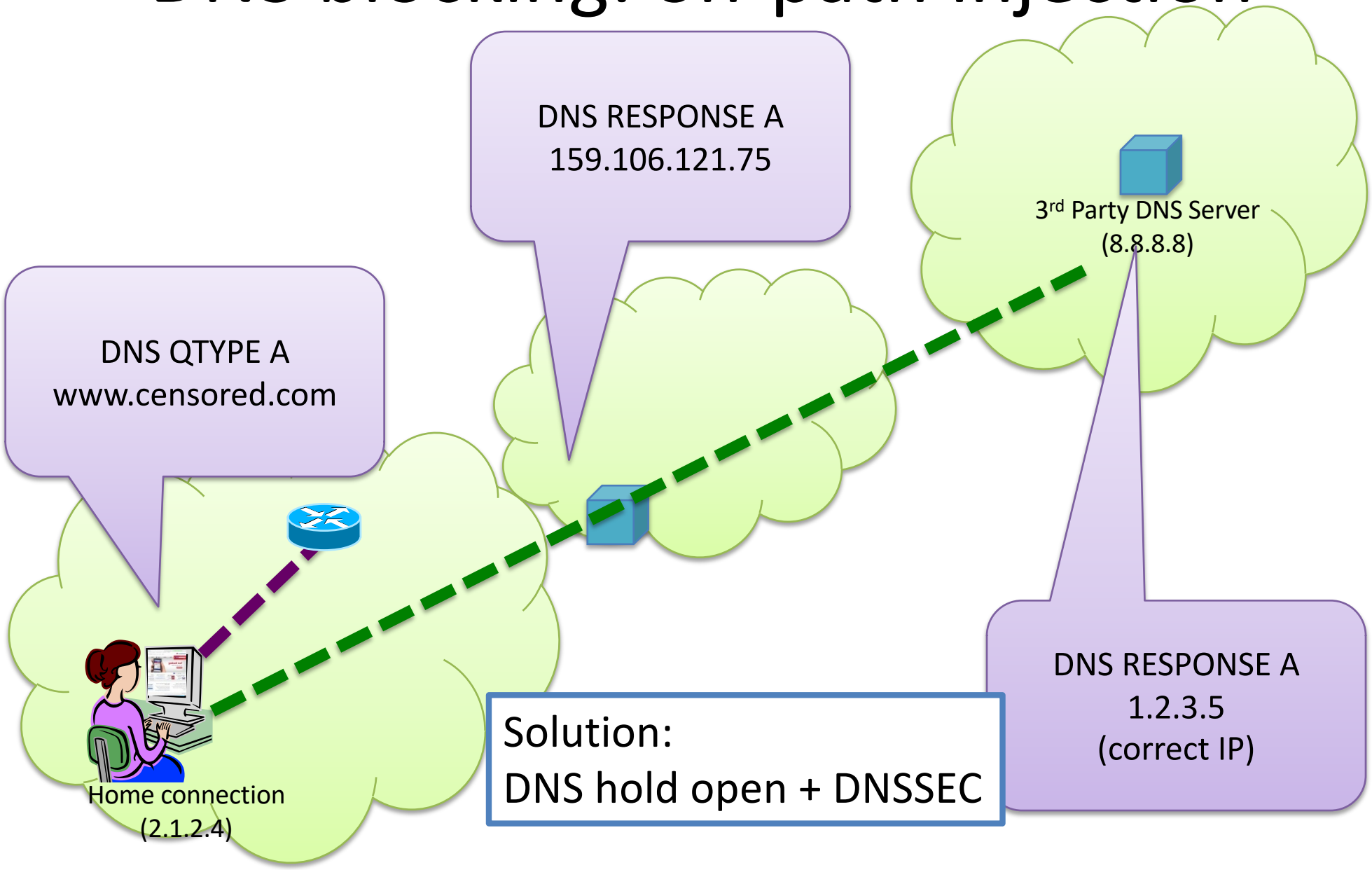
Blocking DNS names

- Can the censor pressure the ISPs?
  - ➡ Just force an entry in the recursive resolver to poison results...
- Clients can trivially evade by using alternate DNS services
  - ➡ But this does require both client changes
  - ➡ And ISPs must not block third-party DNS queries
    - ➡ Collateral damage if you do so, so generally not a good idea
- Initially used by ISPs in the UK to block The Pirate Bay...

DNS blocking at ISP resolver



DNS blocking: on-path injection



Recommended reading

The Collateral Damage of Internet Censorship by DNS Injection *

Sparks
Hovership Nebuchadnezzar
Zion Virtual Labs
zion.vlab@gmail.com

Neo†
Hovership Nebuchadnezzar
Zion Virtual Labs
zion.vlab@gmail.com

Tank
Hovership Nebuchadnezzar
Zion Virtual Labs
zion.vlab@gmail.com

Smith
Hovership Nebuchadnezzar
Zion Virtual Labs
zion.vlab@gmail.com

Dozer
Hovership Nebuchadnezzar
Zion Virtual Labs
zion.vlab@gmail.com

Challenges of Large-Scale Censorship Measurements

- Vantage points
  - Recruiting users is challenging
  - Measuring censorship has (unknown) risks
- Idea: Indirect measurements
  - HTTP cross-origin requests
  - State in TCP/IP protocol stack

