# CS8803 - EMS
# Advanced Network Security and Measurement
# Class 01 — Introduction

Paul Pearce

# Welcome!

# Overview of Today

- Attendance
- Course topic overview
  - Via a taste of my research
- My learning goals for you
- Introductions
- Course logistics
  - Vital stats
  - Format
  - Grading
  - Components
- Tips
- Please interrupt me with questions

# Attendance Discussion

- Attendance
  - In-person attendance is not required. You may join online
  - Attendance either in-person or online IS required
  - This is a discussion-oriented course

- Subject to change

# Internet Attacks

The New York Times

***All 3 Billion Yahoo Accounts Were Affected by 2013 Attack***

The Washington Post
*Democracy Dies in Darkness*

**Hacks of OPM databases compromised 22.1 million people, federal authorities say**

The New York Times

***Mystery of Motive for a Ransomware Attack: Money, Mayhem or a Message?***

The New York Times

***Cyberattack Hits Ukraine Then Spreads Internationally***

The New York Times

***Equifax Says Cyberattack May Have Affected 143 Million in the U.S.***

The Washington Post
*Democracy Dies in Darkness*

**Computer security experts fear second wave of 'biggest ransomware attack ever'**

# Internet Adversaries

**FBI**

**2016 Internet Crime Report**

Loss from cybercrime exceeded $1.3B

The New York Times

*A New Era of Internet Attacks Powered by Everyday Devices*

The New York Times

*Russian Cyberforgers Steal Millions a Day With Fake Sites*

The Washington Post
Democracy Dies in Darkness

36 indicted in global cybercrime ring that stole $530M

The Washington Post
Democracy Dies in Darkness

Asia & Pacific

China's scary lesson to the world: Censoring the Internet works

The Washington Post
Democracy Dies in Darkness

WorldViews

Turkey just banned Wikipedia, labeling it a 'national security threat'

THE VERGE TECH SCIENCE CULTURE CARS REVIEWS LONGFORM VIDEO MORE

Two-thirds of the world's internet users live under government censorship: report

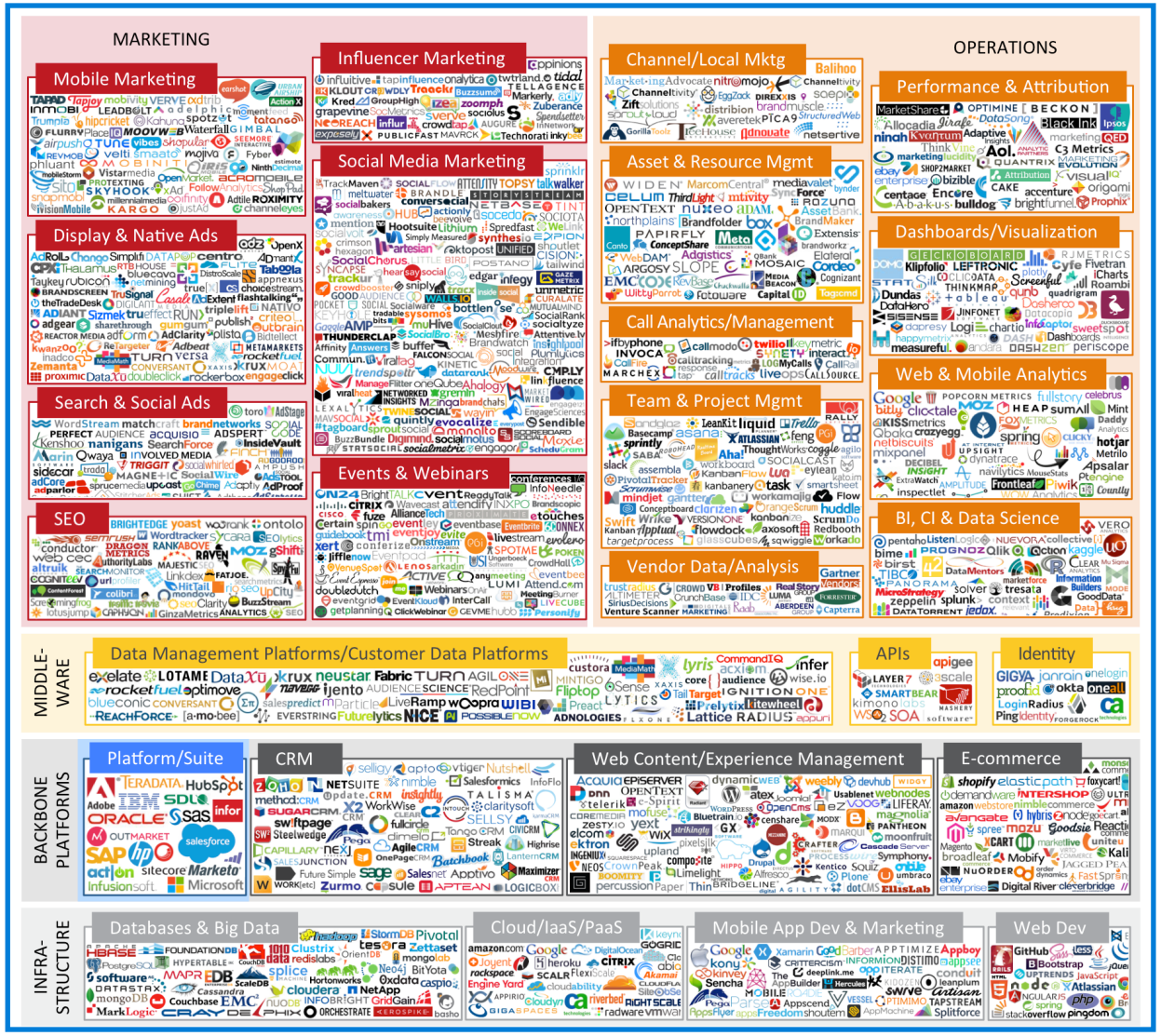
*Web freedom declined across the globe for the sixth consecutive year, according to a new report*

**Criminals** ⟷ **Nation-States**

# Why do these attacks persist?

- Extensive work on defenses
- Yet difficult to remediate?
  - Complexity, layering
  - Difficult to identify
  - Landscape favors the attacker
- How do you develop effective solutions?
- Do we *actually* understand the problems?

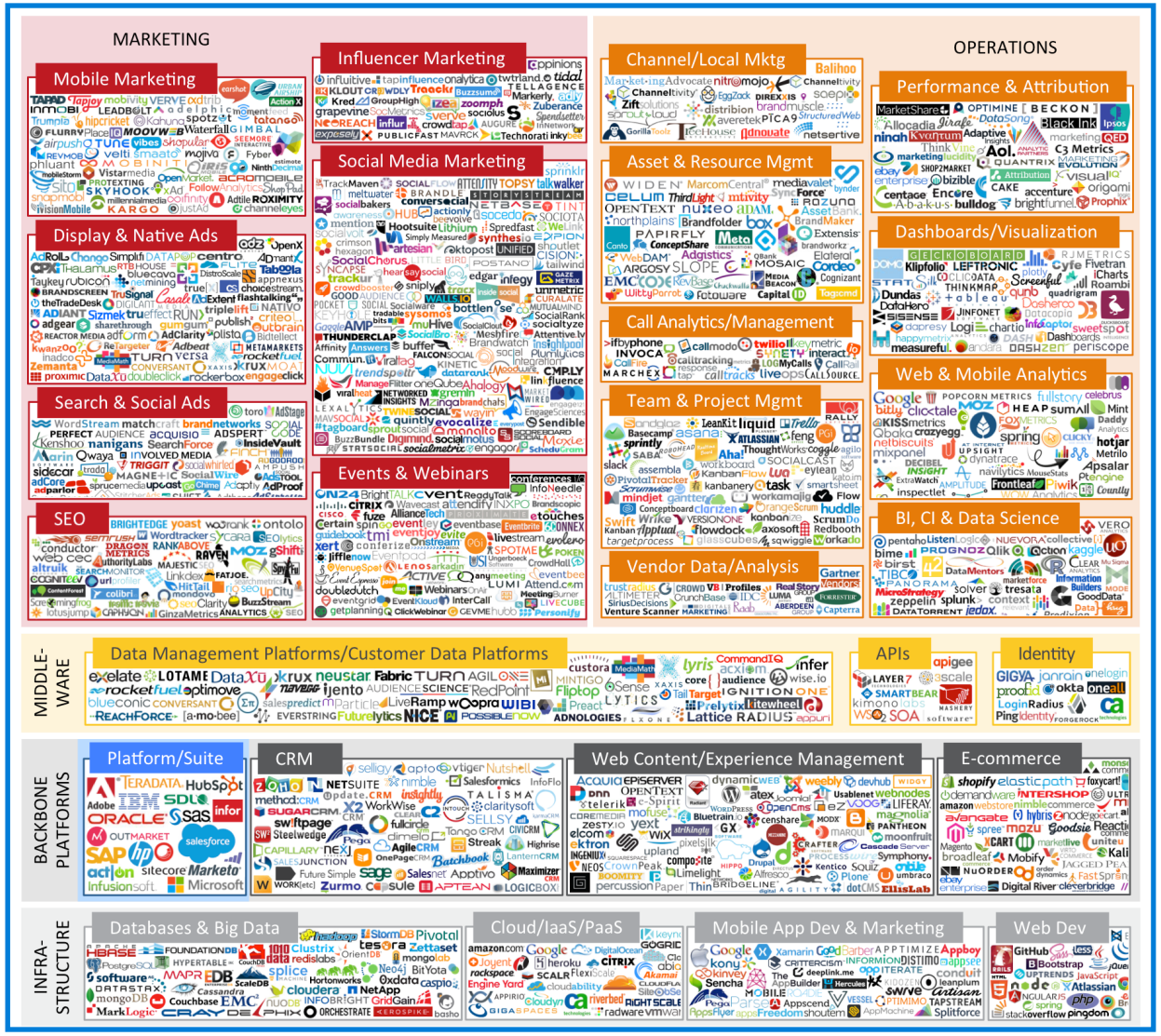
  → Measurement



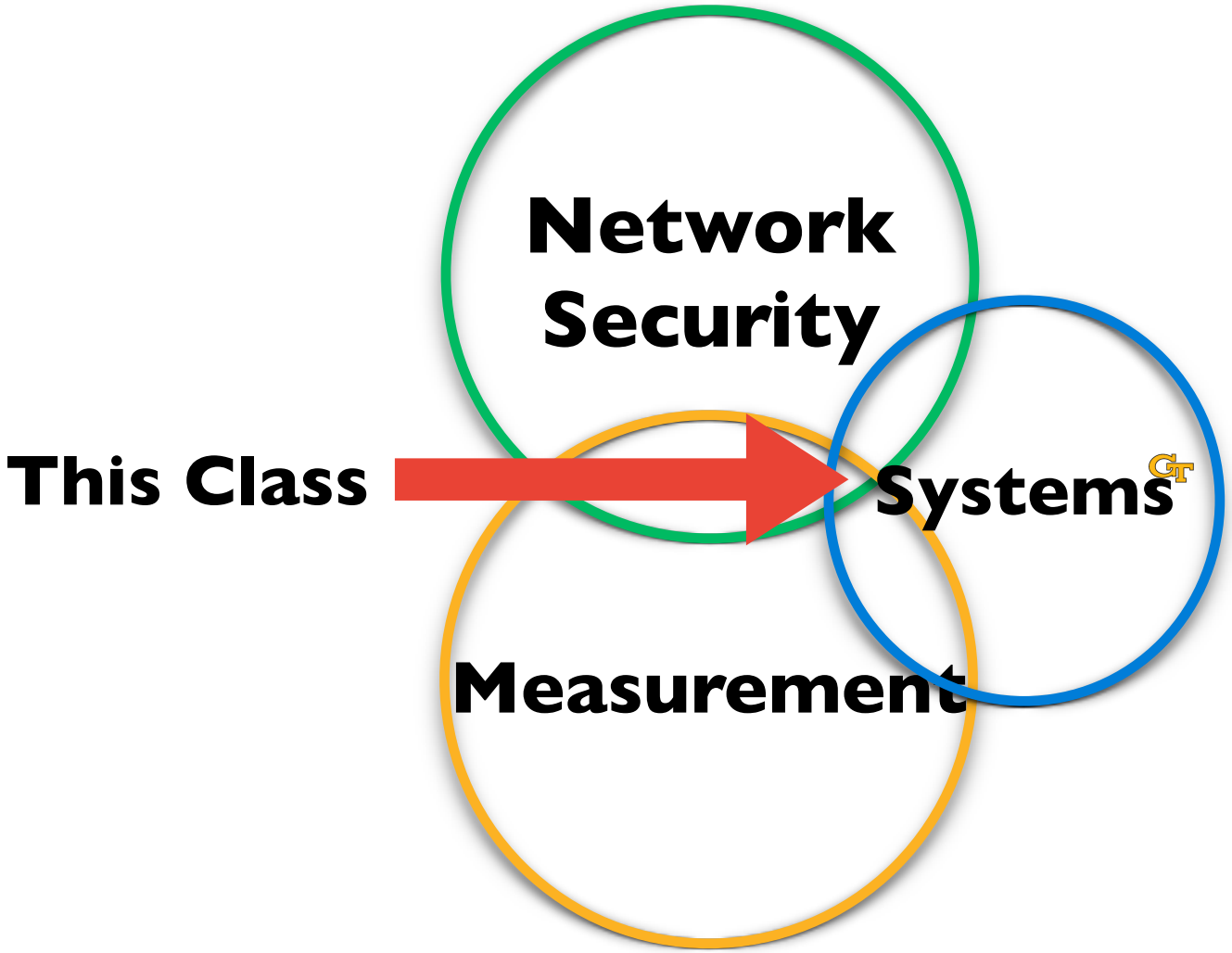
# How Do We Measure?

