Tor and circumvention: Lessons learned

Nick Mathewson
The Tor Project
https://torproject.org/

What is Tor?

Online anonymity 1) open source software, 2) network, 3) protocol

Community of researchers, developers, users, and relay operators

Funding from US DoD, Electronic Frontier Foundation, Voice of America, Google, NLnet, Human Rights Watch, NSF, US State Dept, SIDA, ...

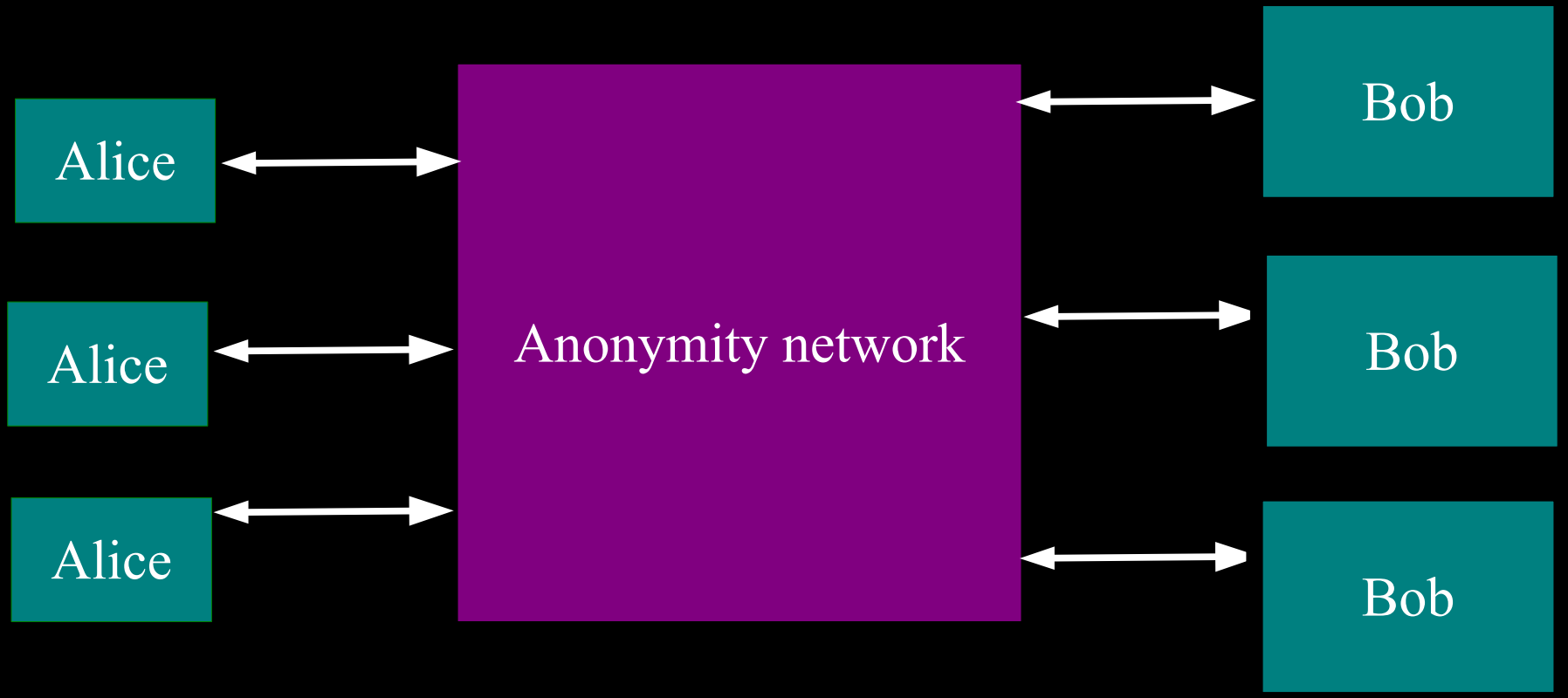
The Tor Project, Inc.

501(c)(3) non-profit organization dedicated to the research and development of tools for online anonymity and privacy

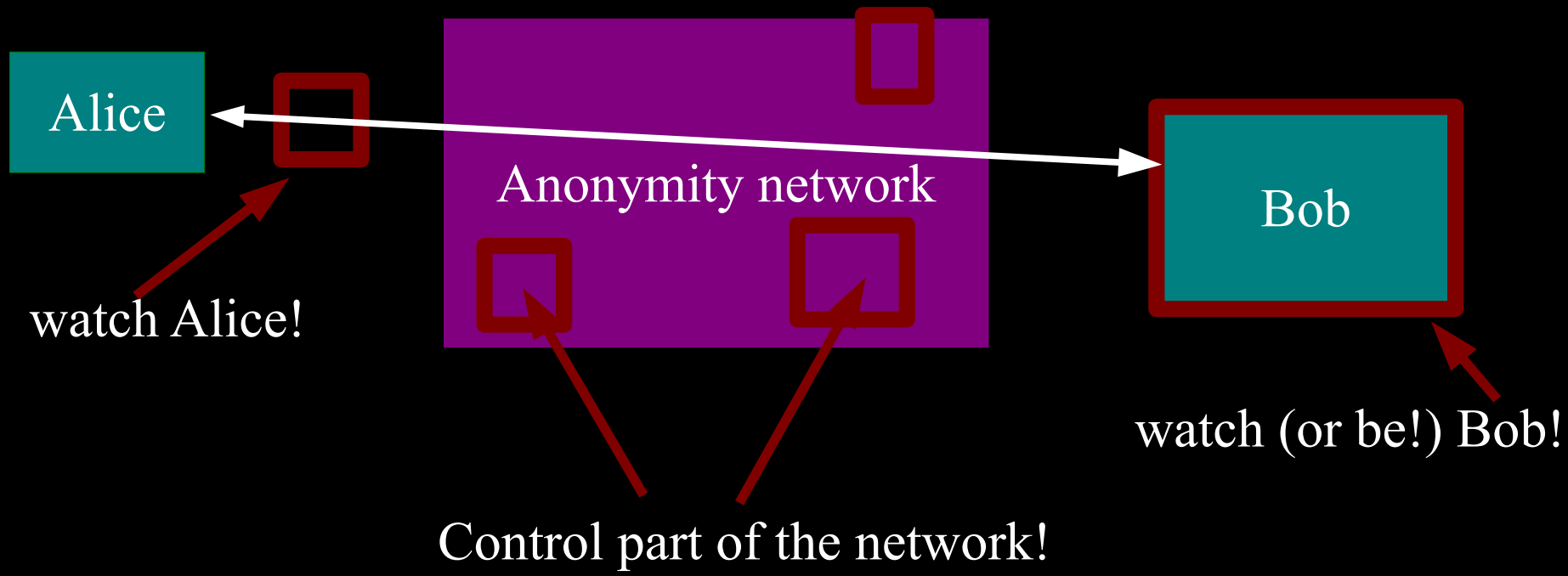
Not secretly evil.


Estimated ~250,000?
daily Tor users

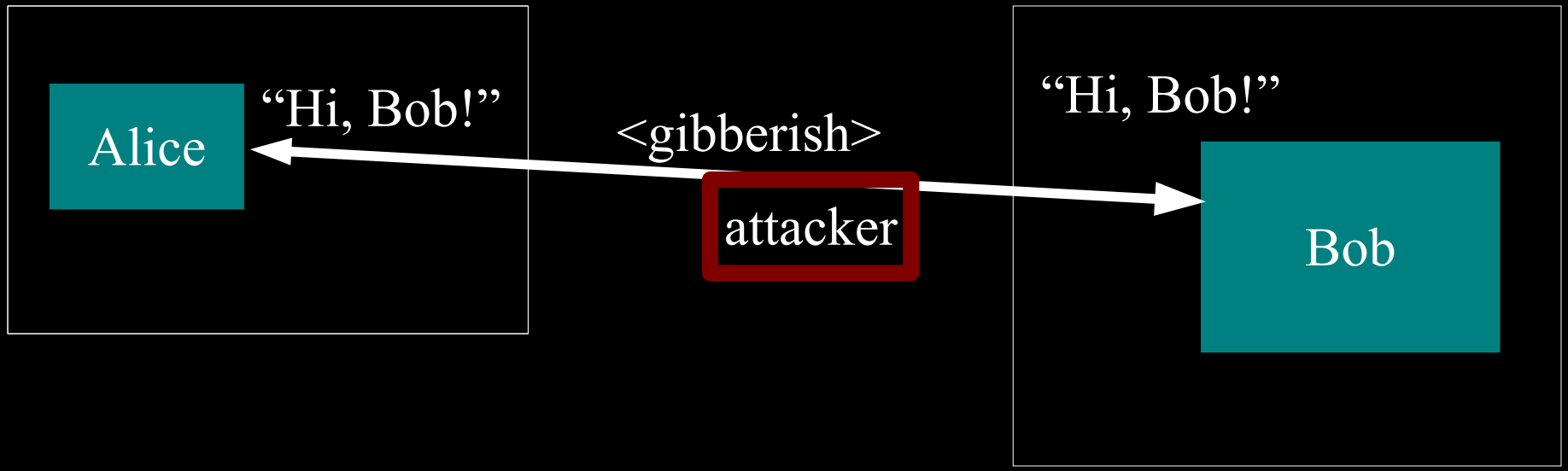
Anonymity in what sense? "Attacker can't learn who is talking to whom."



Threat model: what can the attacker do?



Anonymity isn't cryptography: Cryptography just protects contents.



Anonymity isn't just wishful thinking...

“You can't prove it was me!”

“Promise you won't look!”

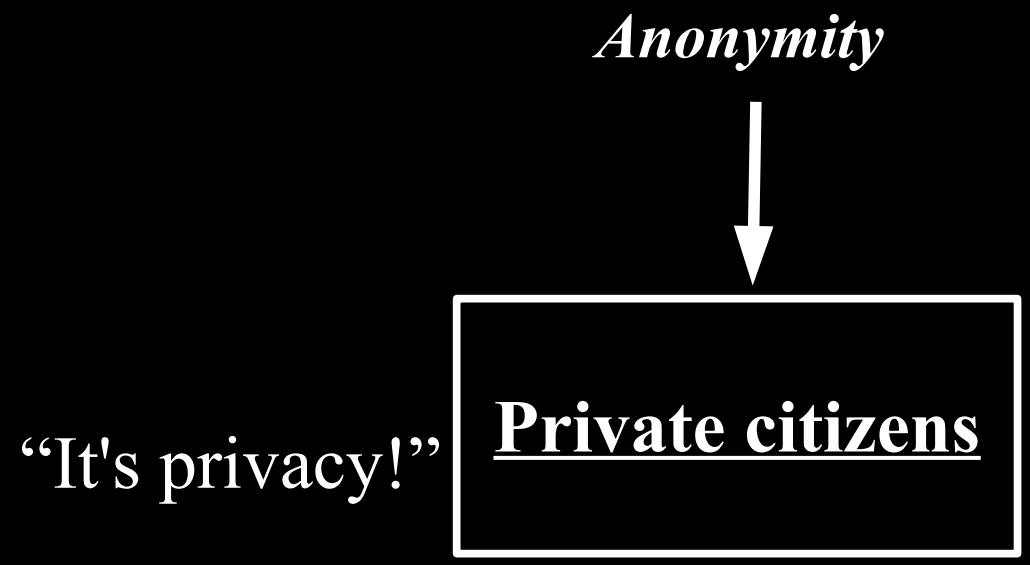
“Promise you won't remember!”

“Promise you won't tell!”

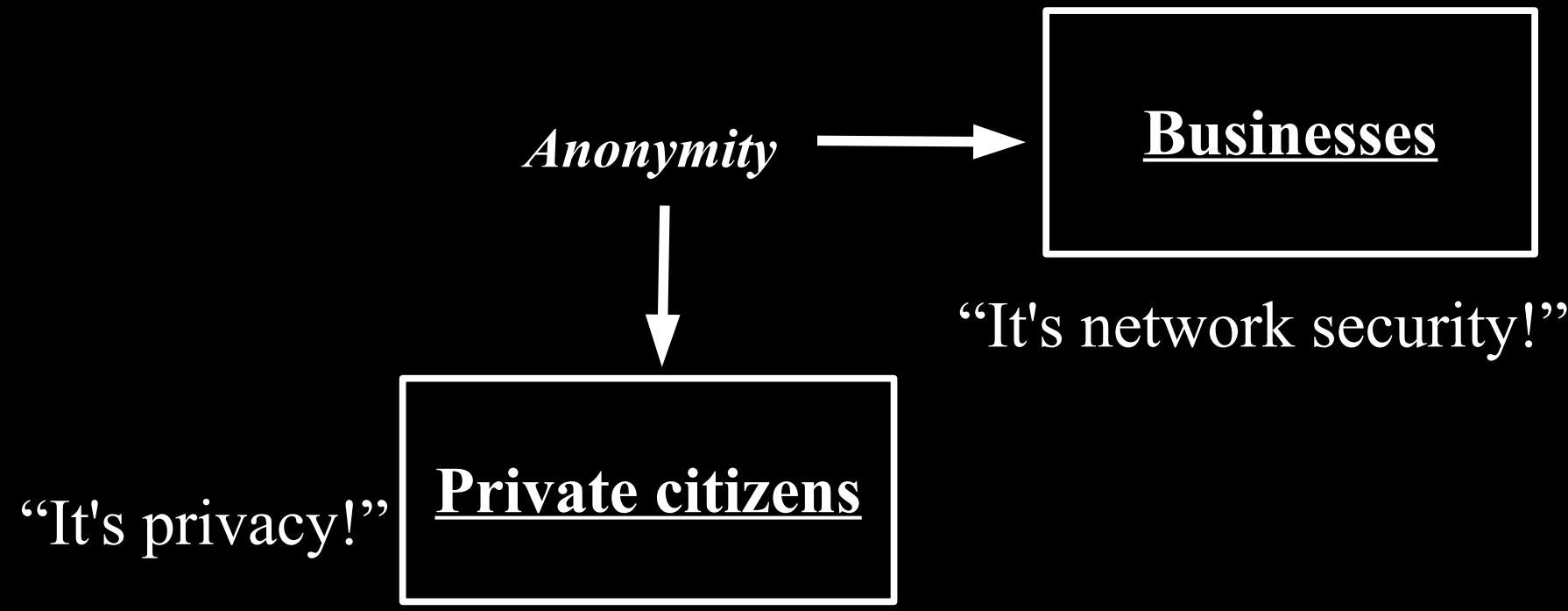
“I didn't write my name on it!”

“Isn't the Internet already anonymous?”

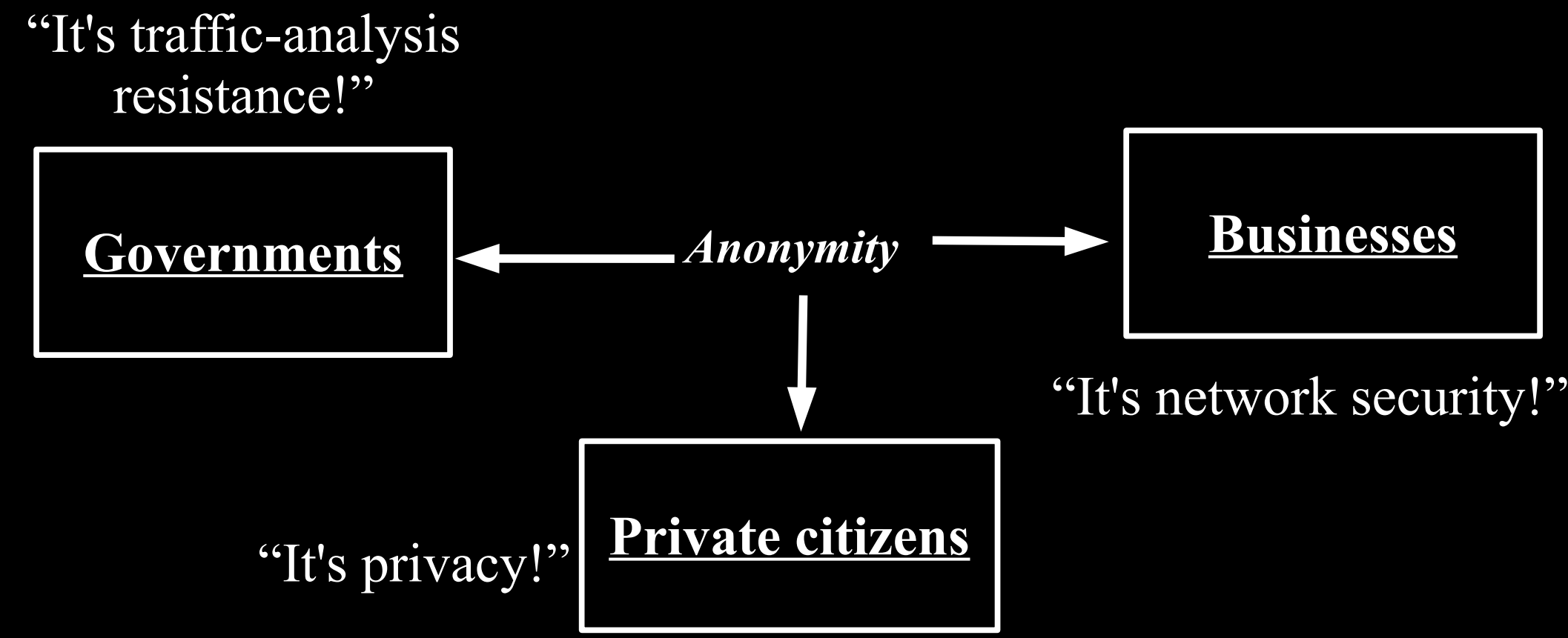
Anonymity serves different interests for different user groups.



