Tor: Anonymous Communications for the Dept of Defense ... and you.

Roger Dingledine
The Free Haven Project
http://tor.eff.org/

Tor: Big Picture

- Freely available (Open Source), unencumbered.
- Comes with a spec and full documentation: German universities implemented compatible Java Tor clients; researchers use it to study anonymity.
- Chosen as anonymity layer for EU PRIME project.
- 200000+ active users.
- PC World magazine named Tor one of the Top 100 Products of 2005.

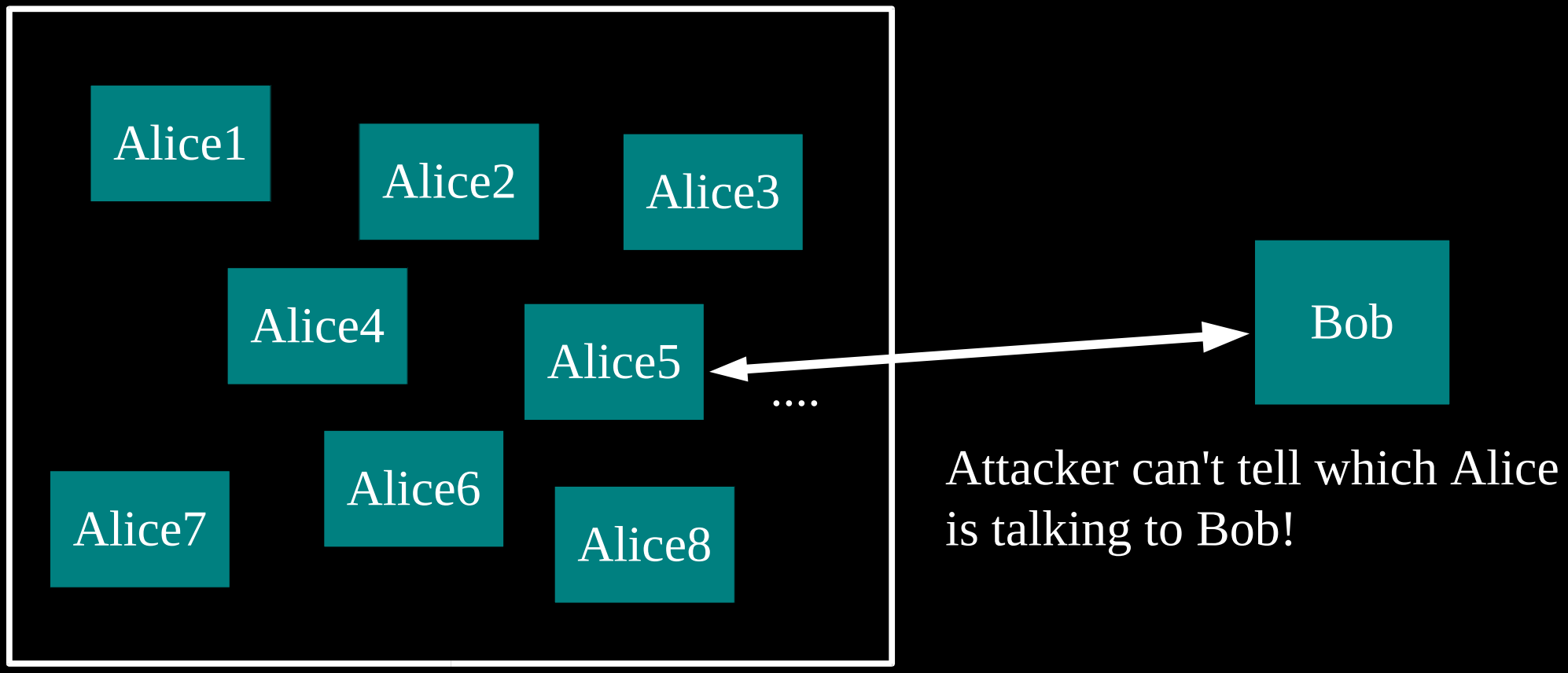
Informally: anonymity means you can't tell who did what

“Who wrote this blog post?”

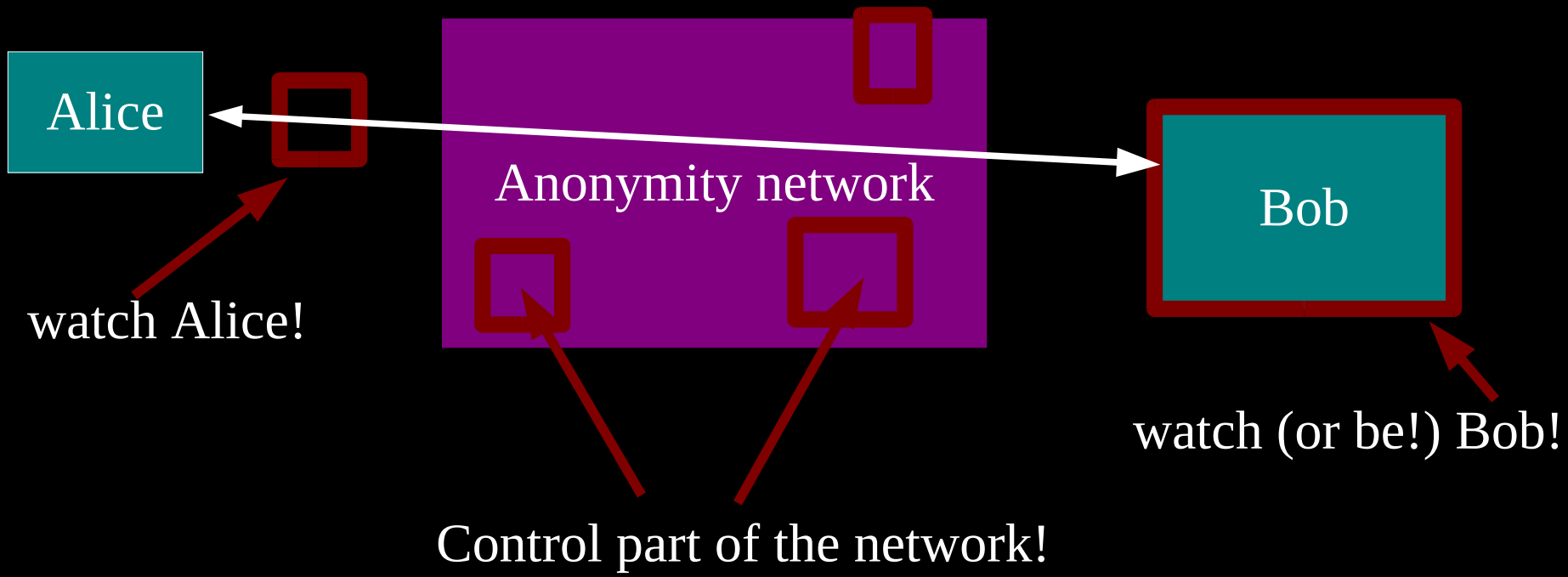
“Who's been viewing my webpages?”

“Who's been emailing patent attorneys?”

Formally: anonymity means indistinguishability within an "anonymity set"

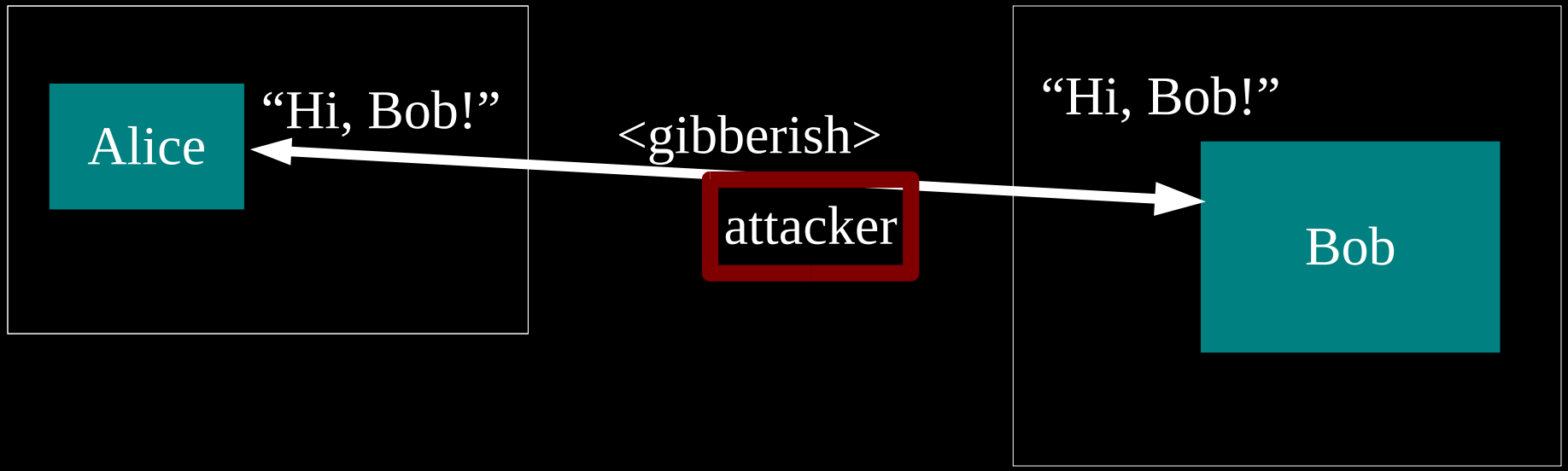


We have to make some assumptions about what the attacker can do.

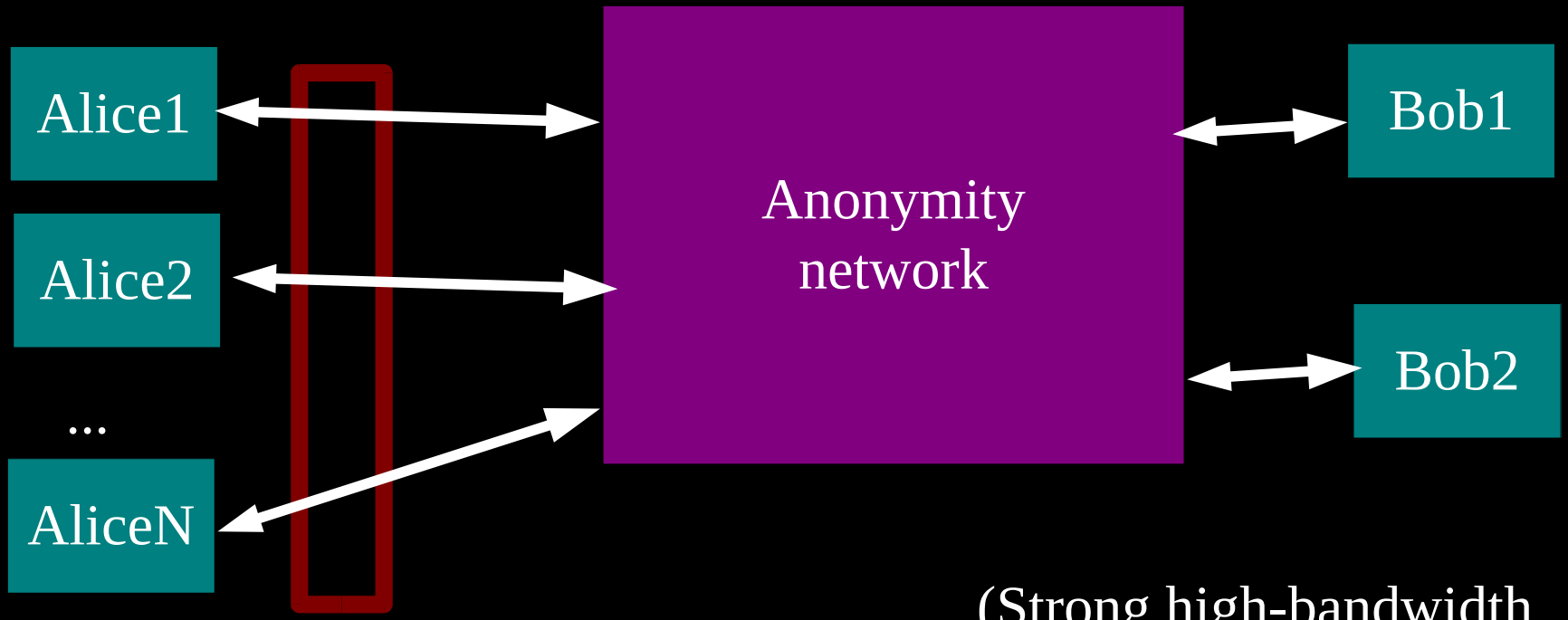


Etc, etc.

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!”~~

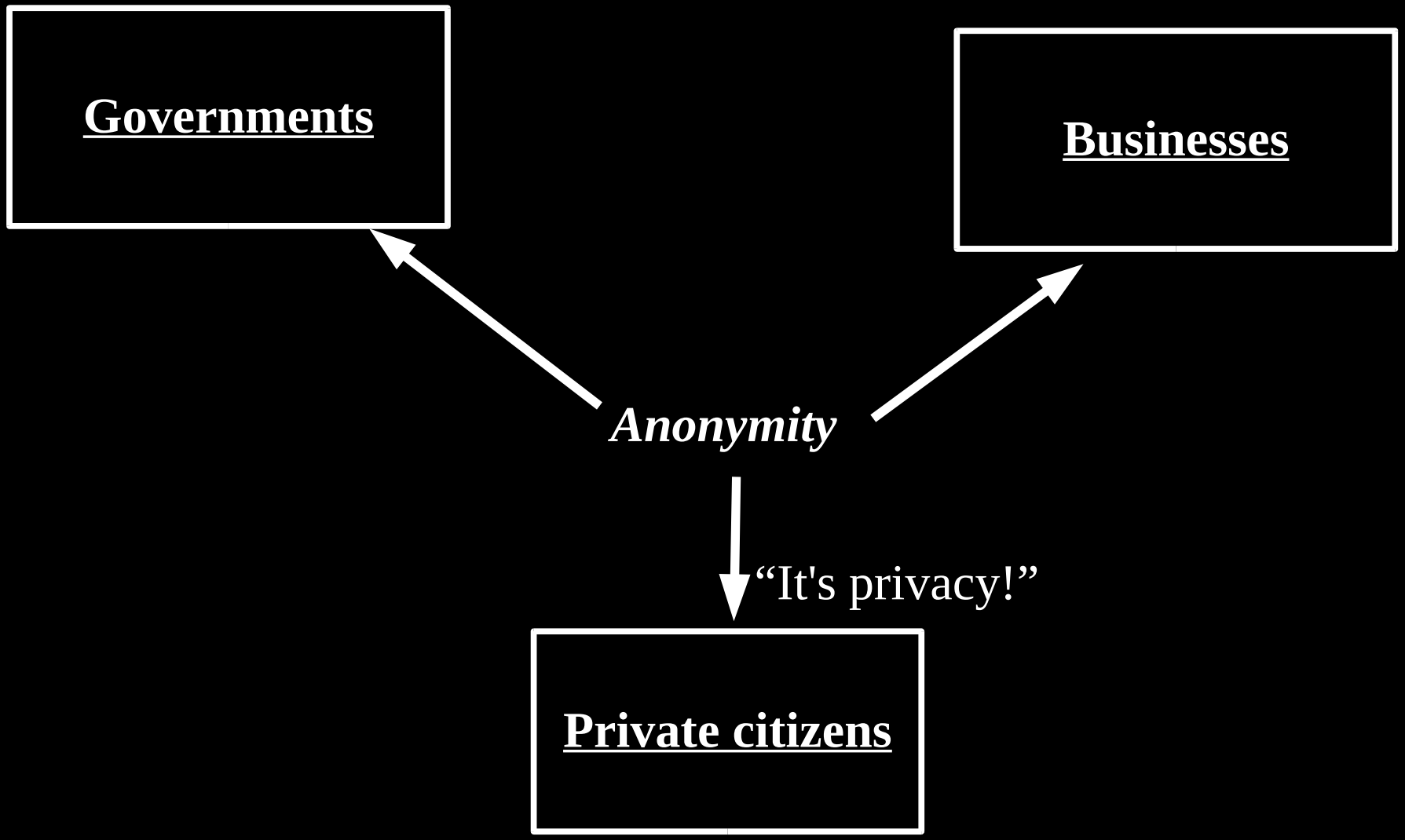
~~“I didn't write my name on it!”~~

Not what we're talking about.

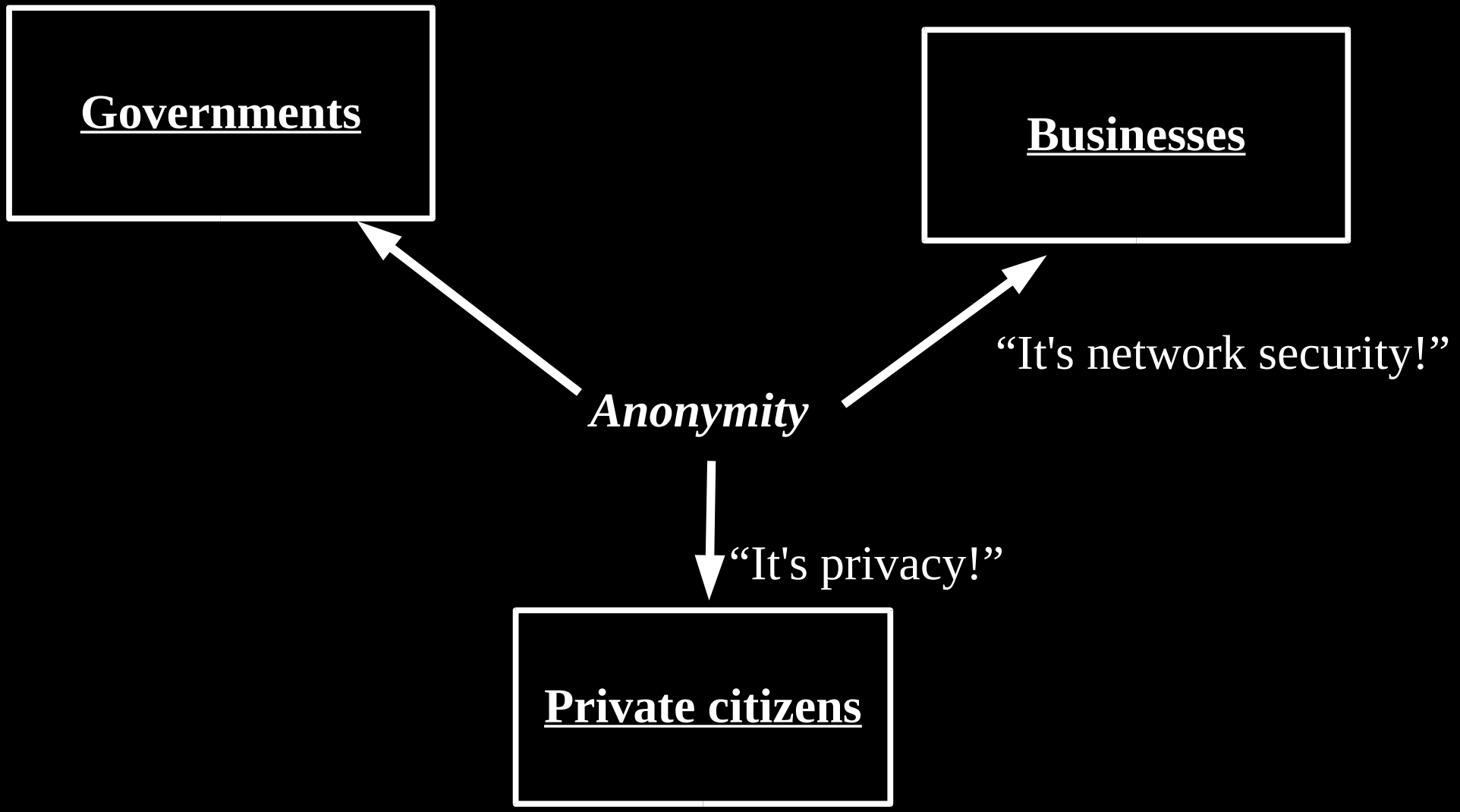
Nope!
(More info later.)