- How do you know *what* to measure?
- Can't measure everything everywhere
  - Layering
  - Location
- Can't measure directly
- My work
  - Infer & derive what you can't measure directly

→ Remediation



# This Class




: Not To Scale

# My Work

## Cybercrime

- Characterizing Large-Scale Click Fraud in ZeroAccess **(ACM CCS)**
- Ad Injection at Scale: Assessing Deceptive Advertisement Modifications **(IEEE S&P Distinguished Paper)**
- To Catch a Ratter: Monitoring the Behavior of Amateur DarkComet RAT Operators in the Wild **(IEEE S&P)**

## Internet Censorship

- Augur: Internet-Wide Detection of Connectivity Disruptions **(IEEE S&P)**
- Global Measurement of DNS Manipulation **(USENIX Security)**
- Characterizing the Nature and Dynamics of Tor Exit Blocking **(USENIX Security)**

# My Work

## Cybercrime

- Characterizing Large-Scale Click Fraud in ZeroAccess **(ACM CCS)**
- Ad Injection at Scale: Assessing Deceptive Advertisement Modifications **(IEEE S&P Distinguished Paper)**
- To Catch a Ratter: Monitoring the Behavior of Amateur DarkComet RAT Operators in the Wild **(IEEE S&P)**

## Internet Censorship

- Augur: Internet-Wide Detection of Connectivity Disruptions **(IEEE S&P)**
- Global Measurement of DNS Manipulation **(USENIX Security)**
- Characterizing the Nature and Dynamics of Tor Exit Blocking **(USENIX Security)**

# Cybercrime

Characterizing Large-Scale Click Fraud in ZeroAccess (ACM CCS)

# Cybercrime and Advertising Abuse

- Monetarily Driven
- Costs $6.5 Billion Annually
- Impacts 10s of millions of users
- Ad losses hurt everyone

THE WALL STREET JOURNAL.

**Russian Hackers Stole Millions From Video Advertisers, Ad Fraud Company Says**

White Ops says Russian hacking operation created fake users and sites to scam online advertisers out of more than $3 million a day

THE WALL STREET JOURNAL.

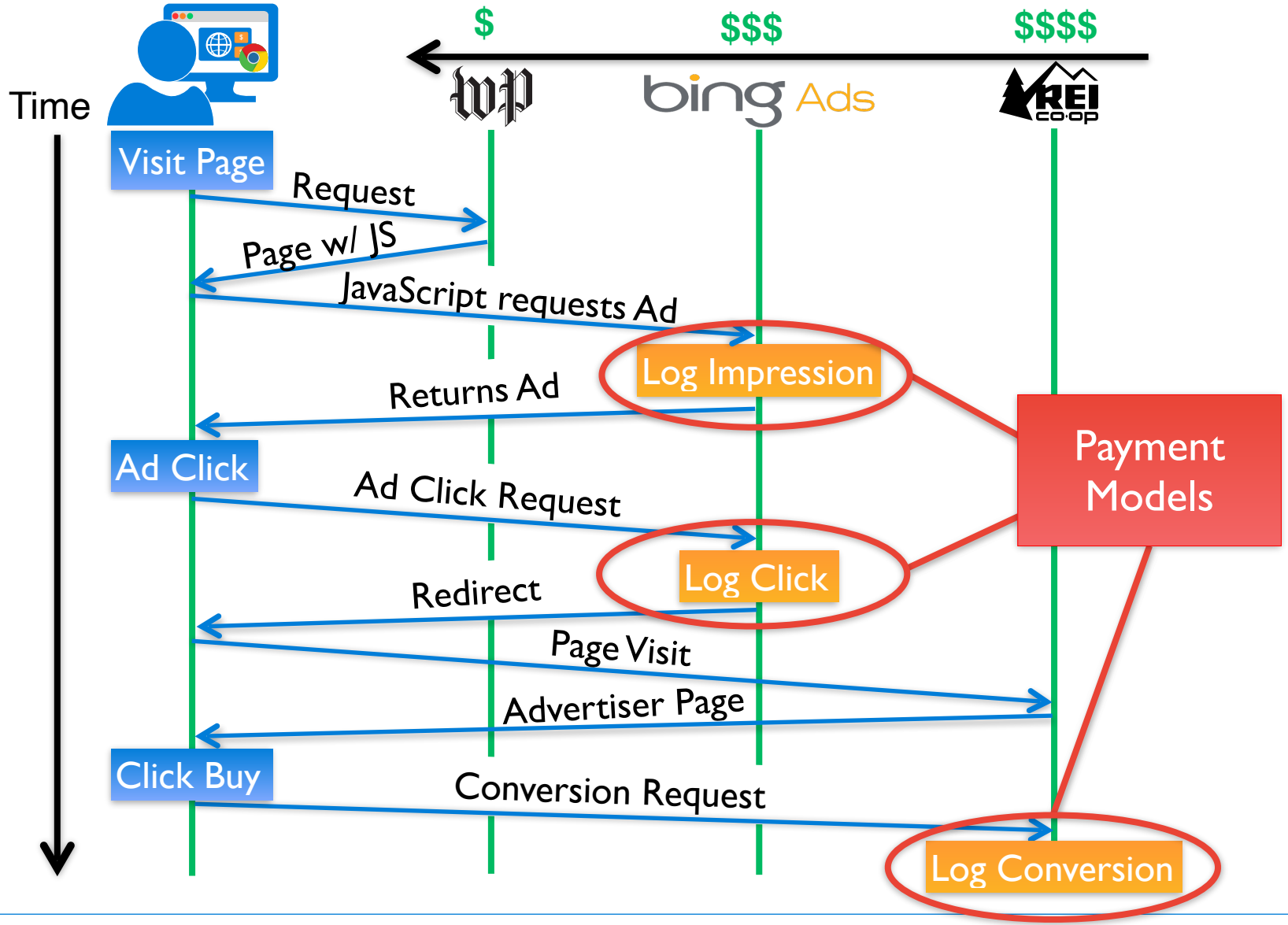
**Fake-Ad Operation Used to Steal From Publishers Is Uncovered**

Adform says 'Hyphbot' scheme created fake websites, nonhuman traffic to scam advertisers of more than $500,000 a day

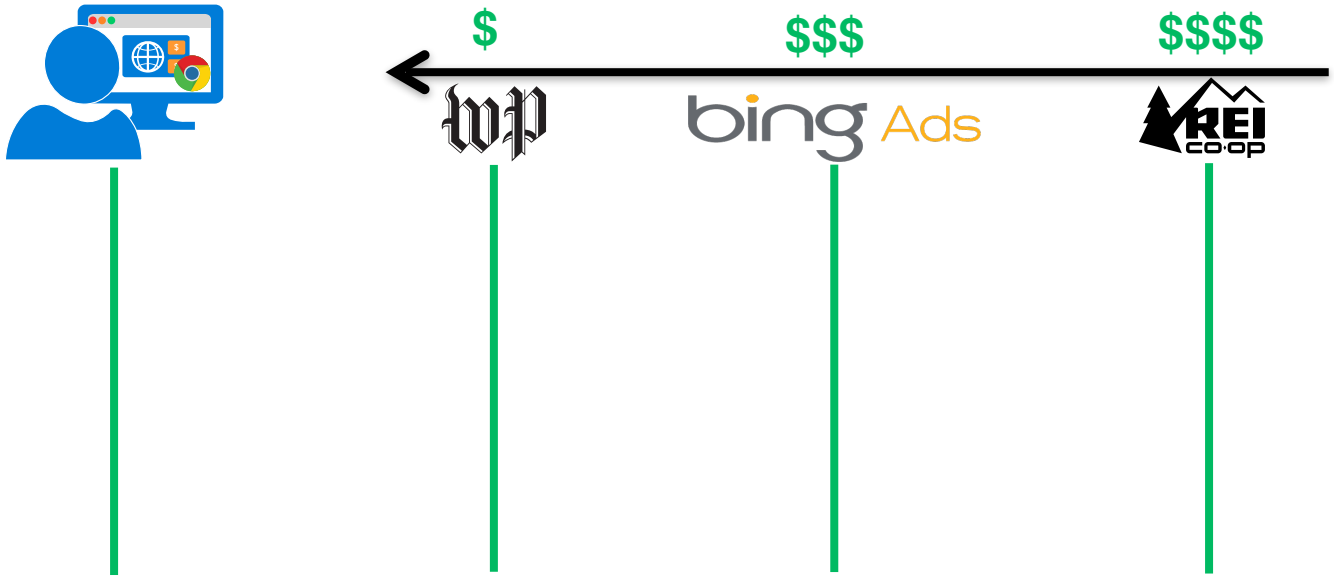
# Ad Abuse Overview

- Goal: Illuminate the nature and behavior of large scale ad abuse
  - How does click fraud look at scale?
  - $$$
  - → Defenses
- Our lens: ZeroAccess
  - Structure and function of the botnet
- Reveal
  - Innovative fraud structure
  - Complex supporting ecosystem
- Remediation and takedown

