Dark Web

Agenda

- Introduction
- Review
- VPN proxies
- Tor
- Tor services
- Tor browser
- Tor history

The dark web

- Pejoratively, the “disreputable” web
- Part of the Web accessible only through an anonymous connection
  - TL;DR: Tor Browser


SURFACE WEB
Google
Bing
Wikipedia
DEEP WEB
Academic Information
Medical Records
Legal Documents
Scientific Reports
Subscription Information
Contains 90% of the information on the Internet, but is not accessible by Surface Web crawlers.
Social Media
Multilingual Databases
Financial Records
Government Resources
Competitor Websites
Organization-specific Repositories
(DARK WEB)
A part of the Deep Web accessible only through certain browsers such as Tor designed to ensure anonymity. Deep Web Technologies has zero involvement with the Dark Web.
Illegal Information
TOR-Encrypted sites
Political Protests
Drug Trafficking sites
Private Communications

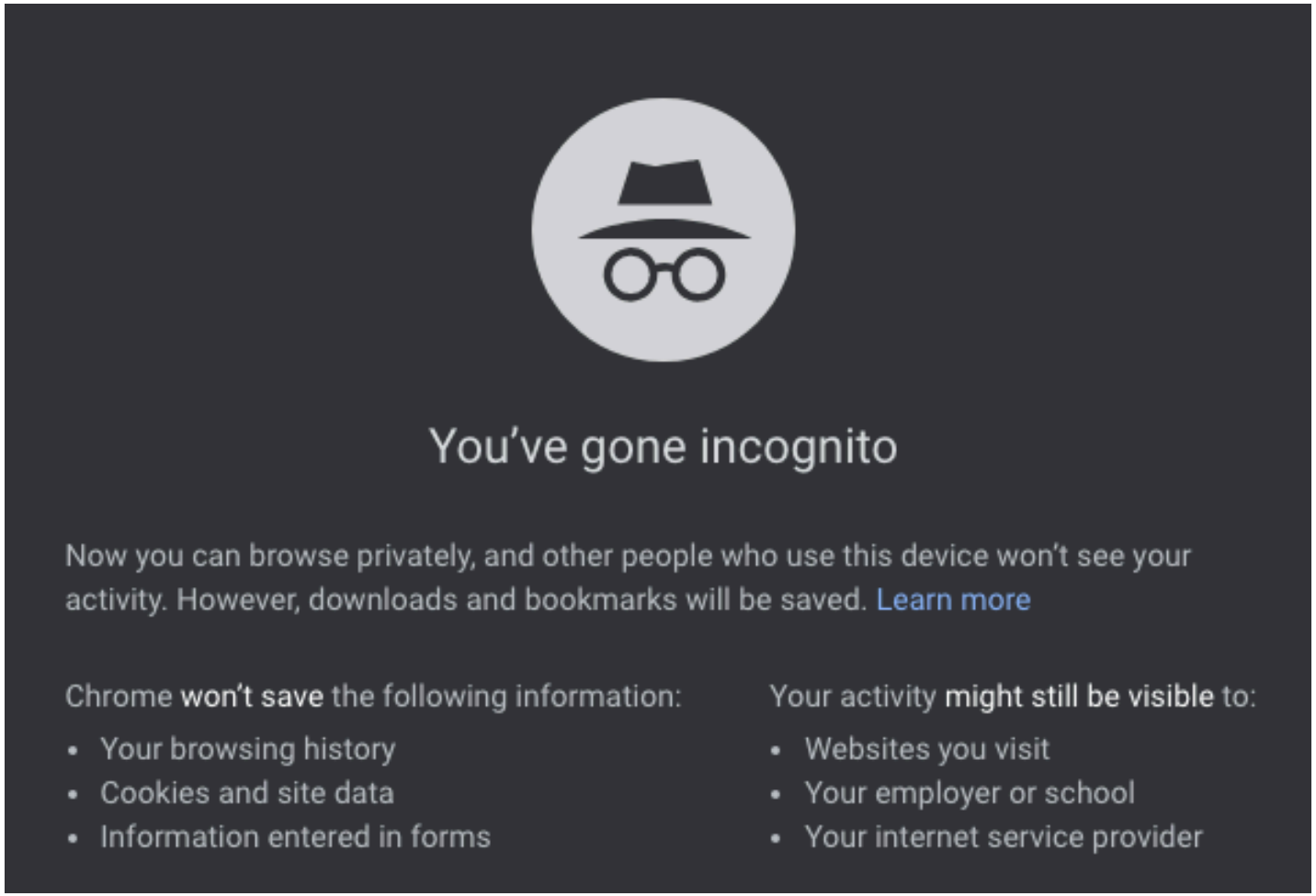
Dark web vs. deep web

- *Deep web*: pages that search engines can't find
  - Pages behind paywalls, e.g., Wall Street Journal
  - Pages behind passwords, e.g., Gmail, Facebook, online banking
- *Dark web*: pages accessible only through an anonymous connection
- The deep web includes the dark web

Anonymity online

- Is anonymous browsing mode enough for anonymity online?
  - Private browsing mode, incognito mode, etc.
- What about HTTPS?
- Is that enough? Why?

Anonymity online



Agenda

- Introduction
- **Review**
- VPN proxies
- Tor
- Tor services
- Tor browser
- Tor history

Internet Protocol (IP)

- How does information travel to the right place?
- Every computer on the Internet has an address
  - Google 172.217.5.14
  - Amazon 205.251.242.103
  - My laptop 141.212.109.1
- Newer version: longer addresses
  - IPv4: 32 bits is 4 billion computers
  - IPv6: 64 bits is way more