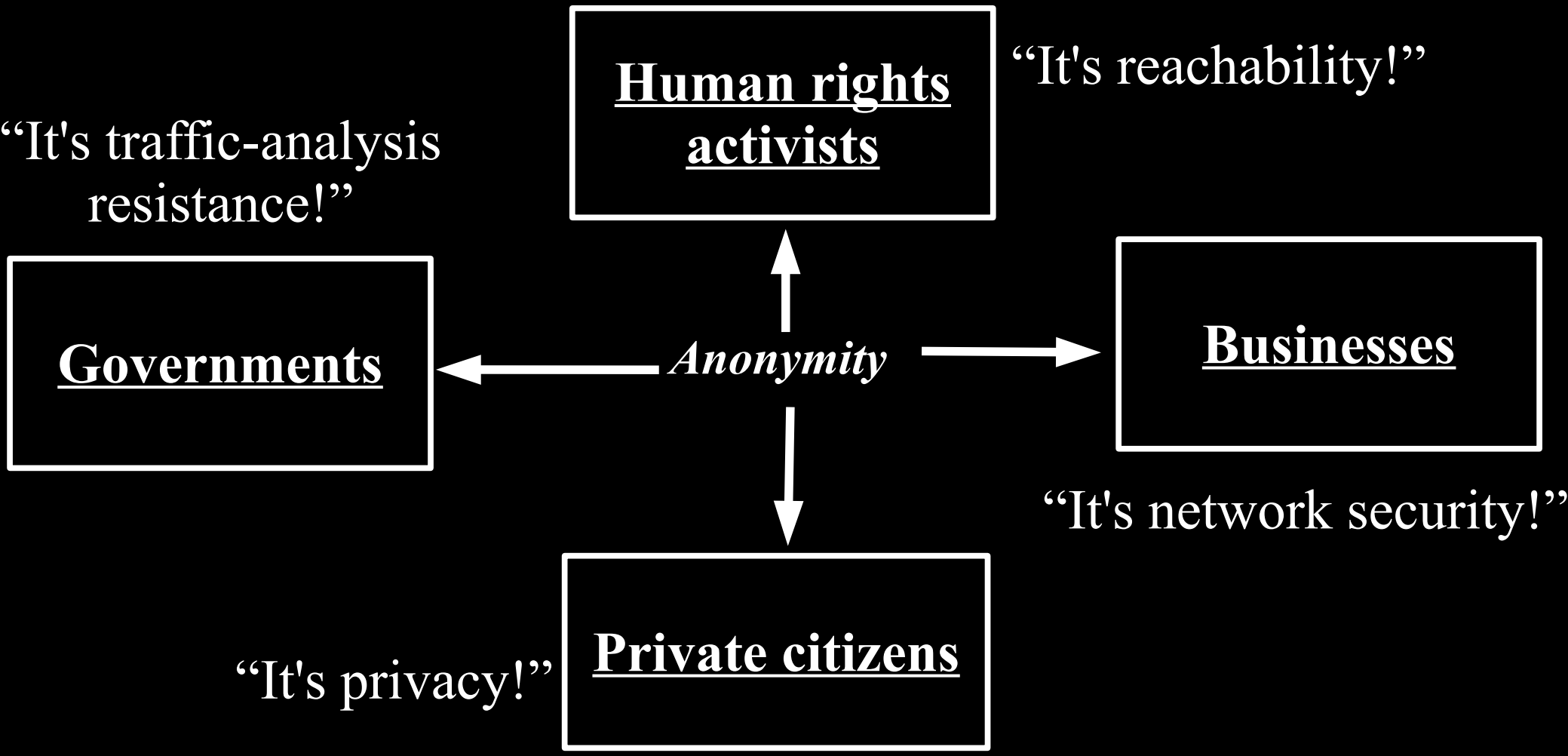
Anonymity serves different interests for different user groups.



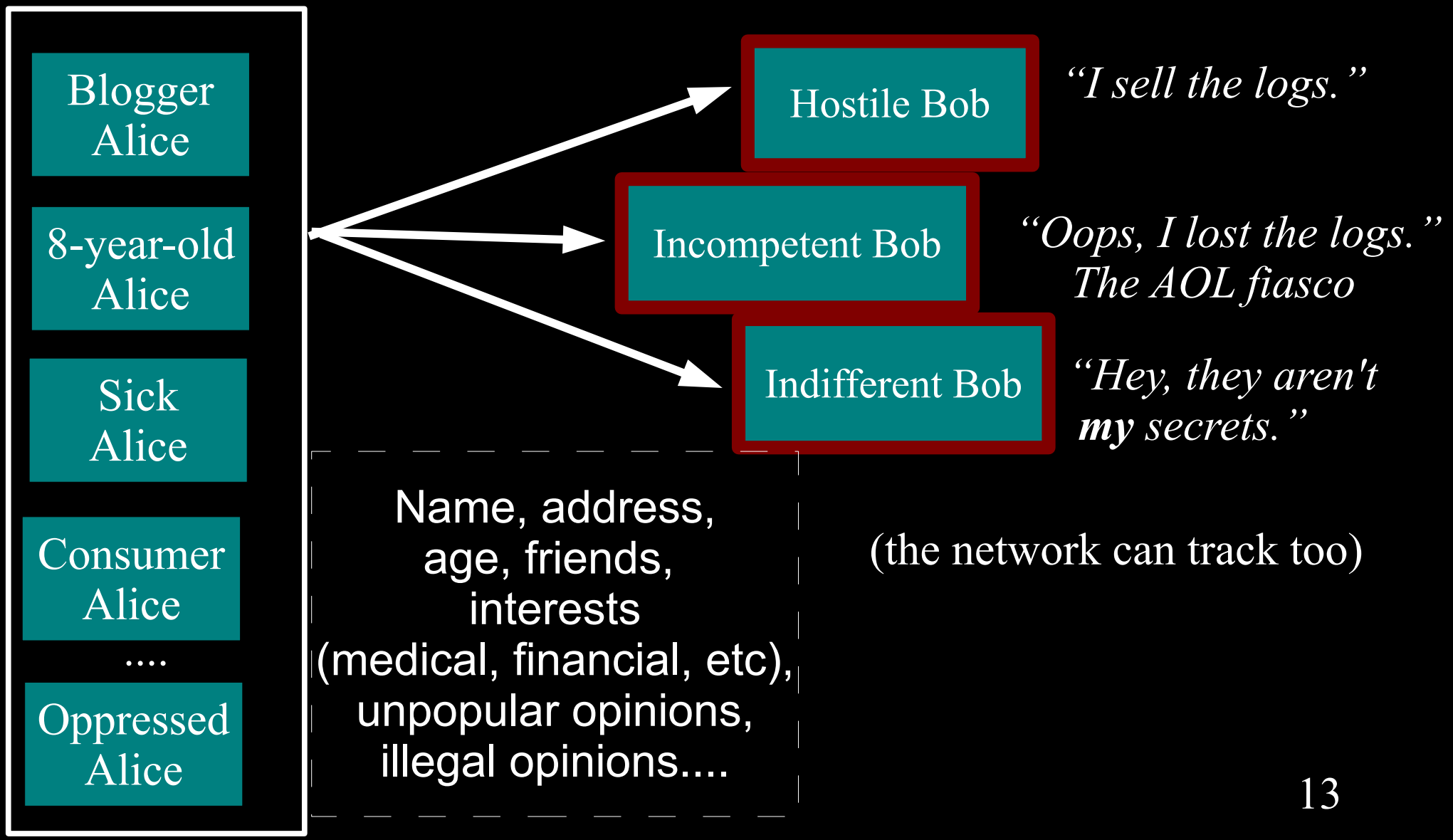
Anonymity serves different interests for different user groups.



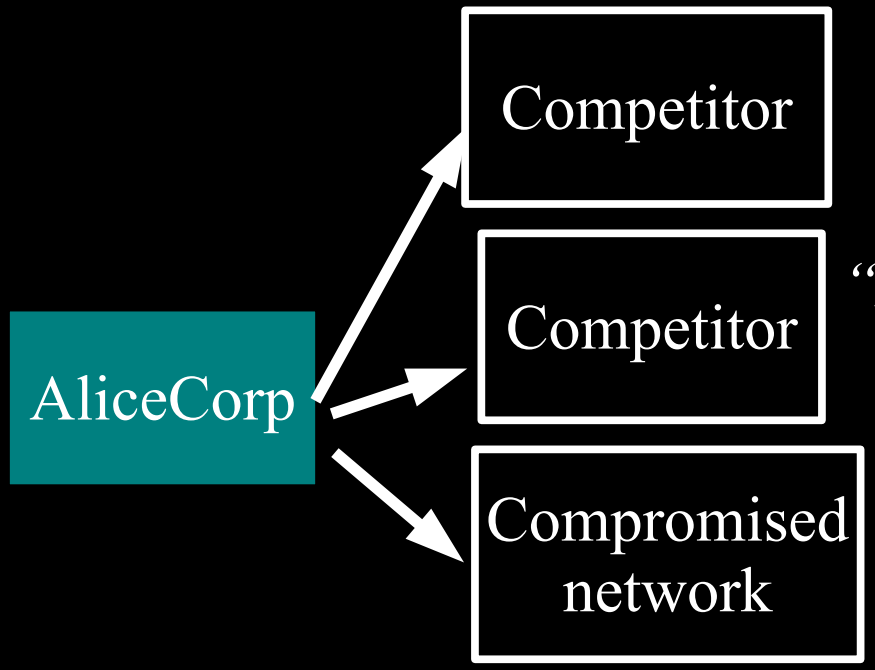
Anonymity serves different interests for different user groups.



Regular citizens don't want to be watched and tracked.



Businesses need to keep trade secrets.

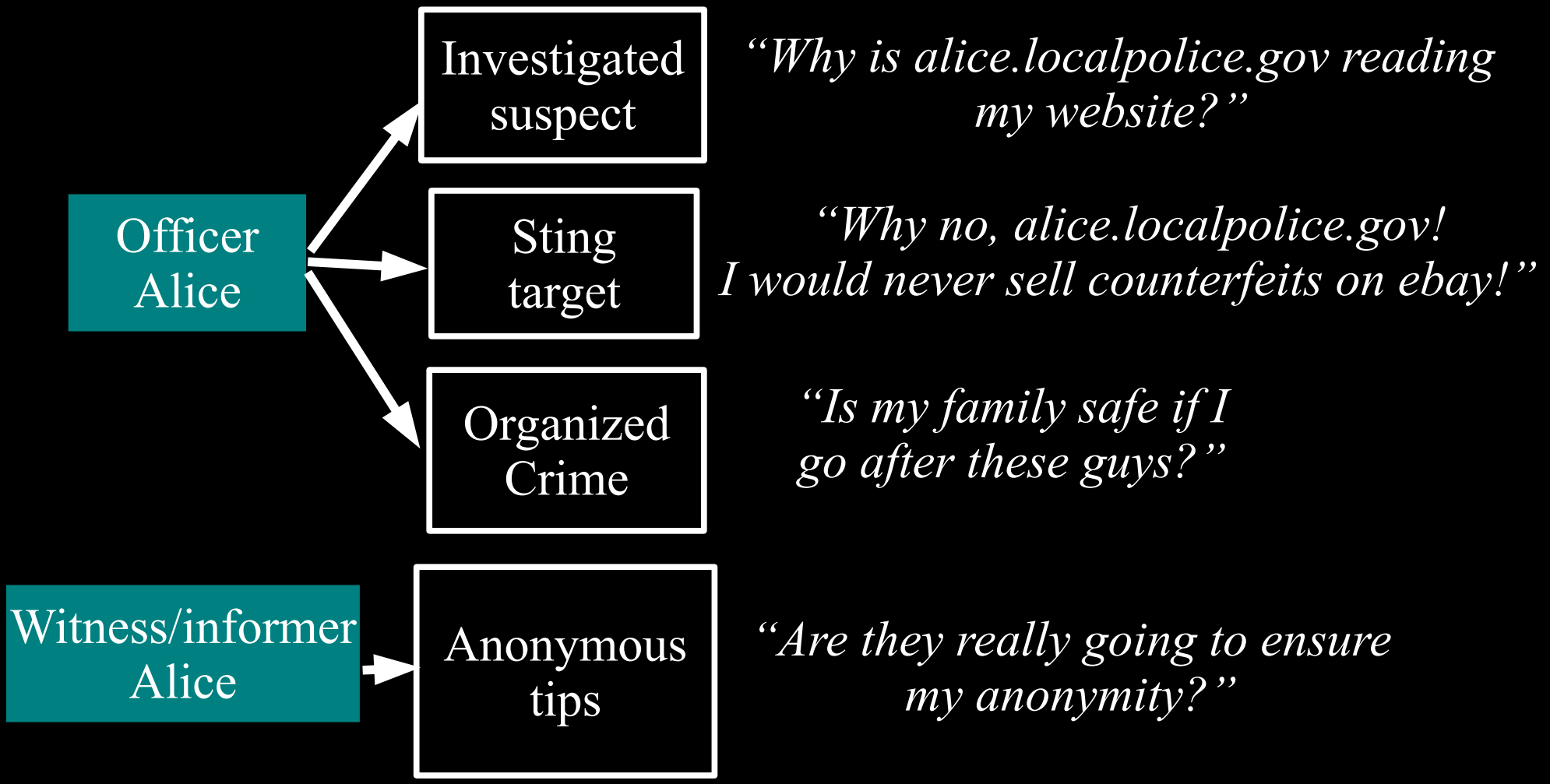


"Oh, your employees are reading our patents/jobs page/product sheets?"

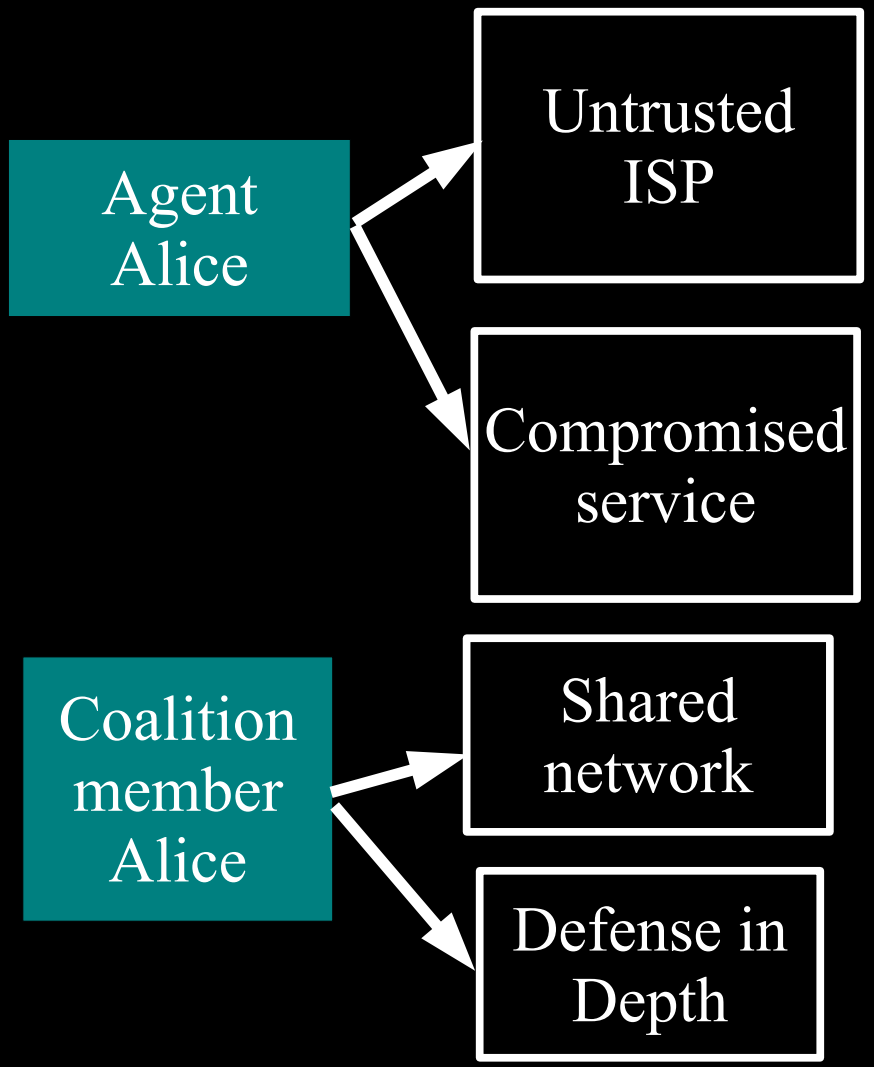
"Hey, it's Alice! Give her the 'Alice' version!"

"Wanna buy a list of Alice's suppliers? What about her customers? What about her engineering department's favorite search terms?"

Law enforcement needs anonymity to get the job done.



Governments need anonymity for their security



"What will you bid for a list of Baghdad IP addresses that get email from .gov?"

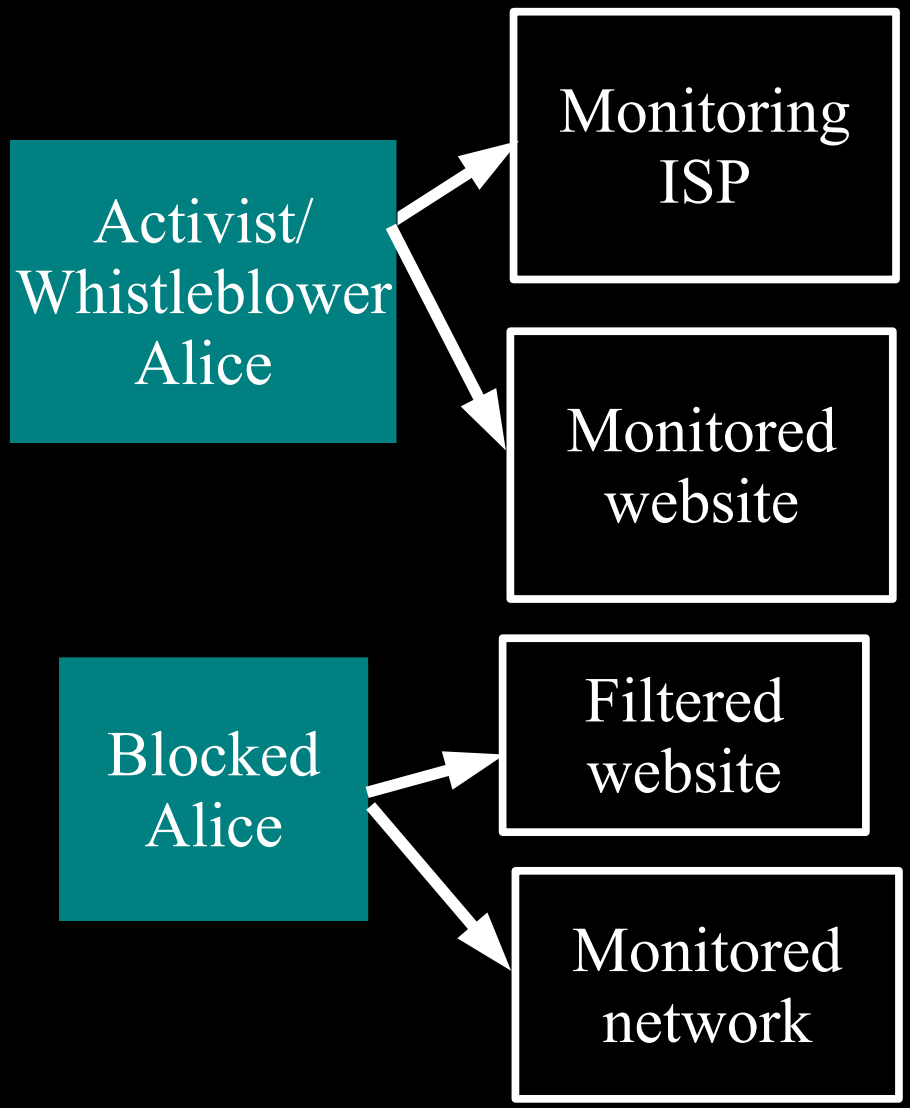
"Somebody in that hotel room just checked his Navy.mil mail!"

"*What* does *FBI Google for?*"

"Do I really want to reveal my internal network topology?"

"What about insiders?"