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

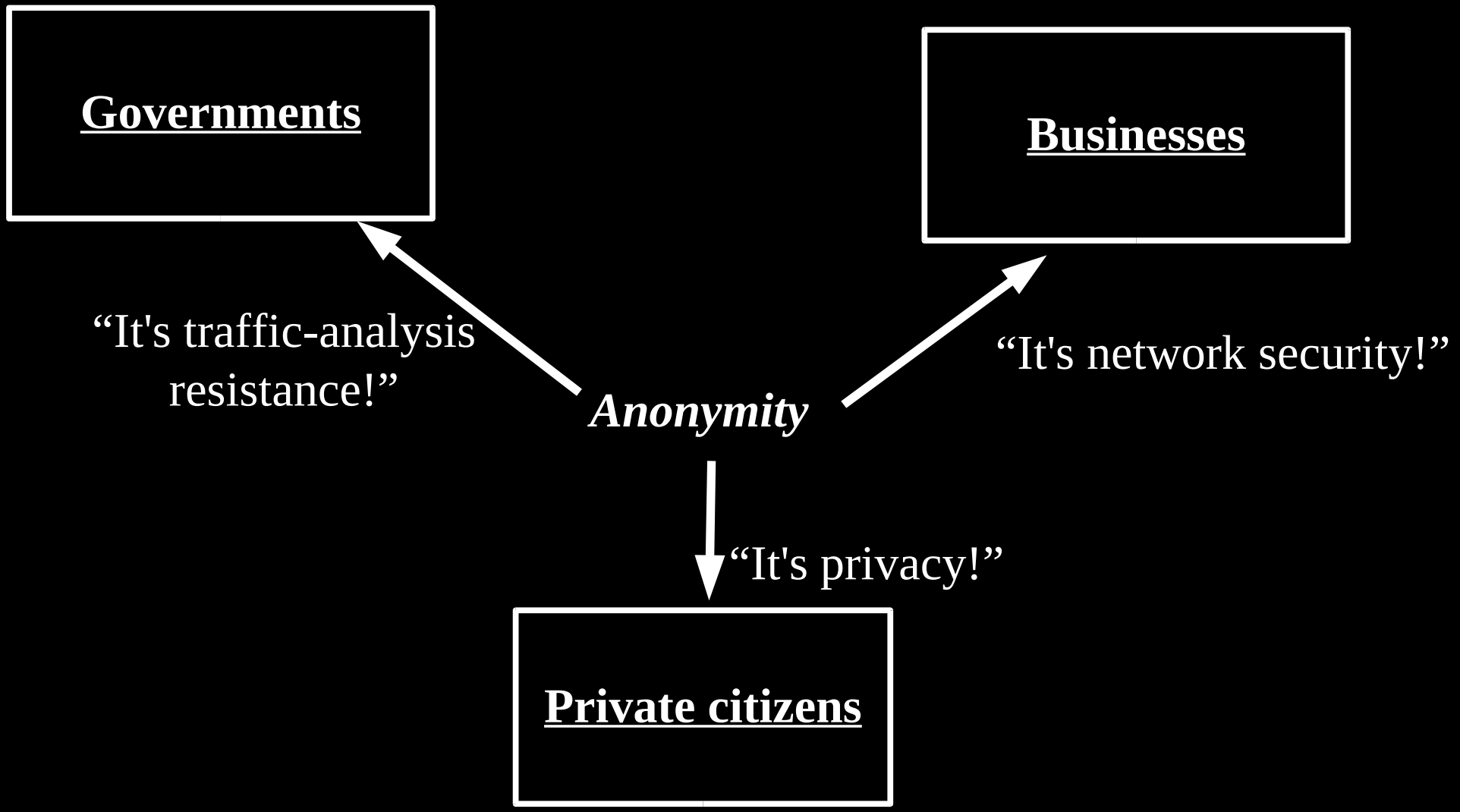
Anonymity serves different interests for different user groups.



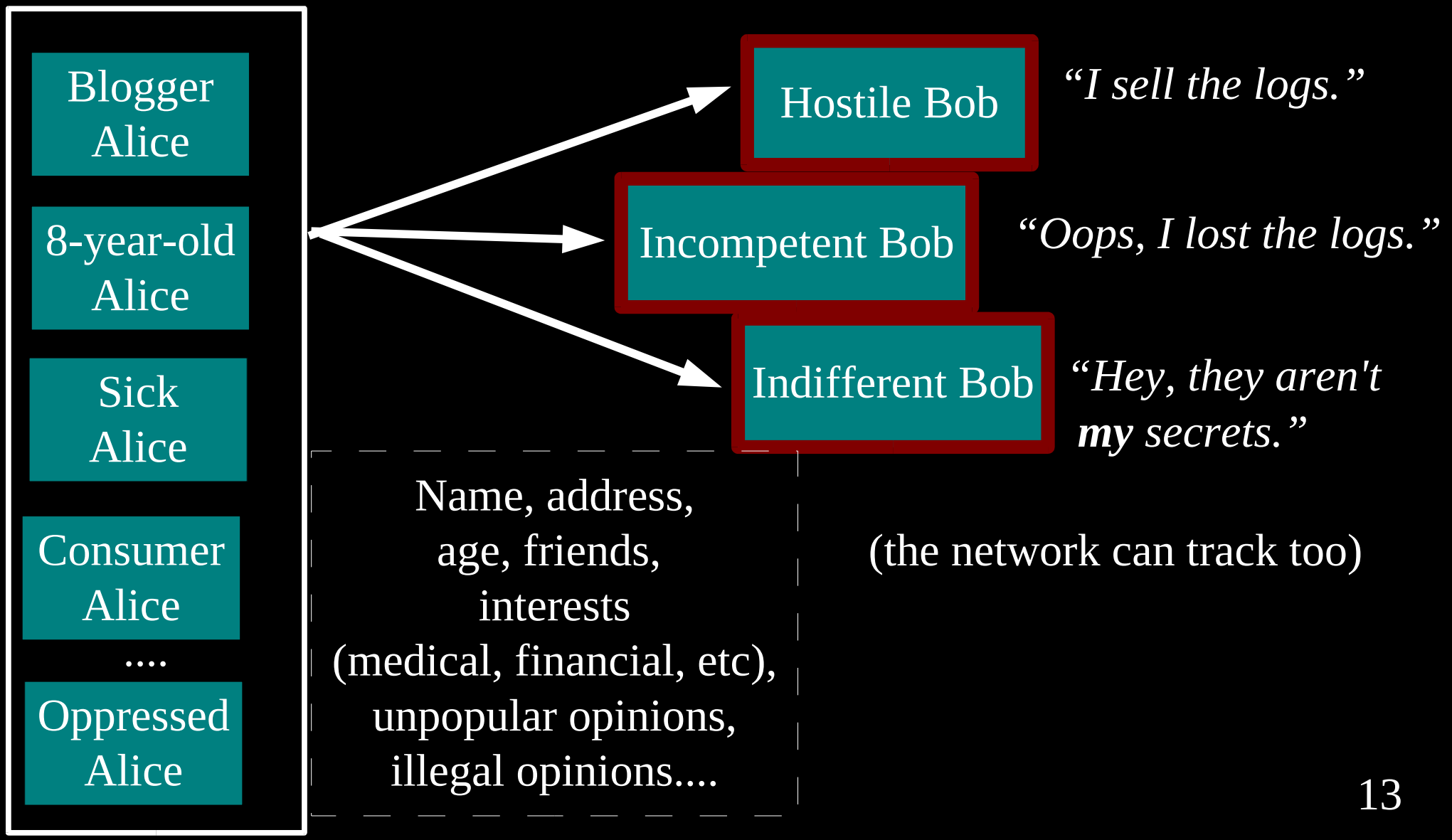
Anonymity serves different interests for different user groups.



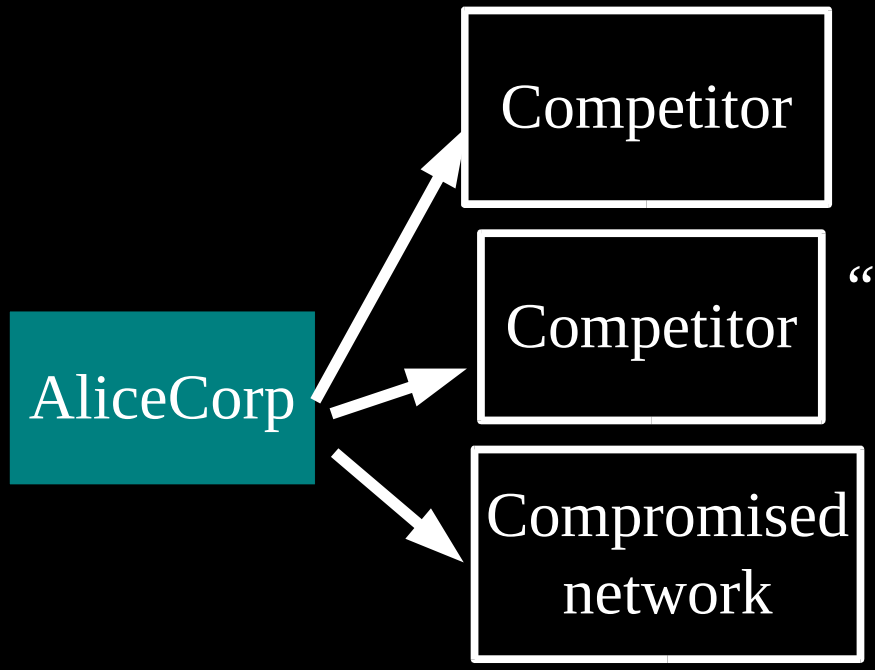
Anonymity serves different interests for different user groups.



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



Businesses need to keep trade secrets.



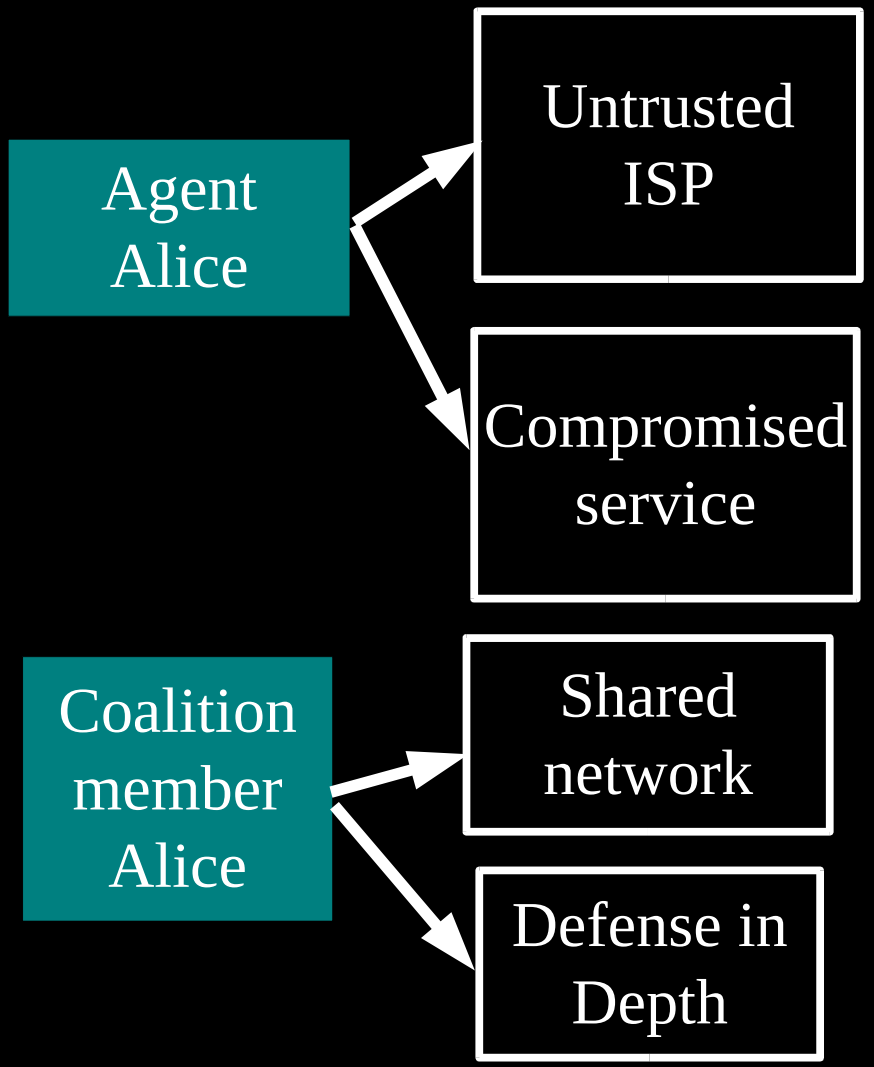
“Oh, your employees are reading our patents/jobs page/product sheets?”

“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



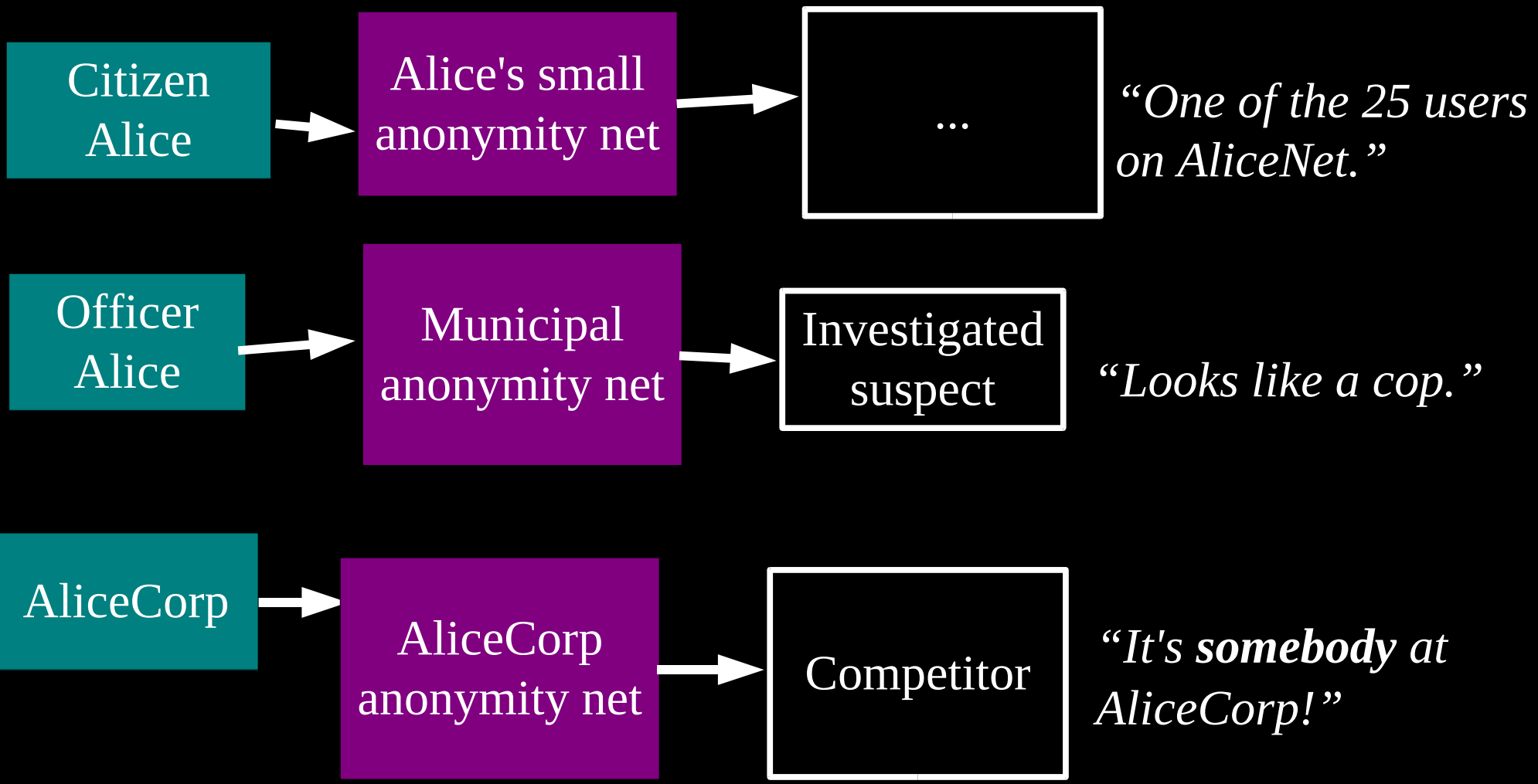
"What will you bid for a list of Baghdad IP addresses that get email from .gov?"

