Privacy & Anonymity

CMSC 23200, Spring 2025, Lecture 14

Grant Ho

University of Chicago, 05/08/2025
(Slides adapted from Peyrin Kao, Vern Paxson, and Zakir Durumeric)

Logistics

- Assignment 5 released either Fri / Sat (May 9 / 10)
  - Due Thursday, May 15 by 11:59pm

- Next week:
  - TA Office Hours as scheduled
  - Final Discussion section next week (May 14)
  - Instructor Office Hours on May 12: cancelled
  - Lecture #15 by David Cash

Outline

- Privacy vs. Anonymity vs. Confidentiality
- Proxies & VPNs
- Tor
  - Overview & Design
  - Tor: Attacks & Additional Defenses/Services
  - Tor in Practice

What is Privacy?

Many different definitions:

- Privacy is control over your own information. Freedom from intrusion into personal matters
- Privacy is a person's right or expectation to control the disclosure of his/her personal information, including activity metadata
- Privacy is the "right to be let alone" — Louis Brandeis

Violations of Privacy

Last class:

- How can web attackers (websites) violate privacy by tracking what sites & content you interact with?

Today:

- How can network attackers track what sites / who you're communicating with on the Internet?

Anonymity: Related Concept

Anonymity (“without a name”): Concealing your identity

- Anonymous communication: the identity of source & destination in communication are concealed
- Anonymity provides some forms of privacy (e.g., unlinkability: prevents attackers from knowing action/information = yours, etc.)

Anonymity is not confidentiality

- Confidentiality hides the contents of the communication
- Anonymity hides the identities of who is communicating with whom

Metadata & Anonymity

TLS only protects content… doesn't offer anonymity or complete privacy

Anonymity often requires protecting metadata:

- Who is visiting what websites? Who is sending messages to whom?
  - Gov't might not like that you're visiting Human Rights Watch website
  - Gov't might not be amused that you're sending messages to Human Rights Watch
- We may want to hide the existence of the message (maybe sending an encrypted message at all is going to cause you problems)

Achieving Anonymity is Difficult

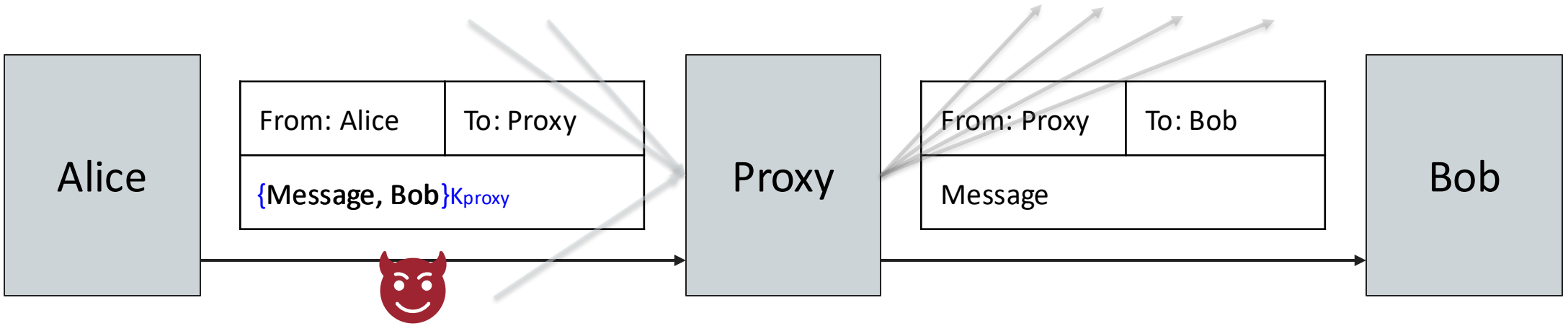
- Difficult, if not impossible, to achieve on your own
  - Source and destination IP address visible in every packet
- Anonymity is easier for attackers
  - An attacker can hack into someone else's computer and/or often spoof messages from fake source addresses!
  - Benign users don't usually do these things
- Main strategy for anonymity: Ask someone else to send messages for you

Outline

- Privacy vs. Anonymity vs. Confidentiality
- Proxies & VPNs
- Tor
  - Overview & Design
  - Tor: Attacks & Additional Defenses/Services
  - Tor in Practice

Proxies

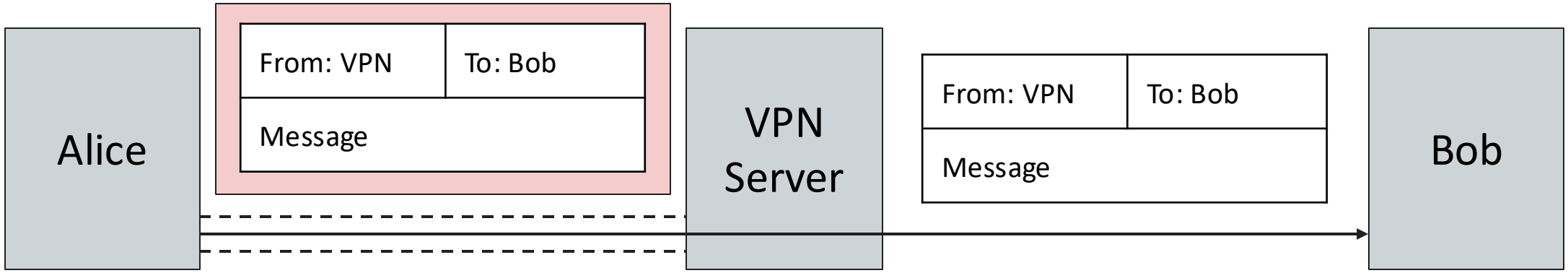
- **Goal & Threat Model:** Alice wants to anonymously send a message to Bob
  - Bob shouldn't know the message is from Alice
  - An eavesdropper (Eve) cannot deduce that Alice is talking to Bob
- **Proxy**: A third party that relays Internet traffic
  - Alice sends the message and the recipient (Bob) to the proxy, and the proxy forwards the message to Bob (along with many other src + dest pairs)
    - The recipient's name (and optionally the message) is encrypted, so an eavesdropper looking at packets can't see both msg src & dest
  - Bob receives the message from the proxy, with no indication it came from Alice



Virtual Private Networks (VPNs)

VPNs: A virtual connection to an internal network

- Creates encrypted “tunnel” to VPN server at the Network / IP layer
- Traffic from client first encrypted & sent to VPN server
- VPN server then decrypts & forwards traffic to final destination
- VPNs act as a proxy into internal network: outbound traffic appears to come from internal network and not Alice



Naive anonymity approach .... VPNs

Naive approach .... VPNs

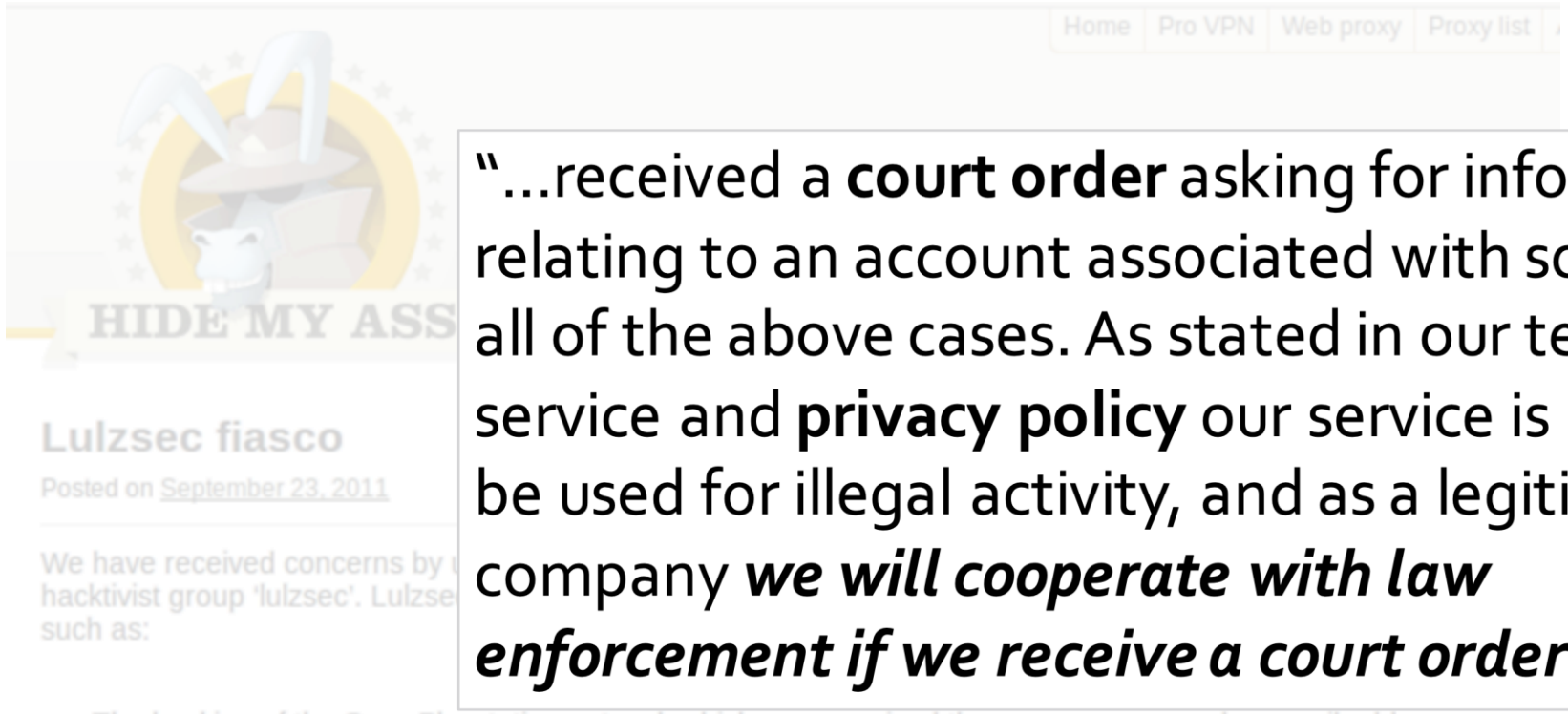
Lulzsec fiasco

Posted on September 23, 2011

We have received concerns by users that our VPN service was utilized by a member or members of the hacktivist group 'lulzsec'. Lulzsec have been ALLEGEDLY been responsible for a number of high profile cases such as:

- The hacking of the Sony Playstation network which compromised the names, passwords, e-mail addresses, home addresses and dates of birth of thousands of people.
- The DDOS attack which knocked the British governments SOCA (Serious Organised Crime Agency) and other government websites offline.
- The release of various sensitive and confidential information from companies such as AT&T, Viacom, Disney, EMI, NBC Universal, and AOL.
- Gaining access to NATO servers and releasing documents regarding the communication and information services (CIS) in Kosovo.
- The defacement of British newspaper websites The Sun & The Times.
- The hacking of 77 law enforcement sheriff websites.