Journalists and activists need Tor for their personal safety



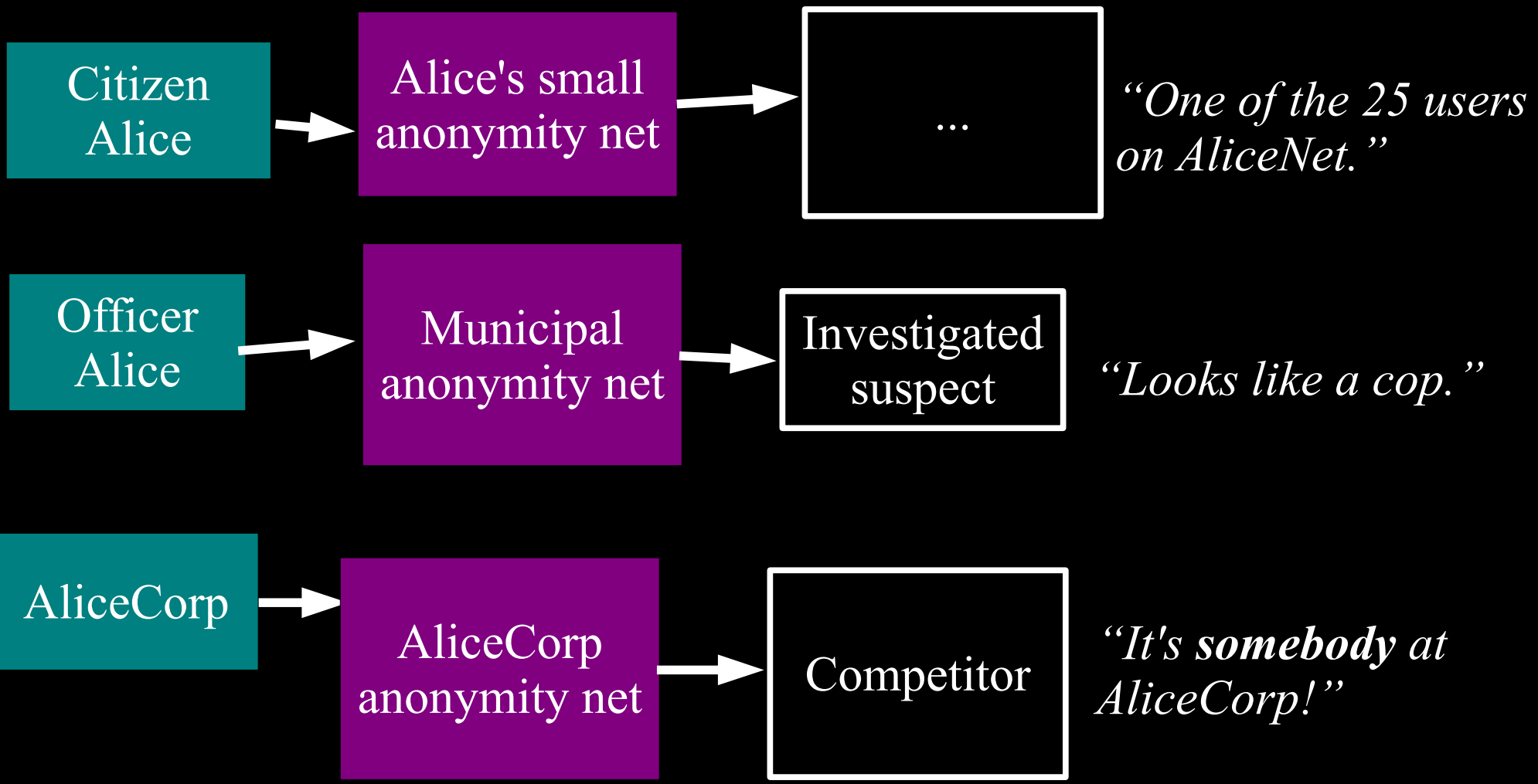
"Did you just post to that website?"

"Where are the bloggers connecting from?"
"I run livejournal and track my users"
"Of course I tell China about my users"

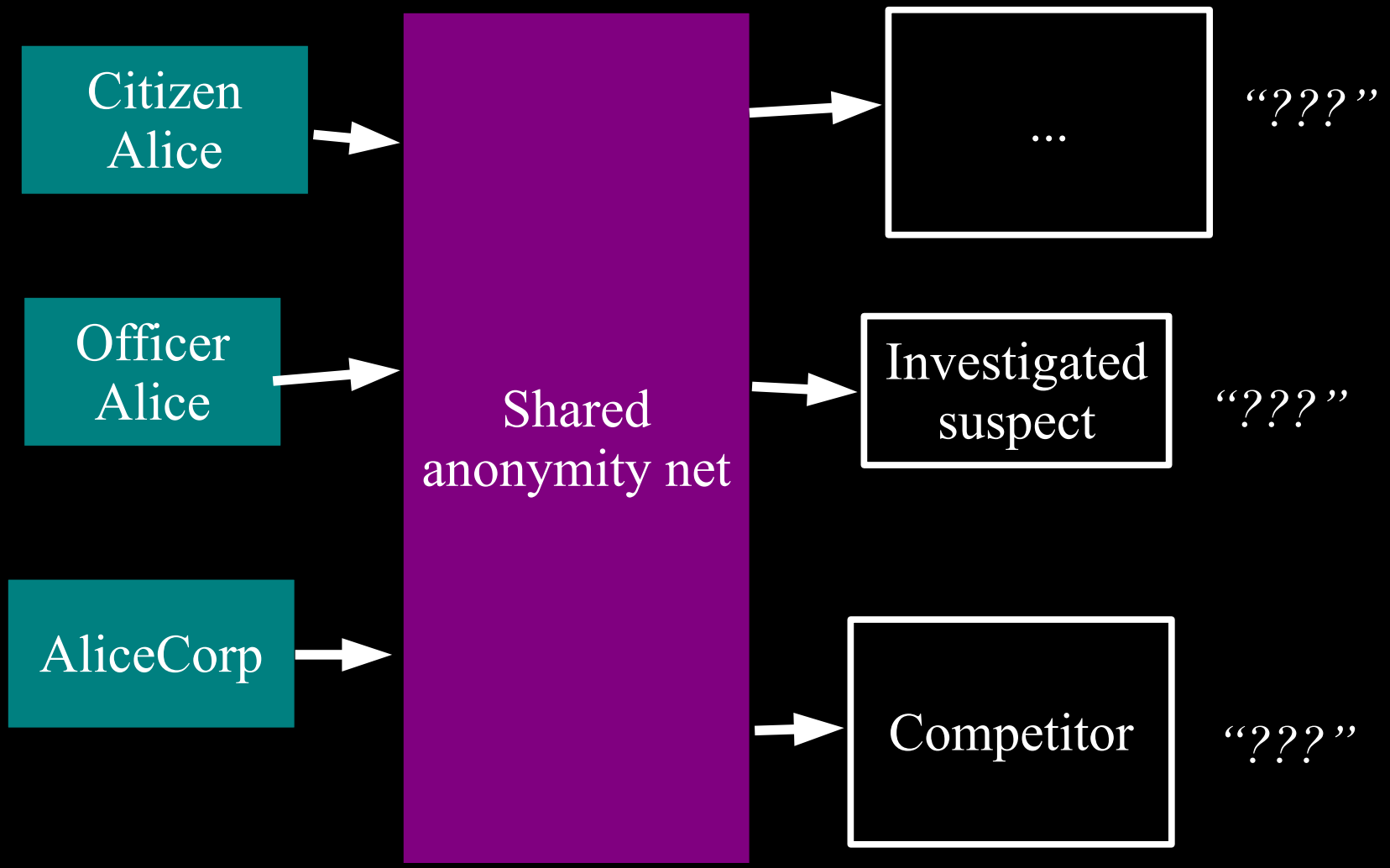
"What does the Global Voices website say today?"
"I want to tell people what's going on in my country"

"I think they're watching. I'm not even going to try."

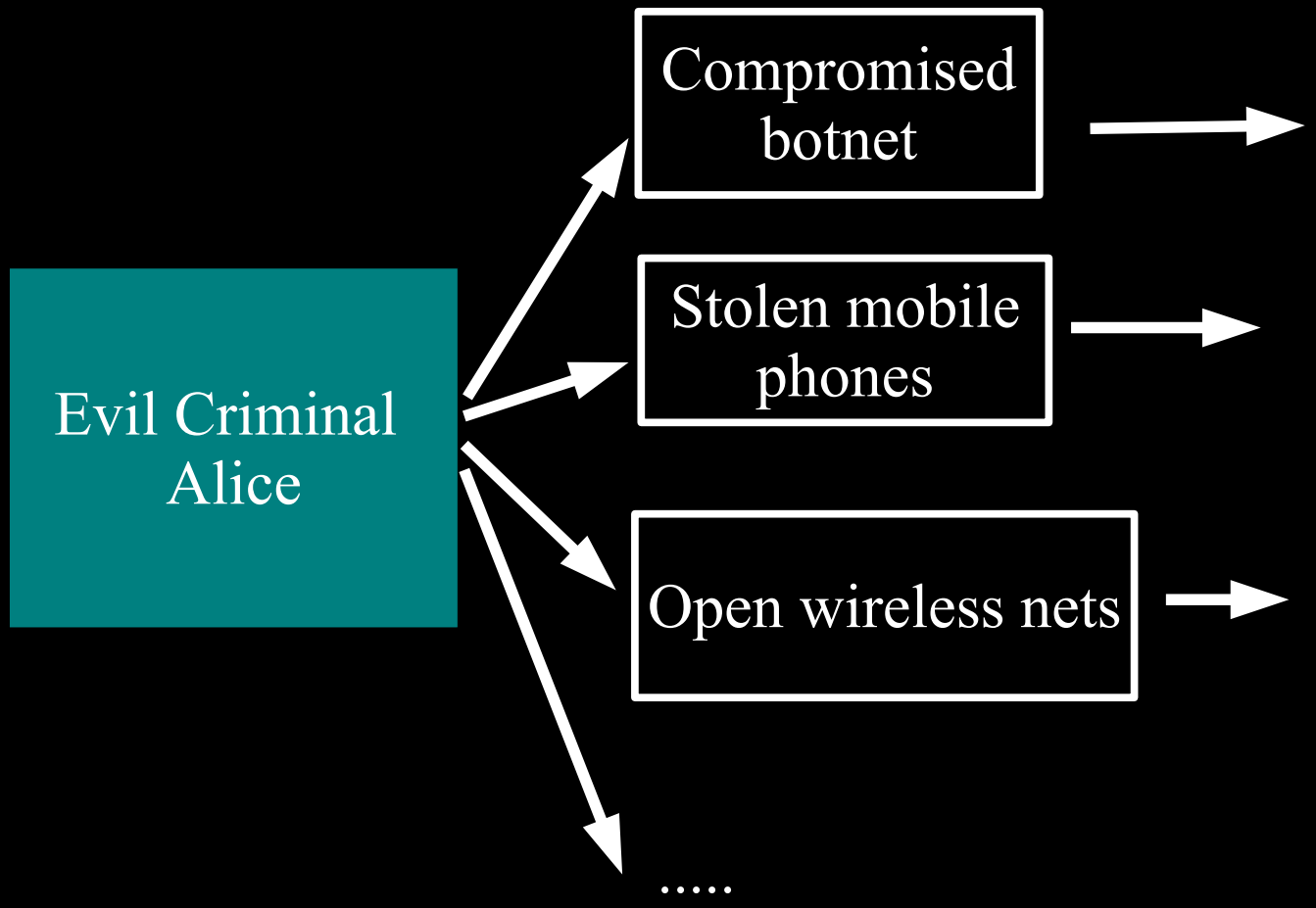
You can't get anonymity on your own: private solutions are ineffective...



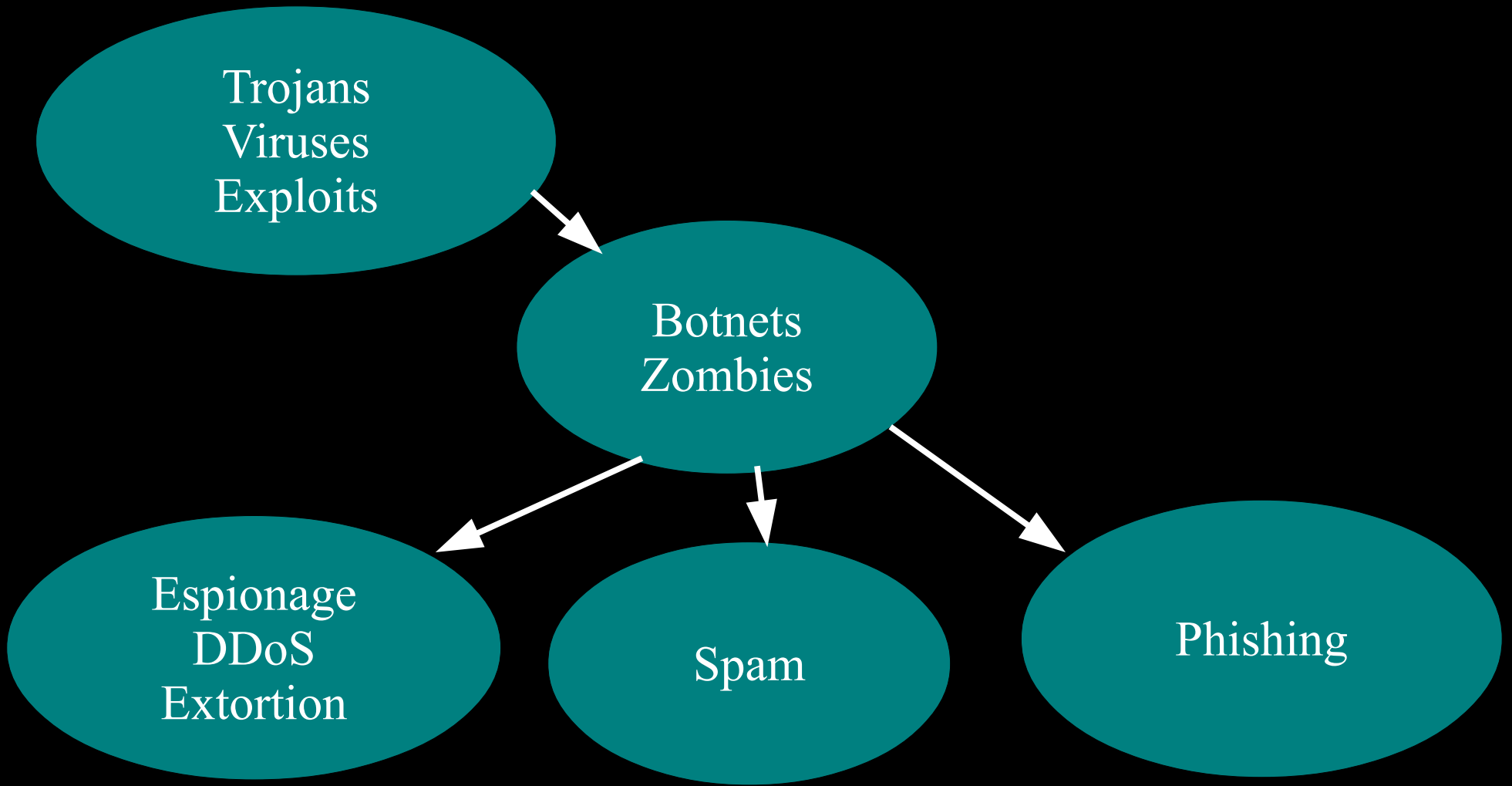
... so, anonymity loves company!



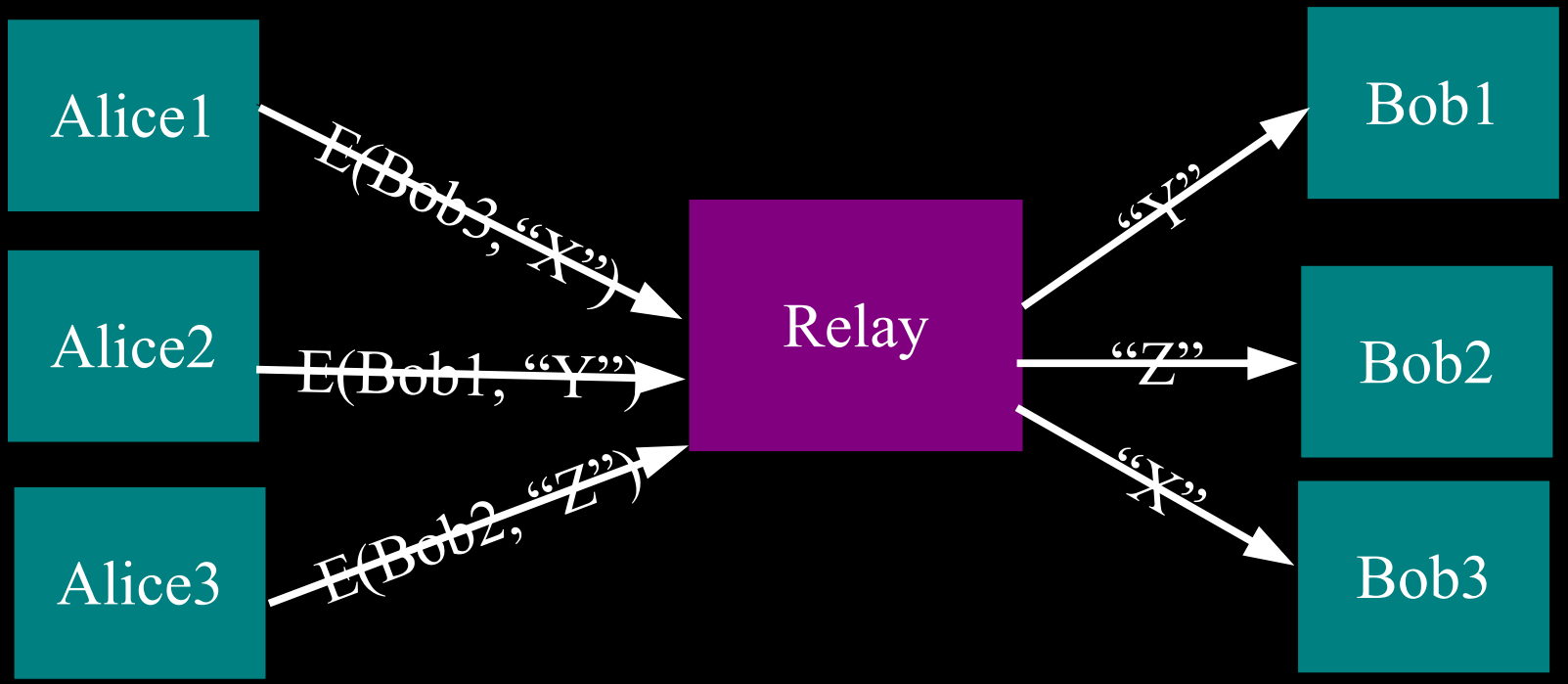
Yes, bad people need anonymity too. But they are *already* doing well.



