Anonymous Communication and Internet Freedom

March 29, 2018

Anonymous Communication

"On the Internet, nobody knows you're a dog."

Goup Discussions:

- When you are browsing on Internet, what may leak your identity?

You Are Not Anonymous

- Your IP address can be linked directly to you
  - ISPs store communications records
  - Usually for several years (Data Retention Laws)
  - Law enforcement can subpoena these records
- Your browser is being tracked
  - Cookies, Flash cookies, E-Tags, HTML5 Storage
  - Browser fingerprinting
- Your activities can be used to identify you
  - Unique websites and apps that you use
  - Types of links that you click

Wiretapping is Ubiquitous

- Wireless traffic can be trivially intercepted
  - Airsnort, Firesheep, etc.
  - Wifi and Cellular traffic!
  - Encryption helps, if it's strong
    - WEP and WPA are both vulnerable!
- Tier 1 ASs and IXPs are compromised
  - NSA, GCHQ, "5 Eyes"
  - ~1% of all Internet traffic
  - Focus on encrypted traffic

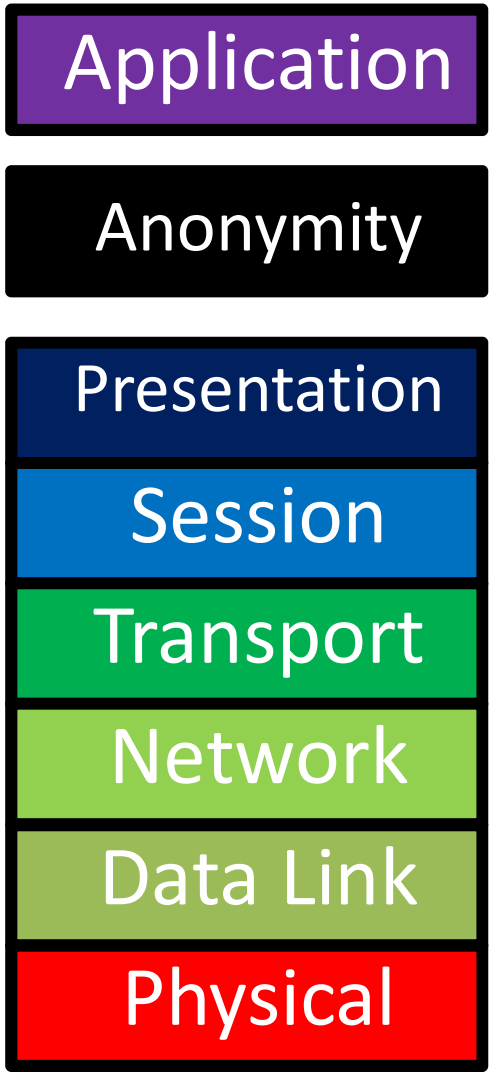
Who Uses Anonymity Systems?

- "If you're not doing anything wrong, you shouldn't have anything to hide."
  - Implies that anonymous communication is for criminals
- The truth: who uses Tor?
  - Journalists
  - Law enforcement
  - Human rights activists
  - Normal people
    - Business executives
    - Military/intelligence personnel
    - Abuse victims
- Fact: Tor was/is developed by the Navy

Why Do We Want Anonymity?

- To protect privacy
  - Avoid tracking by advertising companies
  - Viewing sensitive content
    - Information on medical conditions
    - Advice on bankruptcy
- Protection from prosecution
  - Not every country guarantees free speech
  - Downloading copyrighted material
- To prevent chilling-effects
  - It’s easier to voice unpopular or controversial opinions if you are anonymous

