Can You Hide in an Internet Panopticon?

Bryan Ford – Yale University

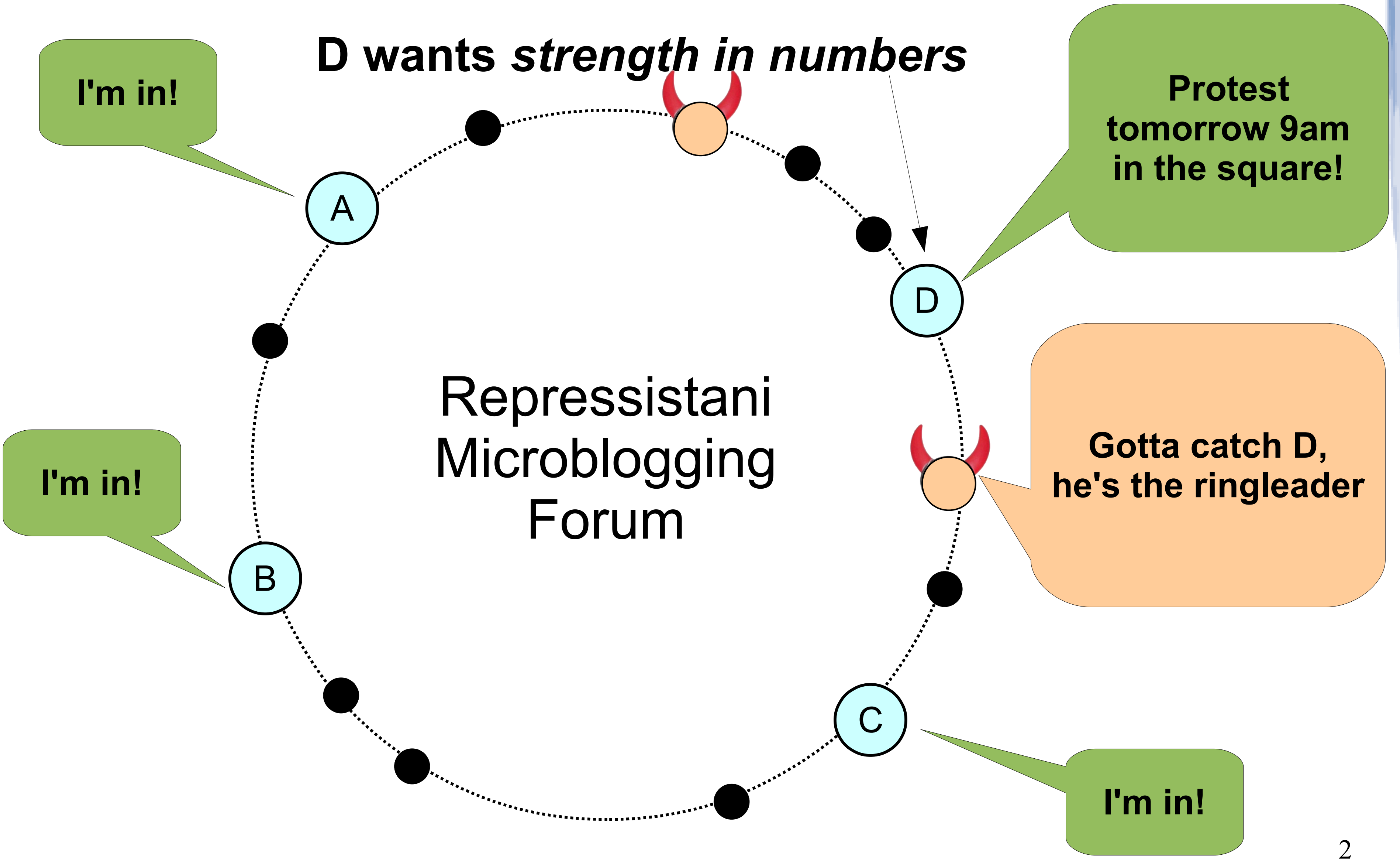
working with:
David Isaac Wolinsky, Joan Feigenbaum,
Henry Corrigan-Gibbs, Ewa Syta, John Maheswaran,
Ramakrishna Gummadi **– Yale**

Vitaly Shmatikov, Amir Houmansadr,
Chad Brubaker **– UT Austin**

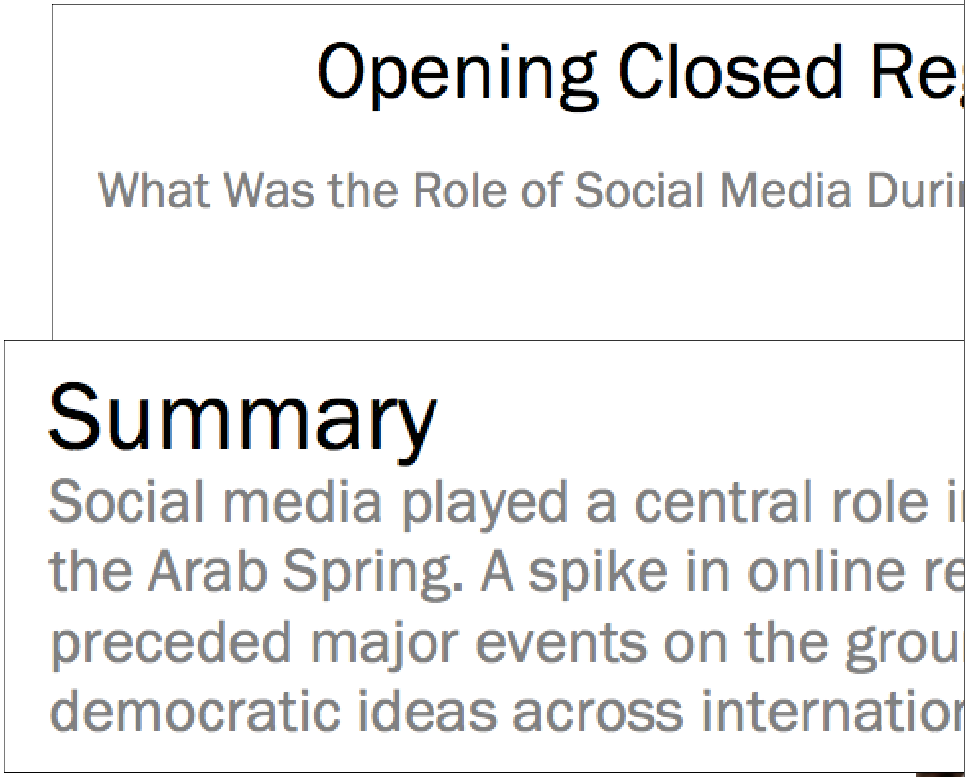
Aaron Johnson **– US Naval Research Lab**

University of Texas at Austin – Oct 24, 2013

A Dissident in Repressistan


D wants *strength in numbers*
I'm in!
A
Protest tomorrow 9am in the square!
D
Repressistani Microblogging Forum
Gotta catch D, he's the ringleader
I'm in!
B
C
I'm in!

Real Situations



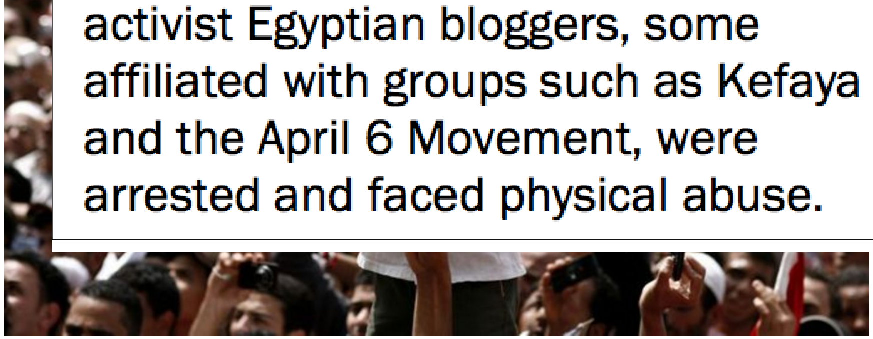
Opening Closed Re

What Was the Role of Social Media Duri

Summary

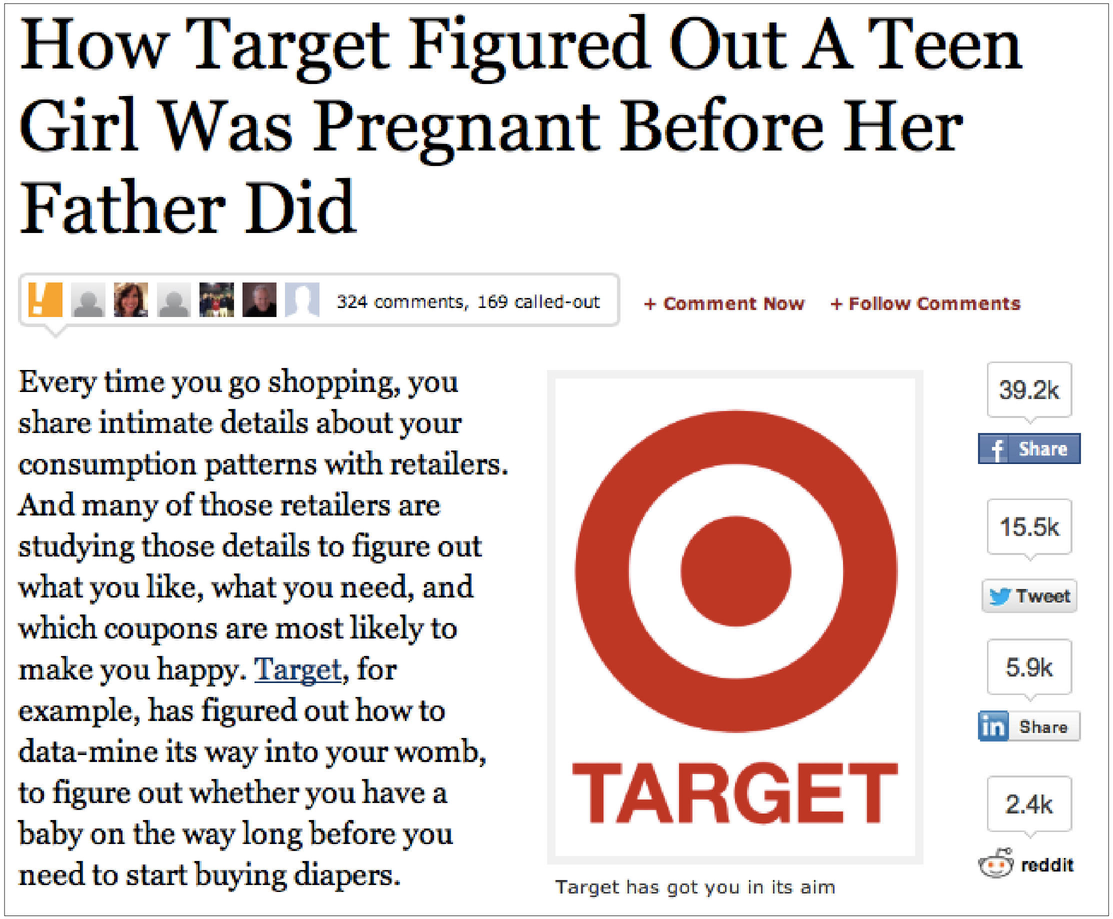
Social media played a central role i the Arab Spring. A spike in online re preceded major events on the grou democratic ideas across internatio

In response, the governments in Tunisia and Egypt arrested bloggers, tracked online conversations, and shuttered Websites and Internet access. For example, in 2005 Egyptian blogger Abdolkarim Nabil Seliman was arrested and imprisoned for four years after criticizing President Hosni Mubarak and the state's religious institutions. In 2007, a number of bloggers were arrested for organizing and covering social protests when the Egyptian parliament approved controversial constitutional amendments. Many activist Egyptian bloggers, some affiliated with groups such as Kefaya and the April 6 Movement, were arrested and faced physical abuse.



Who Wants to Track **You** Online?

- Advertisers (if you ever spend money)



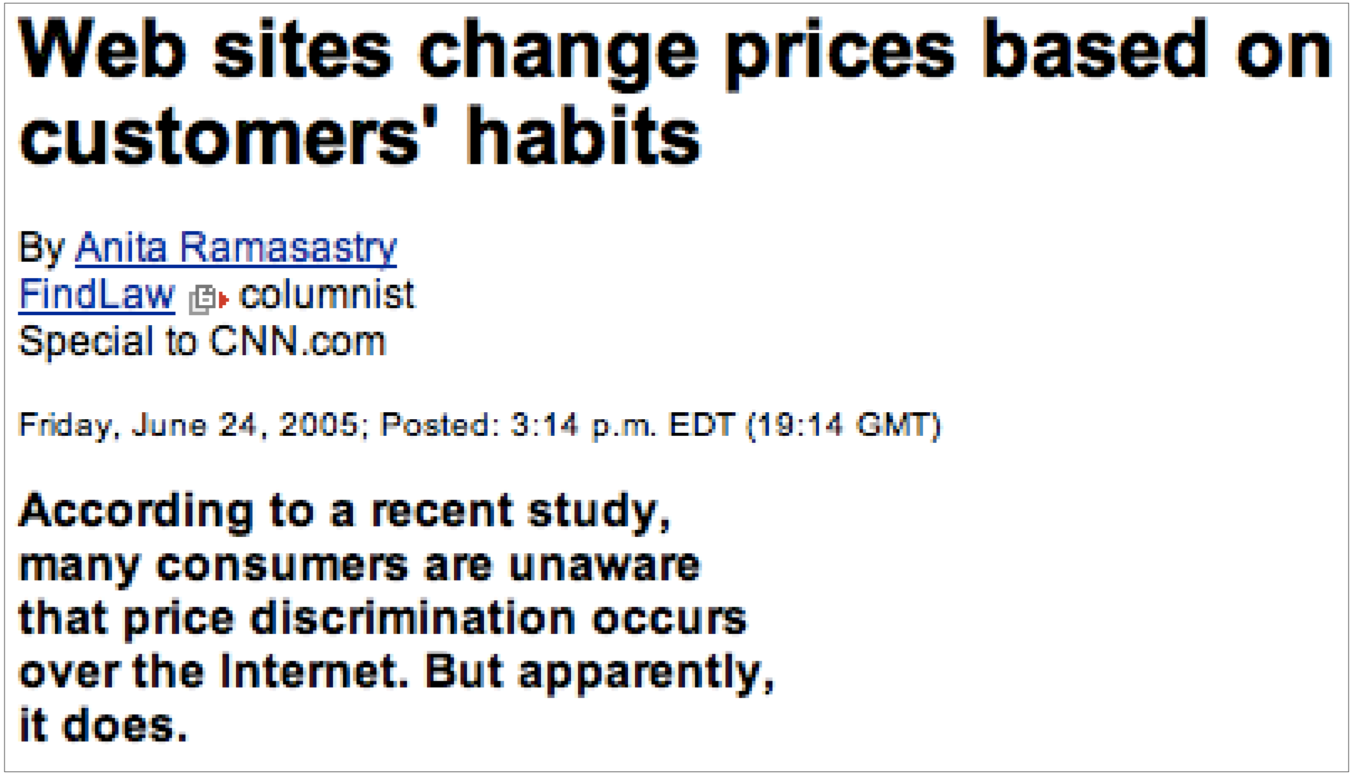
How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did

Every time you go shopping, you share intimate details about your consumption patterns with retailers. And many of those retailers are studying those details to figure out what you like, what you need, and which coupons are most likely to make you happy. Target, for example, has figured out how to data-mine its way into your womb, to figure out whether you have a baby on the way long before you need to start buying diapers.

Target has got you in its aim

Who Wants to Track **You** Online?

- Advertisers (if you ever spend money)
- Vendors (if you ever buy things)



Web sites change prices based on customers' habits

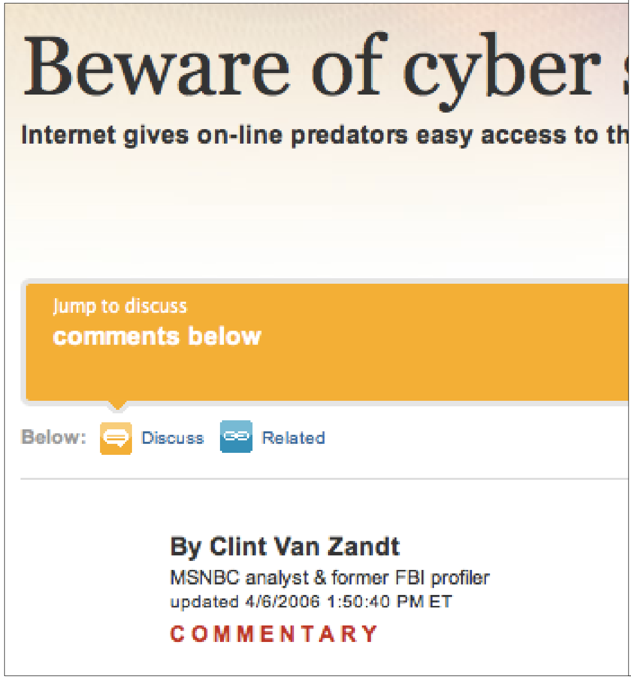
By Anita Ramasastry
FindLaw columnist
Special to CNN.com

Friday, June 24, 2005; Posted: 3:14 p.m. EDT (19:14 GMT)

According to a recent study, many consumers are unaware that price discrimination occurs over the Internet. But apparently, it does.

Who Wants to Track **You** Online?

- Advertisers (if you
- Vendors (if you ev
- Stalkers (if you're



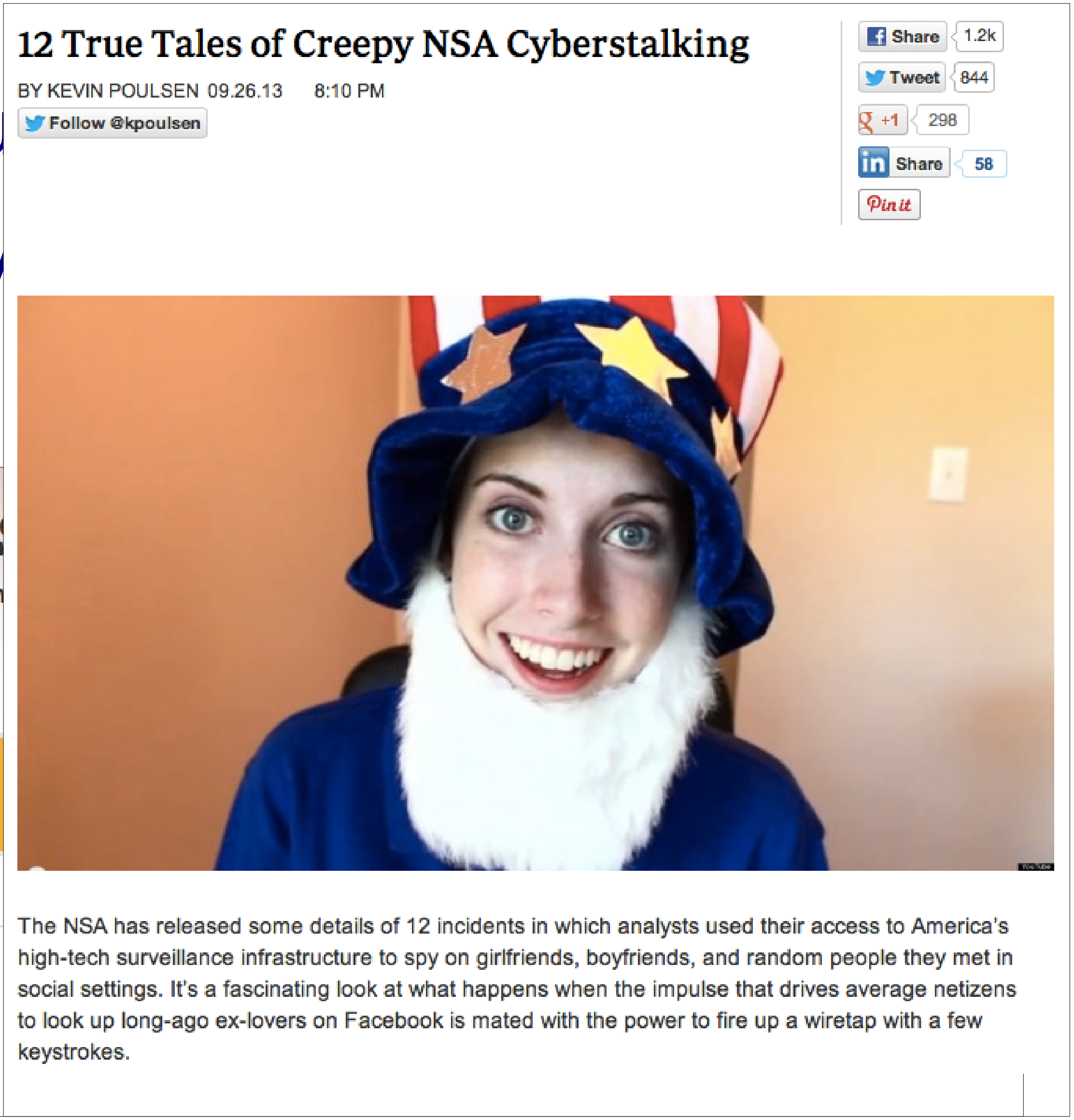
Beware of cyber s

Internet gives on-line predators easy access to th

Jump to discuss
comments below

By Clint Van Zandt
MSNBC analyst & former FBI profiler
updated 4/6/2006 1:50:40 PM ET

COMMENTARY



12 True Tales of Creepy NSA Cyberstalking

BY KEVIN POULSEN 09.26.13 8:10 PM

The NSA has released some details of 12 incidents in which analysts used their access to America's high-tech surveillance infrastructure to spy on girlfriends, boyfriends, and random people they met in social settings. It's a fascinating look at what happens when the impulse that drives average netizens to look up long-ago ex-lovers on Facebook is mated with the power to fire up a wiretap with a few keystrokes.