Naive approach .... VPNs



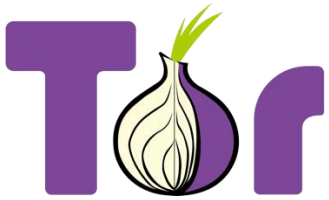
"...received a **court order** asking for information relating to an account associated with some or all of the above cases. As stated in our terms of service and **privacy policy** our service is not to be used for illegal activity, and as a legitimate company ***we will cooperate with law enforcement if we receive a court order***"

Proxies and VPNs: Issues

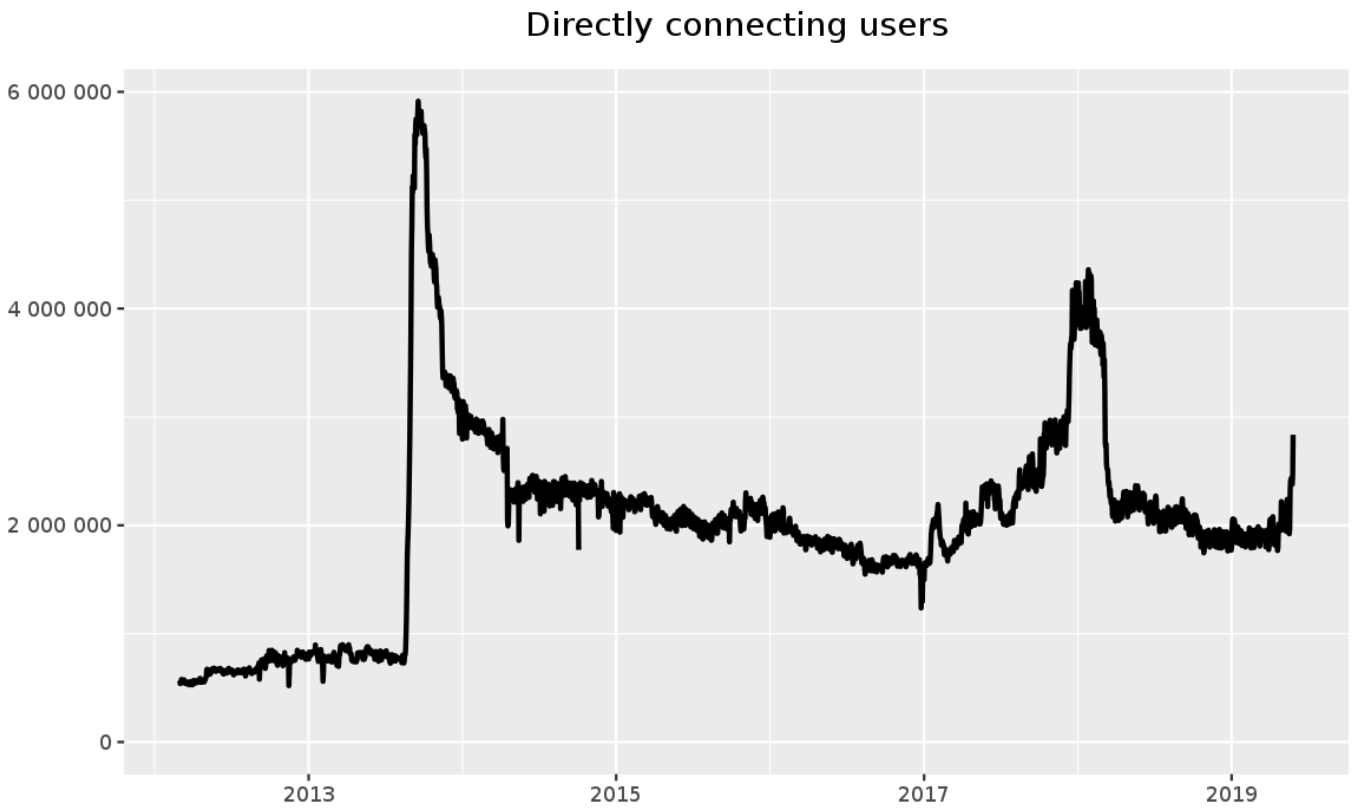
- Trusting the proxy
  - The proxy can see the sender and recipient's identities
  - Attackers might convince the proxy to tell them about your identity – or the proxy itself could be an attacker!
- Performance
  - Sending a packet requires additional hops across the network
- Cost
  - VPNs can cost $80 to $200 per year

Outline

- Privacy vs. Anonymity vs. Confidentiality
- Proxies & VPNs
- Tor
  - Overview & Design
  - Tor: Attacks & Additional Defenses/Services
  - Tor in Practice



- Tor is a successful privacy enhancing technology that works at the transport layer
- Millions of active users.
- Provides anonymous TCP connections (conceals your and/or destination IP address)



Tor ("The Onion Router")

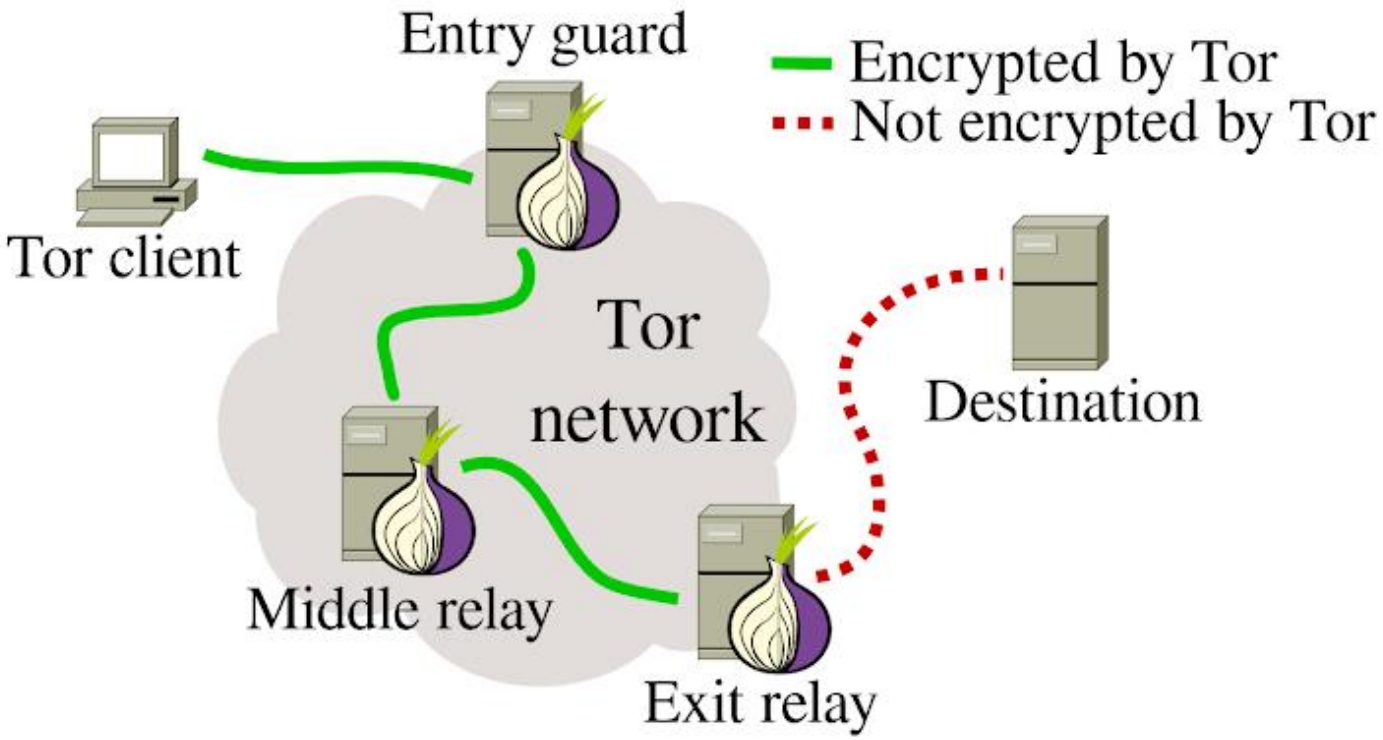
Idea: Send the packet through multiple proxies instead of just one proxy

- **Tor:** A network that uses cryptography + multiple proxies (relays or "onion routers") to enable anonymous communications

Key components of Tor:

- Network of many **Tor relays (proxies)** for forwarding packets
- **Directory server:** Lists all Tor relays and their public keys
- **Tor Browser:** A web browser configured to connect to the Tor network
- **Tor onion services:** Servers that can only be reached through the Tor network
- **Tor bridges:** relays that try to hide the fact that a user is connecting to the Tor network

