CS 5321 Network Security
Week 10: Anonymous communications

Daisuke MASHIMA

http://www.mashima.us/daisuke/index.html

2022/23 Sem 2

(some slides from Lujo Bauer, Bryan Ford)

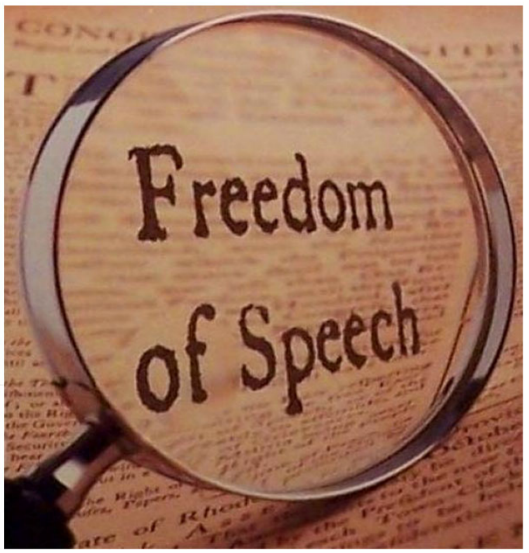
What Is Anonymity?

- Anonymity is the state of being not identifiable within a **set of subjects**
  - You cannot be anonymous by yourself!
  - Hide your activities among others' similar activities

- **Unlinkability** of action and identity
  - For example, sender and his email are no more related after observing communication

Anonymity… why?

- Do we need anonymity on the Internet?



- Accountability?

Applications of Anonymous Communication

- Privacy-preserving web browsing
  - Hide online transactions, Web browsing, etc., from intrusive governments, marketers, etc.
- Untraceable electronic mail
  - Corporate whistle-blowers
  - Political dissidents
  - Socially sensitive communications
- Digital cash
  - Electronic currency with properties of paper money (online purchases unlinkable to buyer's identity)
- Anonymous electronic voting / online survey
- Censorship-resistant publishing

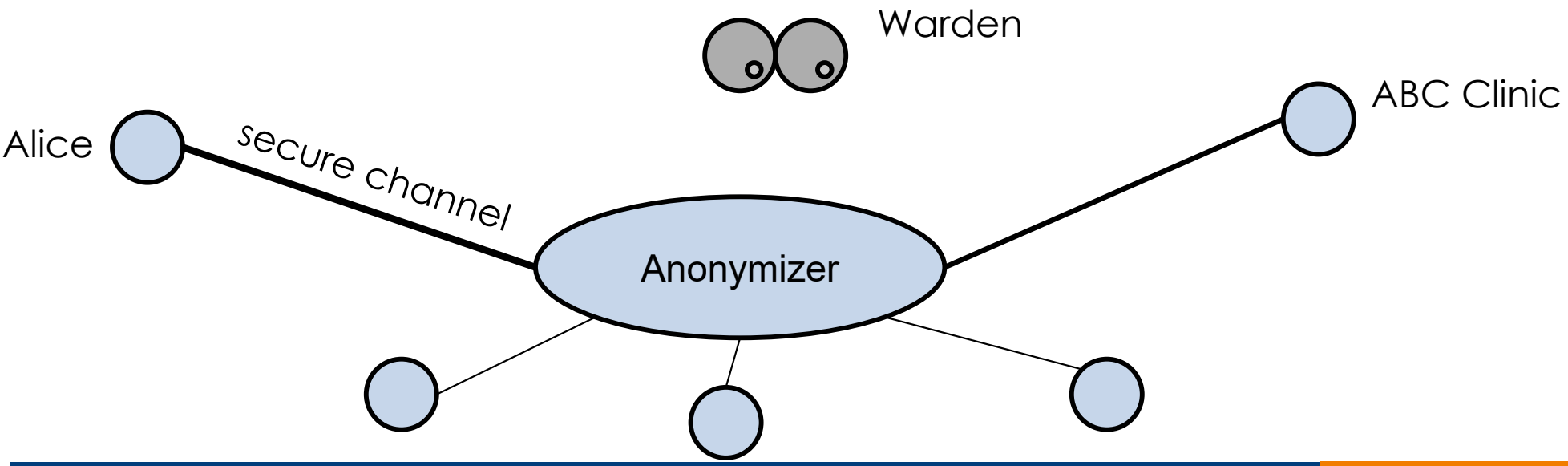
Anonymous Communication on the Internet?

- Internet is designed as a public network
  - Machines on your LAN may see your traffic, network routers see all traffic that passes through them
- Routing information is public
  - IP packet headers identify source and destination
  - Even a passive observer can easily figure out **who is talking to whom**
- Encryption and secure communication do not hide identities
  - Encryption hides payload, but not routing information

Anonymizer (Anonymous Proxy)

- Probably the easiest way to achieve anonymity.

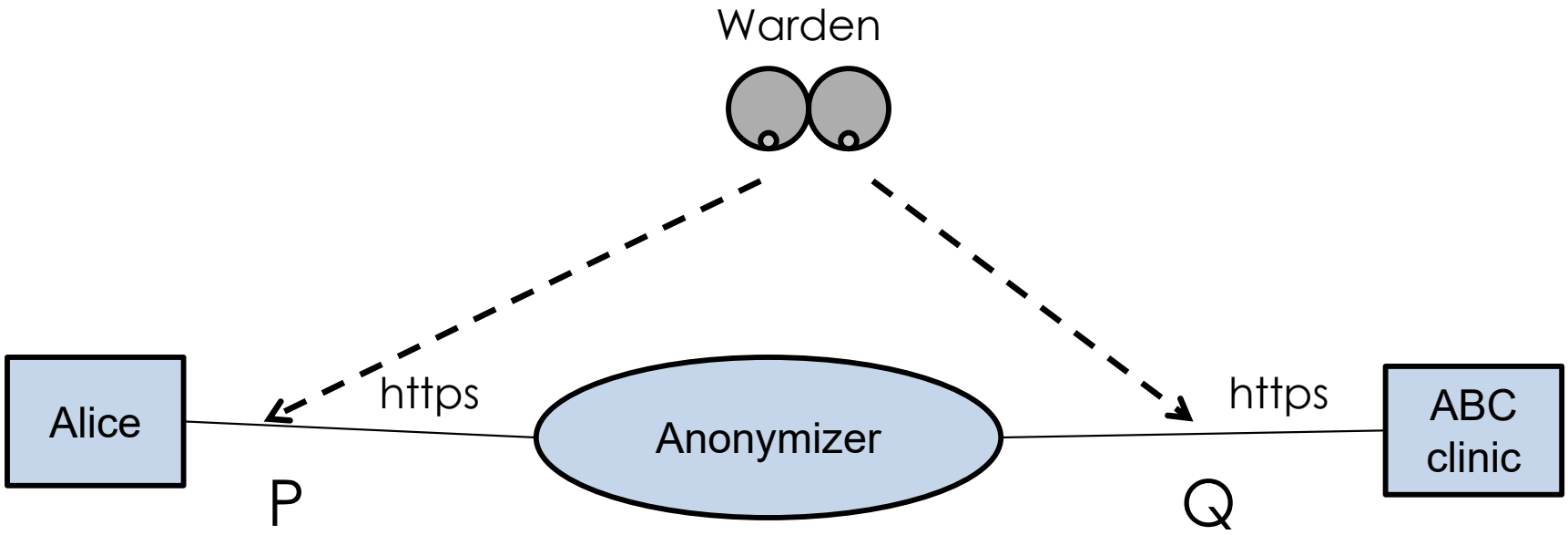
  An anonymizer is simply a **proxy**. It establishes a secure channel with a client and relays/forwards message to-and-from another node