IPv4 routing

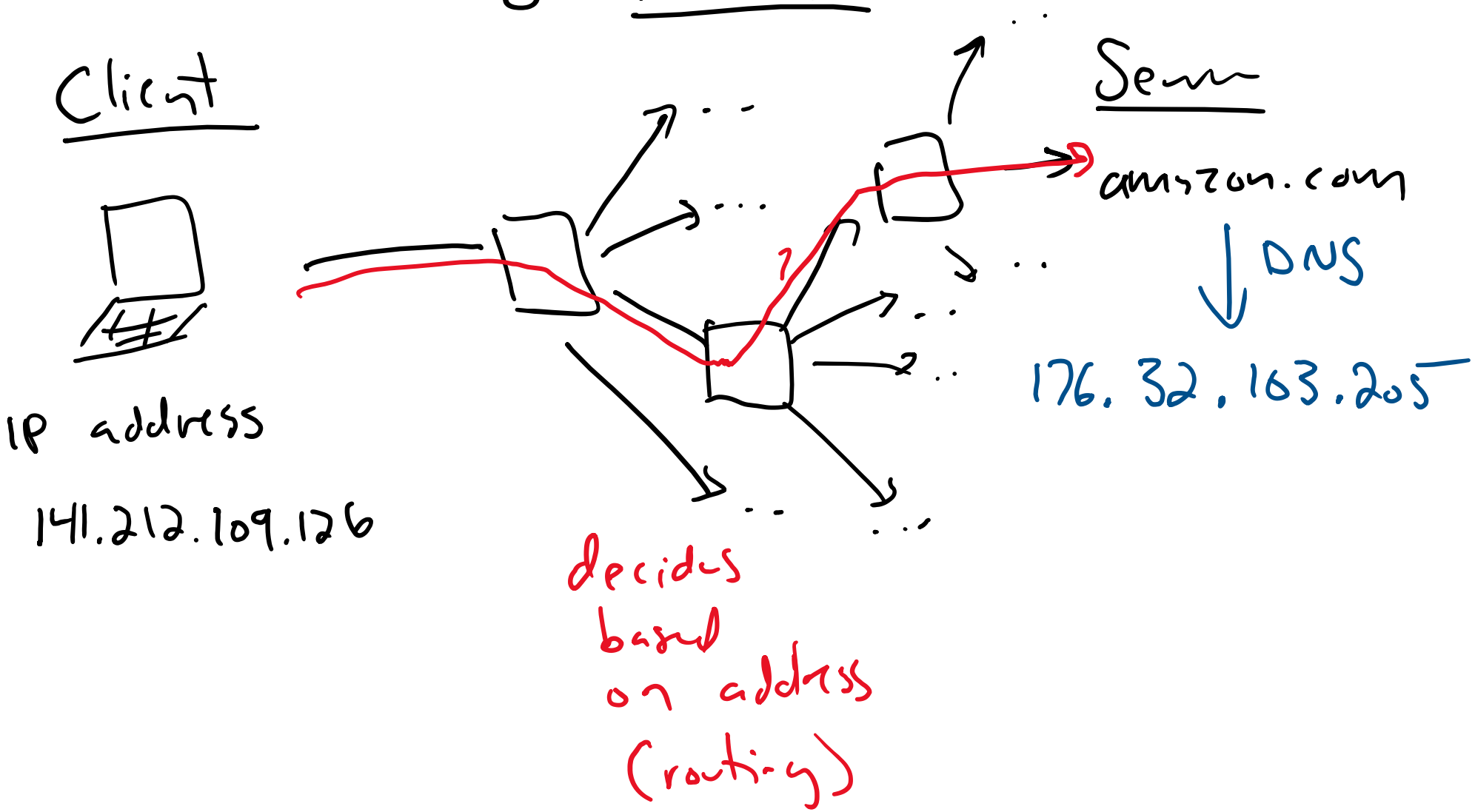
- `traceroute` shows route to destination

```
$ traceroute google.com
traceroute to google.com (74.125.95.99)
141.212.111.1 (141.212.111.1)  1.037 ms (Ann Arbor, MI)
13-caen-bin-arb.r-bin-arbl.umnet.umich.edu
(192.12.80.177)  0.566 (Ann Arbor, MI)
...
7  216.239.48.154 (216.239.48.154)  6.785 ms (Mountain
View, CA)
8  209.85.241.22 (209.85.241.22)  43.101 ms (Mountain
View, CA)
```

traceroute google.com → 9 hops

IPv4 routing

routers



Eavesdropping

- Your packet traverses many routers controlled by third parties
- Attack vector: capture a router and listen in

Whistle-Blower Outs NSA Spy Room

Ryan Singel 04.07.06

AT&T's central office on Folsom Street in San Francisco houses a secret room that allows the National Security Agency to monitor phone and internet traffic, according to former AT&T technician-cum-whistle-blower Mark Klein. View Slideshow

AT&T provided National Security Agency eavesdroppers with full access to its customers' phone calls, and shunted its customers' internet traffic to data-mining equipment installed in a secret room in its San Francisco switching center, according to a former AT&T worker cooperating in the Electronic Frontier Foundation's lawsuit against the company.

Mark Klein, a retired AT&T communications technician, submitted an affidavit in support of the EFF's lawsuit this week. That class action lawsuit, filed in federal court in San Francisco last January, alleges that AT&T violated federal and state laws by surreptitiously allowing the government to monitor phone and internet communications of AT&T customers without warrants.

On Wednesday, the EFF asked the court to issue an injunction prohibiting AT&T from continuing the alleged wiretapping, and filed a number of documents under seal, including three AT&T documents that purportedly explain how the wiretapping system works.

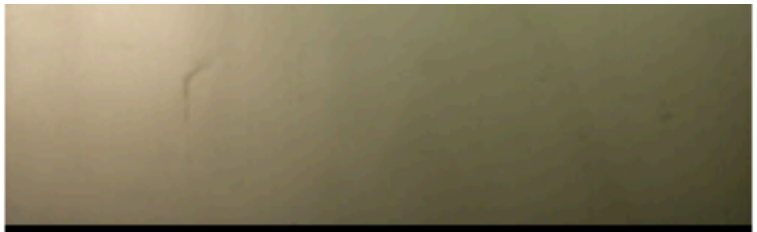
According to a statement released by Klein's attorney, an NSA agent showed up at the San Francisco switching center in 2002 to interview a management-level

Whistle-Blower Outs NSA Spy Room

Ryan Singel 04.07.06

AT&T provided National Security Agency eavesdroppers with full access to its customers' phone calls, and shunted its customers' internet traffic to data-mining equipment installed in a secret room in its San Francisco switching center, according to a former AT&T worker cooperating in

The evidence also shows that the government did not act alone. EFF has obtained whistleblower evidence [PDF] from former AT&T technician Mark Klein showing that AT&T is cooperating with the illegal surveillance. The undisputed documents show that AT&T installed a fiberoptic splitter at its facility at 611 Folsom Street in San Francisco that makes copies of all emails, web browsing, and other Internet traffic to and from AT&T copies to the NSA. This copying includes both domestic and international Internet activitie observed, "this isn't a wiretap, it's a country-tap."



AT&T's central office on Folsom Street in San Francisco houses a secret room that allows the National Security Agency to monitor phone and internet traffic, according to former AT&T technician-cum-whistle-blower Mark Klein. View Slideshow

On Wednesday, the EFF asked the court to issue an injunction prohibiting AT&T from continuing the alleged wiretapping, and filed a number of documents under seal, including three AT&T documents that purportedly explain how the wiretapping system works.

According to a statement released by Klein's attorney, an NSA agent showed up at the San Francisco switching center in 2002 to interview a management-level

Network security

- Two parties communicate over a network
- Assume powerful adversary
  - Can read (eavesdrop on) all data transmitted
  - Can modify or delete any data
  - Can inject new data
- Communication: What properties would you like?