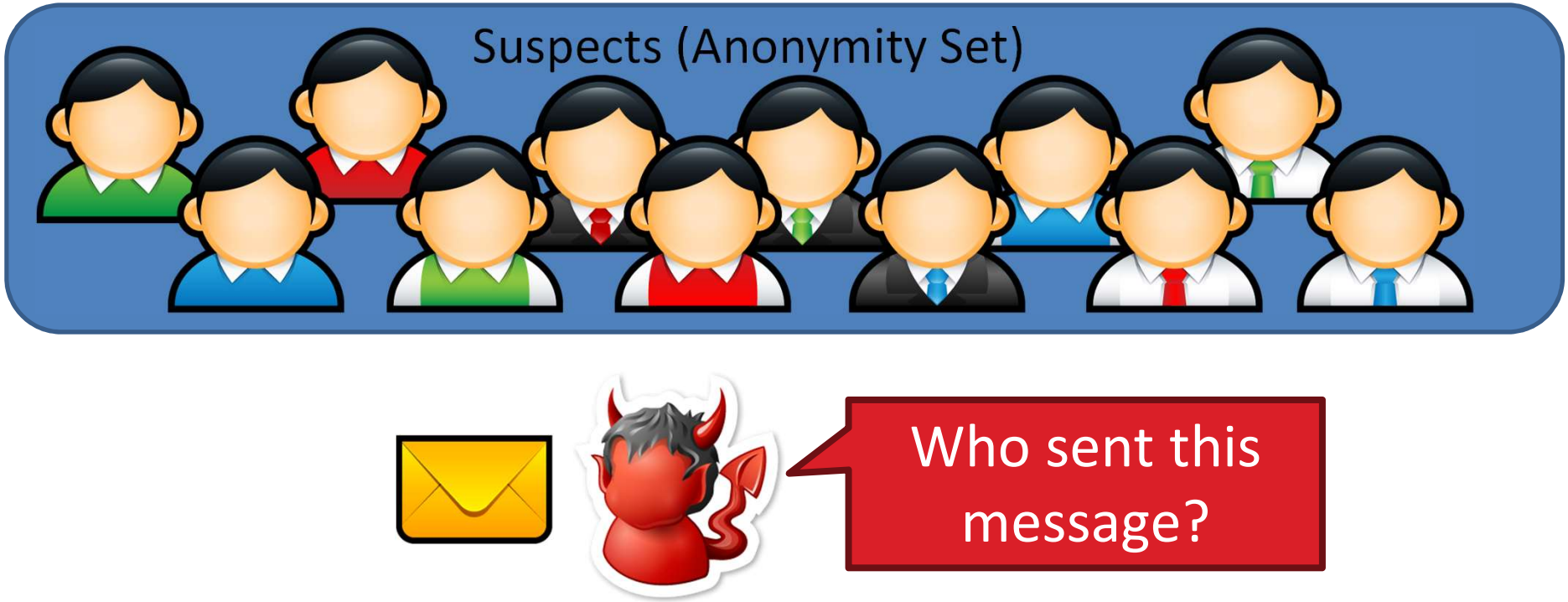
Anonymity Layer



- Function:
  - Hide the source, destination, and content of Internet flows from eavesdroppers
- Key challenge:
  - Defining and quantifying anonymity
  - Building systems that are resilient to deanonymization
  - Maintaining performance