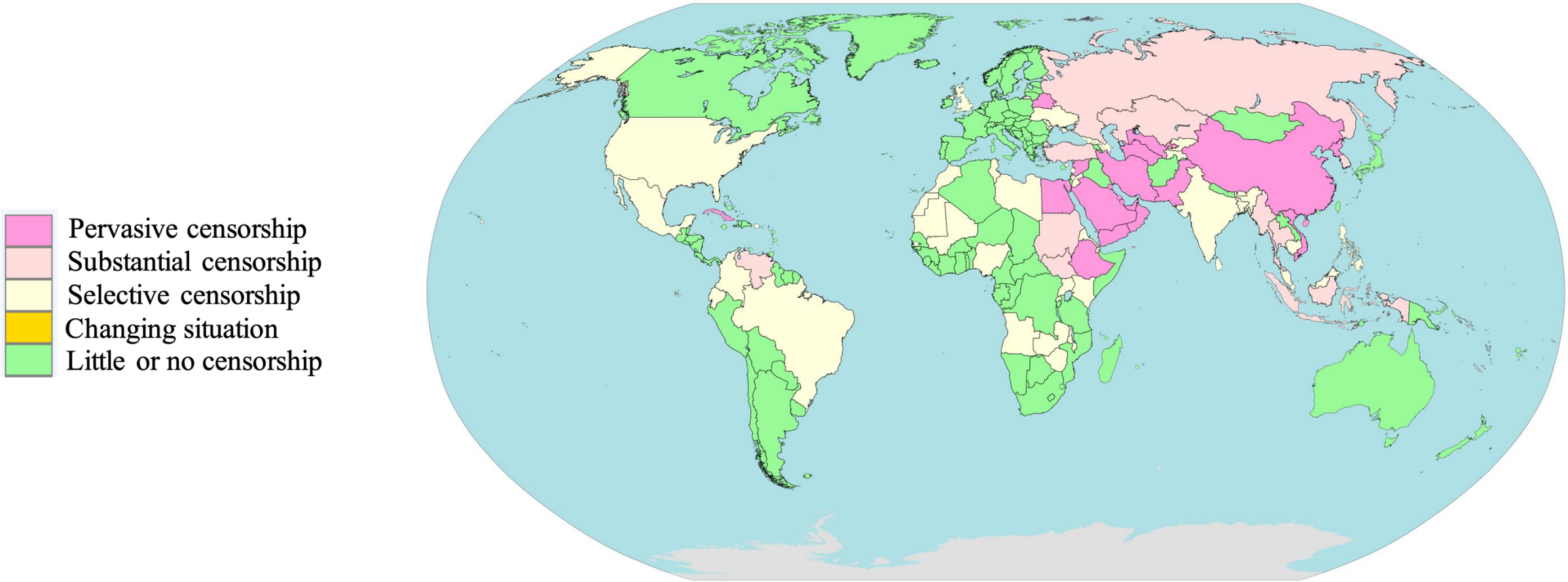
Tor (“The Onion Router”)



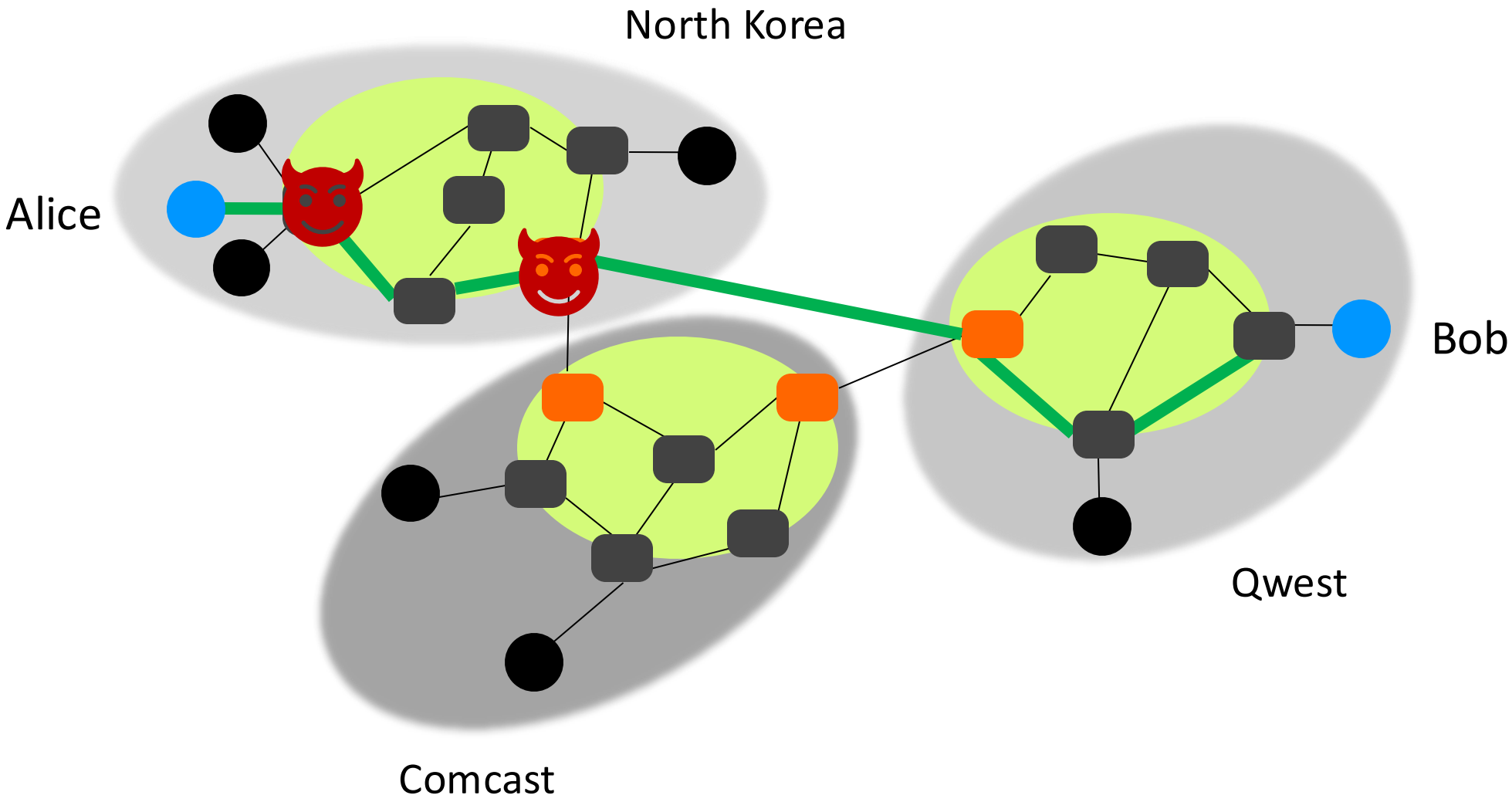
Tor Threat Model & Goals

- **Security:** Client anonymity and censorship resistance
  - Optional: Server anonymity with onion (hidden) services
- Tor preserves anonymity against local adversaries
  - Example: An on-path attacker sees Alice send a message to a Tor relay, but not the final destination of the message
  - Example: The server should not know the identity of the client based solely on network layer info
- **Performance:** Low latency (communication should be fast)

Internet Censorship



Example Censorship Threat Model

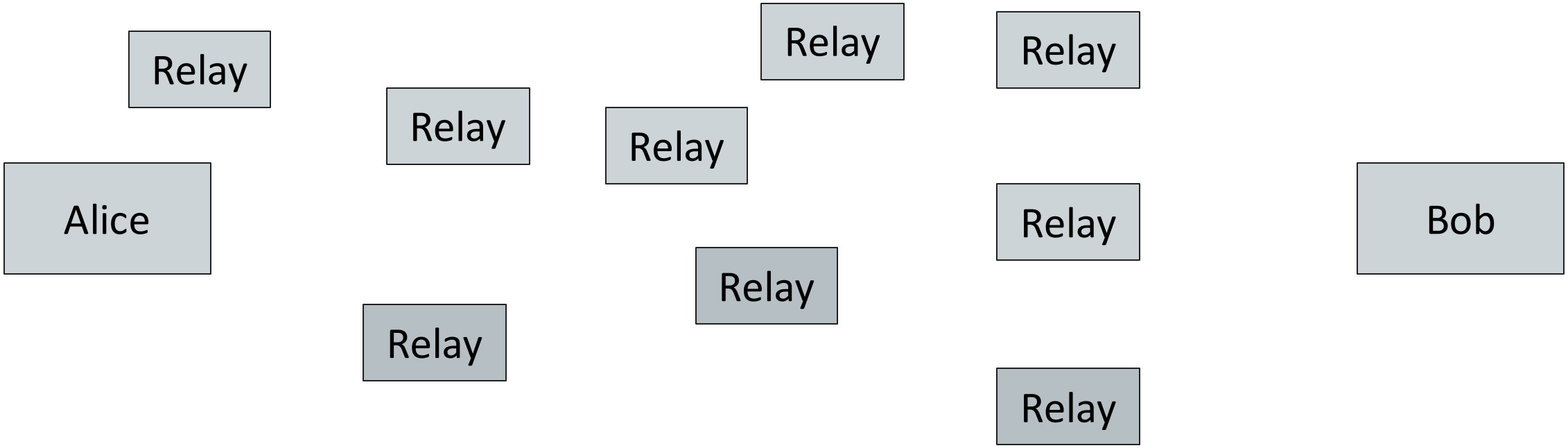


Tor Protocol: Tor Circuits

To communicate anonymously with a server, the Tor client forms a **circuit** consisting of 3 relays (by default)

1. Query the directory server for a list of relays (lists all relays & their PK)
2. Choose 3 relays to form a Tor circuit
3. Connect to the 1st relay, forming an end-to-end TLS connection
4. Connect to the 2nd relay *through* the 1st relay, using end-to-end TLS connection
5. Connect to the 3rd relay *through* the 2nd relay, using end-to-end TLS connection
6. Connect to the web server through 3rd relay using HTTPS (so an end-to-end TLS connection is formed through the third relay)