Desirable properties

- Confidentiality
  - Adversary should not understand message
- Sender authenticity
  - Message is really from the purported sender
- Message integrity
  - Message not modified between send and receive
- Freshness
  - Message was sent “recently”
- Anonymity
  - Attacker should not know that we are communicating

Desirable properties

- Confidentiality
  - Solved with symmetric encryption
- Sender authenticity
  - Solved with asymmetric encryption and public key infrastructure (PKI)
- Message integrity
  - Solved with symmetric encryption and MAC
- Freshness
  - Solved with symmetric encryption and timestamp, sequence number or nonce
- Anonymity
  - Solved with VPN (weak) or Tor (strong)

Agenda

- Introduction
- Review
- **VPN proxies**
- Tor
- Tor services
- Tor browser
- Tor history

Packet inspection

- Each packet has source IP address & destination IP address
- If you change destination, routers won't deliver it to the right place
- Can sometimes change source (spoofing) but not always
- Lots of information by observing ingress and egress links

Packet inspection

- Packet source and destination are visible to anyone observing the network

Image: https://www.howtogeek.com

Metadata

- Packet source and destination are visible to anyone observing the network
  - Your ISP
  - Destination web site
  - Any intermediate router
  - Network eavesdropper
- Simply knowing two parties are communicating reveals info
  - Metadata

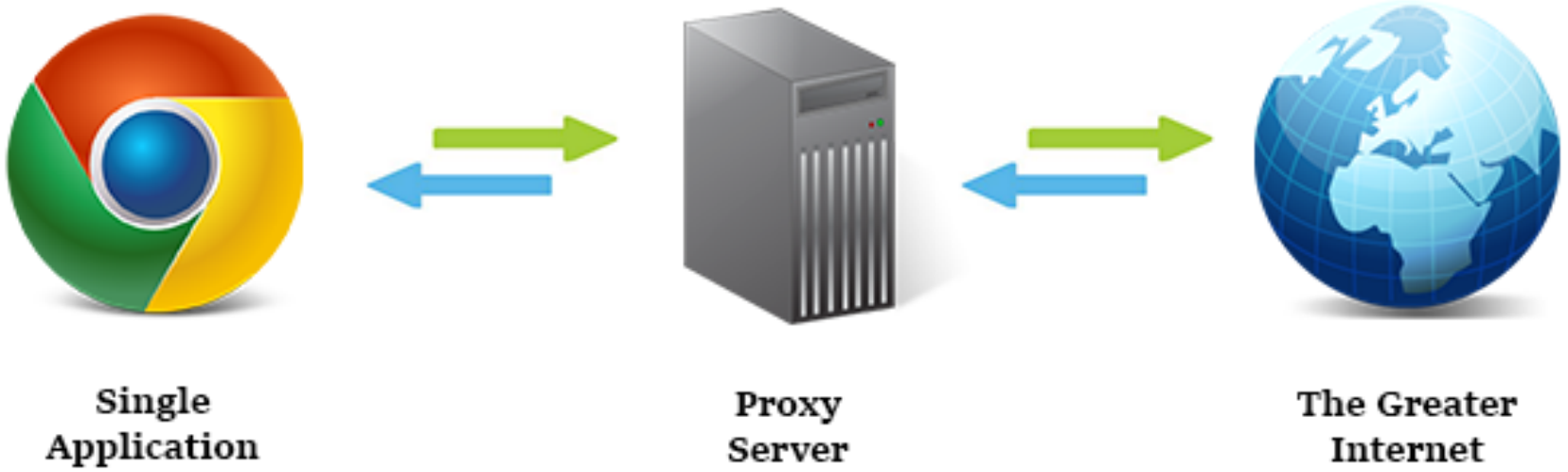
Metadata

- They know you rang a phone sex line at 2:24 am and spoke for 18 minutes. But they don't know what you talked about.
- They know you called the suicide prevention hotline from the Golden Gate Bridge. But the topic of the call remains a secret.
- They know you got an email from an HIV testing service, then called your doctor, then visited an HIV support group website in the same hour. But they don't know what was in the email or what you talked about on the phone.
- They know you received an email from a digital rights activist group with the subject line “Let’s Tell Congress: Stop SESTA/FOSTA” and then called your elected representative immediately after. But the content of those communications remains safe from government intrusion.
- They know you called a gynecologist, spoke for a half hour, and then called the local abortion clinic’s number later that day.

https://ssd.eff.org/en/module/why-metadata-matters

Proxy server

- Proxy server is a middleman
- Makes your internet traffic appear to come from somewhere else
- Hides your IP address



VPN proxy server

- VPN proxy server hides your IP address and encrypts your traffic

VPN proxy server

- Eavesdropper can see client communicate with VPN proxy
- Eavesdropper can see VPN communicate with Web site

VPN proxy server vulnerabilities

- Single point of failure: VPN knows all
- VPN proxy servers vulnerable to subpoenas
  - Court asks proxy server admin for server logs
- Some VPNs have a "destroy logs" policy
  - You'd have to trust your VPN

VPN proxy server vulnerabilities

- VPN proxy servers vulnerable to misconfiguration
- It's easy to screw up VPN server configuration
  - Out-of-date software with security vulnerabilities leads to exfiltrating data
  - Doesn't support IPv6, and IPv6 requests are leaked
  - ... and many others

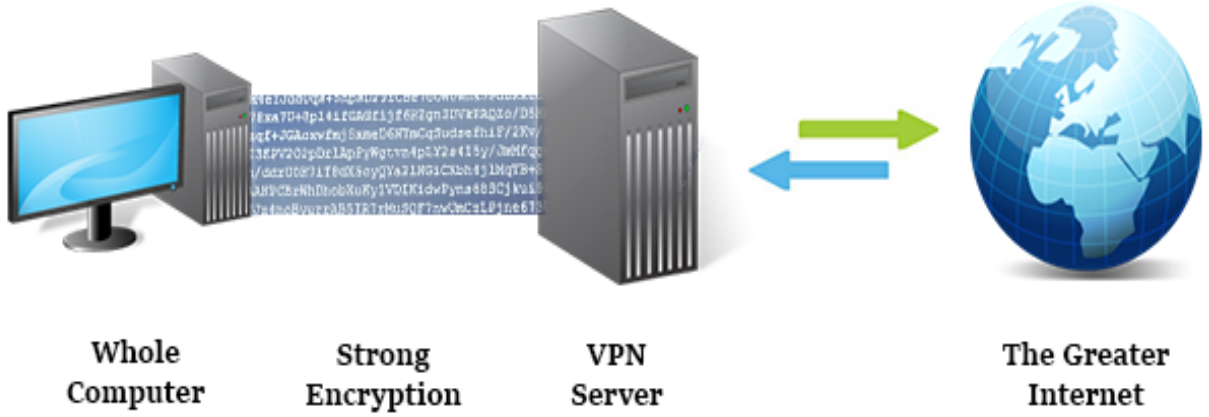
https://authentic8.blog/vpn-privacy-what-nobody-told-you/