Quantifying Anonymity

- How can we calculate how anonymous we are?
  - **Anonymity Sets**

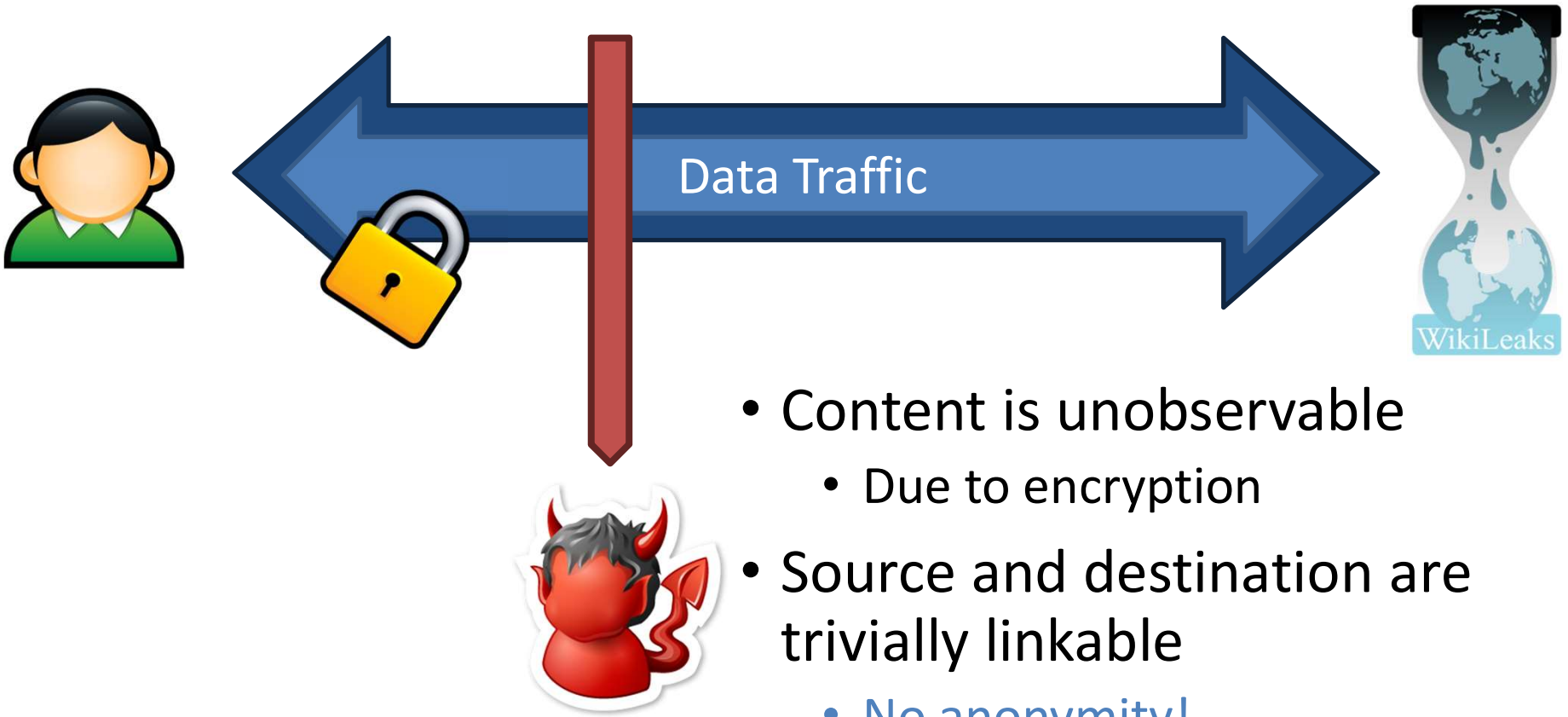


- Larger anonymity set = stronger anonymity

Other Definitions

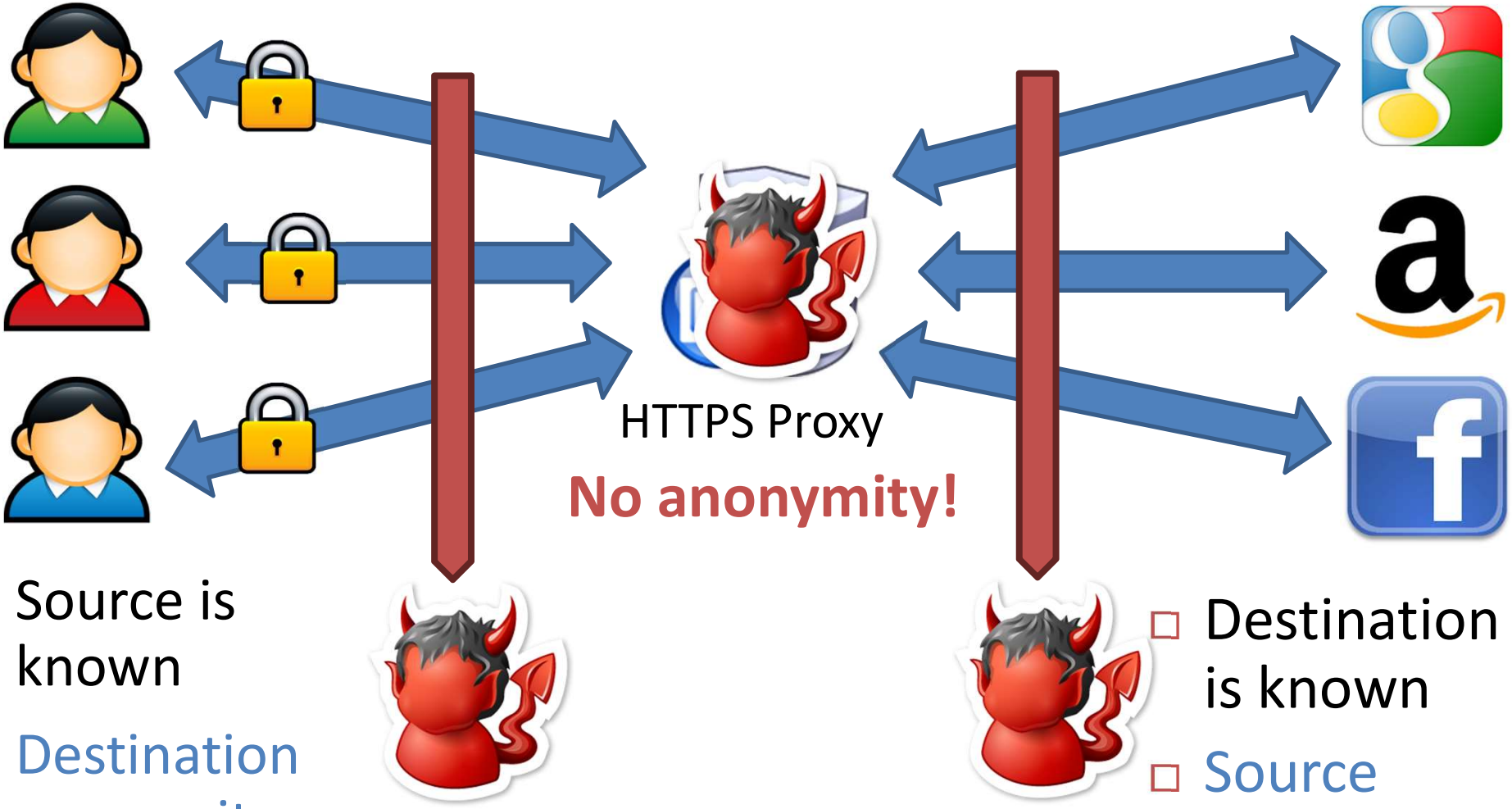
- Unlinkability
  - From the adversaries perspective, the inability the link two or more items of interest
    - E.g. packets, events, people, actions, etc.
  - Three parts:
    - Sender anonymity (who sent this?)
    - Receiver anonymity (who is the destination?)
    - Relationship anonymity (are sender A and receiver B linked?)
- Unobservability
  - From the adversaries perspective, items of interest are indistinguishable from all other items

Crypto (SSL)



- Content is unobservable
  - Due to encryption
- Source and destination are trivially linkable
  - No anonymity!

