Anonymous Communications

Andrew Lewman andrew@torproject.org

December 05, 2012

Who is this guy?

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

https://www.torproject.org

Three hours of this guy talking?

Let's hope not.

Ask questions; early and often.

Agenda

- Definitions and Concepts of Anonymity
- What data?
- Attacks against anonymity
- Deployed Systems (Centralized and Decentralized)

What is Anonymity?

Definitions: Anonymity

- a set of all possible subjects
- state of not being identifiable within anonymity set

Definitions: Unlinkability

- unlinkability of two or more items of interest from the adversary's perspective
  - items can be messages, people, events, actions, etc

Definitions: Unobservability

- state of items of interest being indistinguishable from any items of interest

Definitions: Pseudonymity

- identifiers of sets of subjects

Definitions: Traffic Analysis

- The who, what, when of traffic
- Think of the post office

Definitions: Steganography

- the art and science of writing hidden messages in such a way that no one, apart from the sender and intended recipient, suspects the existence of the message, a form of security through obscurity. – Wikipedia
- alice or bob are talking, but to whom?

Definitions: Cryptography

- protecting content in transit
- does nothing to hide the traffic of items of interest

What data to protect?

- bits of info which put together deanonymize you
- Names of individuals
- location address (street, IP address, zipcode, etc)
- operating system info
- language info
- amount of data sent
- amount of data received
- traffic timing (heartbeats)