VPN proxy server vulnerabilities

- VPN proxy servers are vulnerable to *traffic analysis*
  - Even if you can trust your VPN and it's configured correctly!
- What if every time client transmits, web site soon gets data?
- Your ISP or other eavesdroppers can use traffic analysis even with a VPN and SSL

Traffic analysis

- Every time you transmit, VPN transmits
  - Gather a bunch of data
- Only need to monitor two points on the network
  - Correlate these two

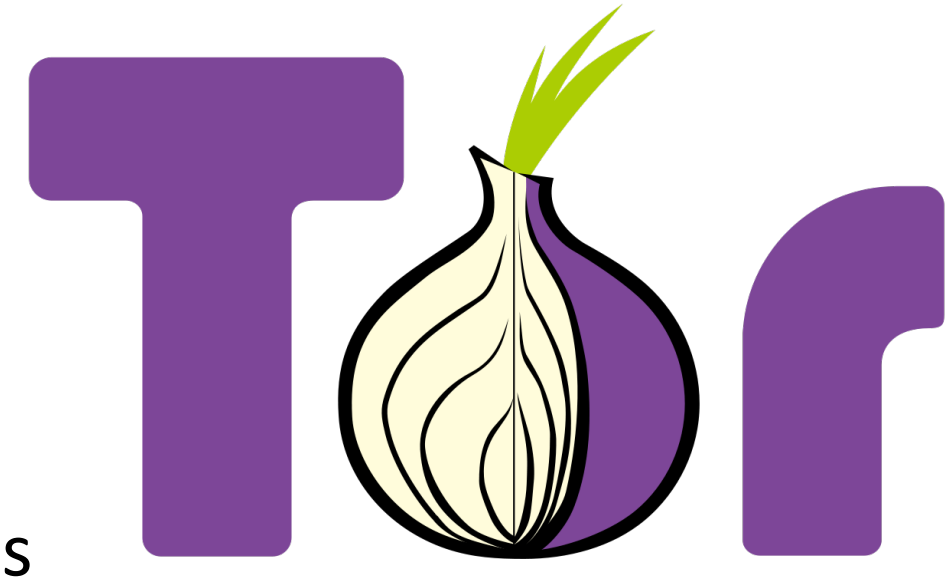


Tor vs. VPN details: https://2019.www.torproject.org/docs/faq.html.en#IsTorLikeAVPN

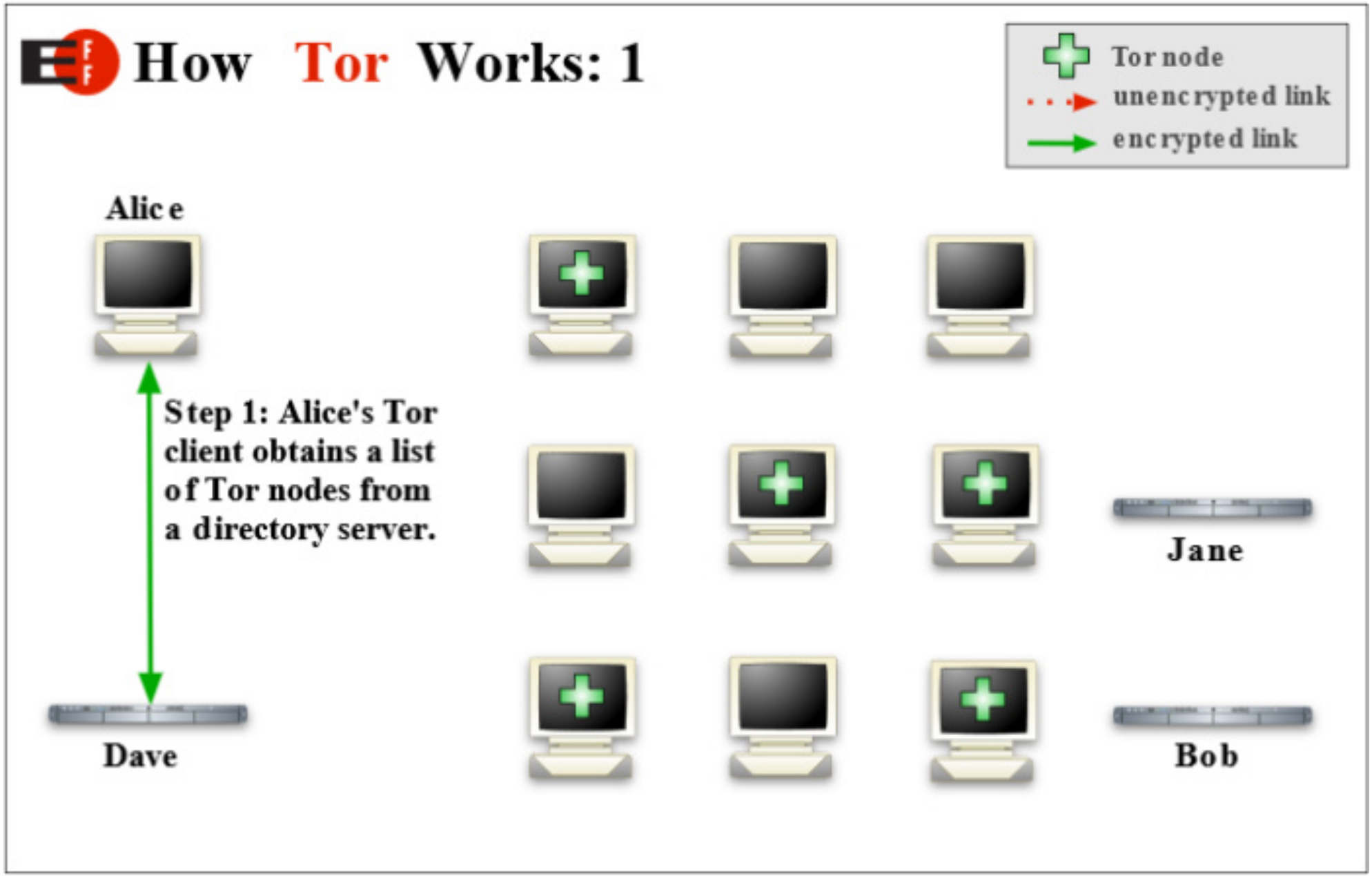
Agenda

- Introduction
- Review
- VPN proxies
- **Tor**
- Tor services
- Tor browser
- Tor history