Current situation: Bad people on the Internet are doing fine

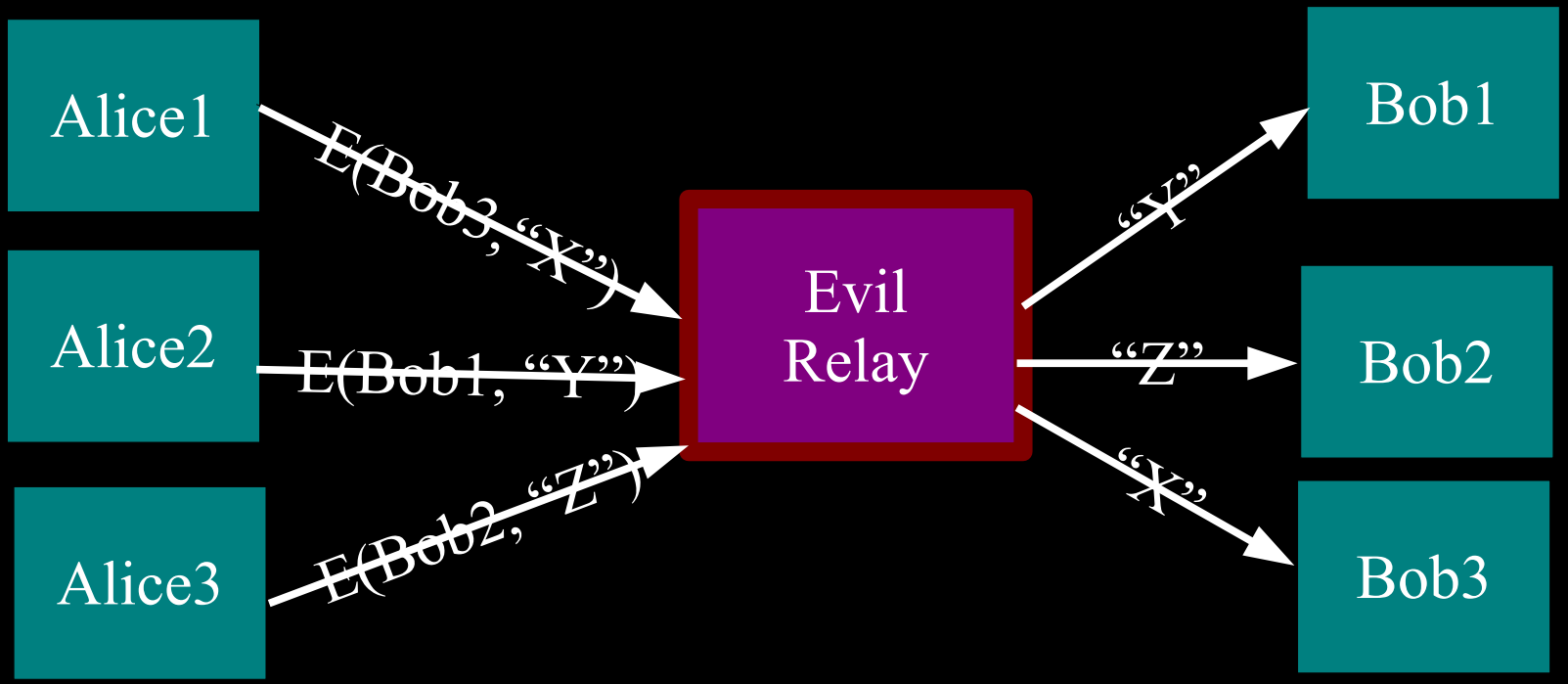


The simplest designs use a single relay to hide connections.

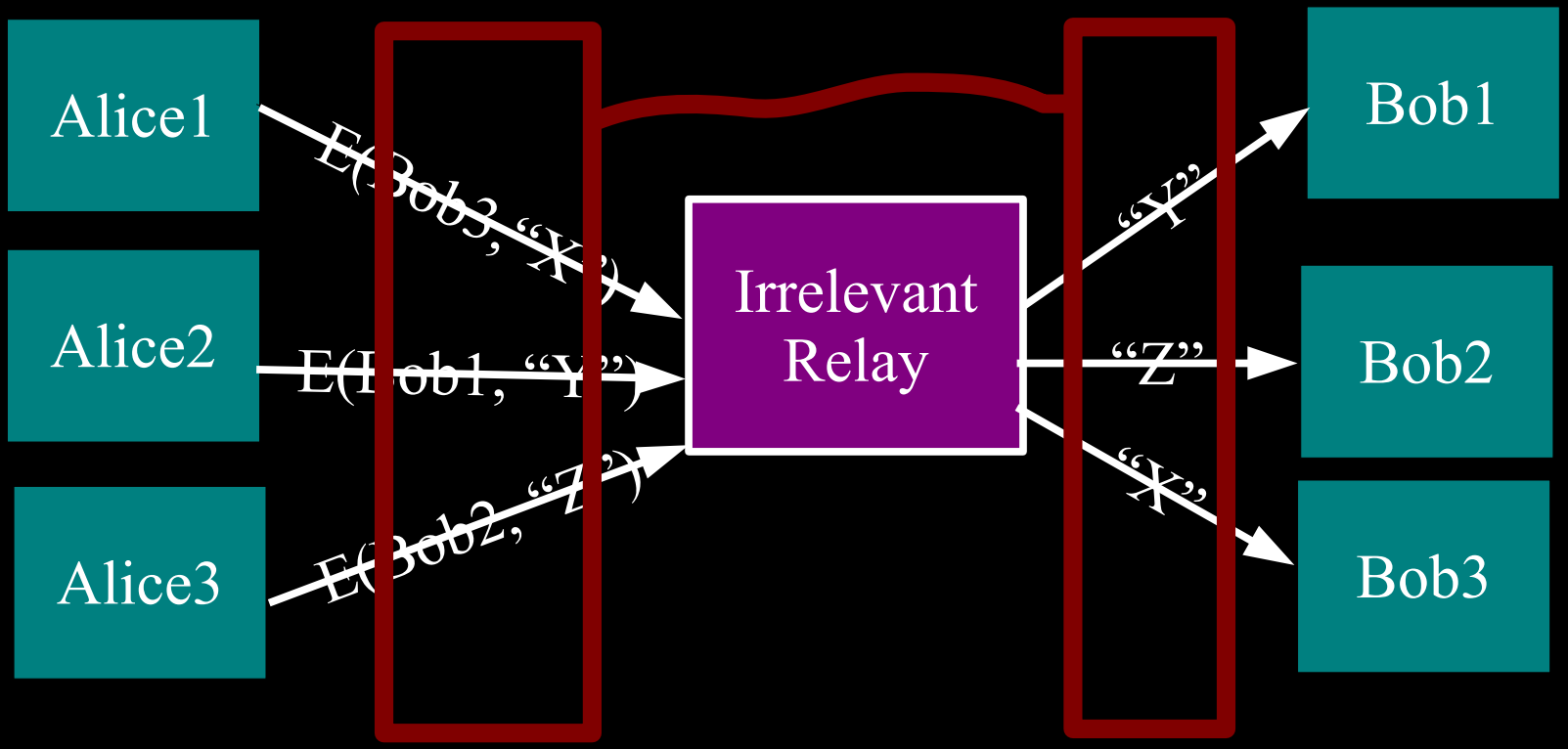


(example: some commercial proxy providers)

But a single relay (or eavesdropper!) is a single point of failure.

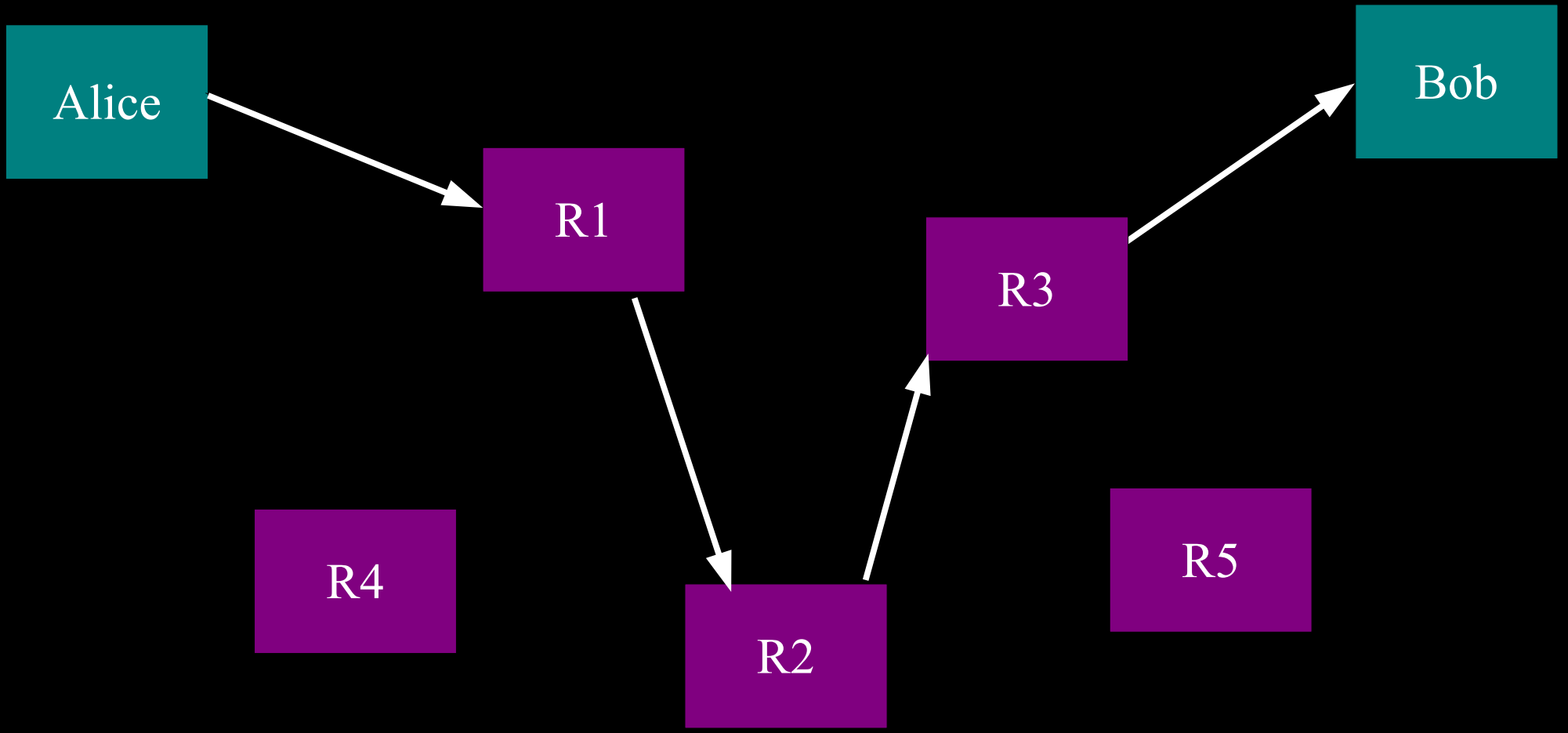


... or a single point of bypass.

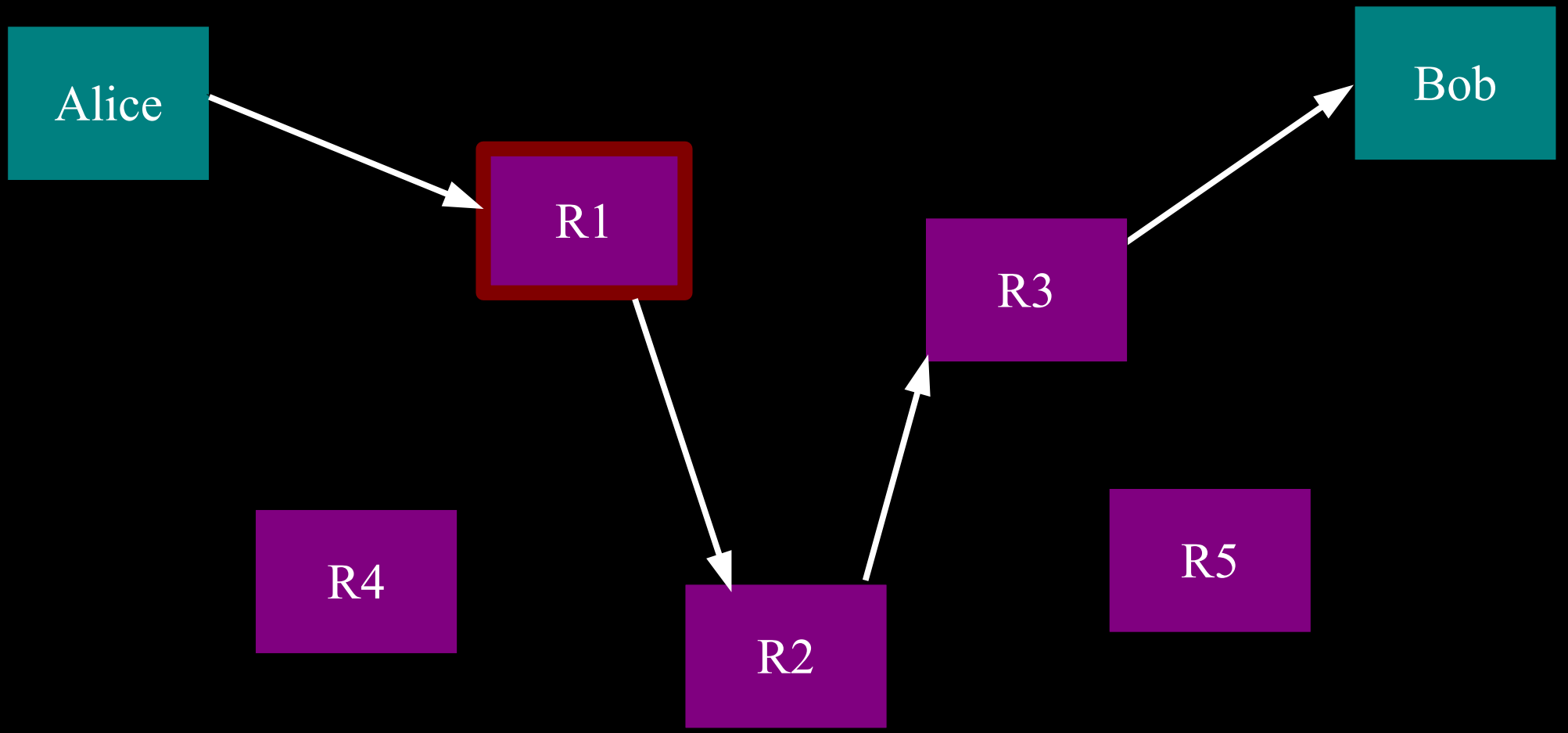


Timing analysis bridges all connections through relay ⇒ An attractive fat target

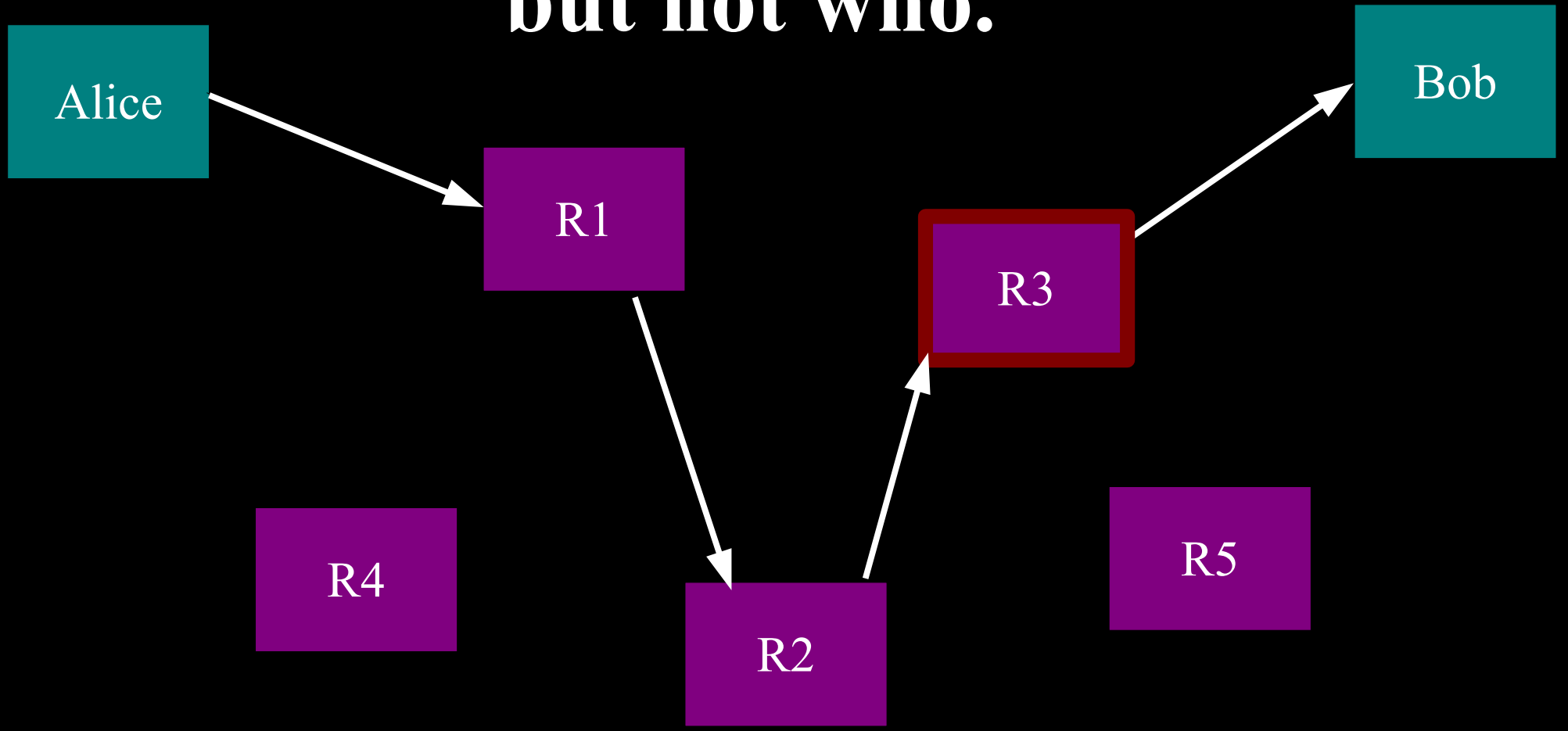
So, add multiple relays so that no single one can betray Alice.