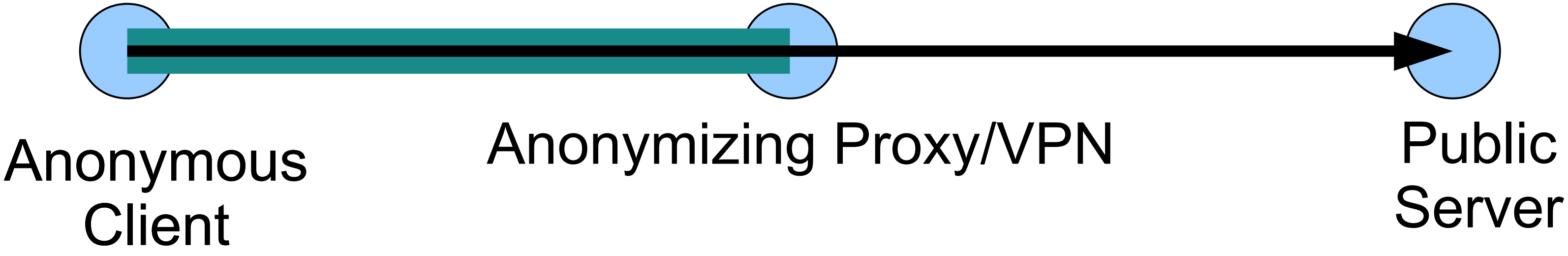
Who Wants to Track **You** Online?

- Advertisers (if you ever spend money)
- Vendors (if you ever buy things)
- Stalkers (if you're a domestic abuse victim)
- Competitors (if you're a business)
- Extremists (if you're minority/gay/pro-choice...)
- The Police (if you're “of interest” w/in 3 hops)
- The Mob (if you're the police)
- …

How Can You Protect Yourself?

Weak defenses:

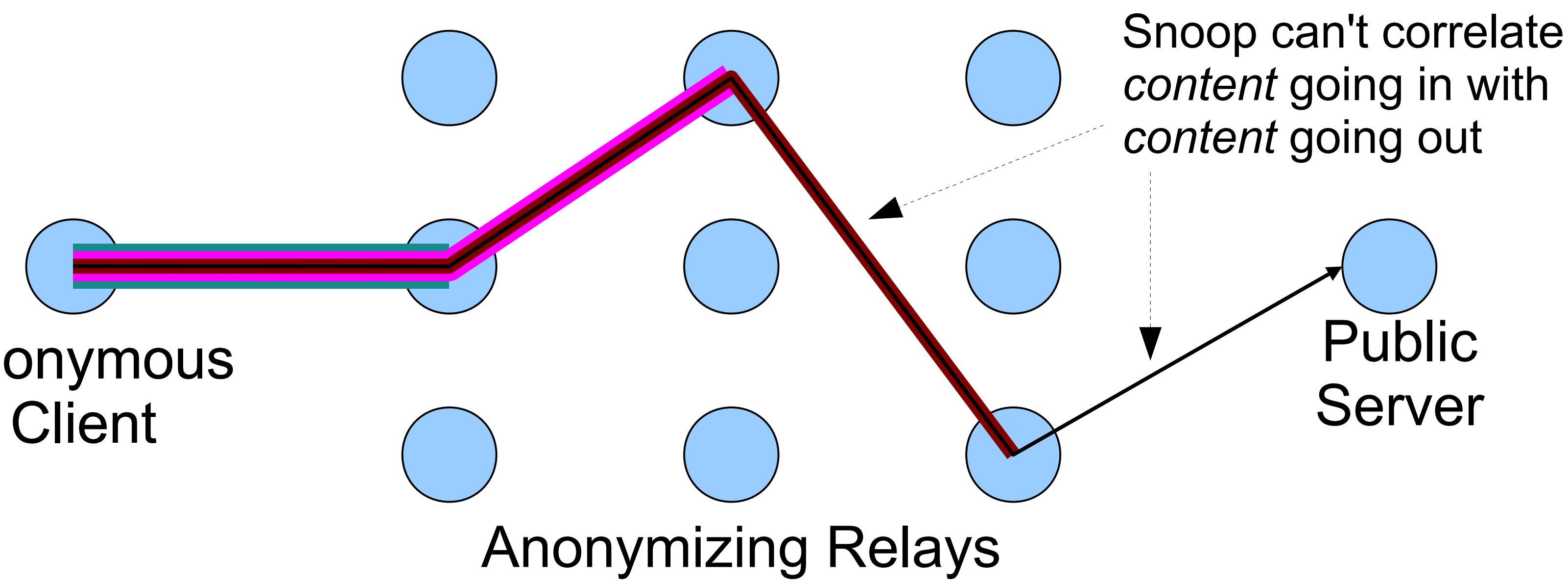
- Disable cookies, browser history, Flash, Java
- “Do-Not-Track” (pretty please) flag
- Hide behind NATs, firewalls, corporate VPNs
- Centralized commercial proxy/VPN services



How Can You Protect Yourself?

Much better defense:
state-of-the-art tools such as **Tor**

- **https://www.torproject.org**



Snoop can't correlate *content* going in with *content* going out

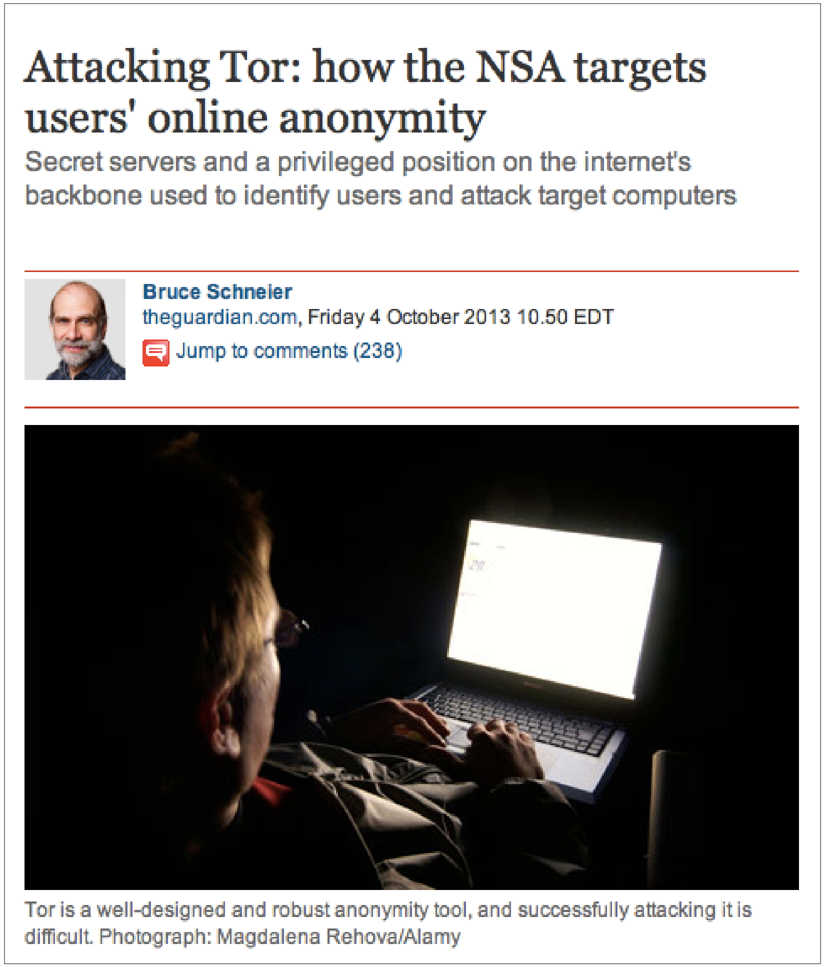
Anonymous Client

Public Server

Anonymizing Relays

The Current State-of-the-Art

- Good News: Tor probably “isn't broken yet”

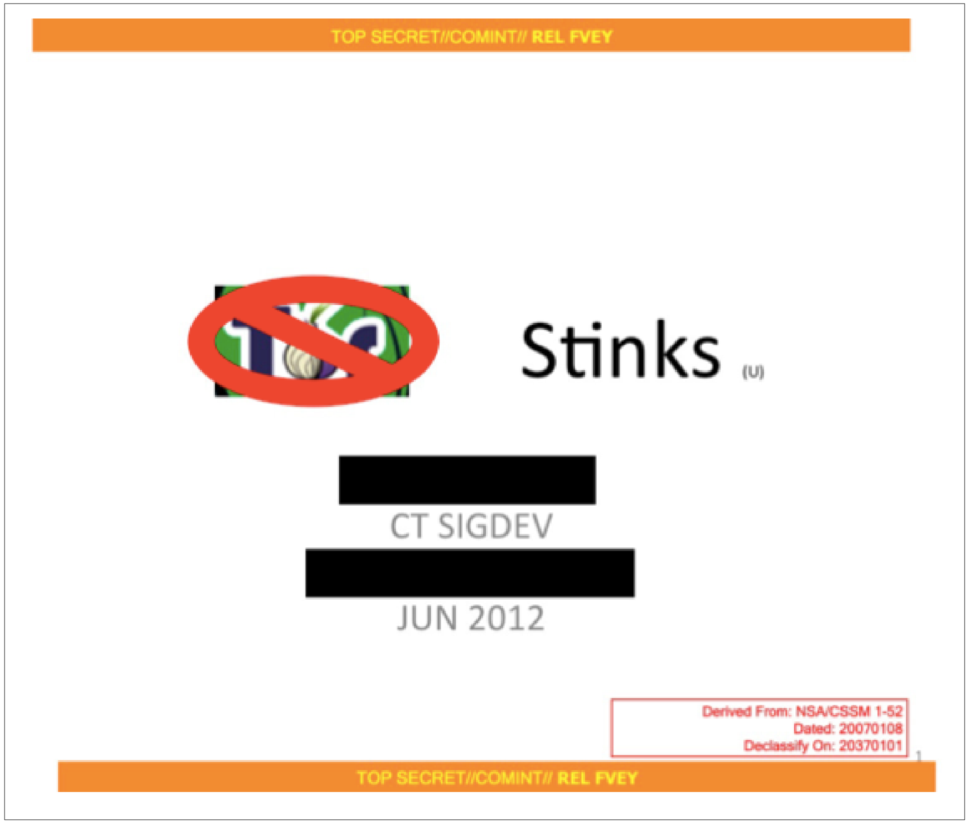


Attacking Tor: how the NSA targets users' online anonymity

Secret servers and a privileged position on the internet's backbone used to identify users and attack target computers

Bruce Schneier
theguardian.com, Friday 4 October 2013 10.50 EDT
Jump to comments (238)

Tor is a well-designed and robust anonymity tool, and successfully attacking it is difficult. Photograph: Magdalena Rehova/Alamy



TOP SECRET//COMINT// **REL FVEY**

Stinks (U)

[redacted]

CT SIGDEV

[redacted]

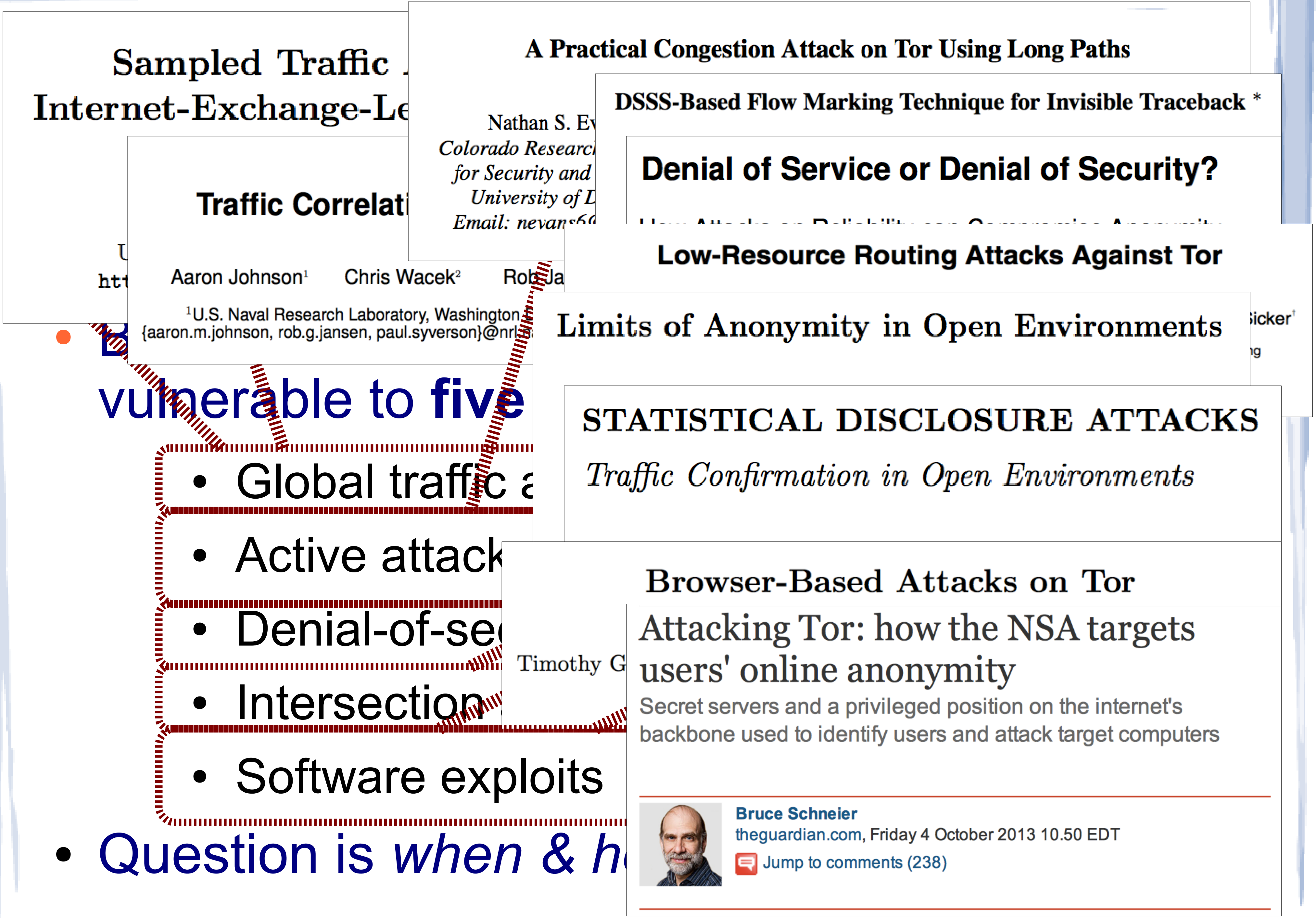
JUN 2012

Derived From: NSA/CSSM 1-52
Dated: 20070108
Declassify On: 20370101

1

TOP SECRET//COMINT// **REL FVEY**

- B
vulnerable to **five**
  - Global traffic a
  - Active attack
  - Denial-of-se
  - Intersection
  - Software exploits
- Question is *when & h*



Sampled Traffic
Internet-Exchange-Le
U
htt

Traffic Correlati

Aaron Johnson[1] Chris Wacek[2] Rob Ja

[1]U.S. Naval Research Laboratory, Washington
{aaron.m.johnson, rob.g.jansen, paul.syverson}@nrl

A Practical Congestion Attack on Tor Using Long Paths

Nathan S. Ev
Colorado Researc
for Security and
University of D
Email: nevans6

DSSS-Based Flow Marking Technique for Invisible Traceback *

Denial of Service or Denial of Security?