Concerns on Anonymizer

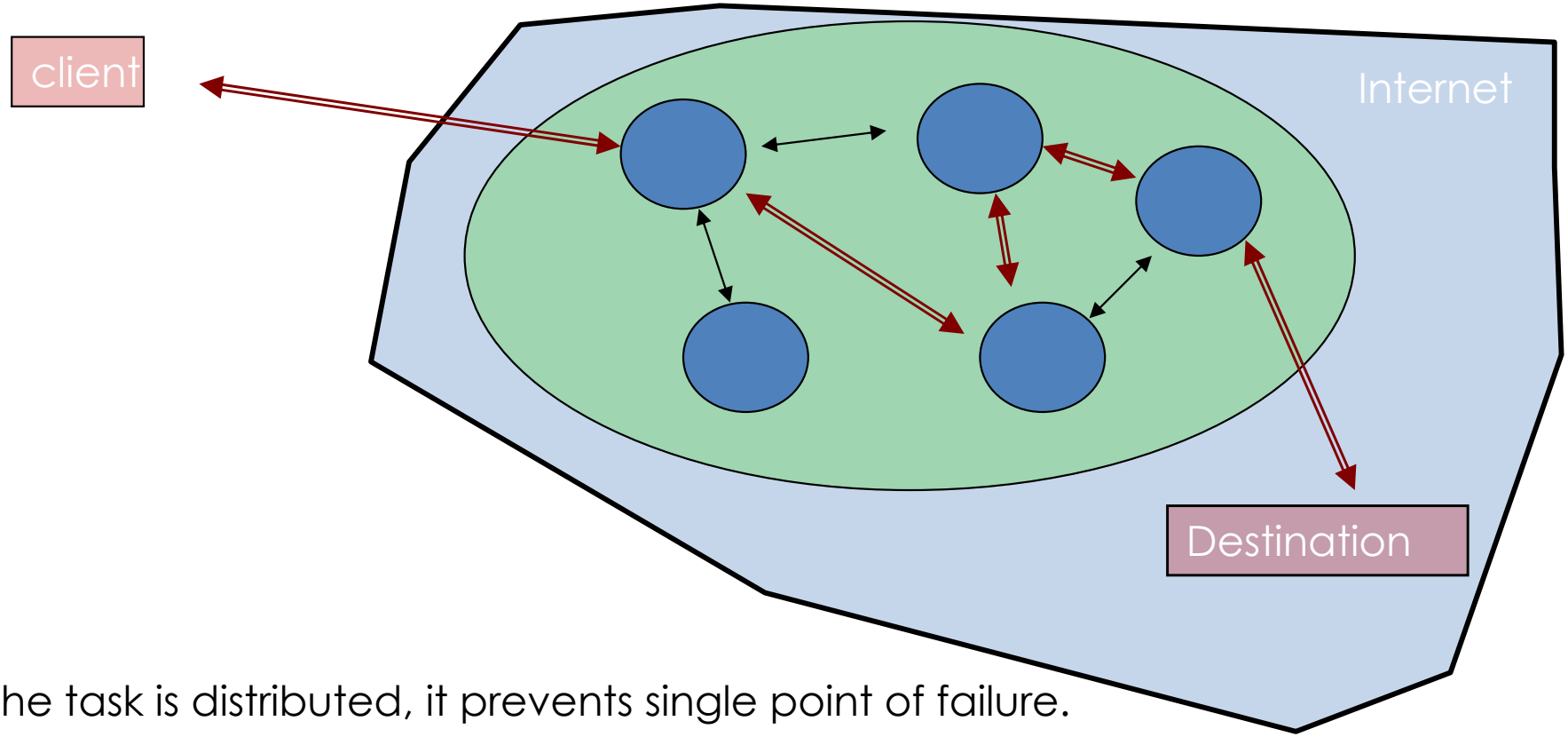
- Are the proxies honest?

- **Traffic analysis**: Consider a powerful warden who has access to the traffic at P and Q obtaining information like **size of packets** and the **timing**. By correlating the traffic at P & Q, he may be able to deduce that Alice is connected to ABC clinic.



Anonymous Routing

- Replace Anonymizer by a **network of Anonymizers**.



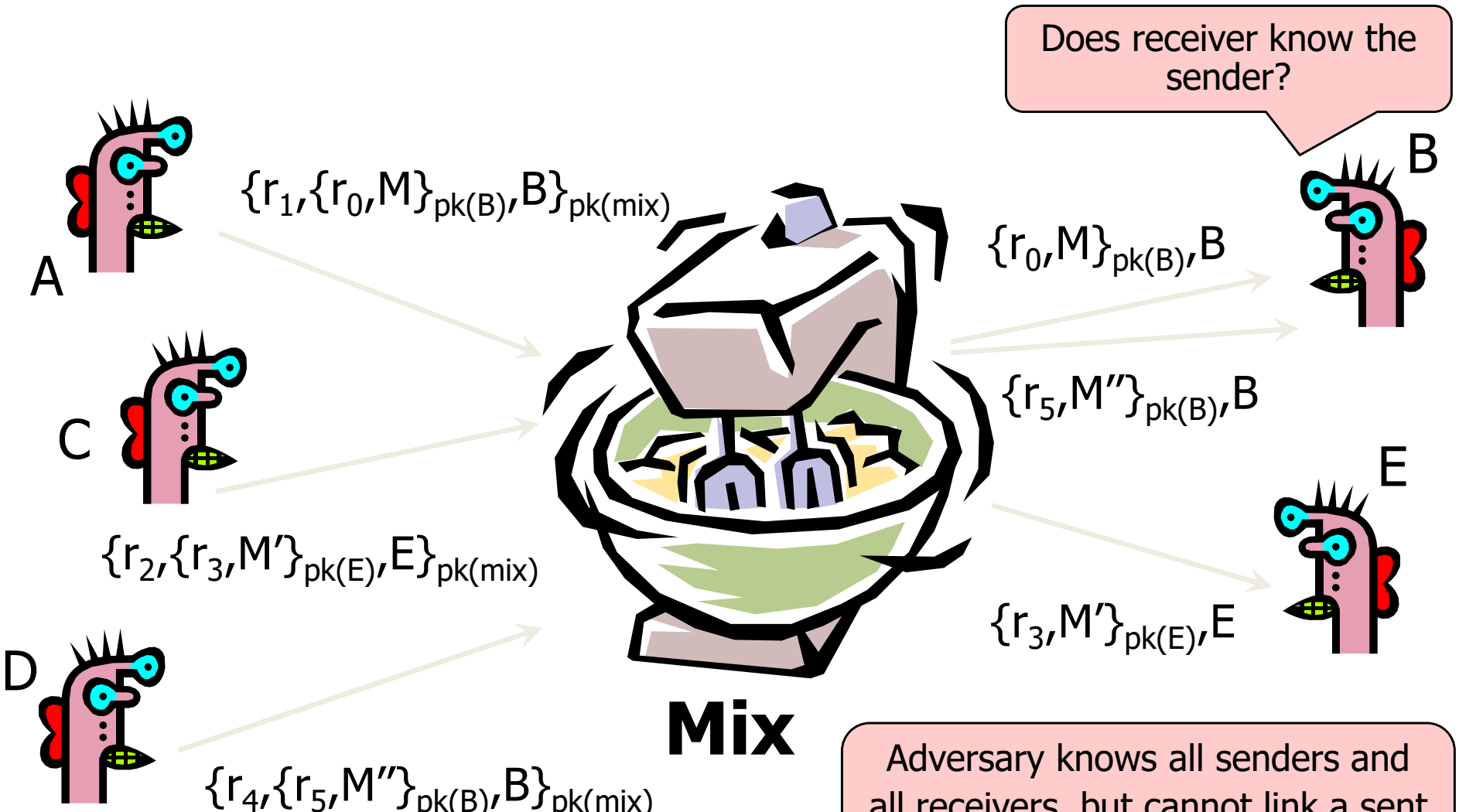
- Since the task is distributed, it prevents single point of failure.
- The network naturally mixes the packets among many connections.
- If the network is large, and there are many users, traffic analysis will be difficult.