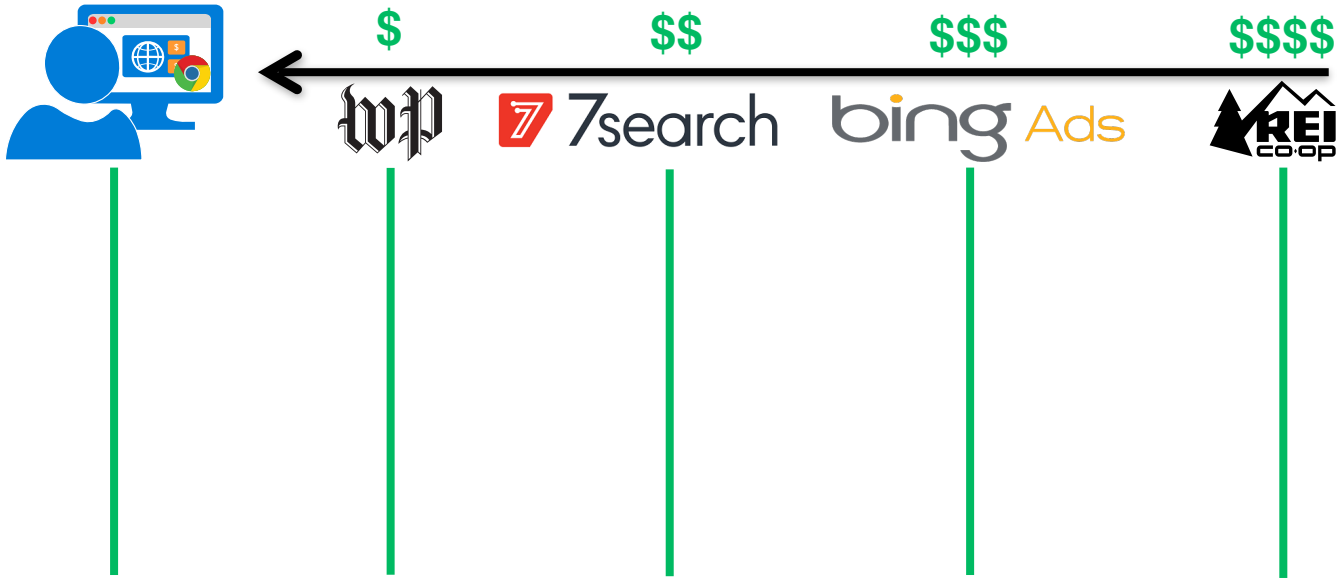
# Online Advertising: Behind The Scenes



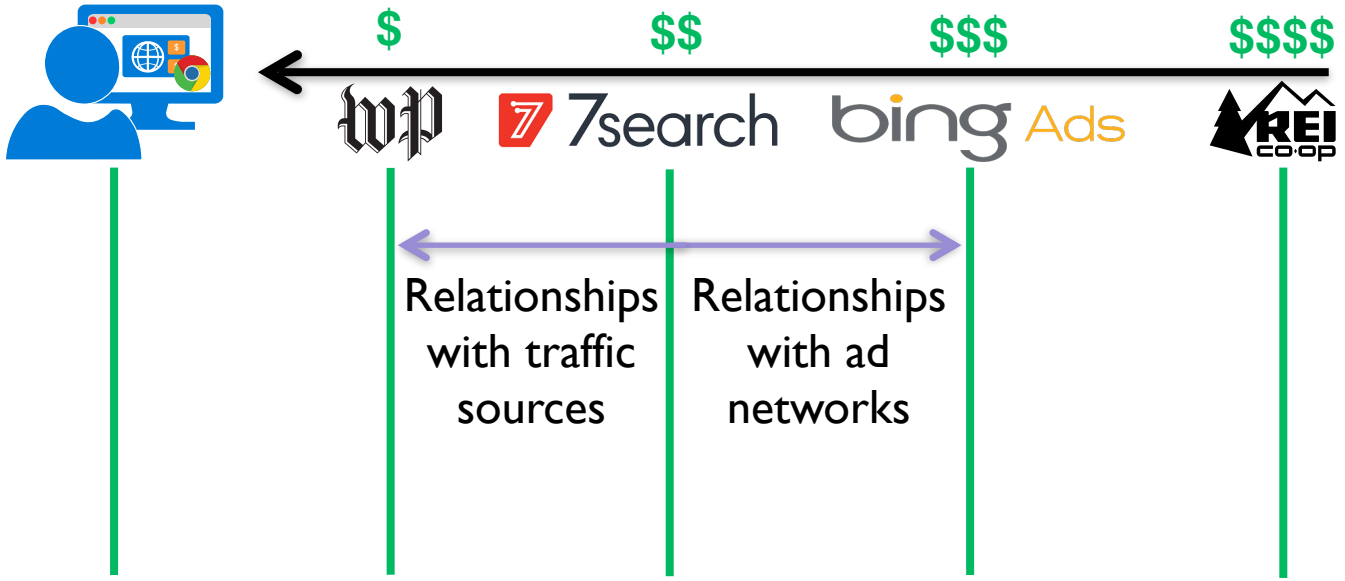
# Online Advertising: Behind The Scenes



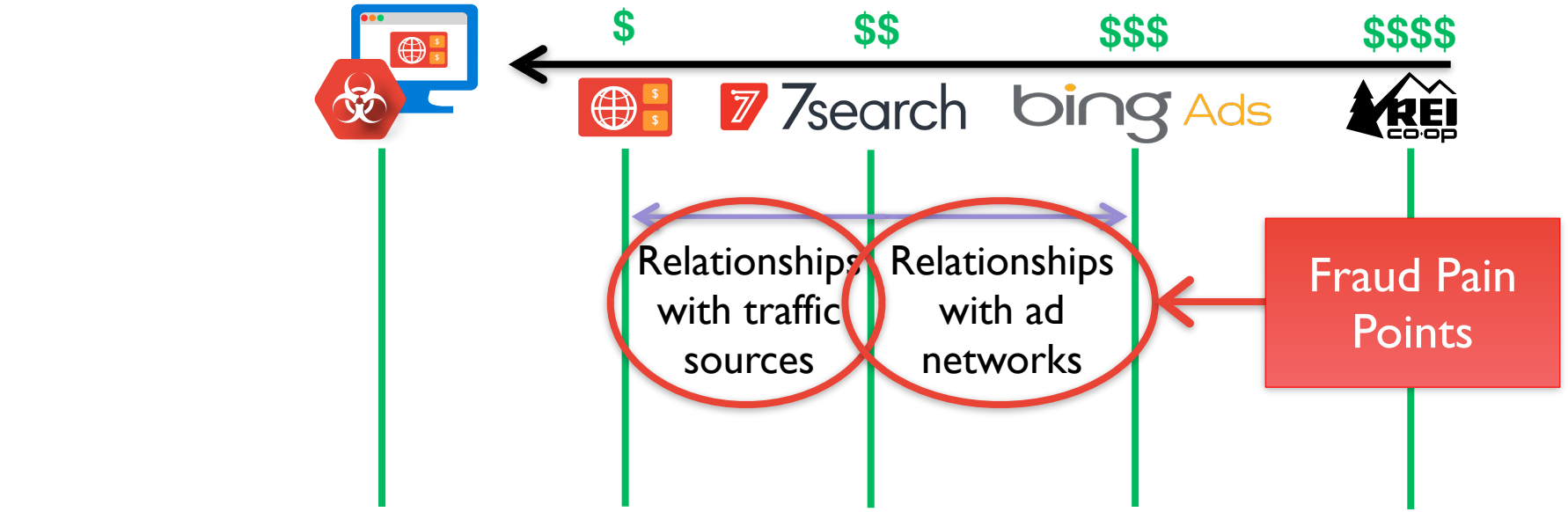
# Online Advertising: Behind The Scenes

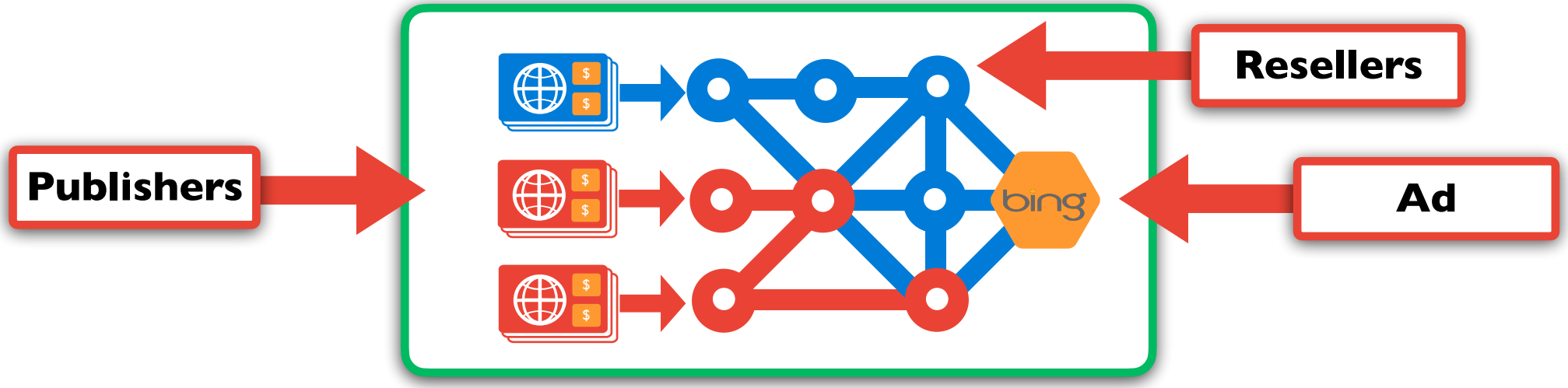


# Online Advertising: Behind The Scenes



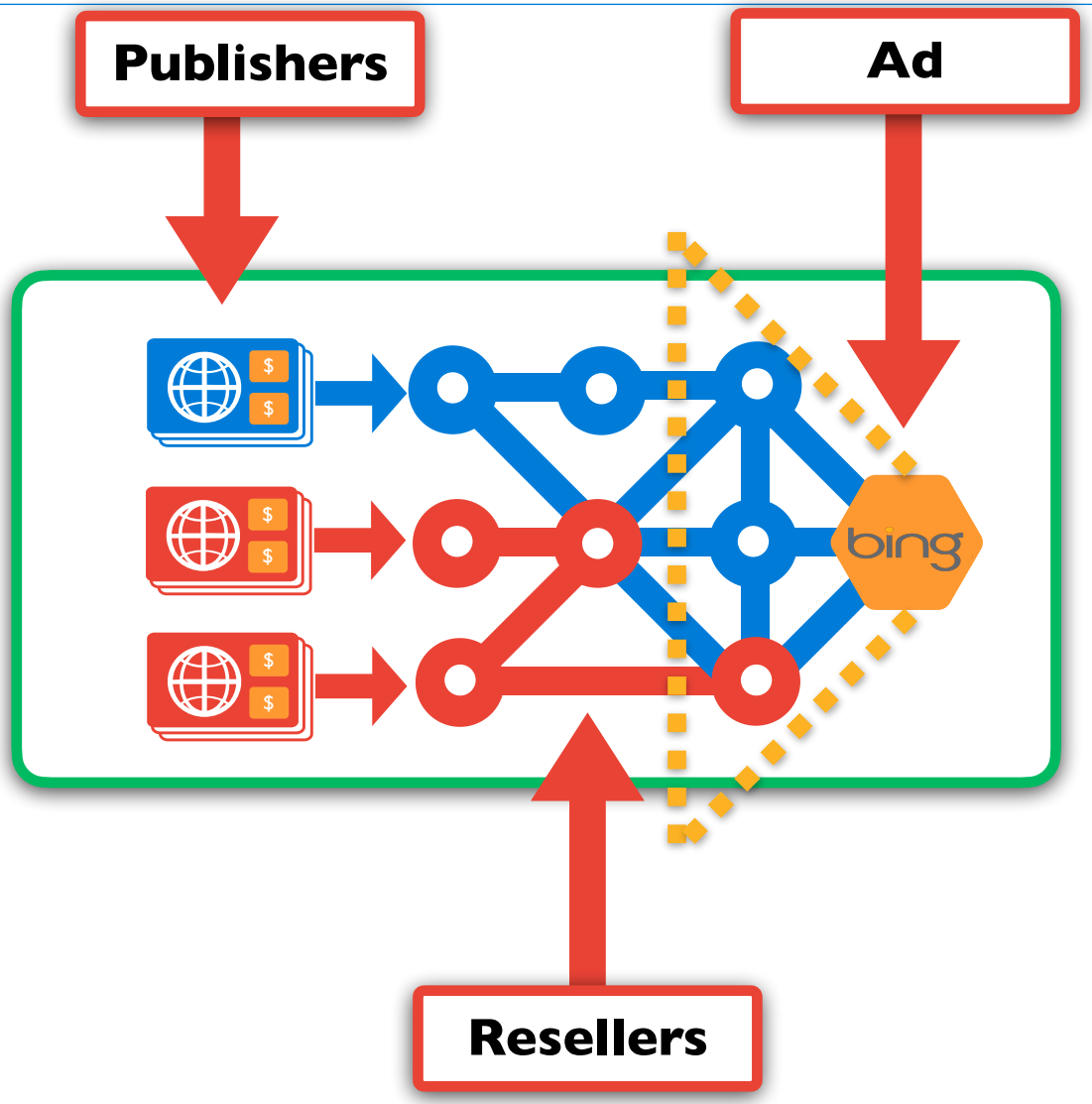
# Online Advertising: Behind The Scenes



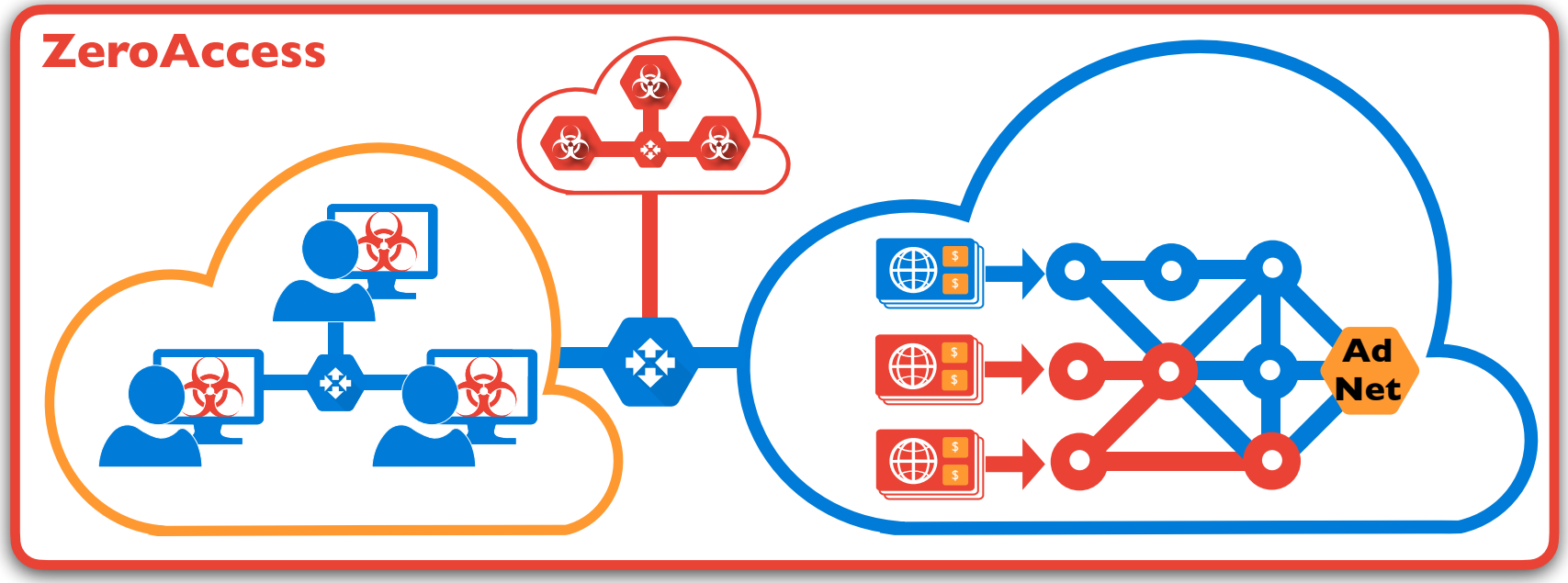
# Combating Fraud

- Fraud is fought at the Ad Network
- All you get is the click-stream
- Difficult to see complete picture
- Fraud is laundered through resellers
  - Resellers mask identity
  - Reseller mix ("cut") fraudulent traffic with real traffic
- Click-stream perspective can't peer beyond resellers