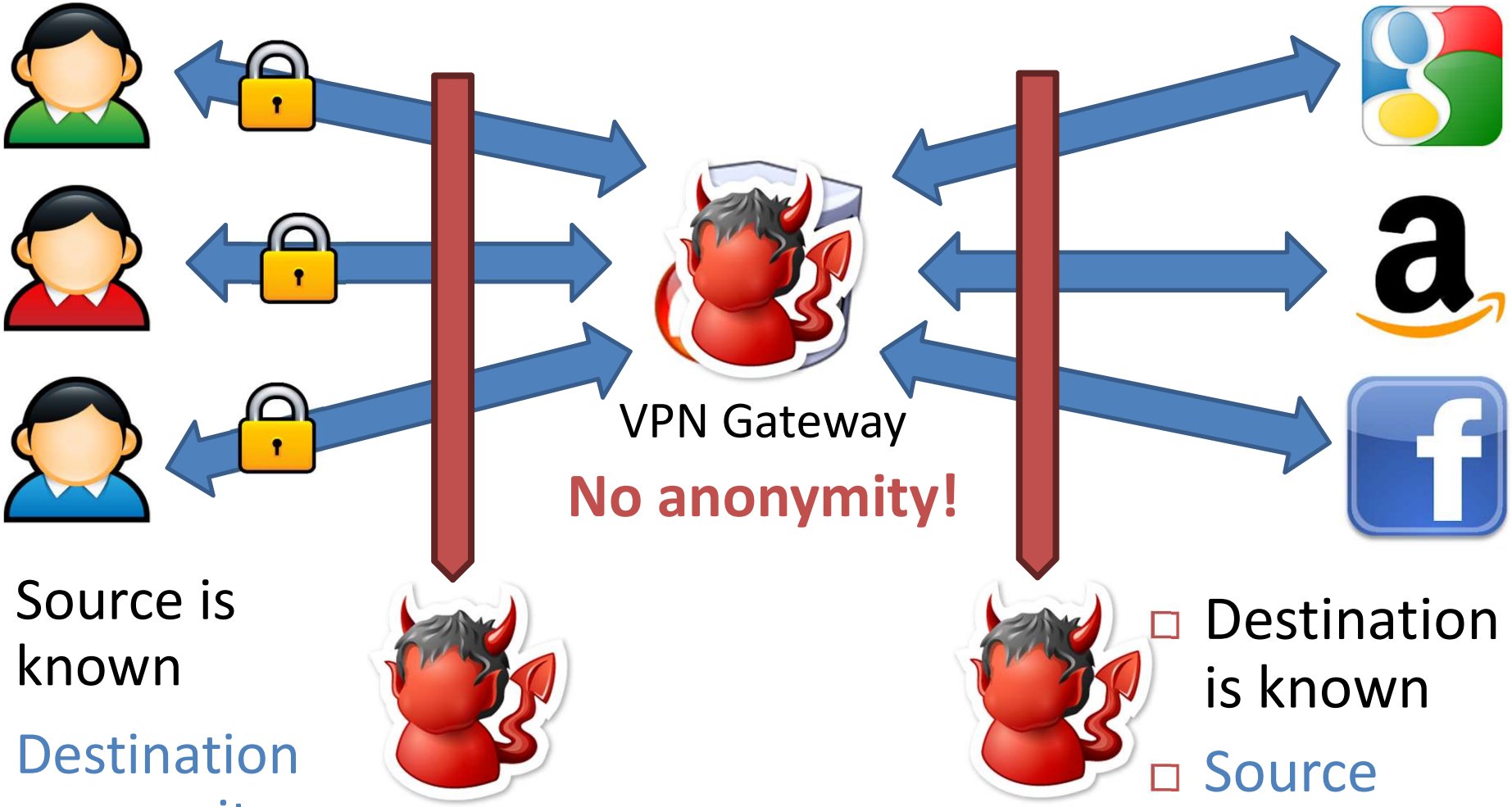
Anonymizing Proxies



- Source is known
- Destination anonymity

- Destination is known
- Source anonymity

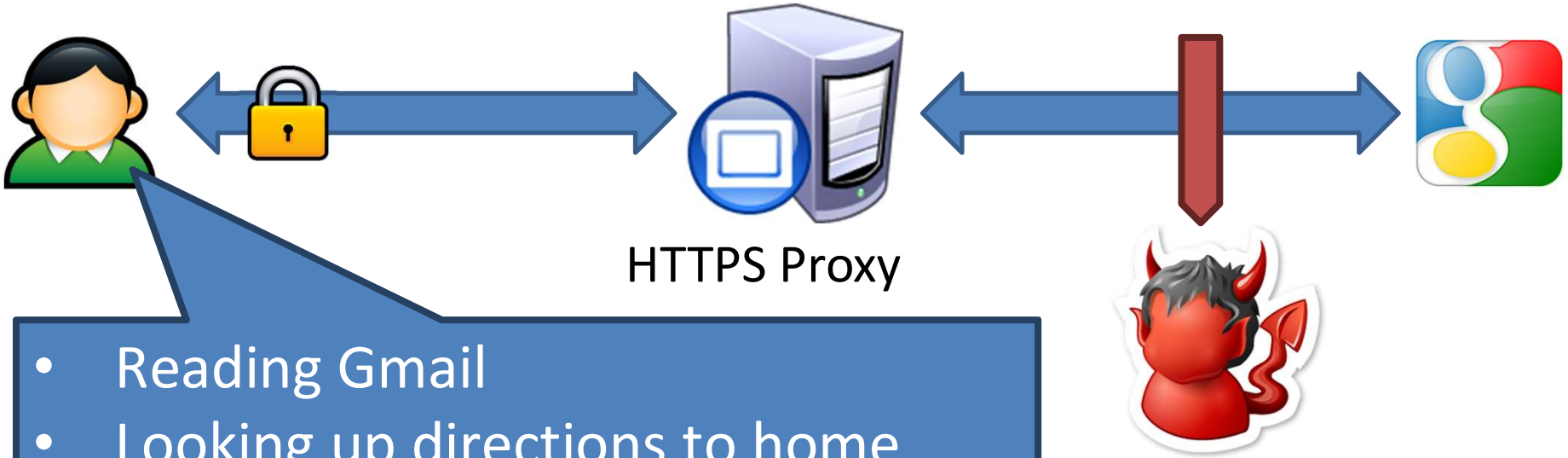
Anonymizing VPNs



- Source is known
- Destination anonymity

- Destination is known
- Source anonymity

Using Content to Deanonymize



- Fact: the NSA leverages common cookies from ad networks, social networks, etc. to track users

Data To Protect

- Personally Identifiable Information (PII)
  - Name, address, phone number, etc.
- OS and browser information
  - Cookies, etc.
- Language information
- IP address
- Amount of data sent and received
- Traffic timing

Anonymity

- Anonymity: Concealing your identity

- In the context of the Internet, we may want anonymous communications
  - **Communications where the identity of the source and/or destination are concealed**

- Not to be confused with confidentiality
  - Confidentiality is about contents, anonymity is about identities

Anonymity

- Internet anonymity is *hard**
  - Difficult if not impossible to achieve on your own
  - Right there in every packet is the source and destination IP address
  - * But it's easy for bad guys. Why?
- You generally need help
- State of the art technique: **Ask someone else to send it for you**
  - (Ok, it's a bit more sophisticated than that…)

Proxies

- Proxy: Intermediary that relays our traffic
- Trusted 3rd party, e.g. …

Hide your IP address with server locations world-wide

Our advanced VPN client enables you to switch server locations at any given time, with servers currently 23+ countries. Our software will hide your IP address (your online 'finger print') and all traffic will be tunneled through our remote servers. Virtually reside in another country with ease. Learn more ».

Learn more | 1 2 3 4 5

Free Proxy

Use our free proxy to surf anonymously online. Proxy to change your IP address, secure your internet connection, hide your internet history and protect your privacy online.

http://www.google.com

Hide My Ass!

Popular sites: YouTube.com | Gmail.com | MySpace.com | FaceBook.com

SSL Encryption

Learn more about our free proxy and how it works.

Our other proxies

Proxies

- Proxy: Intermediary that relays our traffic
- Trusted 3rd party, e.g.… hidemyass.com
  - You set up an encryptedVPN to their site
  - All of your traffic goes through them
- Why easy for bad guys? Compromised machines as proxies.