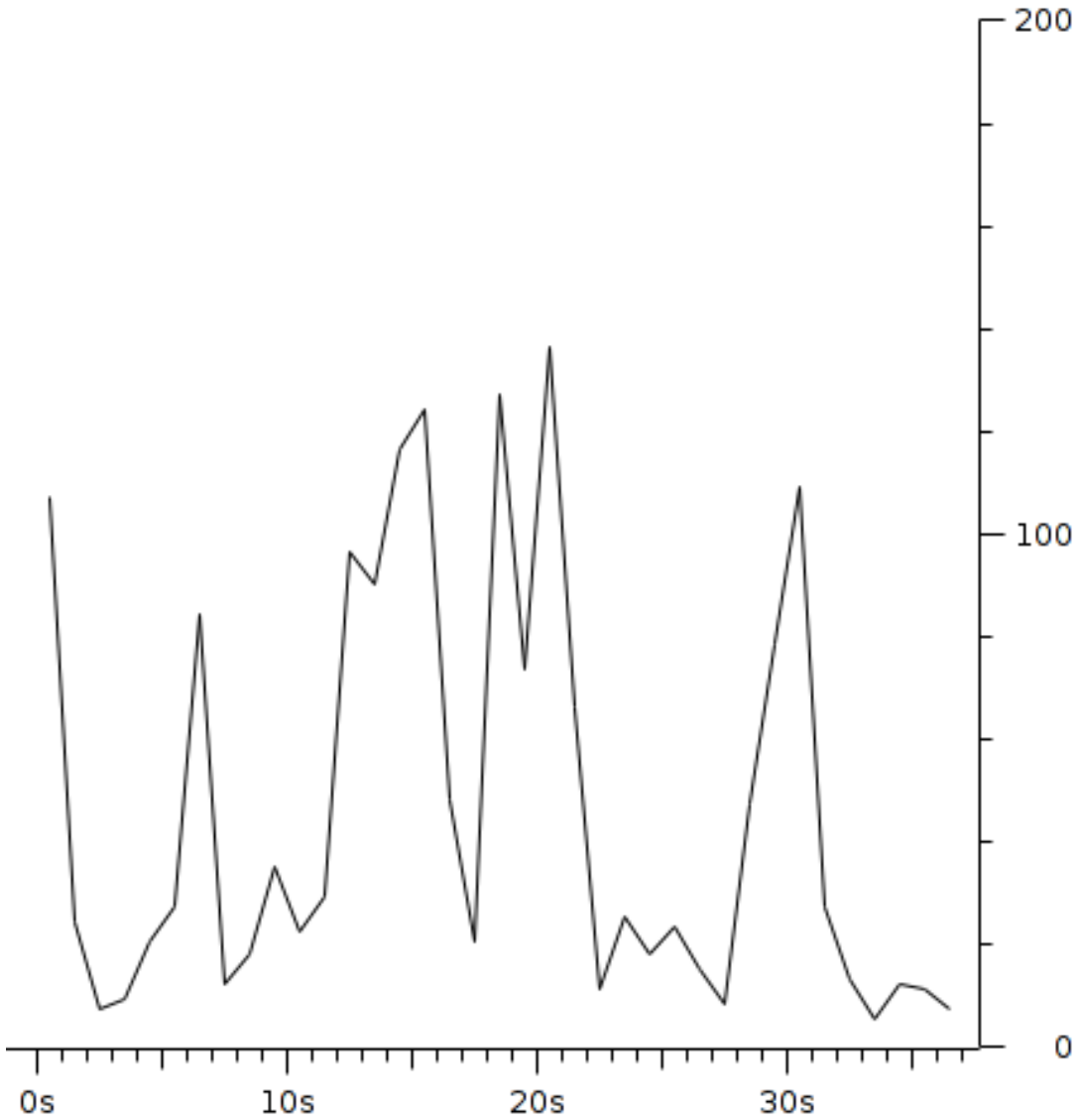
Anonymity Loves Company

- make the set of users as large and coherent as possible to create a large anonymity set

Attacking Anonymity: Timing Analysis

- An attack used to analyze the time properties of data transfer between items of interest.
  - When was data sent?
  - How much was data sent?
  - How long did it take to send the data?
  - When was data received?
- wireshark demo

Attacking Anonymity: Timing Analysis



Attacking Anonymity: Statistical Disclosure

- Also called an intersection attack
- trying to identify mutually disjoint sets of recipients
- exponential time involved per number of messages to be analyzed

Attacking Anonymity: Tagging

- tagging (make one item of interest unique)