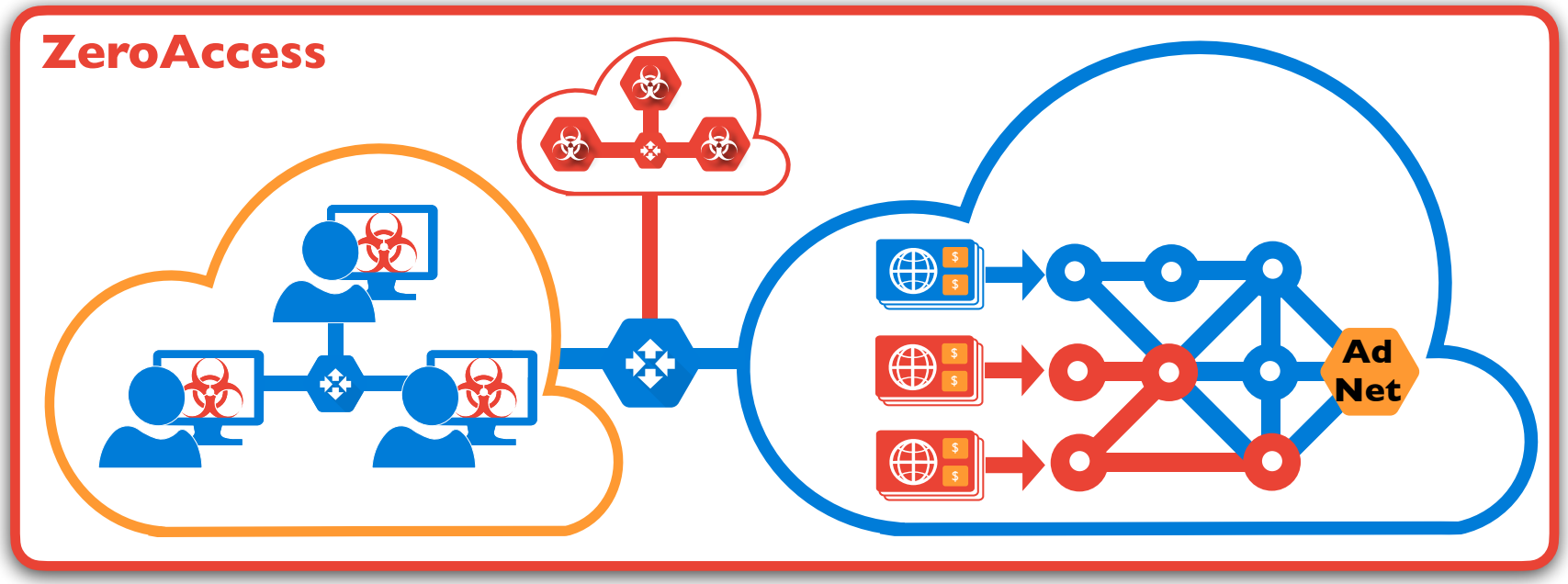
# ZeroAccess

- Malware Driven Botnet
- P2P Control Structure
- Redundant Ad C&C
- Large Ad Footprint



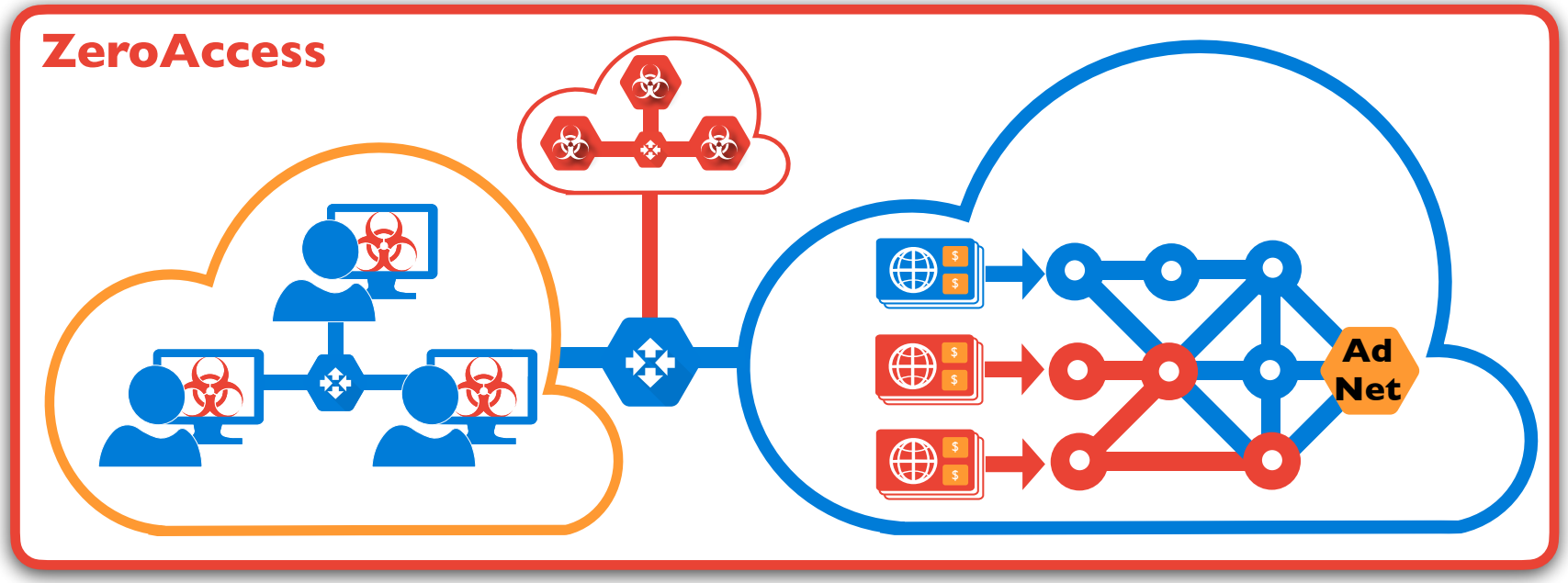
# My Work

- Explored defenses in 2 ways
  - Infiltration
  - Ecosystem



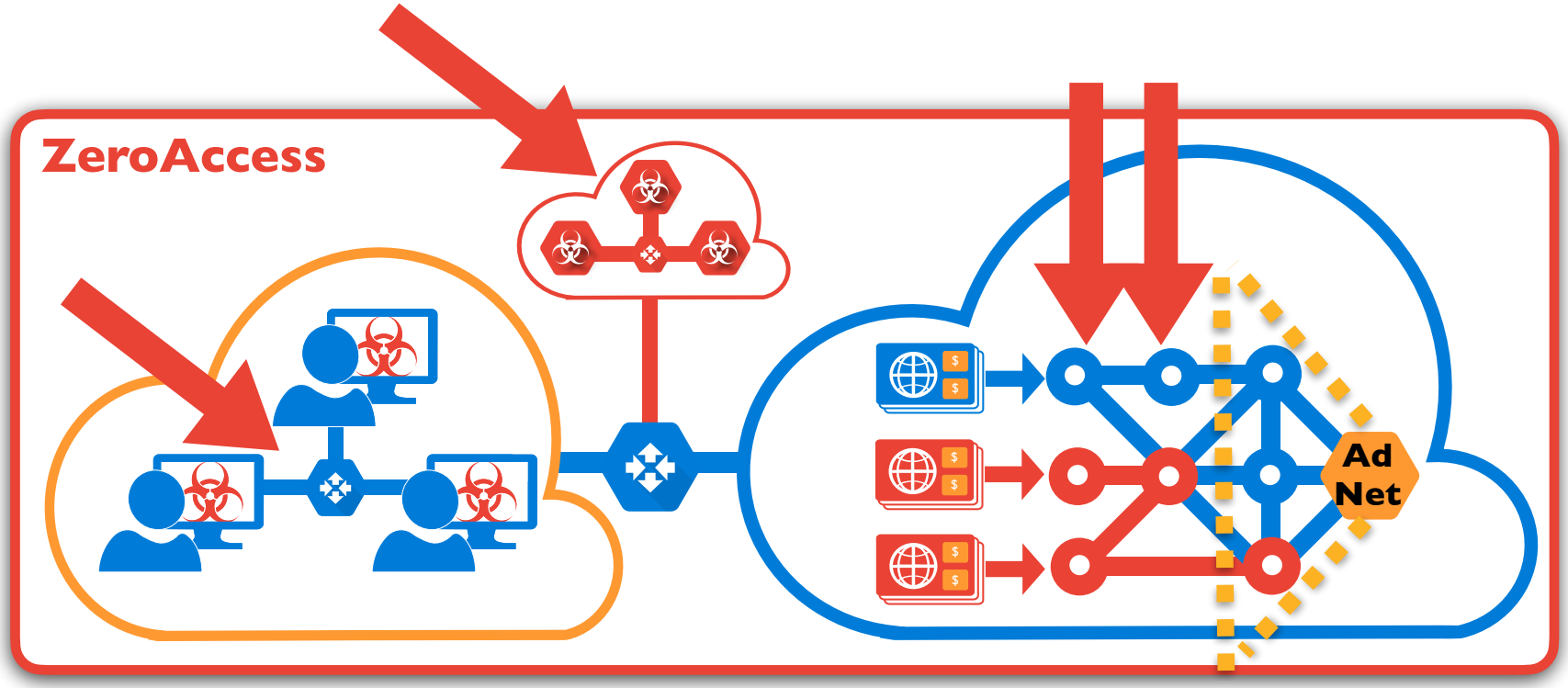
# Botnet Infiltration

- Criminals have information advantage
- With infiltration, get insider perspective



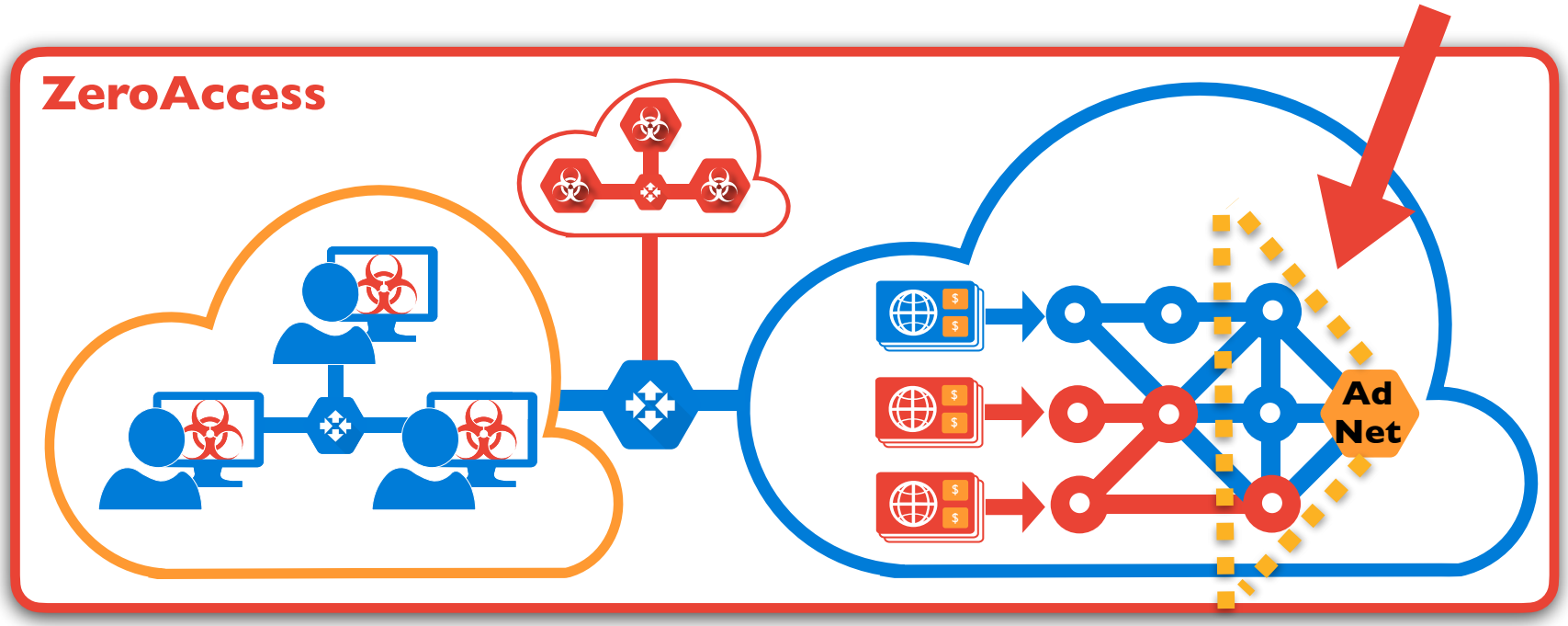
# Botnet Infiltration

- How?
  - Reverse engineer
  - P2P Infiltration
  - C&C Interaction
  - Track Clicks