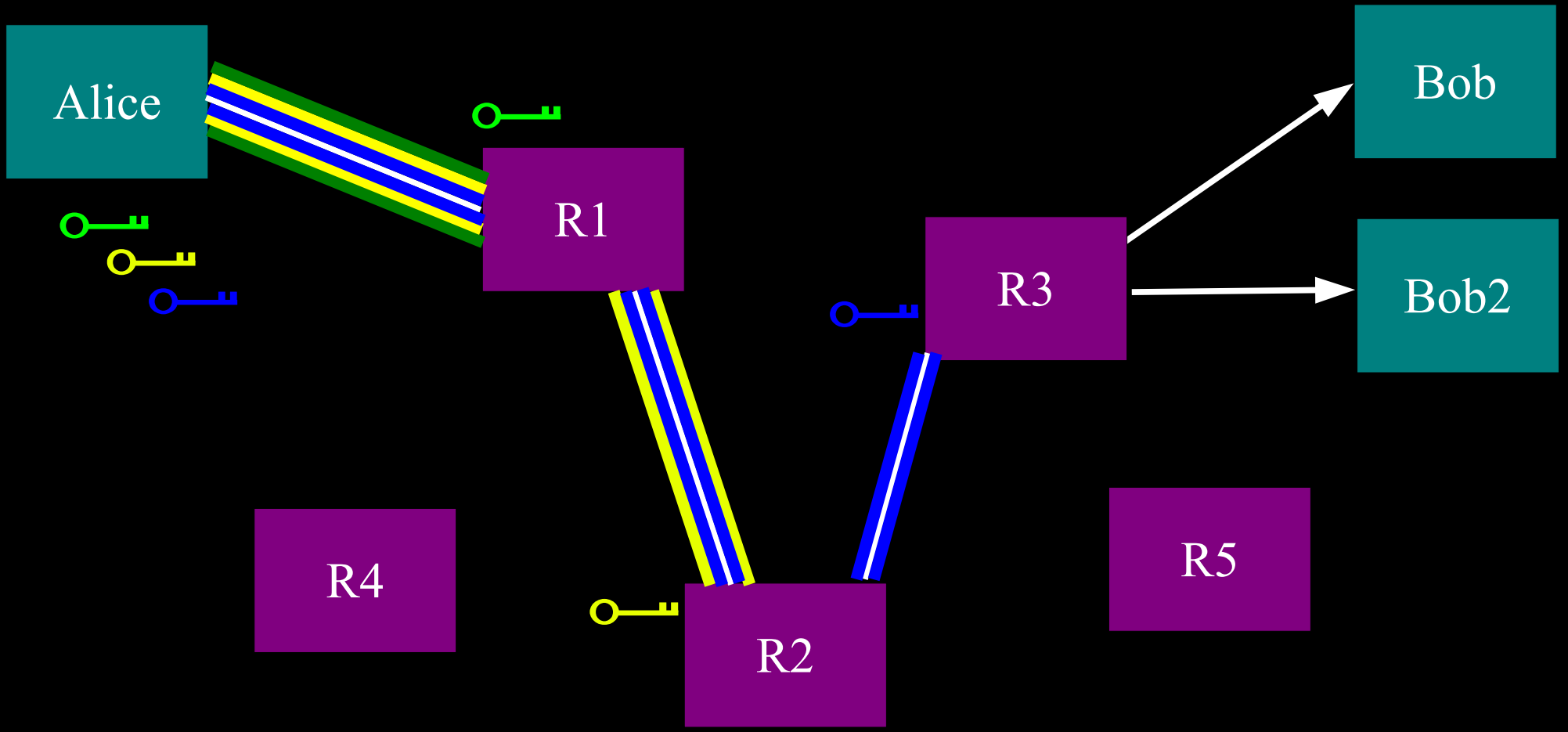
A corrupt first hop can tell that Alice is talking, but not to whom.



A corrupt final hop can tell that somebody is talking to Bob, but not who.



Alice makes a session key with R1 ...And then tunnels to R2...and to R3



What we spend our time on

Performance and scalability

Maintaining the whole software ecosystem

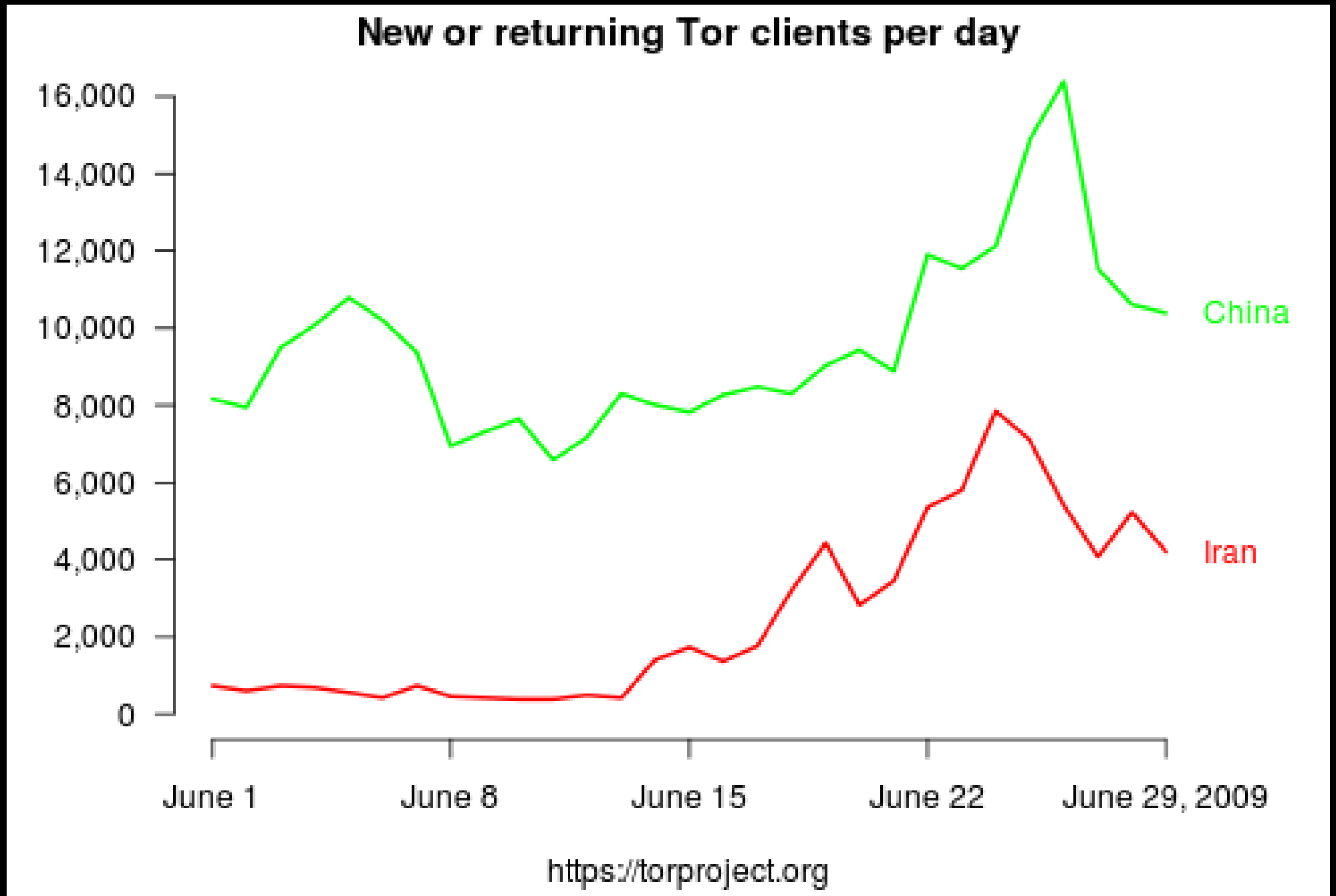
Blocking-resistance (circumvention)

Basic research on anonymity

Reusability and modularity

Advocacy, education, and trainings around the world

Metrics, data, and analysis


New or returning Tor clients per day
16,000
14,000
12,000
10,000
8,000
6,000
4,000
2,000
0
China
Iran
June 1
June 8
June 15
June 22
June 29, 2009
https://torproject.org

Another Iran user count

Talked to chief security officer of one of the web 2.0 social networking sites:

10% (~10k) of their Iranian users in June 2009 were coming through Tor

90% (~90k) were coming from proxies in the Amazon cloud

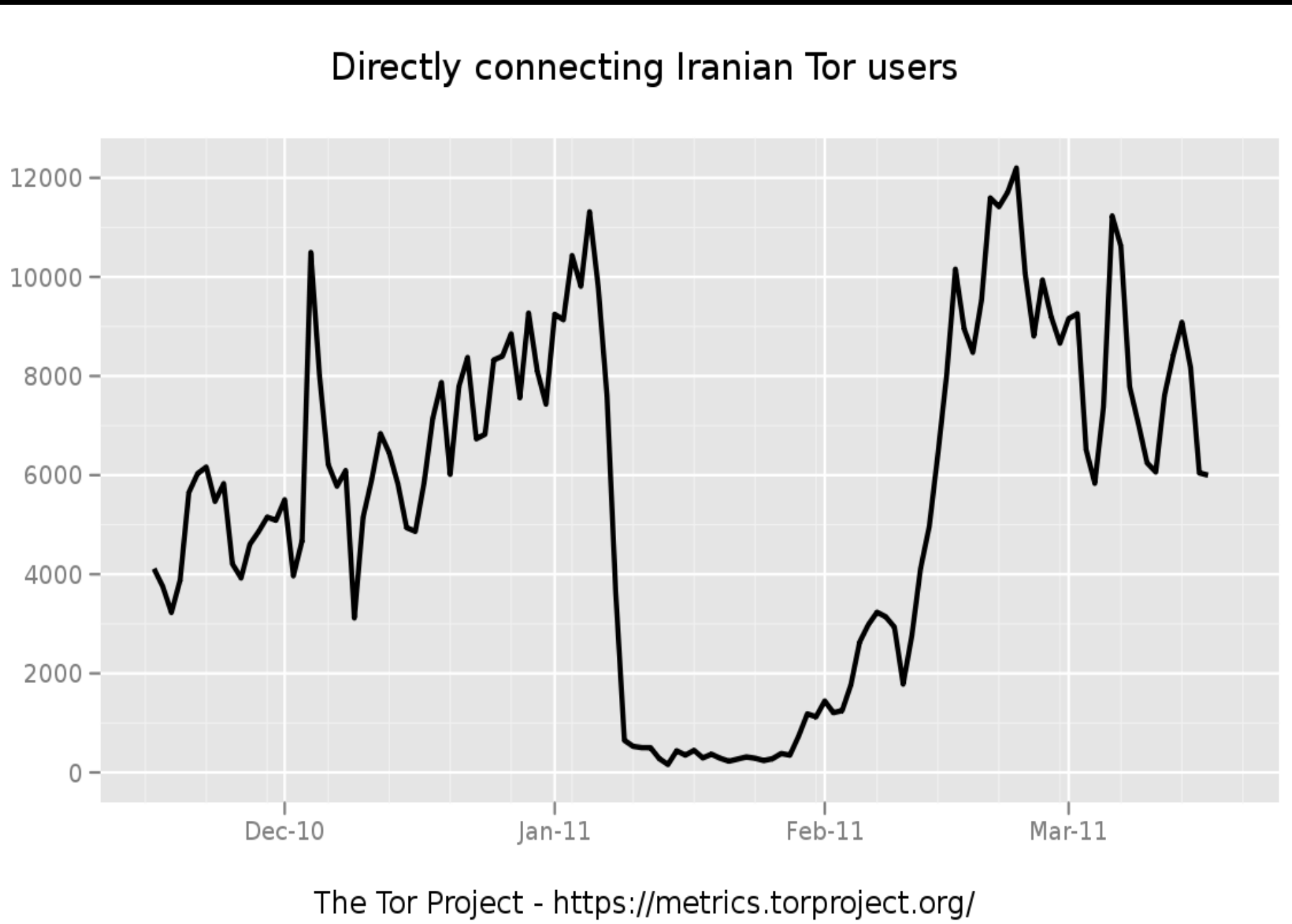
Iran and DPI

We made Tor's TLS handshake look like Firefox+Apache.

When Iran kicked out Smartfilter in early 2009, Tor's old (non-TLS) directory fetches worked again!

Jan 2011, Iran blocked Tor by DPI for SSL and filtering our Diffie-Hellman parameter.

Socks proxy worked fine the whole time.


Directly connecting Iranian Tor users
12000
10000
8000
6000
4000
2000
0
Dec-10
Jan-11
Feb-11
Mar-11
The Tor Project - https://metrics.torproject.org/

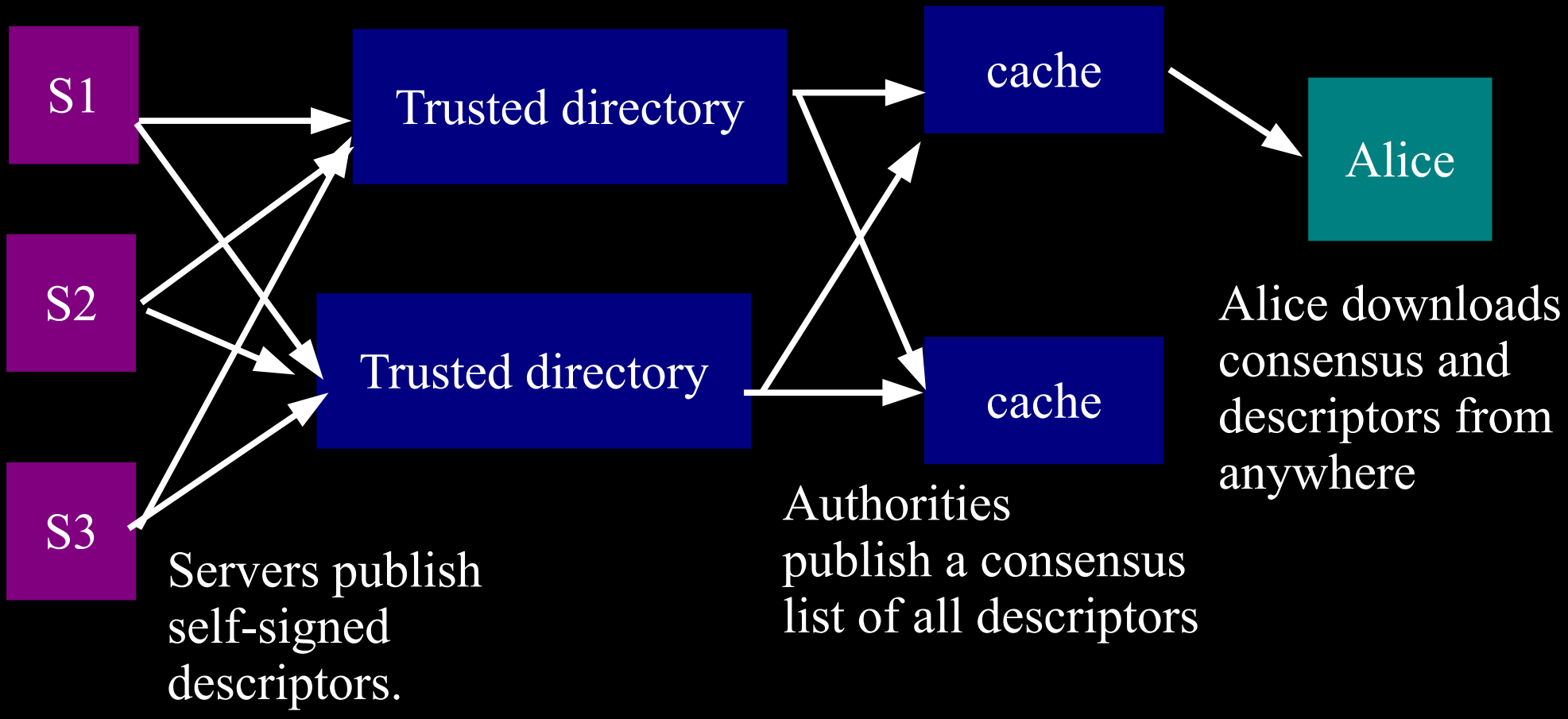
Relay versus Discovery

There are two pieces to all these “proxying” schemes:

a **relay** component: building circuits, sending traffic over them, getting the crypto right

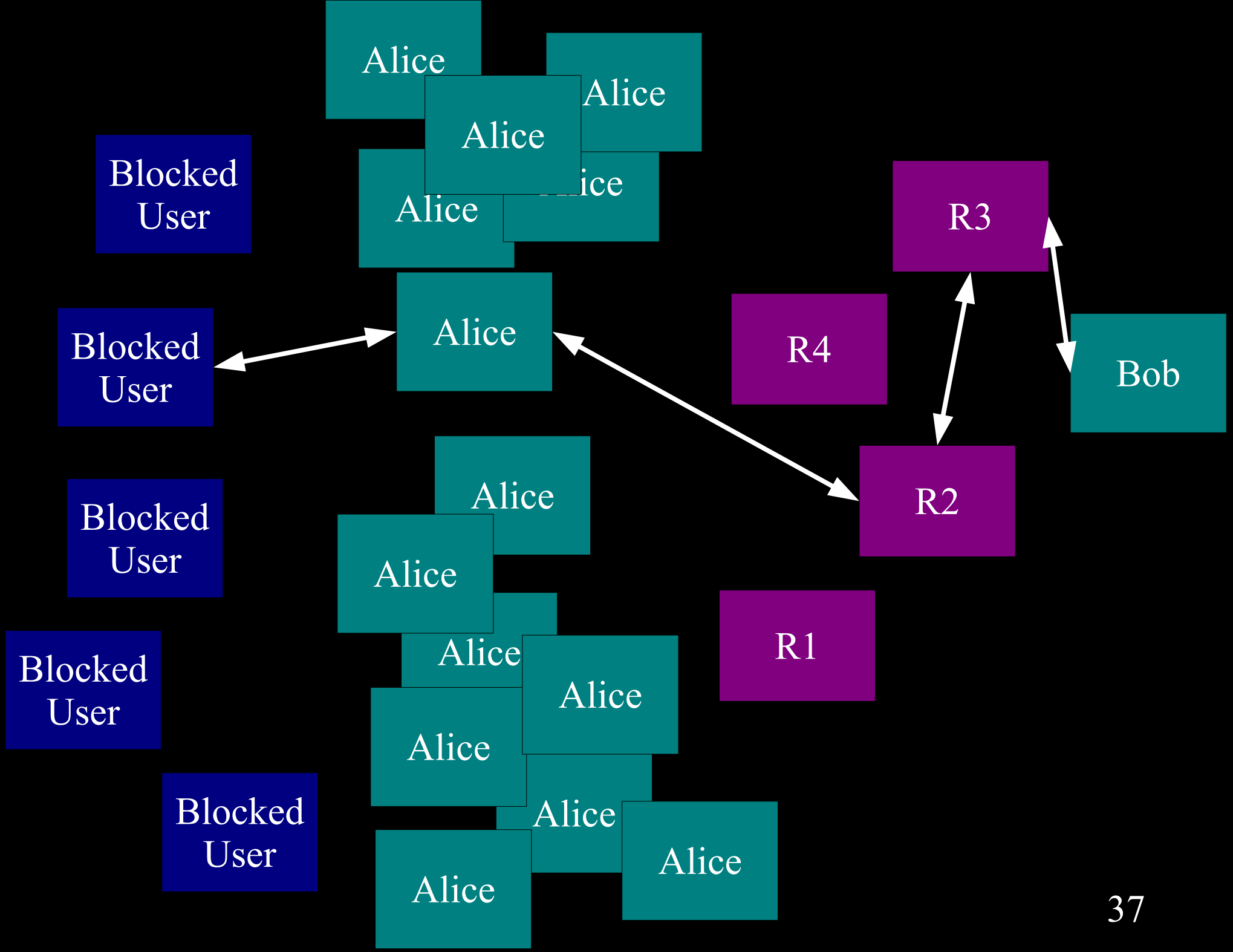
a **discovery** component: learning what relays are available

The basic Tor design uses a simple centralized directory protocol.