"What does *the CIA Google for?"*

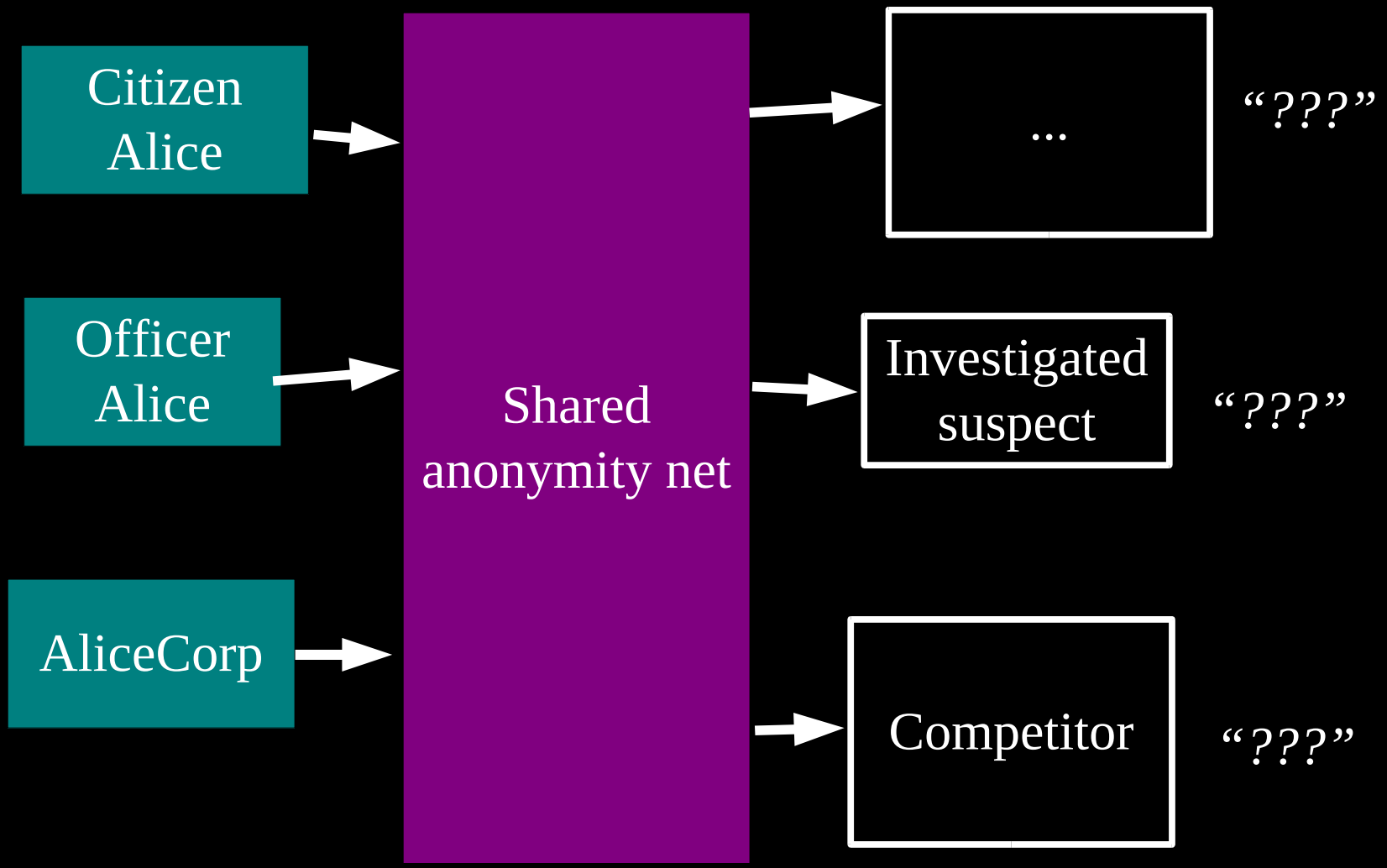
"Do I really want to reveal my internal network topology?"

"What about insiders?"

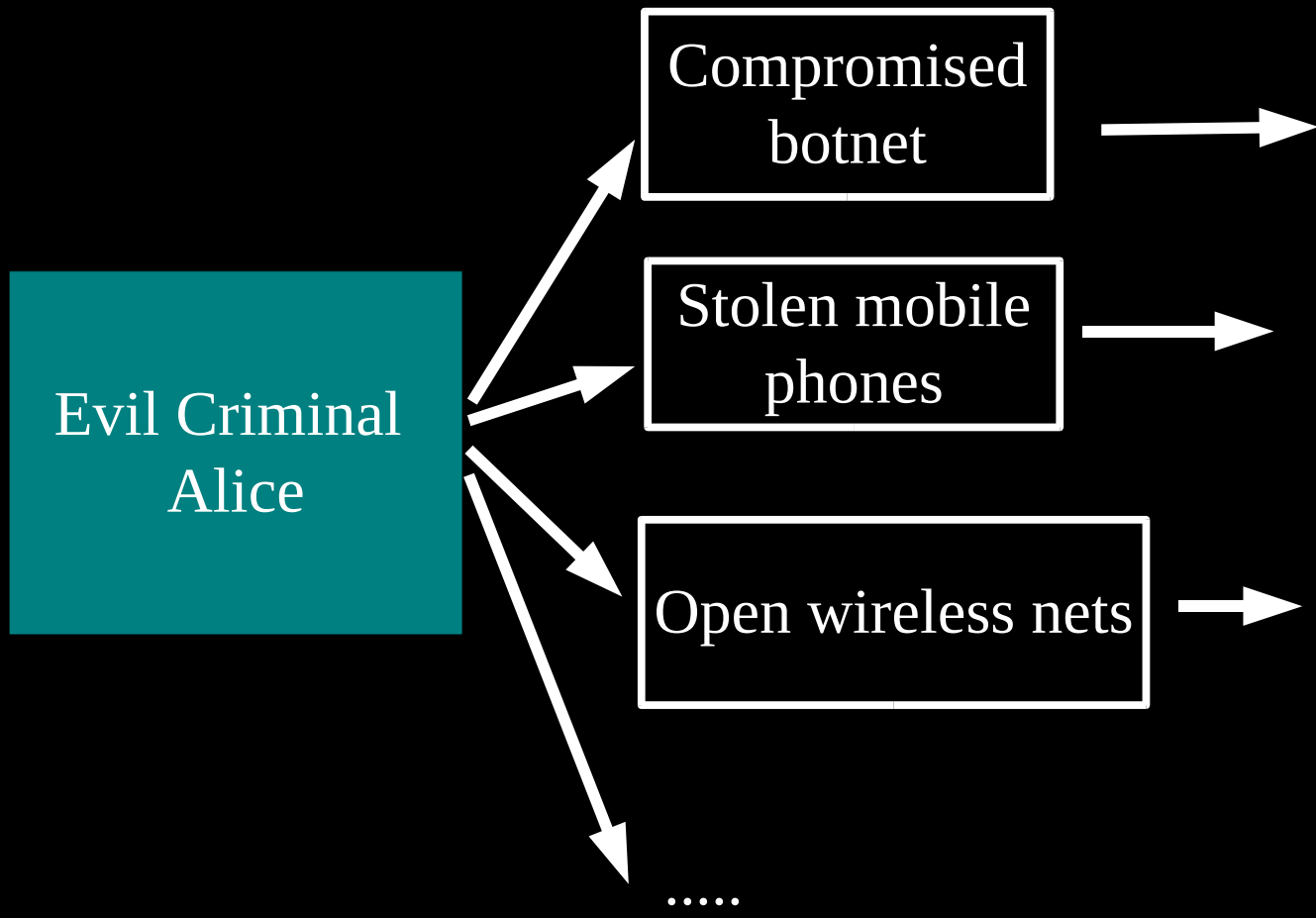
You can't get anonymity on your own: private solutions are ineffective...



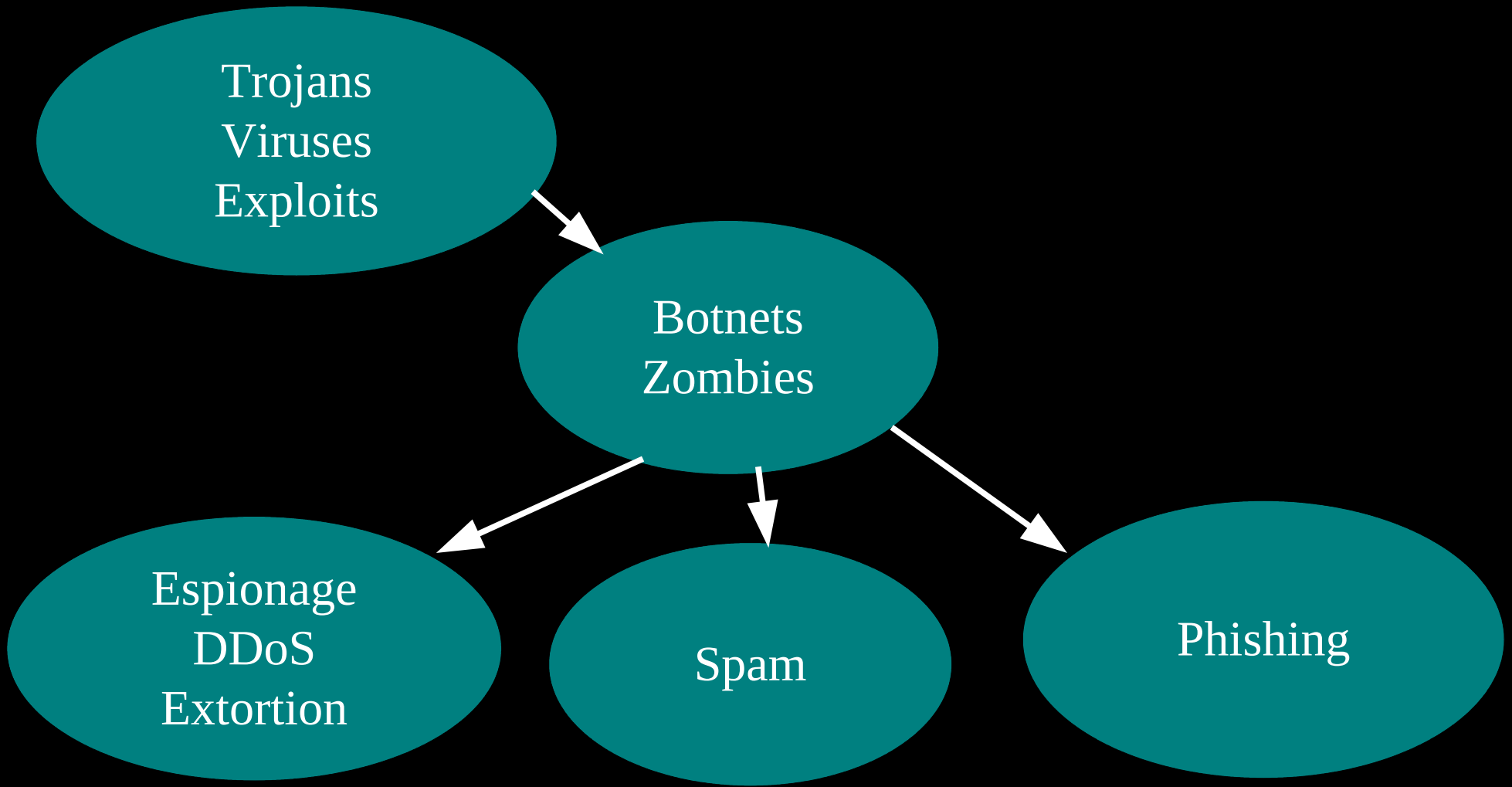
... so, anonymity loves company!



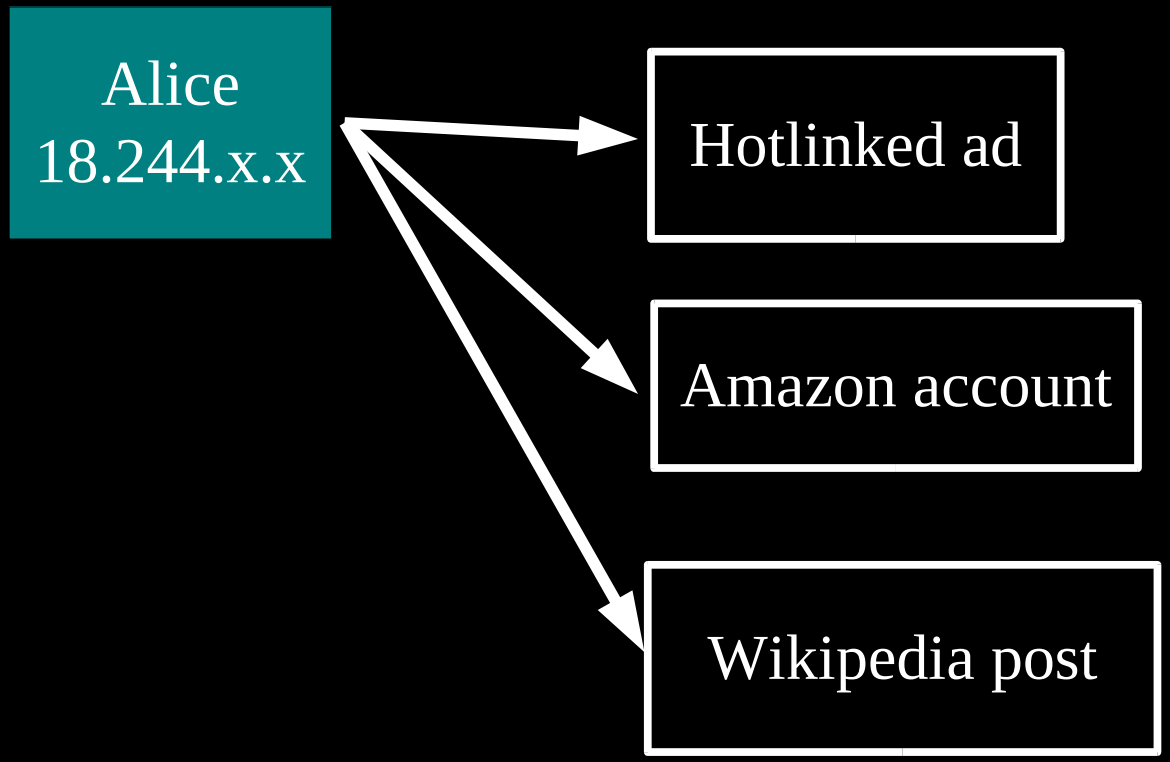
Yes, bad people need anonymity too. But they are *already* doing well.



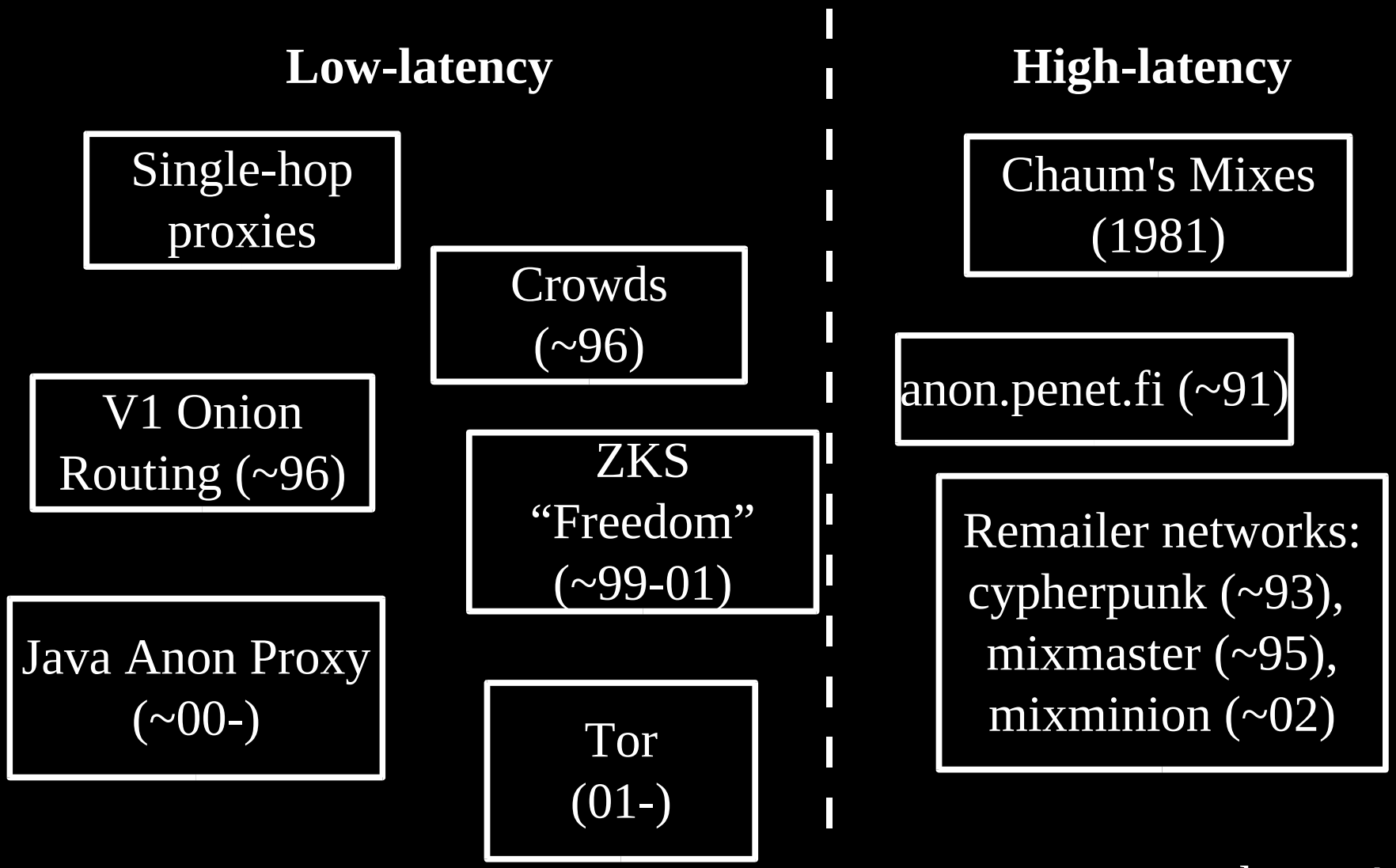
Current situation: Bad people on the Internet are doing fine



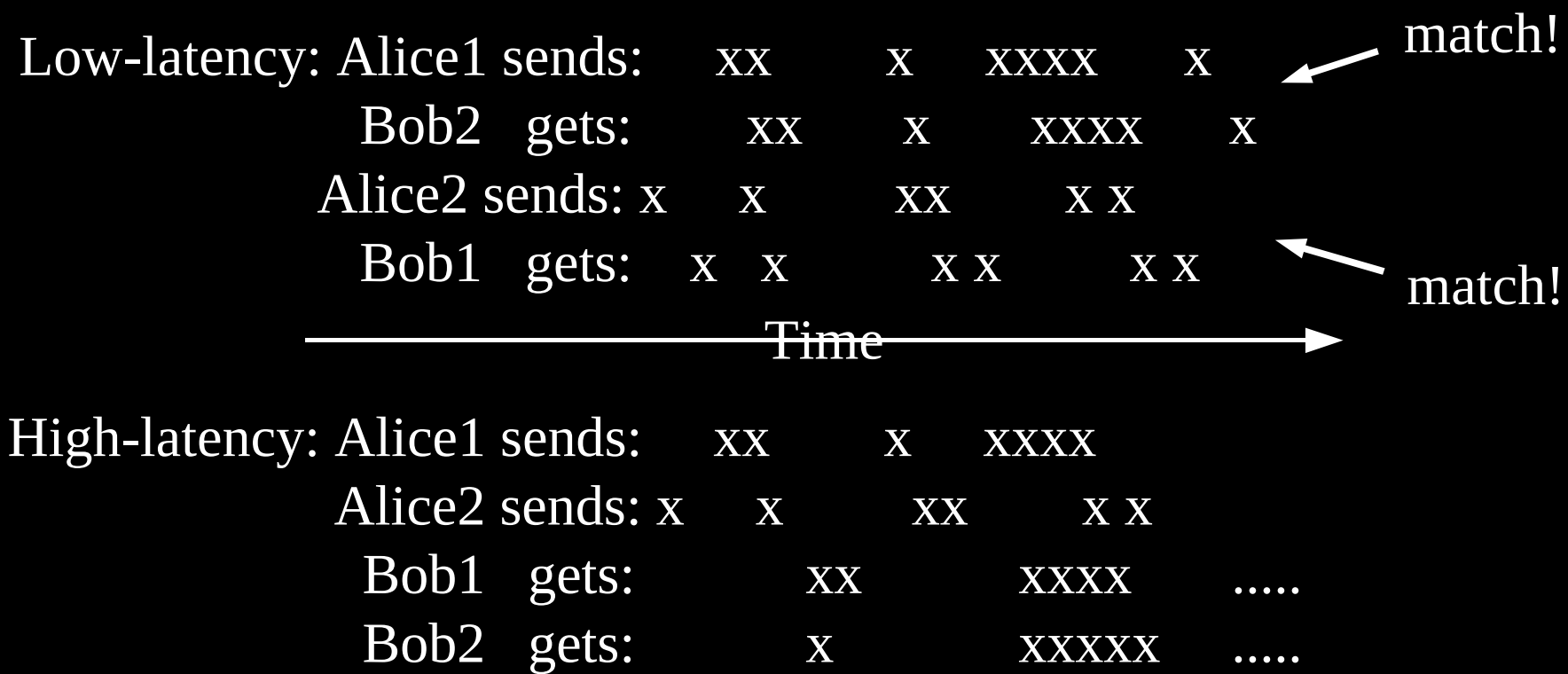
IP addresses can be enough to bootstrap knowledge of identity.