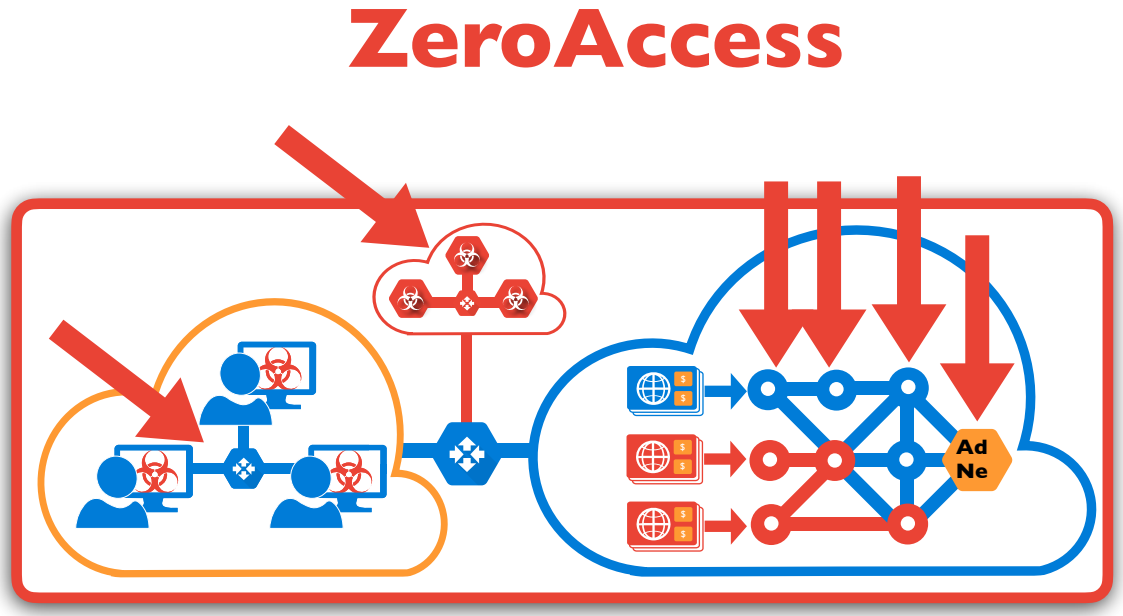
# Ad Network Partnership

- Partner with large top tier ad network
- Get insider view
- **Tie every ad click back with our external data**
- Examining tens of millions in ad data

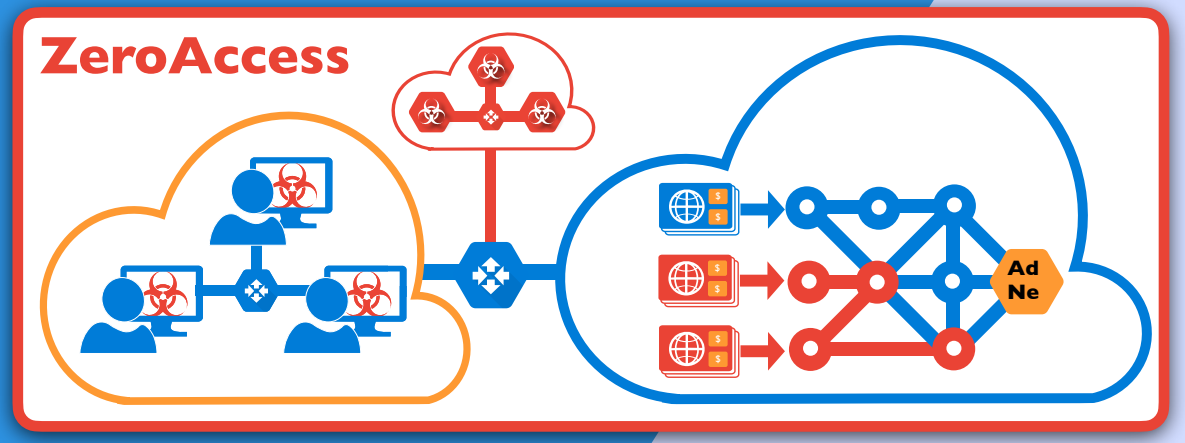


# Pulling the Pieces Together

| Dataset | Size |
|---|---|
| P2P Infiltration | • 260M Commands<br>• 1.2M Victims |
| C&C Infiltration | 16M Commands |
| C&C Interaction | ~2K Click Chains |
| Ad Clicks | > 10TB |
| Ad Affiliate | > 2TB |



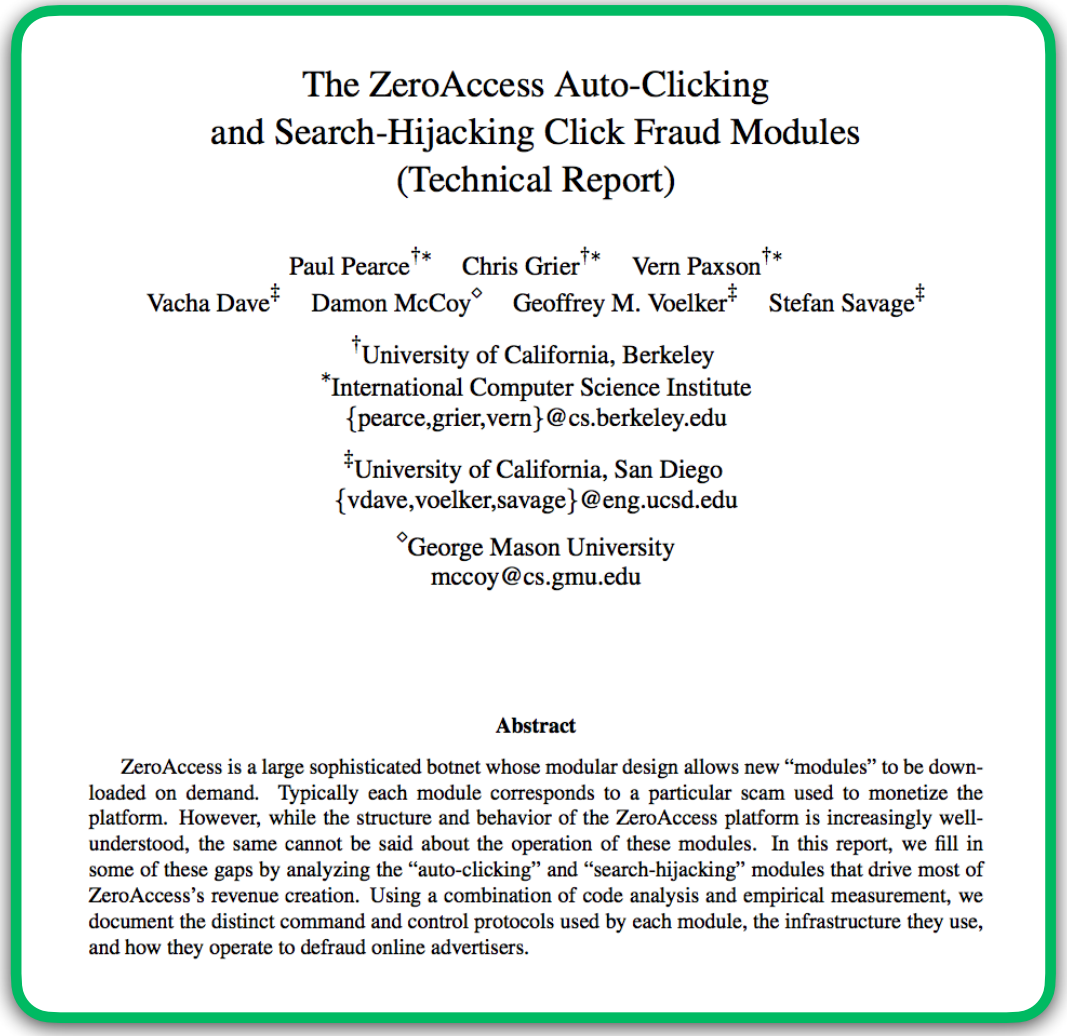
# ZeroAccess: Results




- Identified 54 criminal affiliates
- Byzantine Structure
- Launders Fraud, masks criminals
- Millions of $ in fraud
- Fraud remediated in the ad network

# Results: Takedown

- Collaborated with Microsoft DCU, FBI, and Europol
- Produced a technical report which was Exhibit 1 in legal action
- Technically facilitated a takedown of C&C infrastructure



The ZeroAccess Auto-Clicking
and Search-Hijacking Click Fraud Modules
(Technical Report)


Paul Pearce[†*] Chris Grier[†*] Vern Paxson[†*]
Vacha Dave[‡] Damon McCoy[◇] Geoffrey M. Voelker[‡] Stefan Savage[‡]

[†]University of California, Berkeley
[*]International Computer Science Institute
{pearce,grier,vern}@cs.berkeley.edu

[‡]University of California, San Diego
{vdave,voelker,savage}@eng.ucsd.edu

[◇]George Mason University
mccoy@cs.gmu.edu

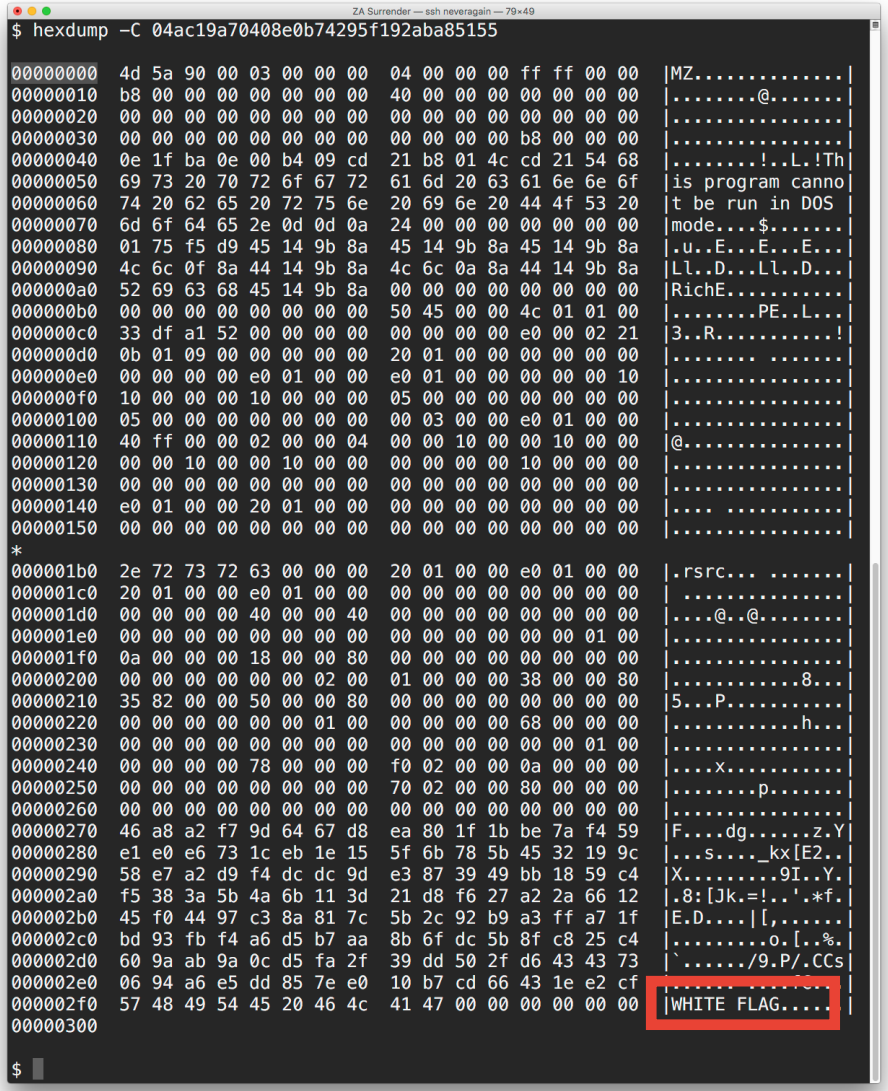


**Abstract**

ZeroAccess is a large sophisticated botnet whose modular design allows new "modules" to be downloaded on demand. Typically each module corresponds to a particular scam used to monetize the platform. However, while the structure and behavior of the ZeroAccess platform is increasingly well-understood, the same cannot be said about the operation of these modules. In this report, we fill in some of these gaps by analyzing the "auto-clicking" and "search-hijacking" modules that drive most of ZeroAccess's revenue creation. Using a combination of code analysis and empirical measurement, we document the distinct command and control protocols used by each module, the infrastructure they use, and how they operate to defraud online advertisers.