Tor Circuits: Walkthrough

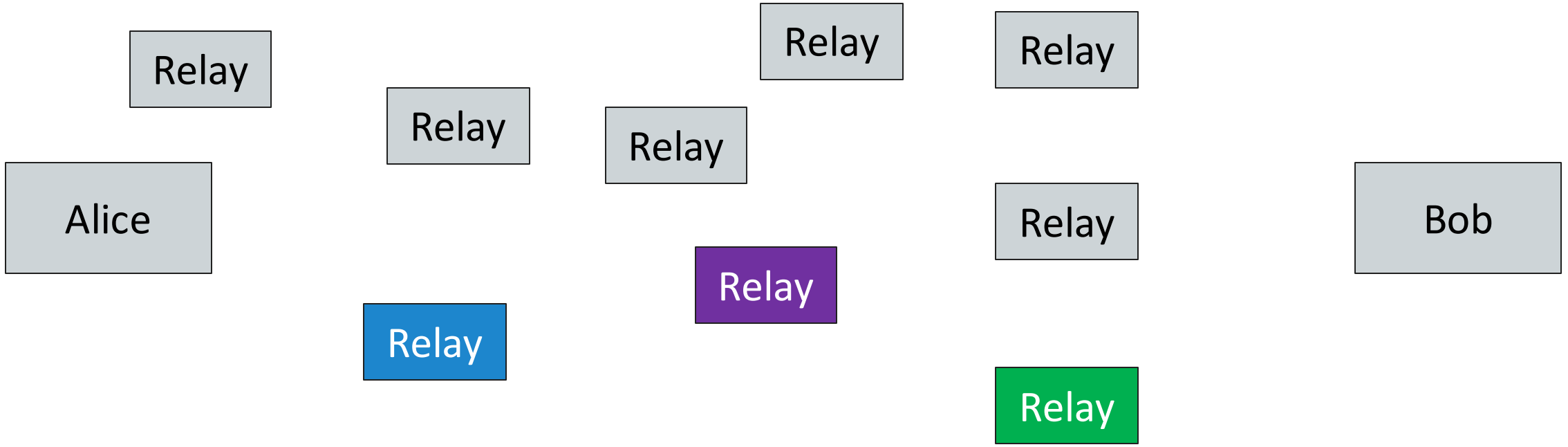


Tor Protocol: What a Relay Does

Runs a Tor relay application (software) that:

1. Listens for someone to initiate a TLS connection
2. When receiving a packet, decrypts using the key obtained through TLS (or encrypts if reverse direction)
3. Forwards the packet to its next hop / destination

Tor Circuits: Walkthrough

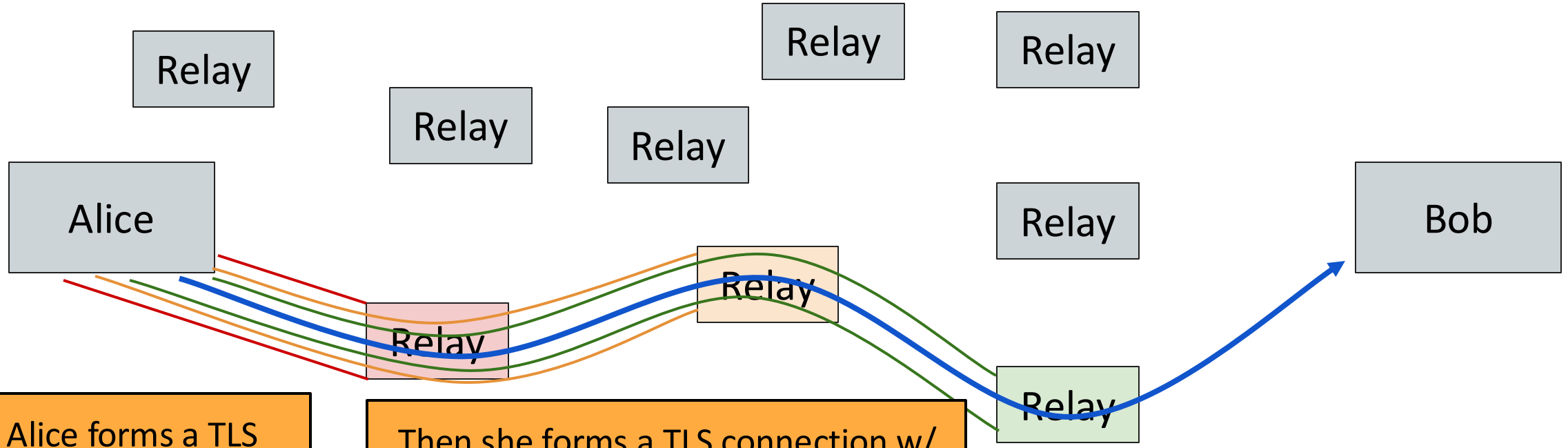


Suppose Alice wants to talk to Bob anonymously.

Alice queries Tor's directory servers and chooses 3 relays: Relay #1 (Entry Node), Relay #2, and Relay #3 (Exit Node)

The directory servers publish a public key for each Relay node

Tor Circuits Setup



Alice forms a TLS connection w/ the entry node (Relay 1)

Then she forms a TLS connection w/ the 2nd node, through the 1st node

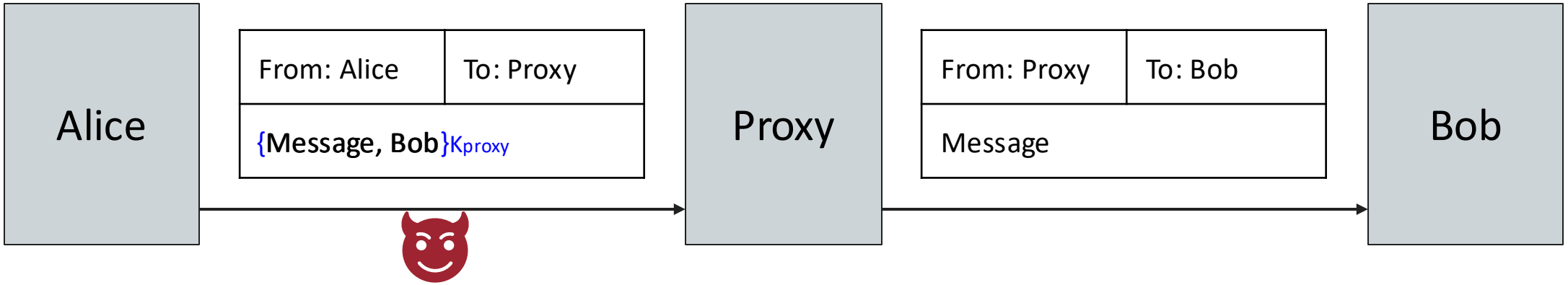
Note: Relay 1 is only relaying TLS packets. It doesn't know the contents of the packets!

Then she forms a TLS connection with the exit node (Relay #3), through the 2nd node

Finally, she opens a TLS connection with Bob via the exit node

Recall: Proxy Message Encryption

- Alice wants to send a message to Bob
- She **encrypts** the recipient's name (and message) so an eavesdropper does not see a packet with both Alice and Bob's identities in plaintext