Chaum's Mix

Before spam, people thought anonymous email was a good idea ☺

- Early proposal for **anonymous email**
  - David Chaum. "Untraceable electronic mail, return addresses, and digital pseudonyms". Communications of the ACM, February 1981.
  - Hide the sender, i.e., return address
- Public key crypto + trusted re-mailer (Mix)
  - Untrusted communication medium
  - Public keys used as persistent pseudonyms
- Modern anonymity systems use Mix as the basic building block

Basic Mix Design



Anonymous Return Addresses

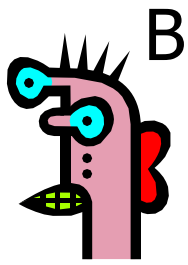
How does B know the return address?

M includes $\{K_1,A\}_{pk(mix)}$, K_2 where K_2 is a fresh public key

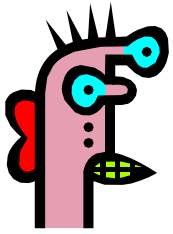
$\{r_1,\{r_0,M\}_{pk(B)},B\}_{pk(mix)}$

MIX

$\{r_0,M\}_{pk(B)},B$



B



A

What if MIX is malicious?

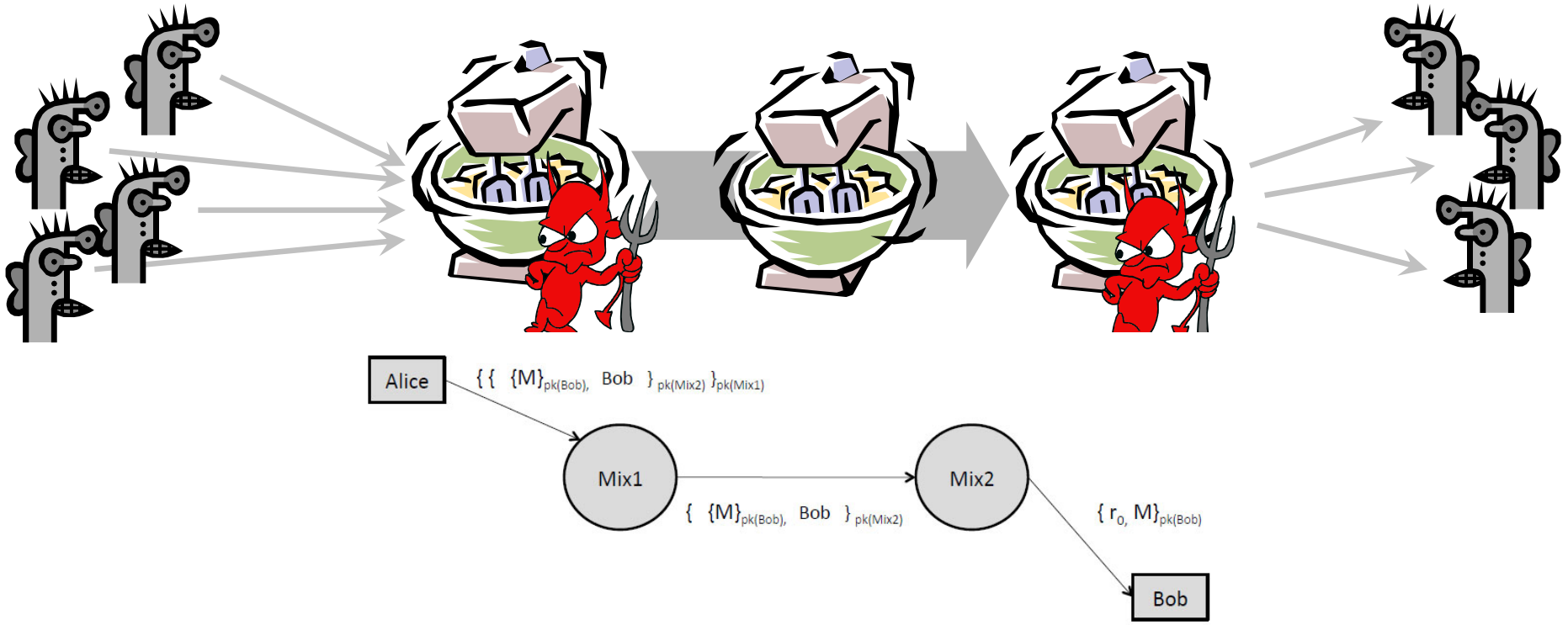
$A,\{\{r_2,M'\}_{K2}\}_{K1}$

Response MIX

$\{K_1,A\}_{pk(mix)}$, $\{r_2,M'\}_{K2}$

Secrecy without authentication

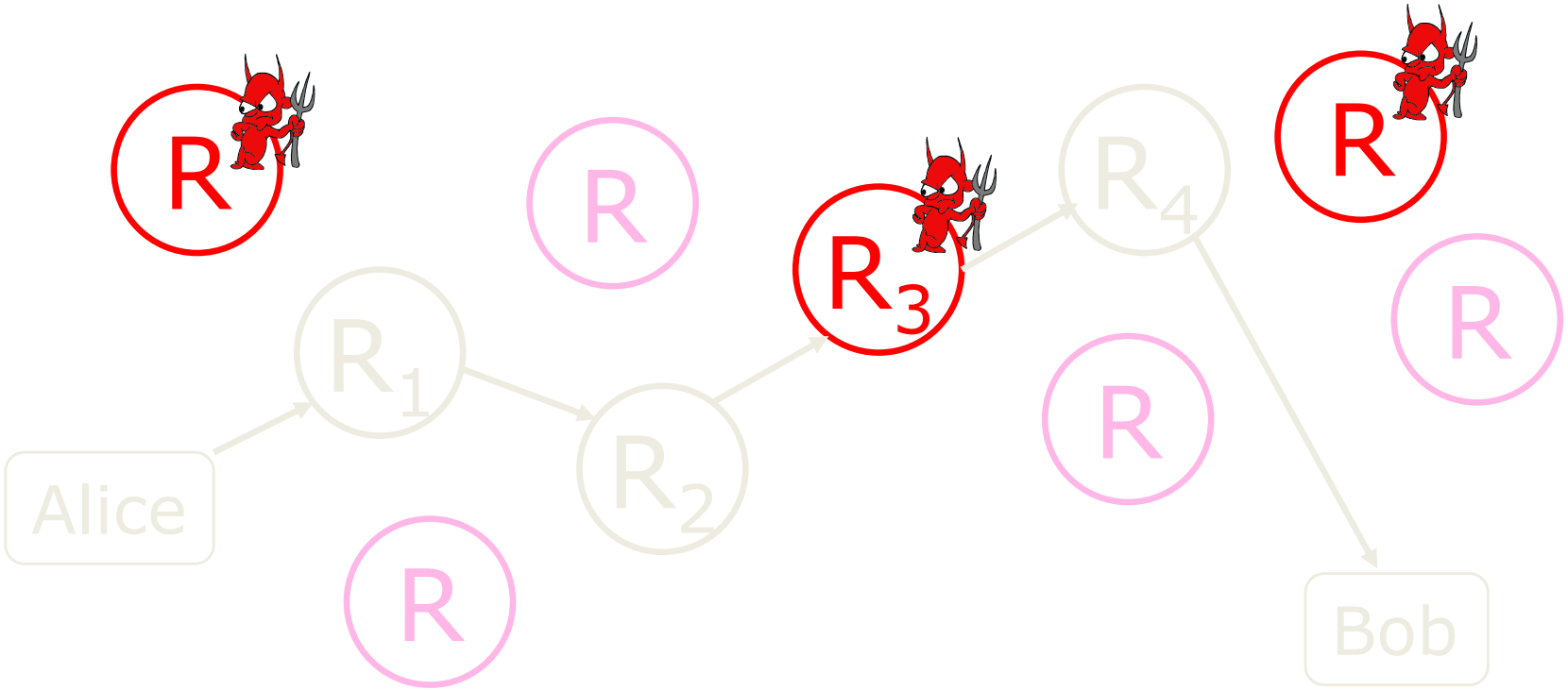
Mix Cascade



- Messages are sent through a **sequence of mixes**
  - Can also form an arbitrary network of mixes ("mixnet")
- Some of the mixes may be controlled by attacker, but even a single good mix guarantees anonymity