Tor



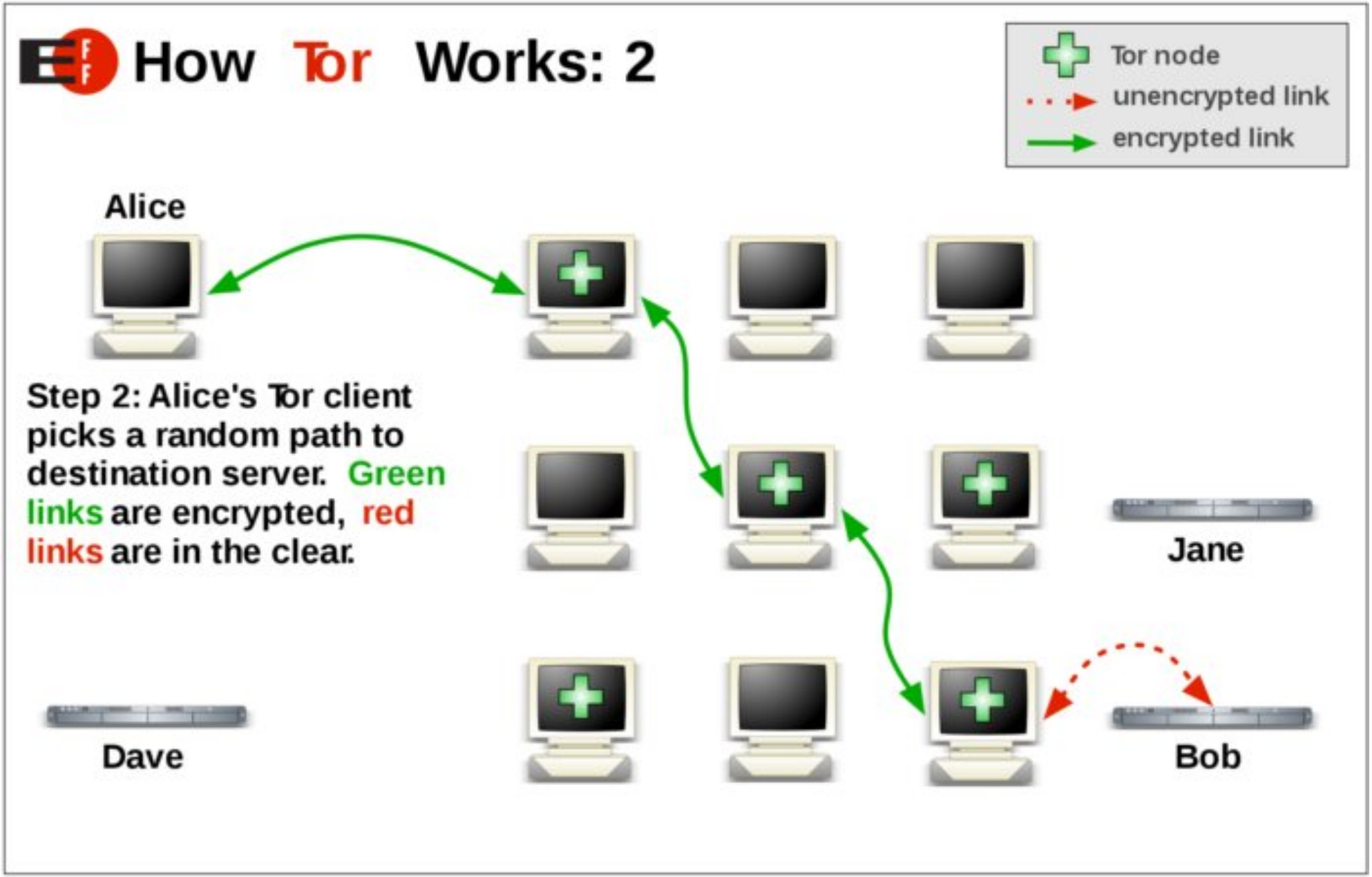
- **T**he **O**nion **R**outer
- Used to fight traffic analysis



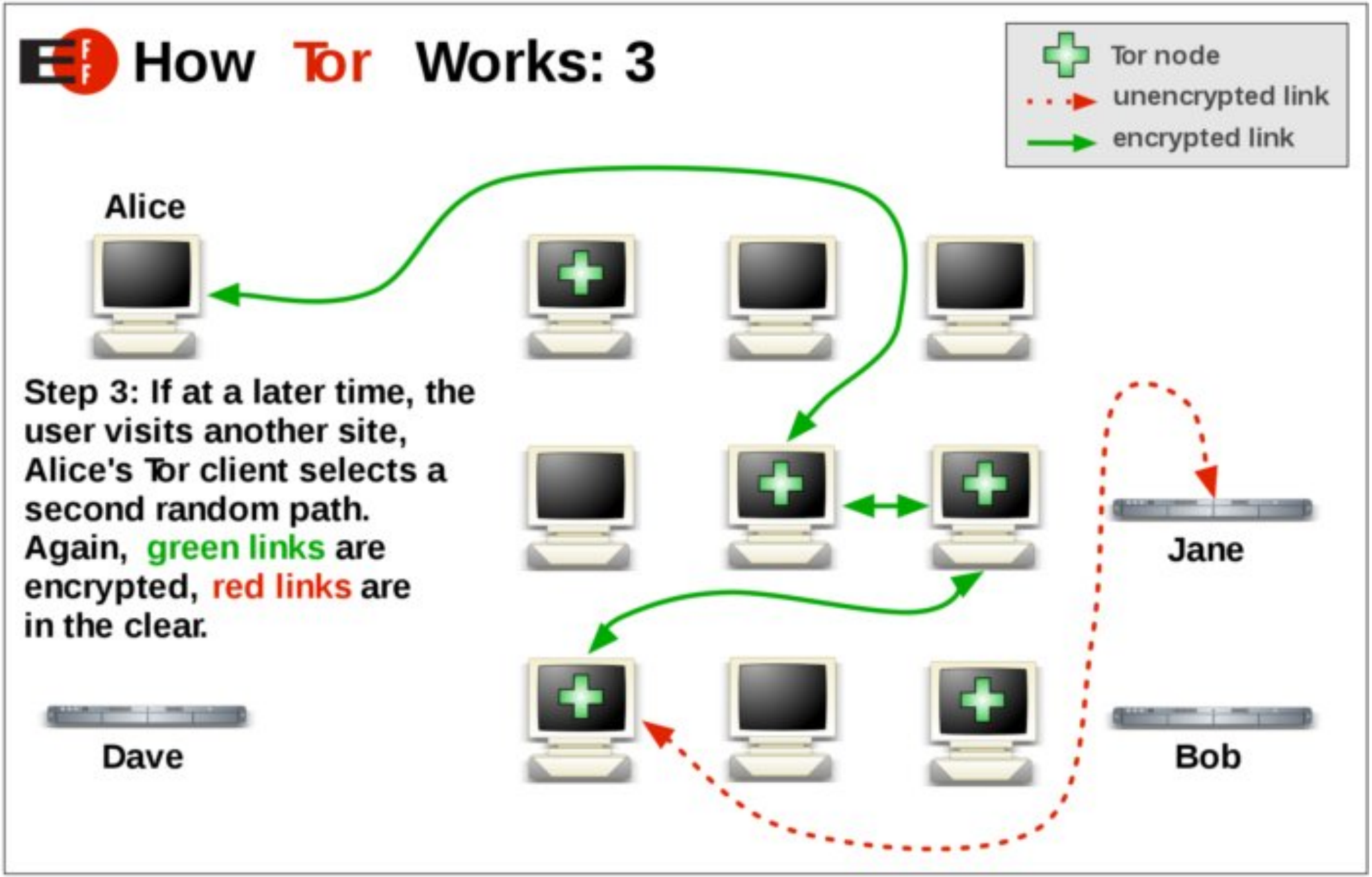
More detail: https://2019.www.torproject.org/about/overview.html.en#thesolution