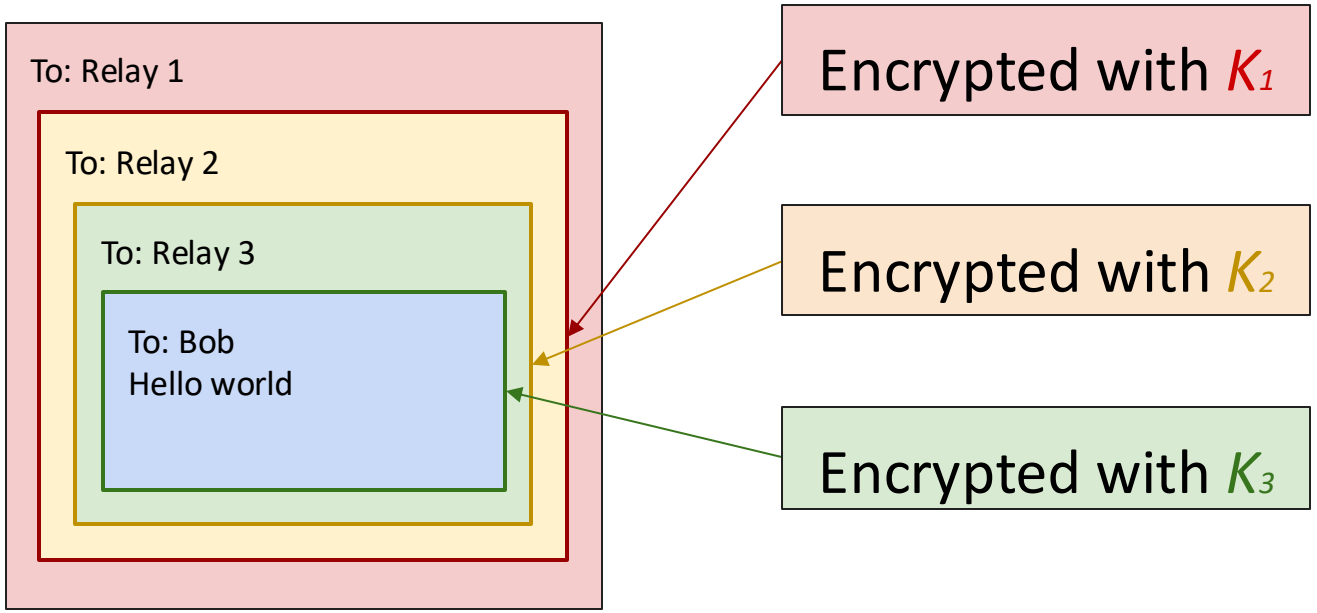


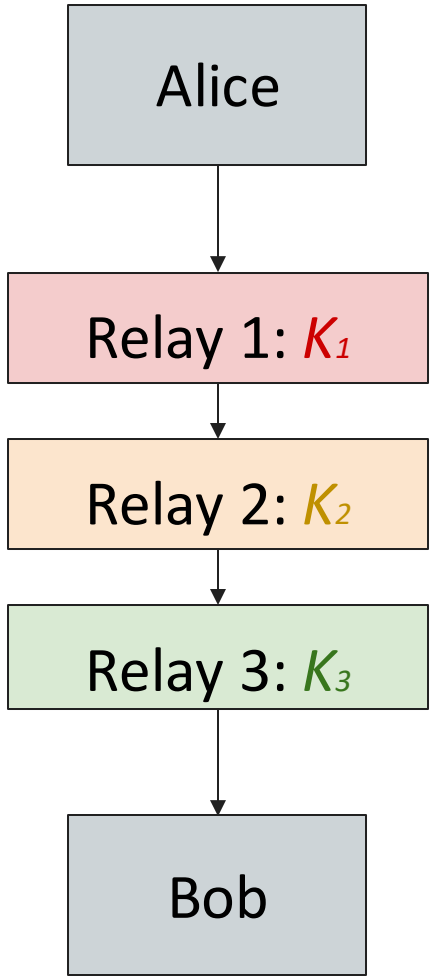
Tor Packet Construction

Wrap the packets via encryption: fixed size "cells" of 512 bytes

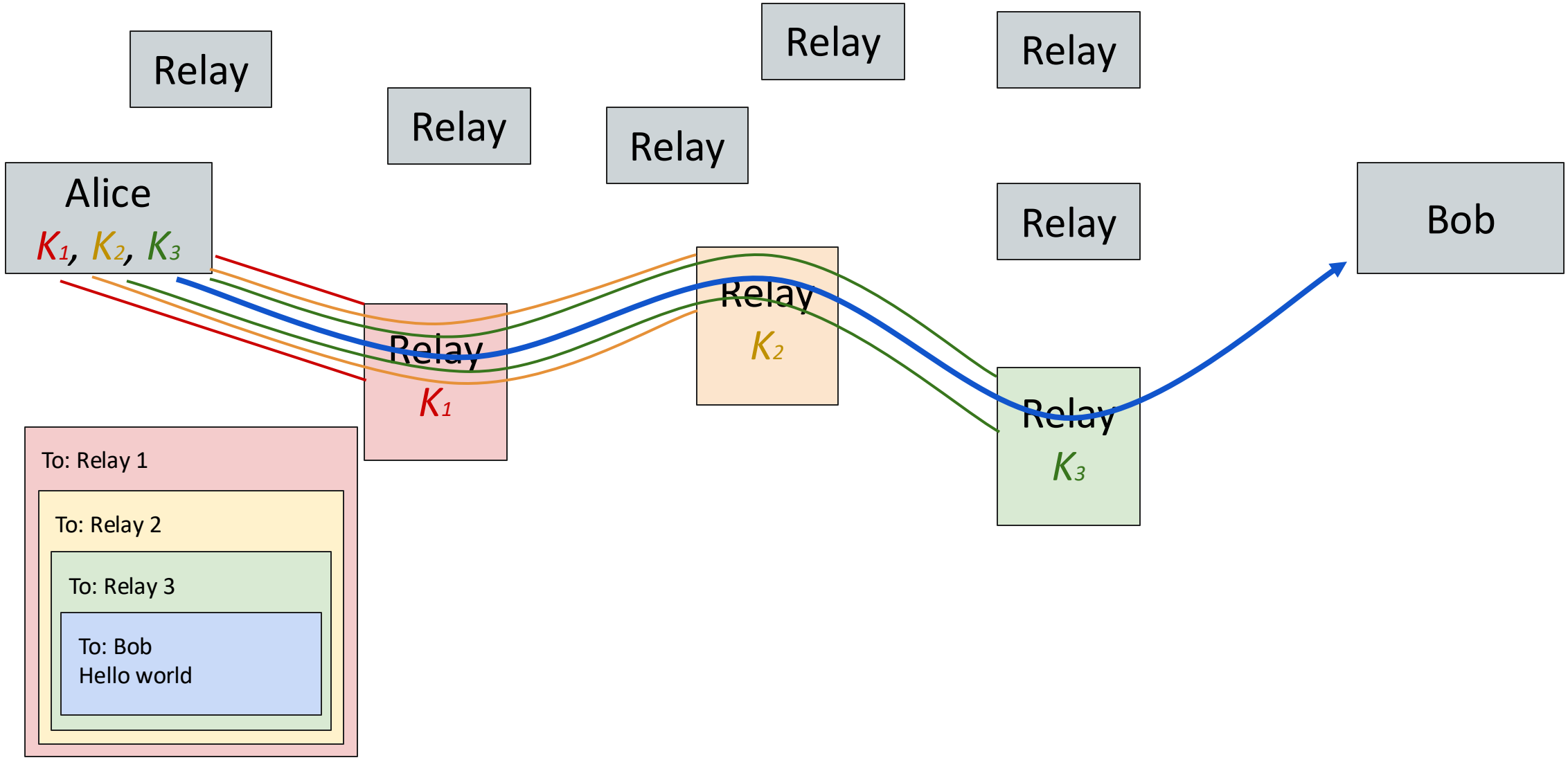
- e.g., the packet sent to Bob is encrypted under K_3 since Relay 3 is the one to forward that information to Bob

Ensures that no one can read or tamper with the messages, since these are all sent over TLS connections