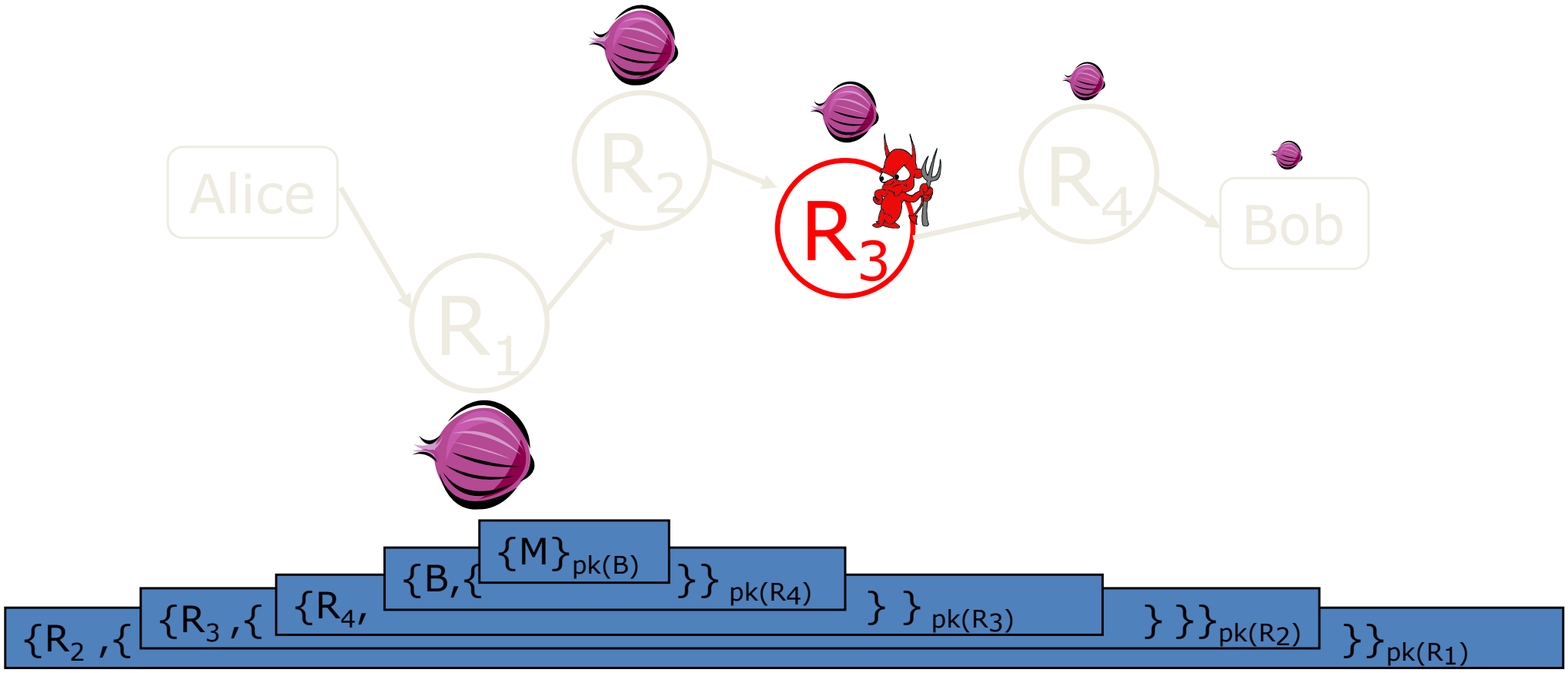
Onion Routing (OR)

[Reed, Syverson, Goldschlag '97]



- Sender chooses a random sequence of routers
  - Each router has public/private key pair
  - Some routers are honest, some controlled by attacker
  - Sender controls the length of the path

Route Establishment



- Routing info for each link encrypted with router's public key
- Each router learns only the identity of the next router

Disadvantages of Basic Mixnets/Onion Routing

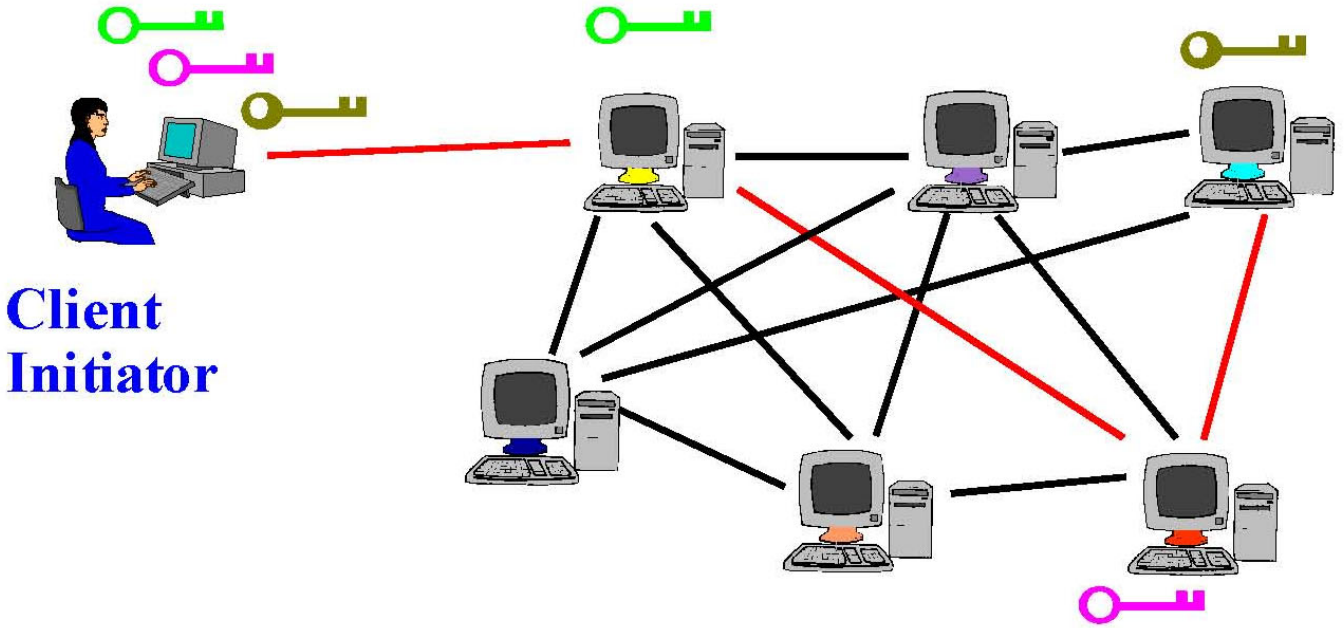
- Public-key encryption and decryption at each mix/router are **computationally expensive**
- Basic Mixnets have **high latency**
  - Ok for email, not Ok for anonymous Web browsing
- Challenge: low-latency anonymity network with good deployability, usability, and flexibility

Tor

- Second-generation onion routing network
  - Developed by Roger Dingledine, Nick Mathewson and Paul Syverson
  - Specifically designed for **low-latency** anonymous Internet communications
- Running since October 2003
- Around 2000 relays
- "Easy-to-use" client proxy
  - Freely available, can use it for anonymous browsing