Attacking Anonymity: Traffic Confirmation

- who sends, how often, and when
- etherape demo

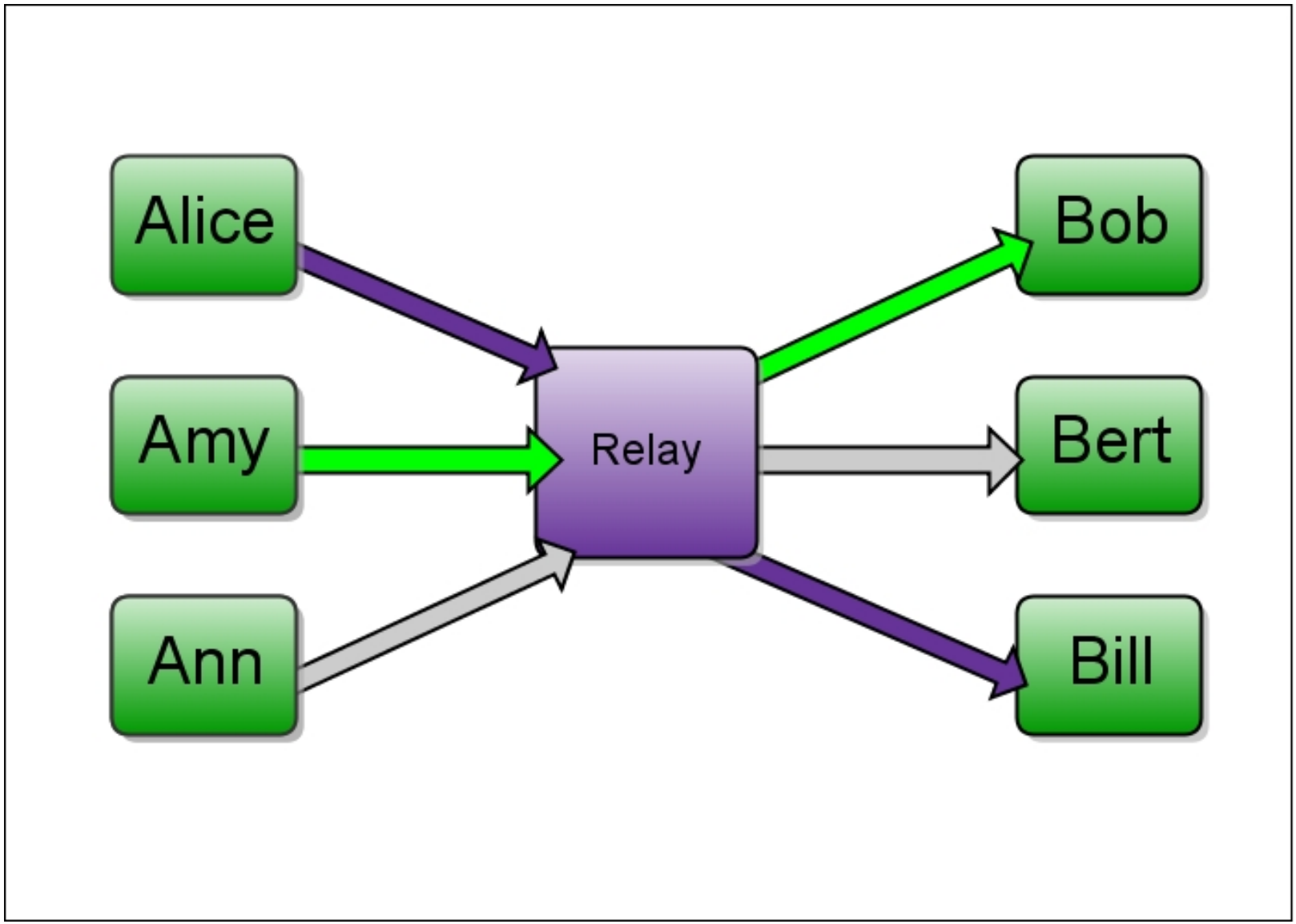
Centralized Systems

- cheap, easy, ubiquitous
- PPTP, IPSec, SSL, SSH, XMPP common protocols

Proxy and VPN Servers

- proxy server works on your behalf
- VPN is virtual private network
  - proxy for the network layers (layers 2 or 3 of OSI model)

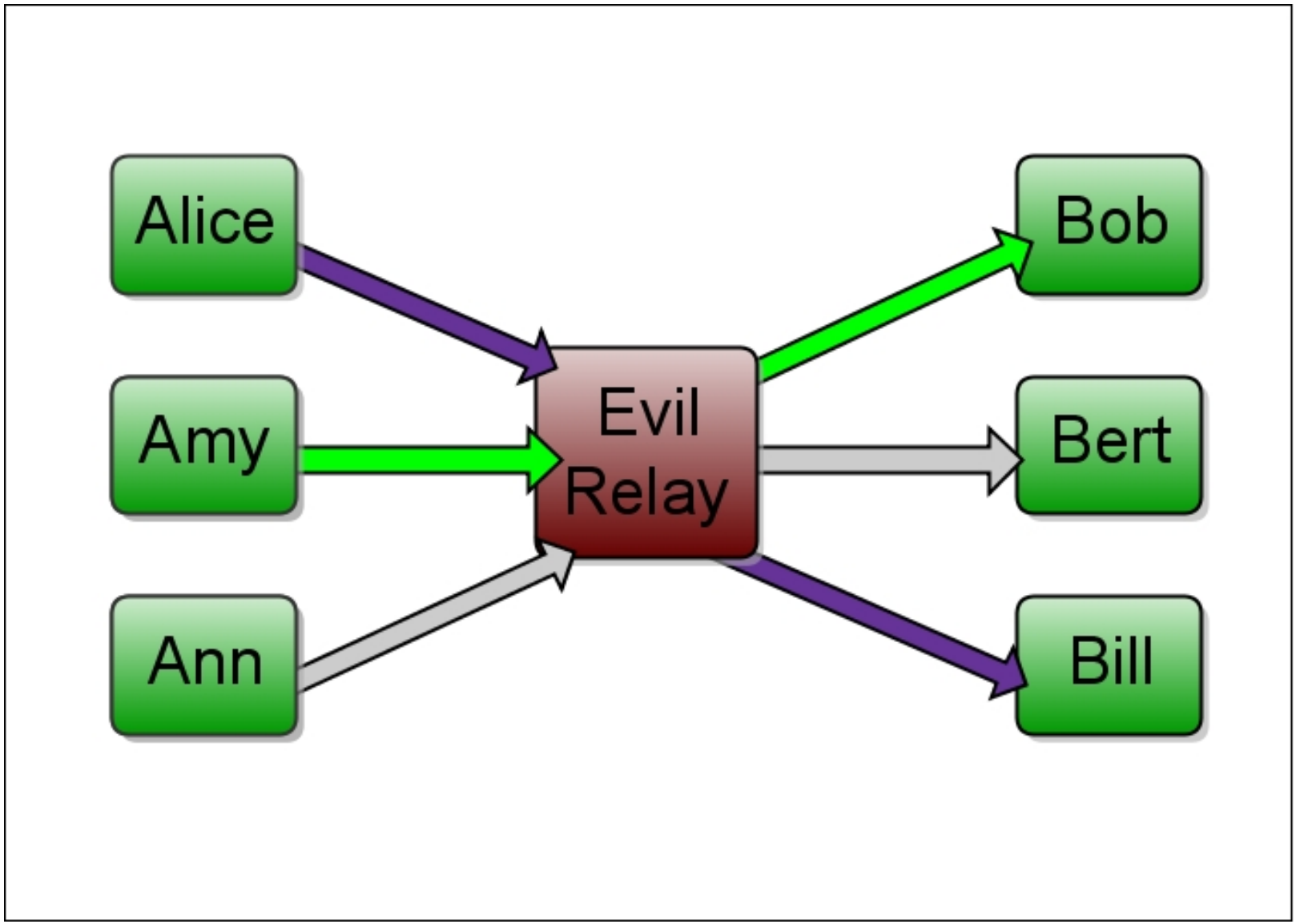
Proxy and VPN Servers



Trusting the provider

- trusting the provider
- promises, contracts, mistakes
- some may filter or clean data before passing on to destination