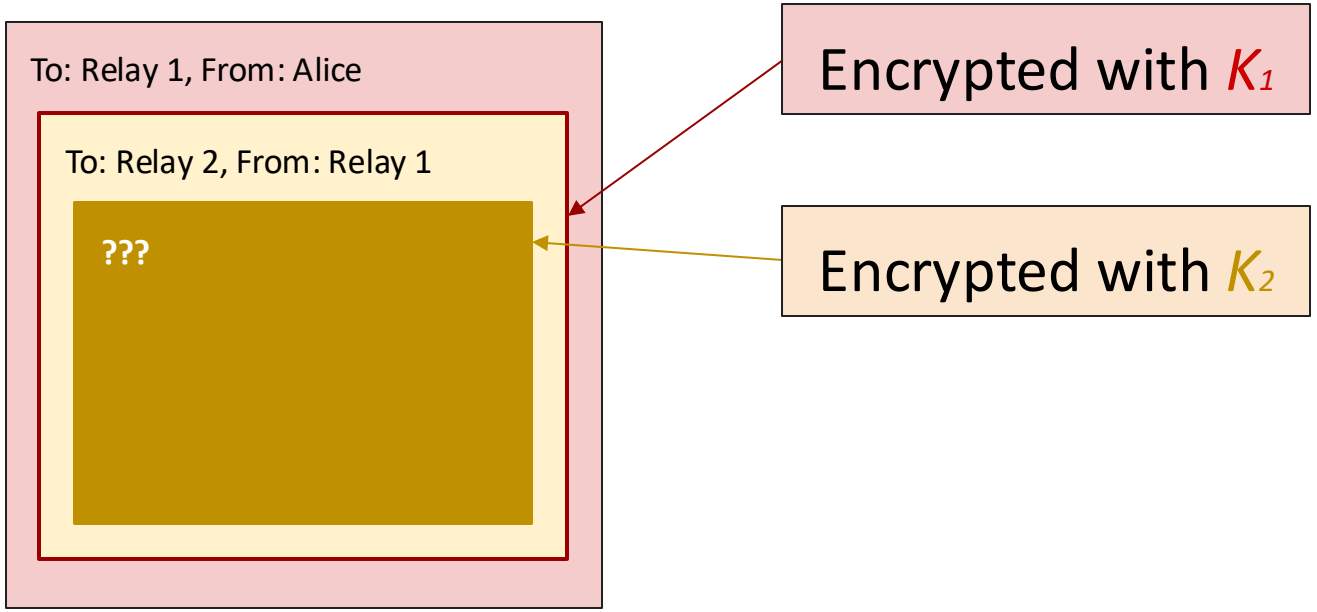


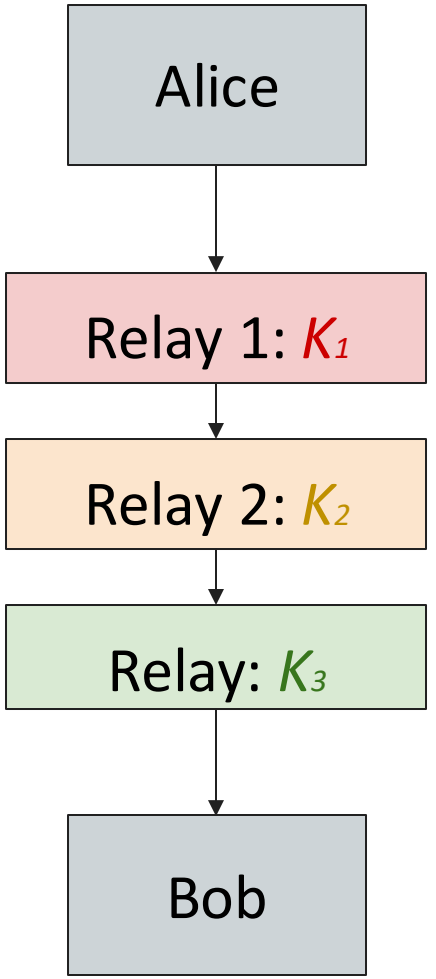
Tor Circuits: Walkthrough



Tor Packet Construction

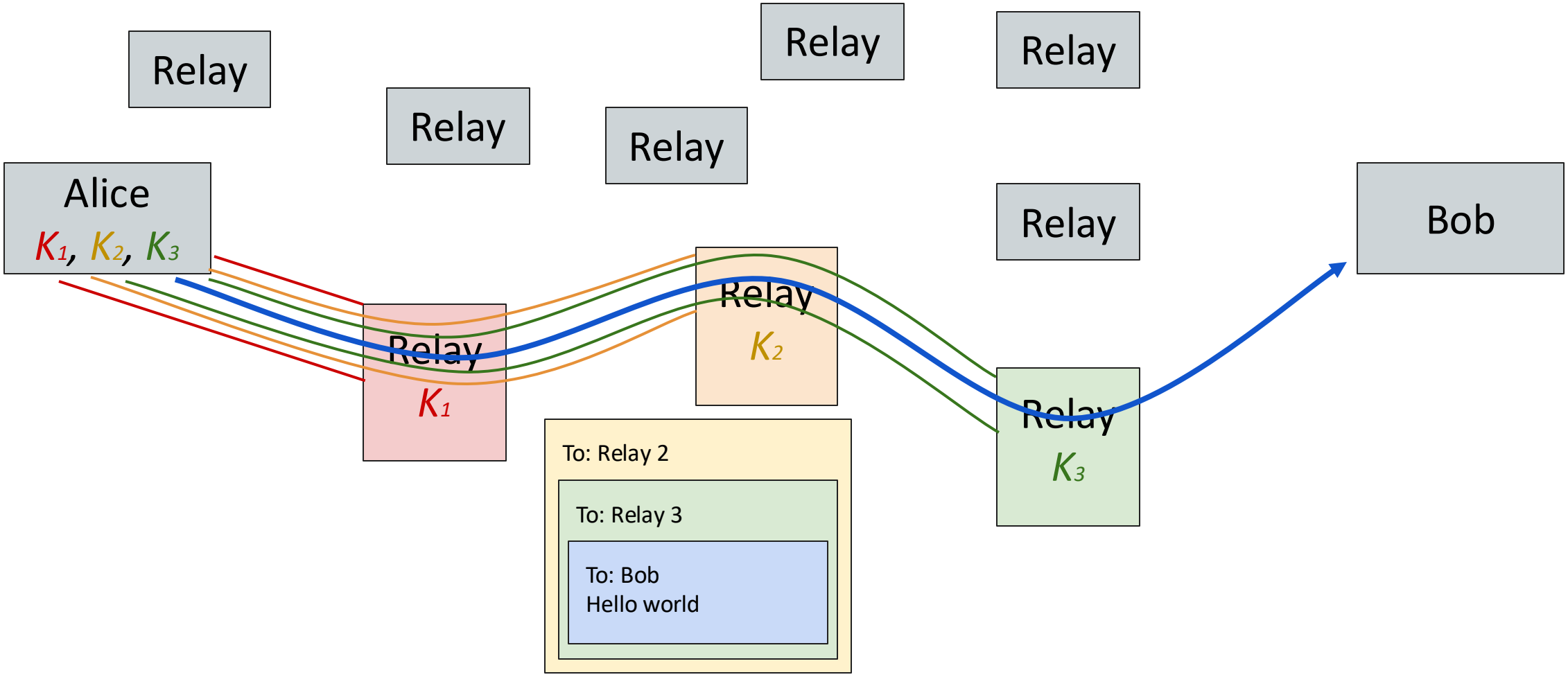
- What does Relay 1 see?





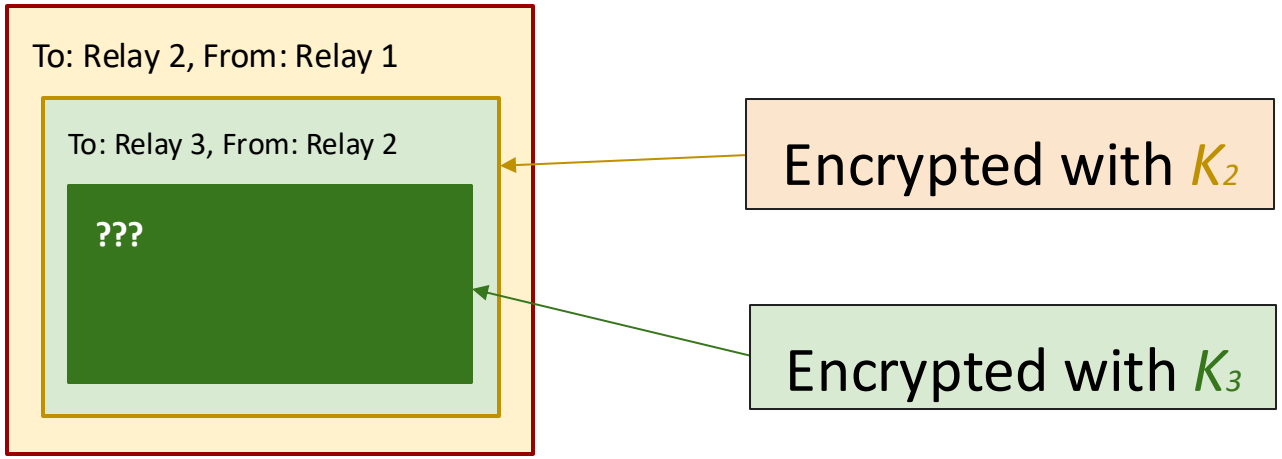
All Relay 1 knows is the message came from Alice and is going to Relay 2. They don't know Alice is talking to Bob!

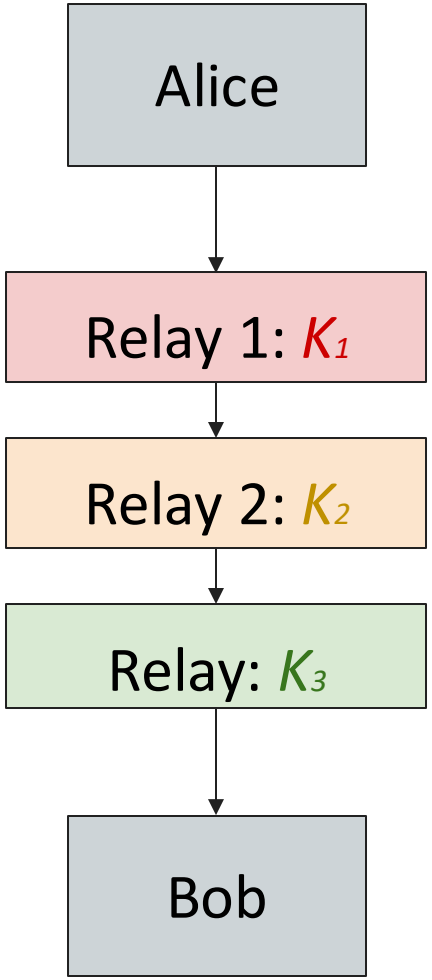
Tor Circuits: Walkthrough



Tor Packet Construction

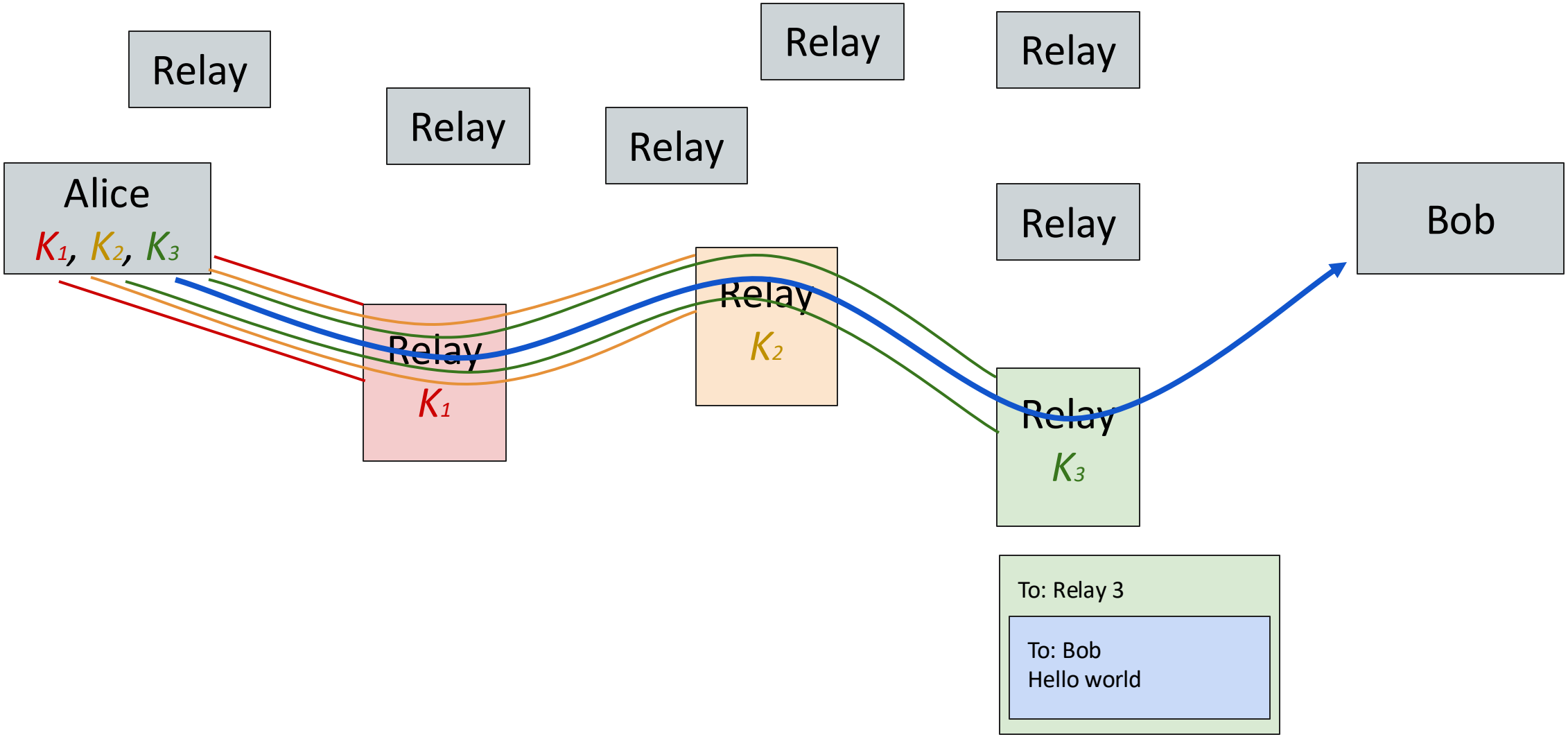
- What does Relay 2 see?