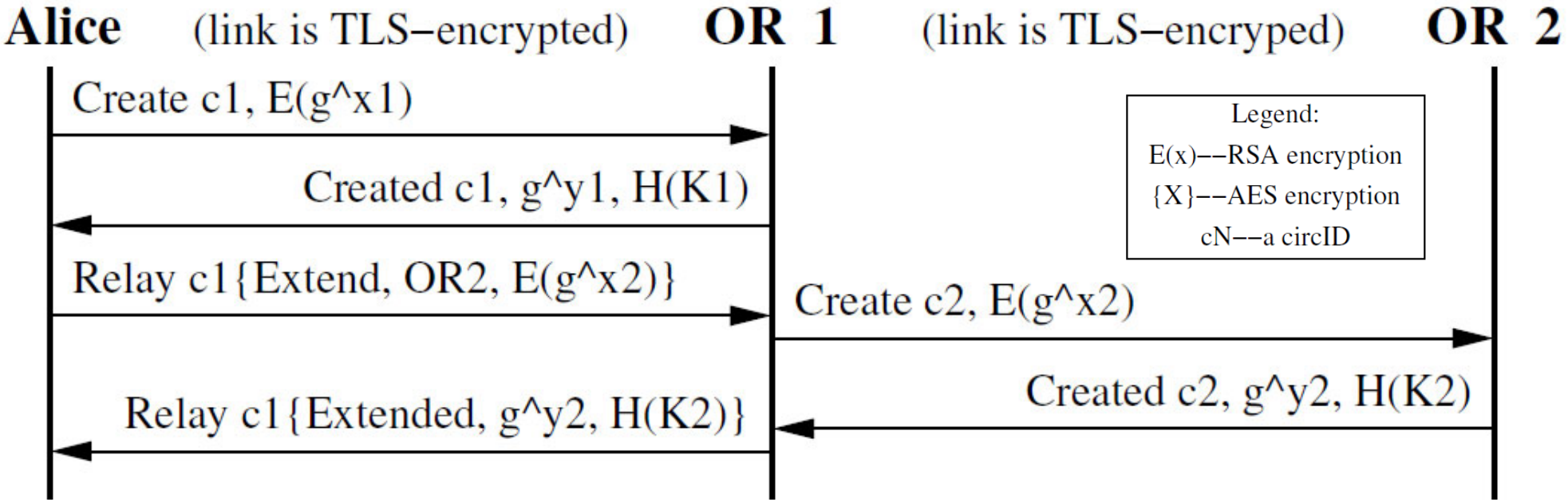
Tor Circuit Setup

- Client proxy establishes **symmetric session keys** with onion routers



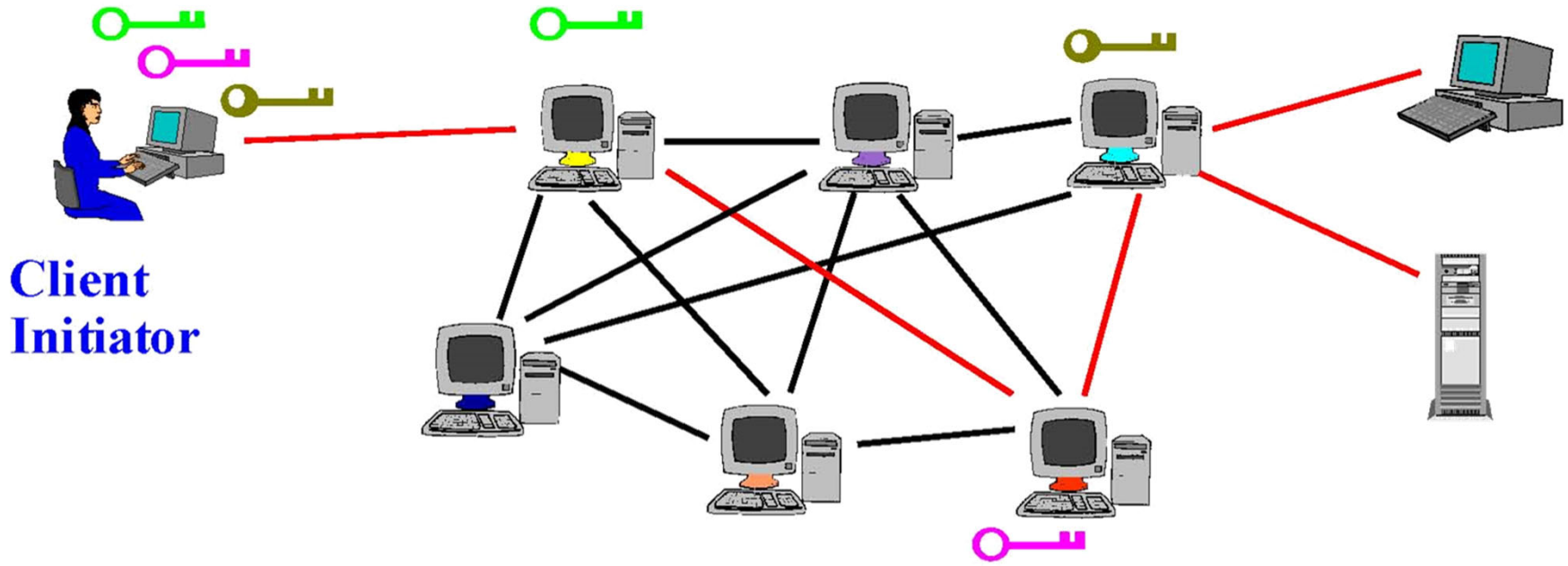
Negotiation of Symmetric Key with OR

- One hop at a time.
- Based on DH key agreement protocol
- Unilateral authentication (Alice authenticates OR, but not the other way)

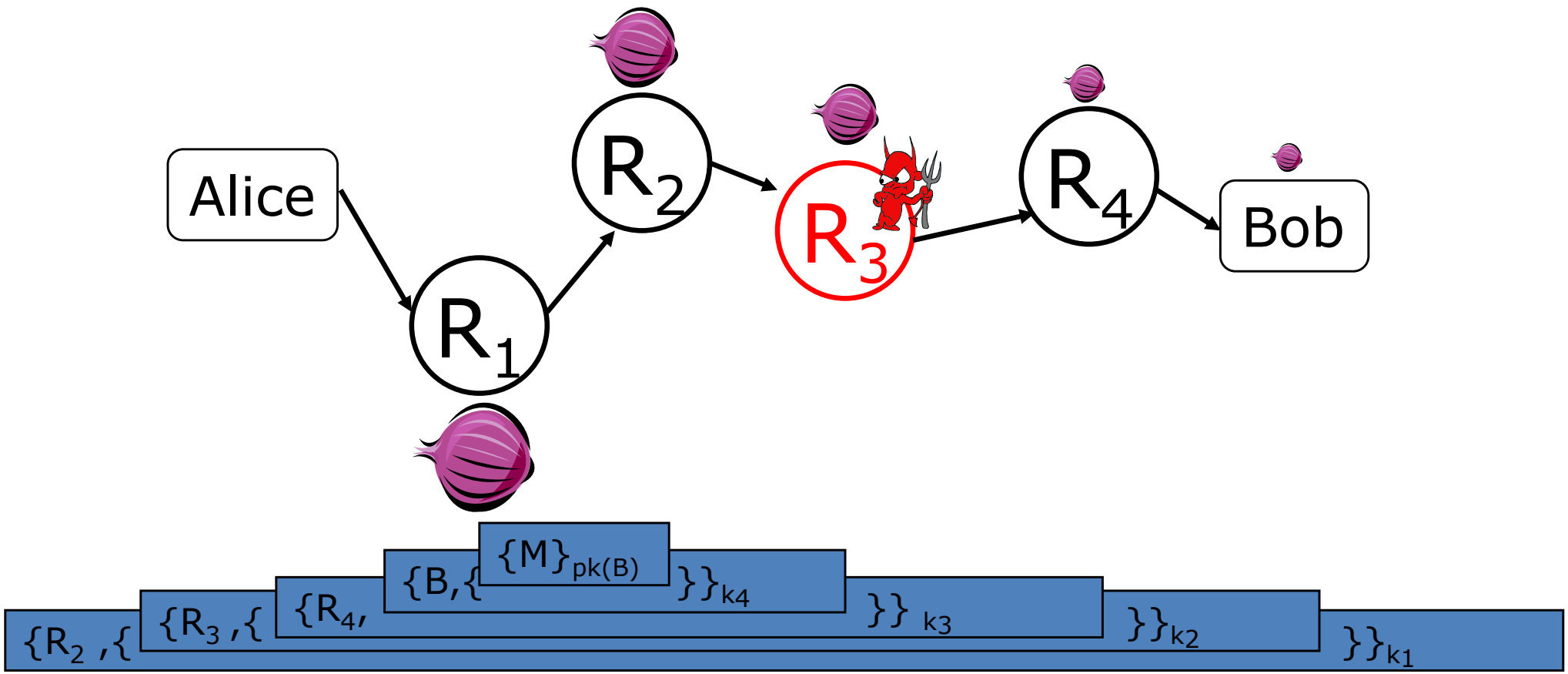


Using a Tor Circuit

- Client applications connect and communicate over the established Tor circuit
  - Note onion now uses only **symmetric keys** for routers