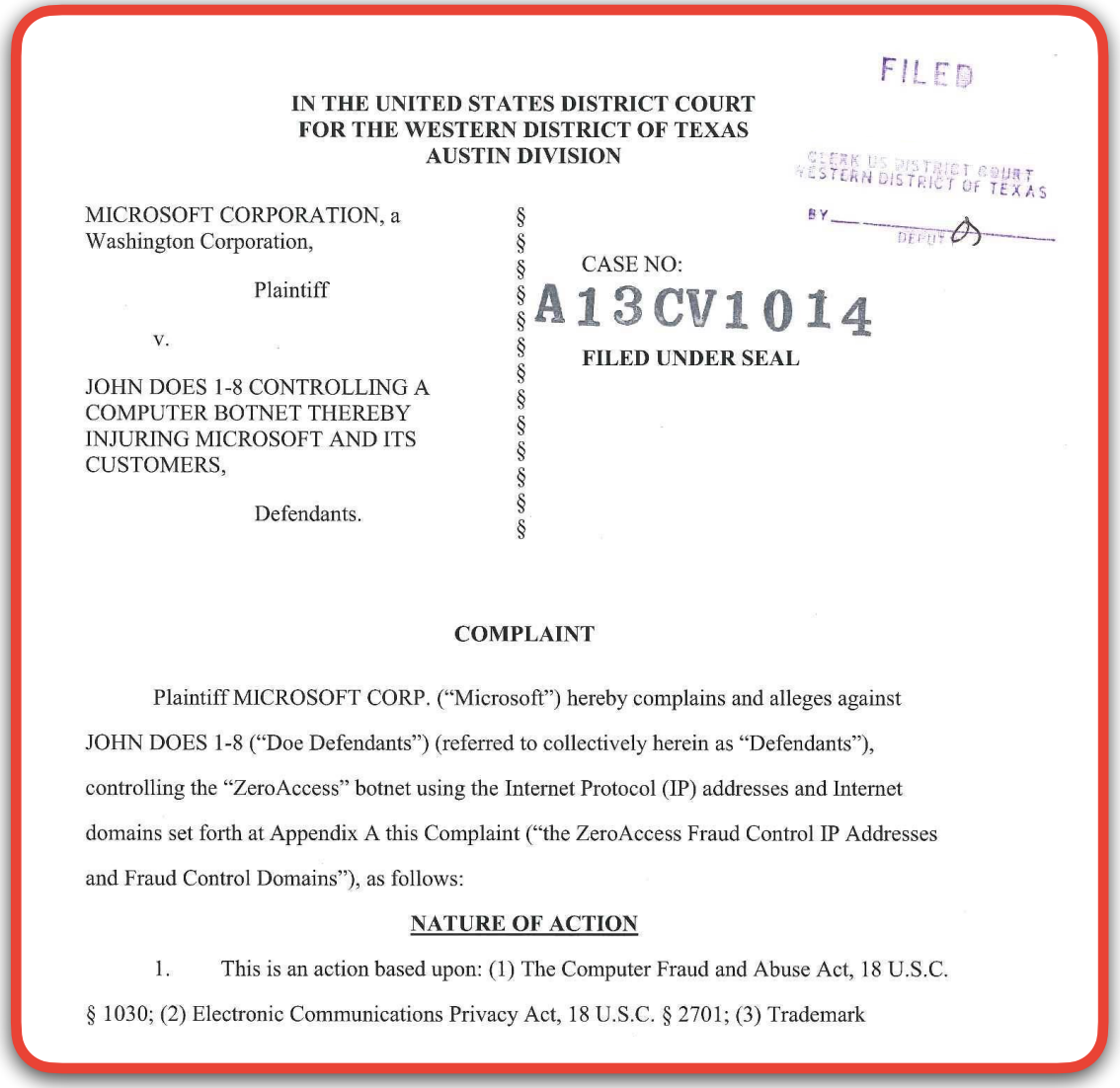
# Results: Cleanup

- Takedown was iterative
- Criminals attempted to revive the botnet
- Technique was so effective they gave a literal white flag of surrender
- This ended ZeroAccess
  - → Millions of users



# Results: *Really* Effective Infiltration (The Fun Bit)

- Infiltration of the botnet was so effective..
- While on takedown ops call…
- Received a court order as John Doe criminal defendants
- Quickly resolved, with first round on DCU



FILED

IN THE UNITED STATES DISTRICT COURT
FOR THE WESTERN DISTRICT OF TEXAS
AUSTIN DIVISION

CLERK US DISTRICT COURT
WESTERN DISTRICT OF TEXAS
BY _____ DEPUTY

MICROSOFT CORPORATION, a Washington Corporation,

Plaintiff

v.

JOHN DOES 1-8 CONTROLLING A COMPUTER BOTNET THEREBY INJURING MICROSOFT AND ITS CUSTOMERS,

Defendants.

CASE NO:
A13CV1014
FILED UNDER SEAL

COMPLAINT

Plaintiff MICROSOFT CORP. ("Microsoft") hereby complains and alleges against JOHN DOES 1-8 ("Doe Defendants") (referred to collectively herein as "Defendants"), controlling the "ZeroAccess" botnet using the Internet Protocol (IP) addresses and Internet domains set forth at Appendix A this Complaint ("the ZeroAccess Fraud Control IP Addresses and Fraud Control Domains"), as follows:

NATURE OF ACTION

1. This is an action based upon: (1) The Computer Fraud and Abuse Act, 18 U.S.C. § 1030; (2) Electronic Communications Privacy Act, 18 U.S.C. § 2701; (3) Trademark

# Cybercrime

# Nation-States

# My Work

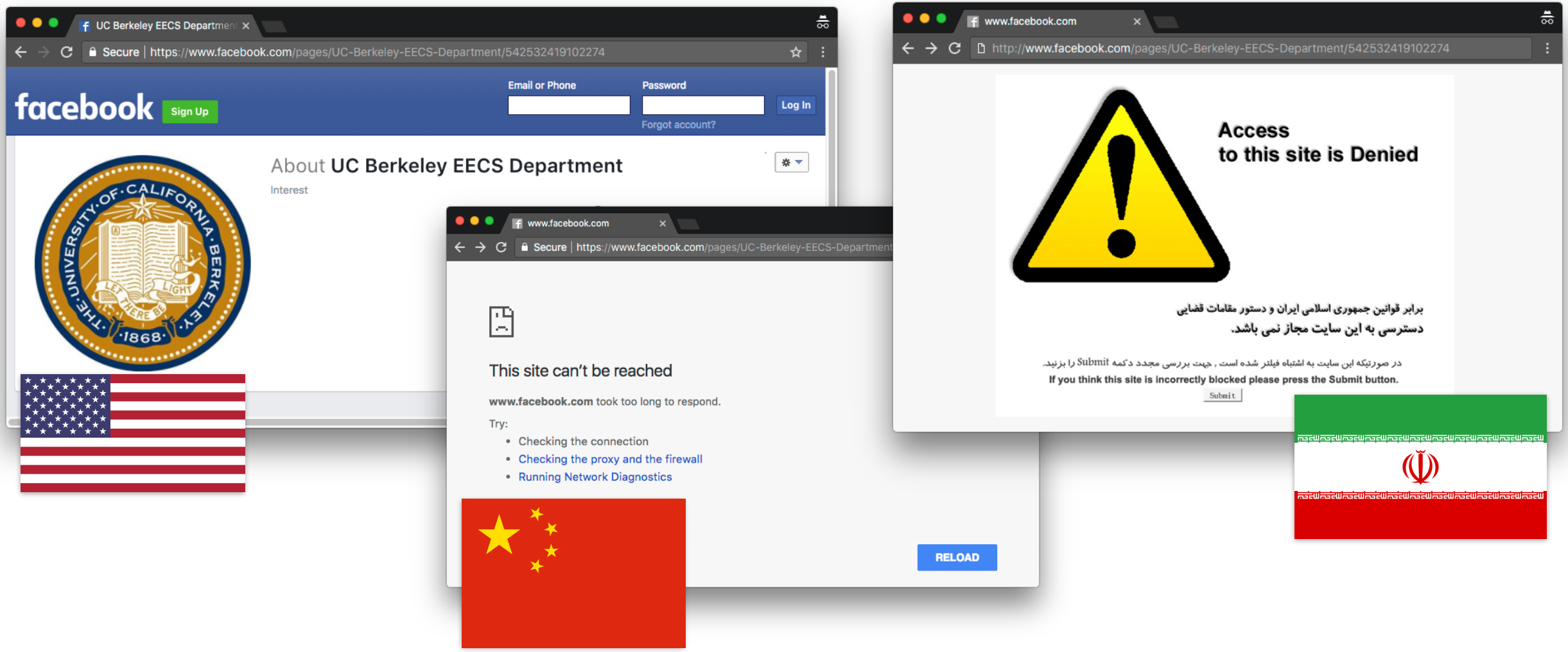
## Cybercrime

- Characterizing Large-Scale Click Fraud in ZeroAccess **(ACM CCS)**
- Ad Injection at Scale: Assessing Deceptive Advertisement Modifications **(IEEE S&P Distinguished Paper)**
- To Catch a Ratter: Monitoring the Behavior of Amateur DarkComet RAT Operators in the Wild **(IEEE S&P)**

## Internet Censorship

- Augur: Internet-Wide Detection of Connectivity Disruptions **(IEEE S&P)**
- Global Measurement of DNS Manipulation **(USENIX Security)**
- Characterizing the Nature and Dynamics of Tor Exit Blocking **(USENIX Security)**

# We don't all see the same Internet



# Understanding Censorship

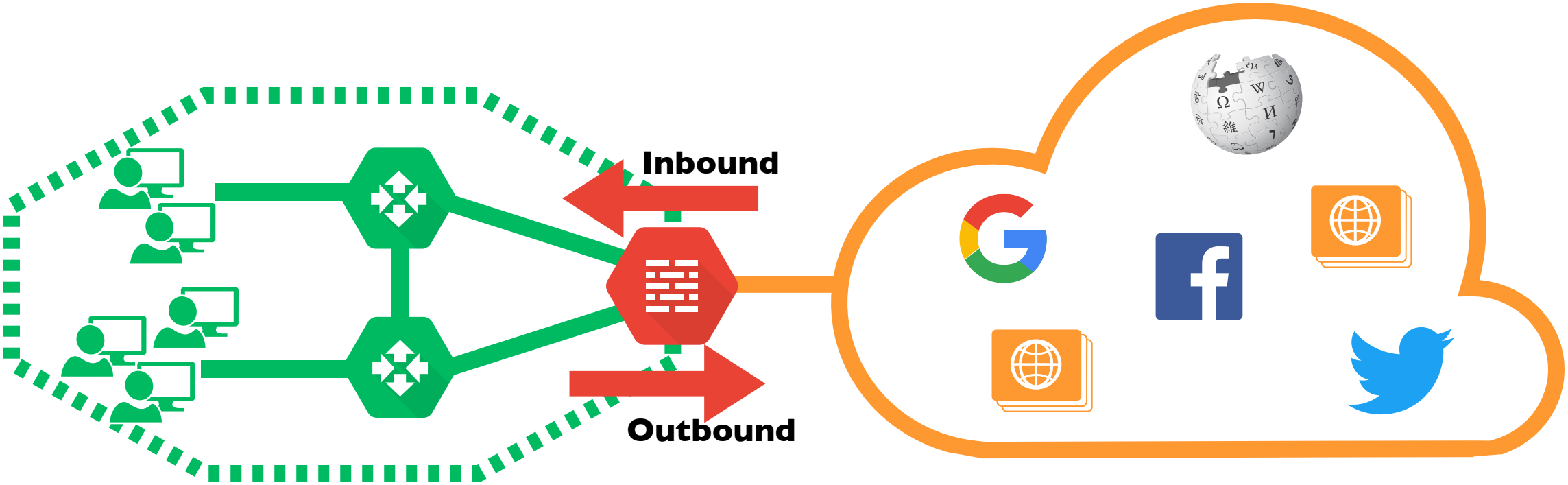
- In order to combat censorship, you need to understand:
  - What's censored
  - Who is censored
  - Where it's censored
  - How it's censored

- Challenges
  - Adversaries don't disclose any of this
  - How do you discover? Measure?
  - From where?

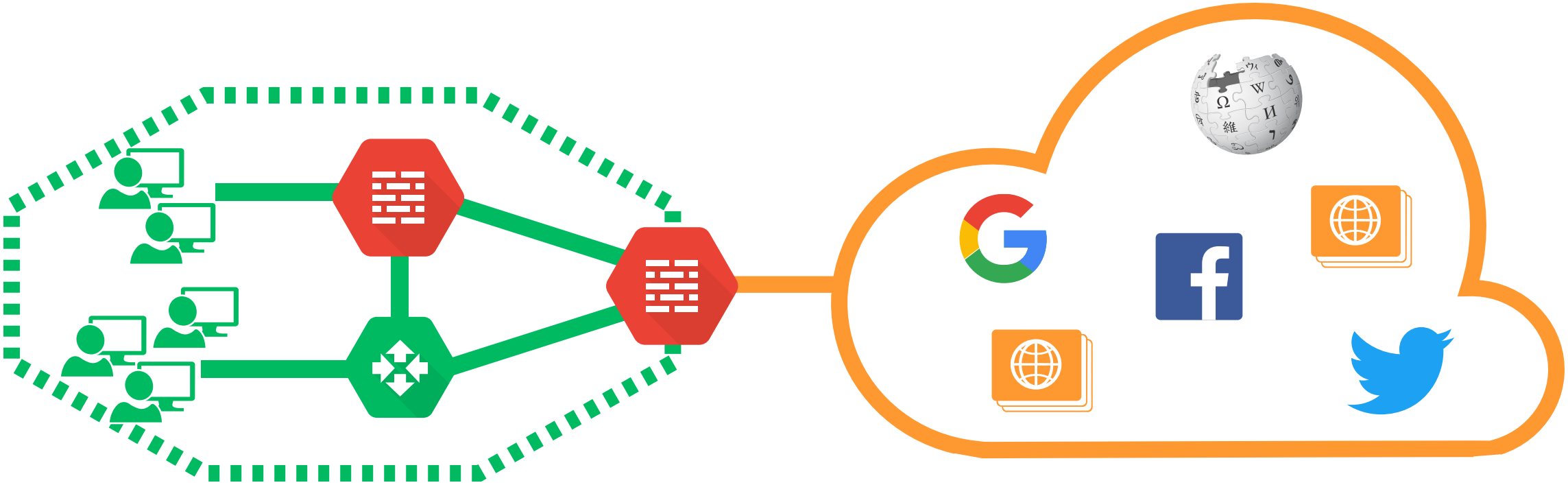
# Censorship Measurement Overview

- Goals:
  - Understand censorship behavior globally
  - Diverse viewpoints within countries
  - Enable longitudinal measurement, without volunteers
  - Remotely
- Augur
  - IPID Side Channels
  - Sequential Hypothesis Testing
- Global Measurement Study across 179 countries and territories