Attackers can block users from connecting to the Tor network

- By blocking the directory authorities
- By blocking all the relay IP addresses in the directory
- By filtering based on Tor's network fingerprint
- By preventing users from finding the Tor software


Alice
Alice
Alice
Alice
Alice
Blocked User
Blocked User
Blocked User
Blocked User
Blocked User
R3
R4
R2
R1
Bob
Alice
Alice
Alice
Alice
Alice
Alice
Alice
Alice

“Bridge” relays

Hundreds of thousands of Tor users, already self-selected for caring about privacy.

Rather than signing up as a normal relay, you can sign up as a special “bridge” relay that isn't listed in any directory.

No need to be an “exit” (so no abuse worries), and you can rate limit if needed

Integrated into Vidalia (our GUI) so it's easy to offer a bridge or to use a bridge

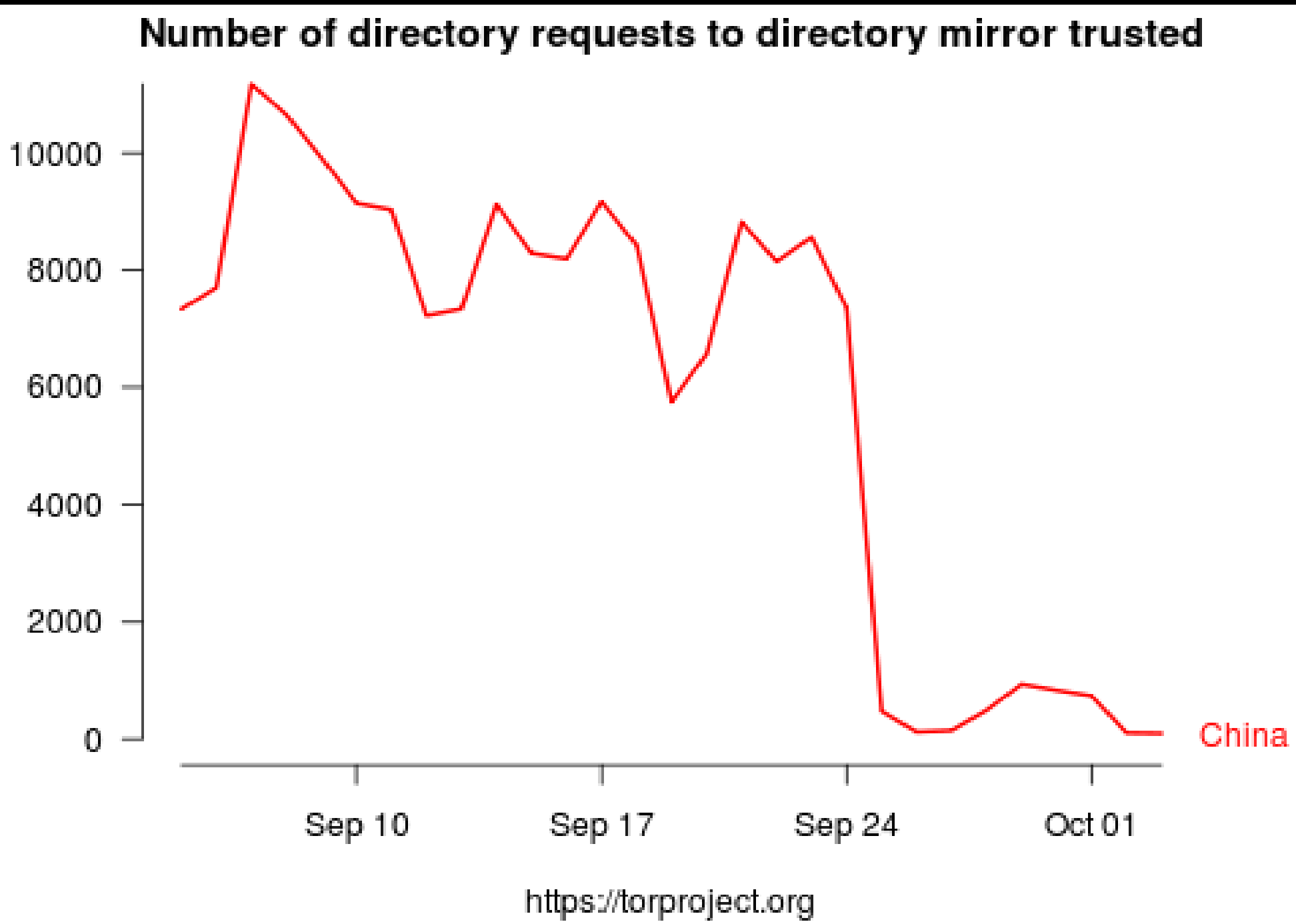
How do you f nd a bridge?

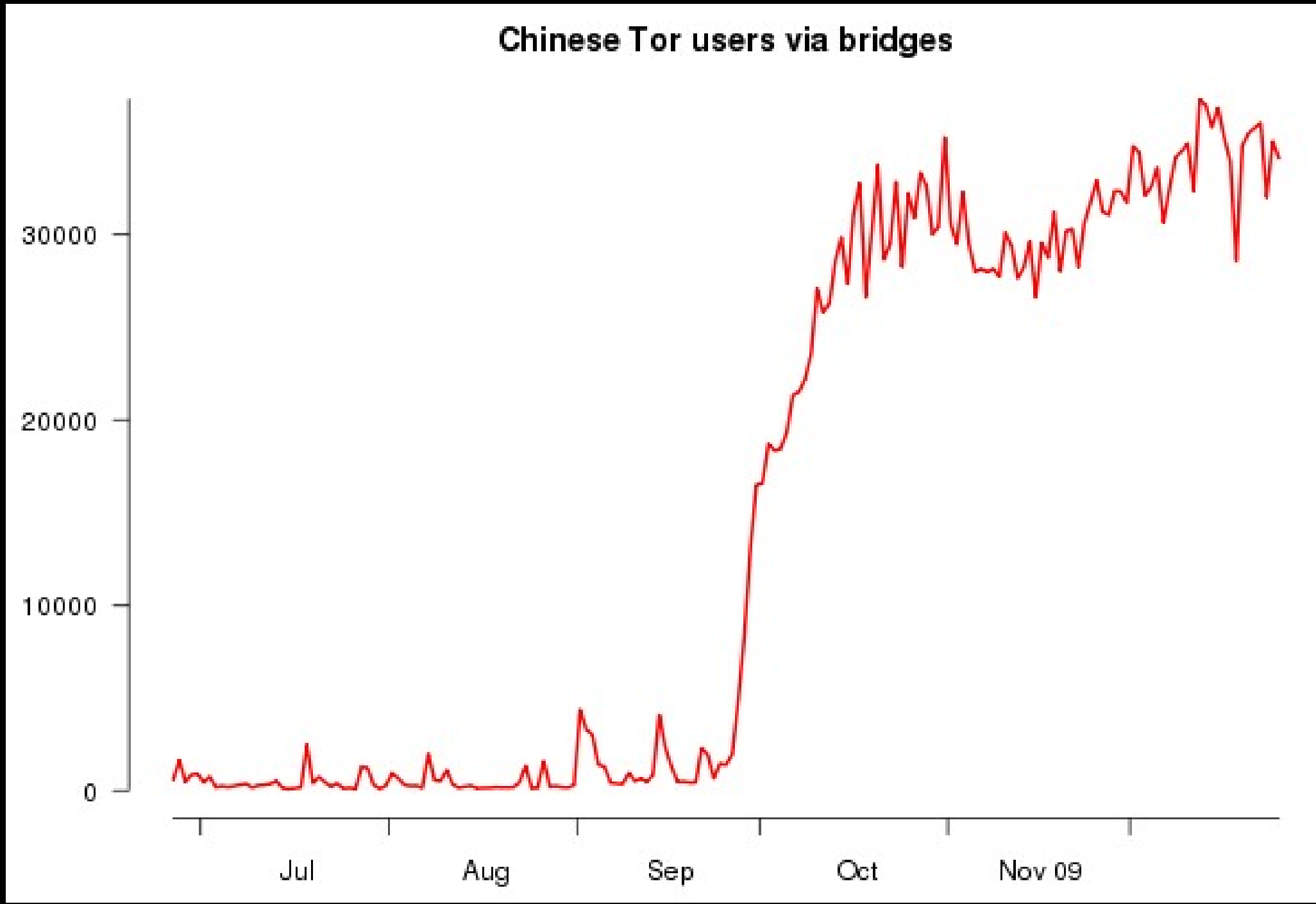
1) **https://bridges.torproject.org/** will tell you a few based on time and your IP address

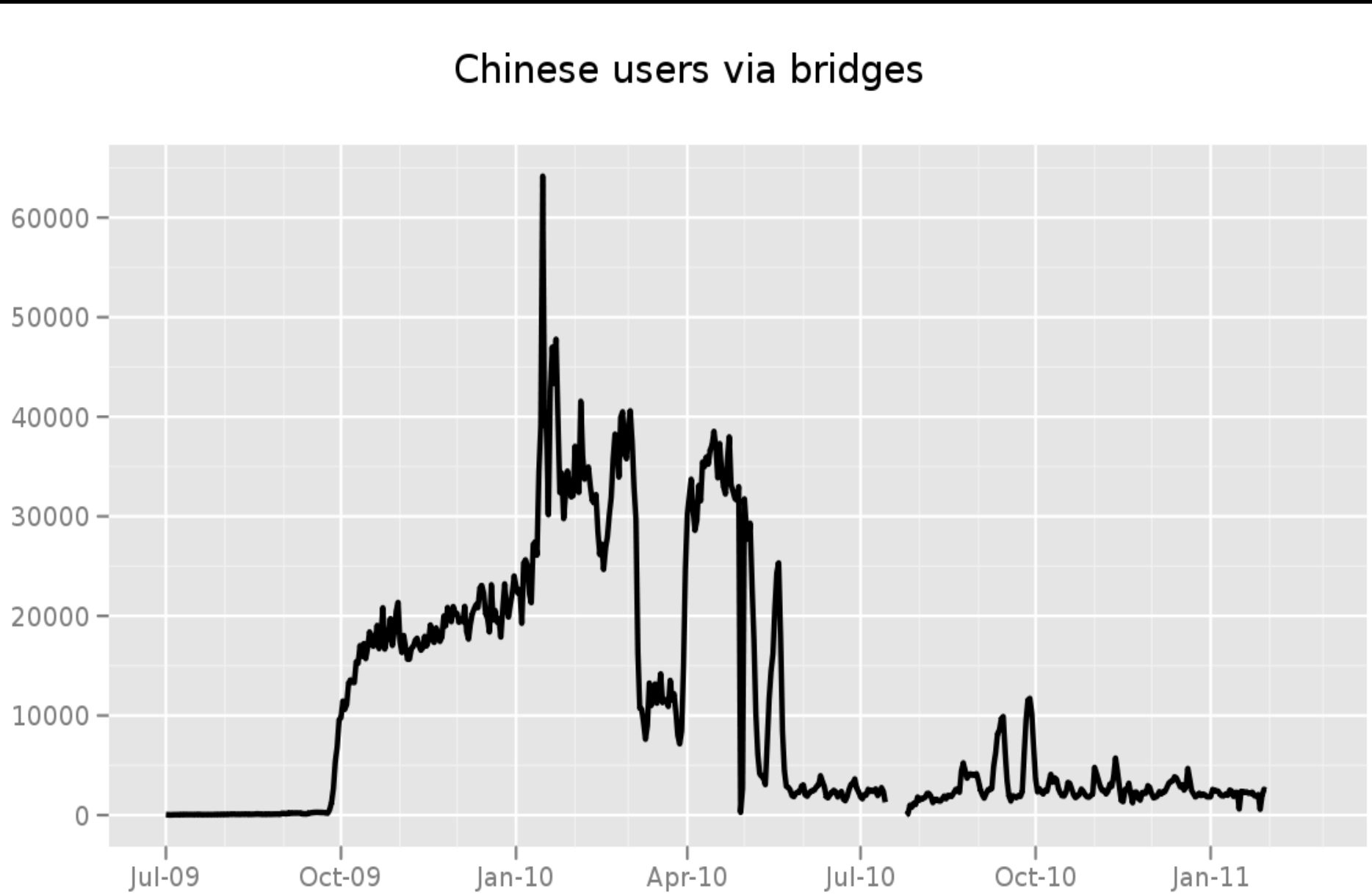
2) Mail bridges@torproject.org from a gmail address and we'll send you a few

3) We mail some to a friend in Shanghai who distributes them via his social network

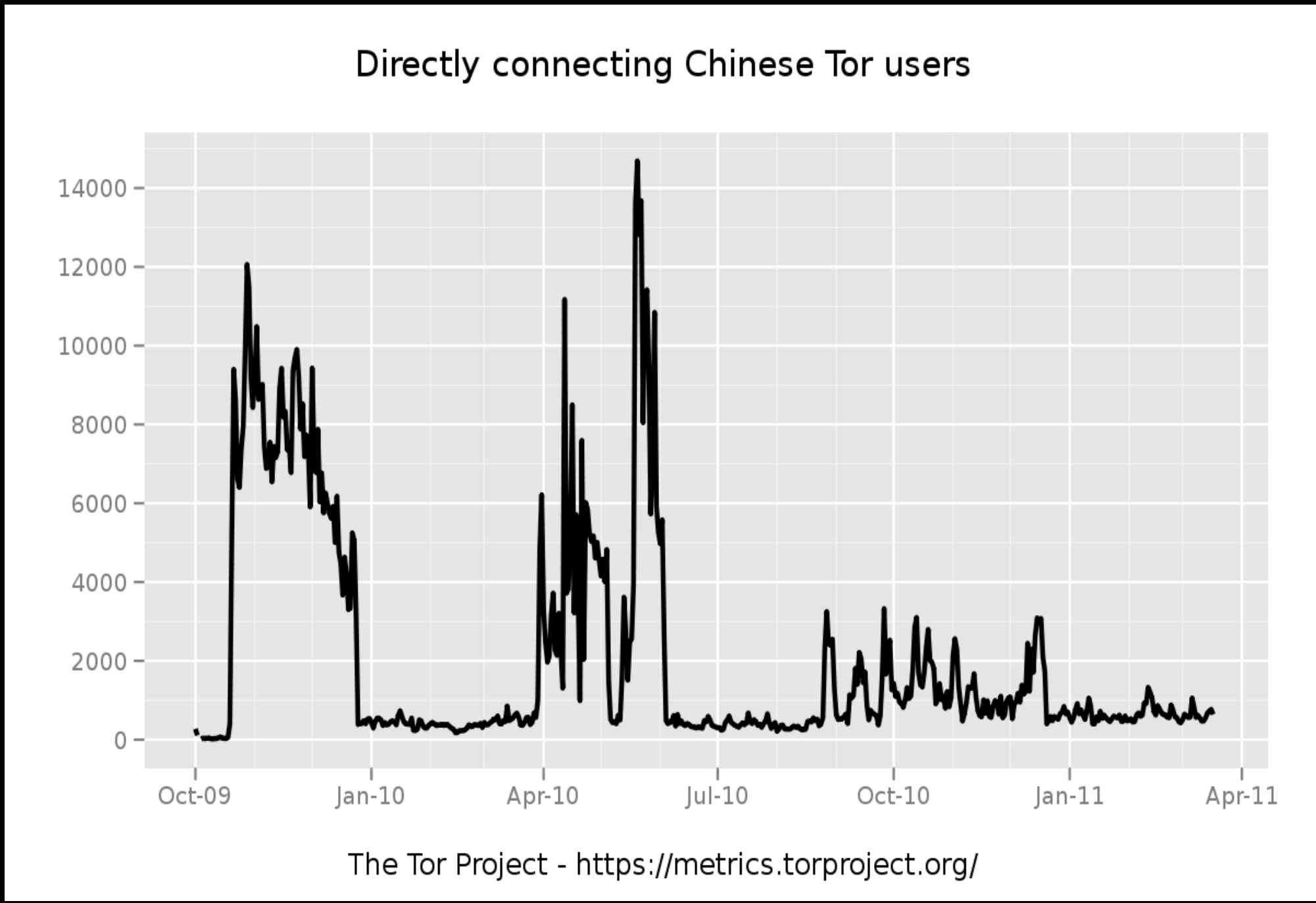
4) You can set up your own private bridge and tell your target users directly

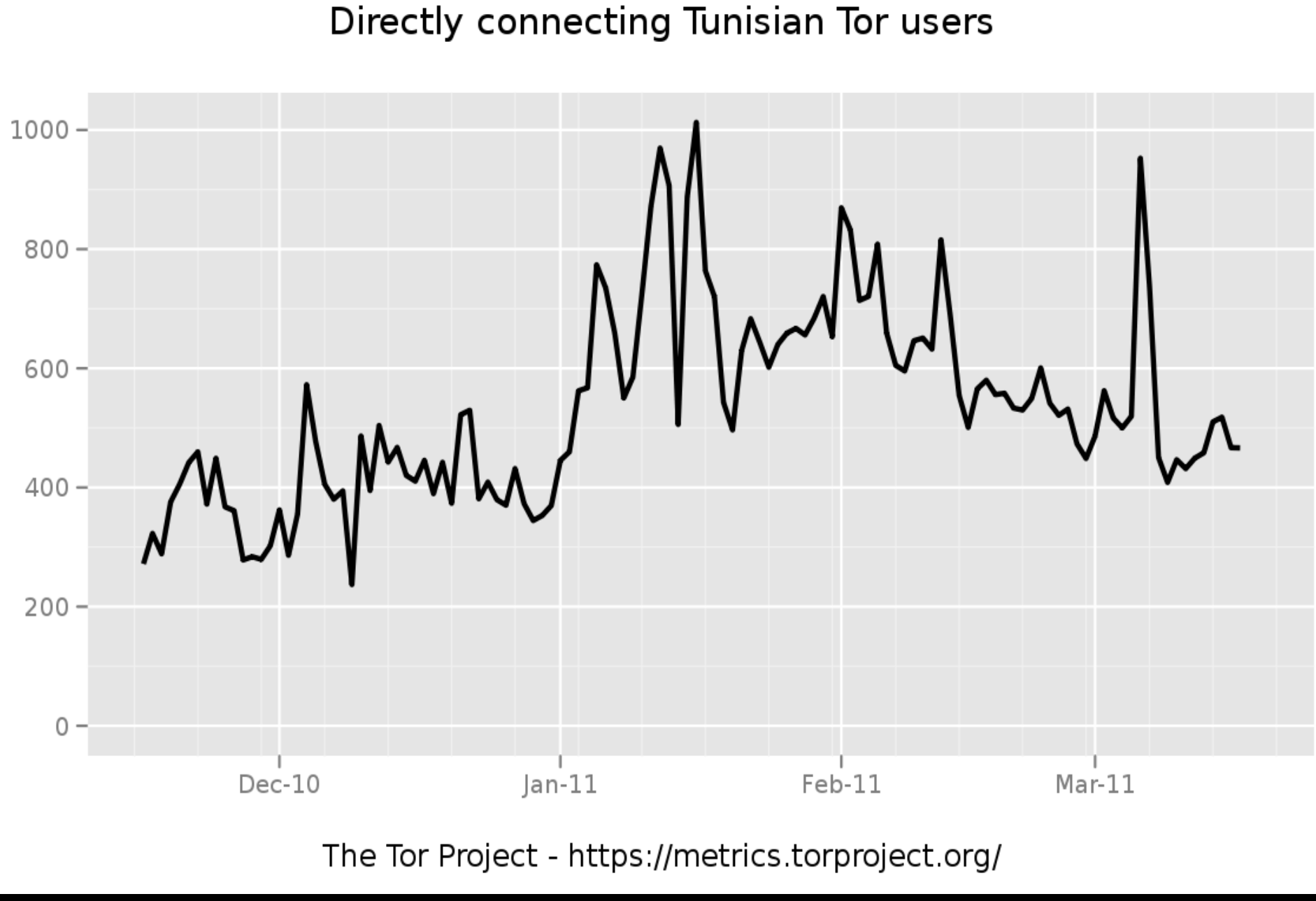

Number of directory requests to directory mirror trusted
10000
8000
6000
4000
2000
0
Sep 10
Sep 17
Sep 24
Oct 01
China
https://torproject.org