Alice wants to send a message M to Bob …

… but ensuring that

- Bob doesn't know M is from Alice, and/or
- Eve can't determine that Alice is indeed communicating with Bob.

Alice wants to send a message M to Bob …

… but ensuring that

- Bob doesn't know M is from Alice, and/or
- Eve can't determine that Alice is indeed communicating with Bob.

Alice wants to send a message M to Bob …

… but ensuring that

- Bob doesn't know M is from Alice, and/or
- Eve can't determine that Alice is indeed communicating with Bob.

Alice wants to send a message M to Bob …

… but ensuring that

- Bob doesn't know M is from Alice, and/or
- Eve can't determine that Alice is indeed communicating with Bob.

Alice wants to send a message M to Bob …

… but ensuring that

- Bob doesn't know M is from Alice, and/or
- Eve can't determine that Alice is indeed communicating with Bob.

Alice wants to send a message M to Bob …

… but ensuring that

- Bob doesn't know M is from Alice, and/or
- Eve can't determine that Alice is indeed communicating with Bob.

HMA accepts messages encrypted for it.
Extracts destination and forwards.

Proxies

- Proxy: Intermediary that relays our traffic
- Trusted 3rd party, e.g.… hidemyass.com
  - You set up an encryptedVPN to their site
  - All of your traffic goes through them
  - Why easy for bad guys? Compromised machines as proxies.
- Issues?
  - Performance
  - $80-$200/year
  - "Trusted 3rd Party"
  - **rubber hose cryptanalysis**
    - Government comes a"calling" (Or worse)
    - HMA knowsAlice and Bob are communicating
- Can we do better?

Mix Networks

- Originally designed for anonymous email
  - David Chaum, 1981
  - Concept has since been generalized for TCP traffic
- Hugely influential ideas
  - Onion routing
  - Traffic mixing
  - Dummy traffic (a.k.a. cover traffic)

David Chaum

- Widely recognized as the inventor of digital cash, he is also known for other fundamental innovations in cryptography, including privacy technology and secure election systems. With a PhD in Computer Science from UC Berkeley, he taught at NYU Graduate School of Business and the University of California, led a number of breakthrough projects as well as founded the International Association for Cryptologic Research, the cryptography group at the Center for Mathematics and Computer Science in Amsterdam, DigiCash, the Voting Systems Institute, and the Perspectiva Fund.