# How It works

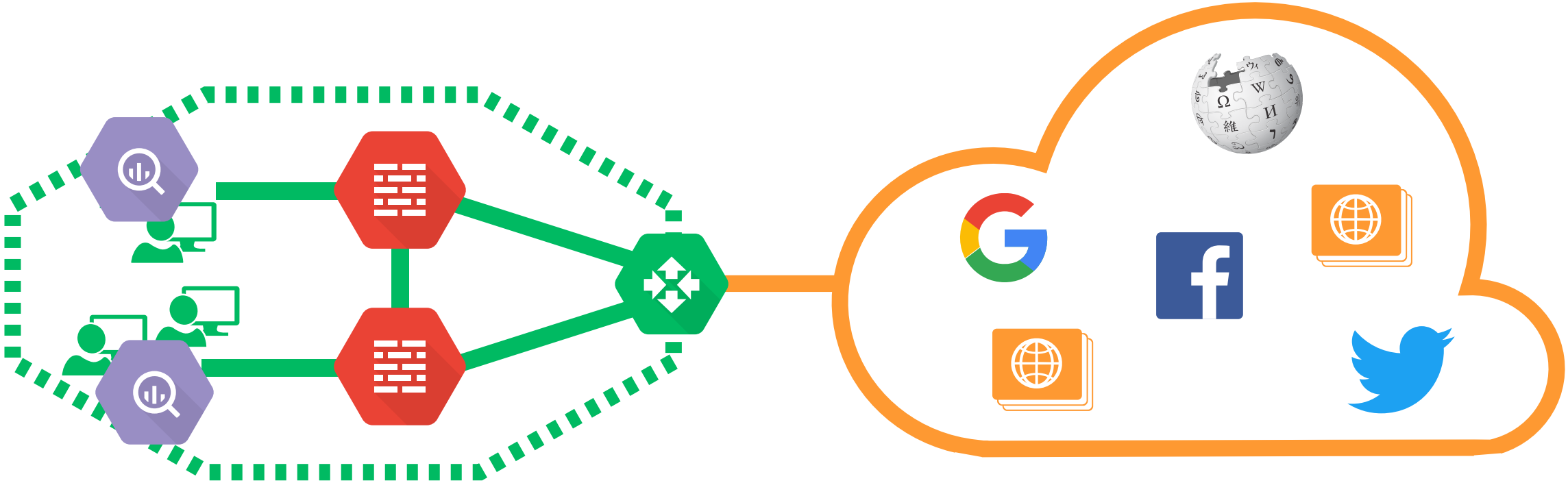


# How It works

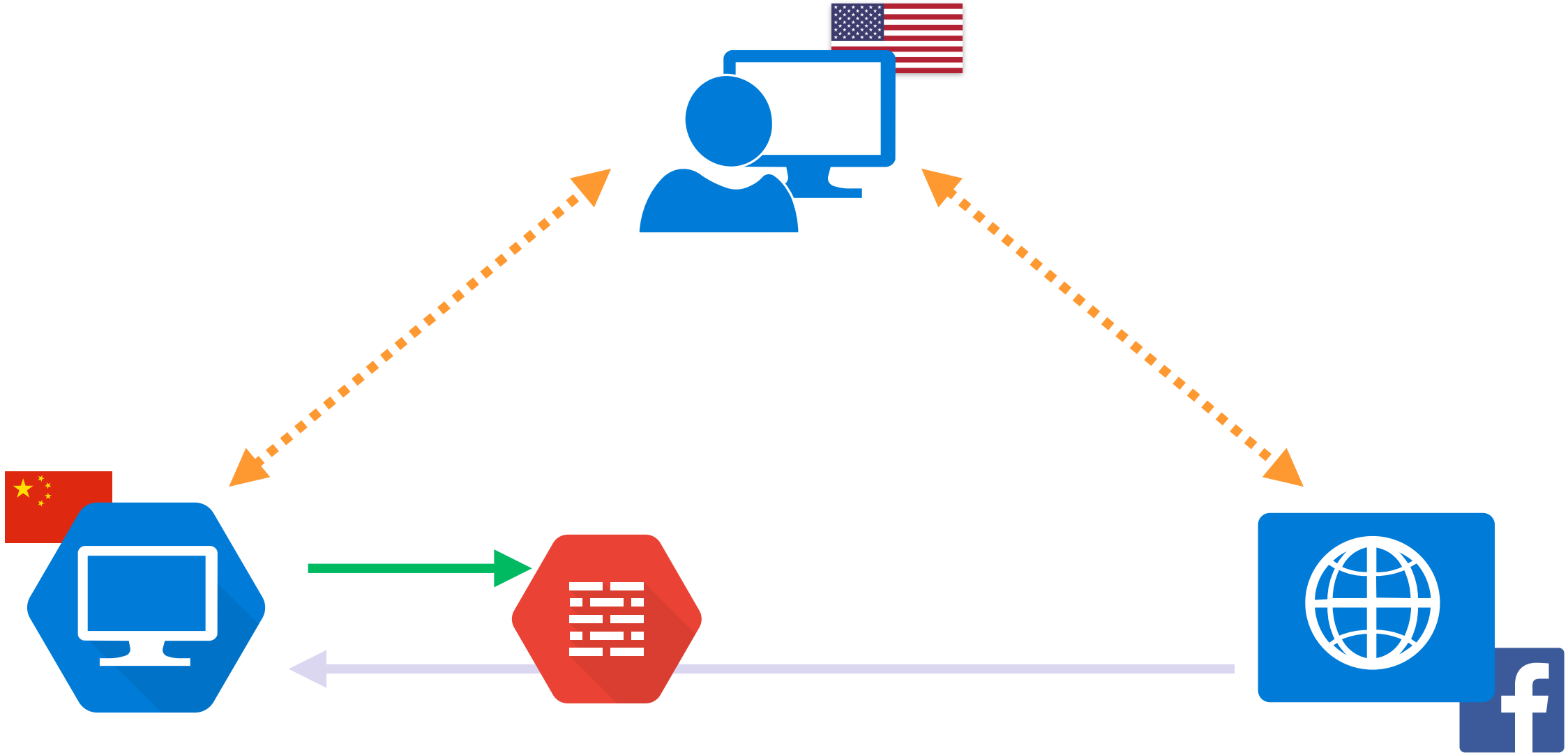


# How To Measure Censorship?

- We need to figure out what is censored, and where?
  - Well, censors won't tell you
- For a comprehensive view, you need stuff at the location



# Our Problem

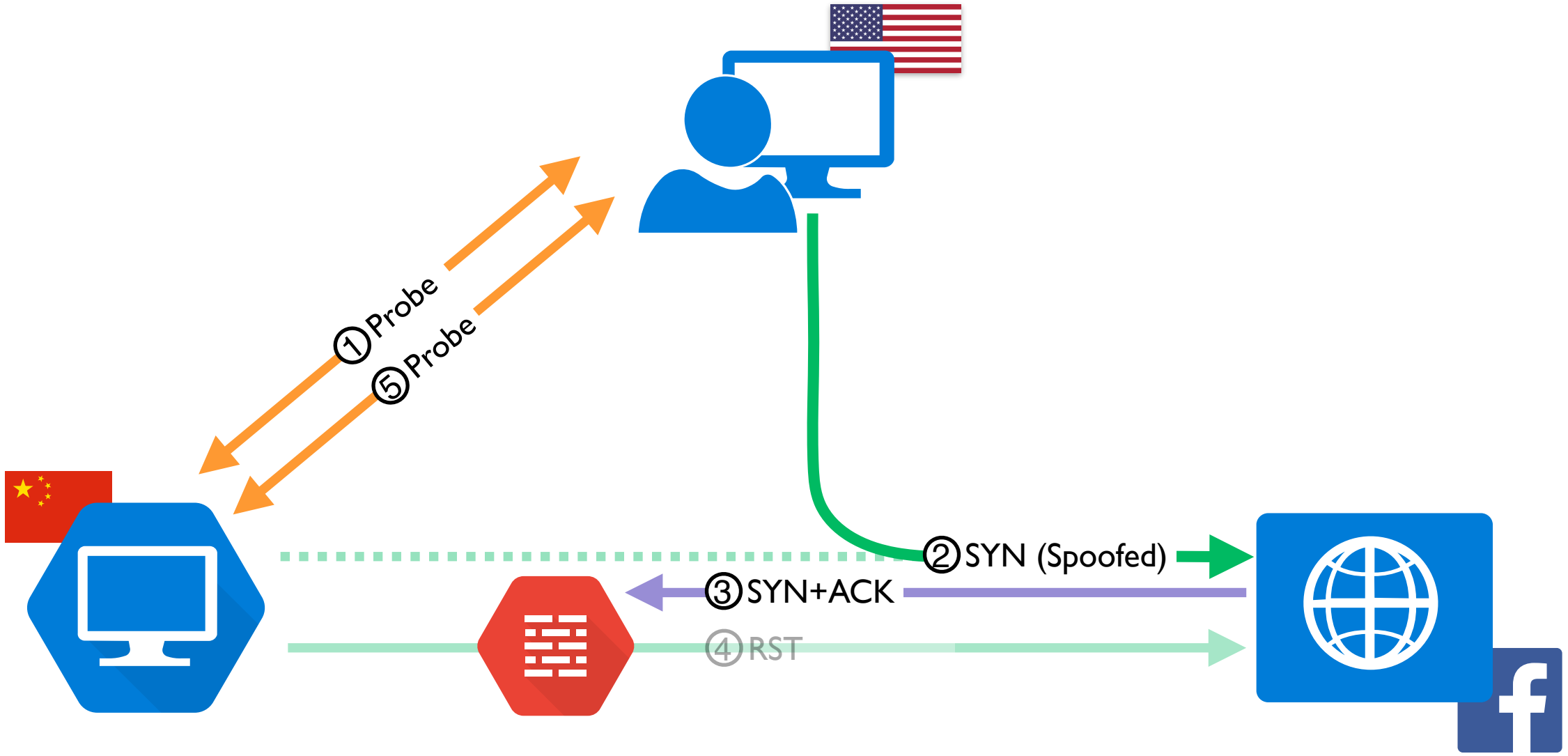


# External Measurement

- **Problem 1:**
  - We need to externally arrange for packets to be sent from FB to China
- **Solution:**
  - Spoofing

- **Problem 2:**
  - We need to externally detect if the packets made it
- **Solution:**
  - IP Identifiers + Sequential Hypothesis Testing

# Low-Level Networking + Side Channels + Stats = Censorship Measurement



# Sequential Hypothesis Testing

**Random Variable**

$$Y_n(\text{src}, \text{dst}) = \begin{cases} 0 & \text{if no IP ID acceleration} \\ 1 & \text{if IP ID acceleration} \end{cases}$$

**Hypotheses**

$$H_0 := \text{No Inbound Blocking}$$

$$H_1 := \text{Inbound Blocking}$$

**Conditional Probabilities**

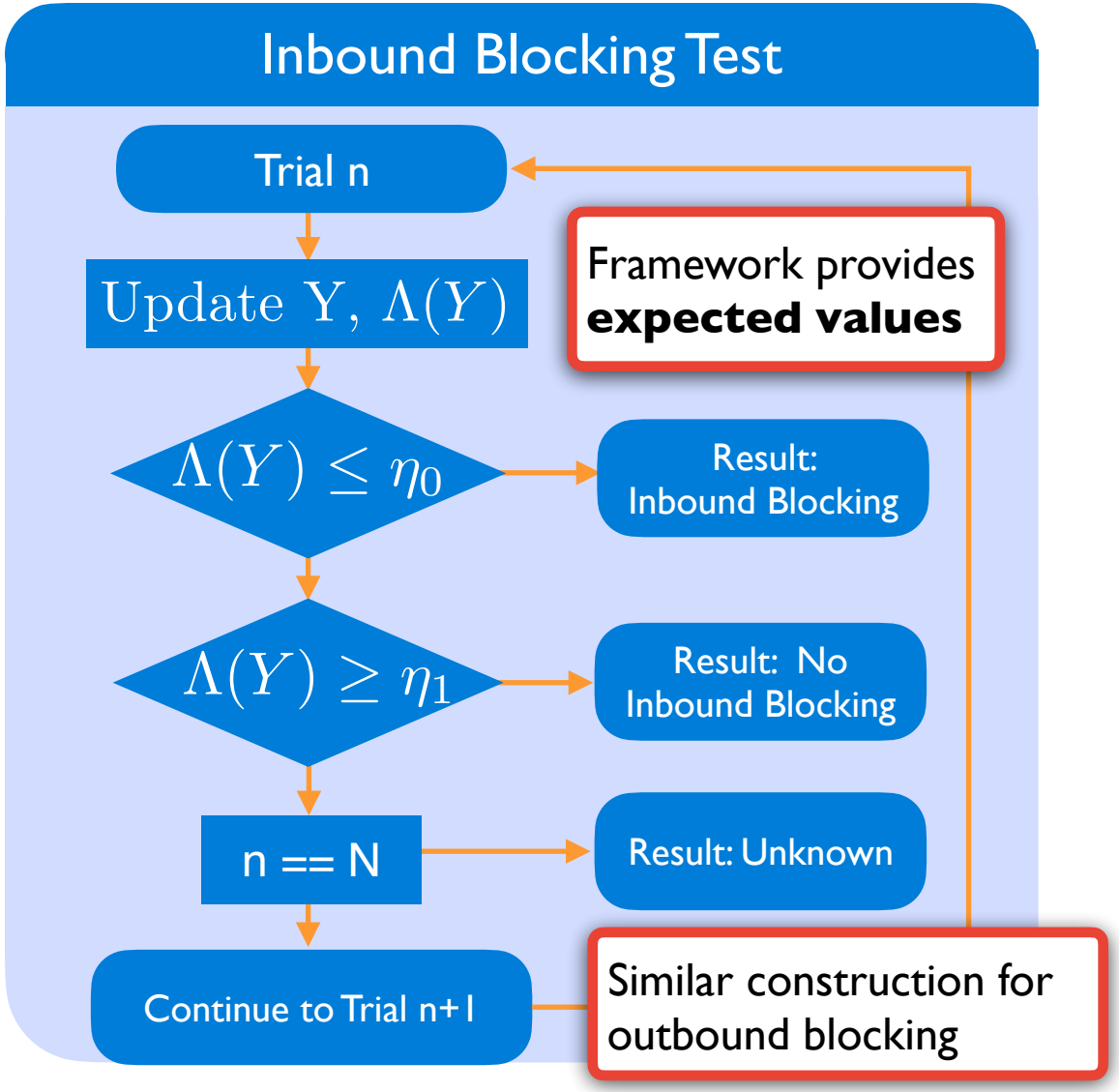
$$\Pr[Y_n = 0|H_0] = \theta_0, \Pr[Y_n = 1|H_0] = 1 - \theta_0$$ Injection