Low-Resource Routing Attacks Against Tor

icker†
ng

Limits of Anonymity in Open Environments

STATISTICAL DISCLOSURE ATTACKS

Traffic Confirmation in Open Environments

Browser-Based Attacks on Tor

Timothy G

Attacking Tor: how the NSA targets users' online anonymity

Secret servers and a privileged position on the internet's backbone used to identify users and attack target computers

Bruce Schneier
theguardian.com, Friday 4 October 2013 10.50 EDT
Jump to comments (238)

The Dissent Project

Goal: rethink the foundations of anonymity

- Offer *quantifiable* and *measurable* anonymity
- Build on primitives offering *provable security*
- Don't just *patch* specific vulnerabilities, but *rearchitect* to address whole *attack classes*

http://dedis.cs.yale.edu/dissent/

Dissent's Contribution

Does not, and *may never* yield
“drop-in replacement” for onion routing

– but –

First anonymity system offering *some*
(imperfect, incomplete, but...)
systematic defense against
all five classes of vulnerabilities

Talk Outline

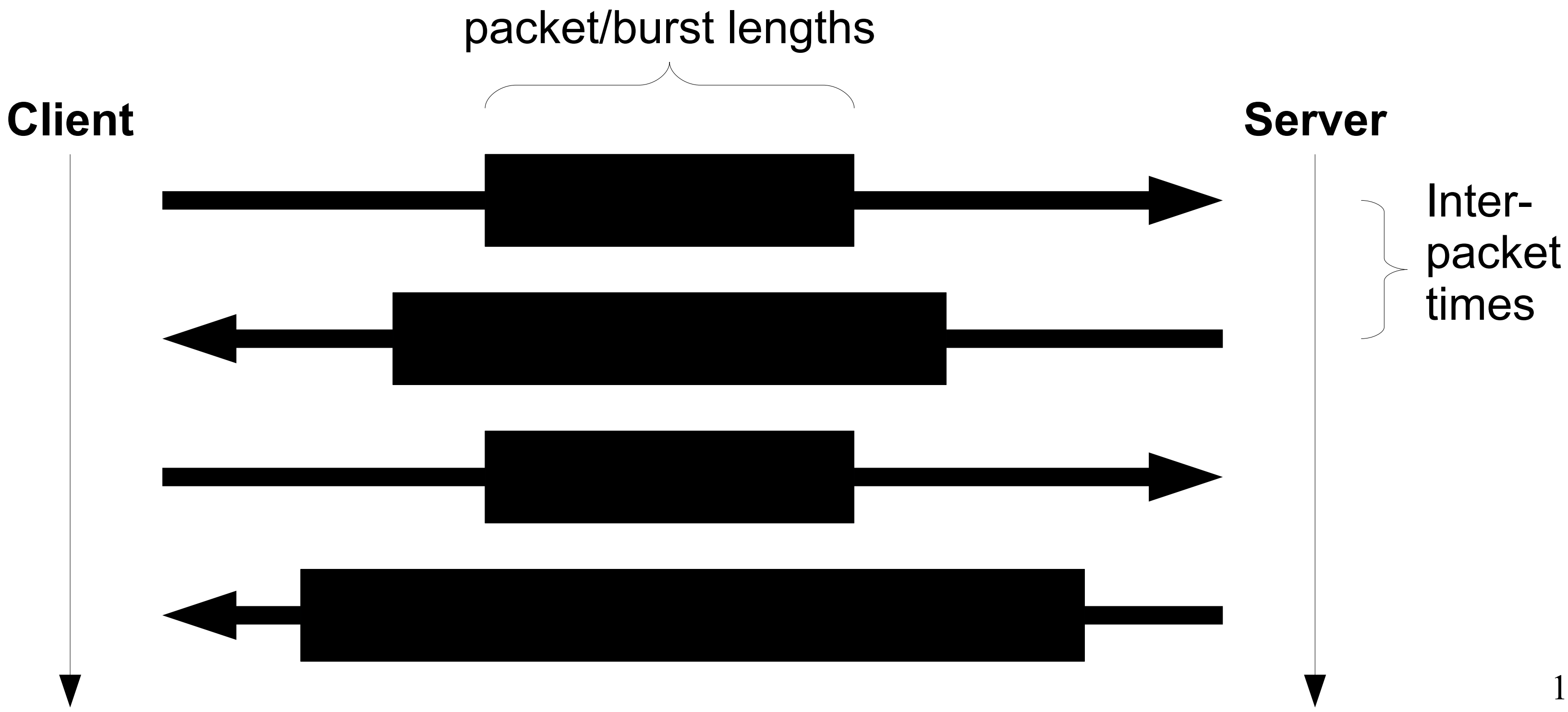
- ✓ Anonymity: Motivation and Background
- *Dissent*, and How It Resists **Strong Attacks**
  - *DC-nets* and *shuffles* resist **global traffic analysis**
  - *Collective control plane* resists **active attacks**
  - *Accountability* resists d**enial-of-security (DoSec)**
  - *Metrics* and *buddies* resist **intersection attacks**
  - *Pseudonym VMs* resist **de-anonymizing exploits**
- Dissent Status: Where We Are, and Aren't
- Conclusion

Talk Outline

- Anonymity: Motivation and Background
- *Dissent*, and How It Resists Strong Attacks
  - ➔ *DC-nets* and *shuffles* resist global traffic analysis
  - *Collective control plane* resists active attacks
  - *Accountability* resists denial-of-security (DoSec)
  - *Metrics* and *buddies* resist intersection attacks
  - *Pseudonym VMs* resist de-anonymizing exploits
- Dissent Status: Where We Are, and Aren't
- Conclusion

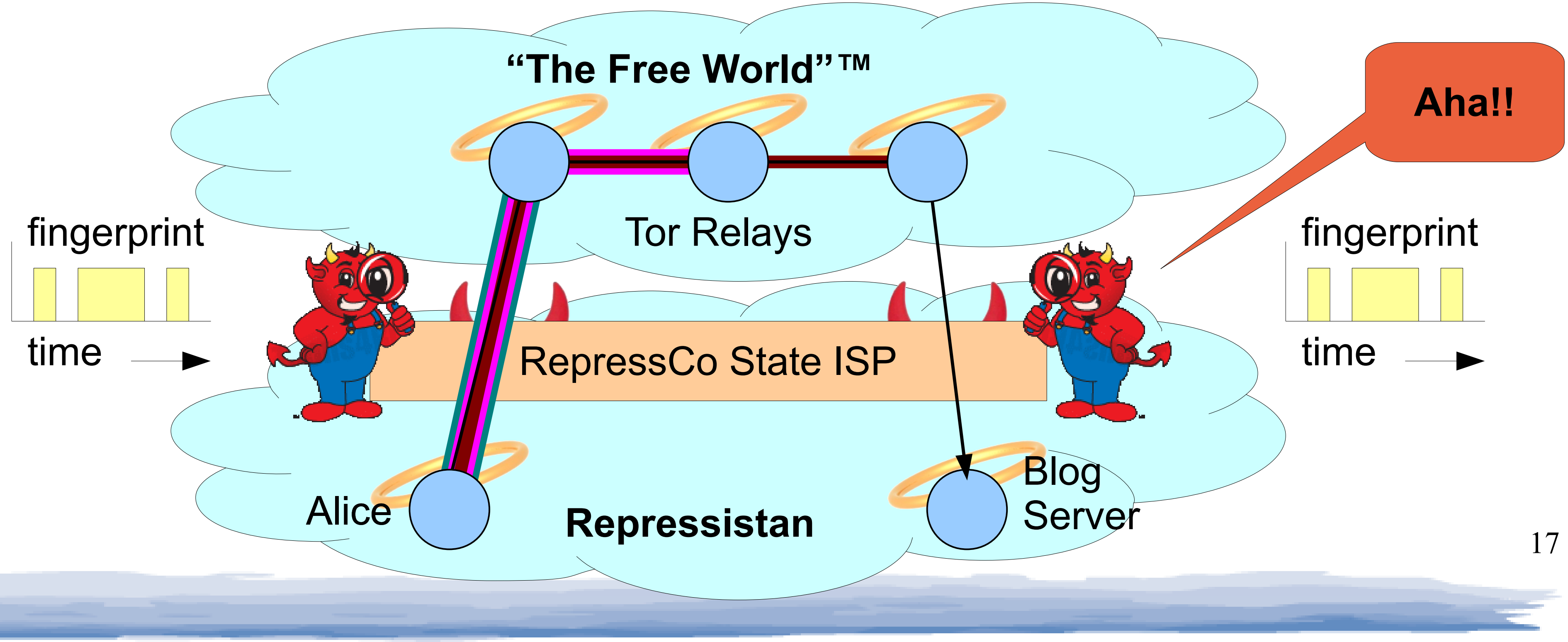
Traffic Analysis Basics

- Most communication has a *traffic pattern*
  - Lengths and timings of packets in each direction
  - Pattern can be *fingerprinted* without seeing content



Tor Traffic Analysis Scenario

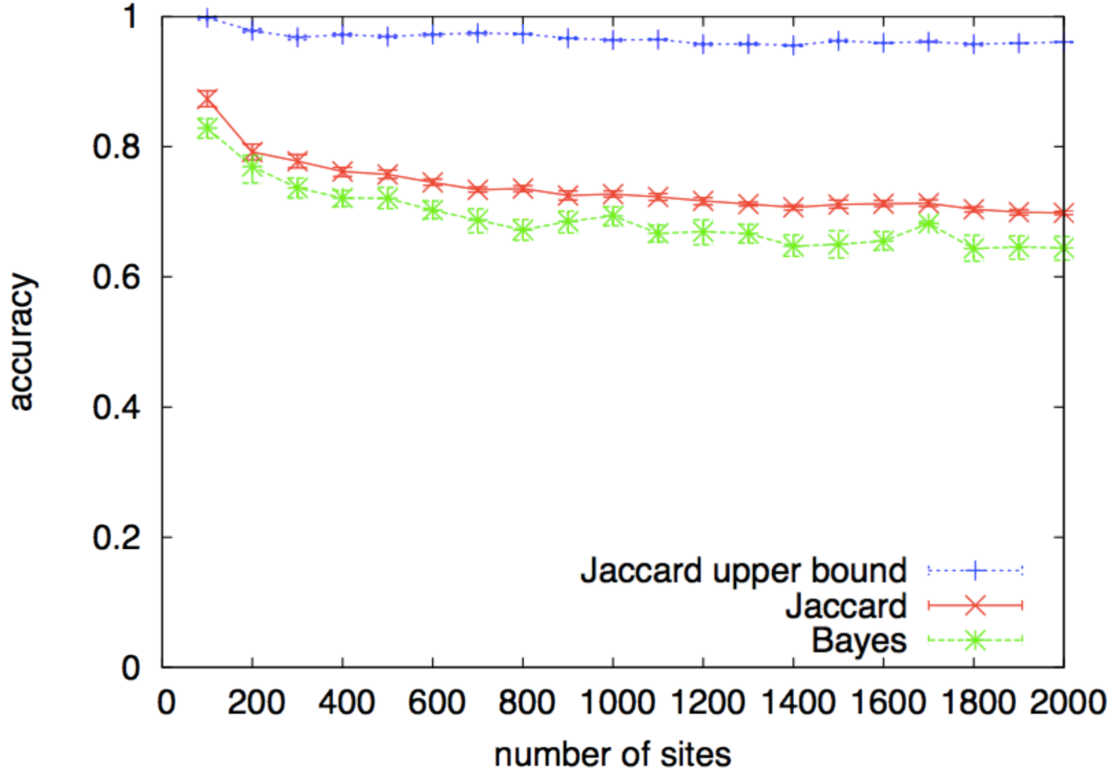
- Alice in Repressistan uses Tor to post on blog server hosted in Repressistan
- State ISP controls *both* entry and exit hops
- Fingerprint & correlate traffic to **deanonymize**



Is Traffic Fingerprinting Practical?

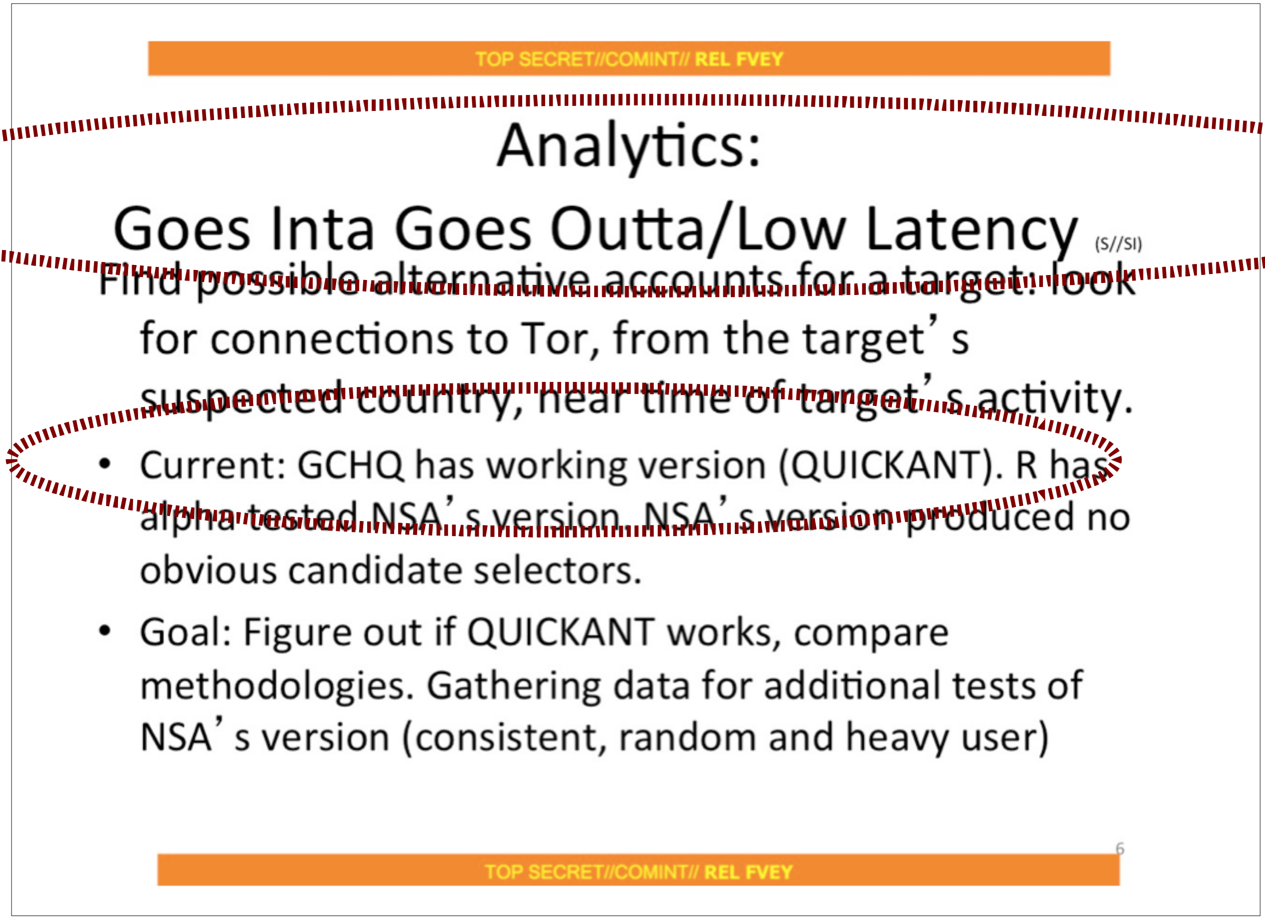
General techniques well-known, scalable

- "Inferring the Source of Encrypted HTTP connections" Liberatore and Levine, CCS '06



Do Attackers Actually *Do* This?

Not sure, but some are *working hard on it...*



TOP SECRET//COMINT// REL FVEY

Analytics:

Goes Inta Goes Outta/Low Latency (S//SI)

Find possible alternative accounts for a target: look for connections to Tor, from the target's suspected country, near time of target's activity.

- Current: GCHQ has working version (QUICKANT). R has alpha-tested NSA's version. NSA's version produced no obvious candidate selectors.
- Goal: Figure out if QUICKANT works, compare methodologies. Gathering data for additional tests of NSA's version (consistent, random and heavy user)

6

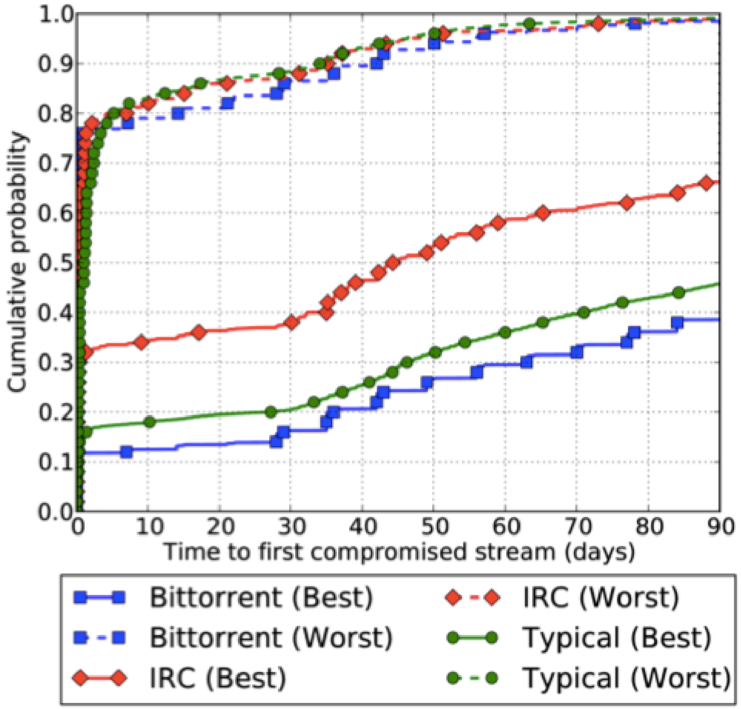
TOP SECRET//COMINT// REL FVEY

("Tor Stinks" slide deck, Guardian 10/4/2013)

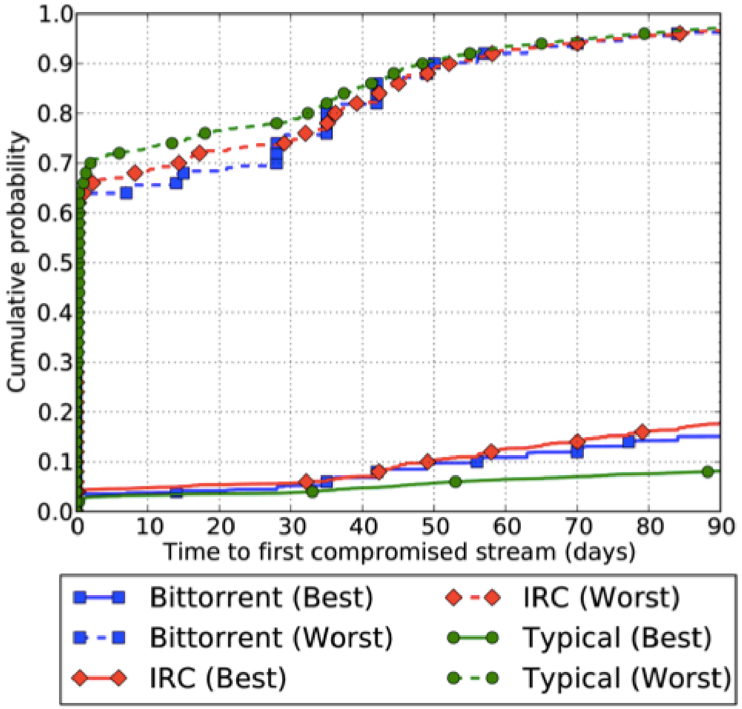
Can De-Anonymize "Real" Users?

Yes, if attacker can monitor an Internet AS or IXP

- "Users Get Routed", Johnson et al. CCS 13



(a) Time to first stream compromised by AS adversary.



(b) Time to first stream compromised by IXP adversary.

How To Resist Traffic Analysis?

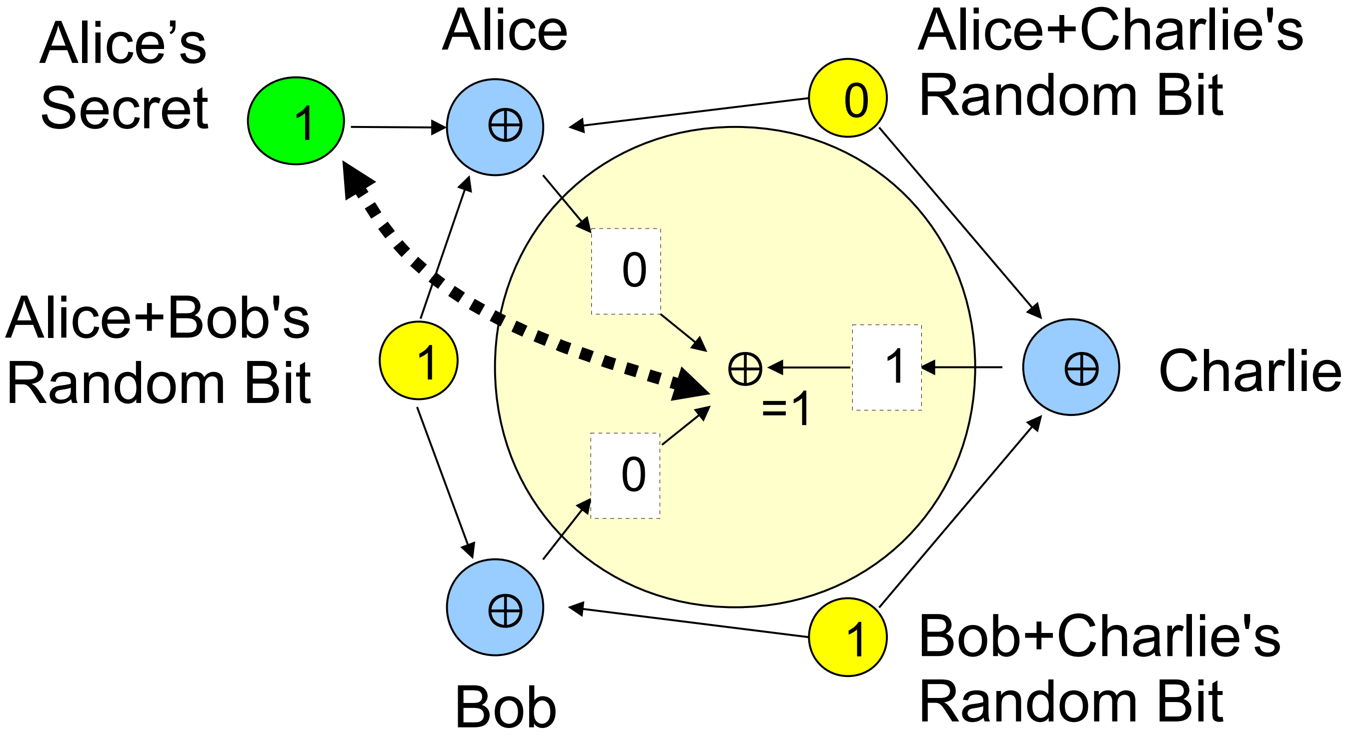
- **Option 1:** “Pad” traffic to uniform rate
  - **Aqua**, Le Blond et al., SIGCOMM 13
  - Works against *passive* attacks, at bandwidth cost
  - Usually fails against *active* attacks

- **Option 2:** Fundamentally different primitive
  - Dining Cryptographers (DC-nets) – Chaum, 88
  - **Herbivore**, Sirer, SIGOPS EW 04
  - **Dissent**, CCS 10, OSDI 12, USENIX Sec 13

Dining Cryptographers (DC-nets)

Another fundamental Chaum invention from the 80s...