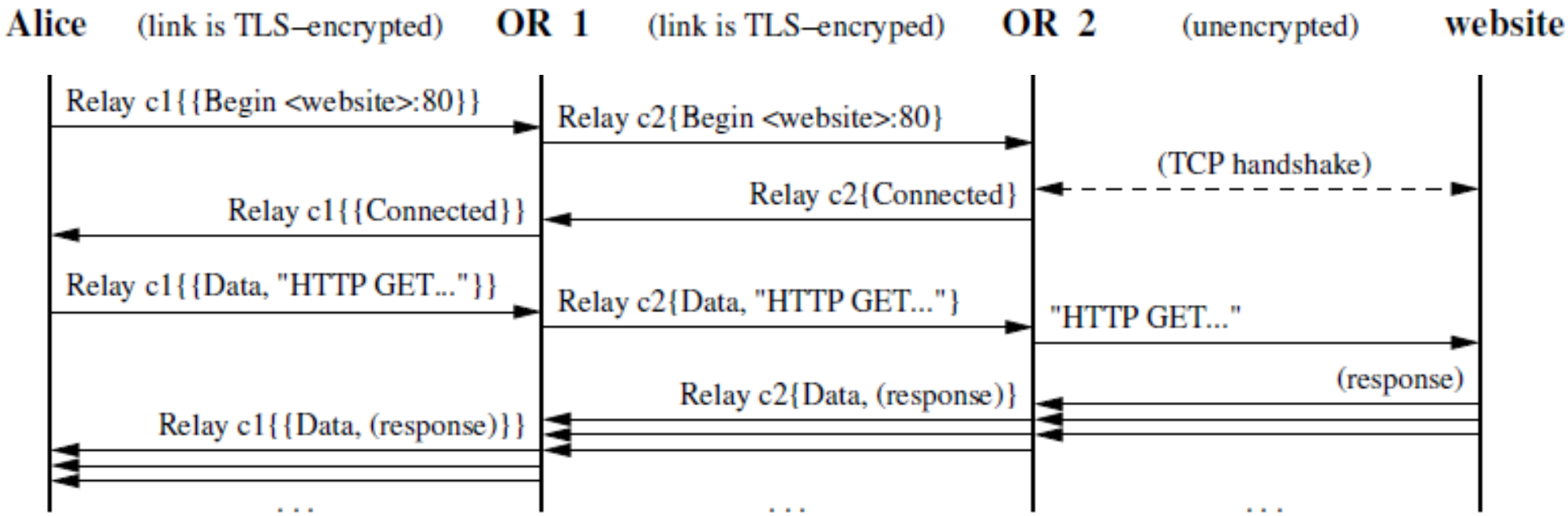


Using a Tor Circuit



Note onion now uses only symmetric keys for routers

Using a Tor Circuit (Example)



Legend:
E(x)––RSA encryption
{X}––AES encryption
cN––a circID

Tor's features

- **Many applications** can share one circuit
  - Multiple TCP streams over one anonymous connection
  - Use of SOCKS proxy interface

- Tor router **doesn't need root privileges or kernel modification**
  - Encourages people to set up their own routers
  - **More participants = better anonymity** (for everyone)

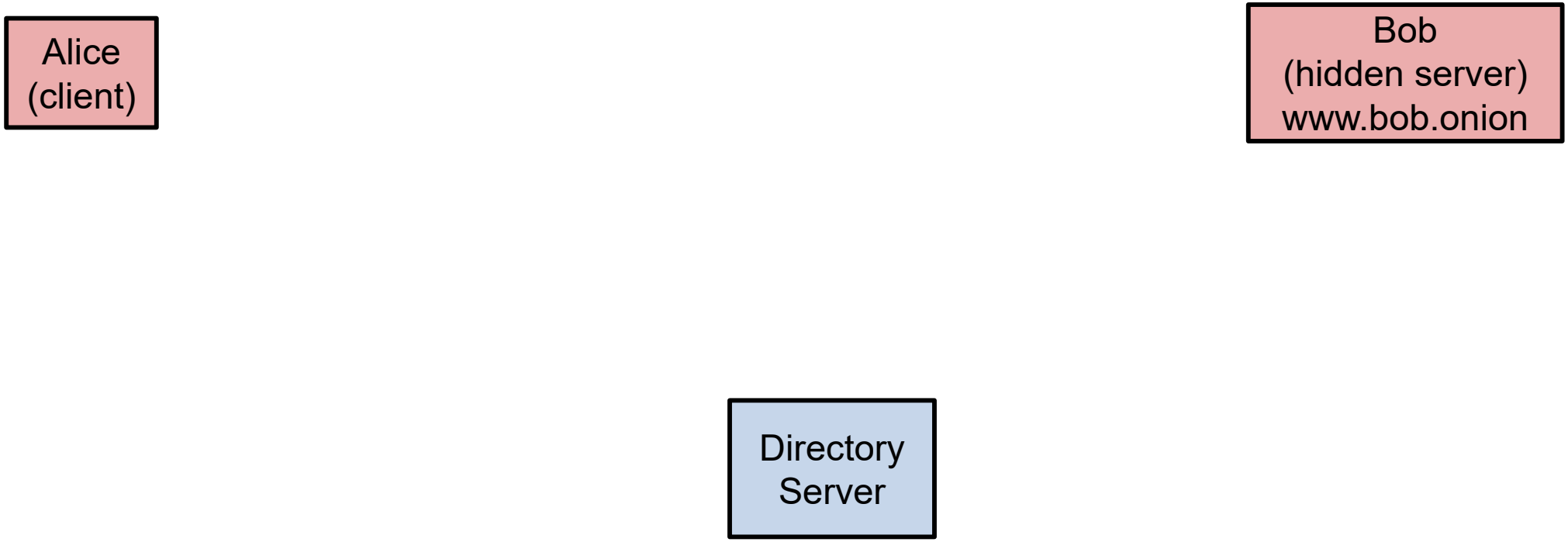
Tor directory servers

- How do clients learn about the onion routers?
  - First generation OR used in-band network status update
    - Expensive flooding required
    - Partitioning attack
- Tor uses nine **directory servers**
  - to maintain lists of active onion routers, their locations, current public keys, etc.
  - Control how new routers join the network
    - "Sybil attack": attacker creates a large number of routers
  - Directory servers' keys ship with Tor code
- Disadvantages?
  - Attacks against directory servers