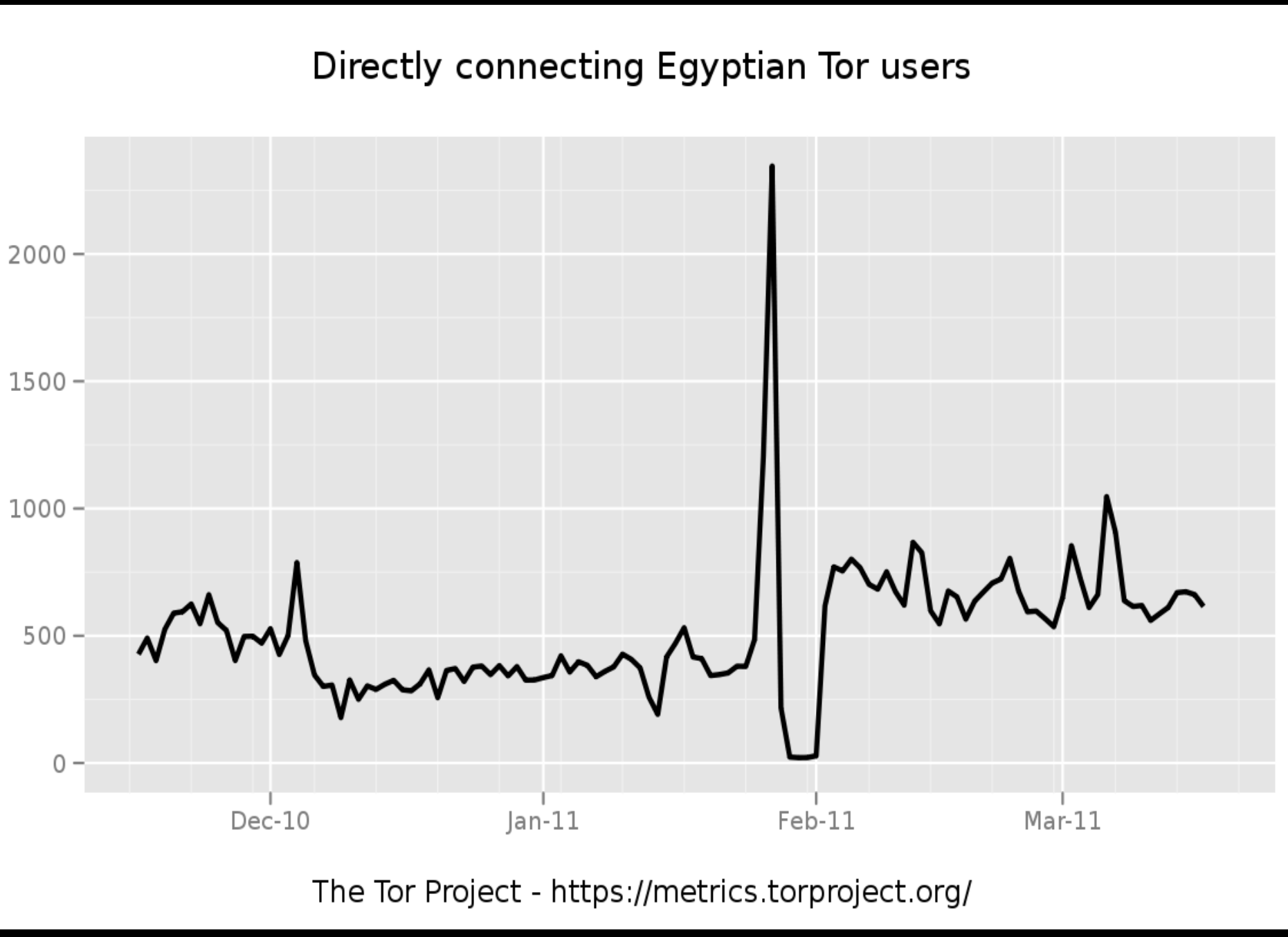

Chinese Tor users via bridges
30000
20000
10000
0
Jul
Aug
Sep
Oct
Nov 09

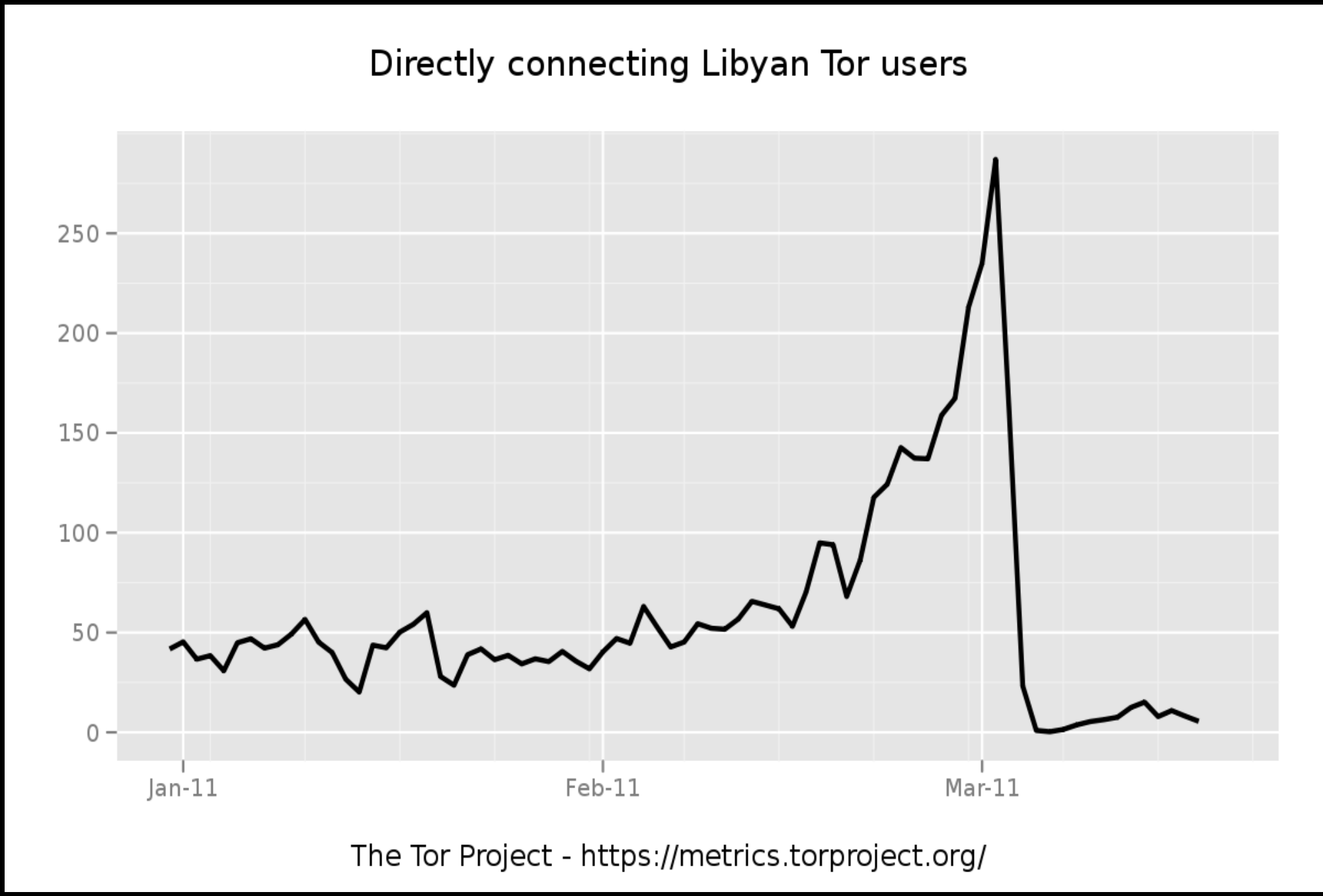


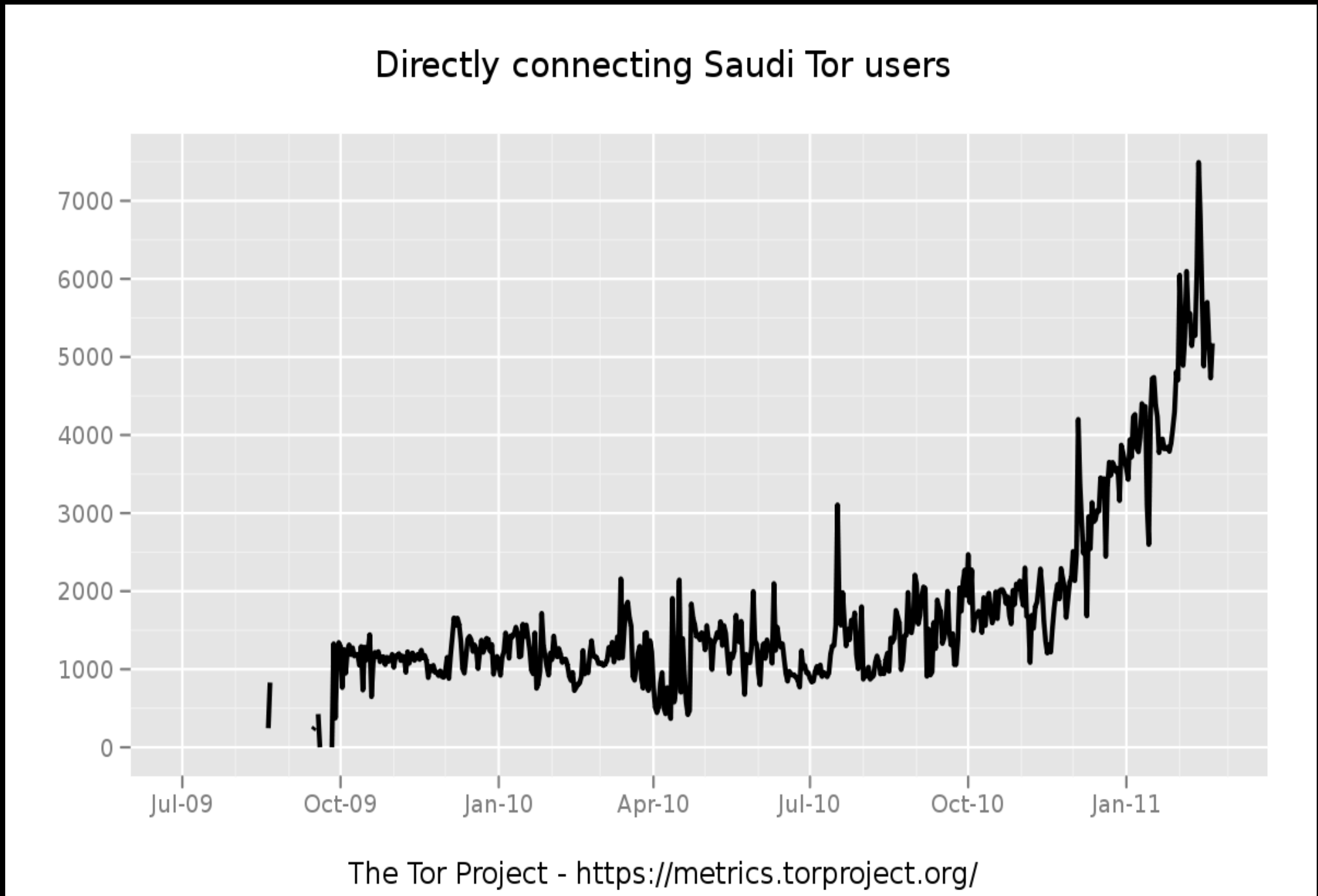
The Tor Project - https://metrics.torproject.org/

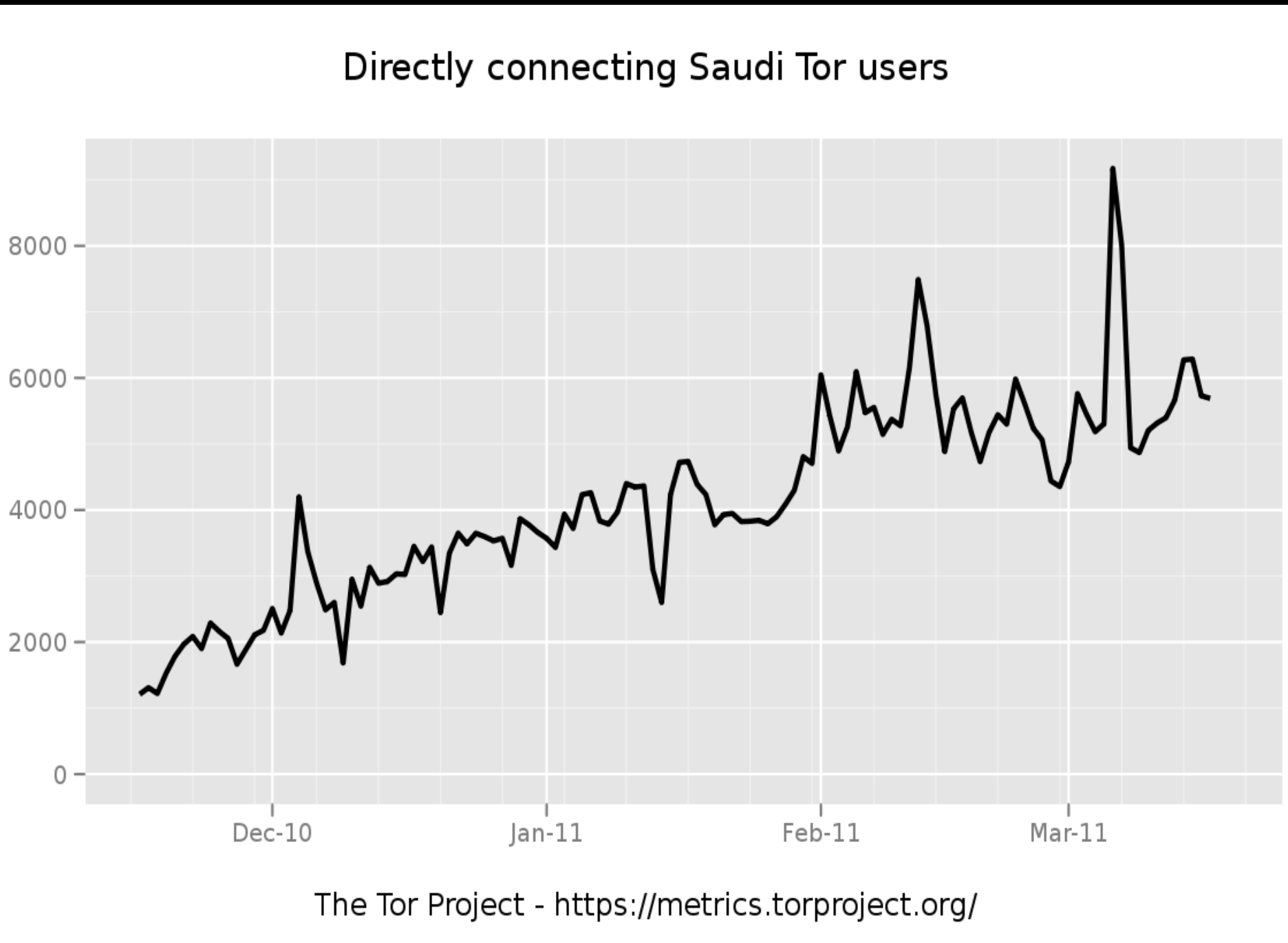

Directly connecting Chinese Tor users
14000
12000
10000
8000
6000
4000
2000
0
Oct-09
Jan-10
Apr-10
Jul-10
Oct-10
Jan-11
Apr-11
The Tor Project - https://metrics.torproject.org/