- Example: anonymity in a 3-member group



Dining Cryptographers (DC-nets)

Attractive:

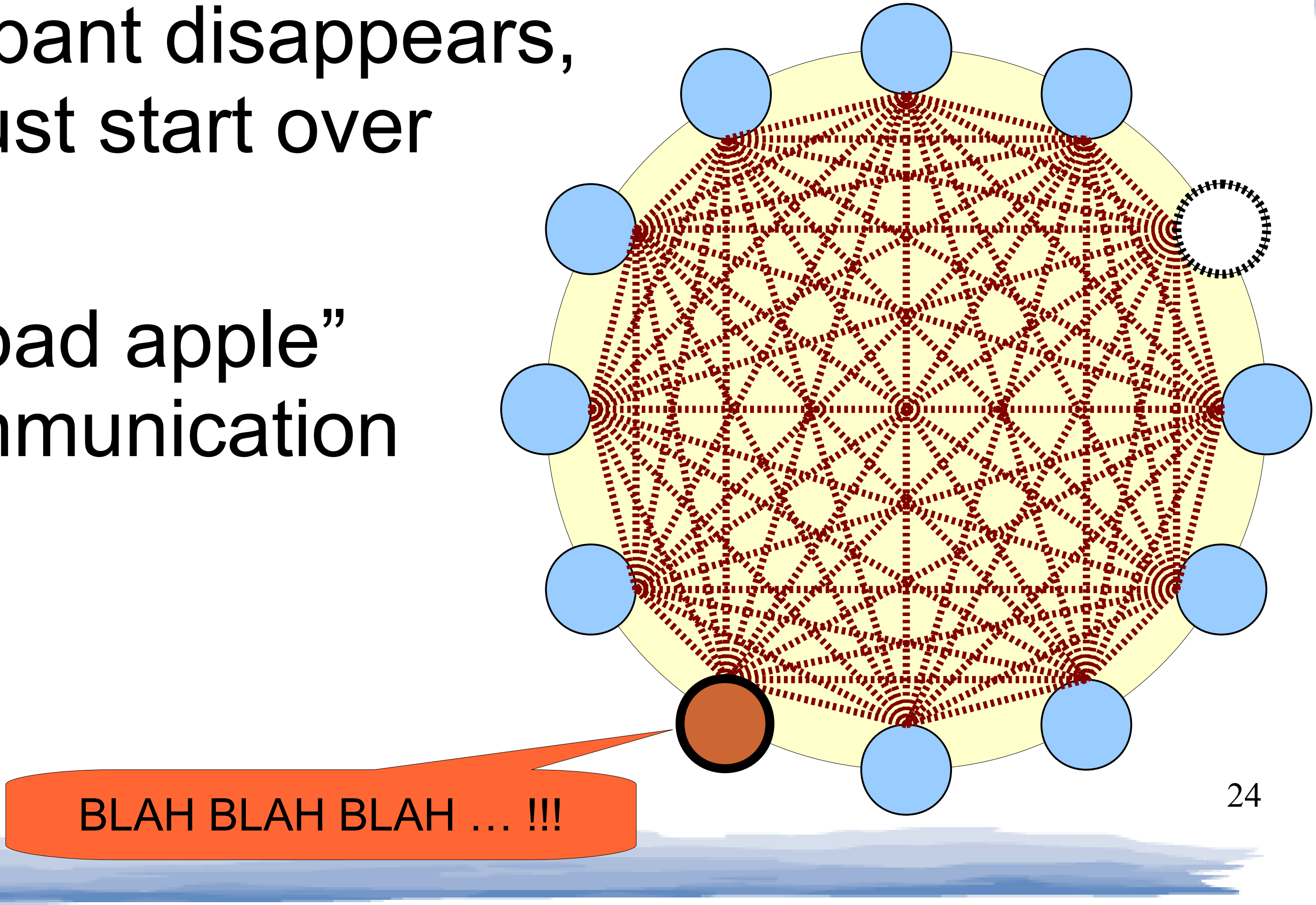
- Provable security against traffic analysis

But never widely used:

- Vulnerable to anonymous disruption
- Hard to scale

Why DC-nets Doesn't Scale

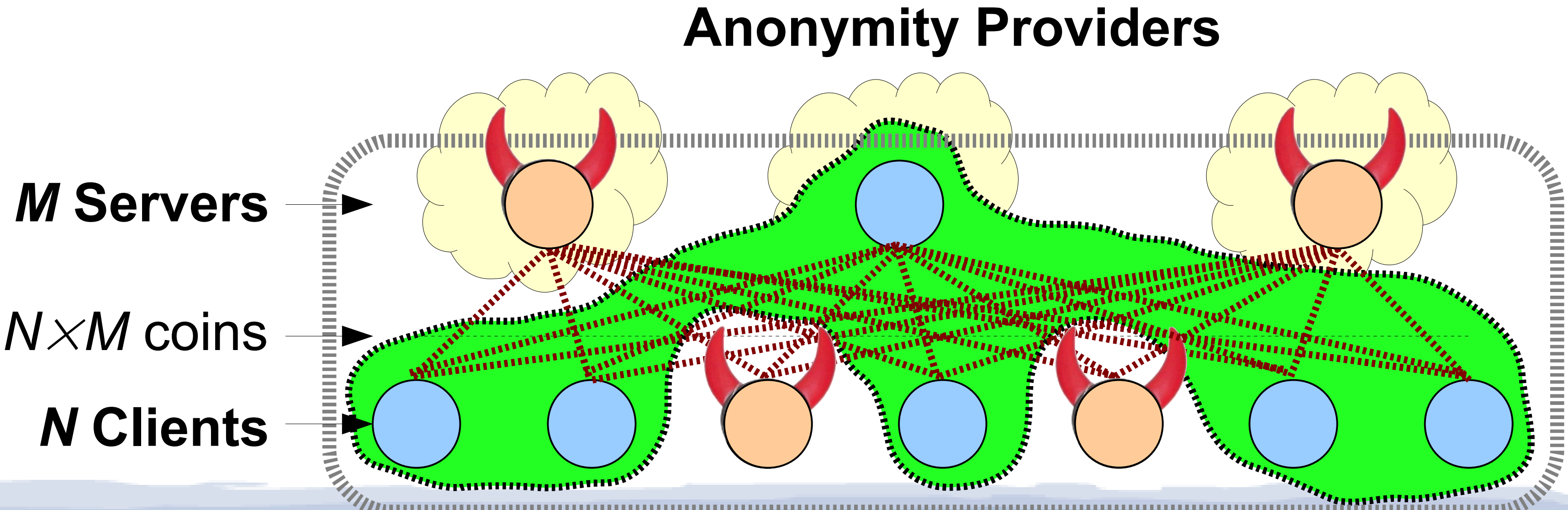
- **Computation cost:** *N*×*N* shared coin matrix
- **Network churn:**
  if *any* participant disappears,
  all nodes must start over
- **Disruption:**
  any single “bad apple”
  can jam communication



“Dissent in Numbers” [OSDI 12]

Many *clients* rely on a few independent *servers*

- Clients share coins *only* with servers
- As long as *at least one* honest server *exists*, yields ideal anonymity among *all honest clients*

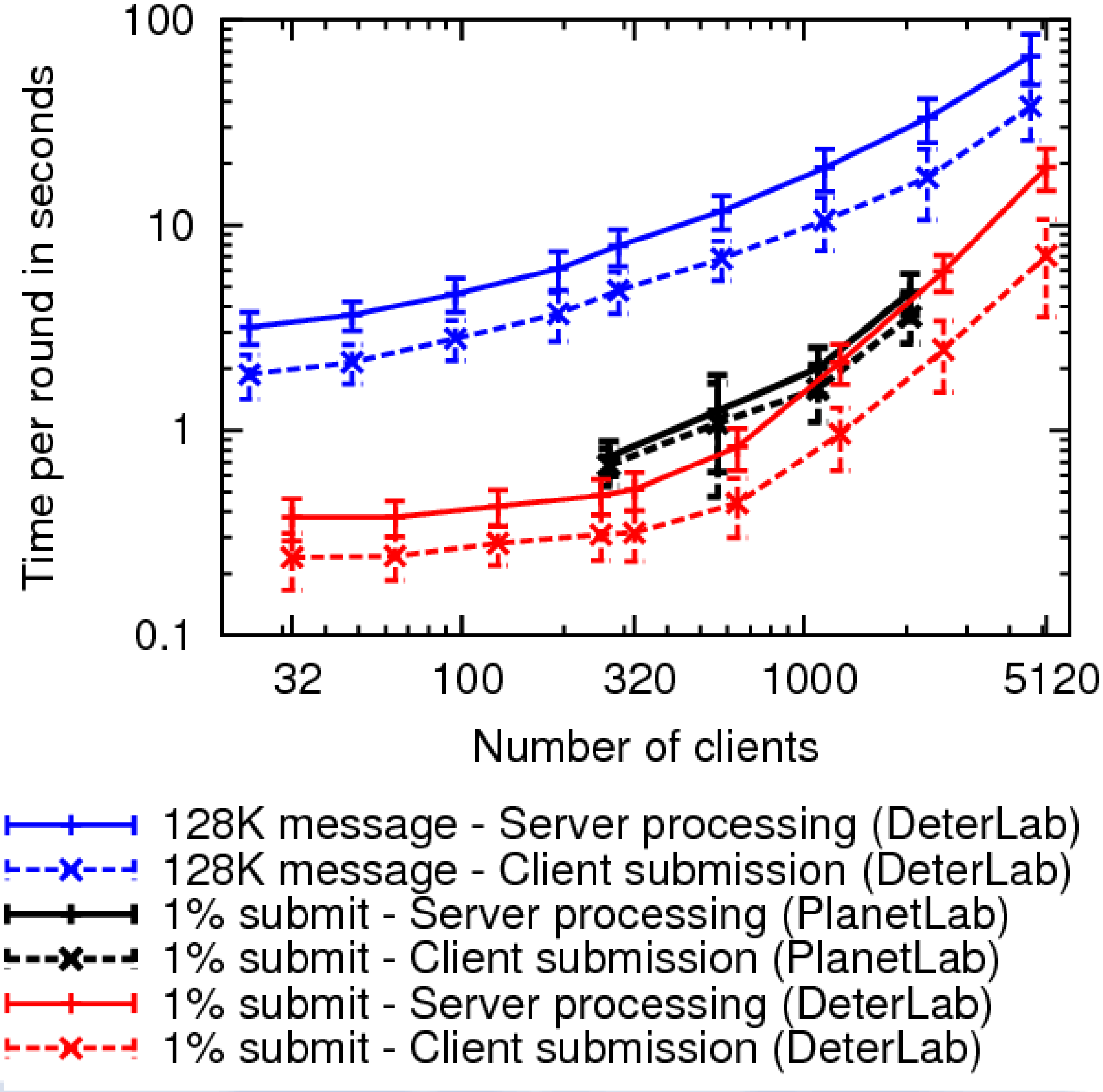


Scaling to Thousands of Clients

100× larger anonymity sets

- (Herbivore, Dissent v1: ~40 clients)

<1 sec latency w/ 1000 clients



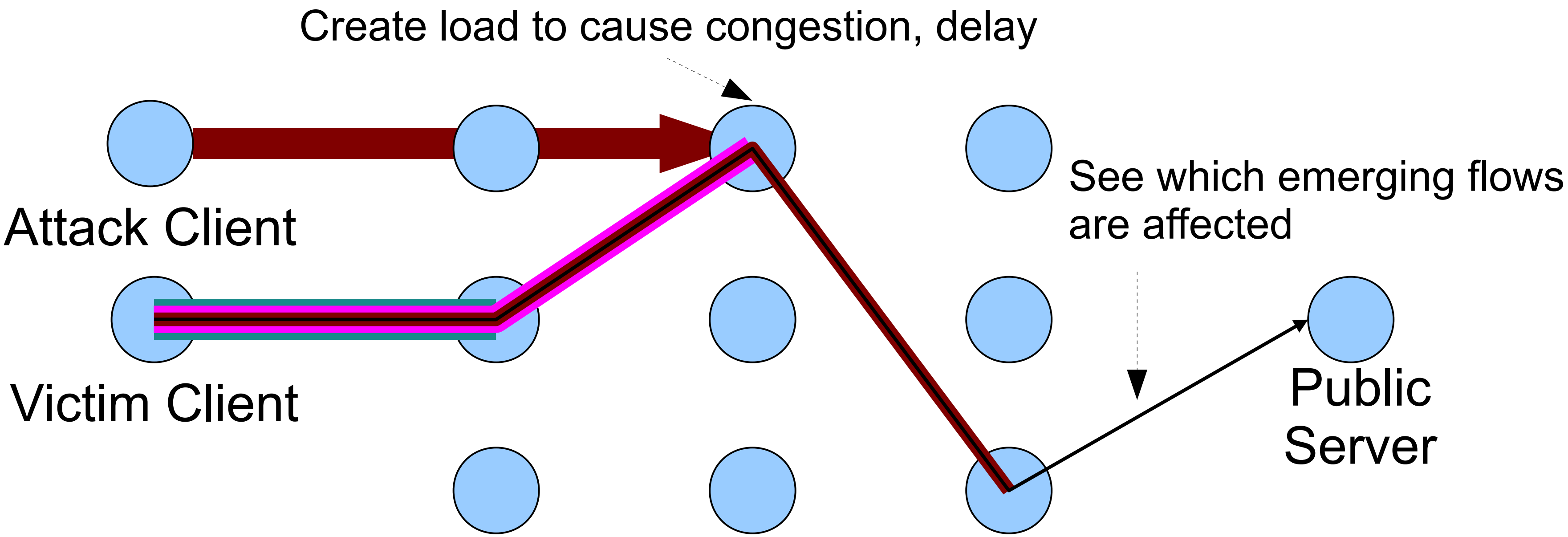
Talk Outline

- Anonymity: Motivation and Background
- *Dissent*, and How It Resists Strong Attacks
  - *DC-nets* and *shuffles* resist global traffic analysis
  - ➔ *Collective control plane* resists active attacks
  - *Accountability* resists denial-of-security (DoSec)
  - *Metrics* and *buddies* resist intersection attacks
  - *Pseudonym VMs* resist de-anonymizing exploits
- Dissent Status: Where We Are, and Aren't
- Conclusion

Active Attacks

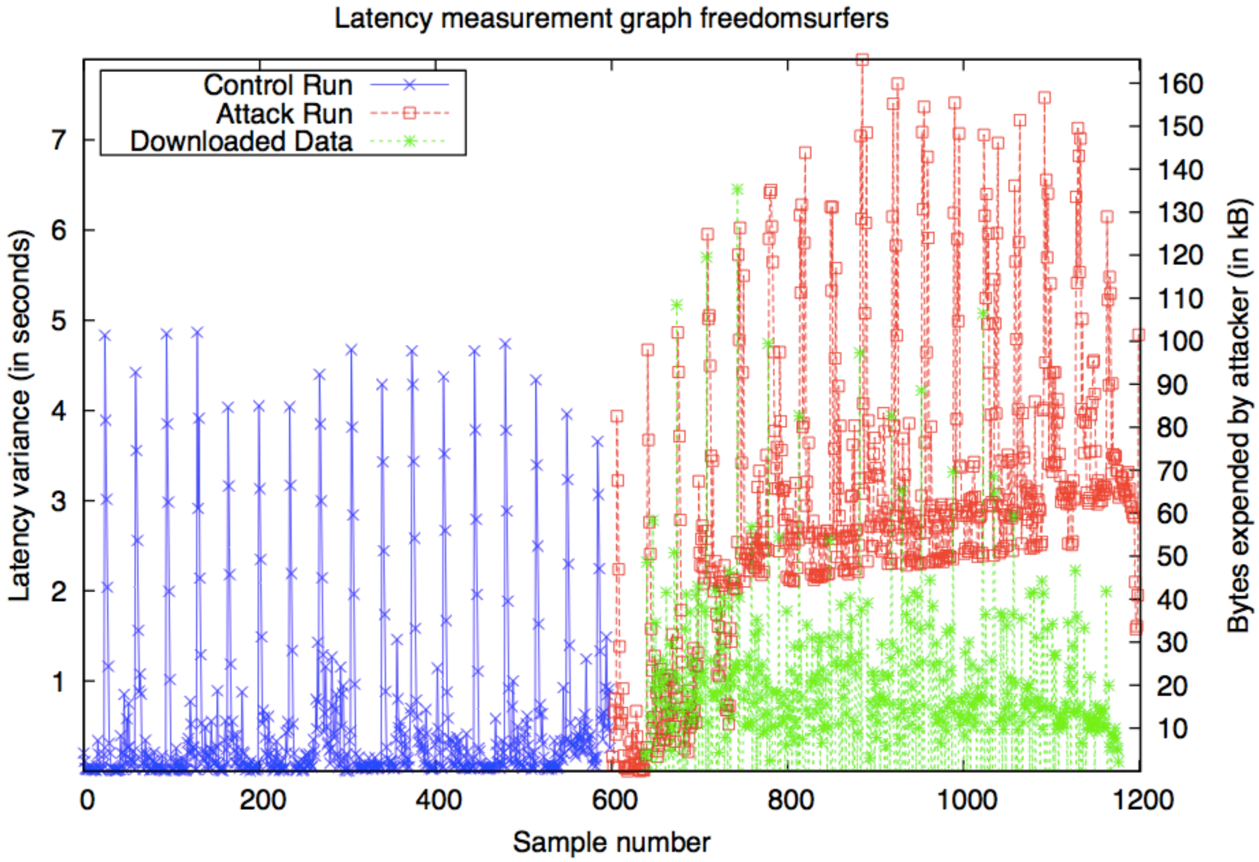
Attacker perturbs performance to inject traceable side-channel “markers” into flows

- Example: “congestion attacks” against Tor (e.g., Murdoch 05, Evans 09)


Create load to cause congestion, delay
See which emerging flows are affected
Attack Client
Victim Client
Public Server

Are Active Attacks Feasible?