Tor is not the first or only design for anonymity.



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



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

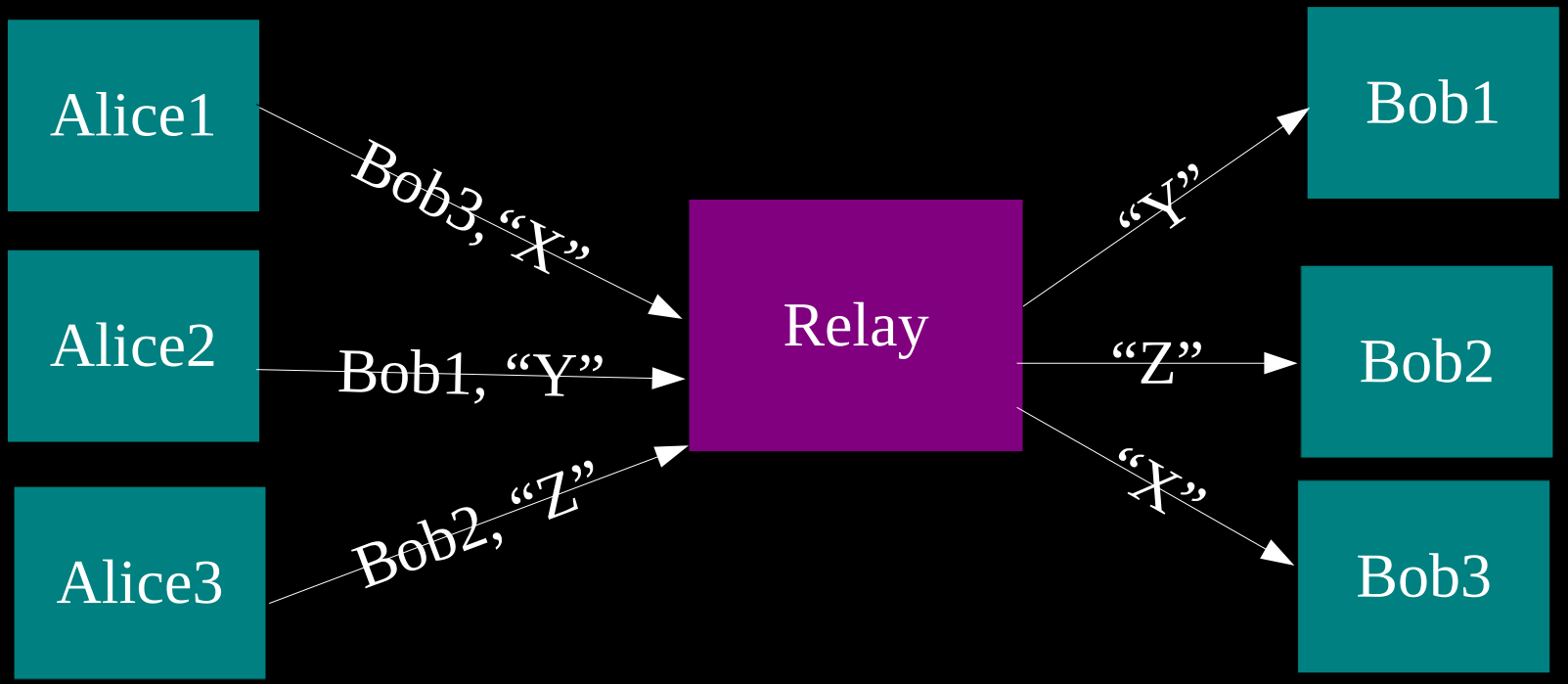
Still, we focus on low-latency, because it's more useful.

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?

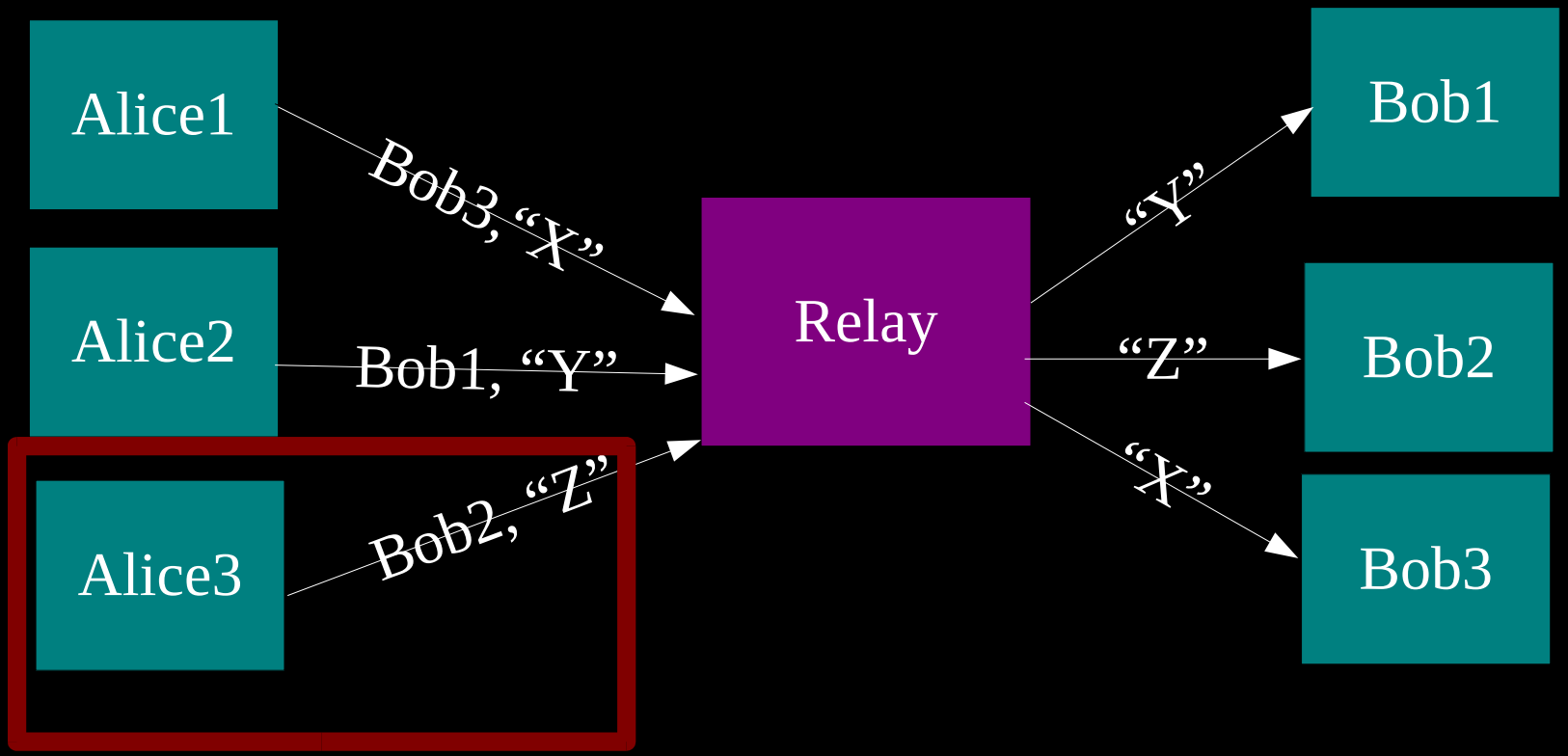
And if anonymity loves company....?

The simplest designs use a single relay to hide connections.

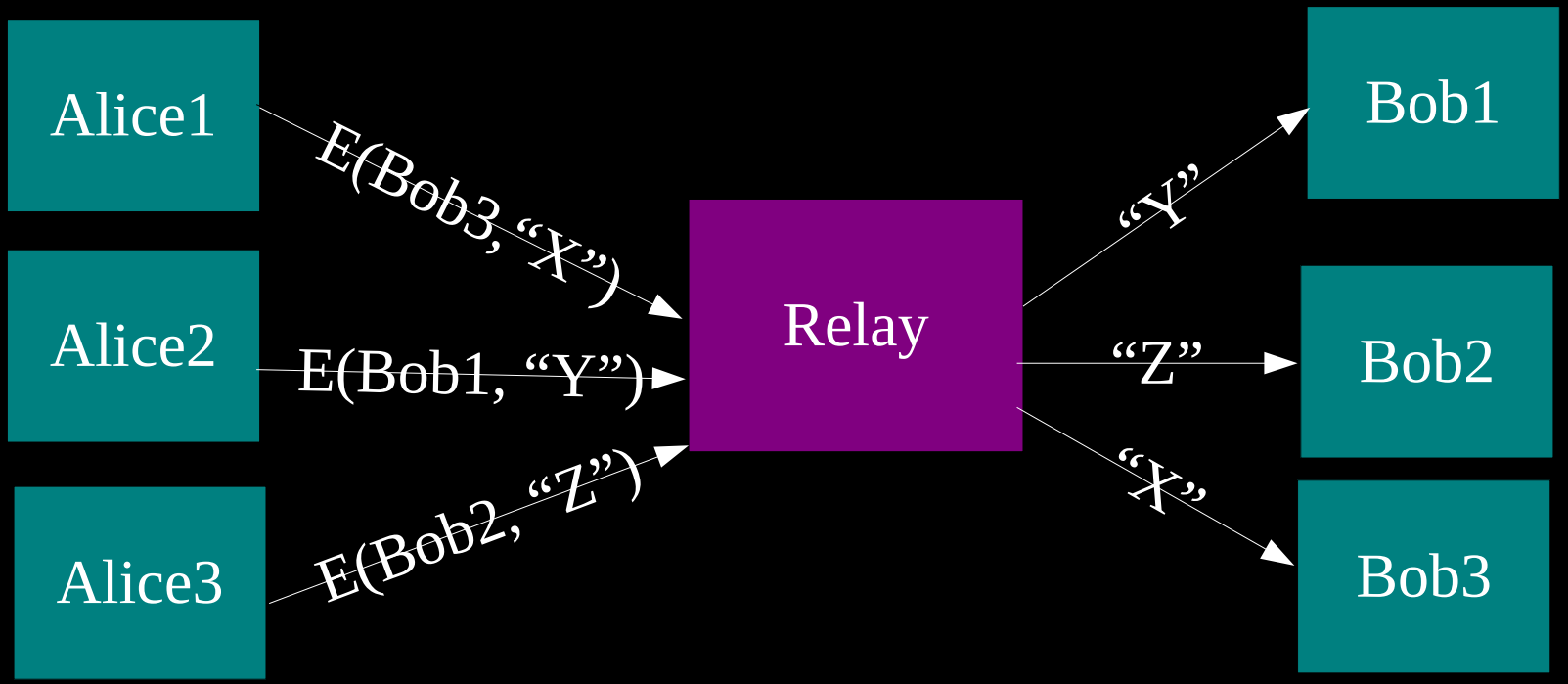


(ex: some commercial proxy providers)

But an attacker who sees Alice can see what she's doing.

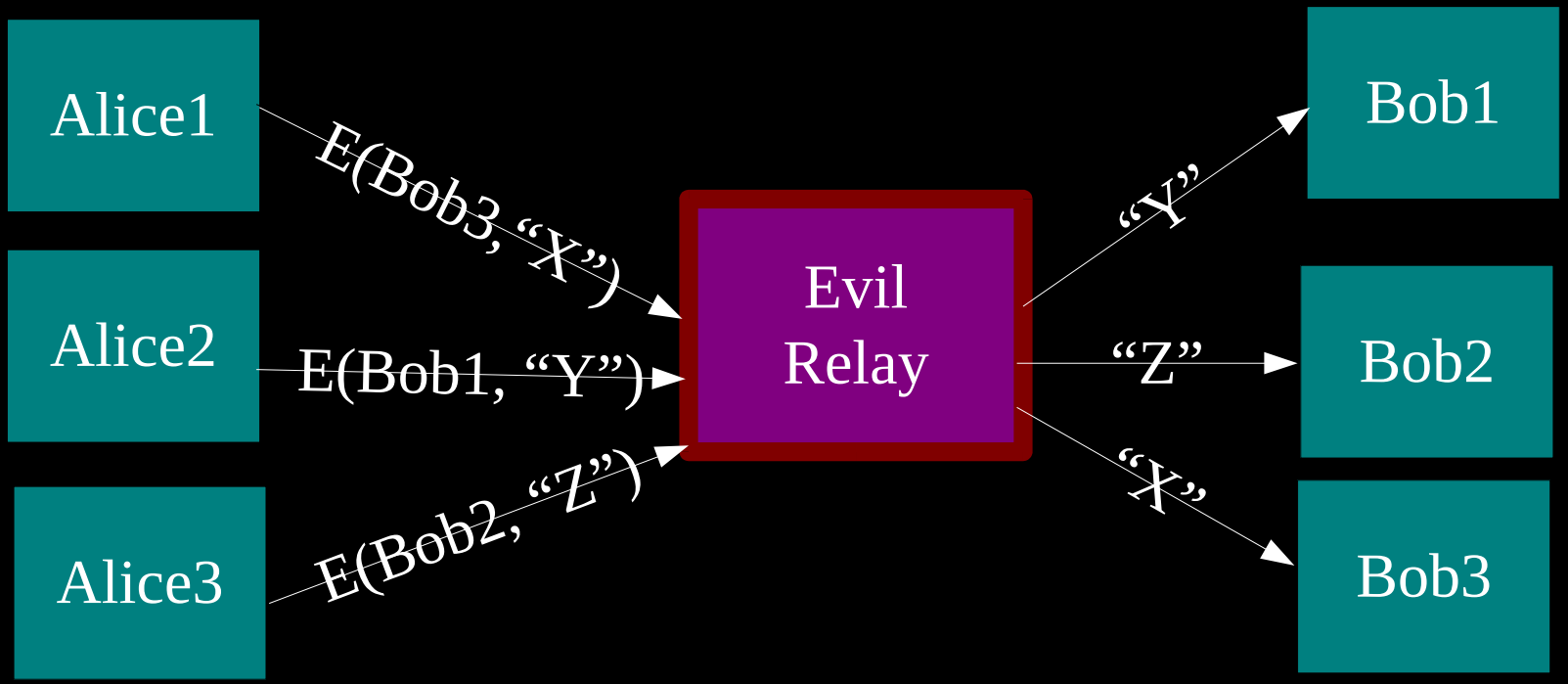


Add encryption to stop attackers who eavesdrop on Alice.



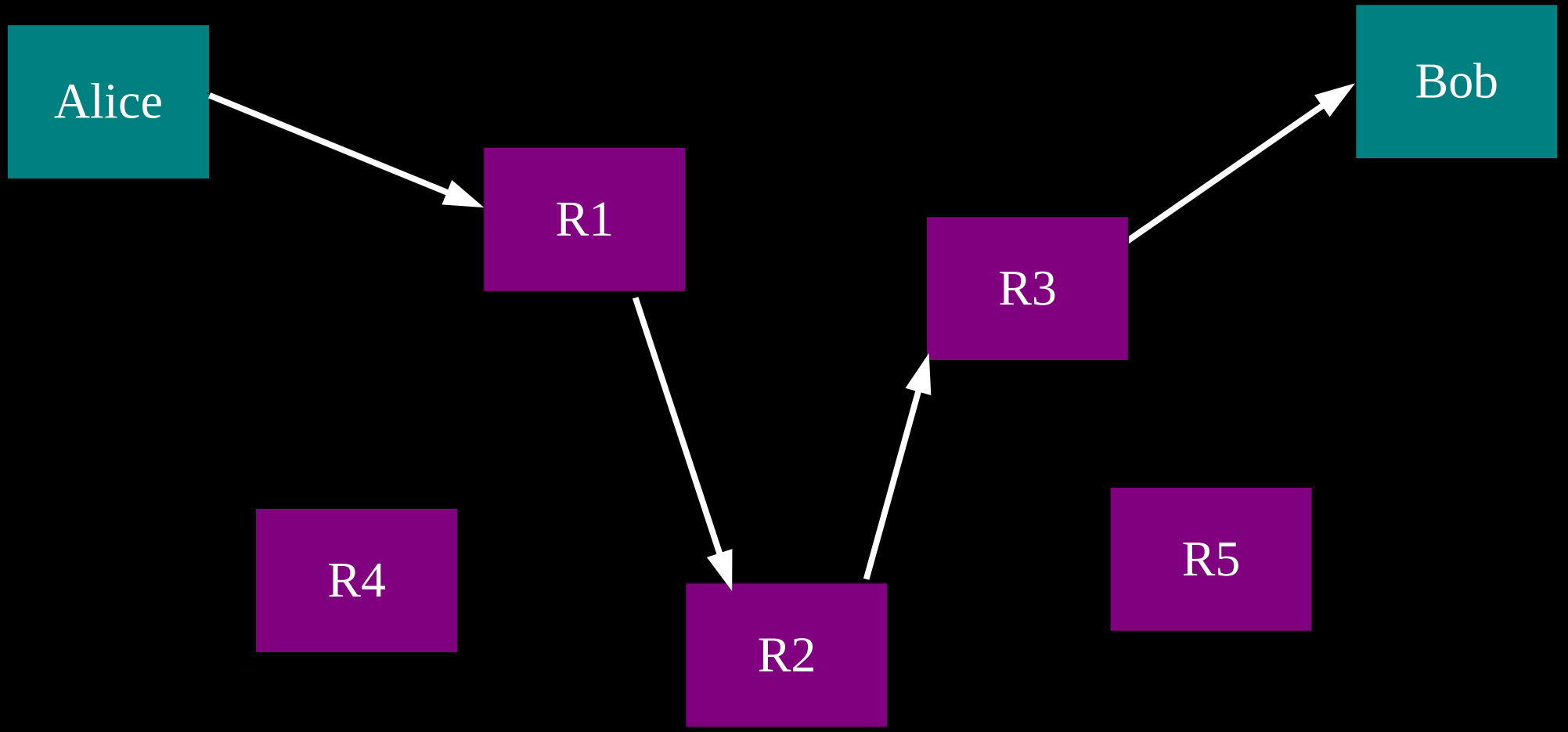
(ex: some commercial proxy providers)

But a single relay is a single point of failure.