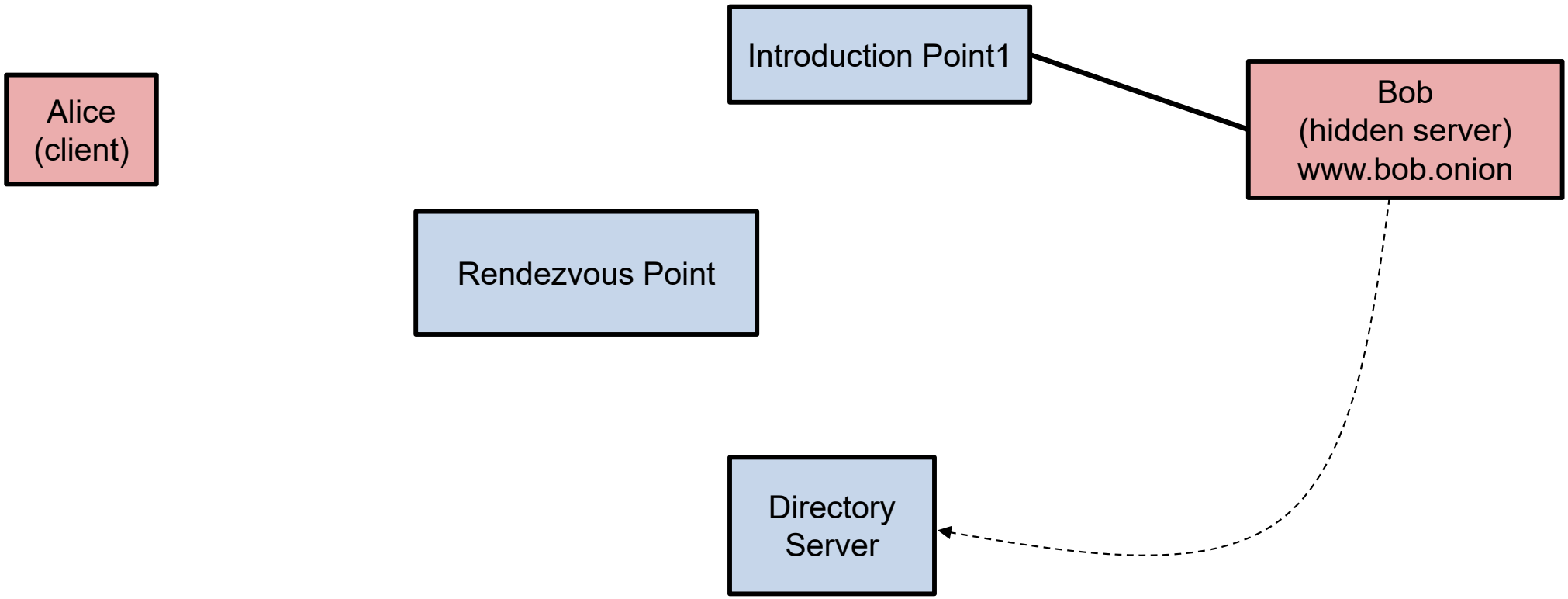
Hidden services

- Tor offers sender anonymity; how about **responder anonymity**?
- *Hidden services (.onion)* via ***rendezvous points***
  - Hidden service selects and announces several ORs as its **introduction points**
  - Hidden service creates circuits to introduction points
  - Client picks one introduction point and construct a circuit

Hidden Service in Tor



1. Hidden Server looks for a node who is willing to be his Introduction Point. Keep trying until at least one is found. There could be more than one.
2. Hidden Server notifies Directory Server. Information on the service (this include name of services, like www.bob.onion , and the IP address of the introduction points) will be stored and managed by the Directory Server.

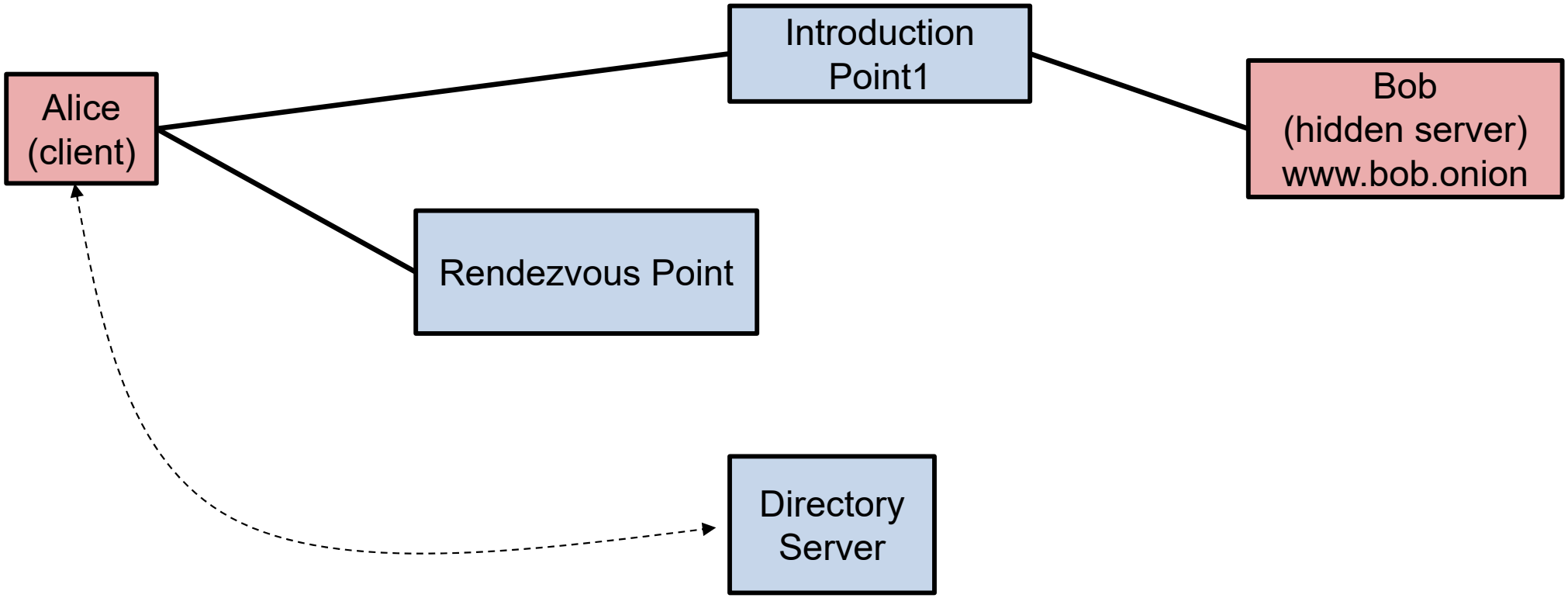


Note that:
Directory server and the introduction points do not know the IP address of the Hidden Server.
Hidden Server knows the IP address of the introduction points.

1. Consider a client who wants the hidden service. From the Directory Server, he get to know of the service and IP address of an introduction point.

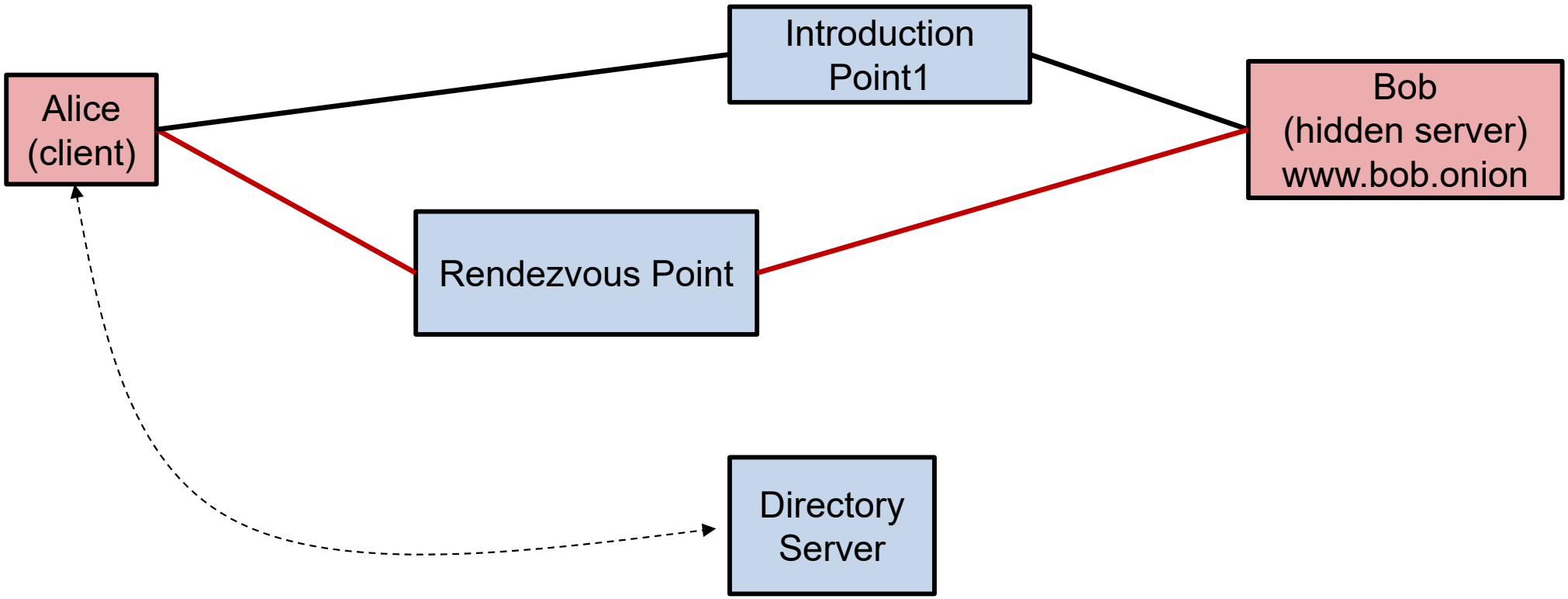
2. The client finds a node that is willing to be the Rendezvous point. Next, the client sends the Rendezvous point's address to the Hidden Server via the Introduction Point1.