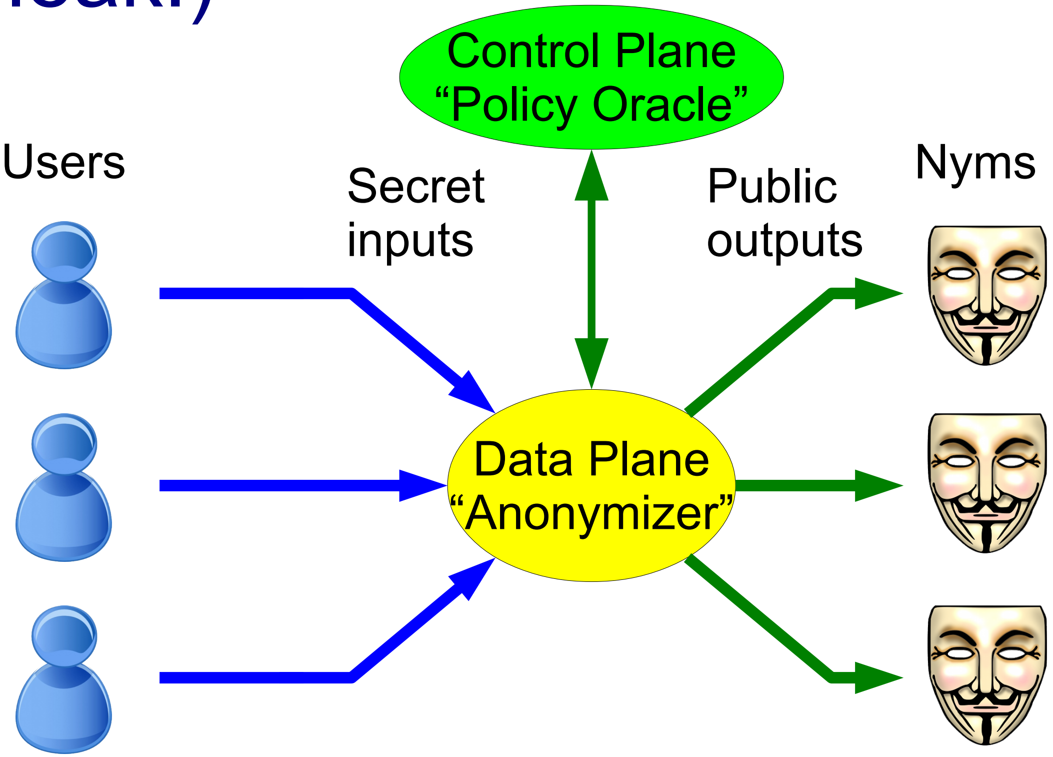
- “A Practical Congestion Attack on Tor” Evans et al. USENIX Security 09



Collective Control Plane (CCC) Model

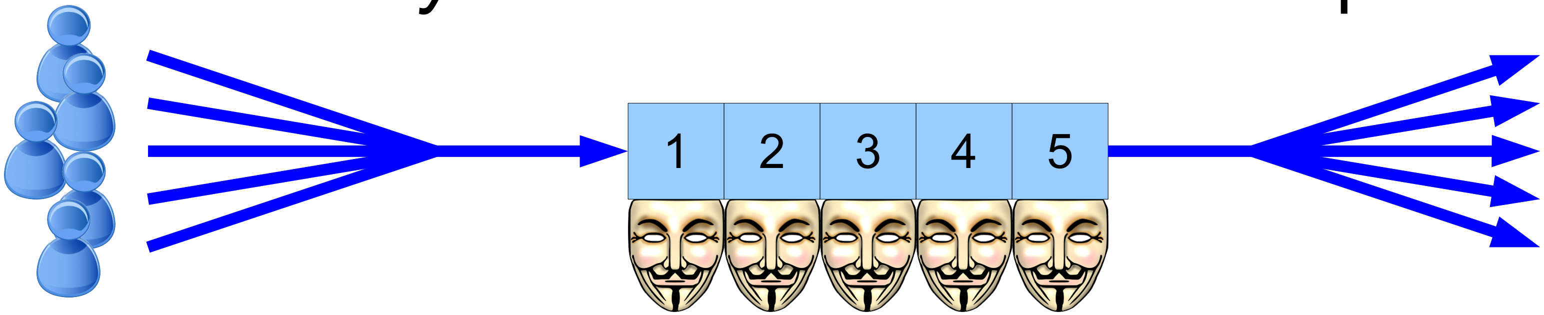
Policy Oracle controls when/how much to send

- But *does not know* who owns which nyms (can't leak!)

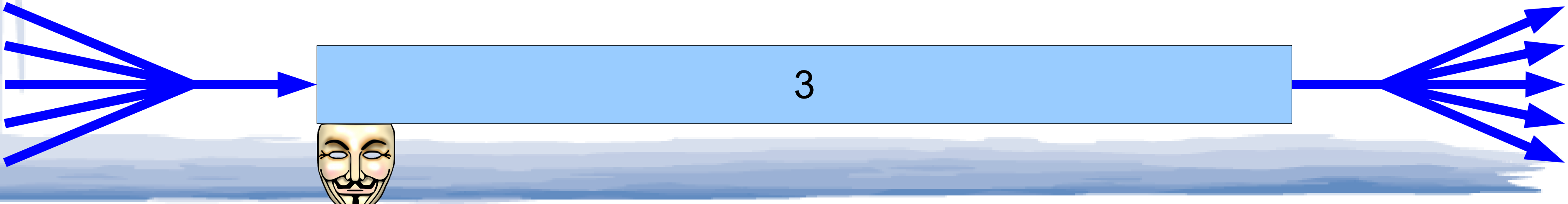


Scheduling Example - “Simon Says”

- Round 1: Policy Oracle (“Simon”) says, “Pseudonyms 1-5 each get 1-bit request slot”
  - *Everyone* sends 5-bit DC-nets ciphertext

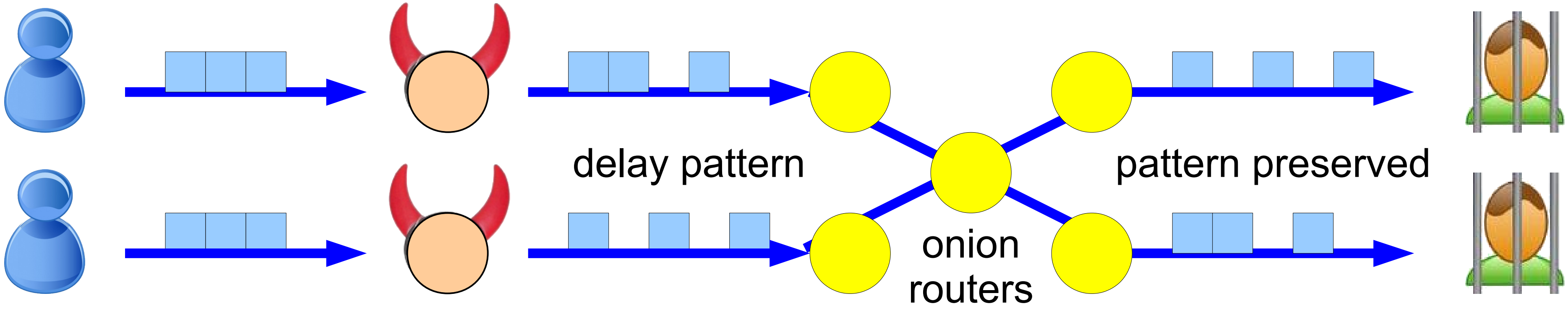


- Round 2: Policy Oracle (“Simon”) says, “Nym 3 wants to send, gets 1024 byte slot”
  - *Everyone* sends 1024-byte DC-net ciphertext

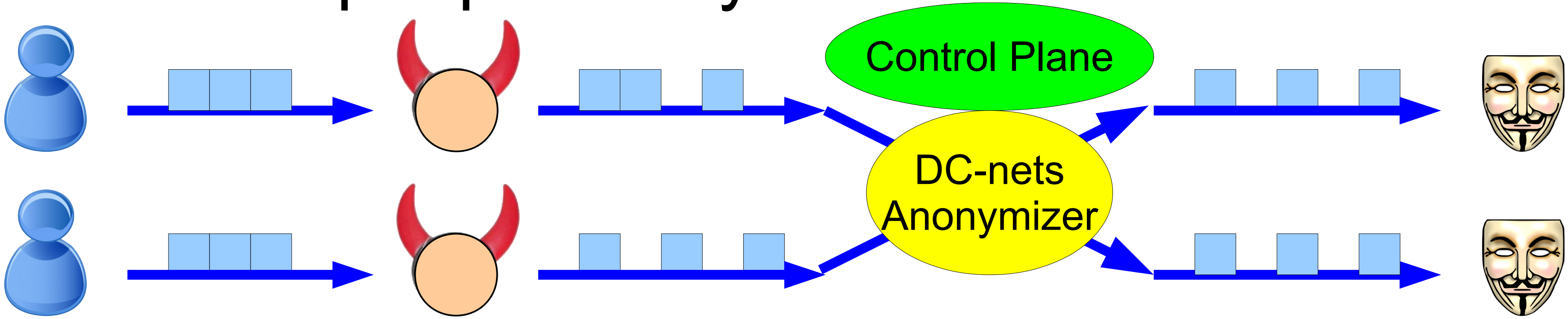


How CCC Counters Active Attacks

Onion routing preserves *individual* flow properties:



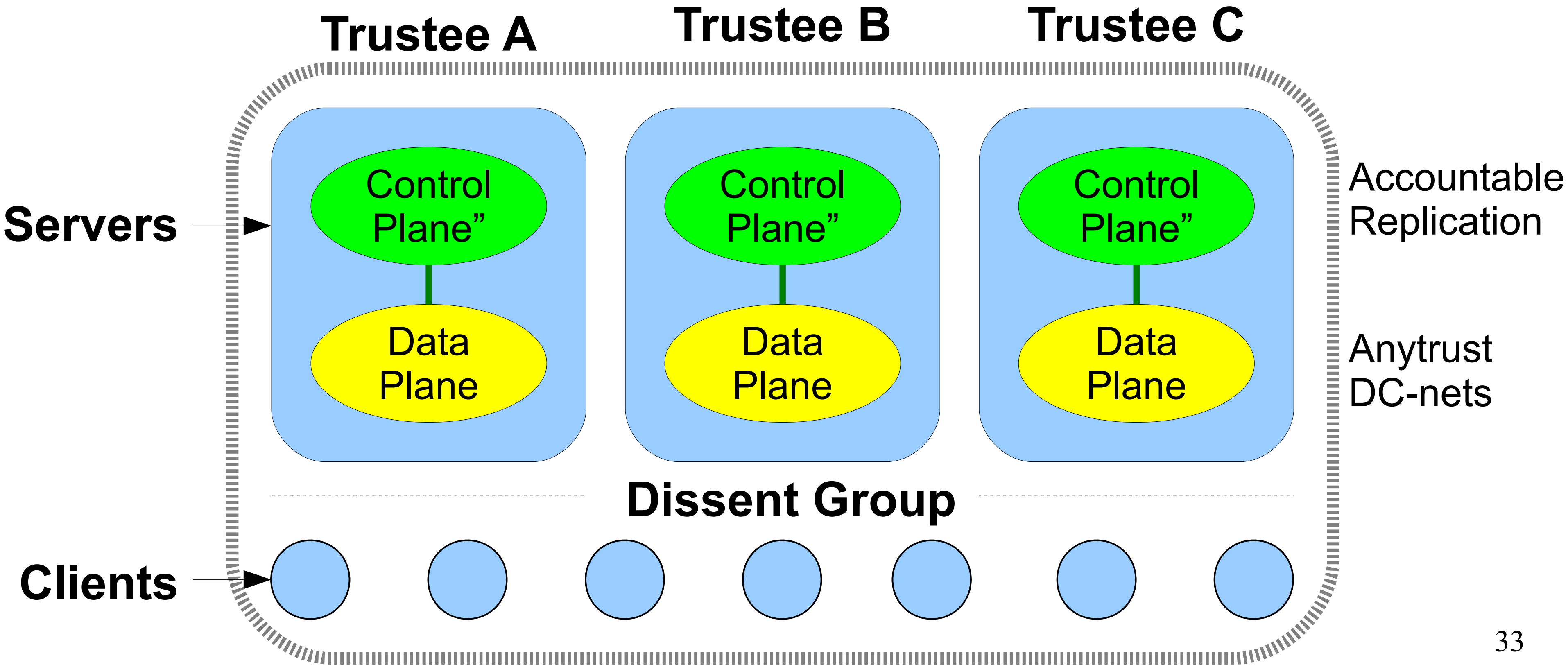
Dissent output paced by *collective* control:



Implementing the CCC

Accountable replication of control plane logic

- Each server implements copy, all must agree



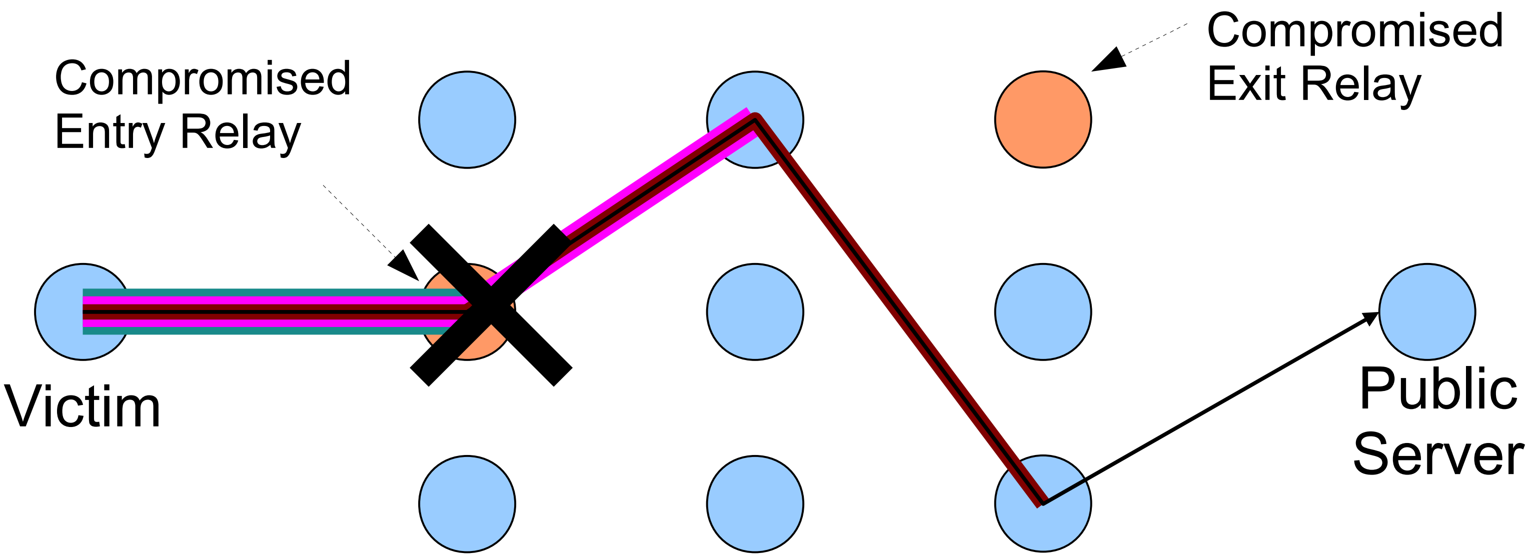
Talk Outline

- Anonymity: Motivation and Background
- *Dissent*, and How It Resists Strong Attacks
  - *DC-nets* and *shuffles* resist global traffic analysis
  - *Collective control plane* resists active attacks
  - ➔ *Accountability* resists denial-of-security (DoSec)
  - *Metrics* and *buddies* resist intersection attacks
  - *Pseudonym VMs* resist de-anonymizing exploits
- Dissent Status: Where We Are, and Aren't
- Conclusion

DoS can Compromise Anonymity

Attacker controls some relays

Step 1: victim chooses *partly* compromised path



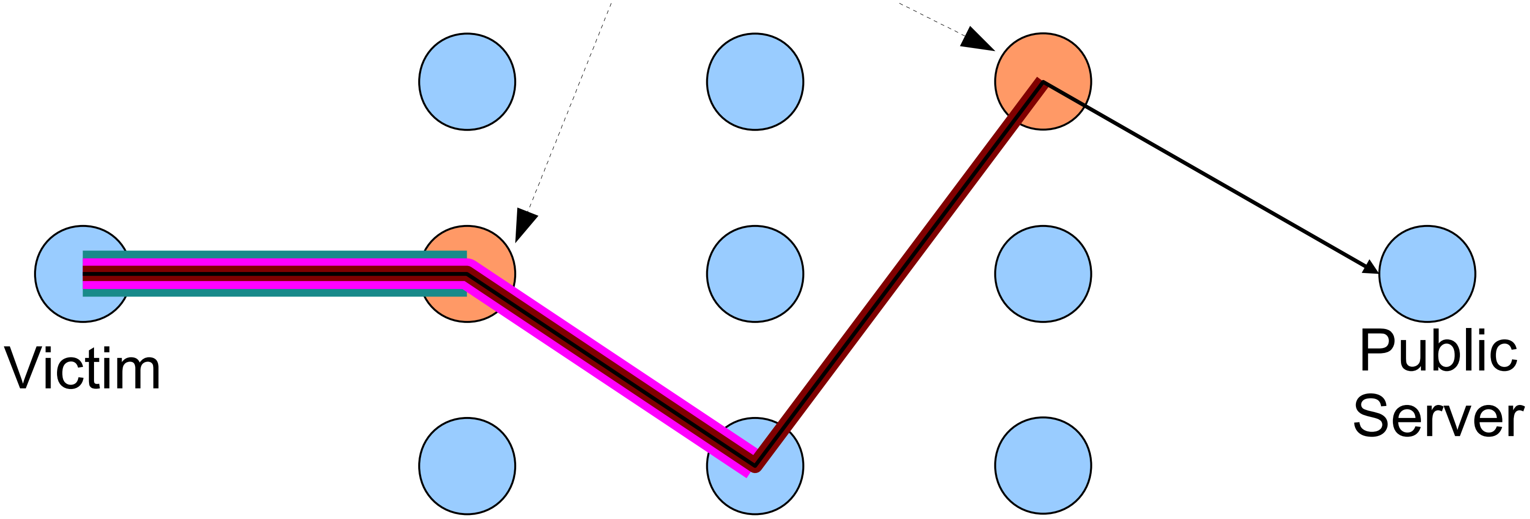
DoS can Compromise Anonymity

Attacker controls some relays

Step 1: victim chooses *partly* compromised path

Step 2: victim re-rolls until path *completely* broken

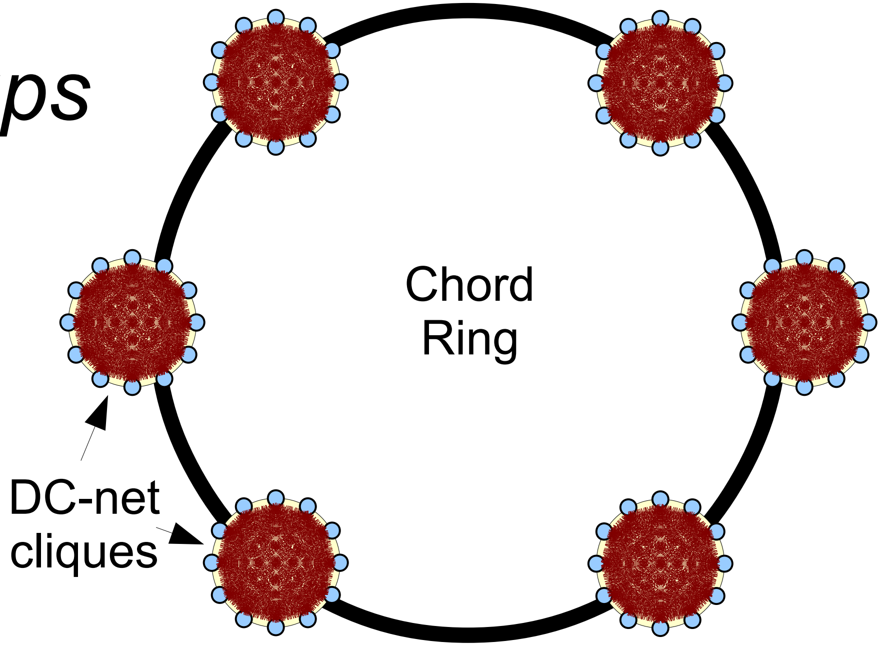
Correlate to de-anonymize, but *offer good service!*



Applies to DC-nets designs too!