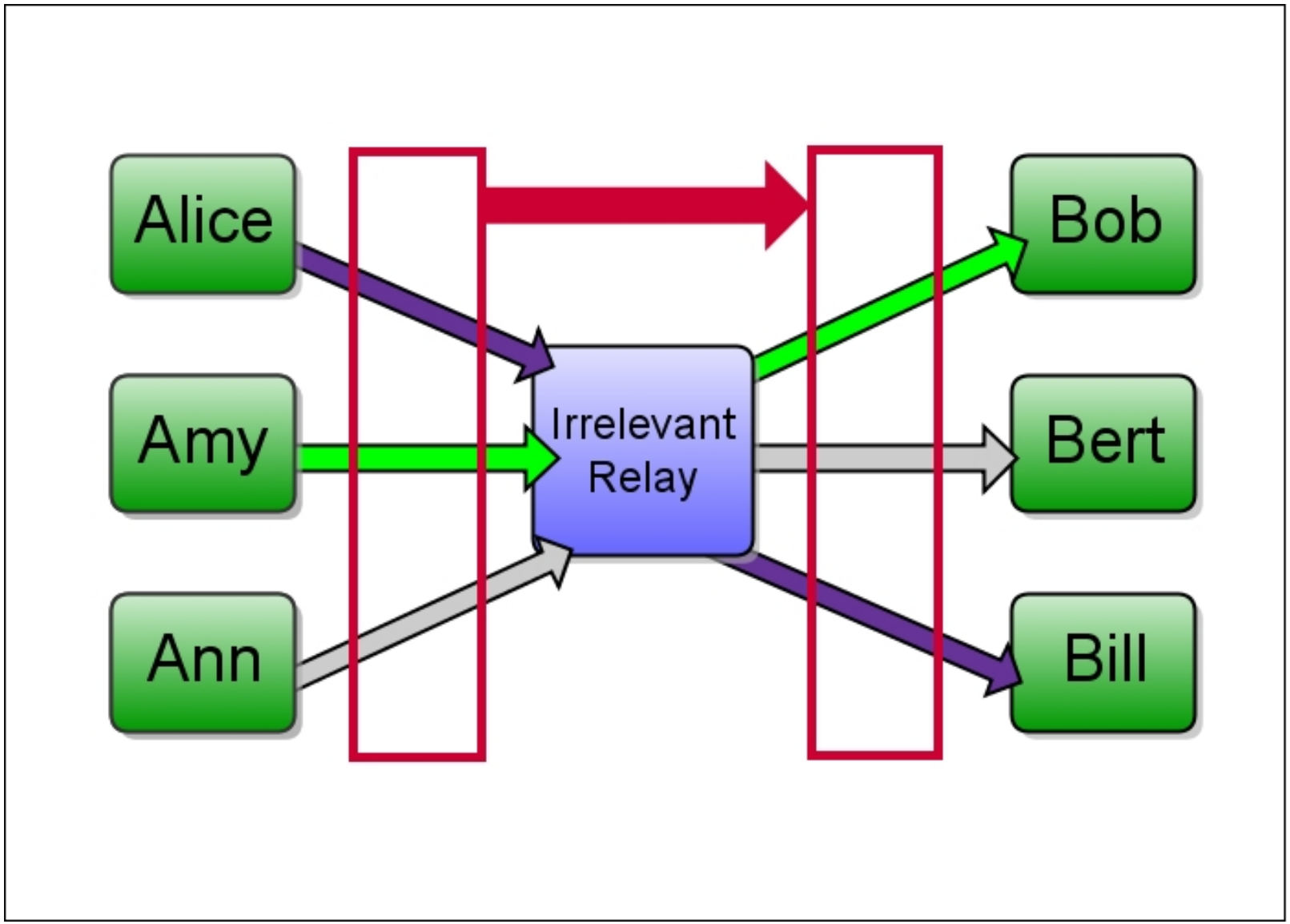
Trusting the provider



Irrelevant provider

- Single machine, or cluster of machines, are connected to a network
- If the proxy provider won't cooperate, use the network around it.

