Online Anonymity

Roger Dingledine, Andrew Lewman
The Tor Project
https://torproject.org/

Outline

- *Why anonymity?*
- Crash course on Tor
- Future

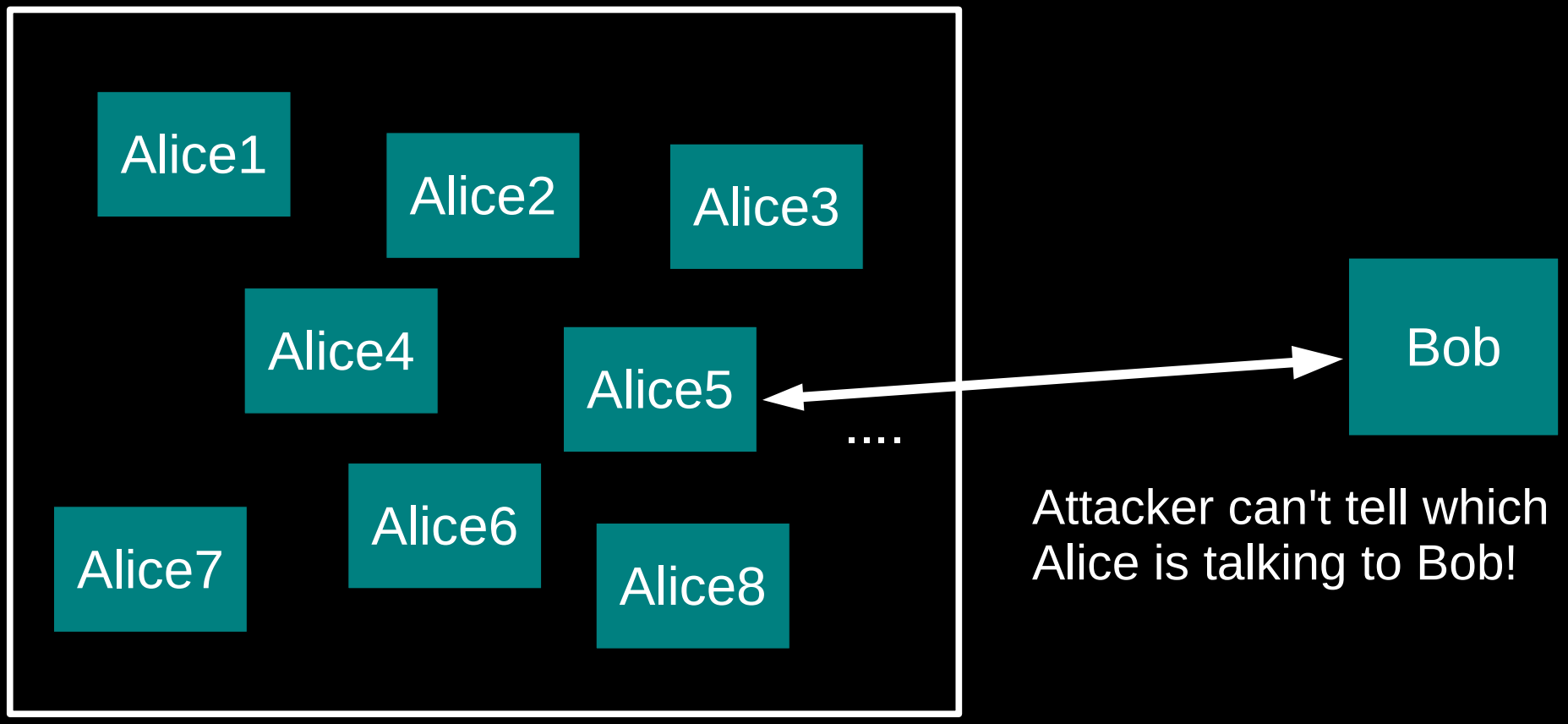
Informally: anonymity means you can't tell who did what

“Who wrote this blog post?”

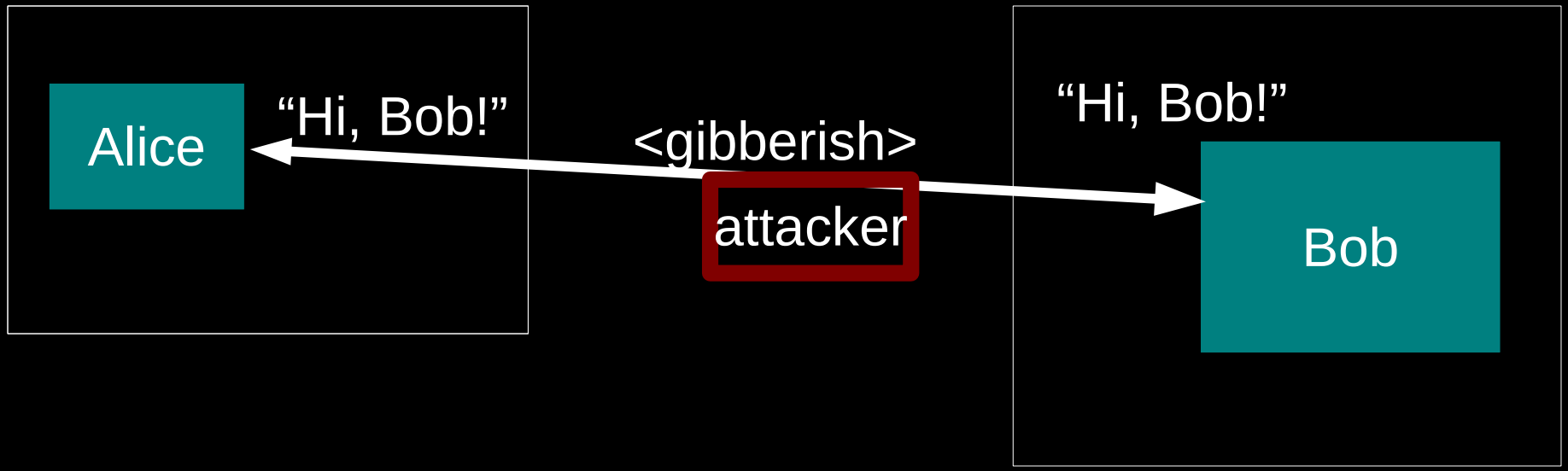
“Who's been viewing my webpages?”

“Who's been emailing patent attorneys?”

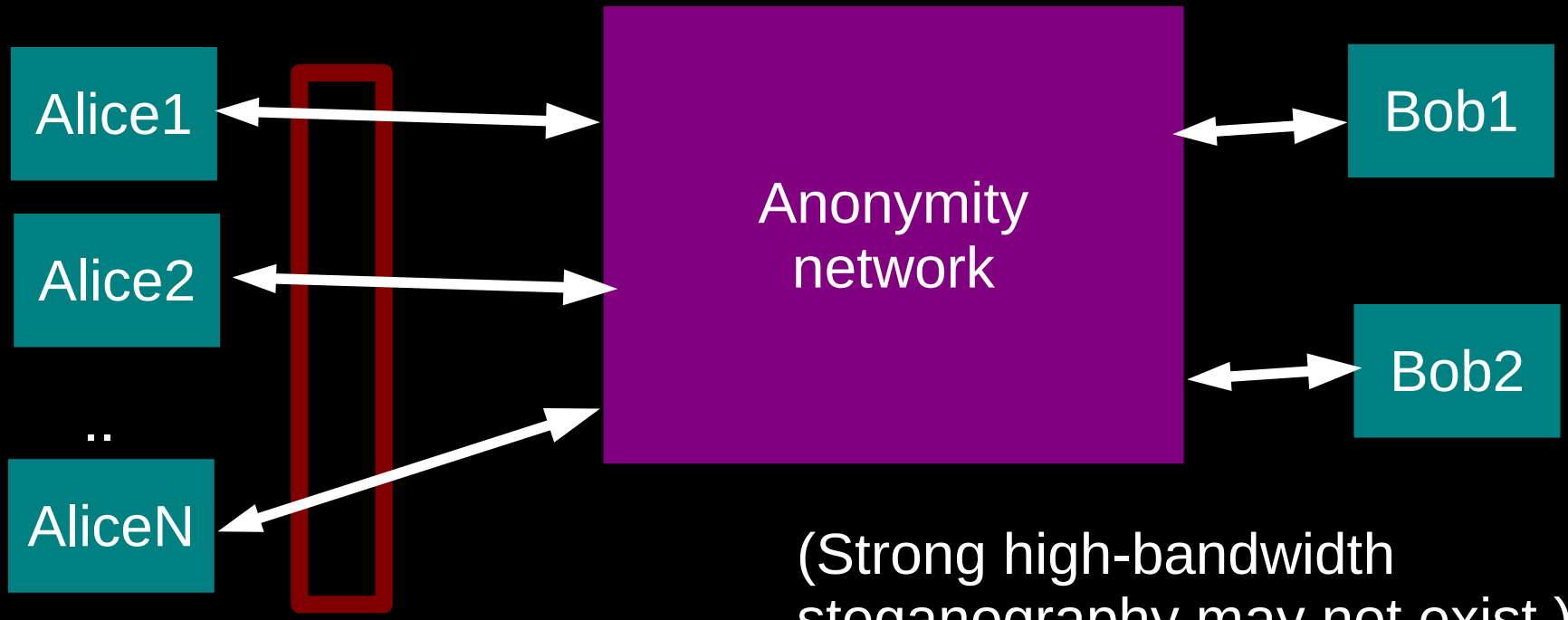
Formally: anonymity means indistinguishability within an "anonymity set"



Anonymity isn't cryptography: Cryptography just protects contents.



Anonymity isn't steganography:
Attacker can tell that Alice is talking; just not to whom.



(Strong high-bandwidth steganography may not exist.)

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?”

…since “weak” anonymity... isn't.

~~“You can't prove it was me!”~~

*Proof is a **very** strong word. With statistics, suspicion becomes certainty.*

Will others parties have the ability and incentives to keep their promises?

~~“Promise you won't look!”~~

~~“Promise you won't remember!”~~

~~“Promise you won't tell!”~~

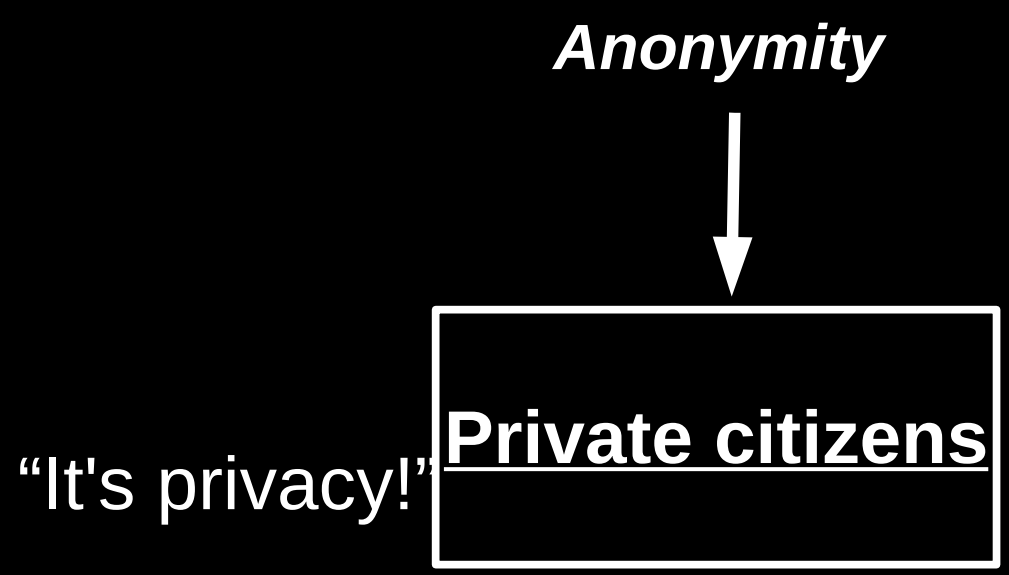
~~“I didn't write my name on it!”~~

Not what we're talking about.