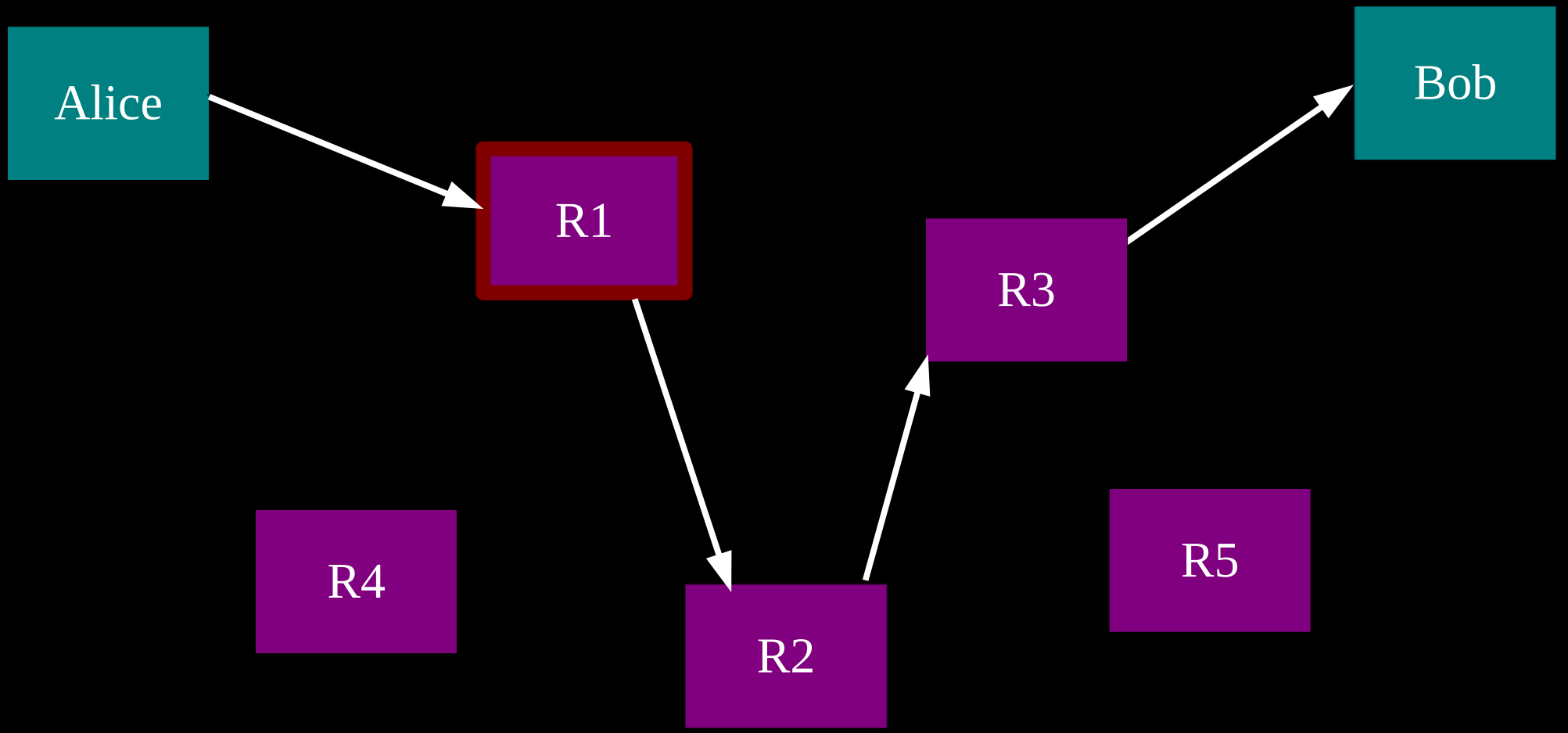


Eavesdropping the relay works too.

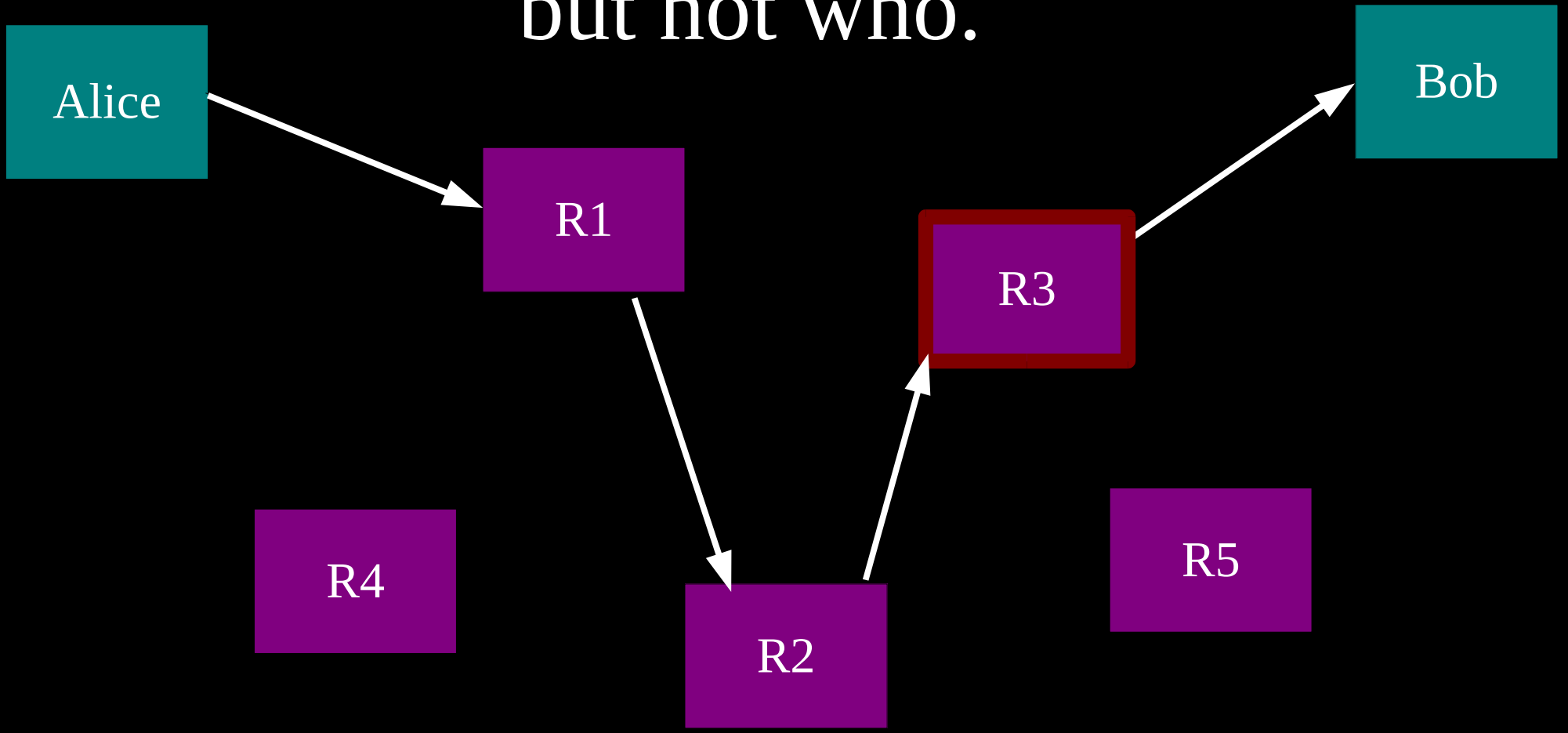
So, add multiple relays so that no single one can betray Alice.



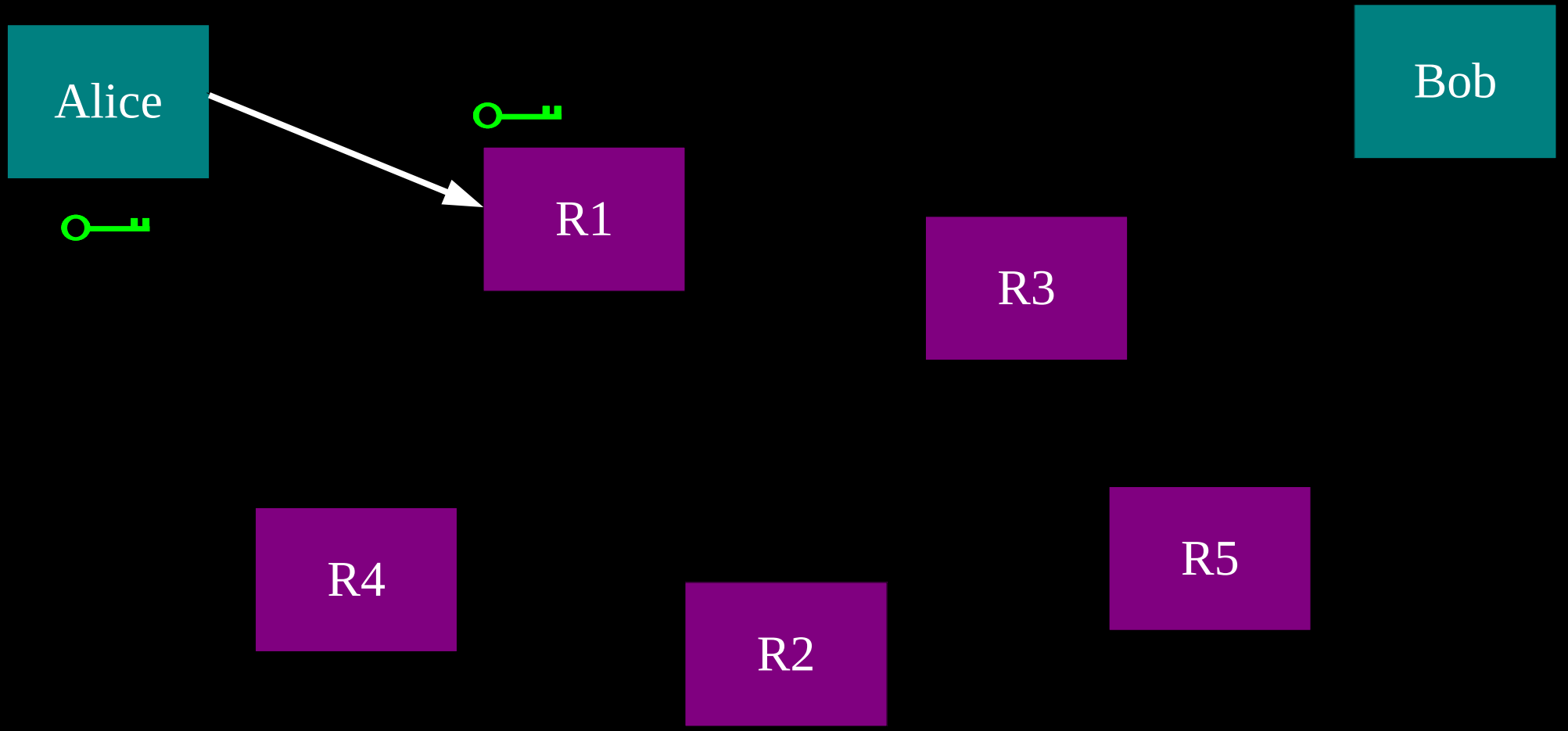
A corrupt first hop can tell that Alice is talking, but not to whom.



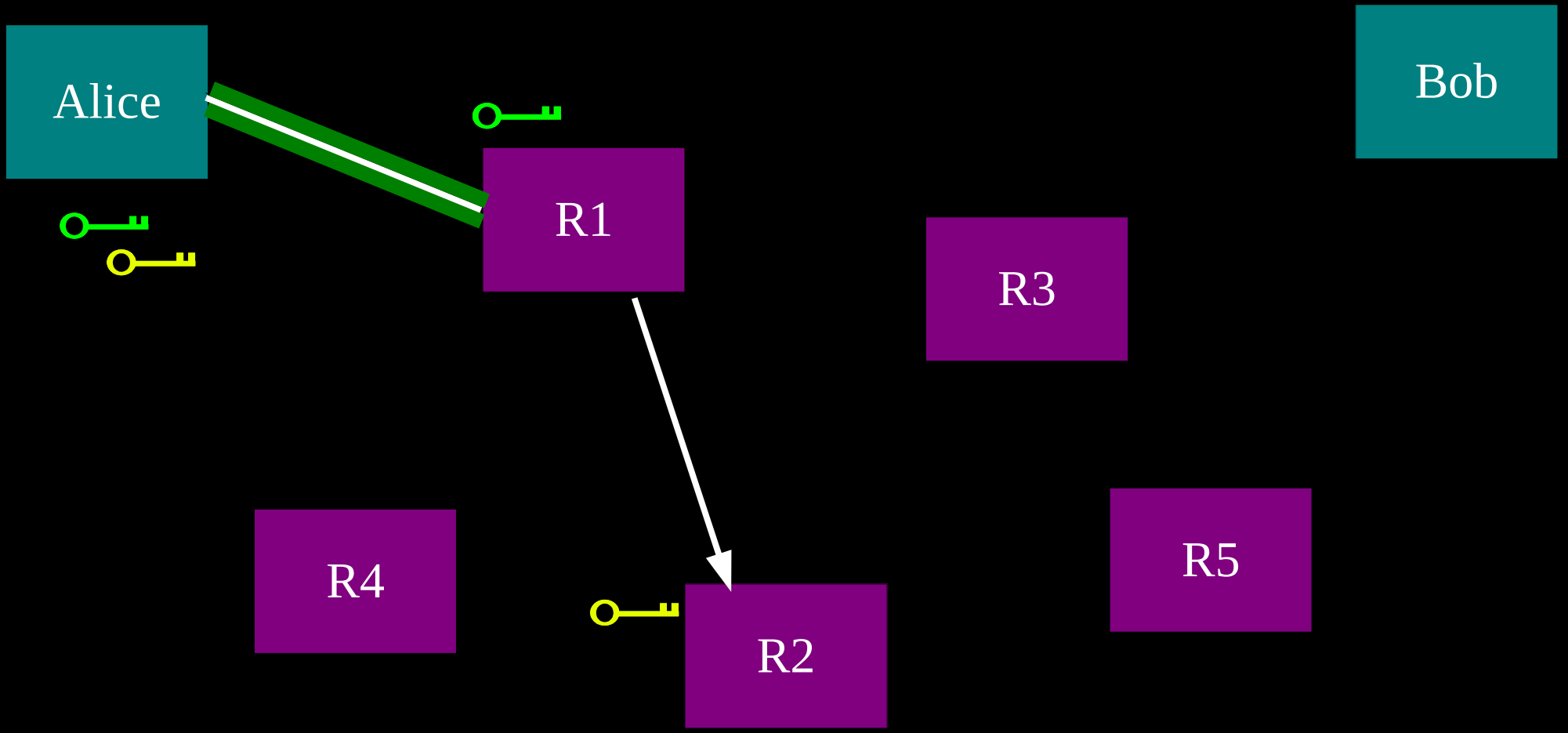
A corrupt final hop can tell that somebody is talking to Bob, but not who.