Irrelevant provider



Decentralized Systems

Mix Networks

- cascades (JonDos/JonDonym)
- routes (tor)

Similar Routing networks

- I2P - Garlic routing, closed network, anonymity and reputation
- Freenet - closed network, anonymity, distributed file storage and sharing
- GNUnet - closed network, anonymity, distributed file storage and sharing

Break?

Anyone need a bio-break for 10 minutes?

What is Tor?

- online anonymity software and network
- open source, freely available (3-clause BSD license)
- active research environment:

Drexel, Univ of Waterloo, Georgia Tech, Princeton, Boston University, University College London, Univ of Minnesota, National Science Foundation, Naval Research Labs, Cambridge UK, Bamberg Germany, MIT...

- increasingly diverse toolset:

Tor, Tor Browser Bundle, Tails Live System, Orbot/OrWeb, Tor Weather, Tor auto-responder, Secure Updater, Arm, Tor2Web, and so on.

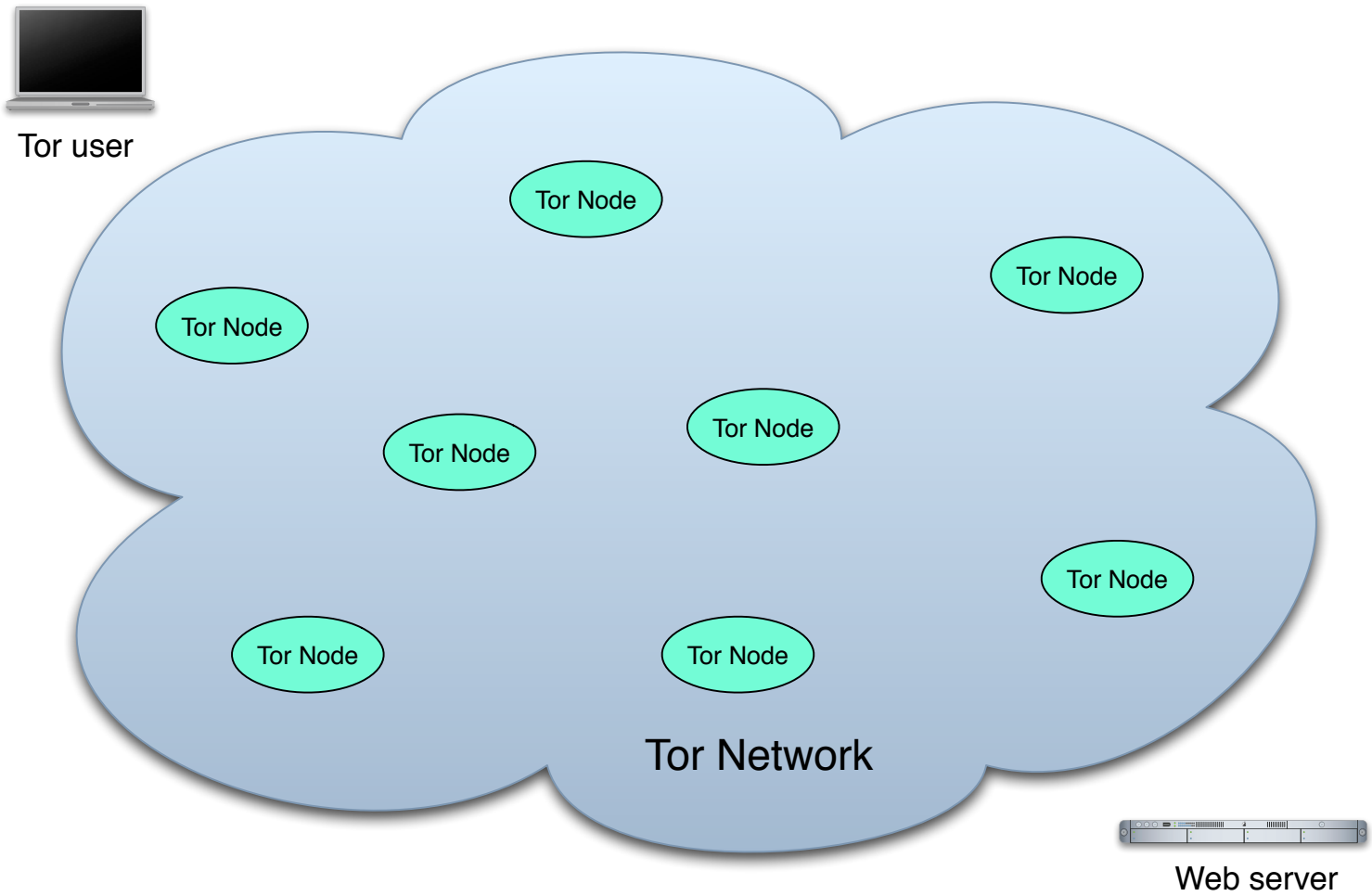
Who uses Tor?