Onion Routing

Onion Routing

- This approach generalizes to an arbitrary number of intermediaries (“mixes”)

Onion Routing

- This approach generalizes to an arbitrary number of intermediaries ("mixes")
- Alice ultimately wants to talk to Bob, with the help of HMA, Dan, and Charlie

Onion Routing

- This approach generalizes to an arbitrary number of intermediaries ("mixes")
- Alice ultimately wants to talk to Bob, with the help of HMA, Dan, and Charlie

Alice

M

Onion Routing

- This approach generalizes to an arbitrary number of intermediaries ("mixes")
- Alice ultimately wants to talk to Bob, with the help of HMA, Dan, and Charlie

Alice

$\{M, Bob\}_{K_{Dan}}$

Onion Routing

- This approach generalizes to an arbitrary number of intermediaries ("mixes")
- Alice ultimately wants to talk to Bob, with the help of HMA, Dan, and Charlie

Alice

$\{\{M, Bob\}_{K_{Dan}}, Dan\}_{K_{Charlie}}$

Onion Routing

- This approach generalizes to an arbitrary number of intermediaries ("mixes")
- Alice ultimately wants to talk to Bob, with the help of HMA, Dan, and Charlie

Alice

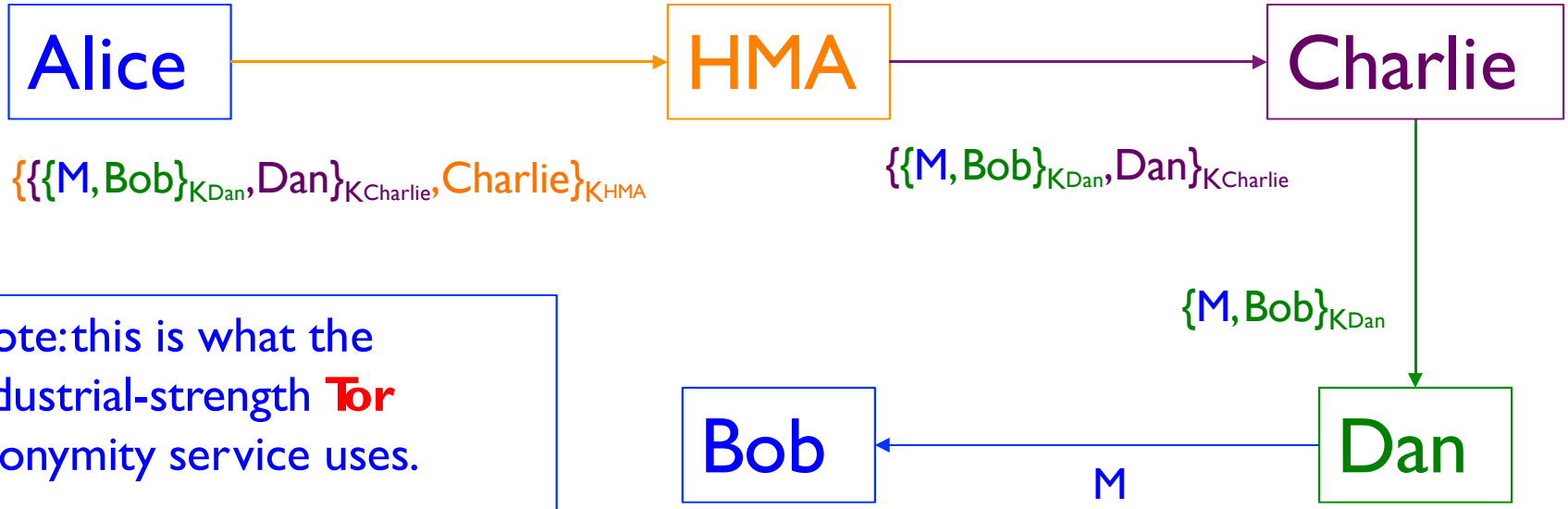
$\{\{\{M, Bob\}_{K_{Dan}}, Dan\}_{K_{Charlie}}, Charlie\}_{K_{HMA}}$

Onion Routing

- This approach generalizes to an arbitrary number of intermediaries ("mixes")
- Alice ultimately wants to talk to Bob, with the help of HMA, Dan, and Charlie

Onion Routing

- This approach generalizes to an arbitrary number of intermediaries ("mixes")
- Alice ultimately wants to talk to Bob, with the help of HMA, Dan, and Charlie
- **As long as any of the mixes is honest, no one can link Alice with Bob**



Note: this is what the industrial-strength **Tor** anonymity service uses.

(It also provides bidirectional communication)

Key concept: No one relay knows both you and the destination!

Onion Routing Issues/Attacks?