Example: **Herbivore** [Sirer'04]

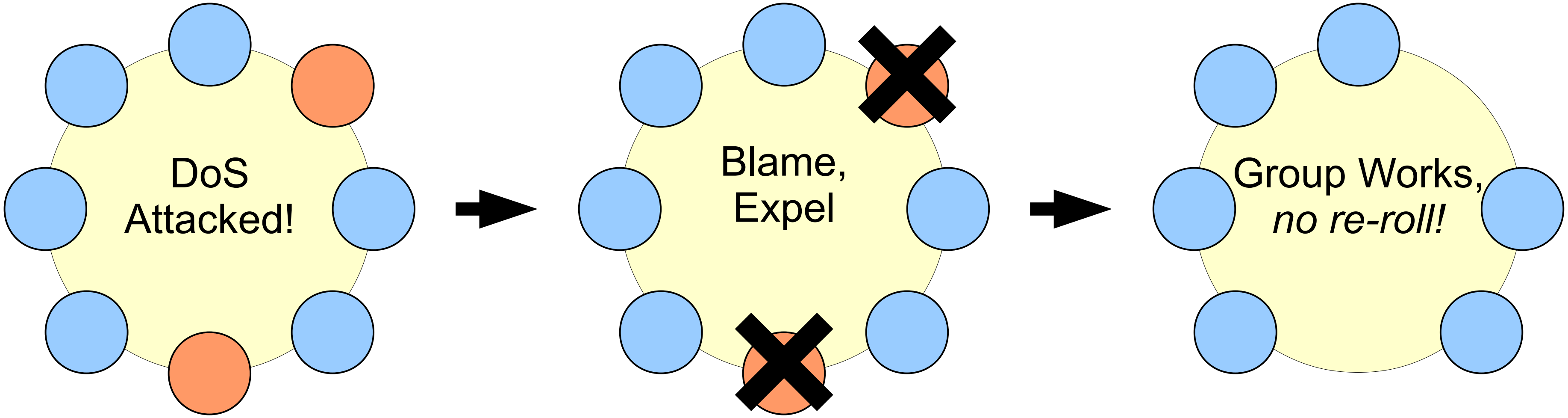
- Divide large network into small groups
    - If one doesn't work, join another
- Smart attacker jams *partly-compromised groups*
- Good service in groups with *only one* honest victim



Why Accountability is Important

Dissent can identify and expel a disruptor

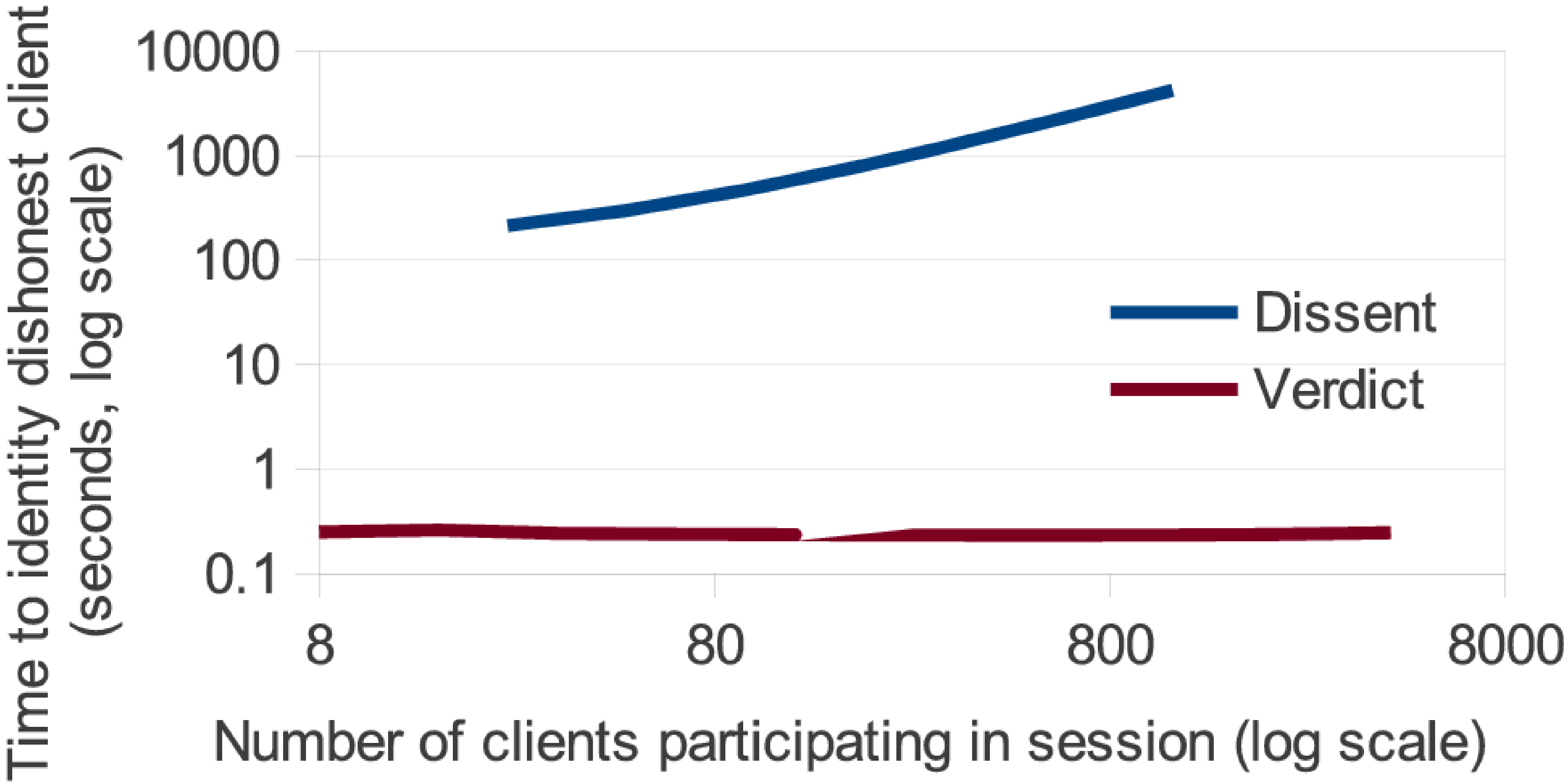
- *Without* forcing victims to re-roll dice
- Existing honest members *remain in group*
    - Attacker can't get *new* attack nodes in new group!



Jam-Proofing DC-nets: 3 Ways

1. **Dissent v1** [CCS'10]:
   use Brickell/Shmatikov shuffle to distribute hash-checked *assignments* before round
   - Simple, but requires expensive shuffle *each* round
2. **Scalable Dissent** [OSDI '12]:
   retroactive disruption-tracing “blame” protocol
   - Complex, efficient when *not* disrupted
3. **Verifiable Dissent** [USENIX Sec 13]:
   proactive verifiability via zero-knowledge proofs
   - Offline possible, lower blame cost *when* disrupted

“Blame” with Verifiability: 2-3 orders of magnitude faster



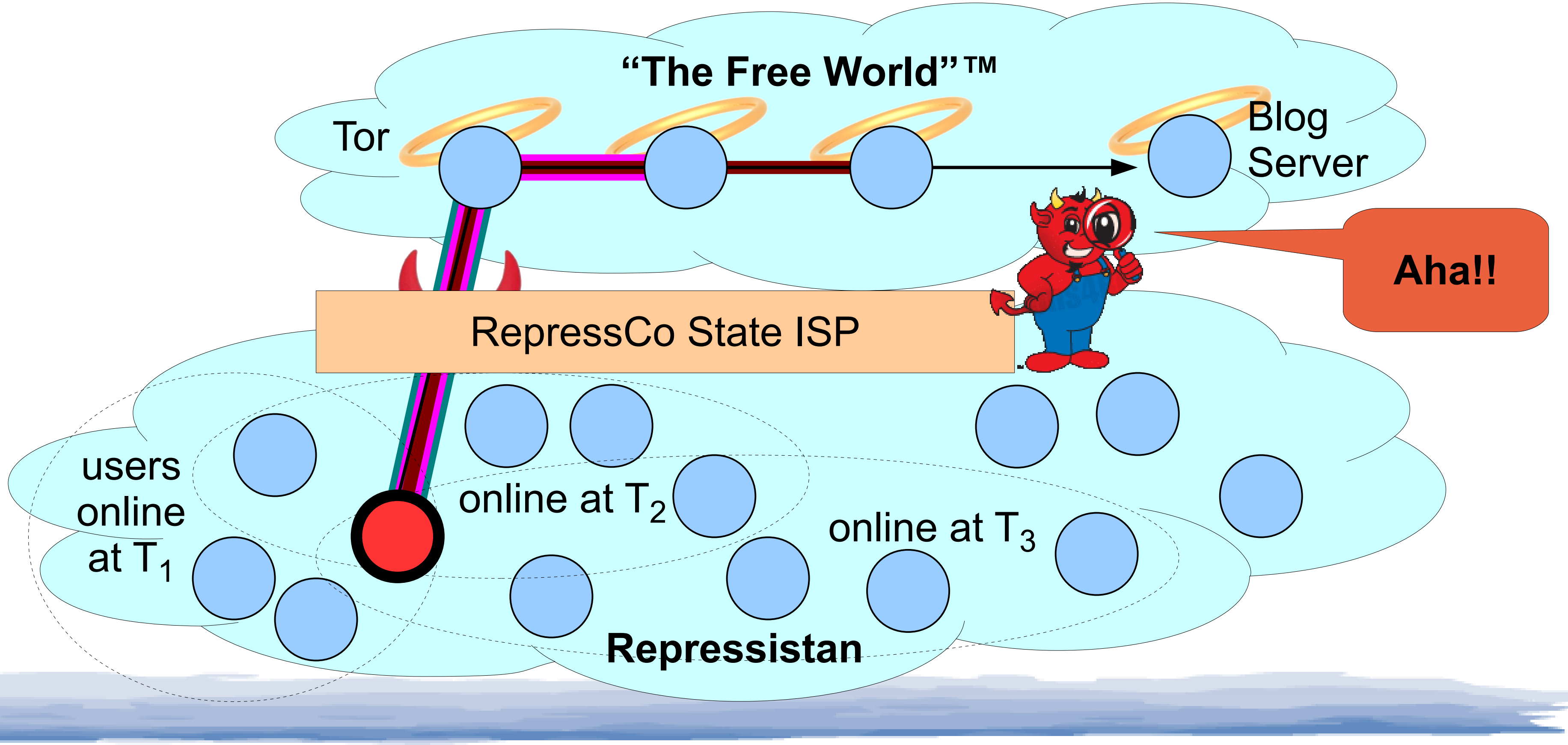
Talk Outline

- Anonymity: Motivation and Background
- *Dissent*, and How It Resists Strong Attacks
  - *DC-nets* and *shuffles* resist global traffic analysis
  - *Collective control plane* resists active attacks
  - *Accountability* resists denial-of-security (DoSec)
  - ➔ *Metrics* and *buddies* resist intersection attacks
  - *Pseudonym VMs* resist de-anonymizing exploits
- Dissent Status: Where We Are, and Aren't
- Conclusion

The Intersection Attack Problem

Kate signs posts with pseudonym "Bob"

- Posts signed messages at times T_1, T_2, T_3
- Police **intersects** user sets online each time



Introducing Buddies

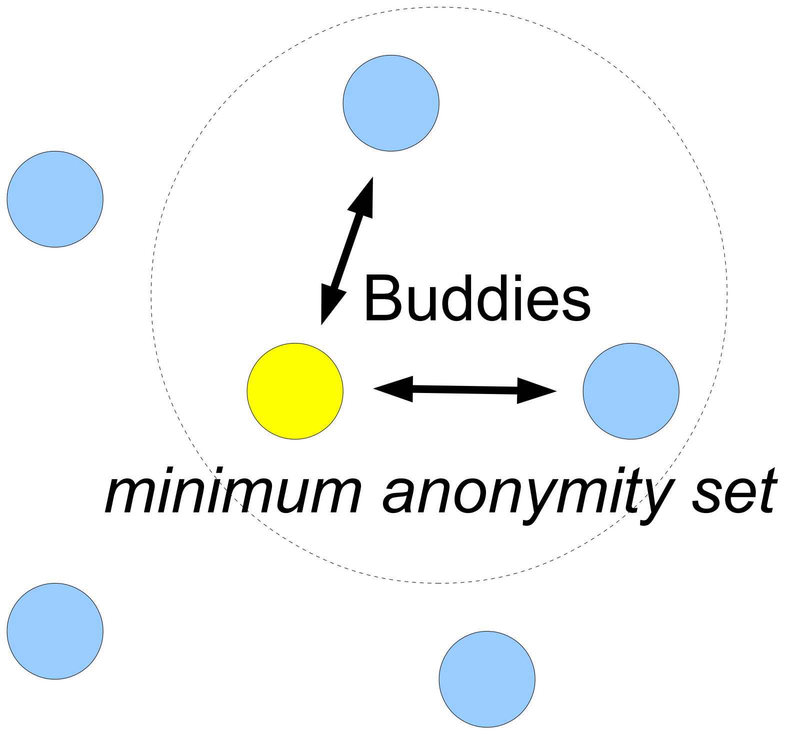
“Hang With Your Buddies to Resist Intersection Attacks” [CCS '13]

Goals:

- *Measure* anonymity under intersection attack
- Actively *mitigate* anonymity loss
- Enforce *lower bounds* by trading availability

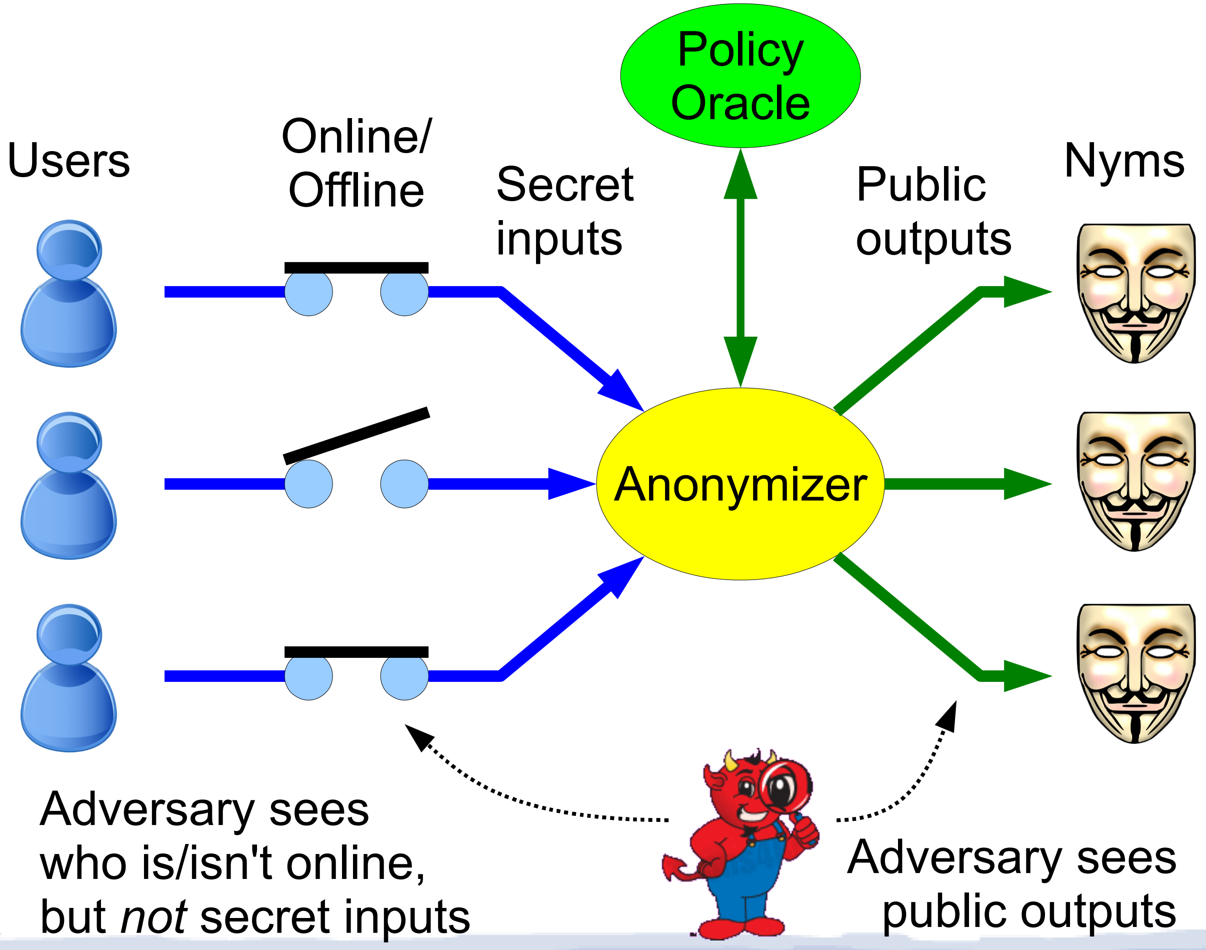
A Strawman Buddy System

- Pick a group of *buddies* for my anonymity set
- *Never* send linkable messages except when *all buddies* are also online (group members)



Buddies Conceptual Model

Focus: what adversary learns from *online status*



Adversary sees who is/isn't online, but *not* secret inputs

Adversary sees public outputs

Computing Anonymity Metrics

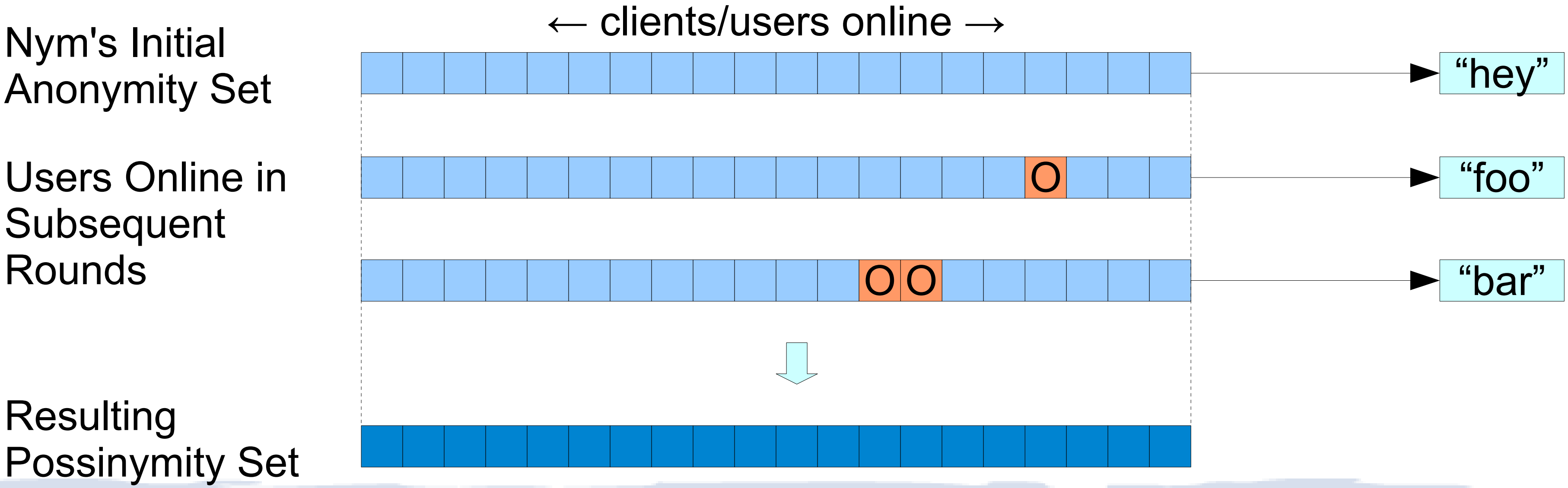
Policy Oracle *simulates an adversary's view*