- Normal people
- Journalists
- Law Enforcement
- Human Rights Activists
- Business Execs
- Militaries
- Abuse Victims

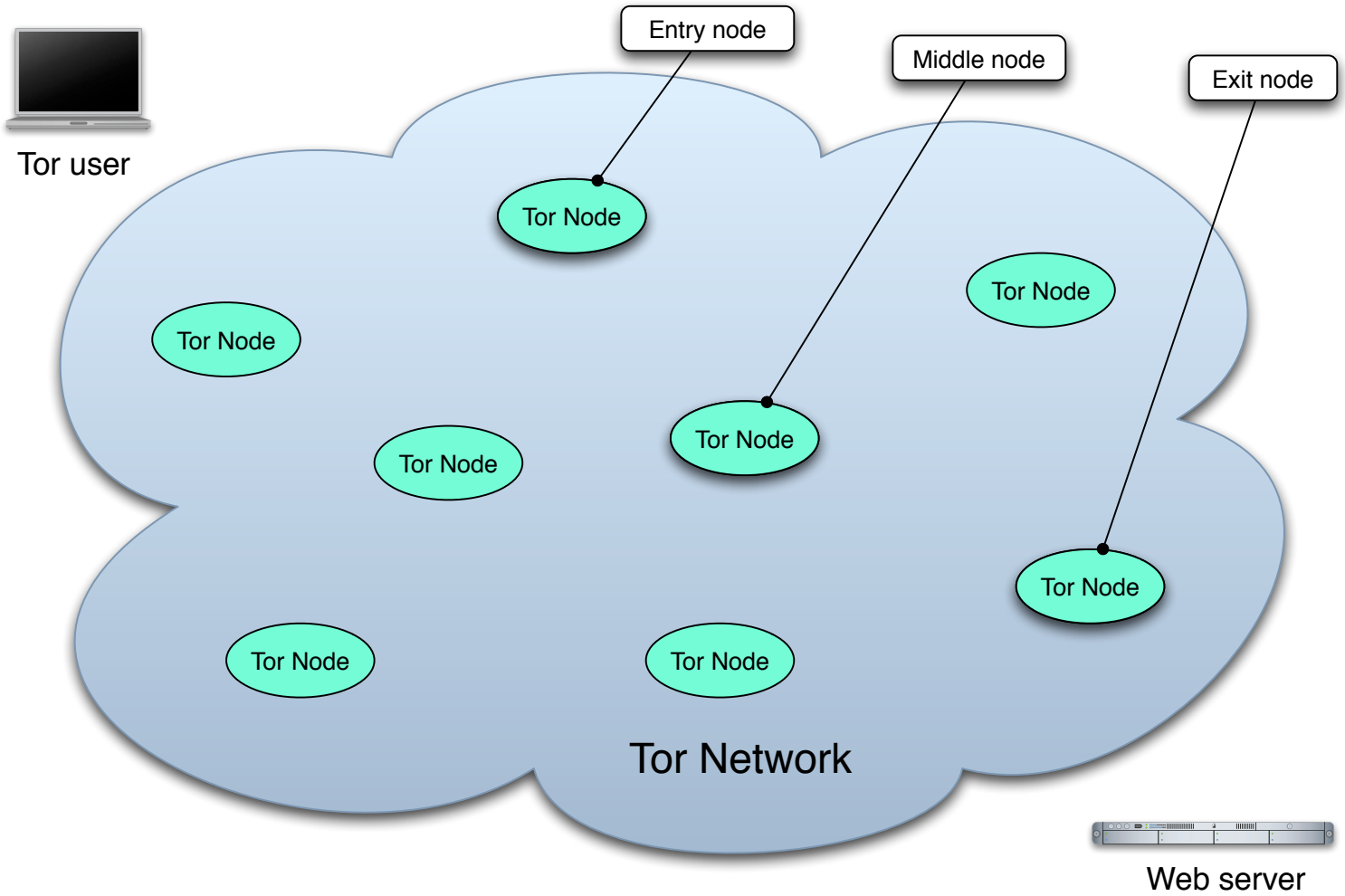
How many people use Tor?