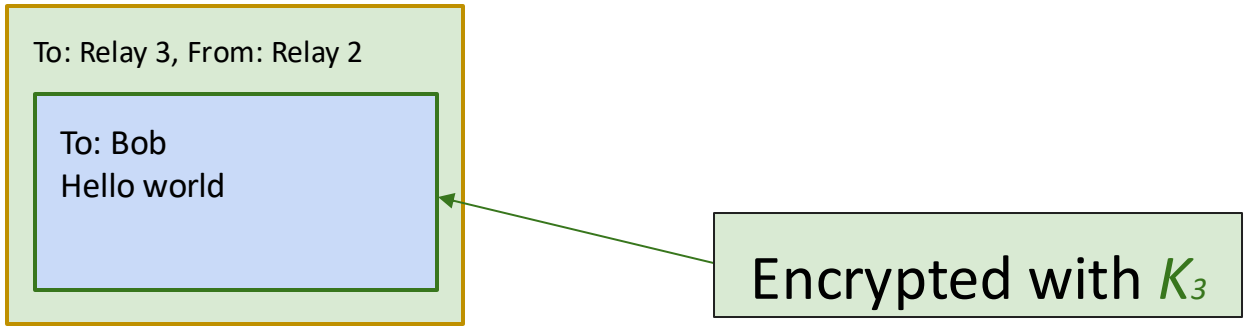
All Relay 2 knows is the message came from Relay 1 and is going to Relay 3. They know nothing about Alice and Bob!

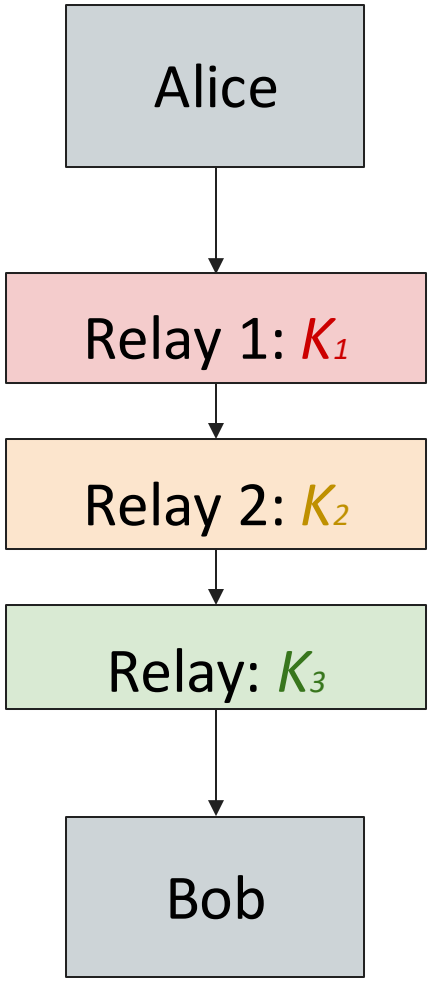
Tor Circuits: Walkthrough



Tor Packet Construction

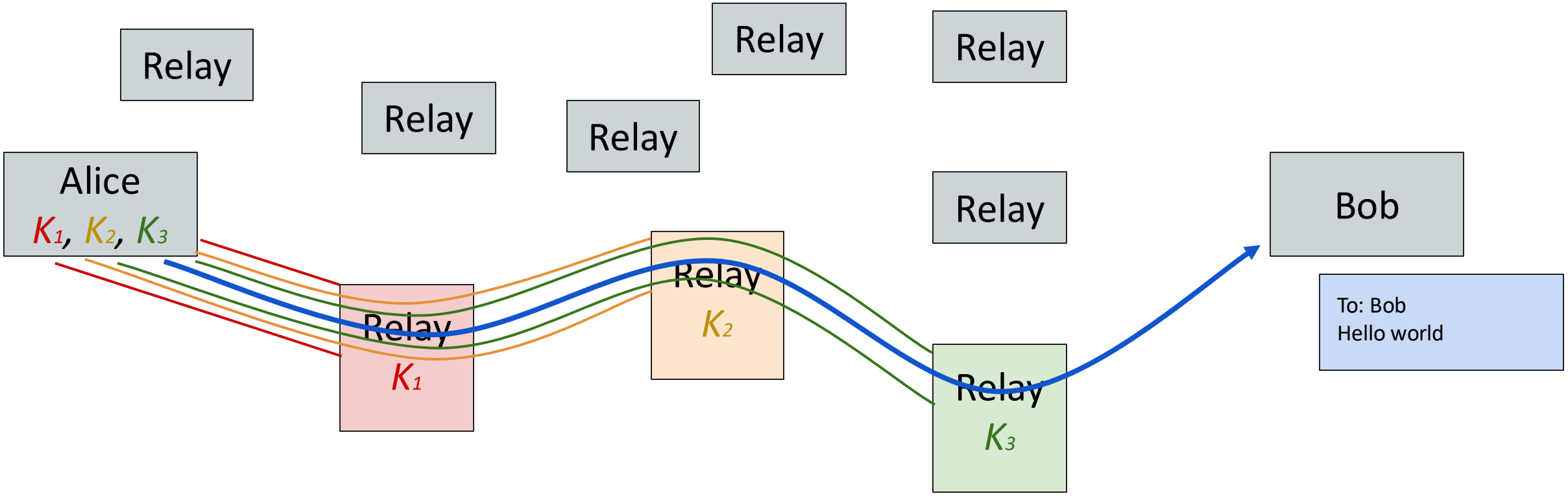
- What does Relay 3 see?



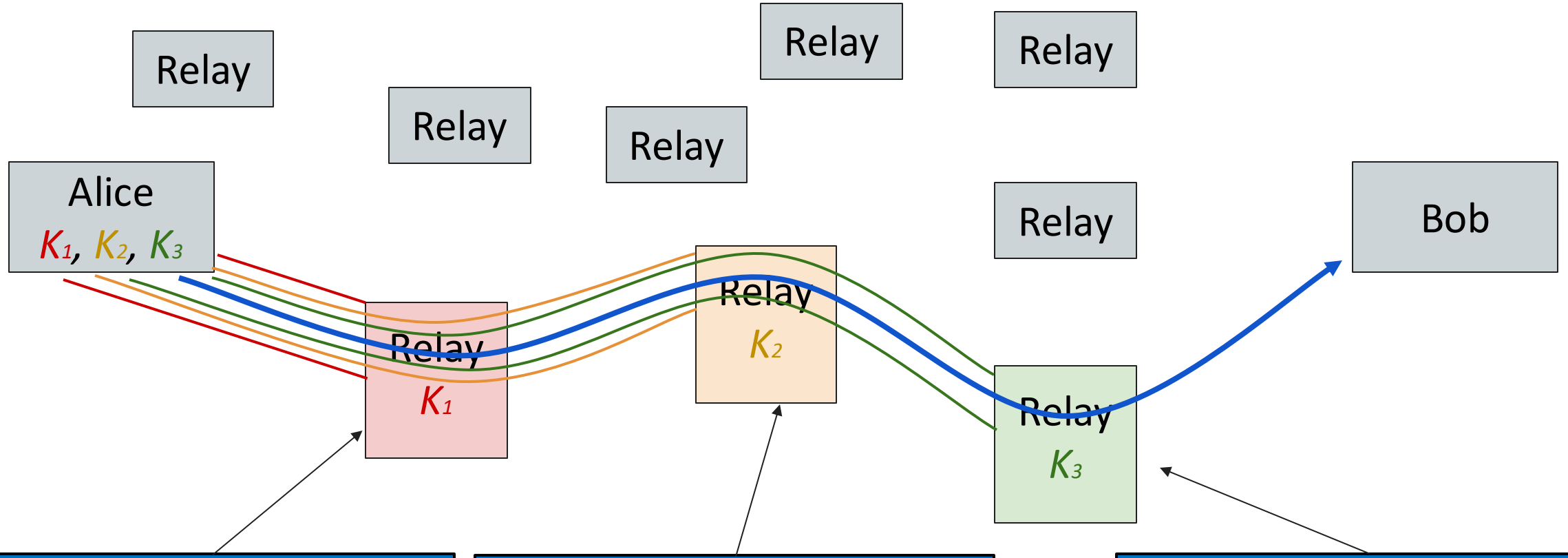


All Relay 3 knows is the message came from Relay 2 and is going to Bob. They don't know Alice sent the message!

Tor Circuits: Walkthrough



Tor Circuits: Walkthrough



Relay 1 knows that Alice is using Tor but *not* who Alice is talking to

Relay 2 knows nothing (other than that someone is using Tor)

Relay 3 knows that someone is talking to Bob, but not who

Tor Exit Nodes

- The exit node can see the message and the recipient (but not the sender)

- The exit node is a man-in-the-middle attacker
  - If the user is not using encryption (TLS) to connect to the end host, the exit node can see and modify the traffic
  - If the user is using TLS (using HTTPS), the exit node cannot see or tamper with the traffic

Tor Exit Nodes in Practice

- Administrators of Tor exit nodes often receive abuse complaints
  - Users complain to the exit node
  - Users complain to the Internet service provider (ISP), which complains to the exit node
  - Legal problems: illegal activity traced to exit node first
- As a result, most Tor relays choose to only be entry or intermediate nodes, not exit nodes
  - Exit node bandwidth is the bottleneck in Tor, not internal bandwidth

Outline

- Privacy vs. Anonymity vs. Confidentiality
- Proxies & VPNs
- Tor
  - Overview & Design
  - Tor: Attacks & Additional Defenses/Services
  - Tor in Practice

Tor Weaknesses: Timing Attacks (Side Channel)

A network attacker who has **a full (global) view** of the network can learn that Alice and Bob are talking

- Timing attack: Observe when Alice sends a message, when Bob receives a message, and link the two together

Global adversaries are *outside* of Tor's threat model and are not defended against

- Tor only defends against local adversaries with partial views of the network
- Timing attacks could be defended against by delaying the timing of packets, but would lead to poor/unusable performance

Tor Weaknesses: Collusion

- **Collusion**: Multiple nodes working together & sharing info
  - Collusion is adversarial (dishonest) behavior
  - If *all* nodes in the circuit collude, anonymity is broken
  - If *at least one* node in the circuit is honest, anonymity is preserved
  - An attacker can create hundreds of nodes in the Tor network to increase the chance that your circuit consists entirely of the attacker's nodes!