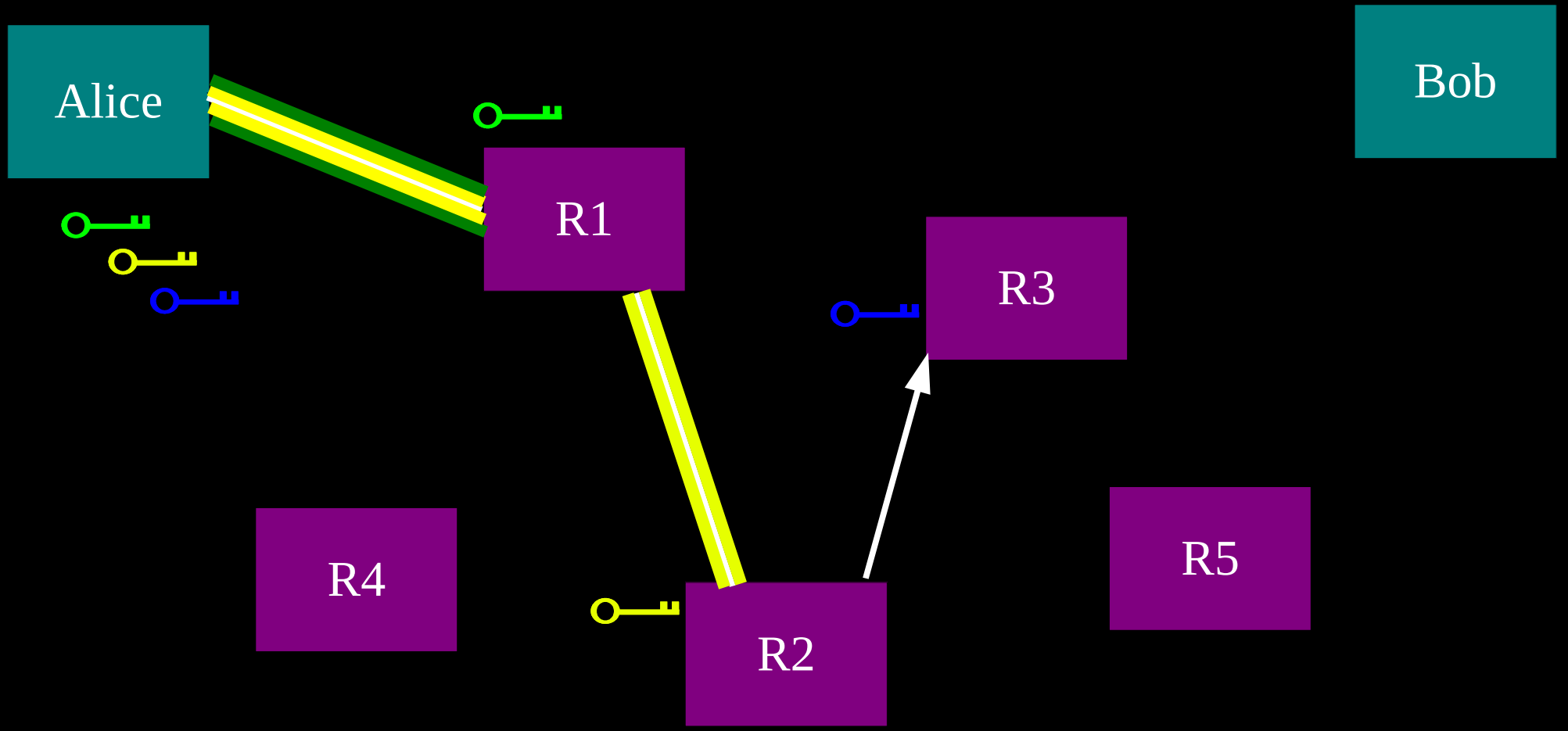
Alice makes a session key with R1



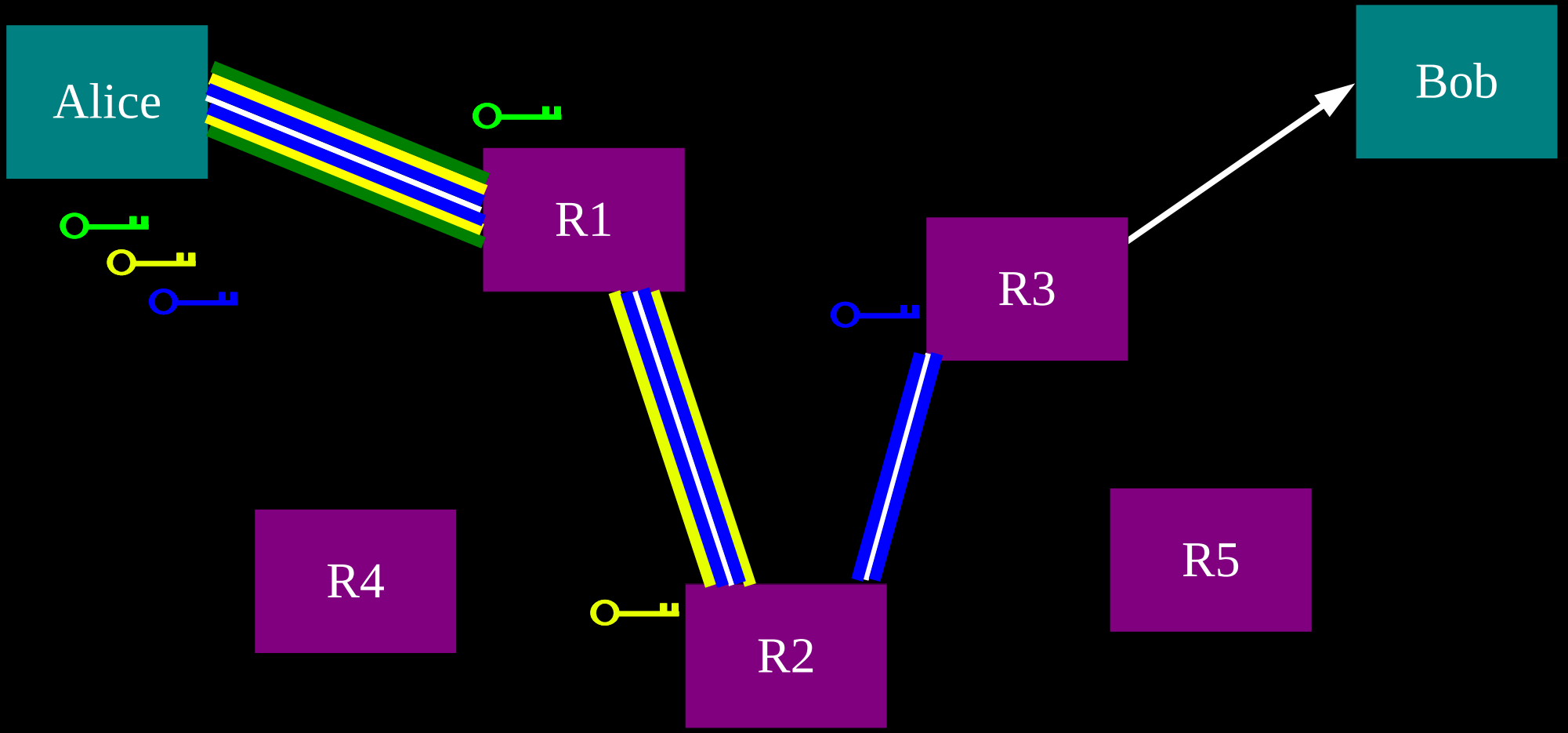
Alice makes a session key with R1 ...And then tunnels to R2



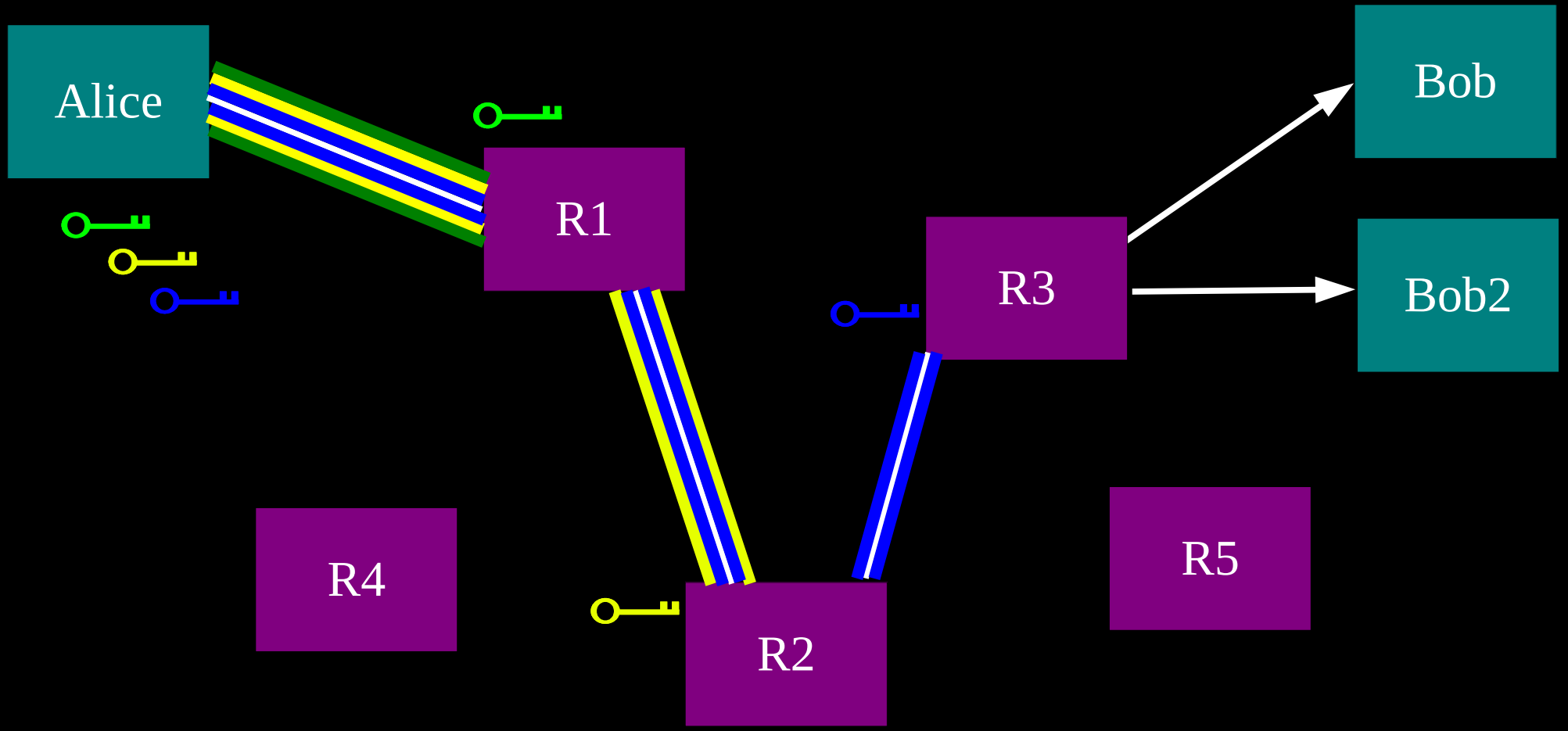
Alice makes a session key with R1 ...And then tunnels to R2...and to R3



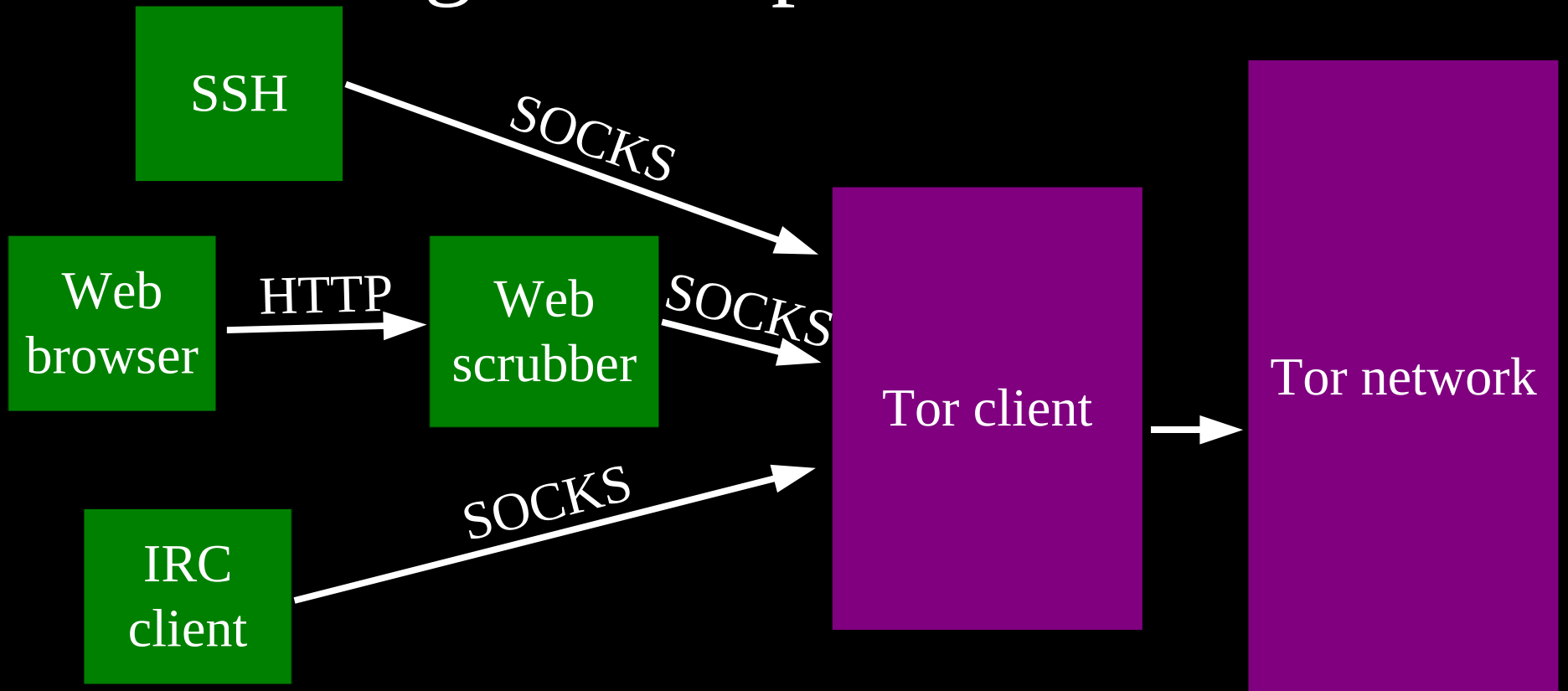
Alice makes a session key with R1 ...And then tunnels to R2...and to R3



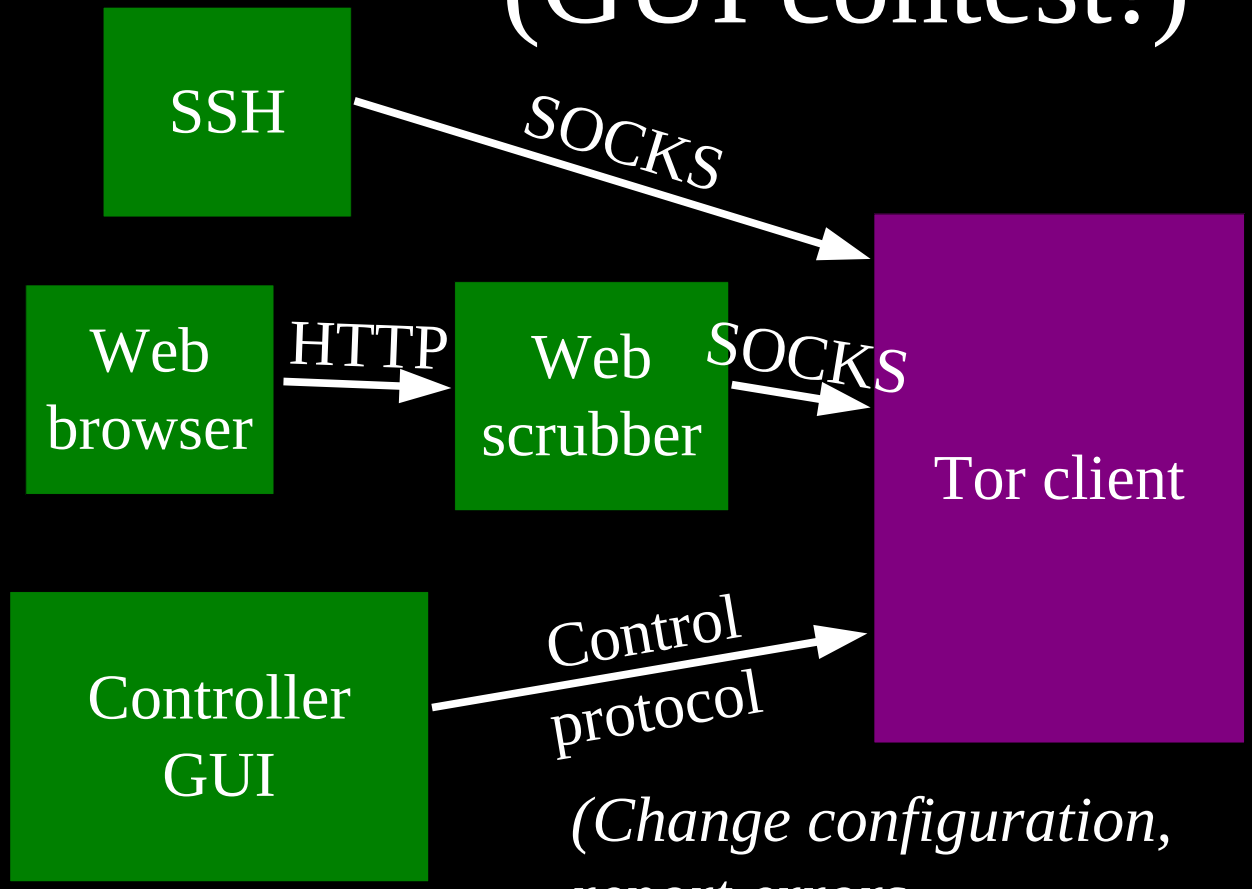
Can multiplex many connections through the encrypted circuit



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



We added a control protocol for external GUI applications. (GUI contest!)

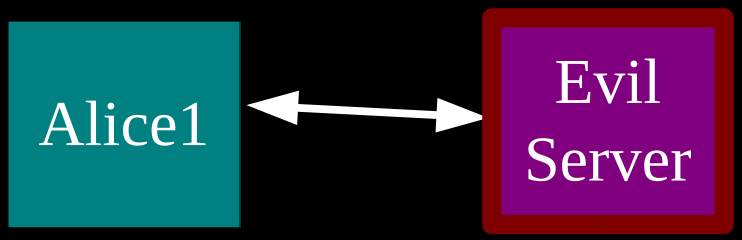


Usability for server operators is key.

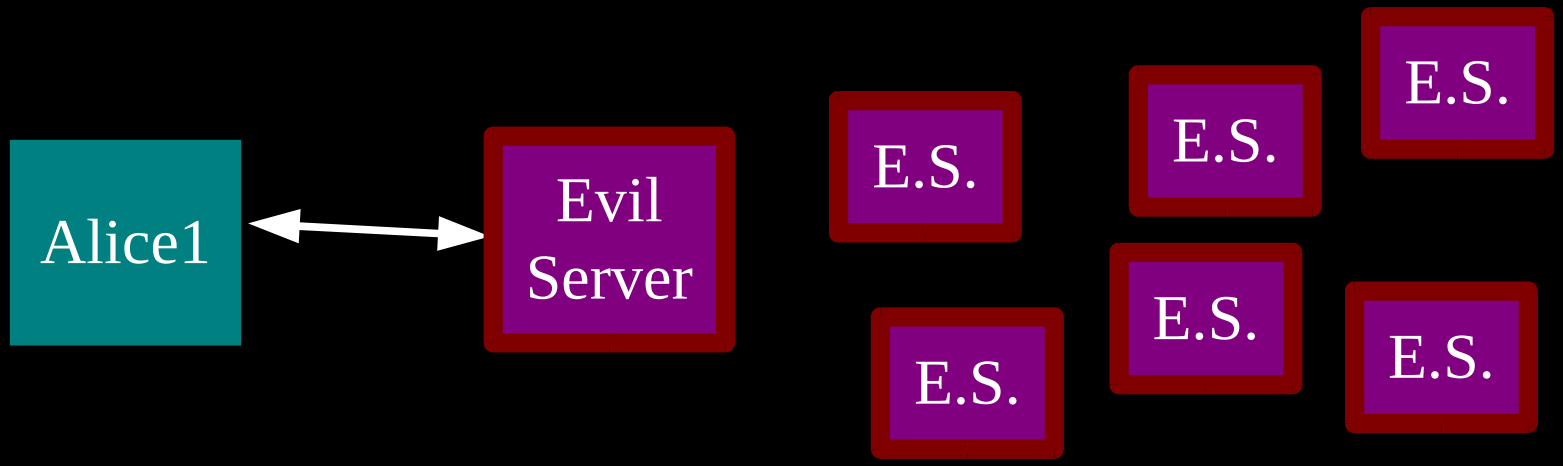
- Rate limiting: eating too much bandwidth is rude!
- Exit policies: not everyone is willing to emit arbitrary traffic.