estimated 500k to 900k daily users

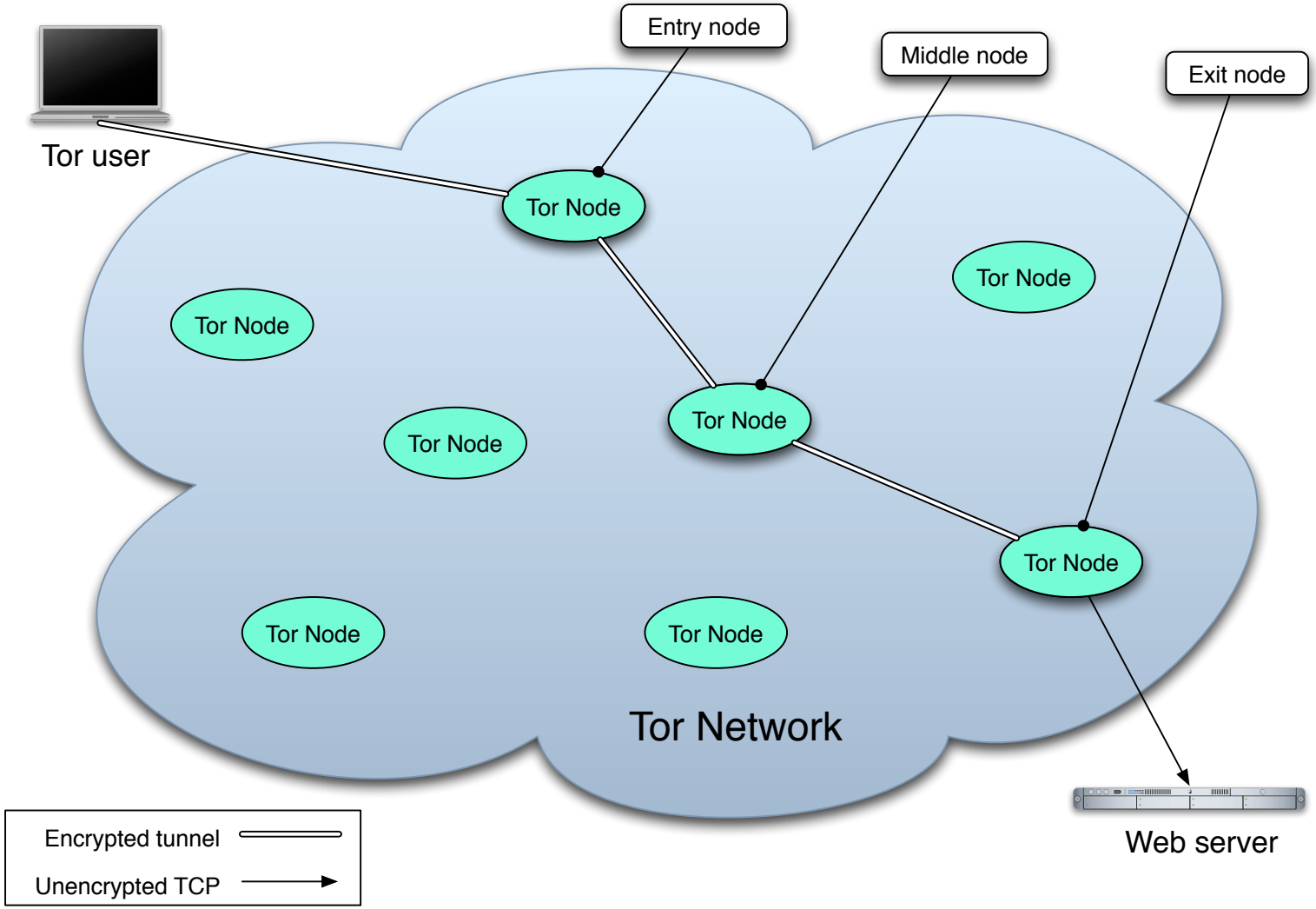
How does Tor work?