Tor servers

- Run by volunteers
- There's one in CSE



More detail: https://2019.www.torproject.org/about/overview.html.en#thesolution



More detail: https://2019.www.torproject.org/about/overview.html.en#thesolution

Tor

- Tor encrypts outgoing data multiple times, like layers of an onion
  - One encryption layer for each step in the relay
- Each layer of encryption is peeled off by a relay node in the network. At each layer:
  - Decrypt the current layer
  - Forward to the next destination

Tor vulnerabilities

- Only the first relay node knows the source IP address
- Only the last relay node knows the destination IP address
- To break anonymity, you need to surveil ALL nodes in the Tor circuit

Statistical correlation attack

- Let's say that Elmo and Abby use Tor regularly
  - You control their ISP and collect frequent traffic logs with timestamps
- Also control the ISP for ILoveBigBird.com and DownWithBigBird.org
  - You also collect frequent traffic logs with timestamps
- How can you tell who is pro-BigBird vs anti-BigBird?
  - With enough log data, you can perform statistical correlation attack

Statistical correlation attack

- How to prevent statistical correlation attack
- Elmo and Abby always transmit data to Tor once per second
- If no data to send, just send NULL or random data
- More Tor users conducting more activity on Tor reduces vulnerability

Agenda

- Introduction
- Review
- VPN proxies
- Tor
- **Tor services**
- Tor browser
- Tor history

Tor services

- Tor allows users to anonymously publish services
  - E.g., web pages, or a chat server
- Service locations (*rendezvous points*) must be known to clients
  - Even though censors may want to locate, take down services
- Key idea: layer of indirection
  - *Introduction points* are Tor nodes that relay traffic from clients to services

More detail: https://2019.www.torproject.org/docs/onion-services.html

Finding Tor services

- Curated lists
  - Reddit
  - Hidden Wiki
  - ... and many others
- Hidden service search engines
  - Ahmia
  - Torch
  - Not Evil
  - ... and many others
- Search engine crawlers on dark web
  - Route crawler GET requests to .onion sites through Tor

Agenda

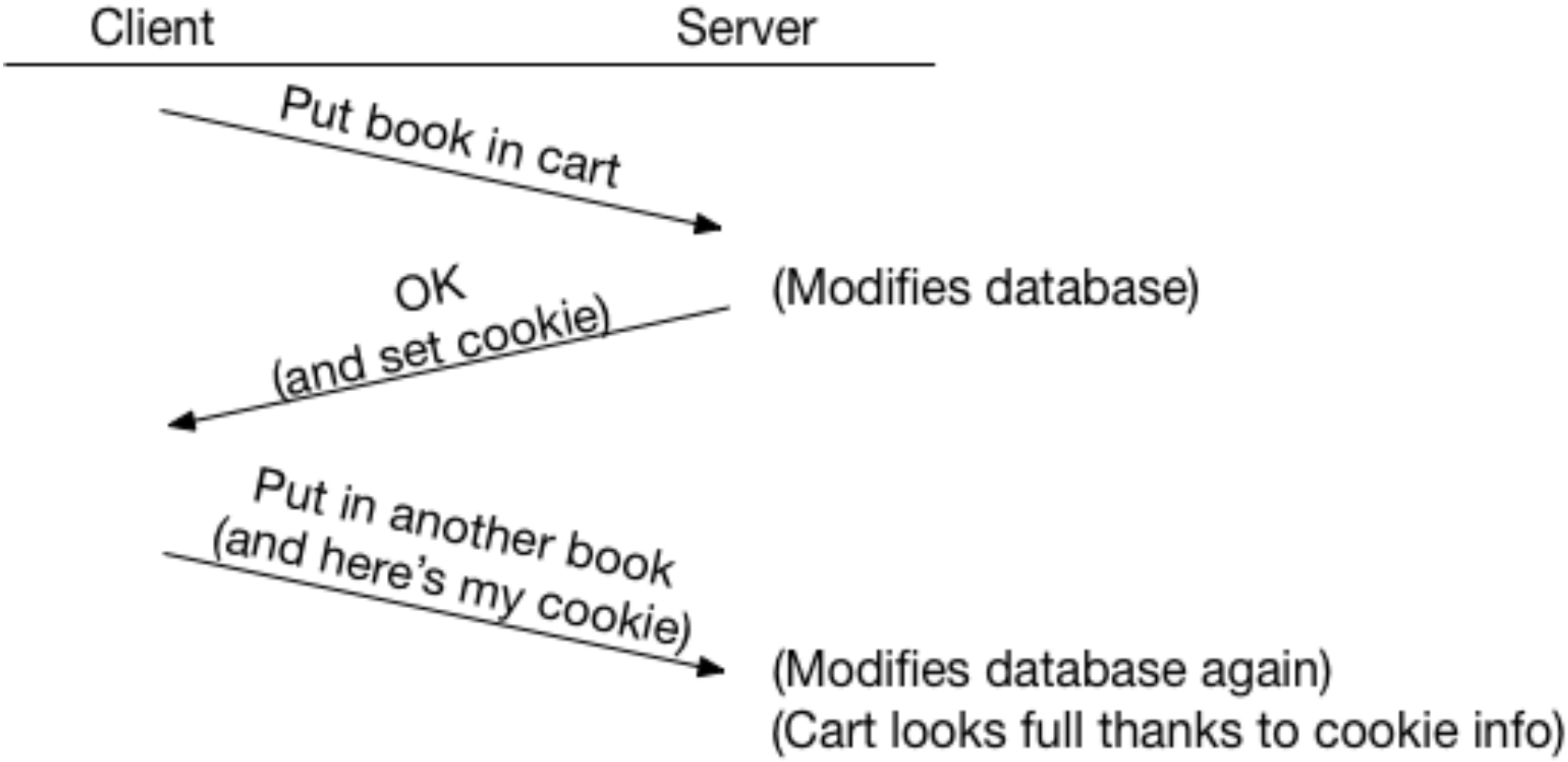
- Introduction
- Review
- VPN proxies
- Tor
- Tor services
- **Tor browser**
- Tor history

Missing link

- Client can browse anonymously with Tor connection
- Server can publish anonymously with Tor services
- Is that enough?