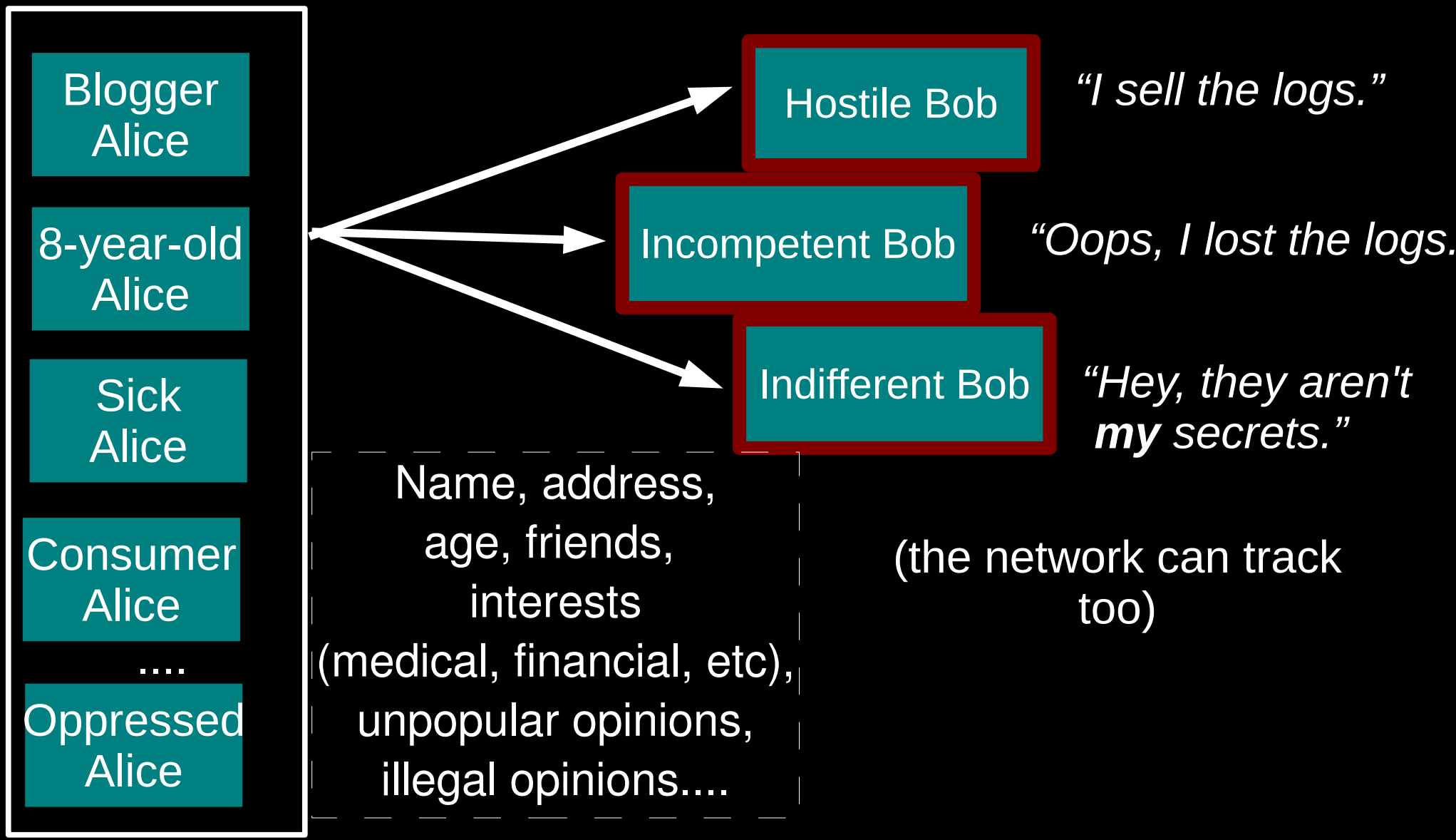
Nope! (More info later.)

~~“Isn't the Internet already anonymous?”~~

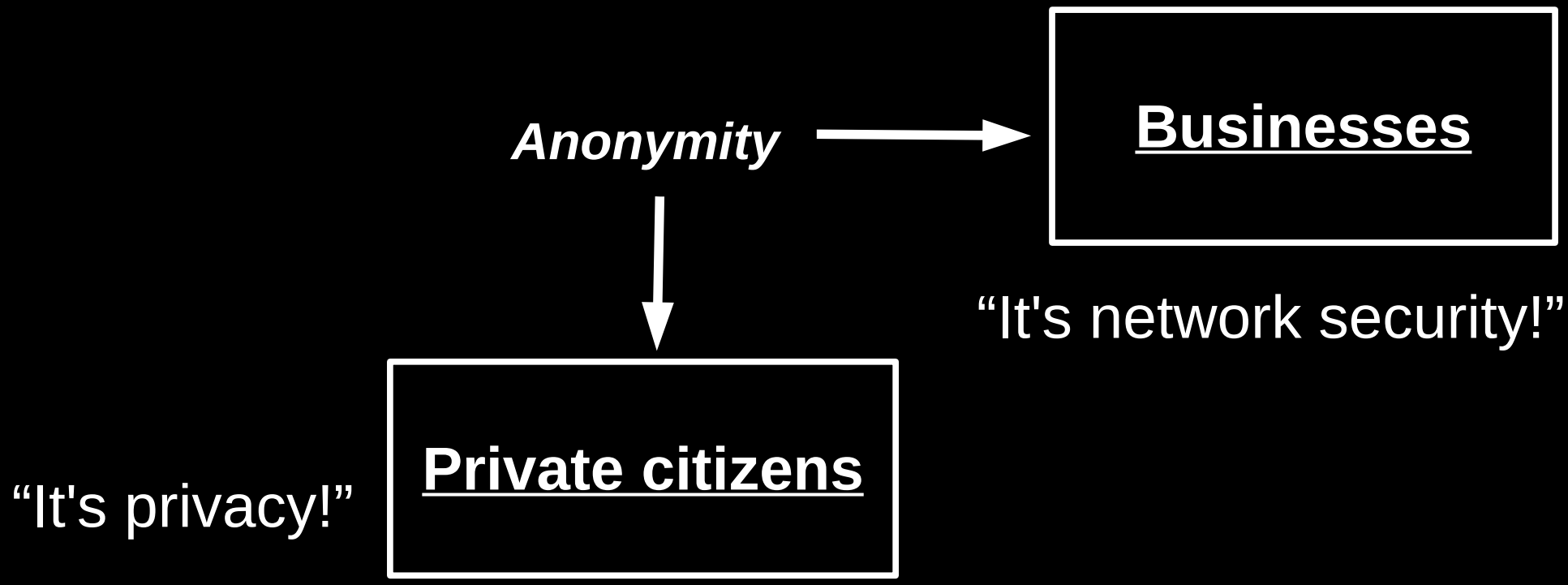
Anonymity serves different interests for different user groups.



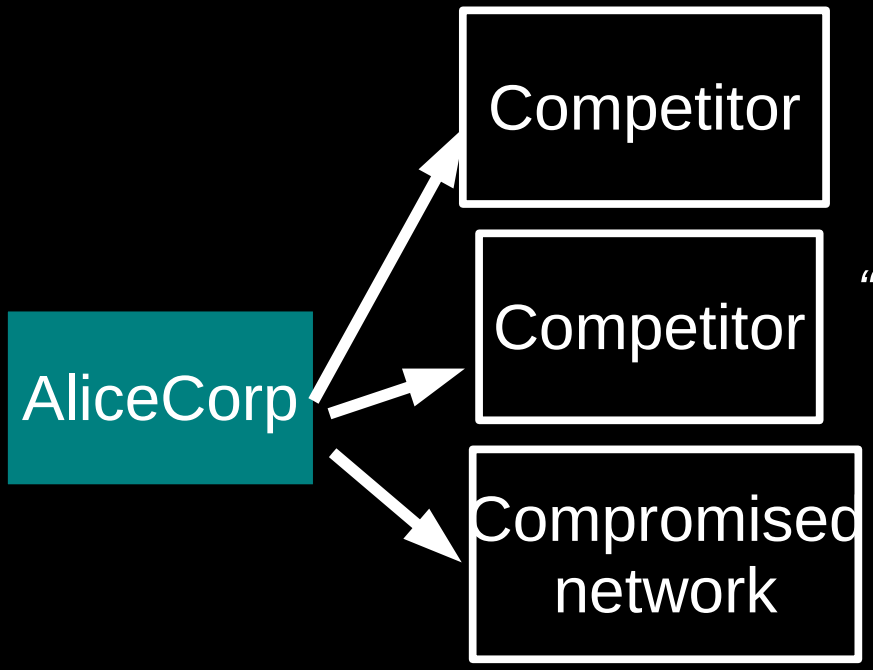
Regular citizens don't want to be watched and tracked.



Anonymity serves different interests for different user groups.



Businesses need to keep trade secrets.

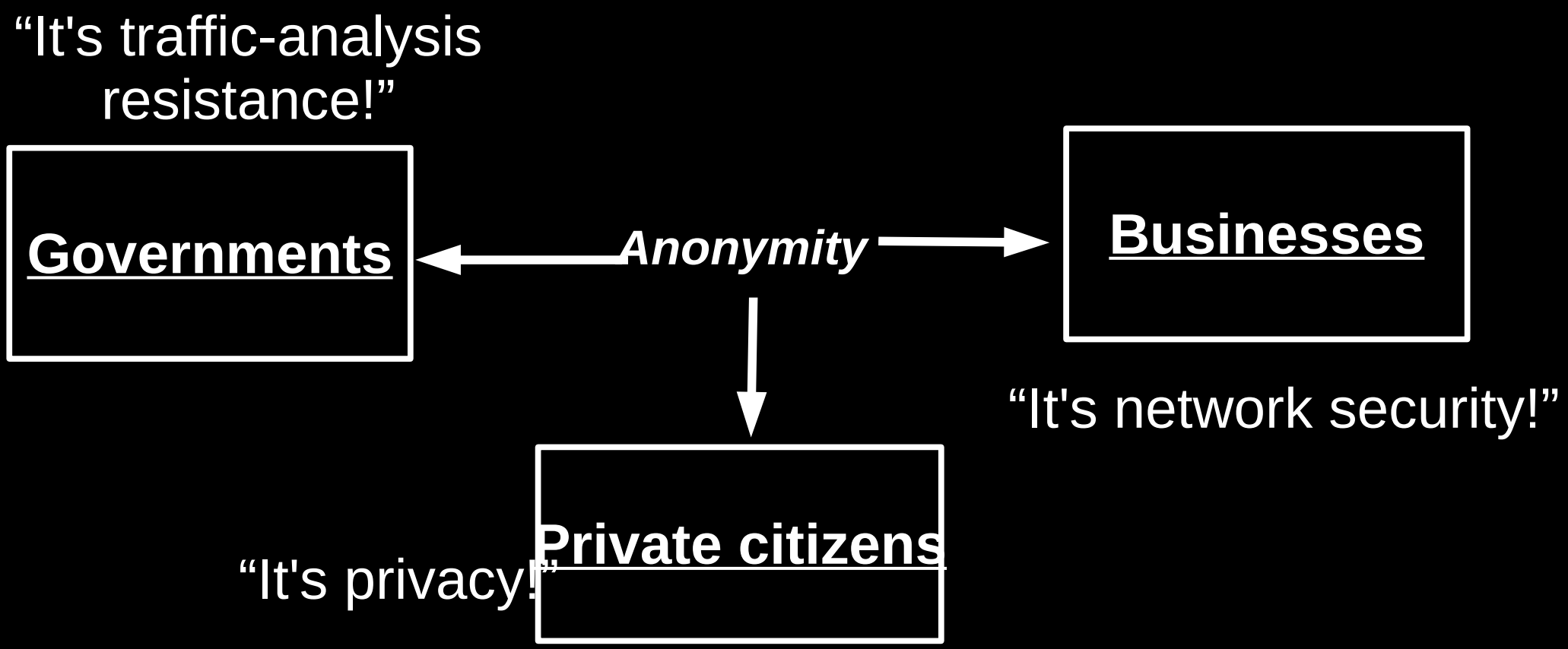


“Oh, your employees are reading our patents/jobs page/product sheets?”

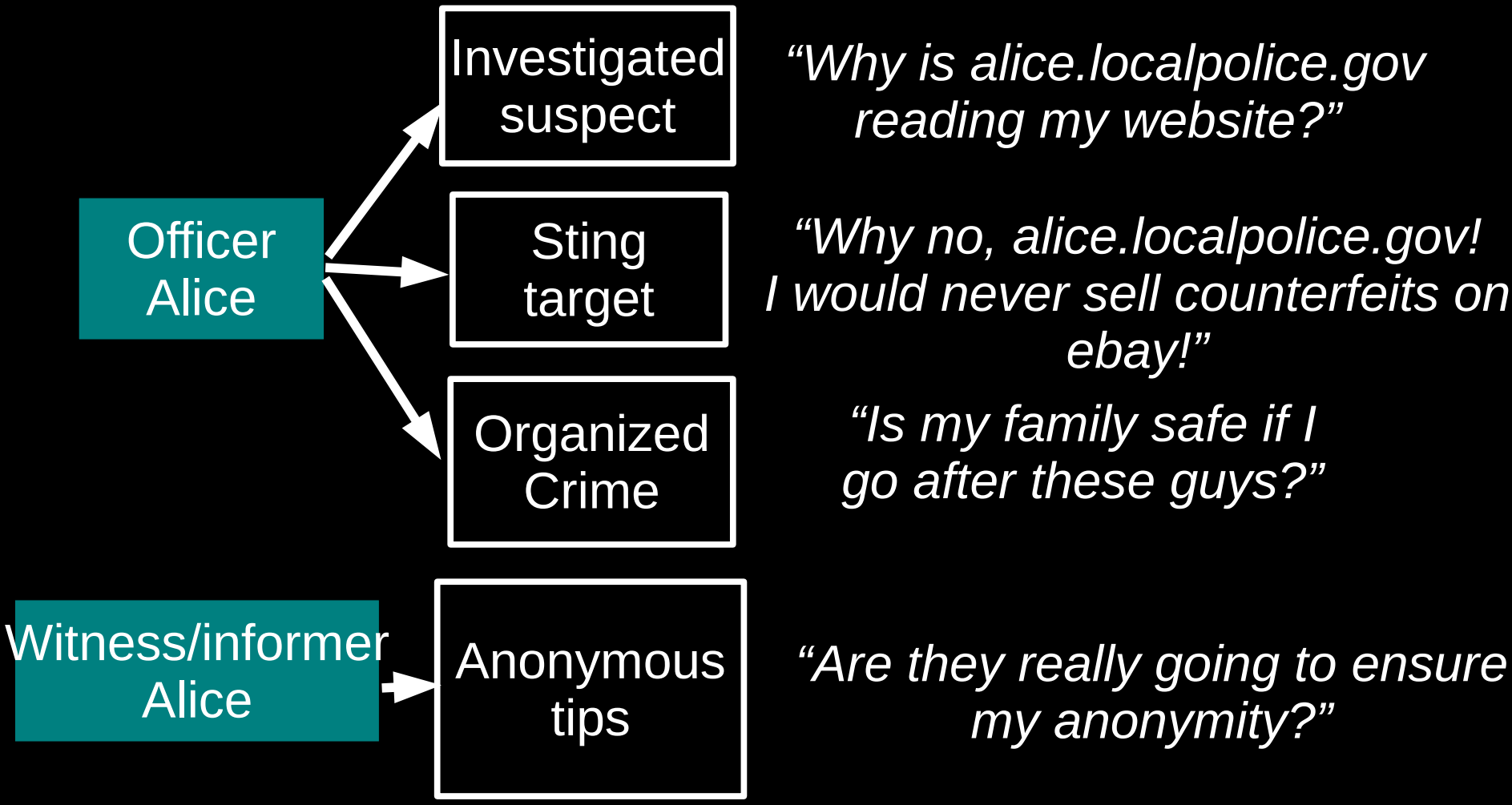
“Hey, it's Alice! Give her the 'Alice' version!”