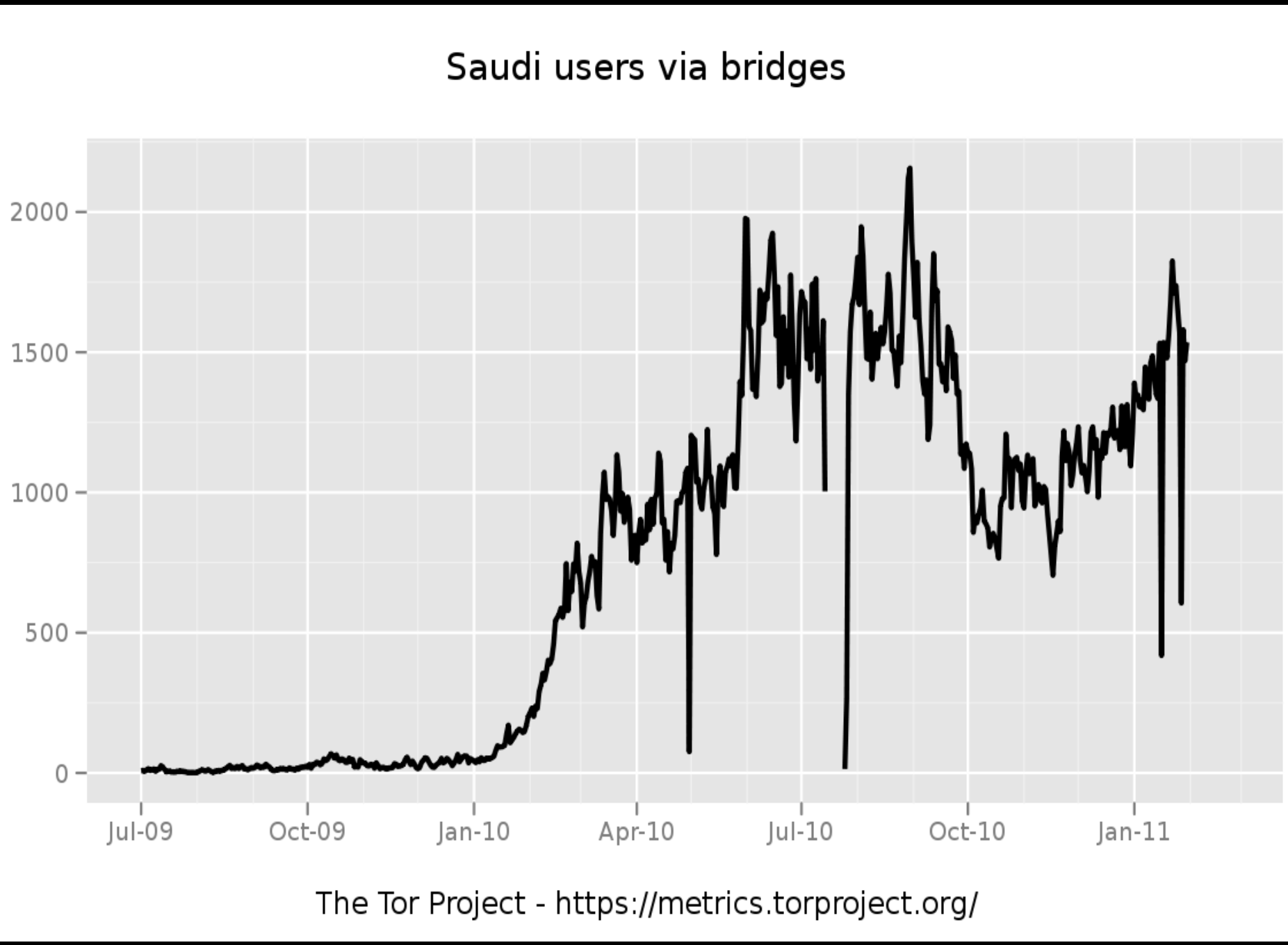

Directly connecting Tunisian Tor users
1000
800
600
400
200
0
Dec-10
Jan-11
Feb-11
Mar-11
The Tor Project - https://metrics.torproject.org/


Directly connecting Egyptian Tor users
2000
1500
1000
500
0
Dec-10
Jan-11
Feb-11
Mar-11
The Tor Project - https://metrics.torproject.org/


Directly connecting Libyan Tor users
250
200
150
100
50
0
Jan-11
Feb-11
Mar-11
The Tor Project - https://metrics.torproject.org/


Directly connecting Saudi Tor users
7000
6000
5000
4000
3000
2000
1000
0
Jul-09
Oct-09
Jan-10
Apr-10
Jul-10
Oct-10
Jan-11
The Tor Project - https://metrics.torproject.org/


Directly connecting Saudi Tor users
8000
6000
4000
2000
0
Dec-10
Jan-11
Feb-11
Mar-11
The Tor Project - https://metrics.torproject.org/


Saudi users via bridges
2000
1500
1000
500
0
Jul-09
Oct-09
Jan-10
Apr-10
Jul-10
Oct-10
Jan-11
The Tor Project - https://metrics.torproject.org/

Censoring attacker's goals

Little reprisal against passive consumers of information.

Producers and distributors of information in greater danger.

Censors (actually, govts) have economic, political, social incentives not to block the whole Internet.

But they don't mind collateral damage.

What we're up against

Govt firewalls used to be stateless. Now they're buying fancier hardware.

Burma vs Iran vs China

New filtering techniques spread by commercial (American) companies :(

How to separate “oppressing employees” vs “oppressing citizens” arms race?

Javascript, cookies, history, etc

Javascript refresh attack

Cookies, History, browser window size, user-agent, language, http auth, ...

Mostly problems when you toggle from Tor to non-Tor or back

Mike Perry's Torbutton Firefox extension tackles many of these

Flash is dangerous too

Some apps are bad at obeying their proxy settings.

Adobe PDF plugin. Flash. Other plugins. Extensions. Especially Windows stuff: did you know that Microsoft Word is a network app?

Choose how to install it

Tor Browser Bundle: standalone Windows exe with Tor, Vidalia, Firefox, Torbutton, Polipo, e.g. for USB stick

Vidalia bundle: Windows/OSX installer

Tor VM: Transparent proxy for Windows

“Net installer” via our secure updater

Amnesia Linux LiveCD

Only a piece of the puzzle

Assume the users aren't attacked by their hardware and software

No spyware installed, no cameras watching their screens, etc

Users can fetch a genuine copy of Tor?

Publicity attracts attention

Many circumvention tools launch with huge media splashes. (The media loves this.)

But publicity attracts attention of the censors.

We threaten their *appearance* of control, so they must respond.

We can affect the pace of the arms race.

Using Tor in oppressed areas

Common assumption: risk from using Tor increases as firewall gets more restrictive.

But as firewall gets more restrictive, more ordinary people use Tor too, for more mainstream activities.

So the “median” use becomes more acceptable?

Trust and reputation

See January 2009 blog post by Hal Roberts about how some circumvention tools sell user data

Many of these tools see circumvention and privacy as totally unrelated goals

Advocacy and education

Unending stream of people (e.g. in DC) who make critical policy decisions without much technical background

Worse, there's a high churn rate

Need to teach policy-makers, business leaders, law enforcement, journalists, ...

Our NSF EAGER

1) Invent and deploy new privacy-preserving algorithms to collect data about the Tor network, its performance, and its users

2) Publish this data, plus tools to analyze it

3) Figure out what else to measure and do it

4) Work with other research groups to make sure they get the data they need to solve the problems Tor actually has

Next steps (policy)

Technical solutions won't solve the whole censorship problem. After all, firewalls are *socially* successful in these countries.

But a strong technical solution is still a critical puzzle piece.

You should run a bridge! We only have ~750.

We'd love to help with some trainings, to help users and to make Tor better.

BridgeDB needs a feedback cycle

- Measure how much use each bridge sees
- Measure bridge blocking
- Then adapt bridge distribution to favor efficient distribution channels
- Need to invent new distribution channels
- Need more and changing bridge addresses
  - Redirecting a whole /16 ?
  - Promote clients to bridges?

Measuring bridge reachability

Passive: bridges track incoming connections by country

Active: scan bridges from within the country

Clients self-report blockage (via some other bridge)

Measure remotely via FTP reflectors

Bridges test for duplex blocking

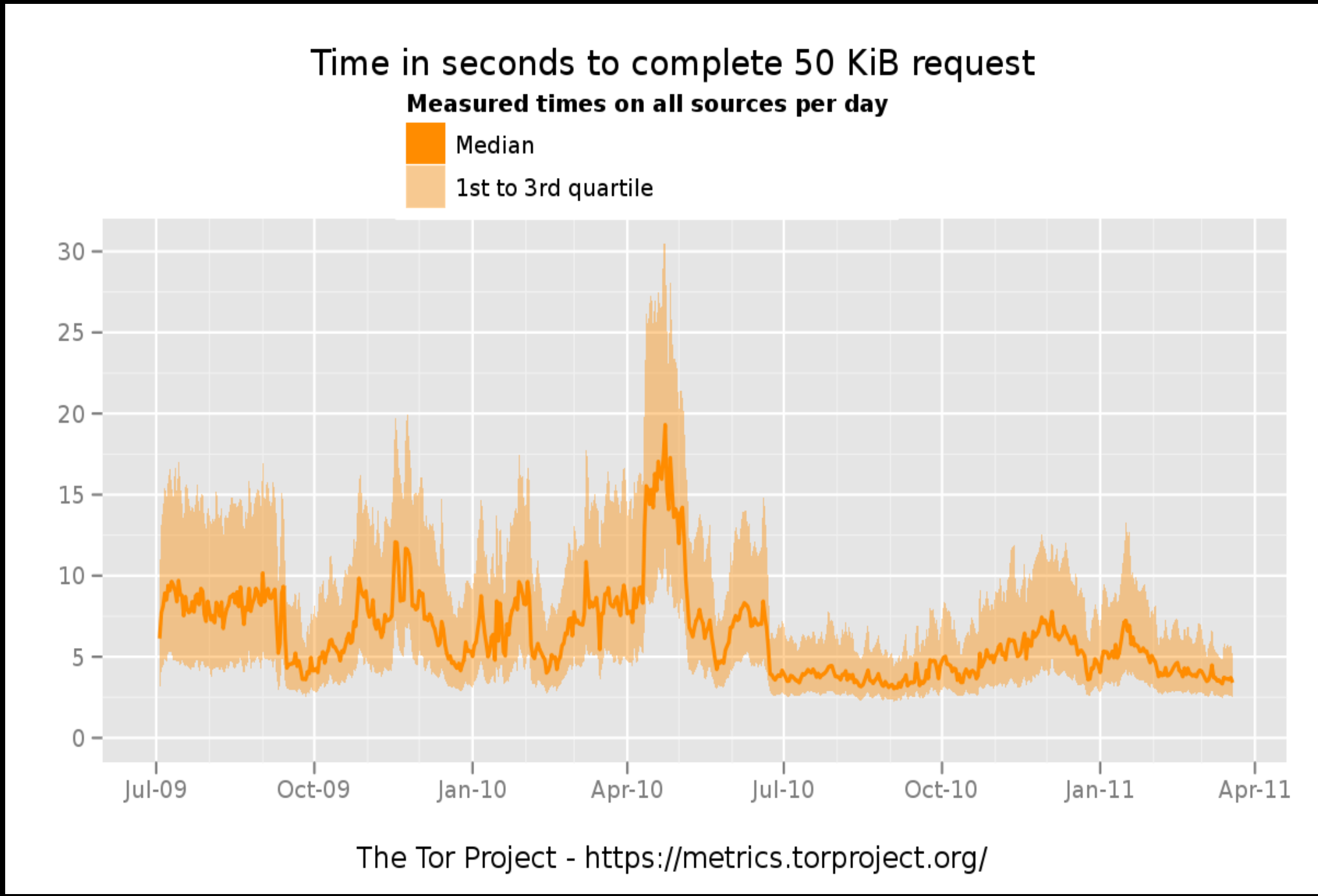
Other components

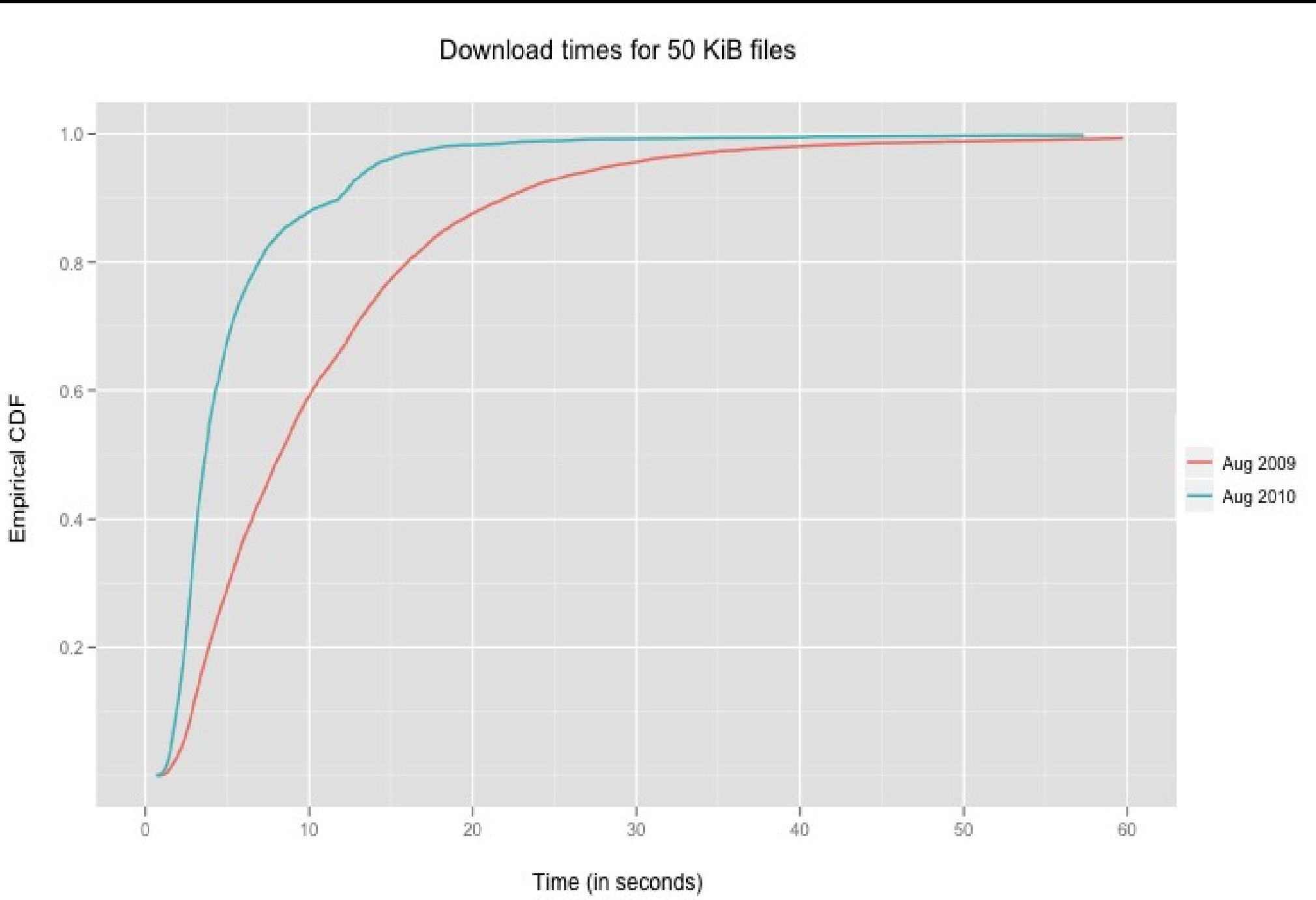
- Traffic camouflaging
  - Superencrypt so no recognizable bytes? Shape like HTTP?
  - We're working on a modular transport API
- Client-side automation for usability
- Performance / scalability
  - Especially for low bandwidth

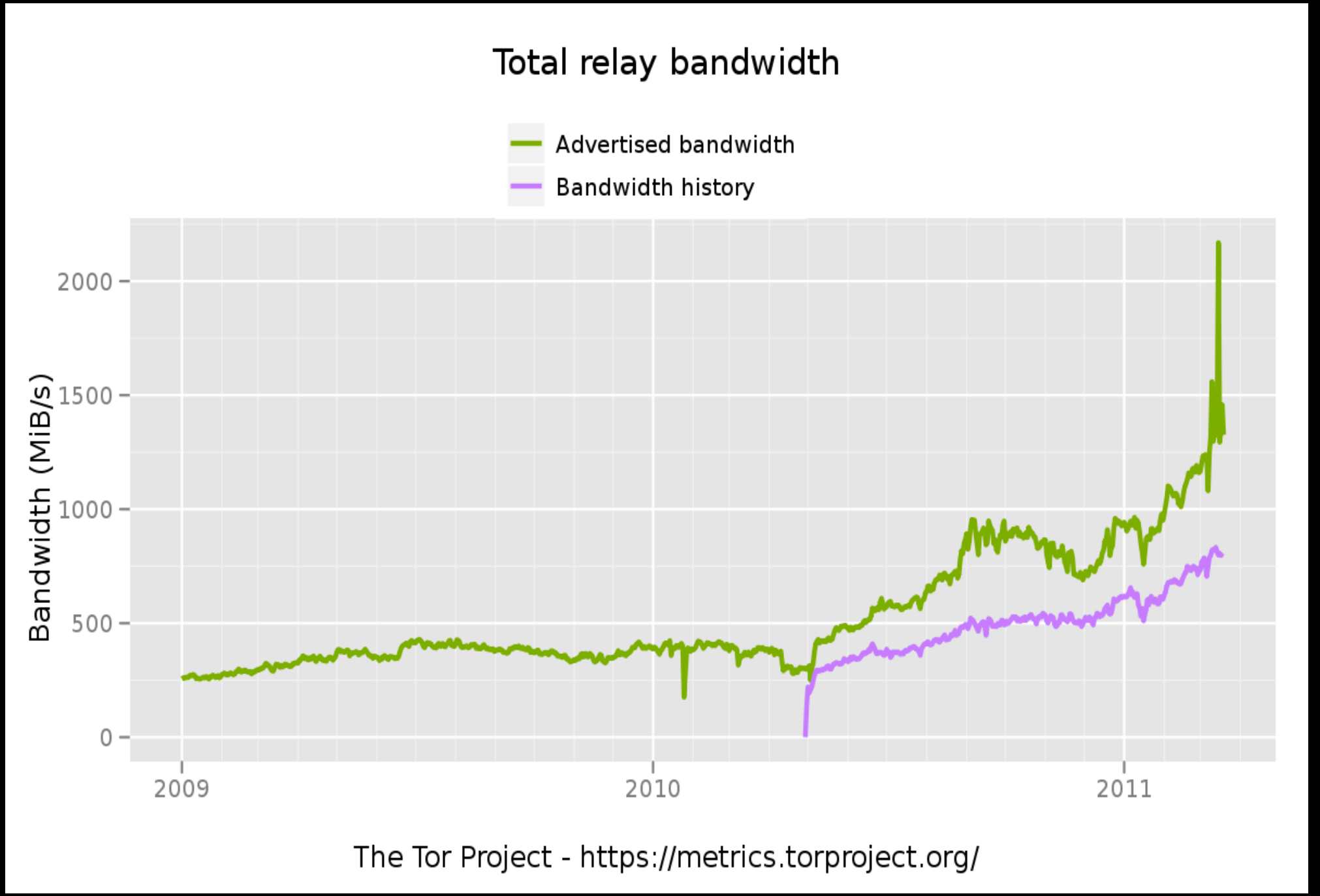
Questions?

See also:

- https://torproject.org/
- http://freehaven.net/anonbib/
- Specs, design papers, open proposals, etc
- Public mailing lists
- The code is open source


Time in seconds to complete 50 KiB request
Measured times on all sources per day
Median
1st to 3rd quartile
30
25
20
15
10
5
0
Jul-09
Oct-09
Jan-10
Apr-10
Jul-10
Oct-10
Jan-11
Apr-11
The Tor Project - https://metrics.torproject.org/


Download times for 50 KiB files
Empirical CDF
1.0
0.8
0.6
0.4
0.2
0
10
20
30
40
50
60
Time (in seconds)
Aug 2009
Aug 2010


Total relay bandwidth
Advertised bandwidth
Bandwidth history
Bandwidth (MiB/s)
2000
1500
1000
500
0
2009
2010
2011
The Tor Project - https://metrics.torproject.org/