We want more than just anecdotes

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

Iran

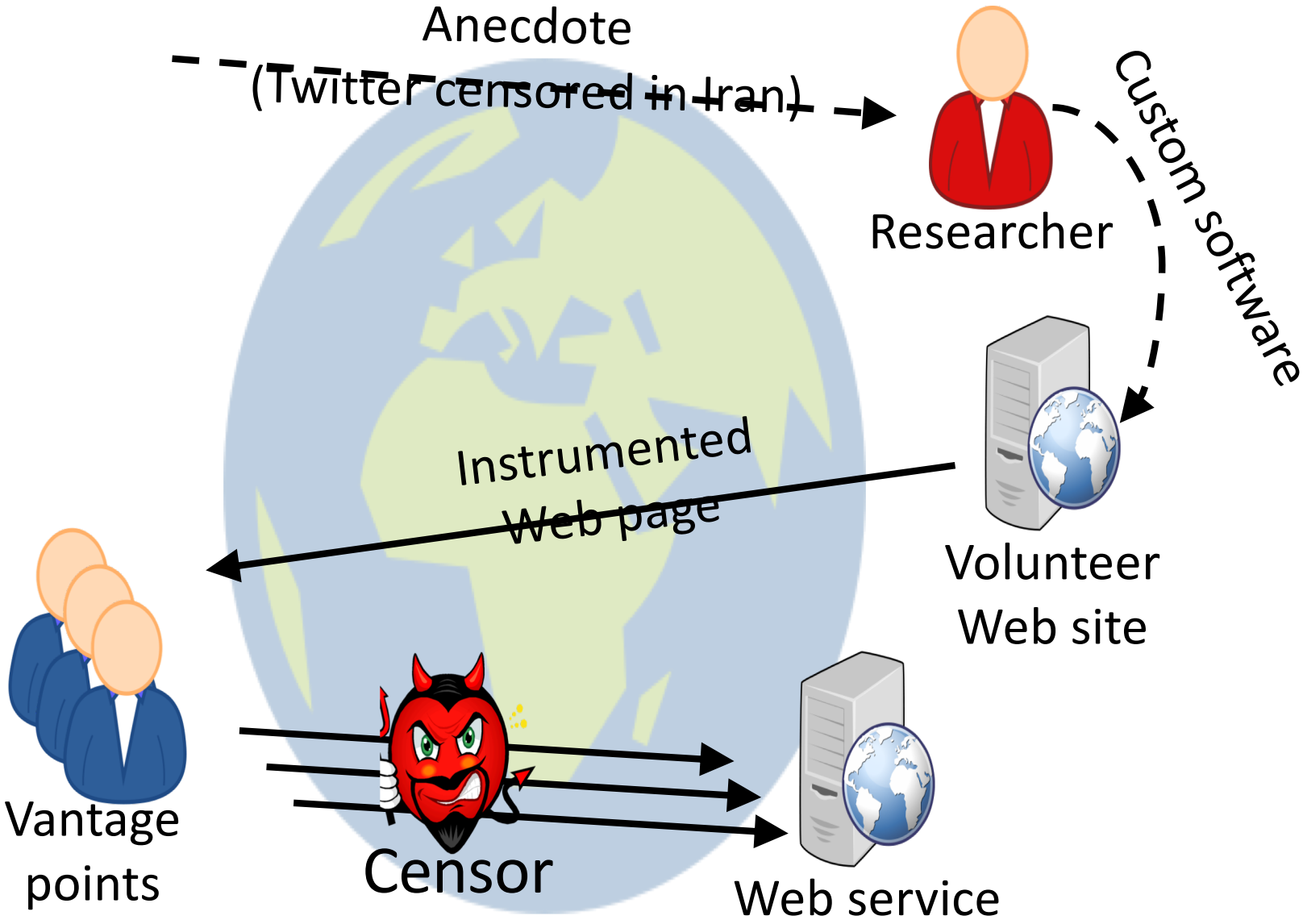
2009

Moving beyond anecdotes is hard

- Many locations, languages, and cultures
  - Recruiting lots of vantage points is necessary but hard
- Many censorship mechanisms
  - Measurements must be flexible

The biggest challenge is **diversity**

Encore: Measure using Web browsers



Ethical implications

- Encore loads potentially censored URLs
  - Possibly dangerous to users
- Informed consent is infeasible
  - Encore opt-out

How to reduce Encore's risks?

Risk depends on usage

- What URLs we ask users to measure
- Which users run those measurements

Low-risk use cases

- URLs often loaded with cross-origin requests
- Scalable network availability monitoring

Recommended reading

- S. Burnett, N. Feamster, “Encore: Lightweight Measurement of Web Censorship with Cross-Origin Requests”, SIGCOMM 2015
  - https://conferences.sigcomm.org/sigcomm/2015/pdf/papers/p653.pdf

Credits

- Slides in this class were taken from
  - Phillipa Gill's censorship class
  - Nick Feamster's measurement tutorial