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

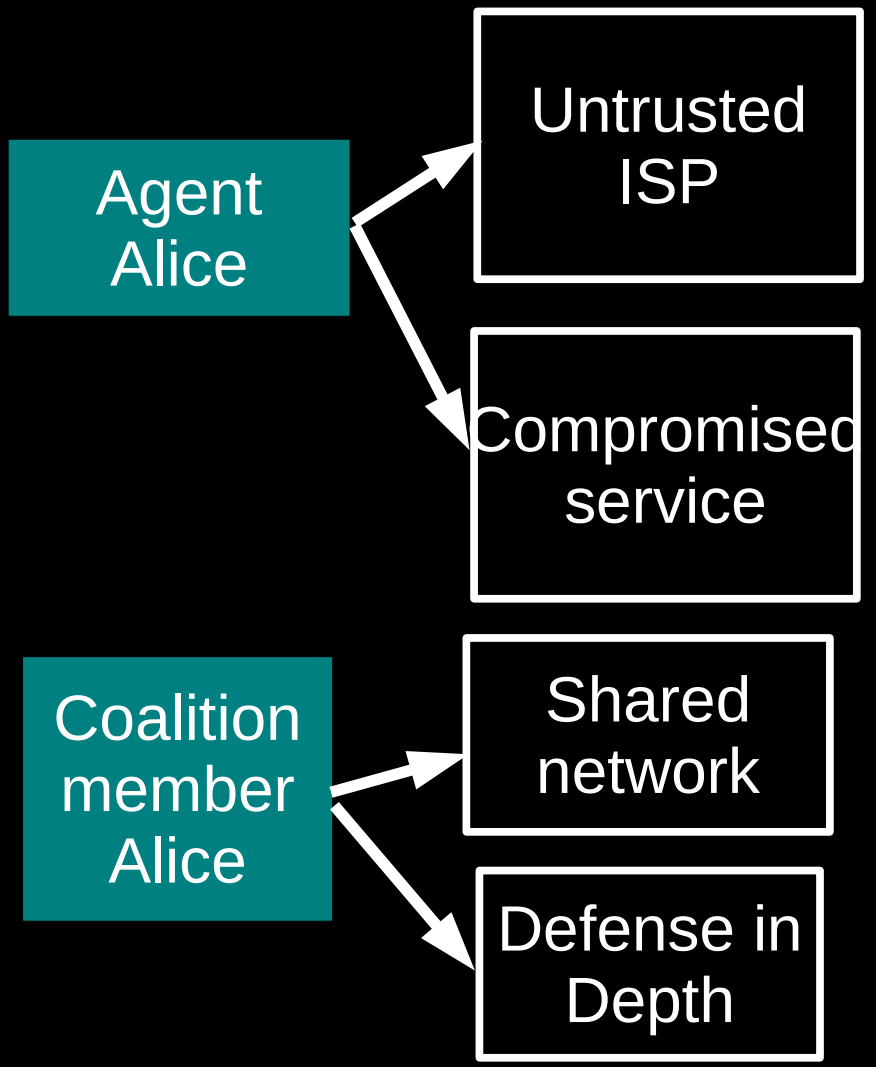
Anonymity serves different interests for different user groups.



Law enforcement needs anonymity to get the job done.



Governments need anonymity for their security



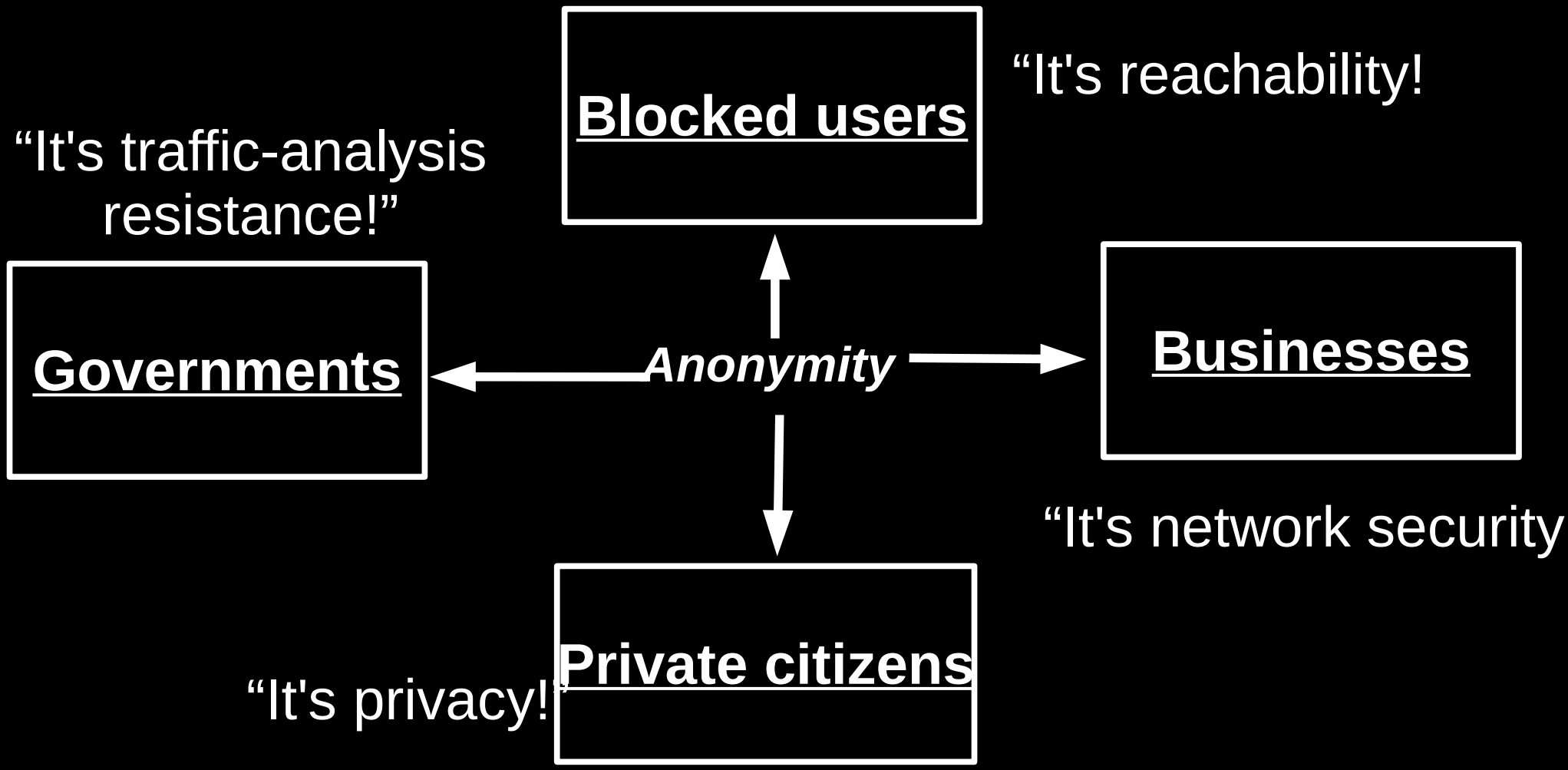
"What will you bid for a list of Baghdad IP addresses that get email from .gov?

"*What* does *the CIA Google for?*"

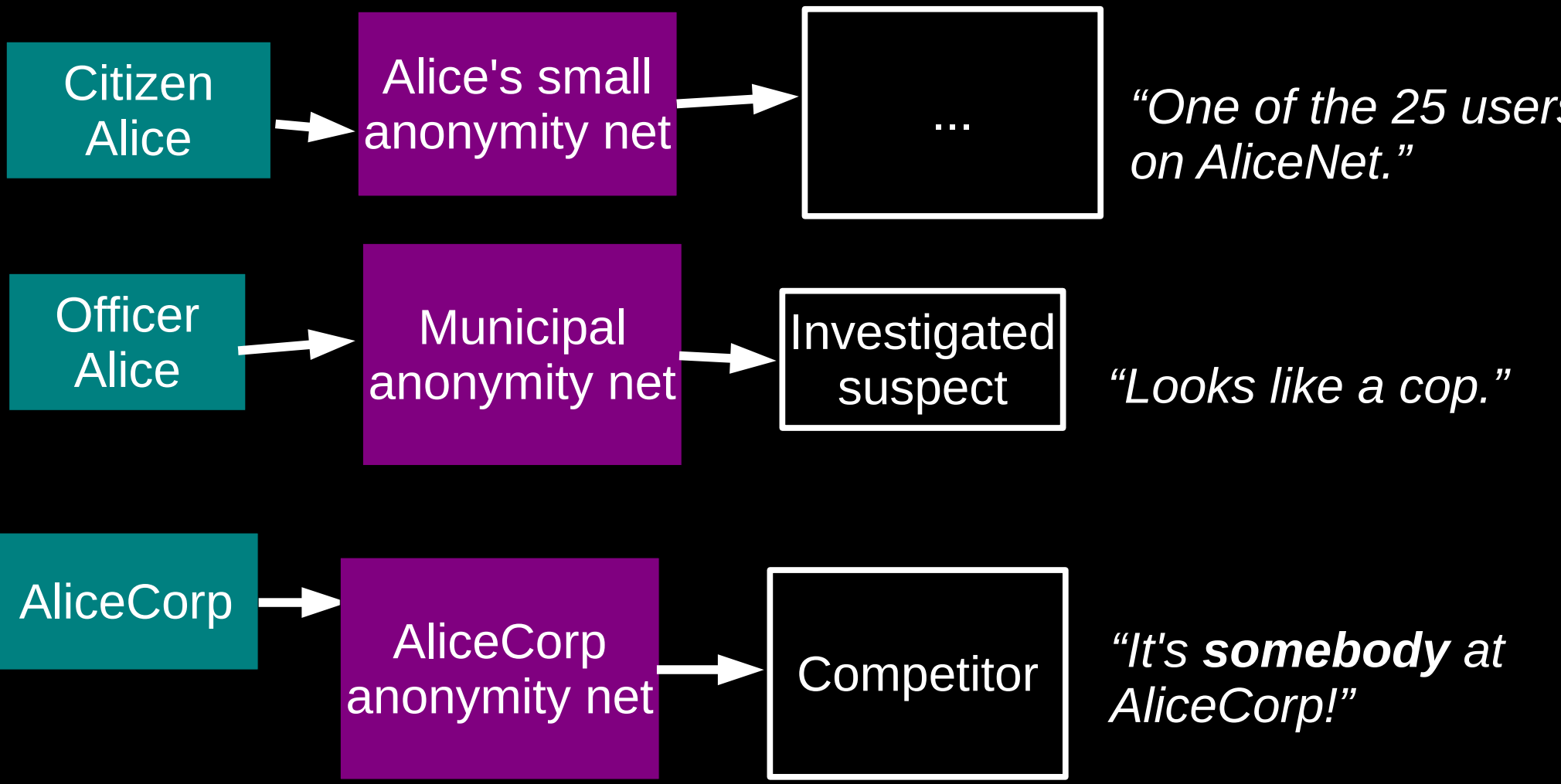
"Do I really want to reveal my internal network topology?"

"What about insiders?"