- Knows who's online each round (via “tags”)
- Performs “intersection attacks” against Nyms
- Computes anonymity metrics
  - **Possinymity:** “possibilistic deniability”
  - **Indinymity:** “probabilistic indistinguishability”
- Reports metrics, uses them in policy decisions

Possinymity: Possibilistic Deniability

Set of users who *could conceivably* own Nym

- Intersection of sets of all users *online and unfiltered* in rounds where *a message appears*
- Simplistic, but may build “reasonable doubt”



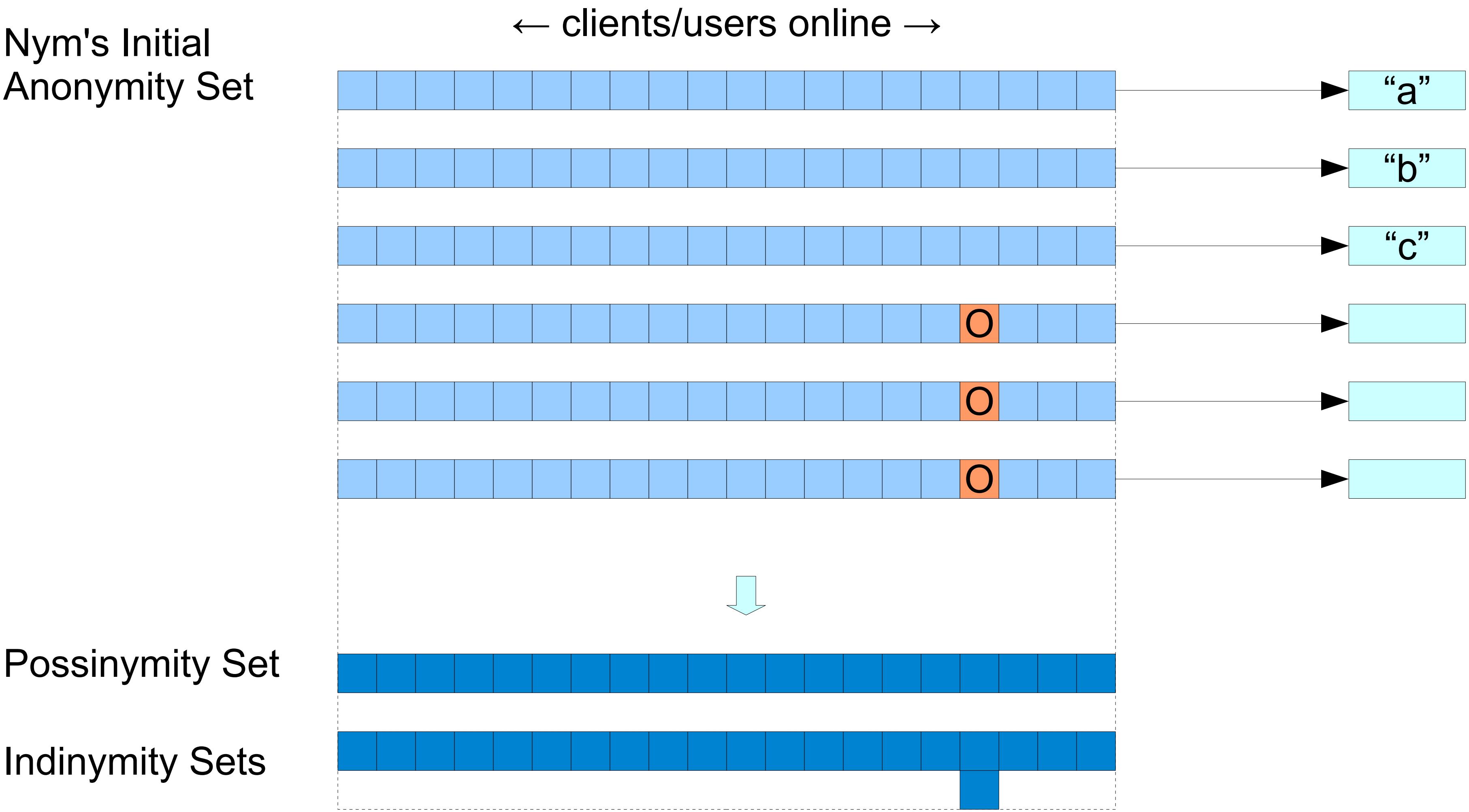
← clients/users online →

Nym's Initial Anonymity Set

Users Online in Subsequent Rounds

Resulting Possinymity Set

The “Statistical Disclosure” Problem



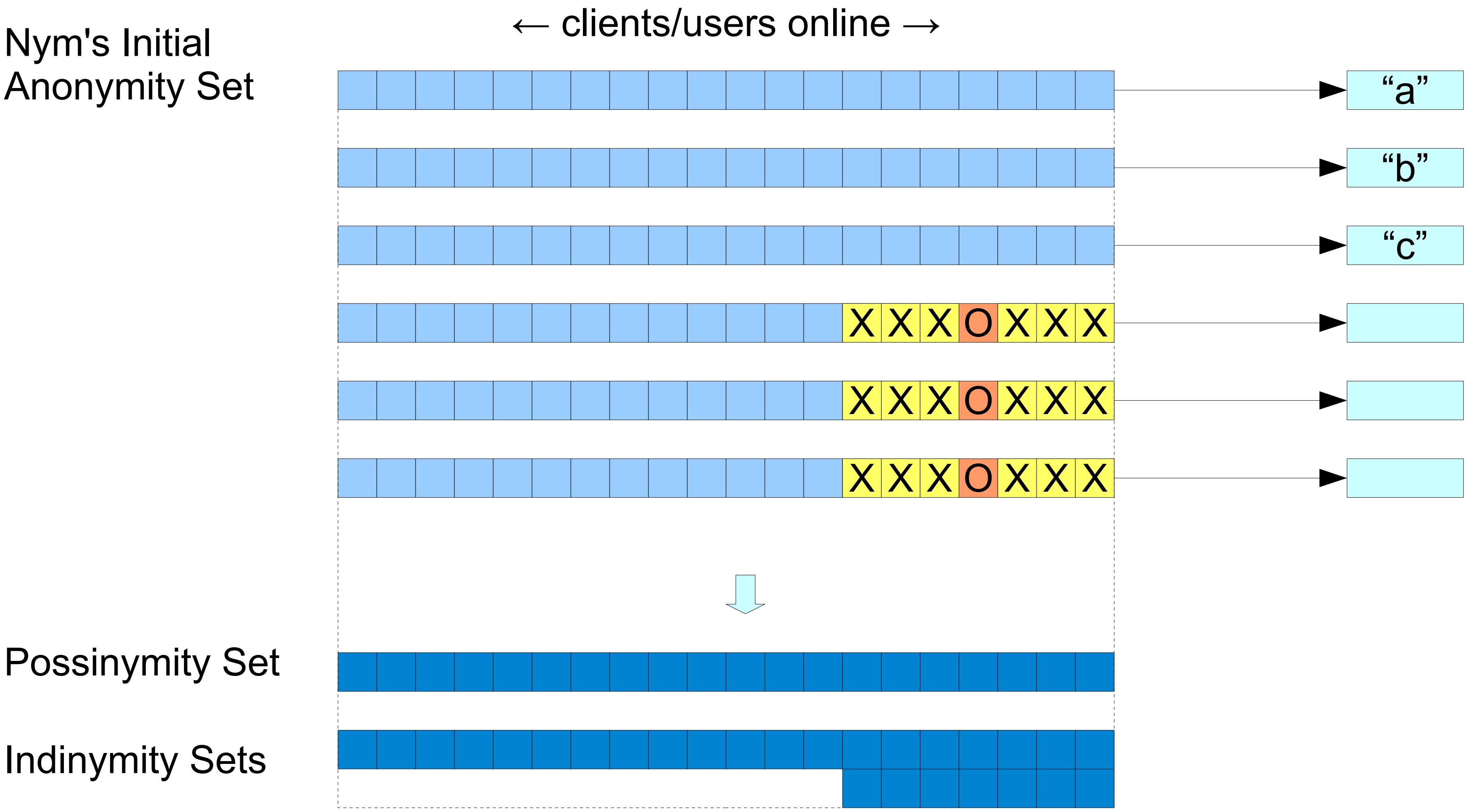
← clients/users online →

Nym's Initial Anonymity Set

Possinymity Set

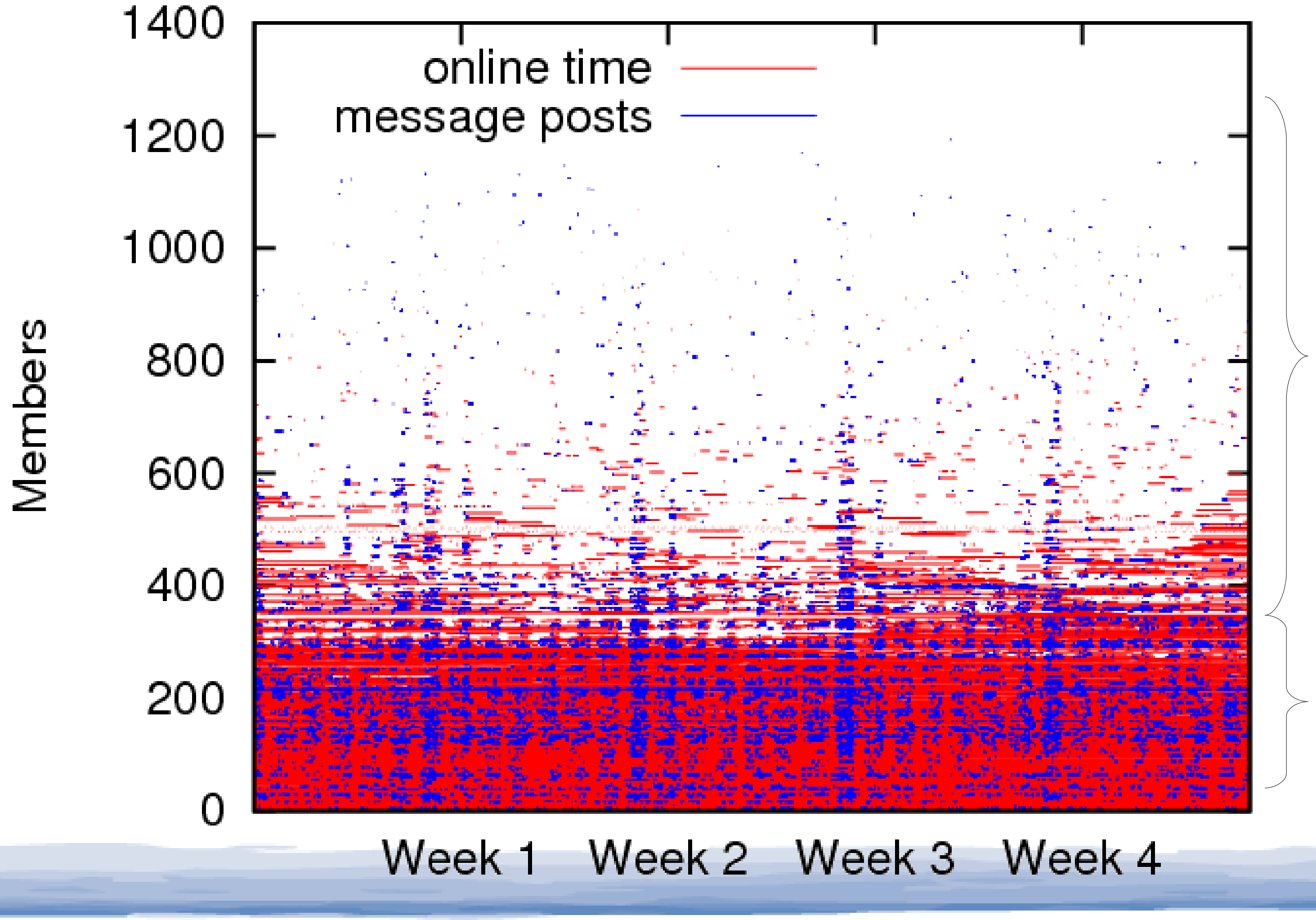
Indinymity Sets

Preserving Indinymity: Example



Is Resistance Futile?

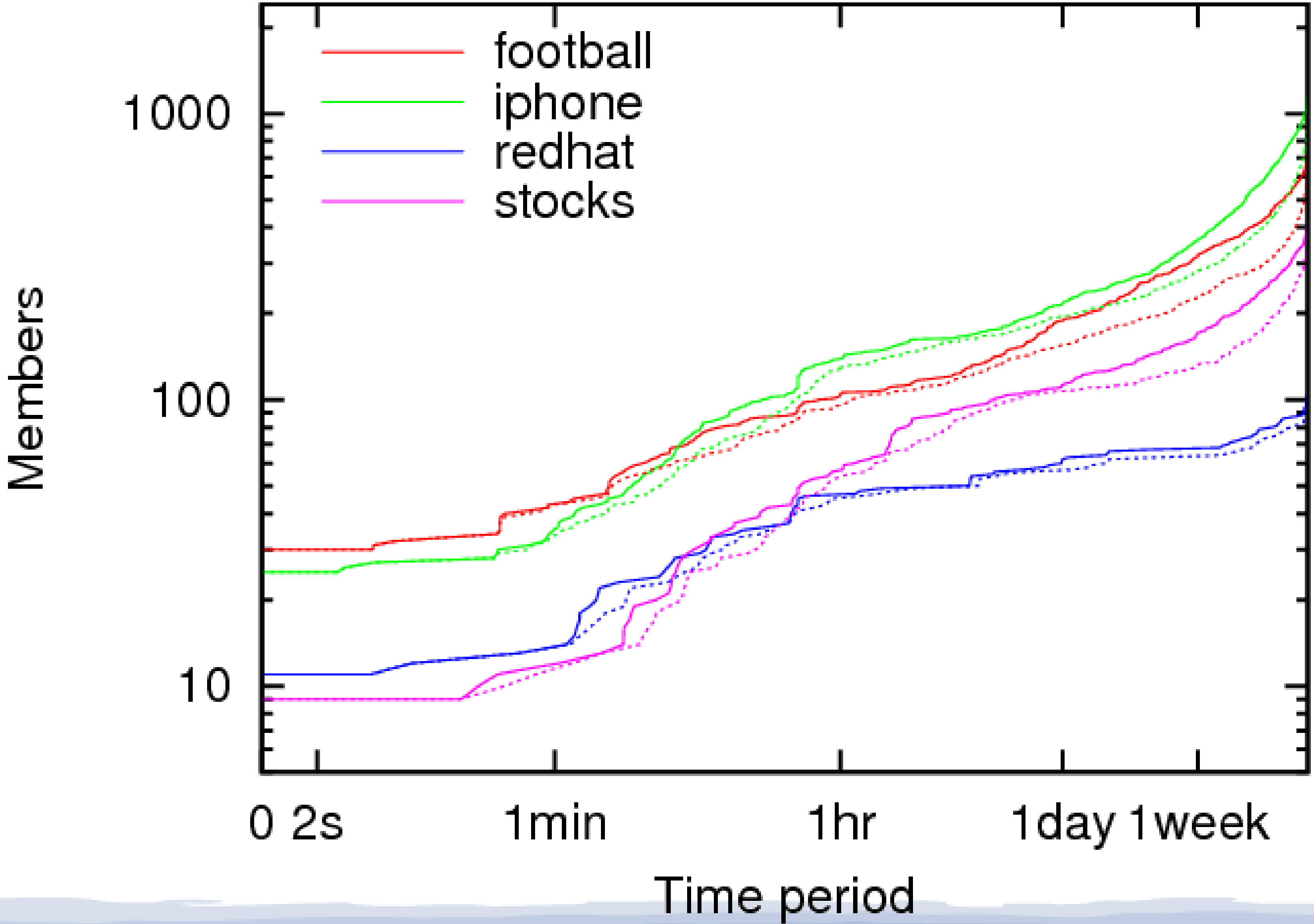
Analysis based on IRC online status traces



Ephemeral users

Where intersection attack resistant anonymity sets may plausibly be found

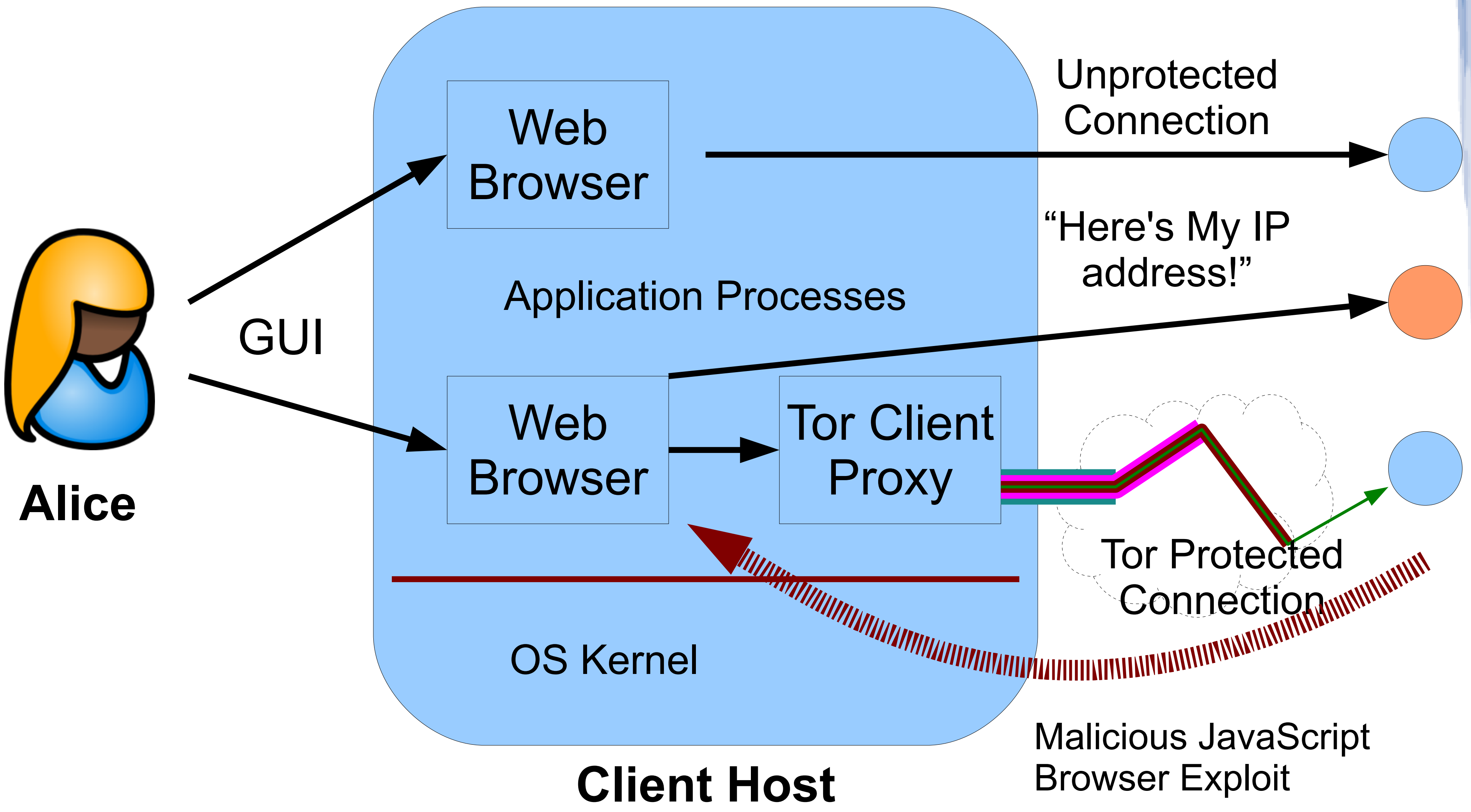
How Much Anonymity Can We Get?