- Performance: message bounces around a lot
- Attack: rubber-hose cryptanalysis of mix operators
  - Defense: use mix servers in different countries
    - Though this makes performance worse :-(
- Attack: adversary operates all of the mixes
  - Defense: have lots of mix servers (Tor today: ~2,000)
- Attack: adversary observes when Alice sends and when Bob receives, links the two together
  - A side channel attack – exploits timing information
  - Defenses: pad messages, introduce significant delays
    - Tor does the former, but notes that it's not enough for defense

Tor: The 2nd Generation Onion Router

- Basic design: a mix network with improvements
  - Perfect forward secrecy
  - Introduces guards to improve source anonymity
  - Takes bandwidth into account when selecting relays
    - Mixes in Tor are called relays
  - Introduces hidden services
    - Servers that are only accessible via the Tor overlay

Deployment and Statistics

- Largest, most well deployed anonymity preserving service on the Internet
  - Publicly available since 2002
  - Continues to be developed and improved
- Currently, ~5000 Tor relays around the world
  - All relays are run by volunteers
  - It is suspected that some are controlled by intelligence agencies
- 500K – 900K daily users
  - Numbers are likely larger now, thanks to Snowden

Celebrities Use Tor

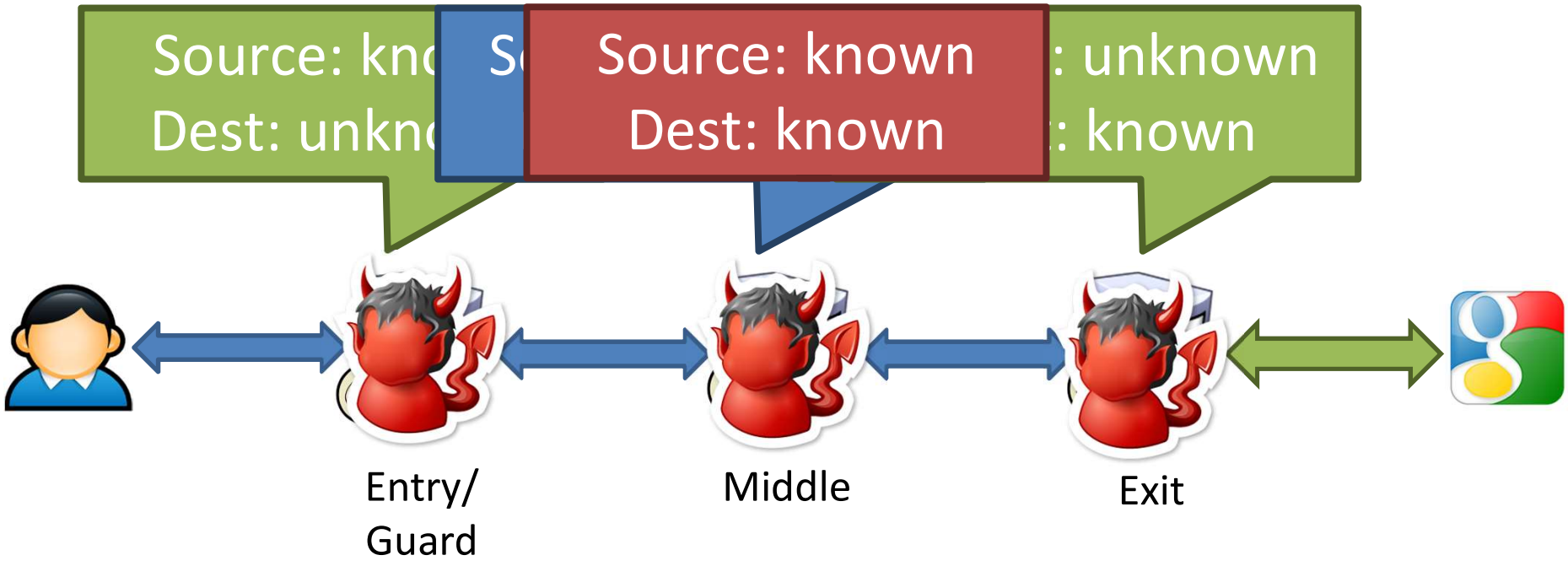
How Do You Use Tor?

1. Download, install, and execute the Tor client
   - The client acts as a SOCKS proxy
   - The client builds and maintains circuits of relays
2. Configure your browser to use the Tor client as a proxy
   - Any app that supports SOCKS proxies will work with Tor
3. All traffic from the browser will now be routed through the Tor overlay

Selecting Relays

- How do clients locate the Tor relays?
- Tor Consensus File
  - Hosted by trusted directory servers
  - Lists all known relays
    - IP address, uptime, measured bandwidth, etc.
- Not all relays are created equal
  - Entry/guard and exit relays are specially labelled
  - Why?
- Tor does not select relays randomly
  - Chance of selection is proportional to bandwidth
  - Why? Is this a good idea?