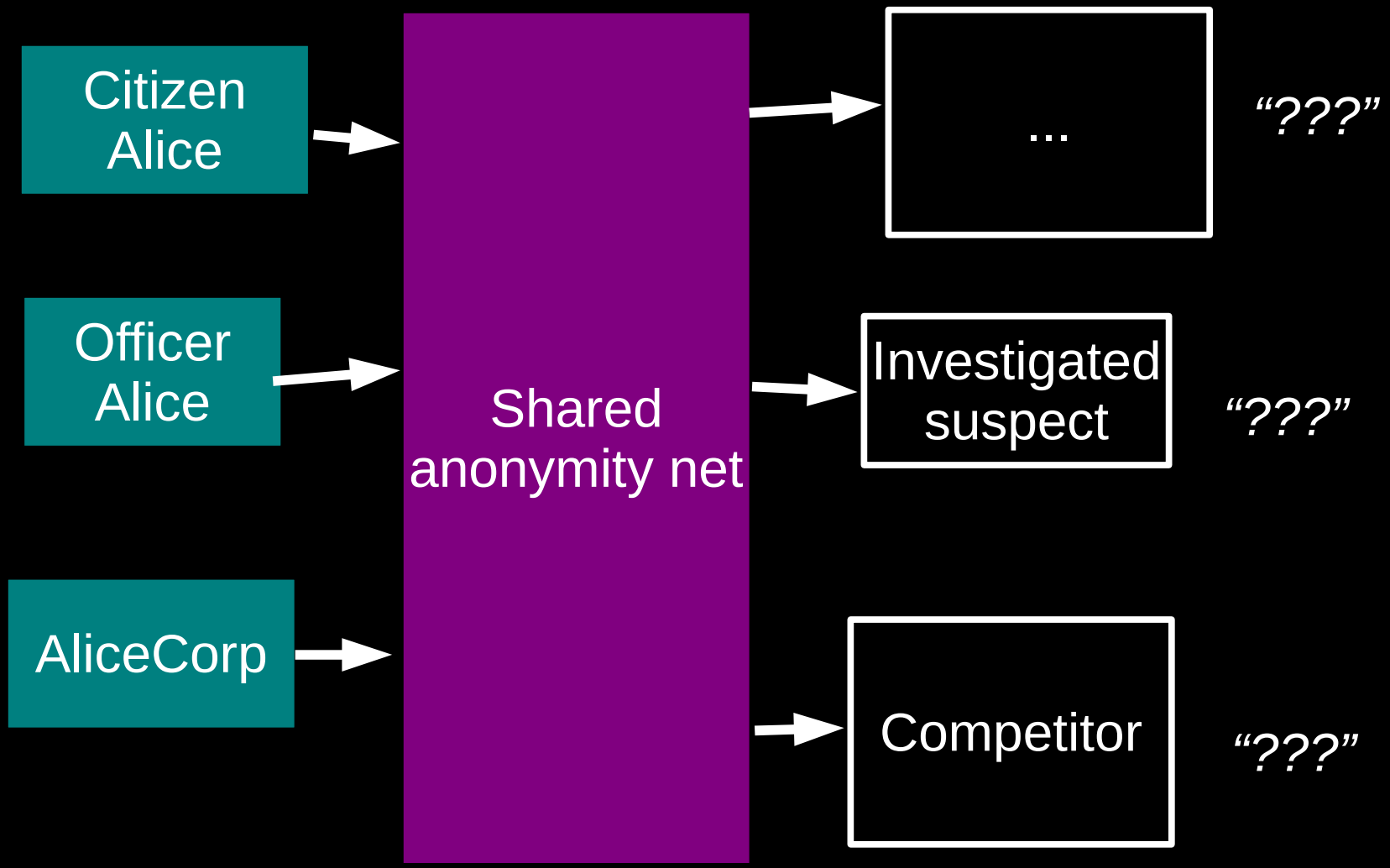
Anonymity serves different interests for different user groups.



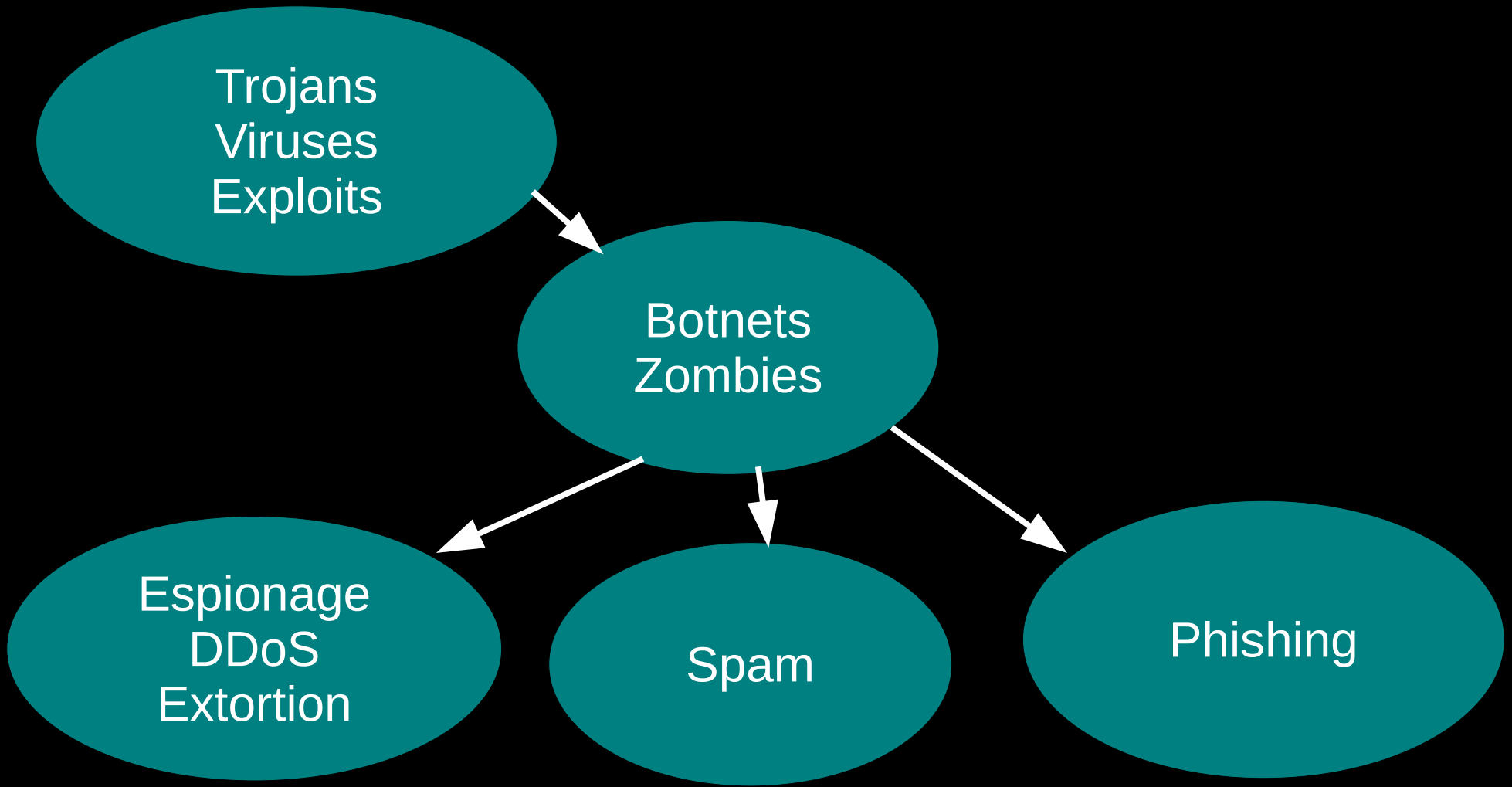
You can't get anonymity on your own: private solutions are ineffective...



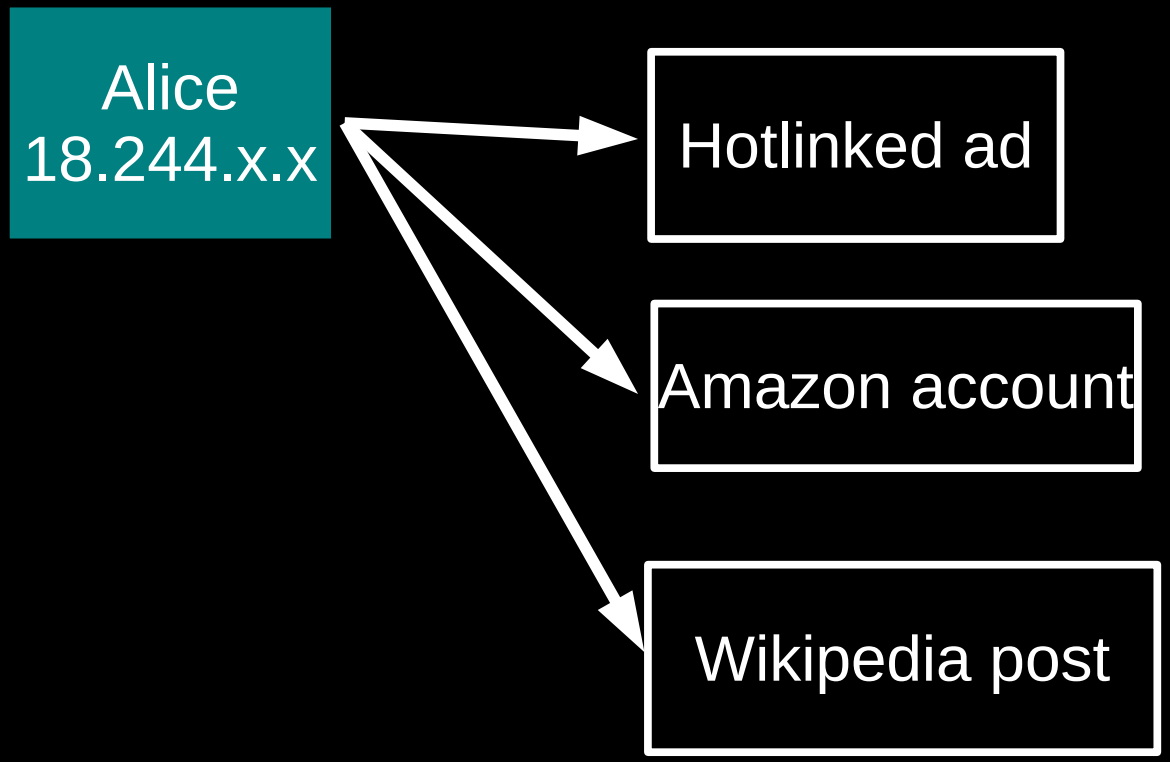
... so, anonymity loves company!



Current situation: Bad people on the Internet are doing fine



IP addresses can be enough to bootstrap knowledge of identity.



Tor is not the first or only design for anonymity.

Low-latency

Single-hop proxies

Crowds (~96)

V1 Onion Routing (~96)

ZKS “Freedom” (~99-01)

Java Anon Proxy (~00-)

Tor (01-)

High-latency

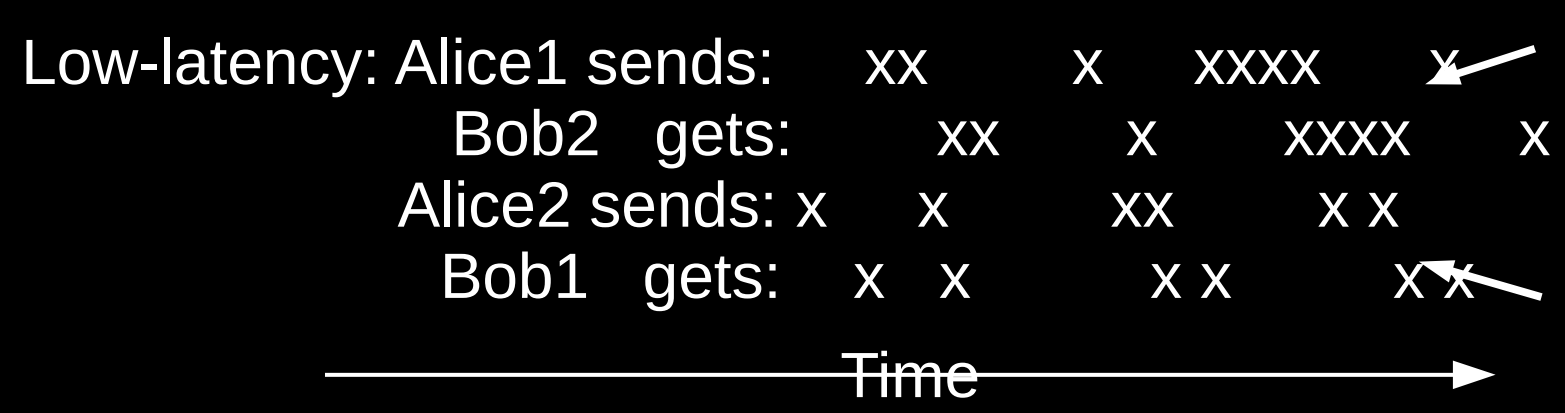
Chaum's Mixes (1981)

anon.penet.fi (~91)

Remailer networks: cypherpunk (~93), mixmaster (~95), mixminion (~02)

...and more!

Low-latency systems are vulnerable to end-to-end correlation attacks.



High-latency: Alice1 sends: xx x xxxx
Alice2 sends: x x xx x x
Bob1 gets: xx xxxx .....
Bob2 gets: x xxxxx .....

These attacks work in practice. The obvious defenses are expensive (like high-latency), useless, or both.

Still, we focus on low-latency, because it's more useful.

Interactive apps: web, IM, VOIP, ssh, X11, ...
 # users: millions?

Apps that accept multi-hour delays and high bandwidth overhead: email, sometimes.
 # users: tens of thousands at most?

And if anonymity loves company....?

Outline

- Why anonymity?
- *Crash course on Tor*
- Future

What is Tor?