- Session cookies, including third party trackers are ubiquitous
- Your browser reveal you

Cookies


Client
Server
Put book in cart
OK
(and set cookie)
(Modifies database)
Put in another book
(and here's my cookie)
(Modifies database again)
(Cart looks full thanks to cookie info)

Third-party cookies

- Page may contain objects from many sources
  - Scripts, images, etc.
- These 3rd-party objects set and get cookies
- Example: nytimes.com

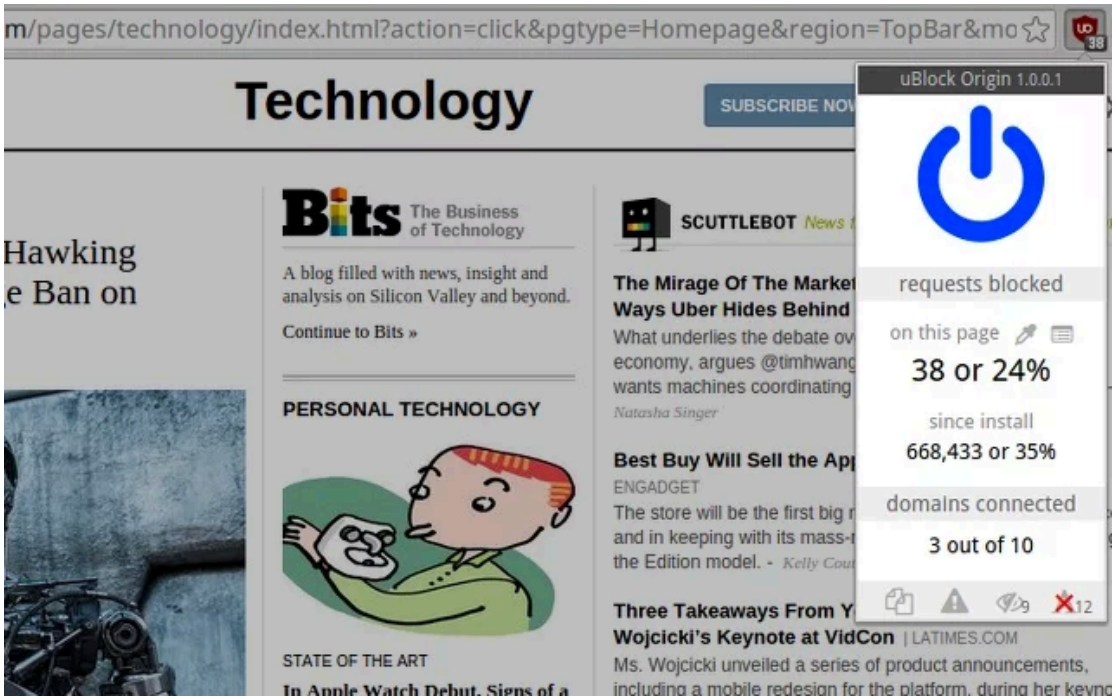
```
<html>
  <head>
    <script
      src="https://tags.bluekai.com/site/50550?ret=js&limit=1"
      type="text/javascript"
    >
    </script>
  </head>
  <body>
    <img src="https://static01.nyt.com/newsgraphics/2018/01/18/shutdown-deal-factions/assets/images/lindsey-graham.png">
  </body>
</html>
```

Example: Google third party cookies

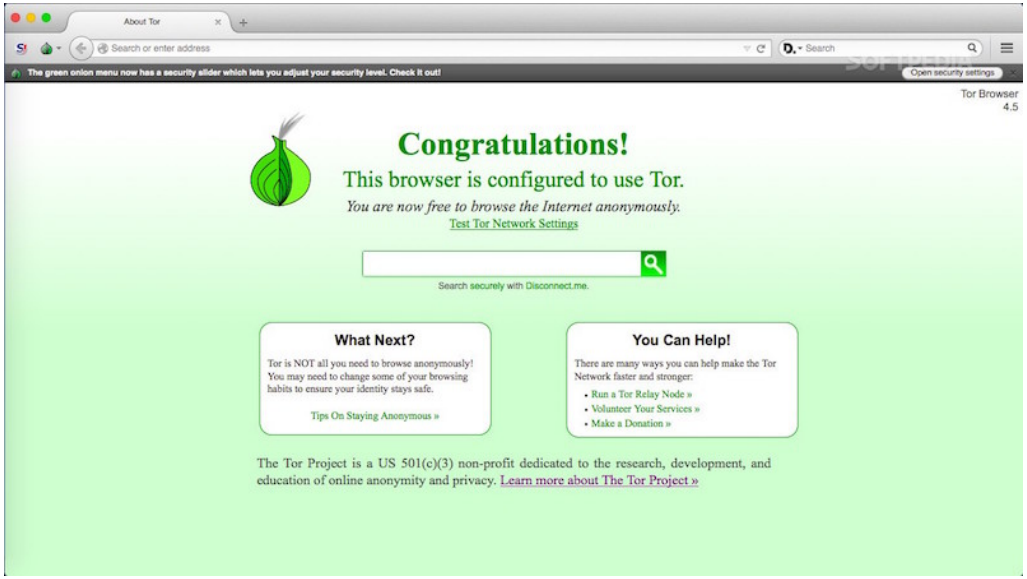
- You type nytimes.com into your browser
- Browser issues `GET` request to nytimes.com
  - Includes nytimes.com cookies
- Browser receives HTML for nytimes.com
  - HTML includes some JavaScript via the `<script>` tag
- Browser executes JavaScript included by nytimes.com
  - JS code figures out you are on nytimes.com, e.g., with `window.location.href`
  - JS codes initiates a request to doubleclick.net
- Browser issues `GET` request to doubleclick.net
  - Includes doubleclick.net cookie
  - Appends your current location (nytimes.com) to the URL
- Now, doubleclick.net (AKA Google) knows you visited nytimes.com

Avoiding trackers

- Add-ons like uBlock Origin, Privacy Badger, Brave, Disconnect or ScriptNo block tracking ads
- Monitors embedded links and blacklists trackers