Attacks Against Tor Circuits

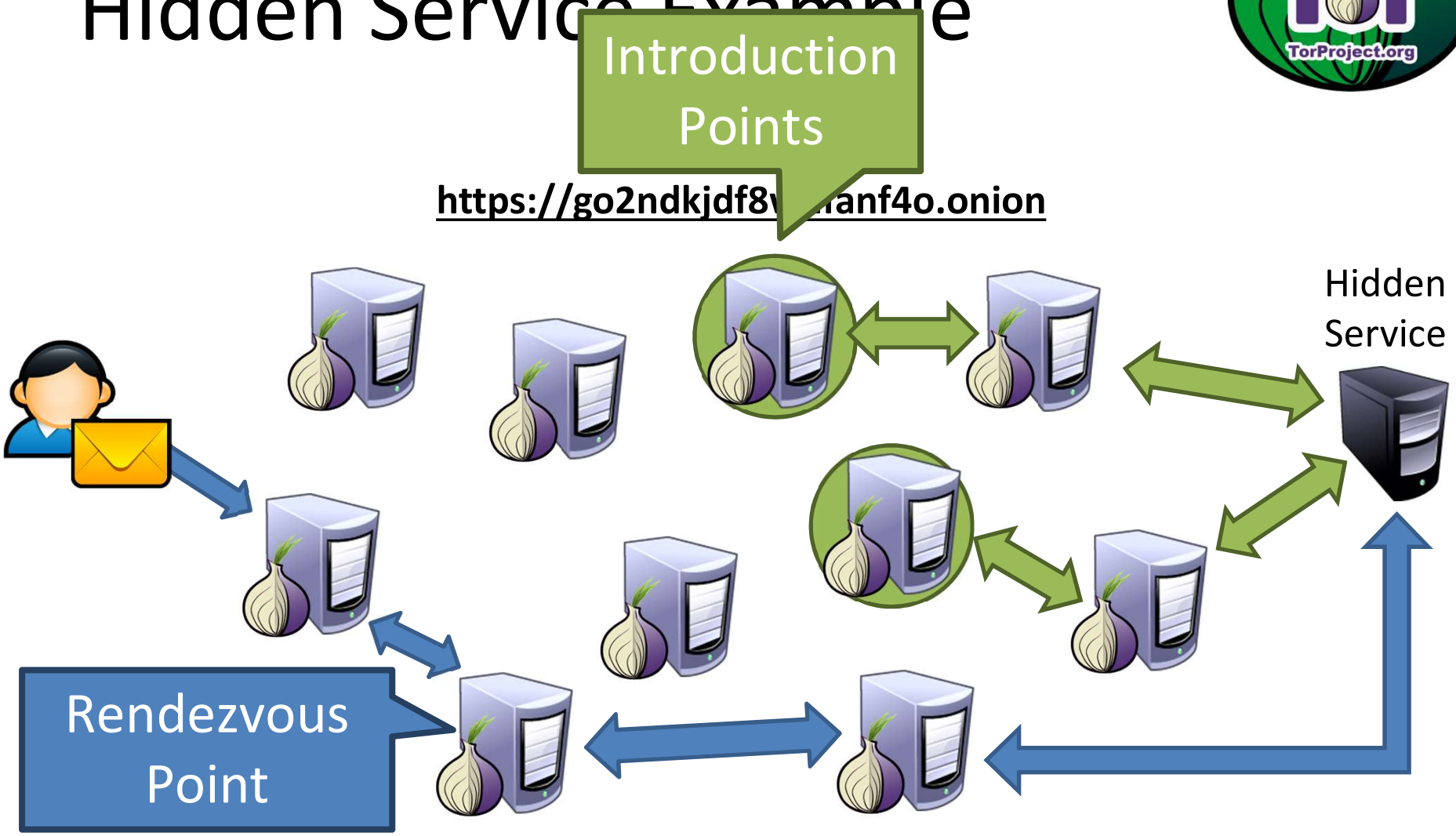


- Tor users can choose any number of relays
  - Default configuration is 3
  - Why would higher or lower number be better or worse?

Hidden Services

- Tor is very good at hiding the source of traffic
  - But the destination is often an exposed website
- What if we want to run an anonymous service?
  - i.e. a website, where nobody knows the IP address?
- Tor supports Hidden Services
  - Allows you to run a server and have people connect
  - … without disclosing the IP or DNS name
- Many hidden services
  - Tor Mail, Tor Char
  - DuckDuckGo
  - Wikileaks
  - The Pirate Bay
  - Silk Road (2.0)

Hidden Service Example



- Onion URL is a hash, allows any Tor user to find the introduction points

Perfect Forward Secrecy

- In
  us
- Pr

- An attacker who compromises a private key can still eavesdrop on future traffic
- … but past traffic is encrypted with ephemeral keypairs that are not stored

- Tor implements Perfect Forward Secrecy (PFS)
  - The client negotiates a new public key pair with each relay
  - Original keypairs are only used for signatures
    - i.e. to verify the authenticity of messages

Tor Bridges

- Anyone can look up the IP addresses of Tor relays
  - Public information in the consensus file
- Many countries block traffic to these IPs
  - Essentially a denial-of-service against Tor
- Solution: Tor Bridges
  - Essentially, Tor proxies that are not publicly known
  - Used to connect clients in censored areas to the rest of the Tor network
- Tor maintains bridges in many countries

Obfuscating Tor Traffic

- Bridges alone may be insufficient to get around all types of censorship
  - DPI can be used to locate and drop Tor frames
  - Iran blocked all encrypted packets for some time
- Tor adopts a pluggable transport design
  - Tor traffic is forwarded to an obfuscation program
  - Obfuscator transforms the Tor traffic to look like some other protocol
    - BitTorrent, HTTP, streaming audio, etc.
  - Deobfuscator on the receiver side extracts the Tor data from the encoding

Conclusions

- Presented a brief overview of popular anonymity systems
  - How do they work?
  - What are the anonymity guarantees?
- Introduced Tor
- Lots more work in anonymous communications
  - Dozens of other proposed systems
    - Tarzan, Bluemoon, etc.
  - Many offer much stronger anonymity than Tor
  - … however, performance is often a problem