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

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

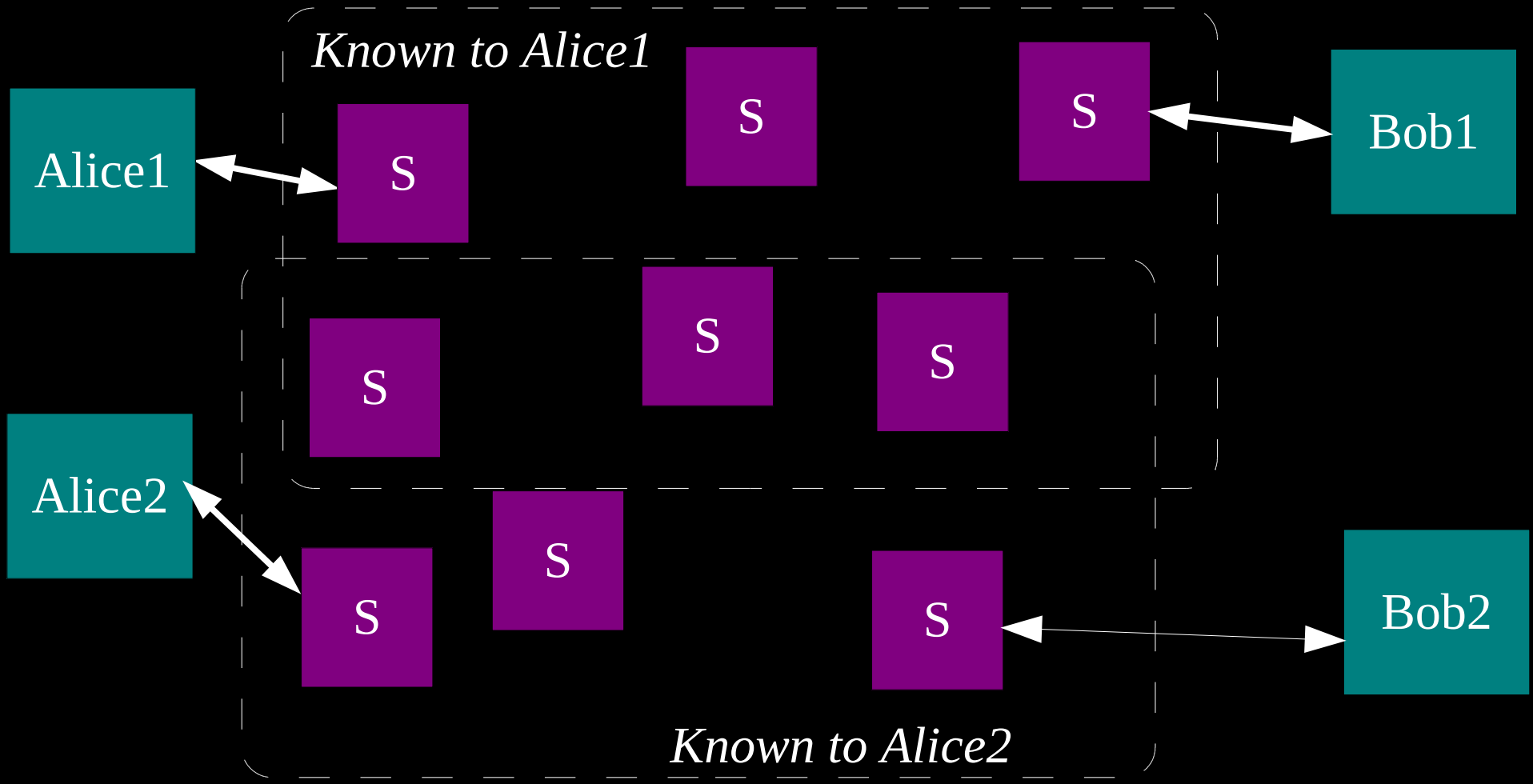


1. Alice says, “Describe the network!”

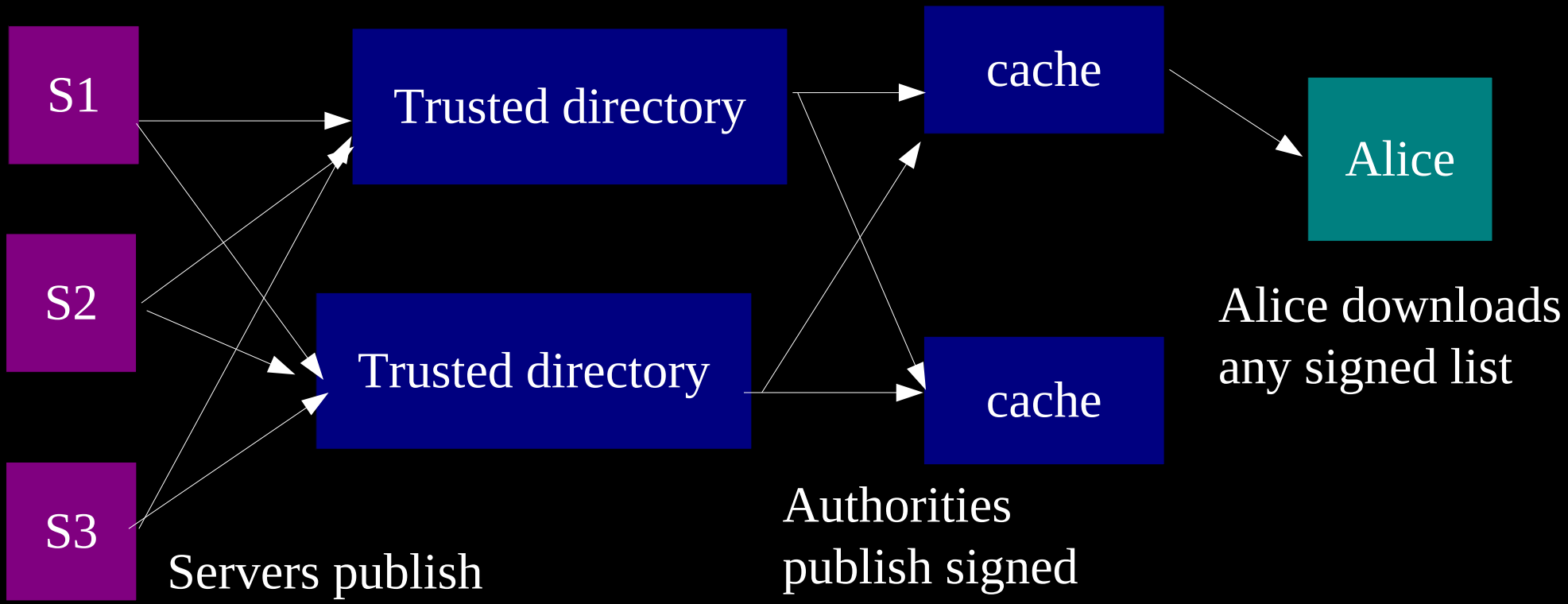


2. Alice is now in trouble.

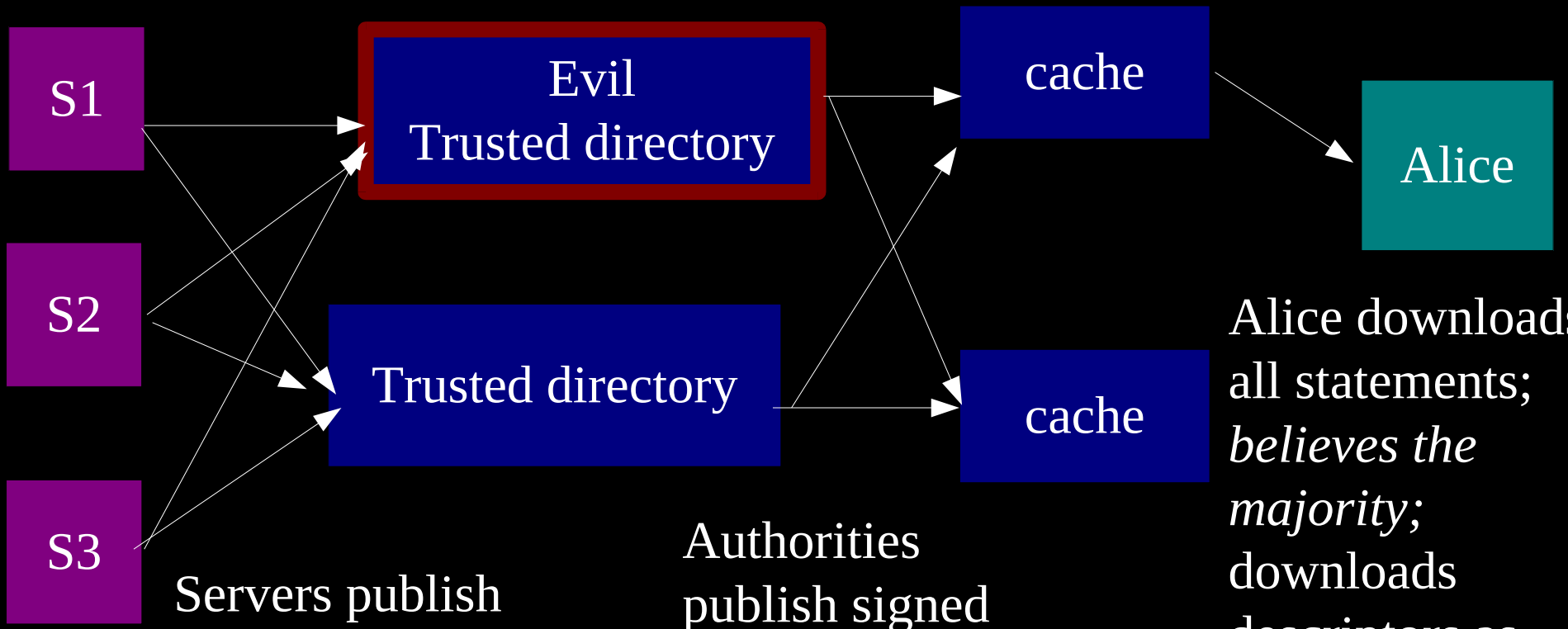
Server discovery is hard because misinformed clients lose anonymity.



Early Tor versions used a trivial centralized directory protocol.



We redesigned our directory protocol to reduce trust bottlenecks.



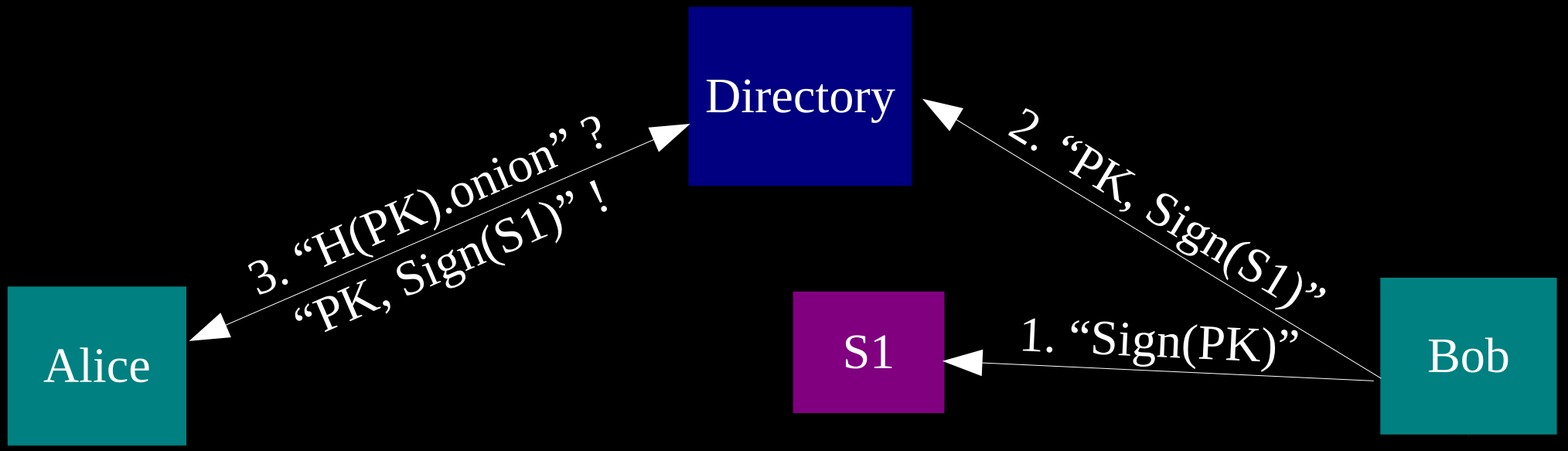
Servers publish self-signed descriptors.

Authorities publish signed statements *about* descriptors.

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

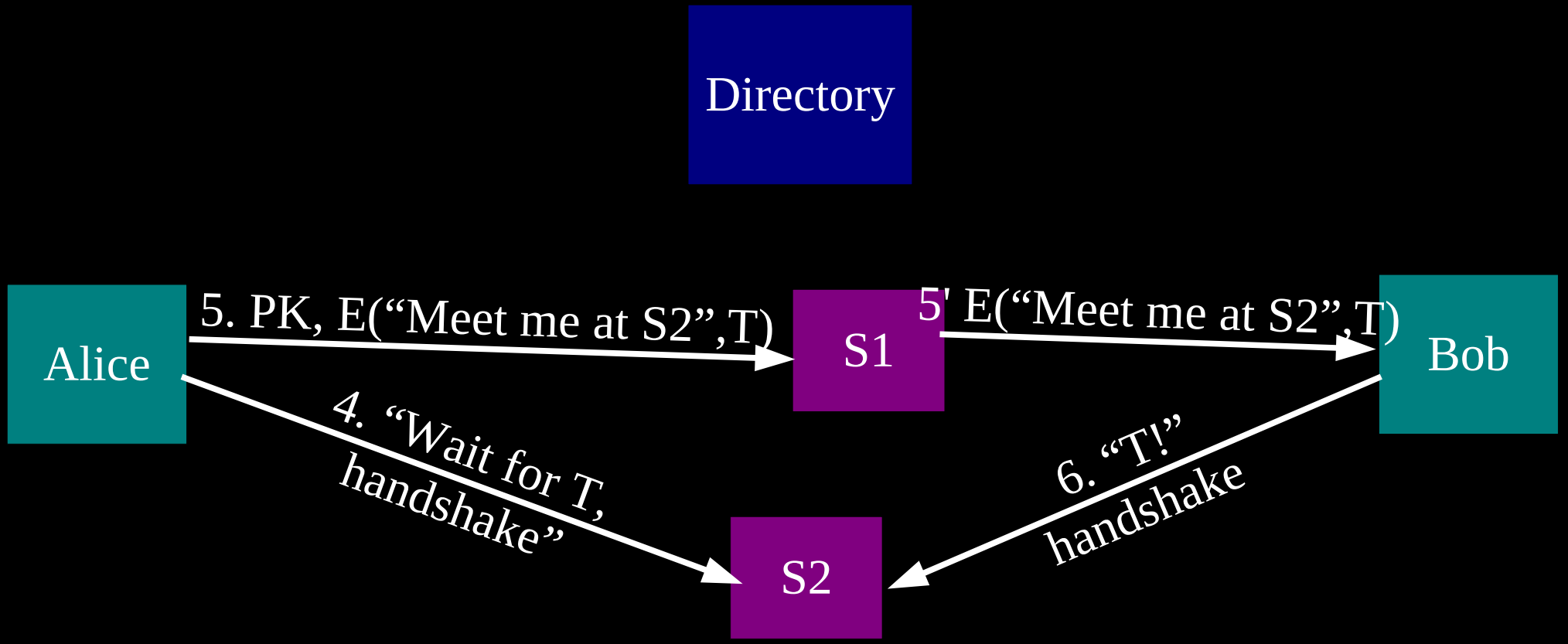
(Also uses less bandwidth!)

Tor implements responder anonymity with hidden services.



All these connections are anonymized.

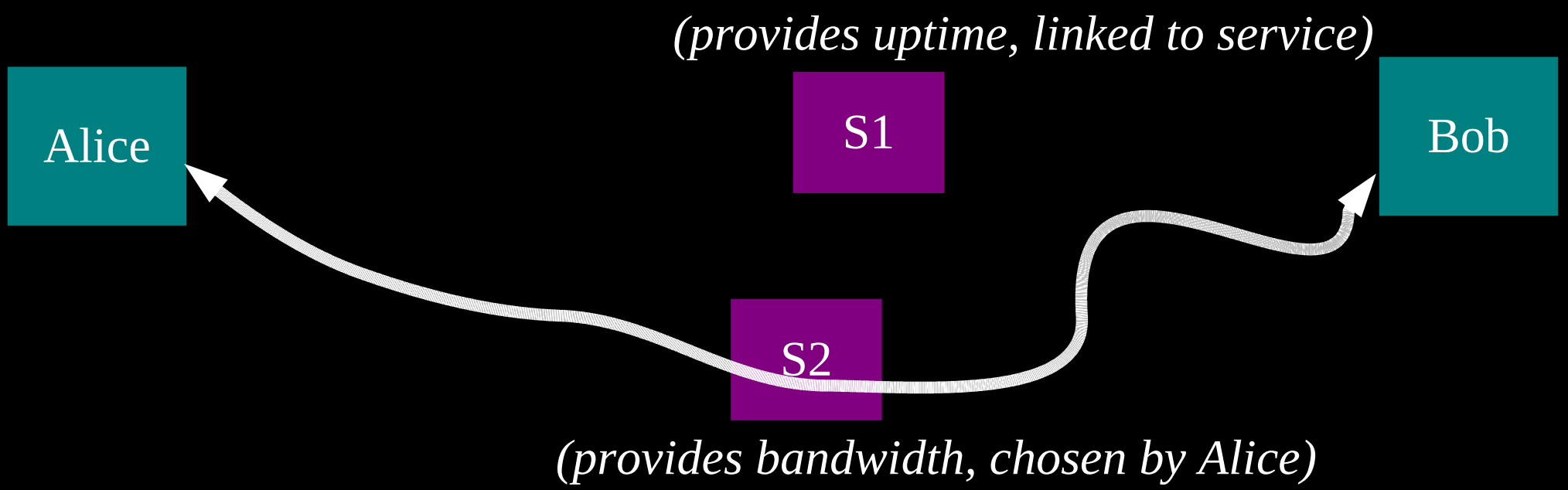
Tor implements responder anonymity with hidden services.