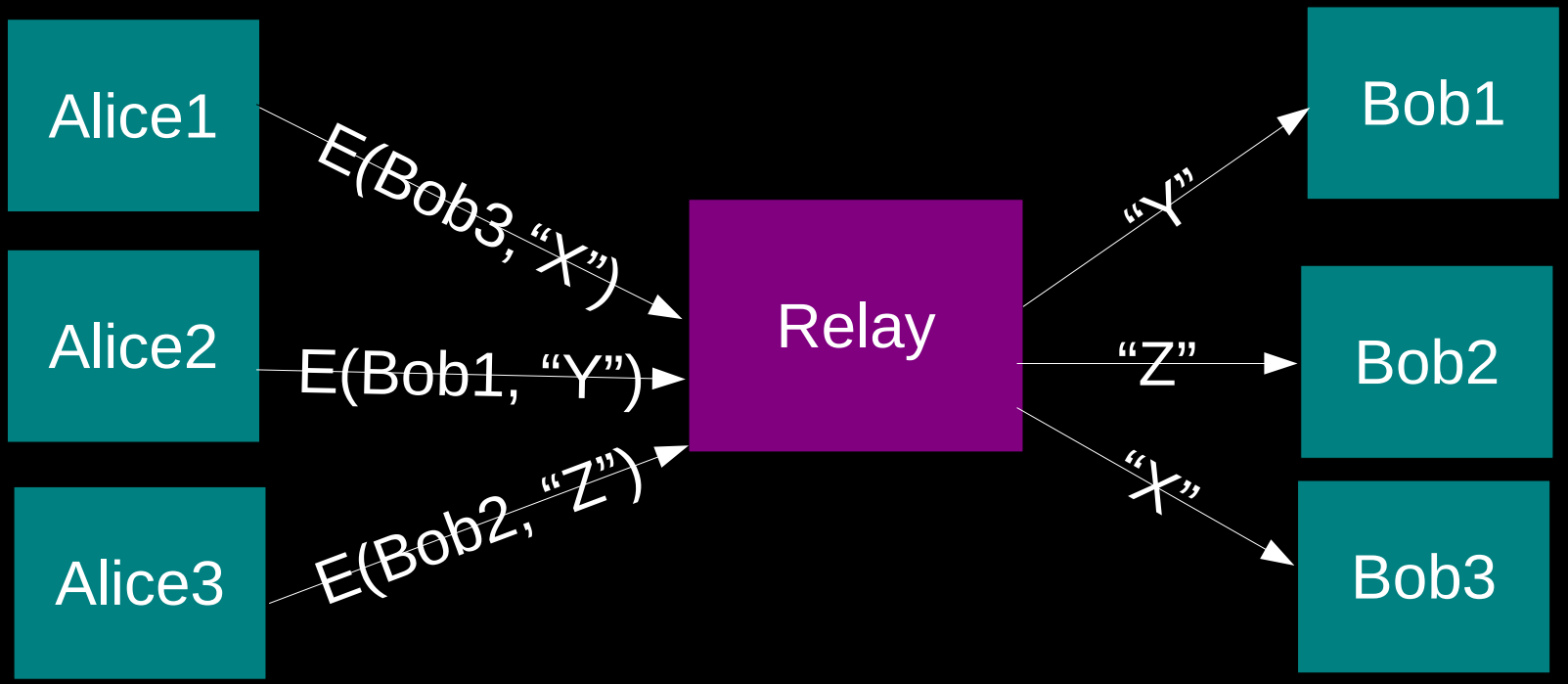
- online anonymity software and network
- open source, freely available
- active research environment

The Tor Project, Inc.

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

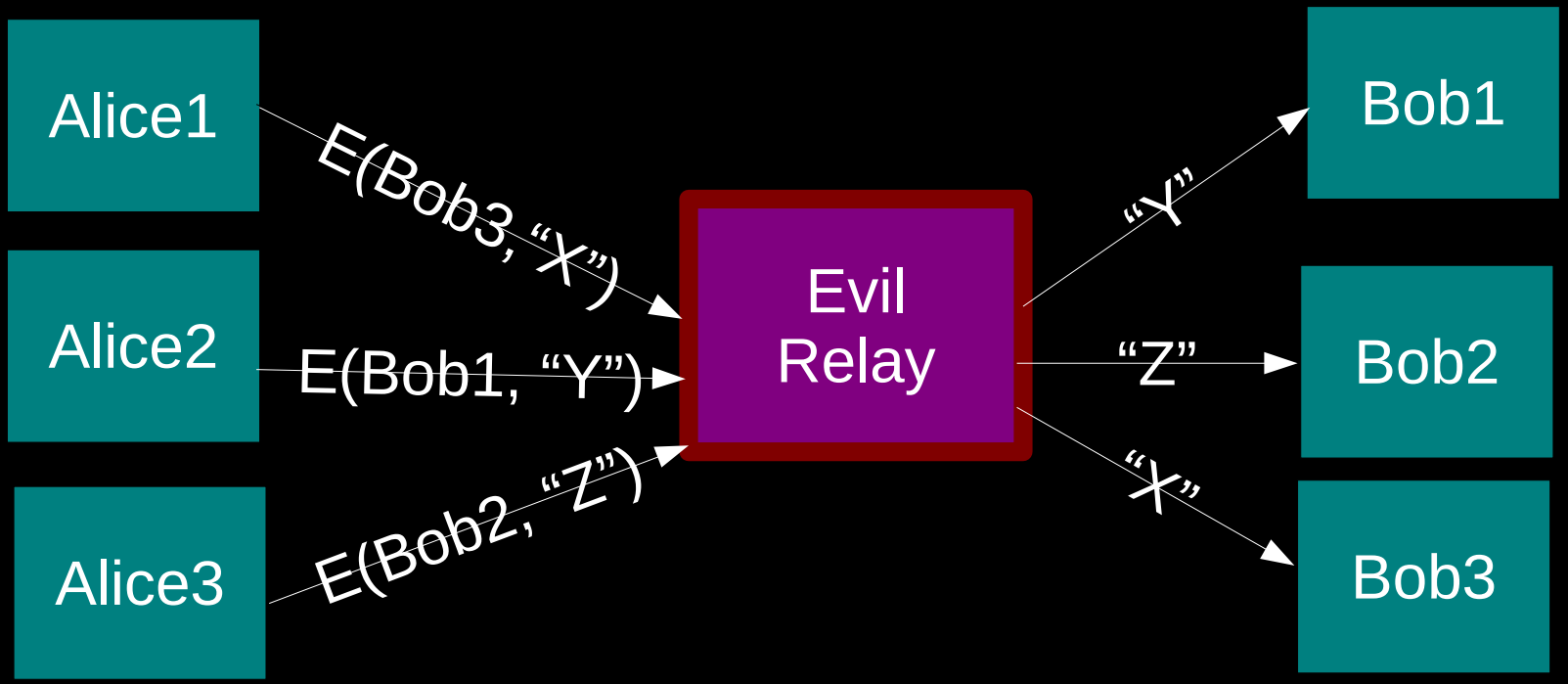

Estimated 300,000
daily Tor users

The simplest designs use a single relay to hide connections.



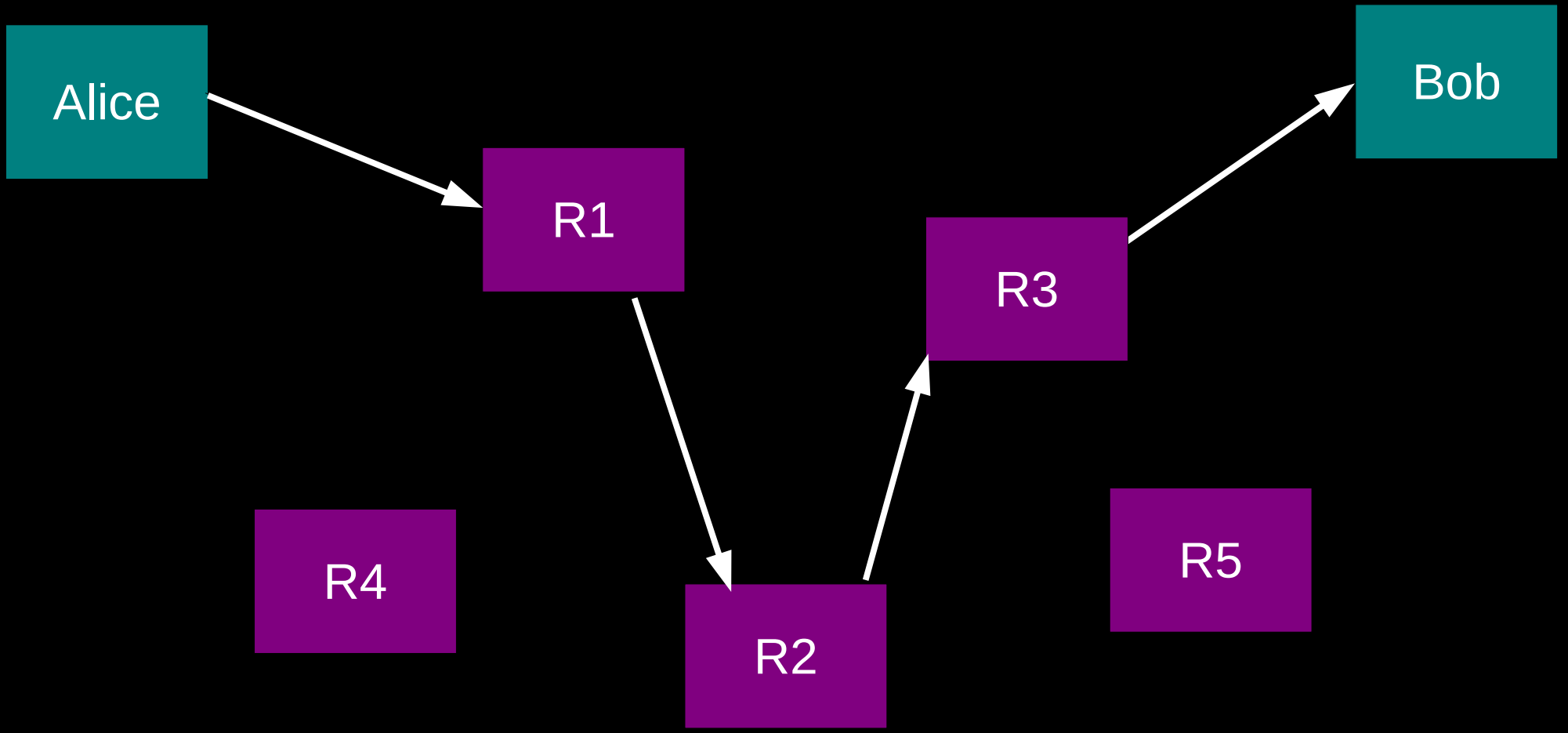
(example: some commercial proxy providers)

But a single relay is a single point of failure.

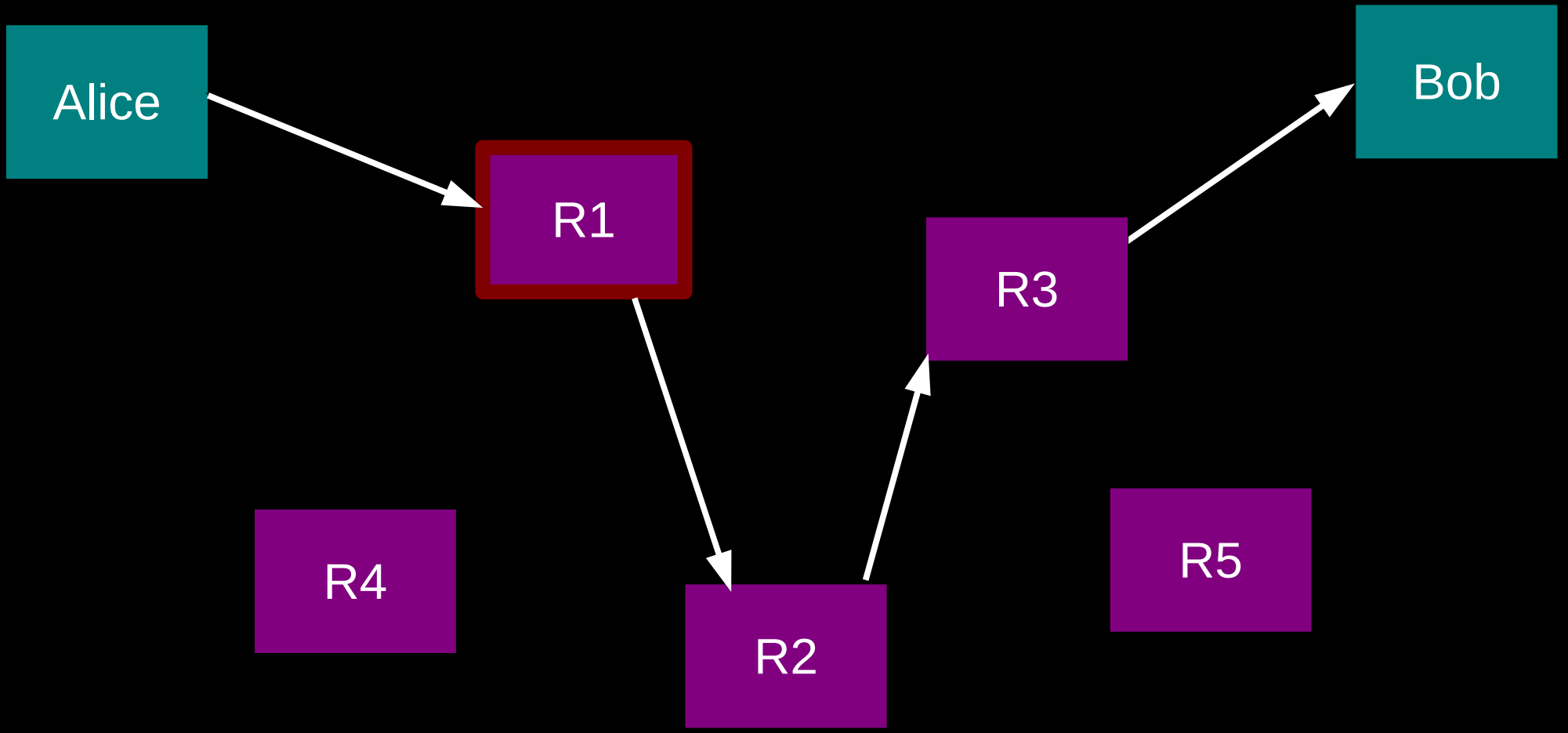


Eavesdropping the relay works too.