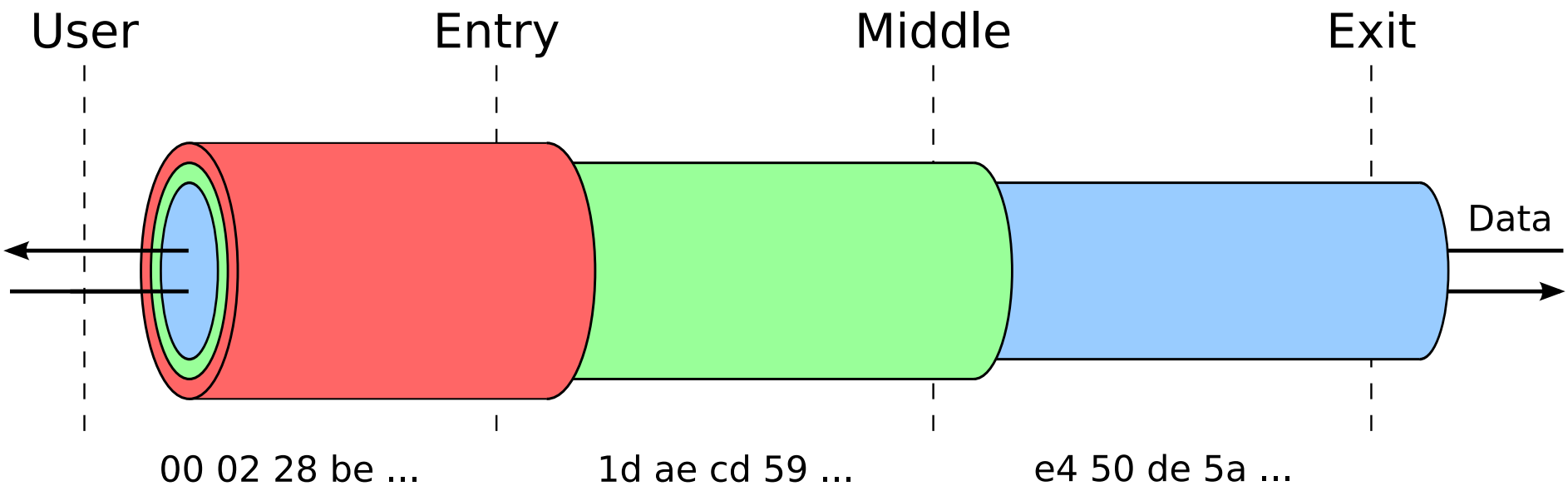
How does Tor work?



How does Tor work?



How does Tor work?



Attacks on Tor

- First hop can learn your IP address.
- Last hop can watch your traffic.

Attacks on Tor

- Adversary can block all Tor nodes by IP address and TCP port
  - our answer is to use non-public relays called Bridges
- Adversary can legally harass last hop; DMCA, Child Abuse Materials, Threats, etc
- Adversary can run relays, use network to restrict access to other relays

Attacks on Tor

- Deep Packet Inspection

Attacks on Tor

```
|0.541    |          https > 50936 [SYN,
|         |(50936)  <------------------  (443)
|0.541    |          50936 > https [ACK]
|         |(50936)  ------------------>  (443)
|0.542    |          Client Hello
|         |(50936)  ------------------>  (443)
|1.030    |          https > 50936 [ACK]
|         |(50936)  <------------------  (443)
|1.033    |          Server Hello,
|         |(50936)  <------------------  (443)
|1.124    |          50936 > https [ACK]
|         |(50936)  ------------------>  (443)
|2.079    |          [TCP Previous segme
|         |(50936)  <------------------  (443)
|2.079    |          [TCP Dup ACK 12#1]
|         |(50936)  ------------------>  (443)
|5.563    |          [TCP Retransmission
|         |(50936)  <------------------  (443)
|5.563    |          50936 > https [ACK]
|         |(50936)  ------------------>  (443)
|6.008    |          [TCP Retransmission
|         |(50936)  <------------------  (443)
|6.008    |          50936 > https [ACK]
|         |(50936)  ------------------>  (443)
|16.025   |          Client Key Exchange
|         |(50936)  ------------------>  (443)
|17.533   |          [TCP Retransmission
|         |(50936)  ------------------>  (443)
|20.735   |          [TCP Retransmission
|         |(50936)  ------------------>  (443)
|21.127   |          [TCP Previous segme
|         |(50936)  <------------------  (443)
|26.447   |          50936 > https [FIN,
|         |(50936)  ------------------>  (443)
|26.743   |          Encrypted Alert
|         |(50936)  <------------------  (443)
|26.743   |          50936 > https [RST]
|         |(50936)  ------------------>  (443)
```

The Future: Usability

Who are our users? What do they understand about anonymity, Tor, and privacy online? Can we guide them to make smarter decisions? How do we educate them before they start?

The Future: Obfsproxy & Pluggable Transports

Obfuscating proxy for network traffic

The Future: Scaling

Why not 10,000 relays? Why not 1 million? 10 million? Need privacy-preserving Scalable Distributed Hash Table designs

The Future: IPv6

Basic support for IPv6 clients and relays works now. Need support for IPv6 destinations and pure IPv6 relays

The Future: UDP

Tor only transports TCP packets now. This limits usable applications
Need to support real-time video and audio chats over Tor.

Thanks!

Visit https://www.torproject.org for more information, links, and ideas.