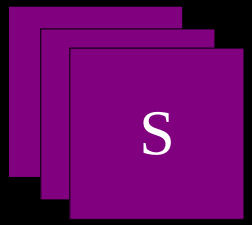
All these connections are anonymized.

Tor implements responder anonymity with hidden services.

Bidirectional anonymity!



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

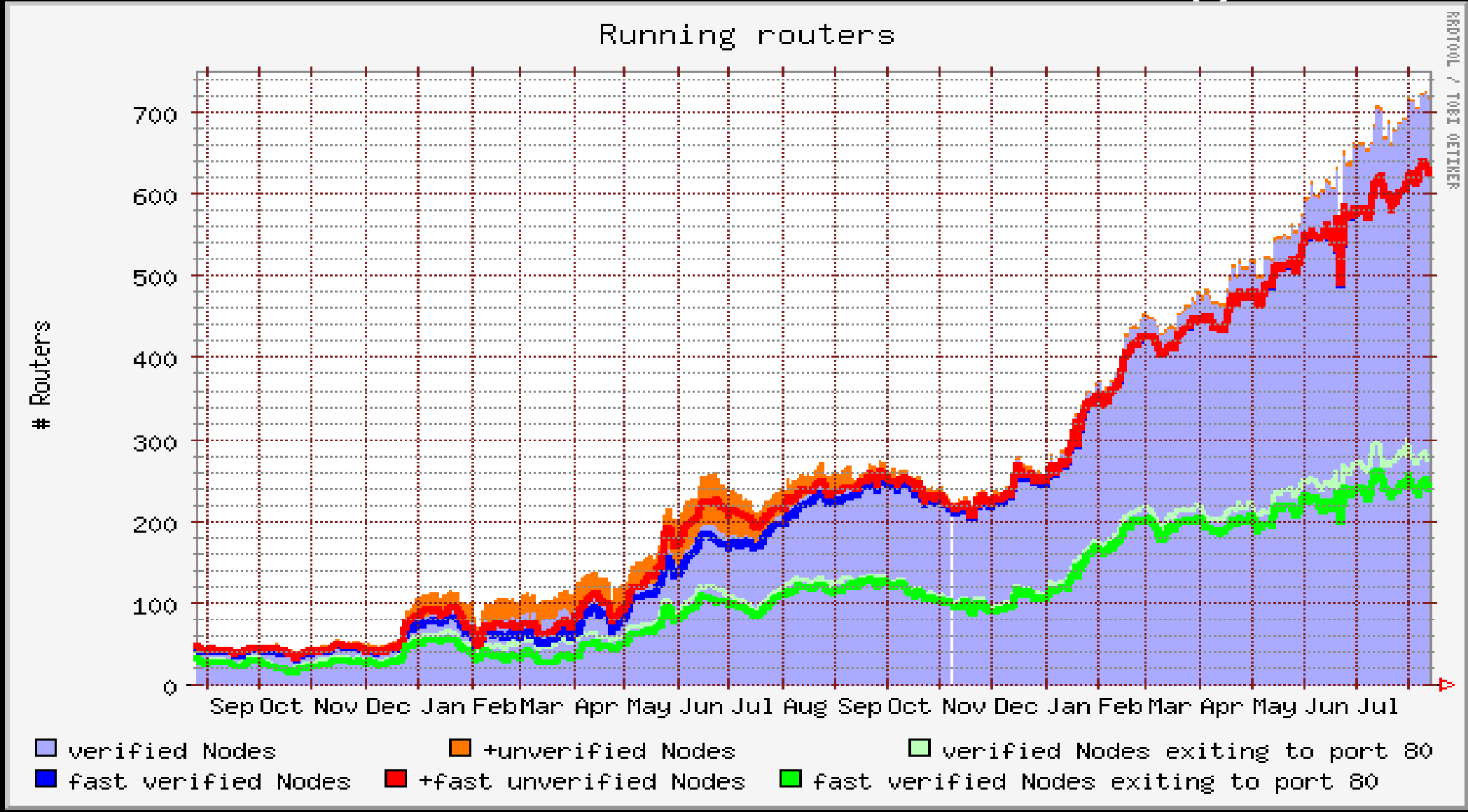


> 700 running

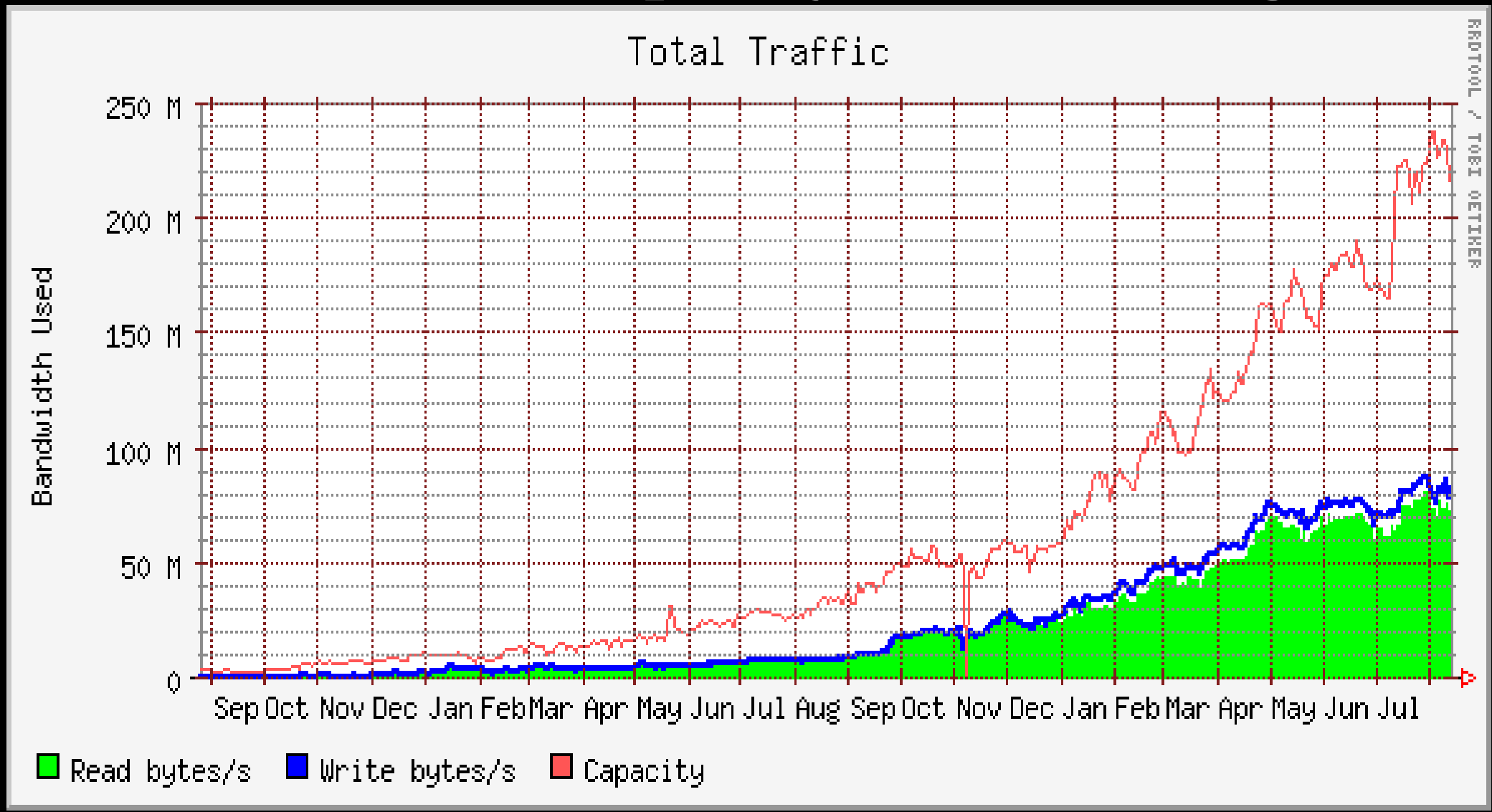
> 200,000 in a week

> 70 MB/sec

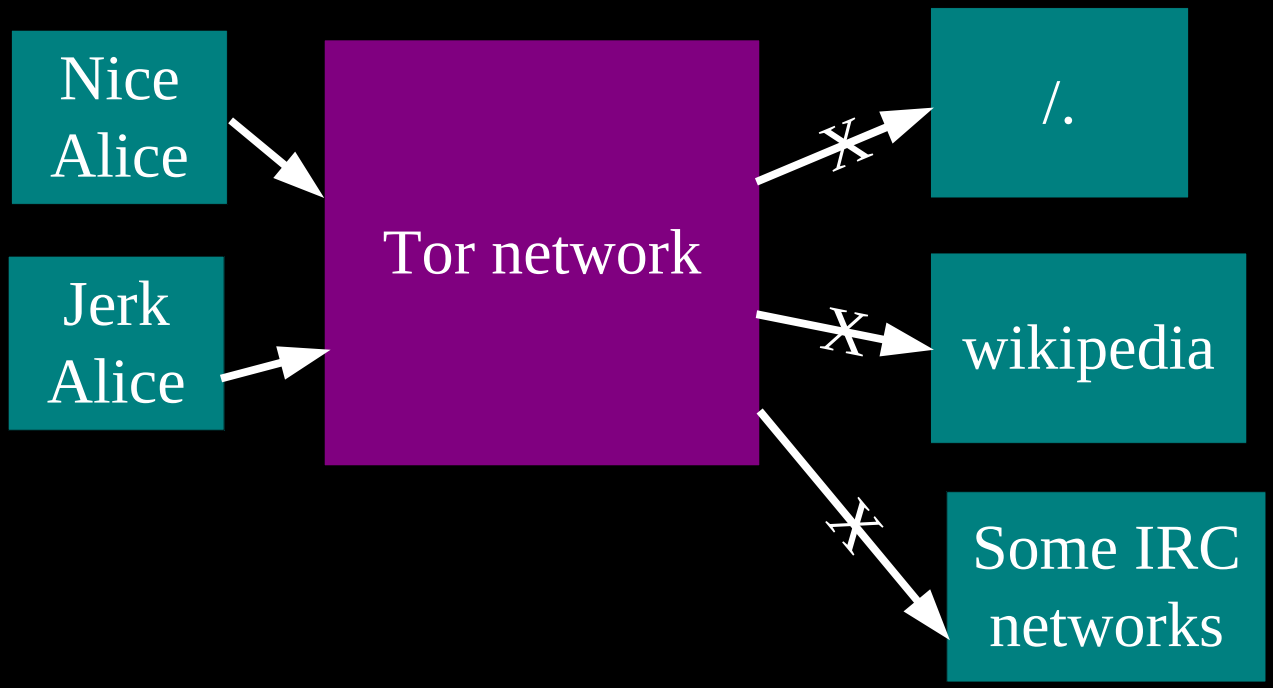
Growth in servers is increasing.



Bandwidth capacity is increasing.



Problem: Abusive users get the whole network blocked.

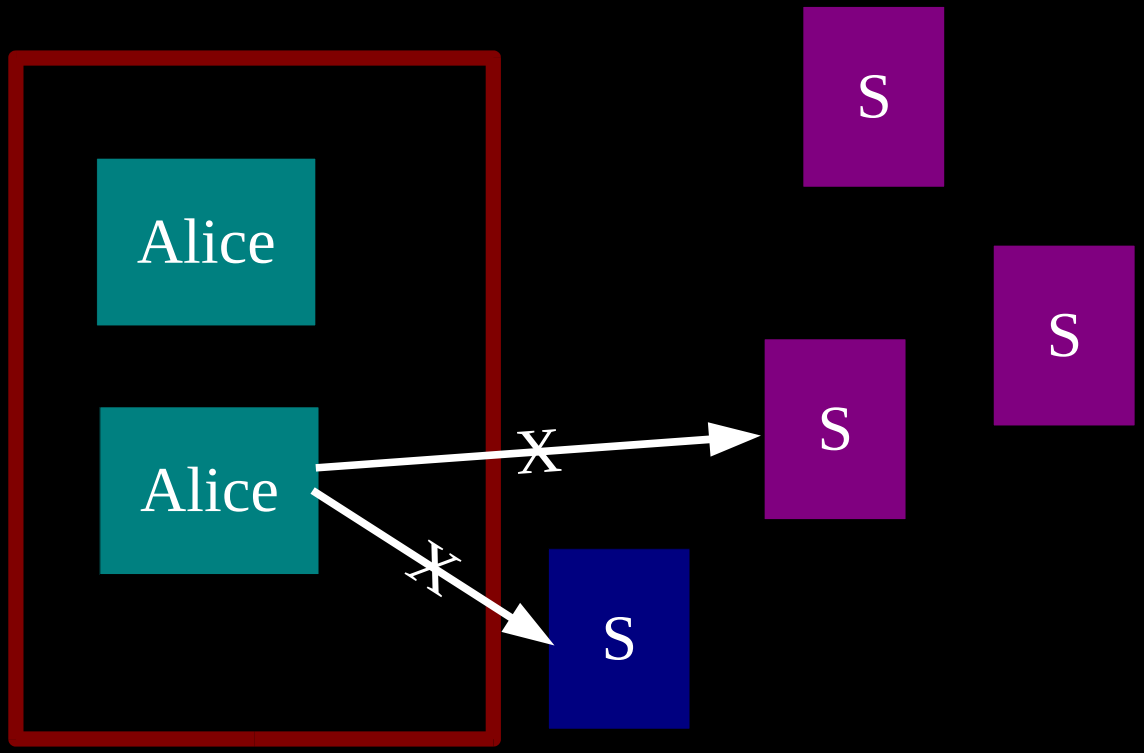


Minimize scope of blocking?

Other common abuses

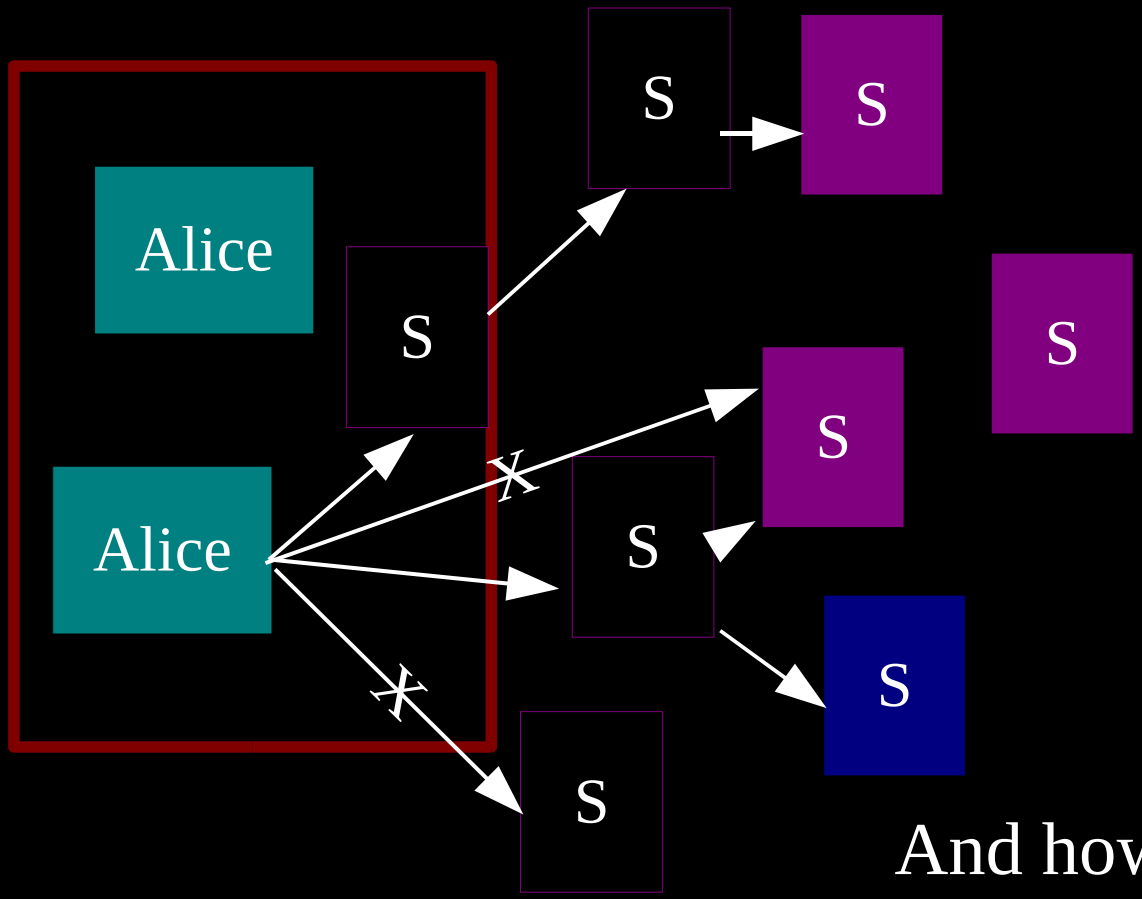
- Somebody connects to Hotmail, and sends an obnoxious mail.
- Somebody connects to IRC and yells -> DDoS on Tor exit server.
- Somebody tries to get you shut down by connecting to Google Groups and posting spam.
- Somebody uses Tor to download a movie, and your ISP gets a DMCA takedown.
- SORBS / Blacklists

Problem: China is hard to beat. They can just block the whole network.



They don't, yet. But when they do...?

Can we get a large number of semi-secret relays for China?



And how to distribute them?

Next steps

- Need to work on Windows stability and usability – including GUI and installers.
- Need to make it easier to be a server; incentives.
- Design for scalability and decentralization – tens of thousands of servers, millions of users.
- Hidden services need to be faster / more stable.
- Enclave-level onion routers (for enterprise/govt).
- Documentation and user support.

University Tor servers

- MIT, RPI, UCLA, Brown, ...
- Harvard, Georgia Tech, CMU

- UMass Amherst
- Dartmouth, Rice, UNC, ...
- Berkeley
- Rose-Hulman
- Michigan Tech

Questions?

- Tor: **http://tor.eff.org/**
  - Try it out; want to run a server?
- Anonymity bibliography: http://freehaven.net/anonbib/