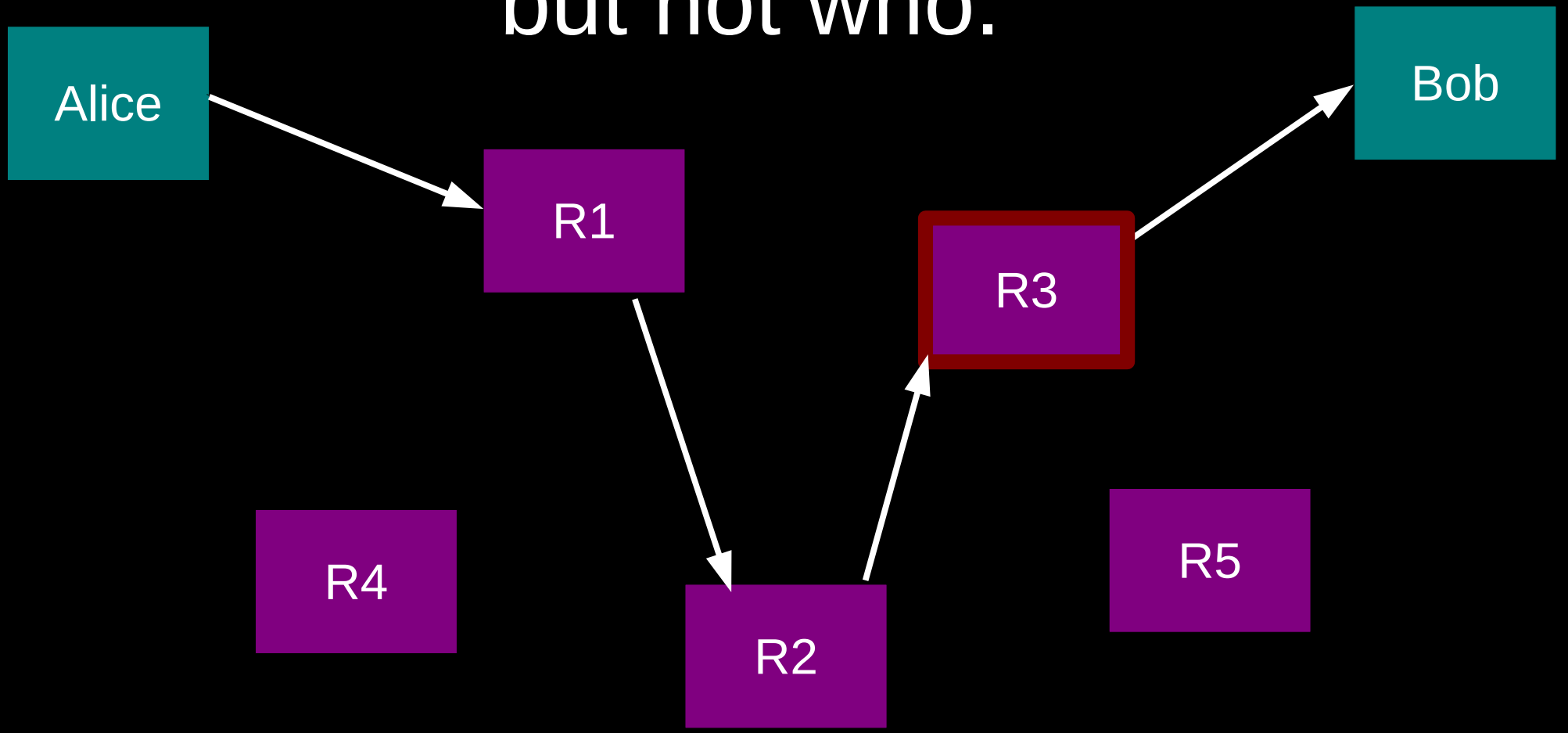
So, add multiple relays so that no single one can betray Alice.



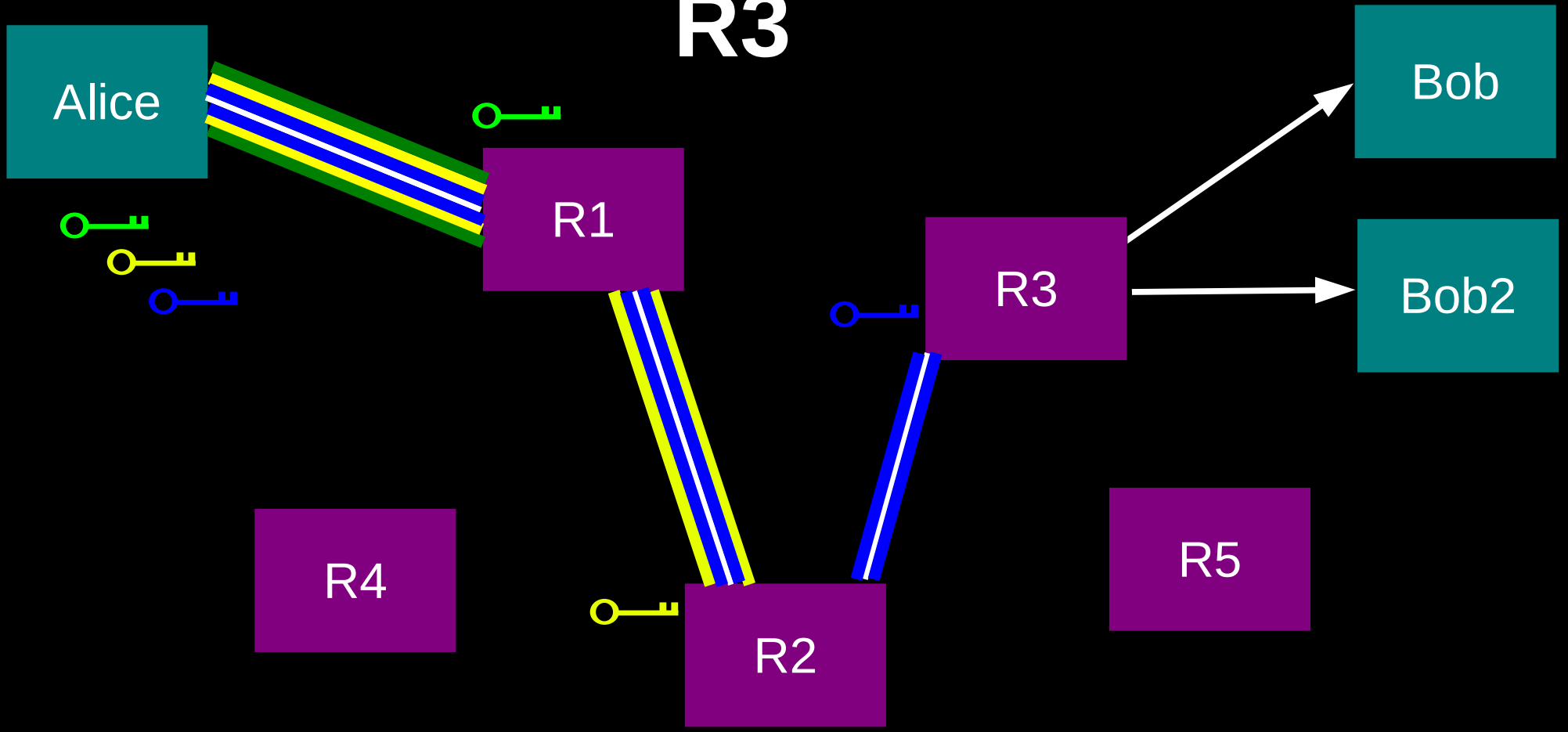
A corrupt first hop can tell that Alice is talking, but not to whom.



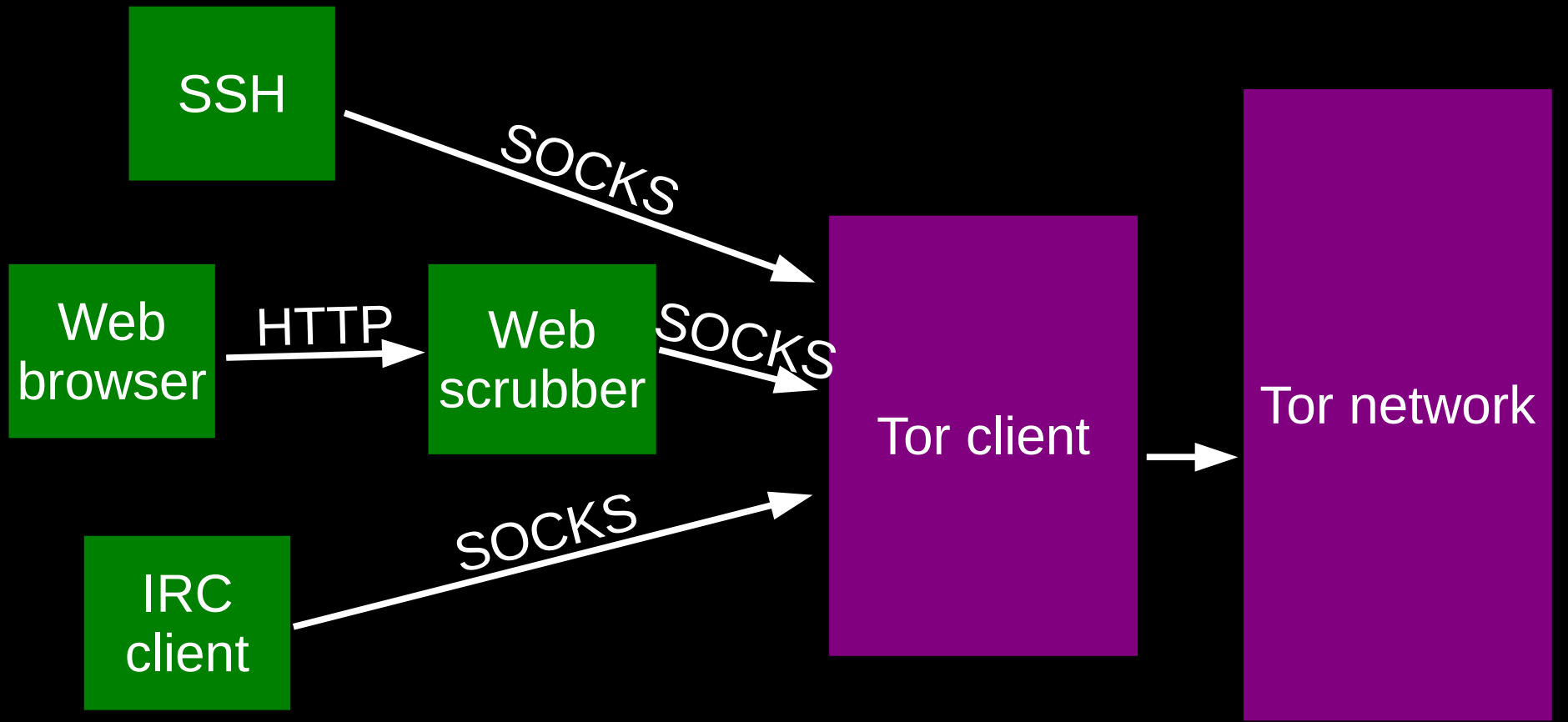
A corrupt final hop can tell that somebody is talking to Bob, but not who.