- Defense: The more nodes we use, less likely they are not all colluding
  - It's much harder for 10 nodes to collude than for 2 nodes to collude
  - 3 nodes is generally considered good enough & is the default

Tor Weaknesses: Collusion Defense

Defense: **Guard nodes**

- Guard nodes must have a high reputation and must have existed for a long time
- Clients will always use a guard node as the entry node (by default) & the same guard node is used for a long period of time
  - Attackers' nodes are unlikely to become guard nodes
  - Because clients use the same guard nodes for a long period of time, there is only a low chance that the client will switch to an attacker's guard node

Tor Weaknesses: Distinguishable Traffic

- Tor does *not* hide the fact that you are using Tor
  - Example: A local adversary can see that you are sending packets to a Tor relay
  - Tor directory publishes all relay nodes for any client

- Anonymity only works in a crowd
  - Example: A Harvard student sent an anonymous bomb threat using Tor. The administrators noticed that only one student on the Harvard network used Tor at that time!

Tor Weaknesses: Distinguishable Traffic

Defense: **Tor bridges**

- Attackers can tell you are using Tor because they can see you are connecting to an entry node
- **Tor bridges**: entry nodes that are not available on public lists
  - Users request bridges from a separate directory, which only gives a few bridges to a user
  - Prevents attackers from enumerating all bridges unless they have many different IP addresses running Tor clients

Tor Weaknesses: Distinguishable Traffic

With Tor bridges, censors can no longer block Tor based on IP addresses of entry/relay nodes

- But they can still distinguish traffic that looks like Tor traffic from normal traffic (fixed sized packets with TLS)

Defense: **Pluggable transports**

- Pluggable transports change the appearance of the client's traffic to the entry node (only for bridges)
- Obfuscates the encrypted traffic to make it "look" more like normal Internet traffic (no longer obvious fixed size packets)

Tor Hidden (Onion) Services

- Sometimes, the *server* wants to be anonymous, so no one knows where the server is located
- **Tor onion services**: Websites that are only accessible through the Tor network
  - Gives the server anonymity protection
  - Sometimes called the **dark web**
- **Idea:** Route the server's traffic through the Tor network so that no one knows who the server is

Tor Onion Services

- Connecting to onion services is a little more complicated:
  - Client has to know where to send packets, but server is trying to be anonymous
- Solution: rendezvous point – a relay node that will connect two circuits from different directions
  - Client connects to rendezvous point over a Tor circuit
  - Server connects to rendezvous point over a Tor circuit
  - Rendezvous point relays packets between these two circuits
  - Security: rendezvous point does not learn the identity of the client or of the server, so can't reveal either identity to the other

Outline

- Privacy vs. Anonymity vs. Confidentiality
- Proxies & VPNs
- Tor
  - Overview & Design
  - Tor: Attacks & Additional Defenses/Services
  - Tor in Practice

Tor Tradeoffs

Benefit: Free to use

- Tor is mostly funded by the US government
- Users “pay” by providing traffic for other users to hide in (recall: you don’t want to be the only user on the network using Tor)

Drawback: Performance

- Latency is significantly worse: Packets need to make more hops across the network

Drawback: Full anonymity requires usability tradeoffs

- No cookies by default (even final server doesn’t know you)
- They even recommend keeping the browser window size constant, which can be annoying!

Internet Censorship & Tor

- Government censors
  - Block websites containing “offensive” content
  - Commonly employ blocklist approach
- Because Tor hides the sites a user is connecting to, it is useful & popular for bypassing censorship
  - Functions similarly to bypassing censorship using a VPN or proxy
- **Problem:** Constant arms race b/t Tor & censors

Arms Race: Tor vs. Censorship

- Censors can easily block access to all public Tor entry points
  - Bridge services provide a set of entry points that aren't listed publicly anywhere, so they can't be blocked by IP
- Censors can block traffic that looks like Tor traffic
  - Pluggable transports make traffic look more like normal web traffic
- Censors can pretend to be a Tor client to see if a host is a Tor entry/bridge node & then block connections to it
  - Some pluggable transports use cloud services (like Google Cloud Platform, Amazon Web Services, etc.): harder to block

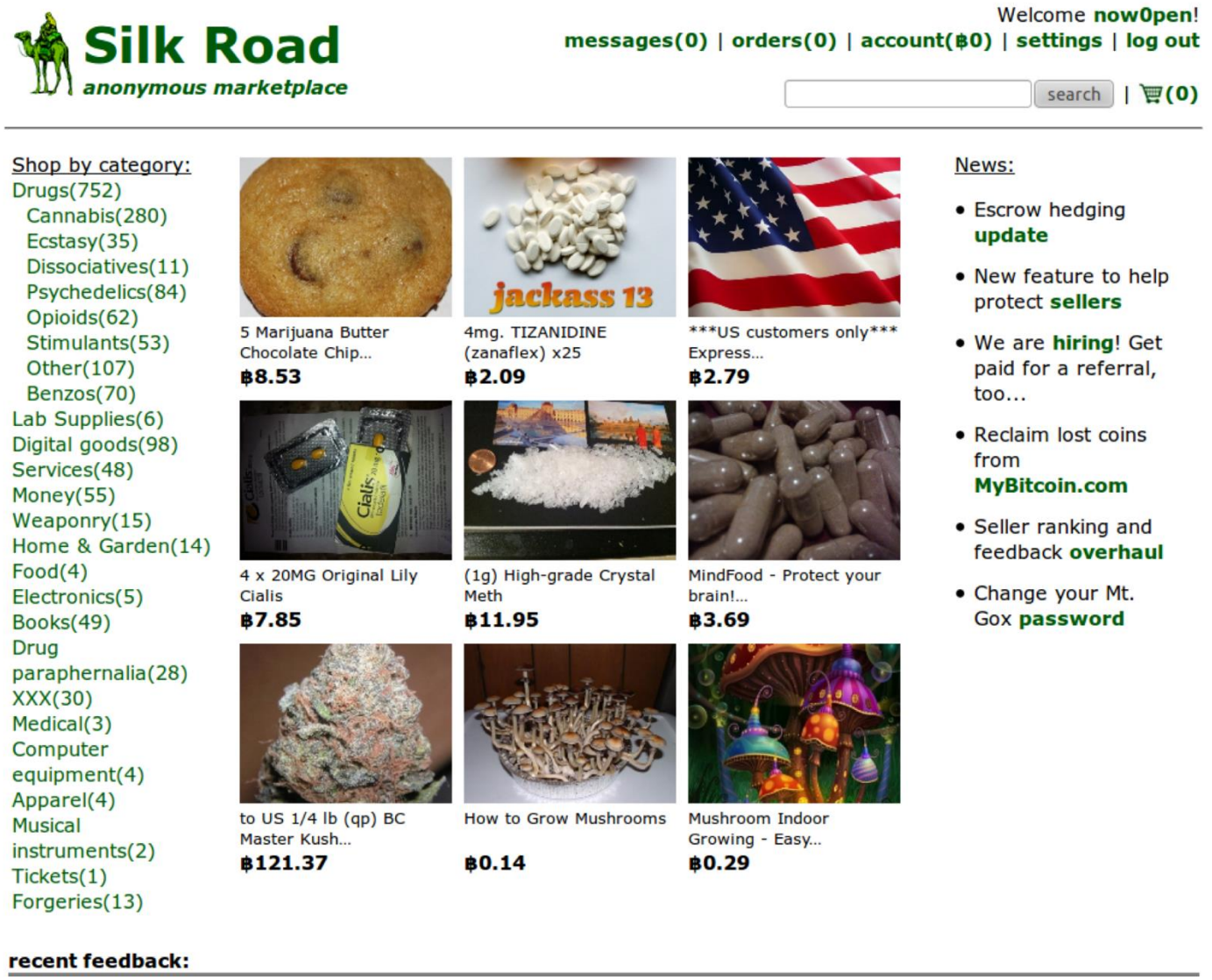
Hosting Illegal Services on Tor

Tor onion service… …und the world

- – Legitimate … …ervices
- – Most coun… …ed on regular web

Dark markets: M…

- – Transactio… …Bitcoin
  - • Service…
- – Ratings sys…
- – Escrow ser…
- – Can only b…



Modern Dark Markets

Hard to find information about where dark markets are located

- Legitimate websites (e.g., Reddit) will remove dark market links
- Legitimate websites with information about dark markets (e.g., DeepDotWeb) get taken down
- Information about dark markets is usually available through Tor onion services (e.g., Dread, a Reddit clone)

Summary: Anonymity & Tor

- Anonymity (concealing one's identity) can be difficult to achieve on the web
  - Different from standard confidentiality
- Proxies and VPNs relay traffic through a single machine: weak anonymity
  - The proxy knows who you are and what you are doing: not anonymous!
- Tor encrypts & routes your traffic through multiple machines
  - Circuits are established by performing TLS handshakes with three nodes, nested onion of encryption (no one knows full end-to-end)

Summary: Anonymity & Tor

Tor does have a few weaknesses

- Weakness: Timing attacks + global adversaries (not defended)
- Weakness: Collusion between nodes can deanonymize users by working together
  - Defense: Guard relays & multiple relays in circuits
- Weakness: Tor traffic is distinguishable from normal traffic, allowing it to be censored and blocked
  - Defense: Bridges and pluggable transports
- Worse performance & Tor itself/usage sometimes has poor reputation

Summary: Anonymity & Tor

Onion services provide anonymity for the server, in addition to the client

Tor in practice

- Often used to evade censorship -- Tor and censors are in a constant arms race
- Illegal services often use Tor because it conceals their identity from authorities