Tor browser



- Routes all traffic via Tor
- Essential privacy plugins
  - NoScript and HTTPS Everywhere
- Non-tracking search engine
  - DuckDuckGo instead of Google, Bing and friends
- Modified version of Firefox

https://www.torproject.org/download/

Browser fingerprinting

- Torbrowser avoids browser fingerprinting
- Browser fingerprinting: collect enough information from a client, you can (probably) identify them uniquely
  - User-Agent string
  - Screen size
  - Browser plugin details
  - Language
  - Platform
  - Etc.

https://www.torproject.org/download/

Not anonymous browsing

- We've talked about *more* anonymous browsing
  - Tor Browser and Duck Duck Go are a good start
- What about *less* anonymous browsing?
  - Google Chrome tracks you, even in incognito mode
  - Google search engine tracks you with Google login, third party cookies *and* Chrome!
- Middle ground: Firefox and Duck Duck Go

Firefox > Chrome

- Firefox browser

- Duck Duck Go search engine

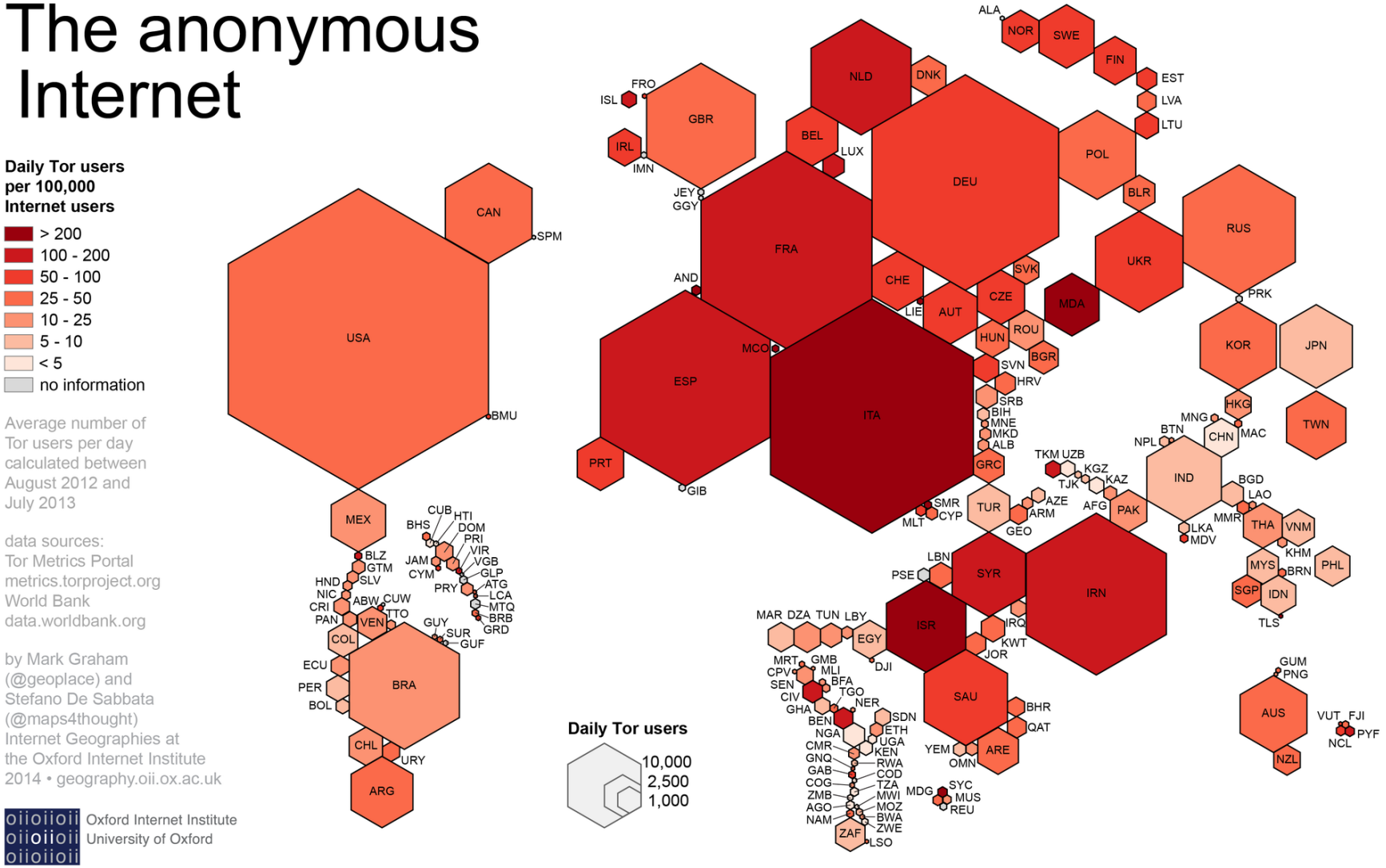
- WAY WAY WAY better privacy protections, even without Tor

Agenda

- Introduction
- Review
- VPN proxies
- Tor
- Tor services
- Tor browser
- **Tor history**

Tor usage

The anonymous Internet



Tor history

- Developed by US Naval Research Laboratory
- Goal: protect US intelligence communications online
- Now a nonprofit
  - US Government is a major funder

Silk Road and Ross Ulbricht

- Silk Road was a darknet market website
  - Access with Tor, pay with Bitcoin
  - Known as a platform for selling illegal drugs
- FBI shut down in 2013
- Creator Ross Ulbricht AKA "Dread Pirate Roberts"
- Got caught because he used the same username (altoid) when asking for programming help. Included email and full name.

https://en.wikipedia.org/wiki/Silk_Road_(marketplace)

https://en.wikipedia.org/wiki/Ross_Ulbricht

Silk Road and Ross Ulbricht

- "To prevent Ulbricht from encrypting or deleting files on the laptop he was using to run the site as he was arrested, two agents pretended to be quarreling lovers. When they had sufficiently distracted him, a third agent grabbed the laptop.
- Good read: American Kingpin by Nick Bilton

https://en.wikipedia.org/wiki/Silk_Road_(marketplace)

https://en.wikipedia.org/wiki/Ross_Ulbricht

Tor Mail

- Tor Mail was an anonymous email service
  - Think GMail as a hidden Tor service
- "No emails, logs or personal data were stored on those servers, thus it doesn't matter if they are seized or shut down."
- 2013, saying "Down for Maintenance"
  - Never came back online
- Zero day JavaScript attack injected into Tor Browser Bundle
  - Exploited to send users' IP addresses and Windows computer names to a server in Virginia

More uses for Tor

- Break censorship barriers
  - E.g., in places that censor web content
- Political activism
  - E.g., in places that discourage free speech
- Net neutrality
  - E.g., ISP speeds up some traffic, slows down other traffic
- Reporting abuse or corruption
  - E.g., whistleblower concerned for safety, Tor enables anonymous communication

Summary

- Tor enables anonymity online
- Not perfect, but vastly more anonymous than traditional browsing