Talk Outline

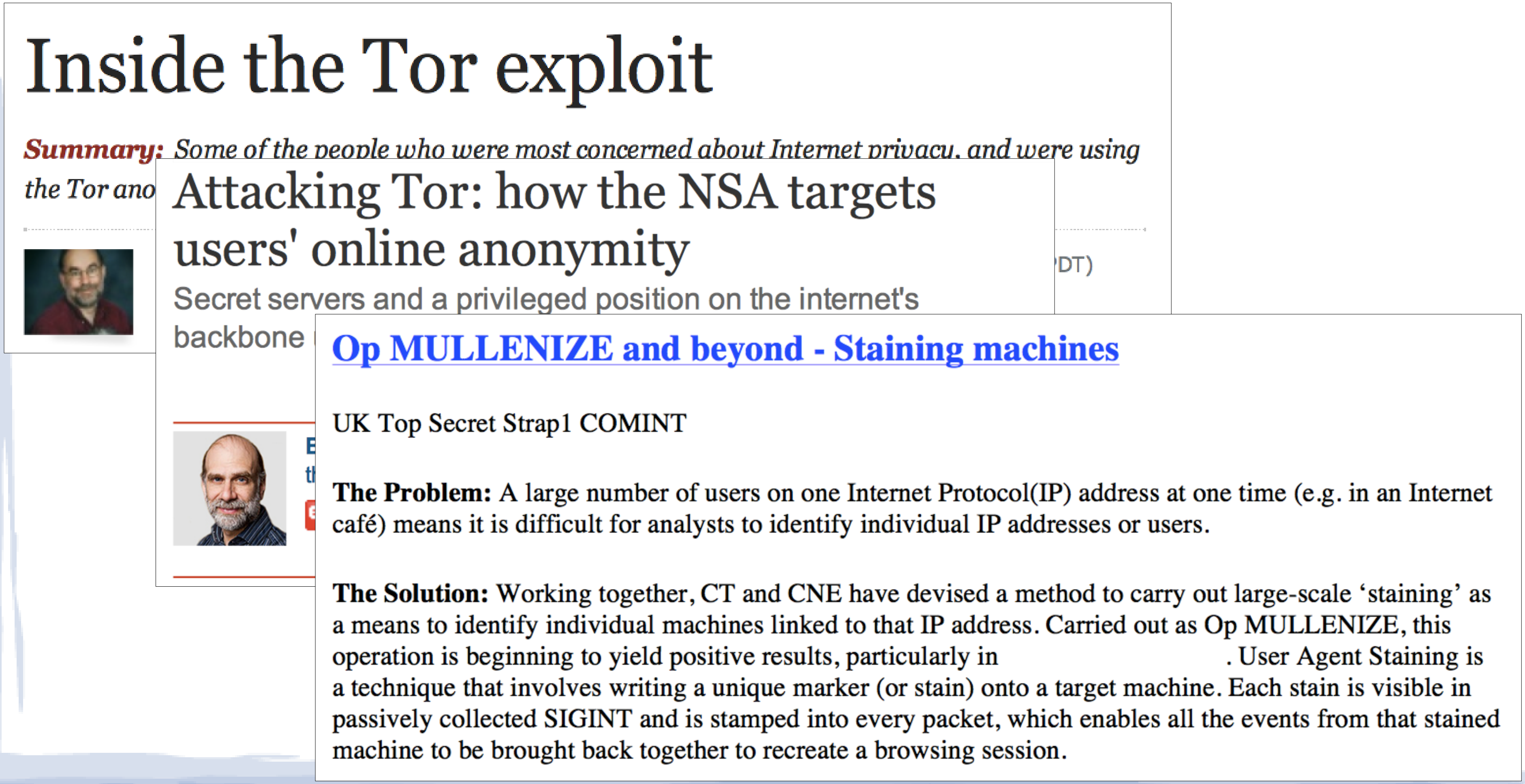
- Anonymity: Motivation and Background
- *Dissent*, and How It Resists Strong Attacks
  - *DC-nets* and *shuffles* resist global traffic analysis
  - *Collective control plane* resists active attacks
  - *Accountability* resists denial-of-security (DoSec)
  - *Metrics* and *buddies* resist intersection attacks
  - ➔ *Pseudonym VMs* resist de-anonymizing exploits
- Dissent Status: Where We Are, and Aren't
- Conclusion

Typical System Model



Exploits: The Low-Hanging Fruit

Circumvent the Anonymizer, Attack the Browser



Inside the Tor exploit

Summary: Some of the people who were most concerned about Internet privacy, and were using the Tor ano

DT)

Attacking Tor: how the NSA targets users' online anonymity

Secret servers and a privileged position on the internet's backbone

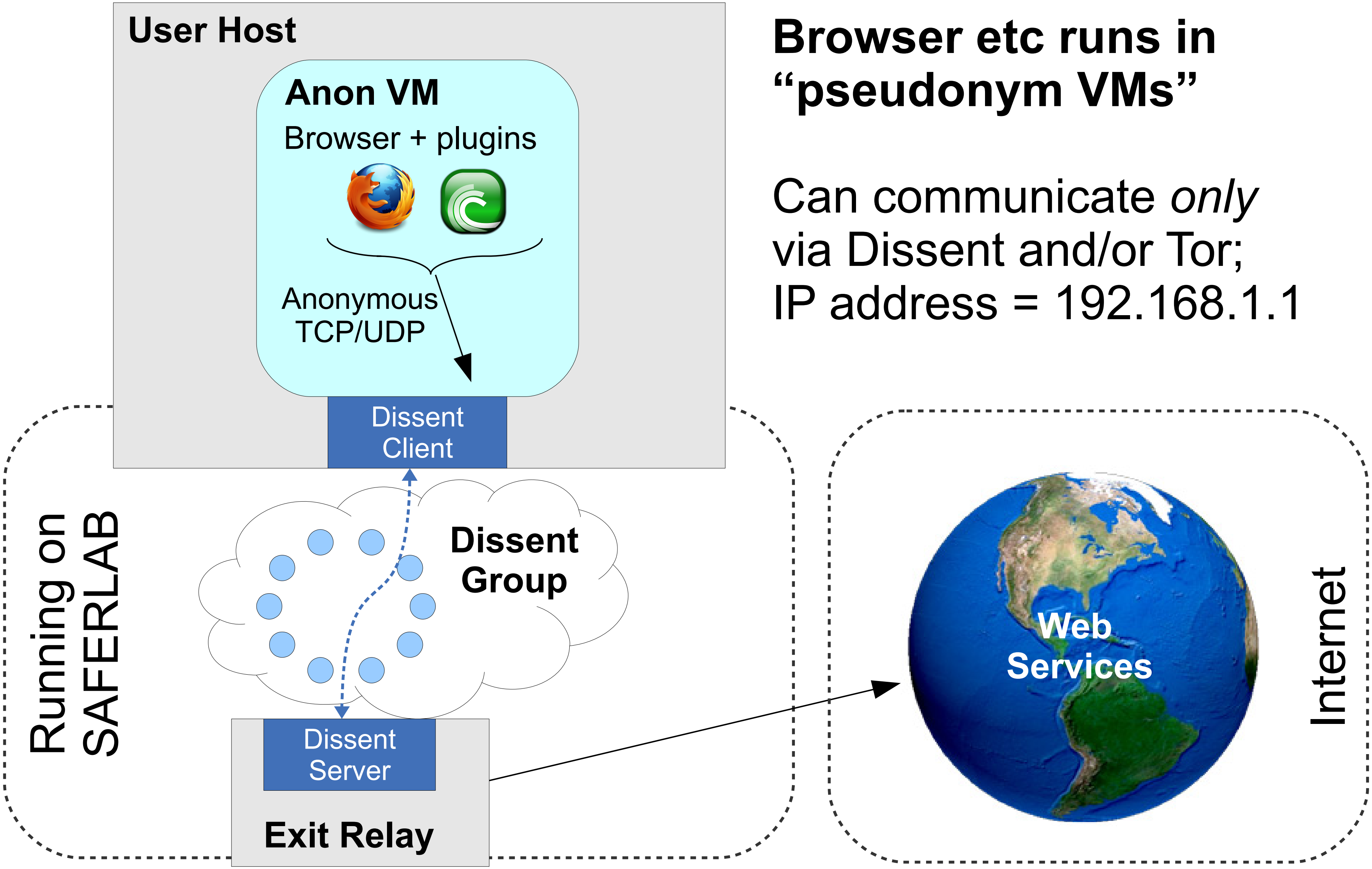
Op MULLENIZE and beyond - Staining machines

UK Top Secret Strap1 COMINT

The Problem: A large number of users on one Internet Protocol(IP) address at one time (e.g. in an Internet café) means it is difficult for analysts to identify individual IP addresses or users.

The Solution: Working together, CT and CNE have devised a method to carry out large-scale 'staining' as a means to identify individual machines linked to that IP address. Carried out as Op MULLENIZE, this operation is beginning to yield positive results, particularly in . User Agent Staining is a technique that involves writing a unique marker (or stain) onto a target machine. Each stain is visible in passively collected SIGINT and is stamped into every packet, which enables all the events from that stained machine to be brought back together to recreate a browsing session.

WiNon: VM-hardened Anonymity



Browser etc runs in "pseudonym VMs"

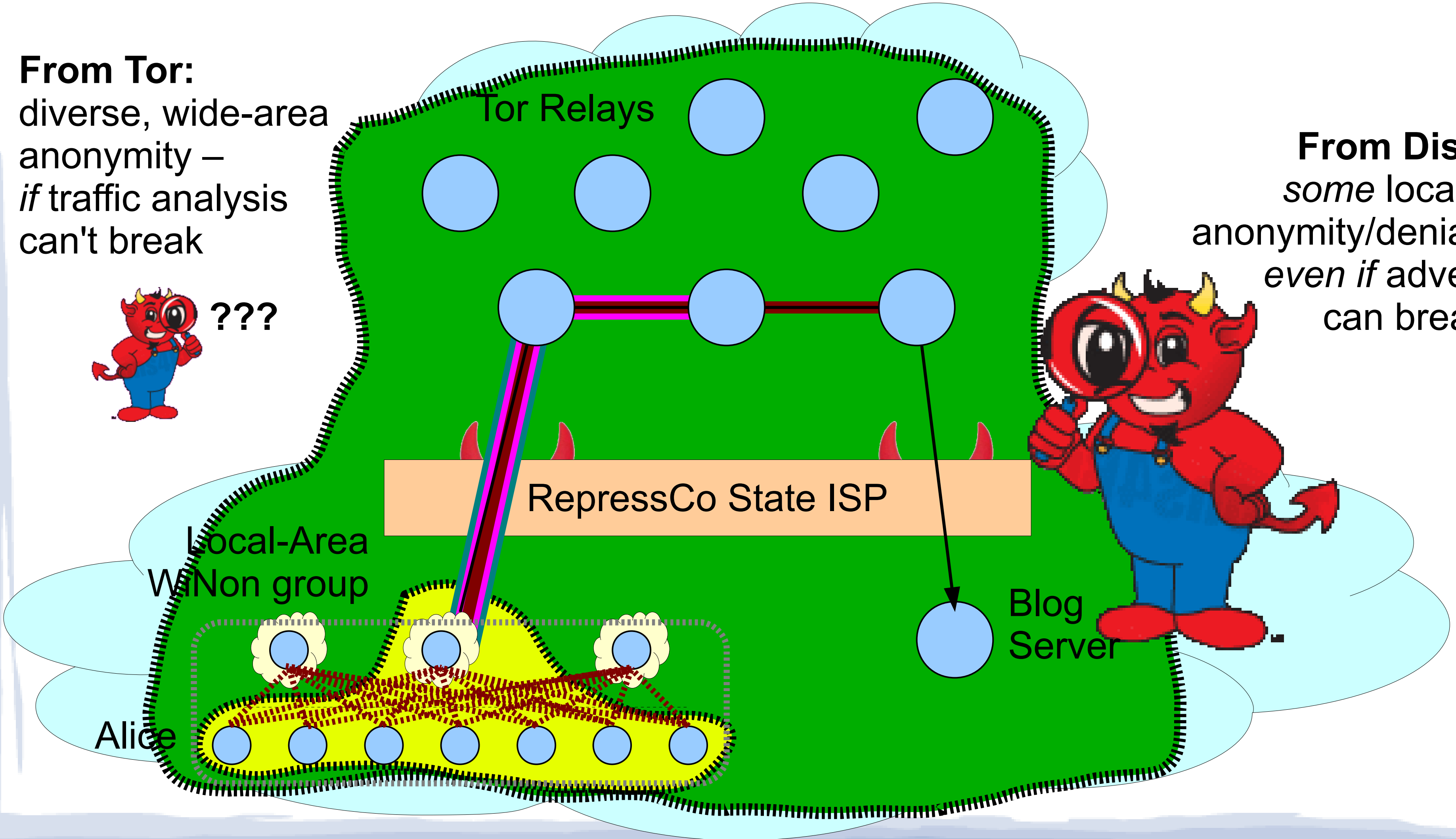
Can communicate *only* via Dissent and/or Tor; IP address = 192.168.1.1

Best of Both Worlds: **Dissent+Tor**

Defend against "Little Brother" *and* "Big Brother"

From Tor:
diverse, wide-area
anonymity –
if traffic analysis
can't break

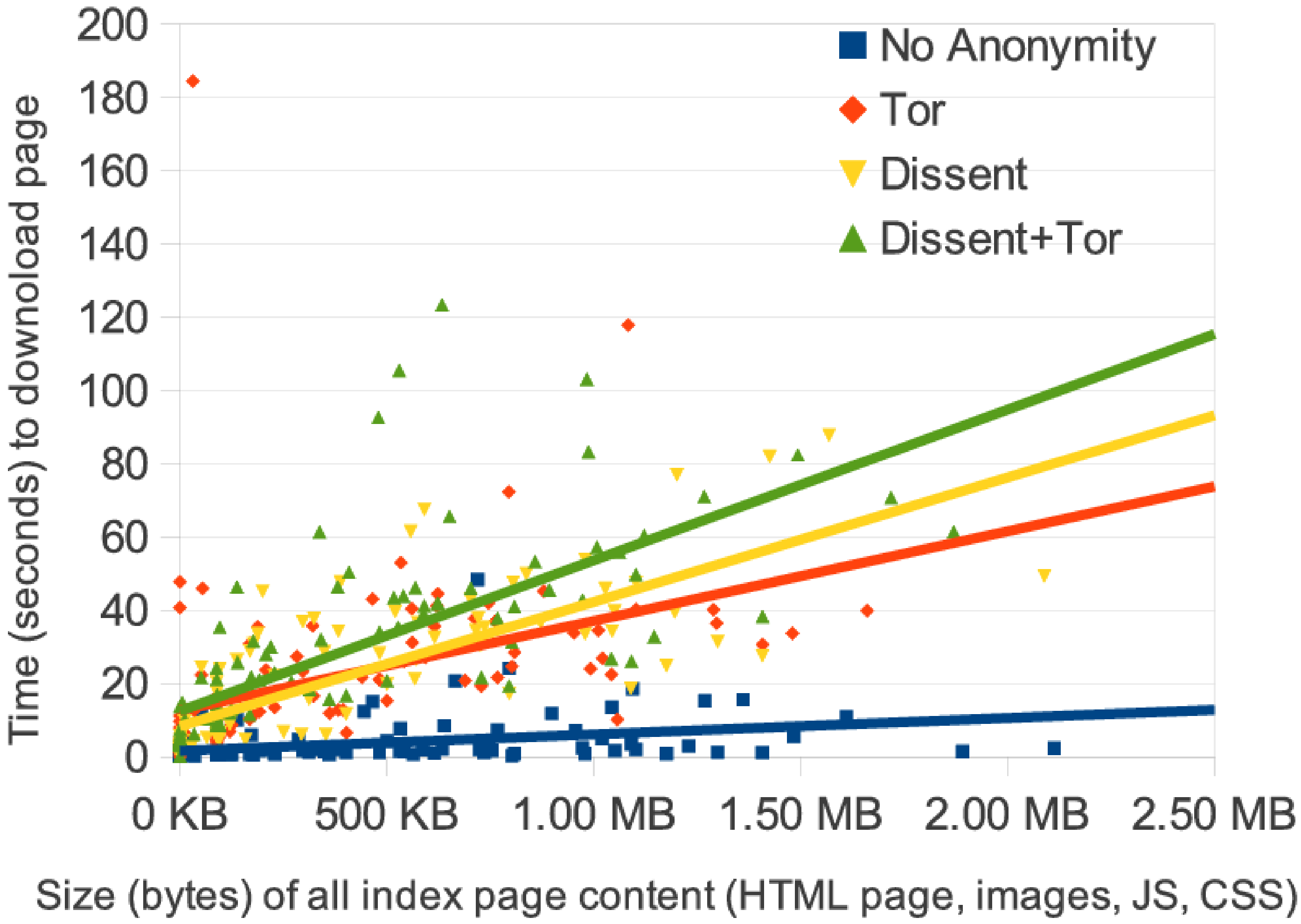
From Dissent:
some local-area
anonymity/deniability,
even if adversary
can break Tor



WiNon Browsing Latency

5 servers,
24 clients,
WiFi LAN
→ usability
comparable
to Tor

Illustrative only **–**
“apples-to-oranges”



Talk Outline

- Anonymity: Motivation and Background
- *Dissent*, and How It Resists Strong Attacks
    - *DC-nets* and *shuffles* resist global traffic analysis
    - *Collective control plane* resists active attacks
    - *Accountability* resists denial-of-security (DoSec)
    - *Metrics* and *buddies* resist intersection attacks
    - *Pseudonym VMs* resist de-anonymizing exploits
- ➔ Dissent Status: Where We Are, and Aren't
- Conclusion

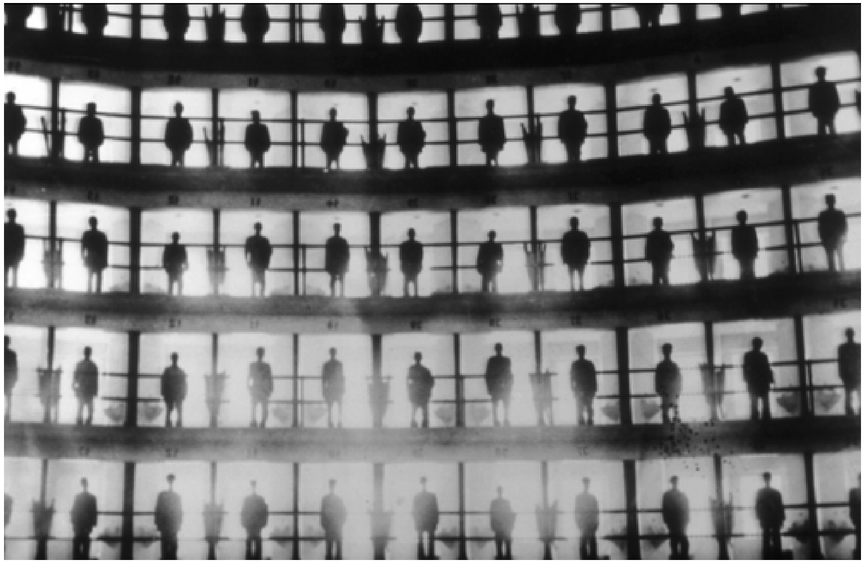
Current Status

- Proof-of-concept works, available on github
  - **Preliminary:** not at all feature-rich, user-friendly
  - **Don't** use it [yet] for security-critical activities!
- Long-term applicability questions
  - How well can we make it perform, scale?
  - Broadcast limits scalability for “point-to-point” use
  - *Might* be very efficient for multicast applications
    - Anonymous chat/microblogging, “town hall” meetings
- Time (and further development) will tell!

Conclusion

Can you hide in an Internet panopticon?
It's hard! – due to “five deadly attack classes”

- Global traffic analysis
- Active attacks
- Denial-of-security
- Intersection attacks
- Software exploits



Dissent: is first ground-up anonymity architecture with *any plausible solution* to all five classes

http://dedis.cs.yale.edu/dissent/