$$\Pr[Y_n = 0|H_1] = \theta_1 \quad \Pr[Y_n = 1|H_1] = 1 - \theta_1$$

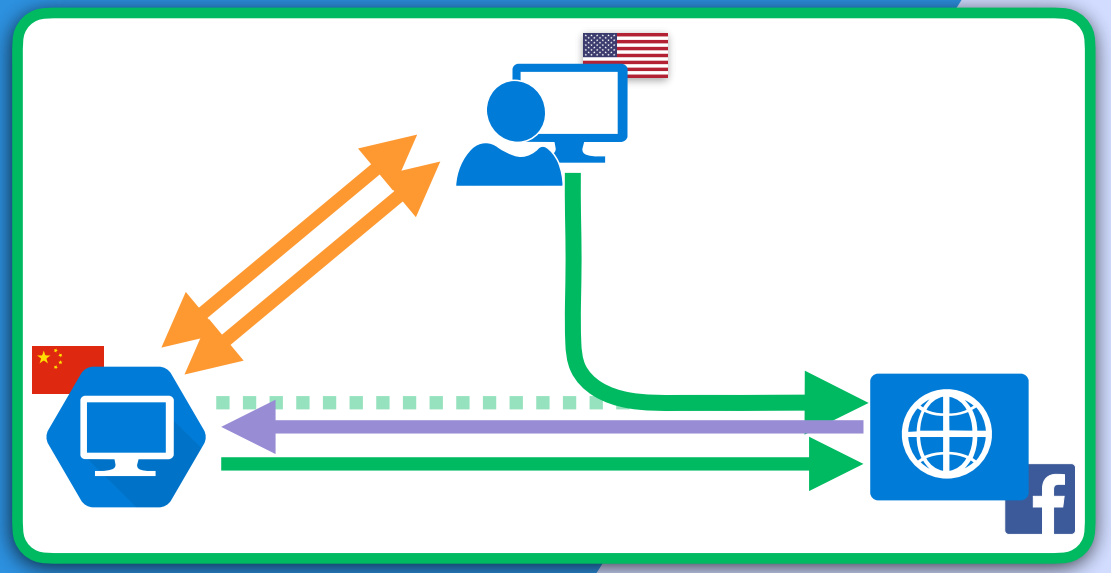
Control

**Likelihood Ratio**

$$\Lambda(Y) \equiv \frac{\Pr[Y|H_1]}{\Pr[Y|H_0]} = \Pi_{n=1}^{N} \frac{\Pr[Y_n|H_1]}{\Pr[Y_n|H_0]}$$

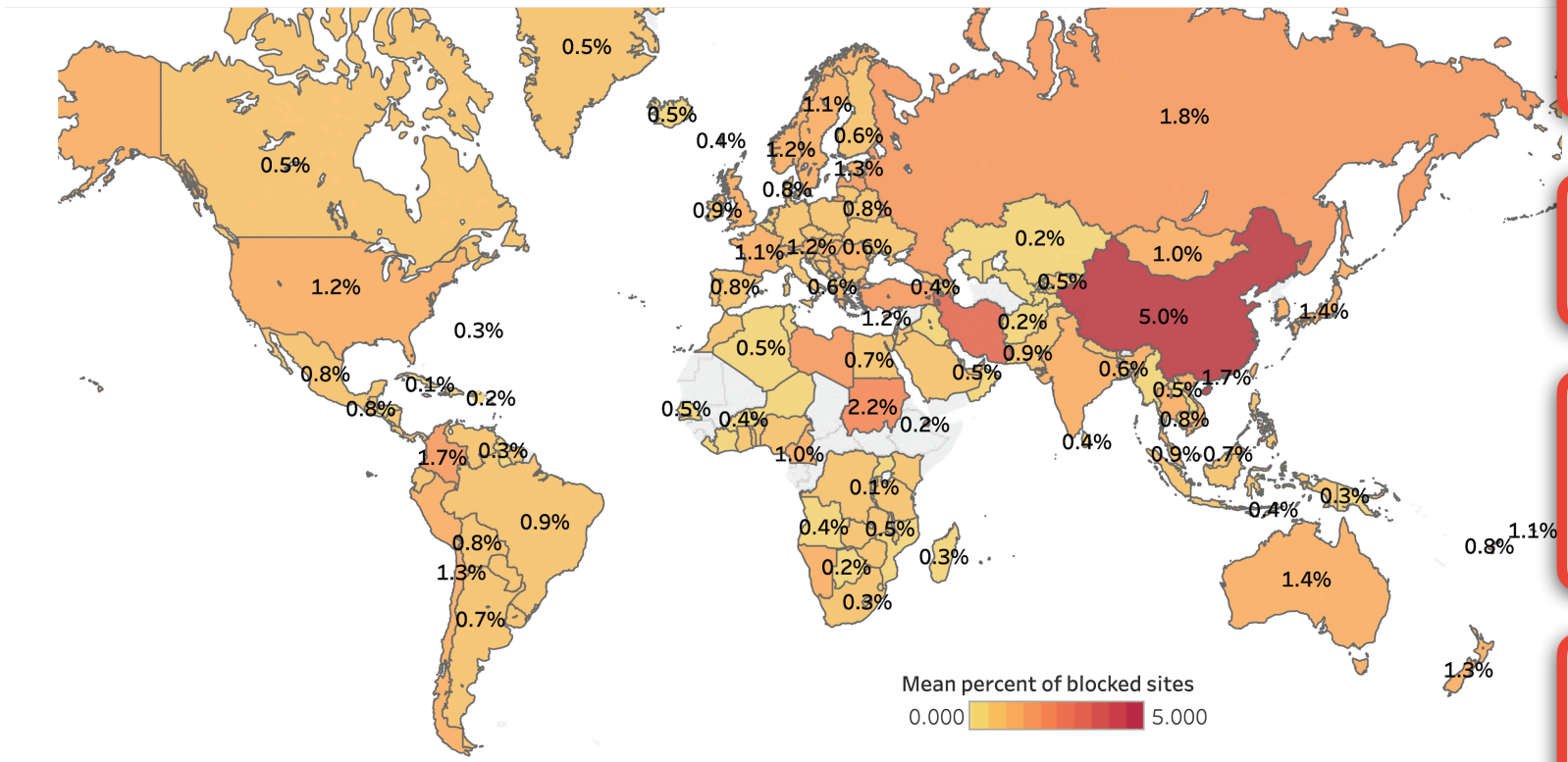


# Augur: IPID + SHT



- Remotely from an external vantage
- IP-Level Network Blocking
- With statistical confidence

# Augur: A Global Picture




- Results consistent with prior direct measurement
- Significant heterogeneity within countries (e.g., China)
- Identify differences in blocking of public and private Tor Bridges
- Detect Corporate blocking (e.g., US)

# Back to Course Logistics

# My course goals for you

- Understand the role of data-driven security research
- Develop skills to critique top-tier research
- Identify interesting questions and research topics
- Execute a research project to the level of a workshop paper
- Have an enjoyable semester learning about fun topics

# Introductions

- Me
  - Paul Pearce
  - Assistant Professor, School of Cybersecurity and Privacy, School of Computer Science
  - PhD UC Berkeley Computer Science, 2018
    - Advised by Vern Paxson
    - Worked closely with UC San Diego and Princeton
  - Spent a year as a visiting researcher Facebook
  - MS and BS also UC Berkeley
    - (Go Bears!)

# Introductions

- You
  - Your name
  - Your program
  - What do you hope to get out of this class?

# This Course

- CS8803 - EMS
  - Advanced Network Security and Measurement
- IC 105, T/Th 330pm - 445pm
  - OH: TBA, probably Tuesdays after this class
    - Starts next week. If you want to meet this week email me
    - Location TBA
- https://www.cc.gatech.edu/~pearce/courses/cs8803-ems/
- Course webpage is the syllabus
- Focuses on the intersection of network security and measurement
- Cutting edge and classical research

# Format

- **Absolutely nothing like today**
- This is meant to be a discussion driven course
- We read papers, talk about them critically
  - I guide
  - I don't talk at you
- You eventually apply what you've learned as a project

# Prerequisites

- Undergraduate Security
  - CS 4235 Undergraduate Introduction to Information Security or equivalent
- Undergraduate Networking
  - CS {3|4}251 Undergraduate Computer Networking or equivalent
- Helpful
  - CS 4237 Undergraduate Computer and Network Security or equivalent
  - Any graduate security course

# Class components and weight

- Participation (10%)
- Discussion Leads (10%)
- Paper reviews (10%)
- Course project (70%)
- Subject to change as the semester progresses
  - But with ample notice

# Paper Reviews

- 10% of your grade
- Brief paper summary of each classes paper(s)
- Submit (via email, mailing address coming soon) the summary by noon the day before each lecture
  - Starting Monday Aug 29
- 2-3 will be selected at random and evaluated by course staff
- Structure
  - What are the paper's main contributions? (3-5 sentences max)
  - What parts of the paper are questionable? (3-5 sentences max)
    - E.g., methodology, omissions, relevance, presentation, ethics.
  - What parts of the paper do you find unclear? (Optional)
- Most papers will include an additional specific 1-2 questions regarding the topic, such as challenging you to come up with and defend a proposed solution. (3-5 sentences max)

# Participation

- 10% of your grade
- Expectations:
  - You attend class regularly
  - You have read* the paper
  - You have answered the question(s)
  - You constructively participate in discussions
- **I will never cold-call anyone, it's up to you to join in**
  - Awkward silence may ensue
- Good:
  - "I didn't understand X"
  - "I thought Y was neat"
- Bad:
  - "This author is stupid"
  - "This work is pointless"
  - I never see you again after today but you appear on the roster in December

# Discussion Leads

- 10% of your grade
- Lead the discussion of 1 paper with a group (depending on enrollment) of our papers
- This begins next Tuesday
- Signups will be posted on Thursday
- Structure:
  - Assume students have read the paper and answered the questions
  - 10min at most of presentation
  - Lead a discussion similar to the review format
  - Make sure you have at least 5 specific points of discussion about the paper
  - Will be done in groups
- Model your engagement after my Thursday lecture

# Project

- 70% of your grade
  - Proposal Presentations (10%)
    - Pre Proposal and Proposal
  - Final Presentation (20%)
  - Writeup (40%)
    - Dec 15th
    - Will be evaluated as on-par with an average workshop submission
- More details on formatting on webpage and as we progress

# Project Timeline

- [Sept 27]: Project pre-proposal
  - 5min talk on you, your interests, and an idea
  - Non-binding, get the ball rolling, establish groups (~=2-3)
- [Early Oct]: Brainstorming sessions
  - Meet with me (and potentially fellow students) to chat about the idea
- [Oct 11, 13]: In-class Proposal
  - 5 min presentation (w/ your laptop) + 5 min of Q&A
- [Dec 1, Dec 6]: Presentation days
  - 10 min presentation (w/ your laptop, including Q&A, strict timing)
- [Dec 15]: Final report write-up
  - Final submission by Dec 15 9pm (strict)
- Report: no longer than 10 pages
- Demo: screenshot or video (optional)

# Tips

- Ask questions
- Come to my office hours
  - "Office hours are my most productive hours. Everyone leaves me alone" — Dr Senior Professor, PhD
  - I have to be there, might as well ask me questions
- Engage critically with the reading
  - This is an acquired skill
  - My questions and class discussion are meant to help you acquire that skill
- The project requires actual engagement, please don't wait until November 30th to begin

# Feedback

- I want to make this course better
- I want you to do well

- If you have feedback, please provide it at any time
  - Canvas anonymous feedback?

- I will arrange for a mid-semester anonymous review

# Action Items

- Read the webpage
  - The front page will have updated info and news
- Look over the schedule and start thinking about topics you want to lead discussion on
- Read the Thursday paper
  - No homework due for it
  - But read it anyway please :)

# Rest of the course

- Thursday Aug 25:
  - Questions for class 3 will go online
  - Sign-up form for discussion leads will go online

- Monday Aug 29:
  - Paper summary and questions due at noon

- Tuesday Aug 30
  - First student lead discussion
  - It's a great paper to present!

# Thank You

# Questions?

Paul Pearce
https://cc.gatech.edu/~pearce/