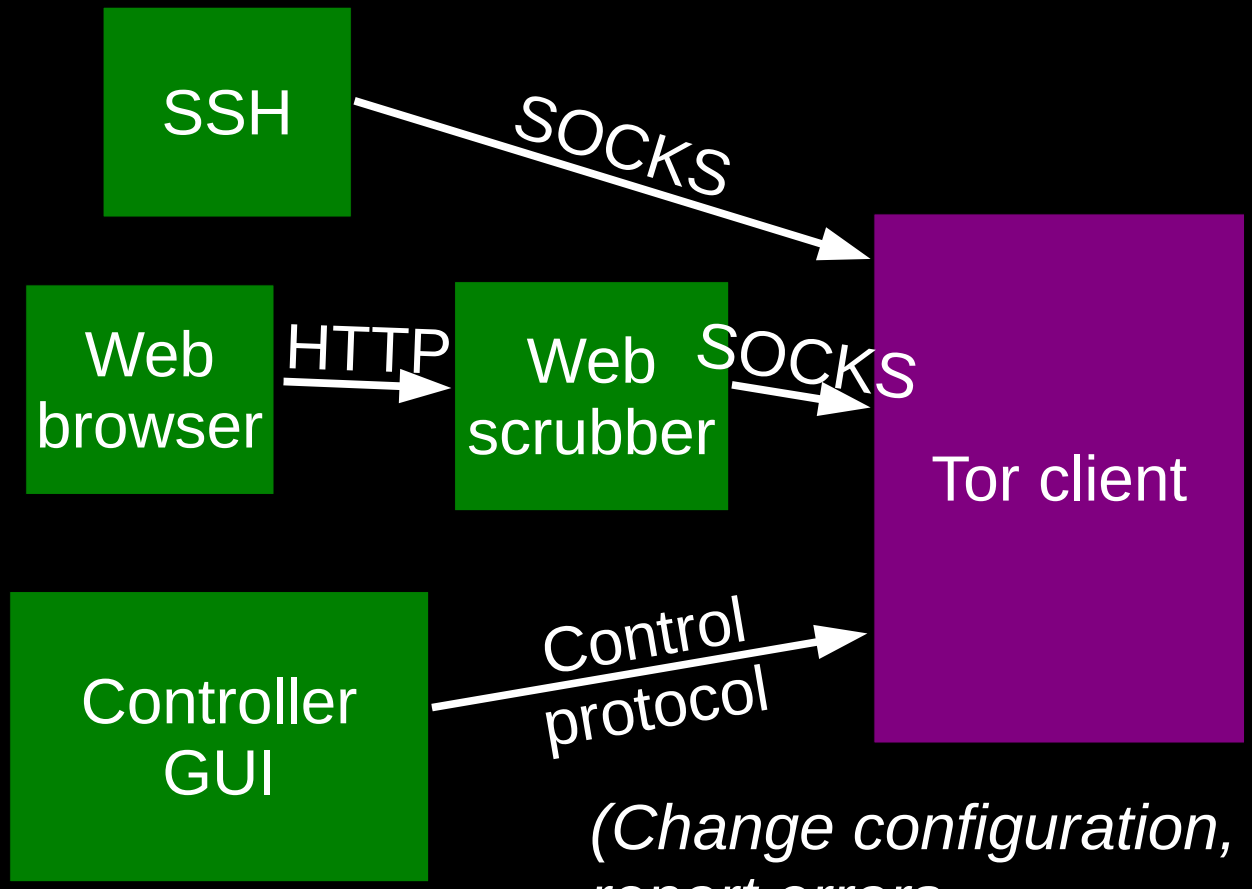
Alice makes a session key with R1
...And then tunnels to R2...and to R3



Tor anonymizes TCP streams only: it needs other applications to clean high-level protocols.



We added a control protocol for external GUI applications.



Refresh | Zoom In | Zoom Out | Zoom To Fit | Full Screen | Help | Close

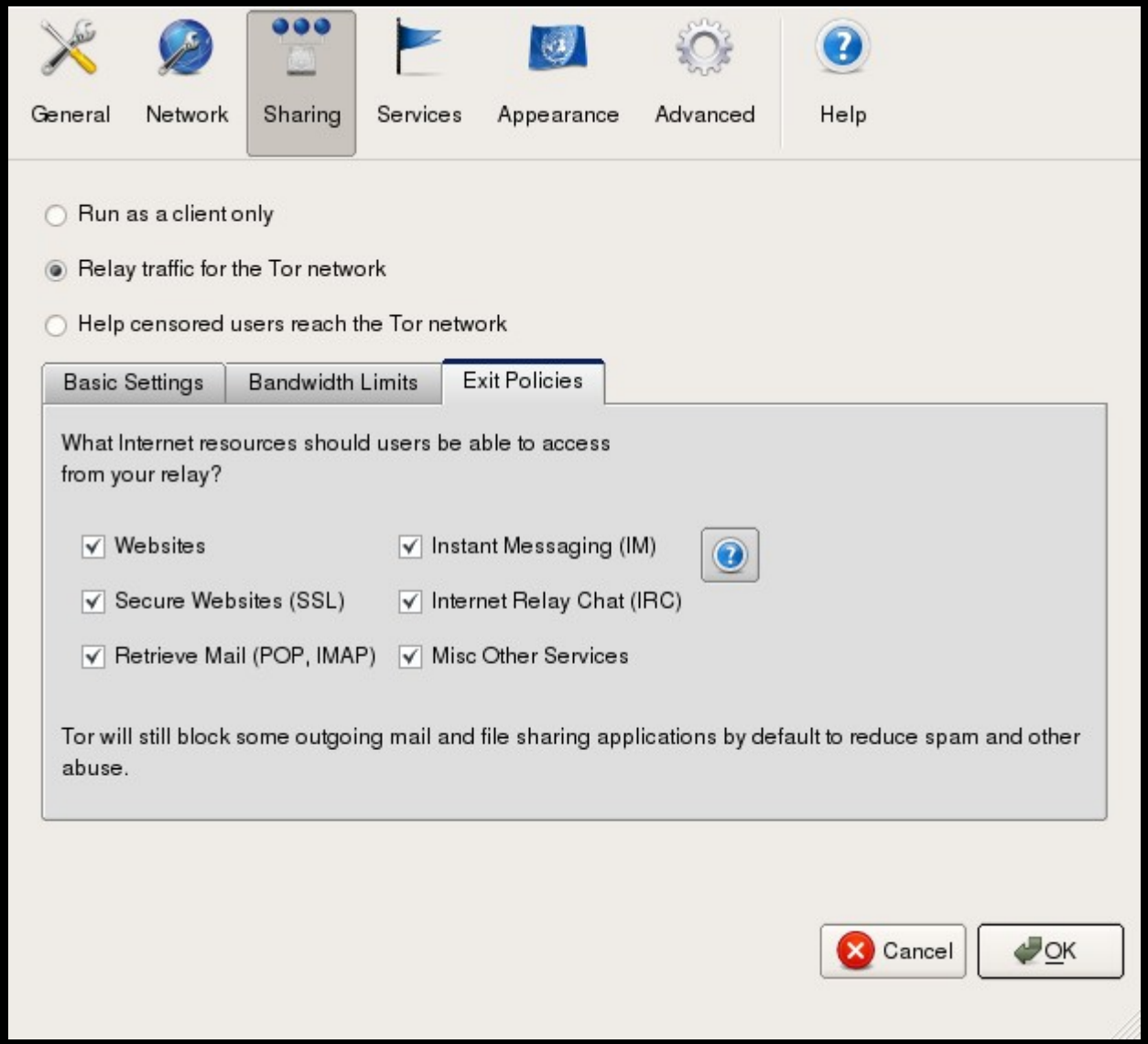
Relay
Tor1
trusted
threeletteragency
myrnaloy
UBIT2
jalopy
Piratenschatzi
Tonga
jceaovh
Lifuka
nixnix
LocomortoFTGR
SEC
teunTest
FoeBuD3
UK2
gpfTOR3
TORelay
gpfTOR4
tornodeviennasil
Butterfly
desync
BostonUCompSci
dizum
tischfuehrer
MopperSmurf
vallenator
CriptoLabTOR...
anansainthesp...
c03d9ebf
Atlantis
XS2Loli
blutmagie

Connection	Status
necrid,bmwanon2,trusted	Open
necrid,teunTest,kurac	Open
necrid,mailus,nixnix	Open
necrid,jgilje,Unnamed	Open
necrid,apogee,iria	Open
necrid,mailus,thebuckflowshere	Open
necrid,BDDF6D5EFAE2FAAF,trithnt	Open
necrid,f5c8dd93013406e7,conf555...	Open
necrid,C8063D26,jms1	Open

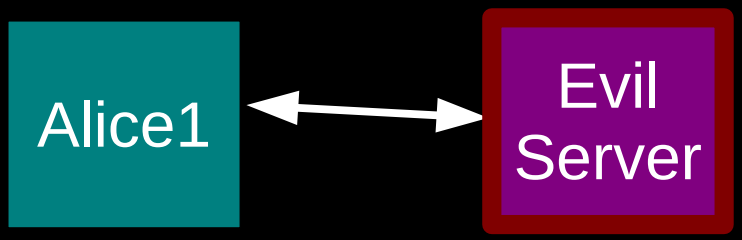
Usability for server operators is key.

- Rate limiting: eating too much bandwidth is rude!
- Exit policies: not everyone is willing to emit arbitrary traffic.

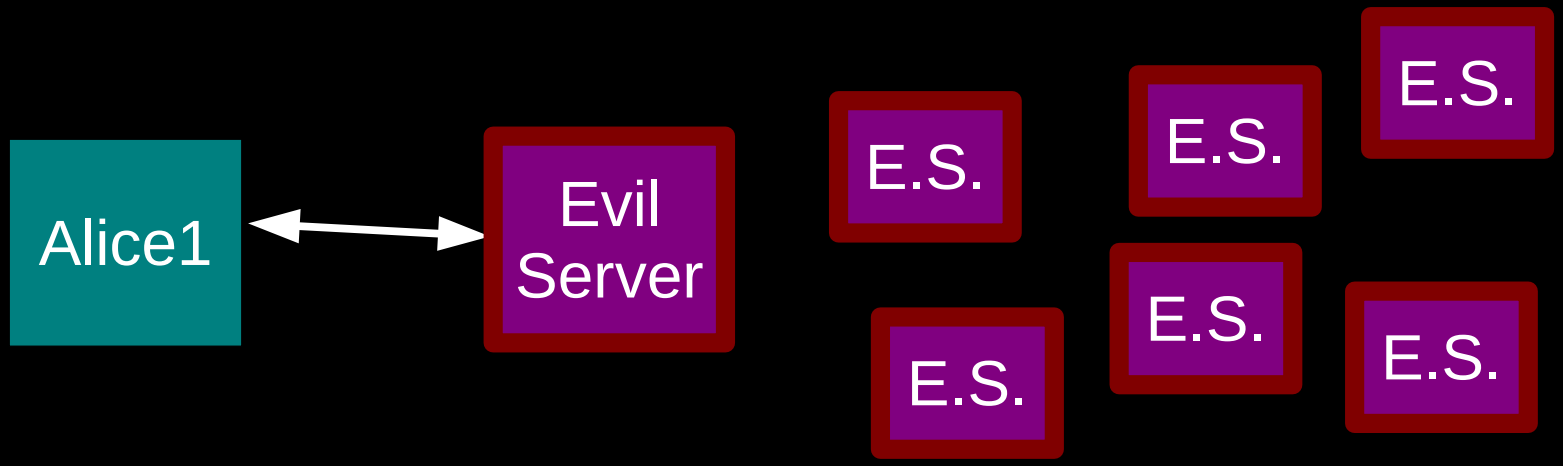
```
allow 18.0.0.0/8:*
allow *:22
allow *:80
reject *:*
```


General
Network
Sharing
Services
Appearance
Advanced
Help
Run as a client only
Relay traffic for the Tor network
Help censored users reach the Tor network
Basic Settings
Bandwidth Limits
Exit Policies
What Internet resources should users be able to access from your relay?
Websites
Instant Messaging (IM)
Secure Websites (SSL)
Internet Relay Chat (IRC)
Retrieve Mail (POP, IMAP)
Misc Other Services
Tor will still block some outgoing mail and file sharing applications by default to reduce spam and other abuse.
Cancel
OK

Server discovery must not permit liars to impersonate the whole network.

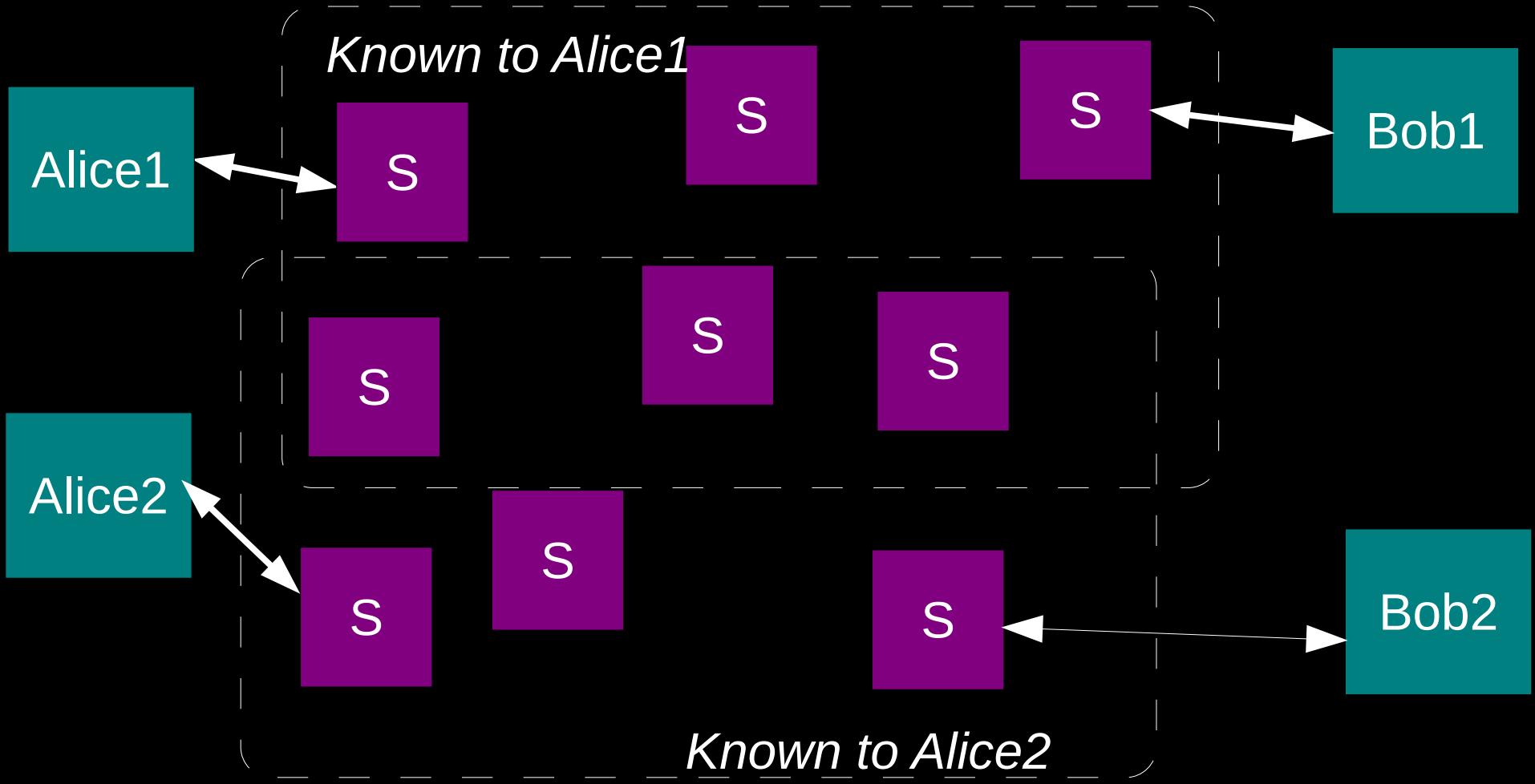


1. Alice says, “Describe the network

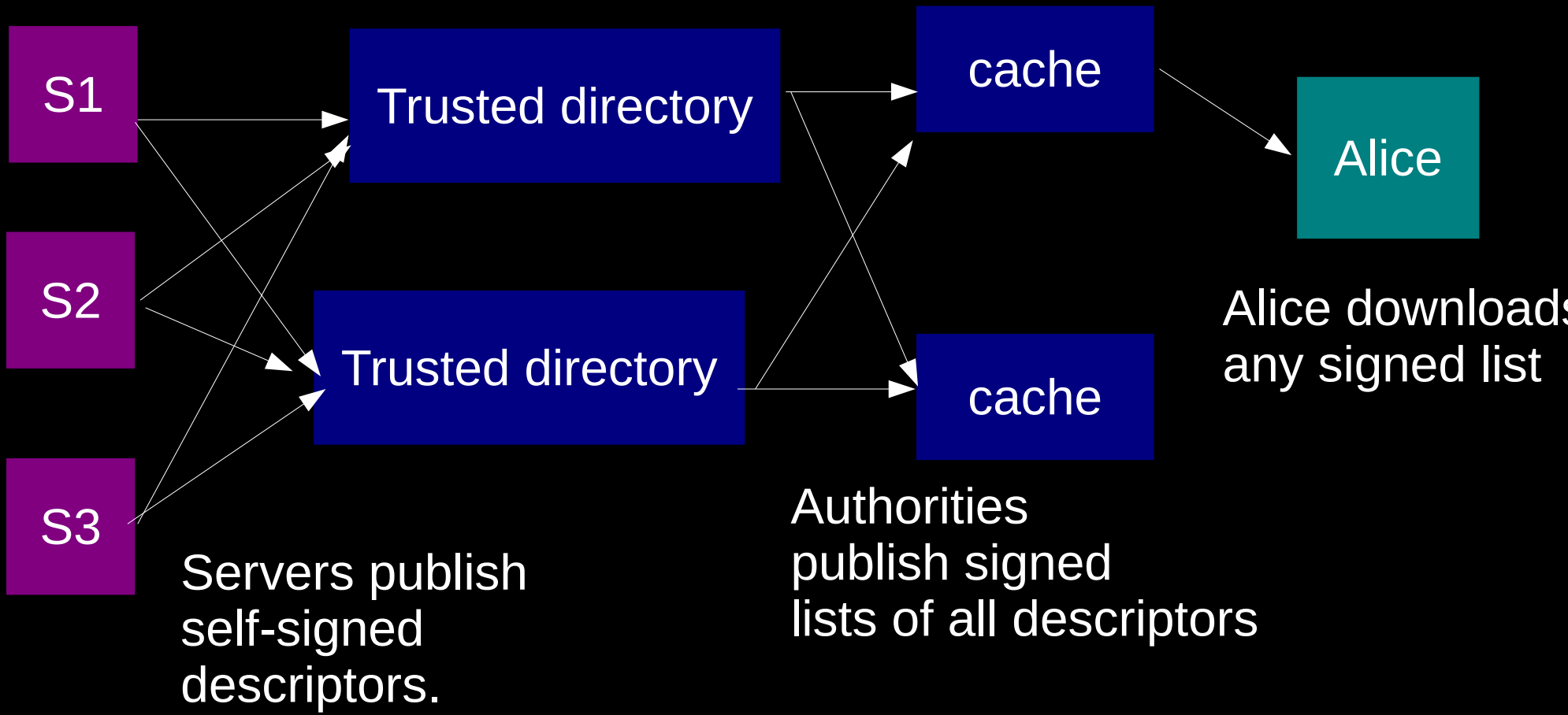


2. Alice is now in trouble.

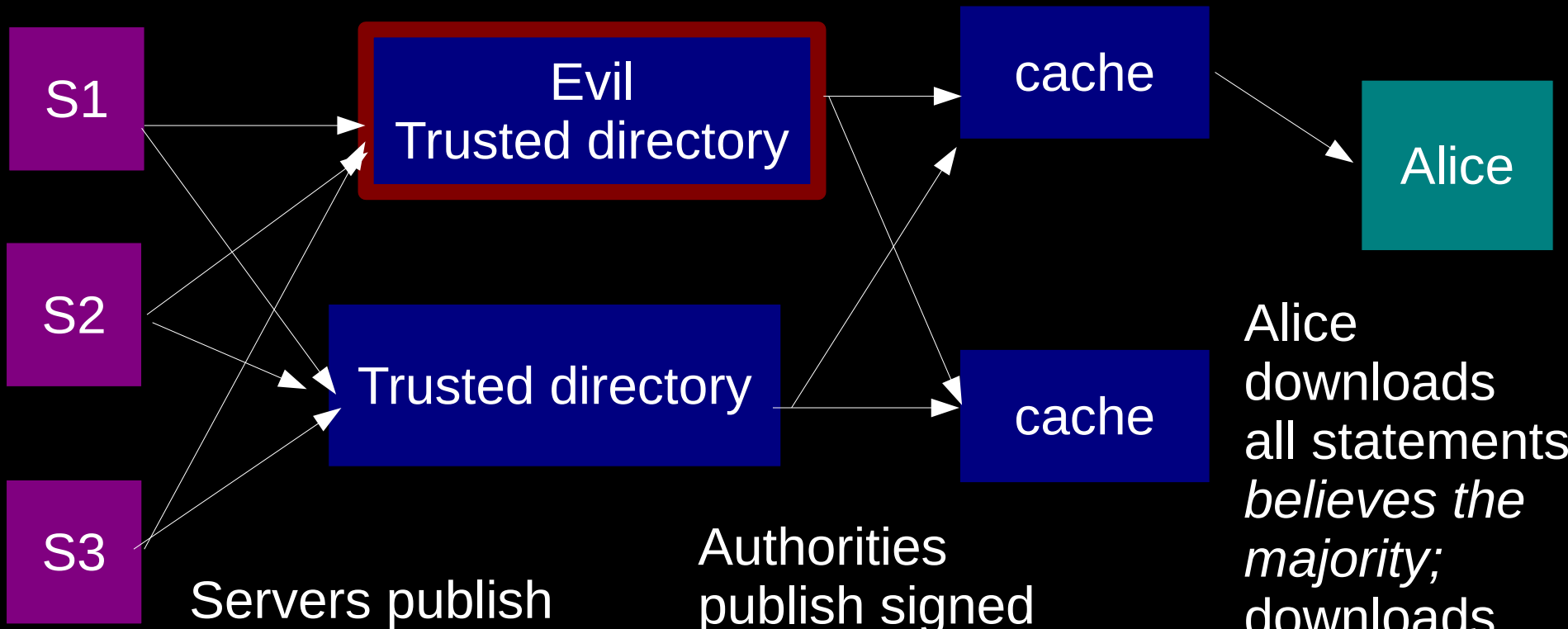
Server discovery is hard because misinformed clients lose anonymity.



Early Tor versions used a trivial centralized directory protocol.



We redesigned our directory protocol to reduce trust bottlenecks.



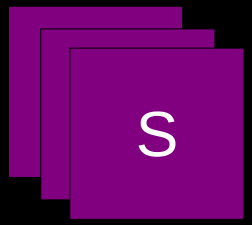
Servers publish self-signed descriptors.

Authorities publish signed statements *about* descriptors.

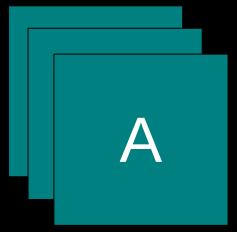
Alice downloads all statements; *believes the majority;* downloads descriptors as needed.

(Also uses less bandwidth!)

We're currently the largest strong anonymity network ever deployed.



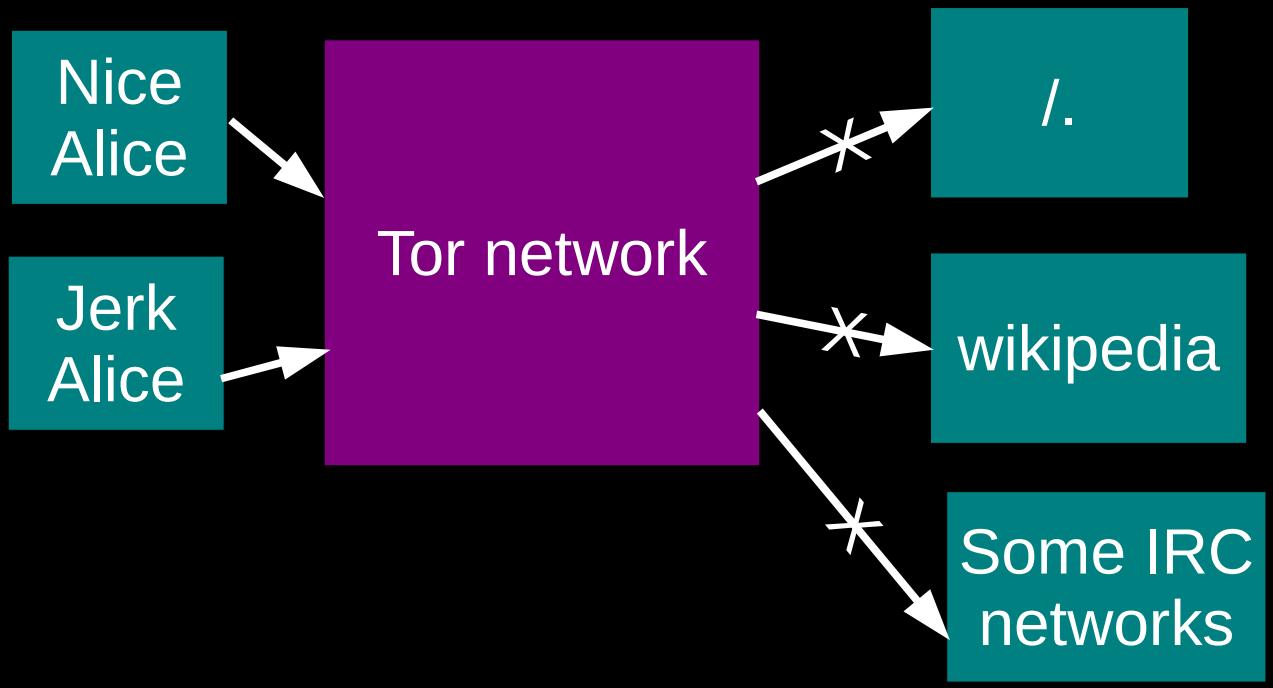
> 1500 running



> 300,000 in a day

> 125 MB/sec

Problem: Abusive users get the whole network blocked.



Minimize scope of blocking?

Other common abuses

- Somebody connects to Hotmail, and sends an obnoxious mail.
- Somebody connects to IRC and yells -> DDoS on Tor exit server.
- Somebody tries to get you shut down by connecting to Google Groups and posting spam.
- Somebody uses Tor to download a movie, and your ISP gets a DMCA takedown.

Who uses Tor?

- Normal people
- Law Enforcement
- Human Rights Activists
- Business Execs
- Militaries
- Abuse Victims
- https://torproject.org/torusers

- Tor doesn't magically encrypt the Internet
- Operating Systems and Applications leak your info
- Browser Plugins, Cookies, Extensions, Shockwave/Flash, Java, Quicktime, and PDF all conspire against you

Outline

- Why anonymity?
- Crash course on Tor
- *Future*

Community

- Many tools make a big splash in the press
  - Censors need to feel in control; publicity removes the appearance of control
- Increase community diversity
  - Strong social network
- Funding
  - Donations, grants, contracts

3-Year Development Roadmap

- Improve Performance
- Client Safety
- Ease of Use and Understanding
- Core Research & Development

https://torproject.org/press/ for details

Copyrights

- who uses tor? http://www.flickr.com/photos/mattw/2336507468/siz , Matt Westervelt, CC-BY-SA
- danger!, http://flickr.com/photos/hmvh/58185411/sizes/o/ , hmvh, CC-BY-SA
- 300k, http://flickr.com/photos/tochis/1169807846/sizes/o/ , tochis, CC-BY-NC