All communications are carried out anonymously using onion routing.



Note that,

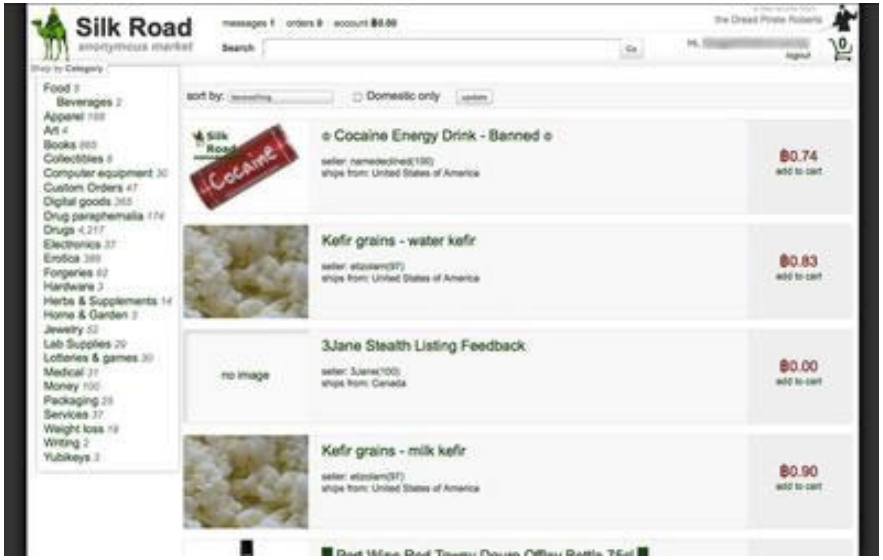
1. IP addresses of Introduction Point1 and Rendezvous Point are known to both Client and Hidden Server.
2. IP addresses of Client and Hidden Server remain secret to both Introduction point and Rendezvous Point.



Hidden Server provides service to Client via Rendezvous Point.

Example of Hidden Service

- Silk Road / Silk Road 2.0
  - An online black market and the first modern darknet market
  - The website was known for its illegal drug marketplace, among other illegal and legal product listings.

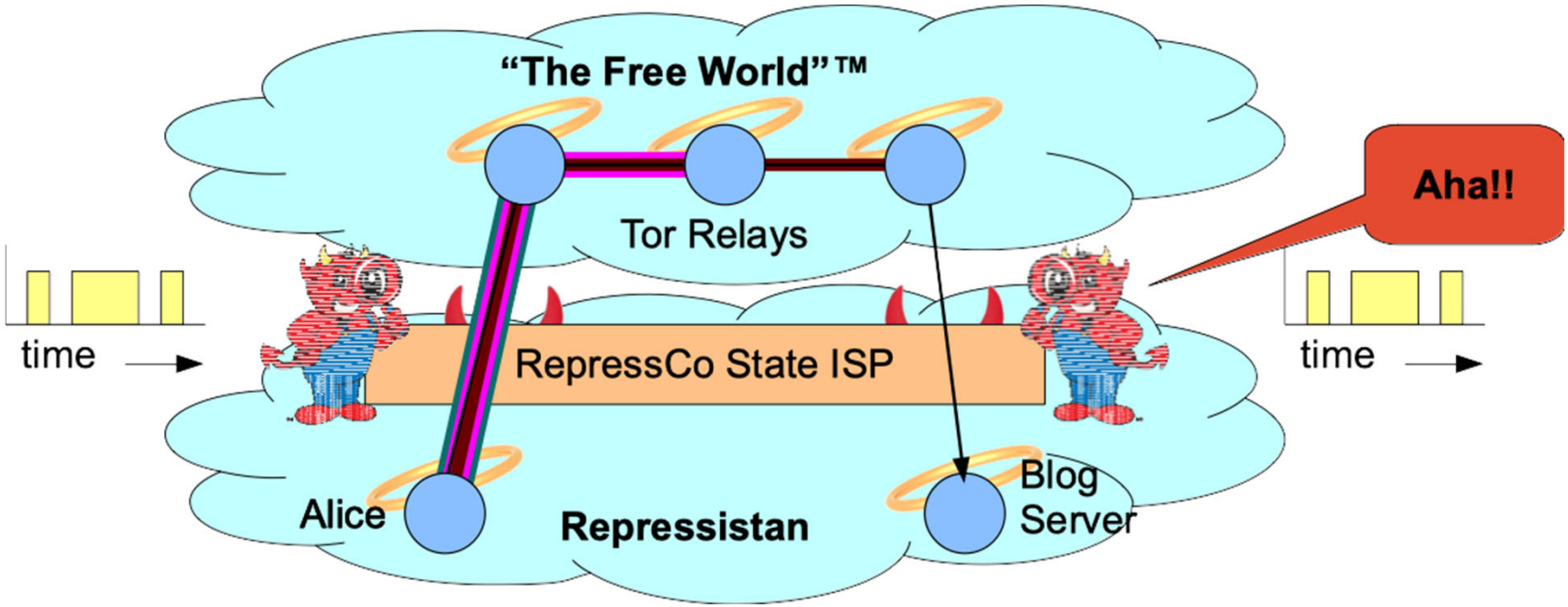


Anonymity against Network Attackers

- Network attackers
  - Partial or global network access (e.g., tapping ISPs by three-letter agencies)
- Traffic correlation attacks
  - Passively monitor cells entering/leaving onion routers
  - Identify client-server pairs
- Traffic confirmation attacks
  - Actively mark flows to identify client-server pairs
- Known solutions: mixing, padding, traffic shaping
  - Design choice: anonymity vs. latency

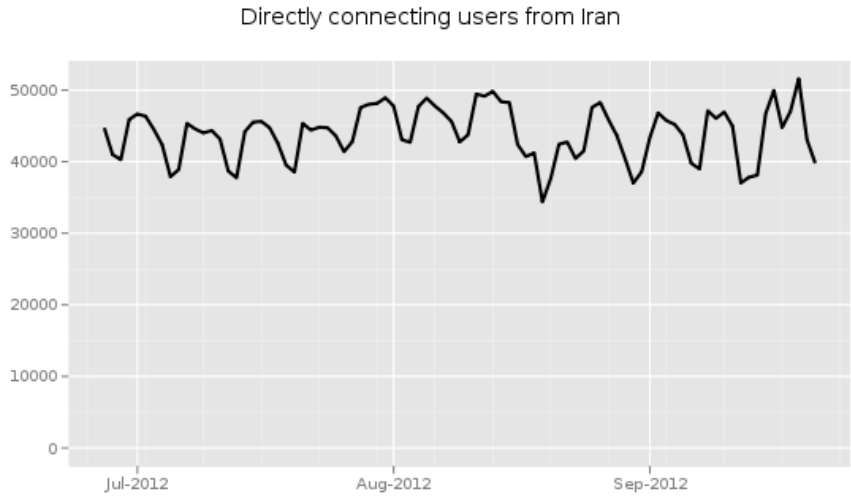
Traffic Analysis: Example 1

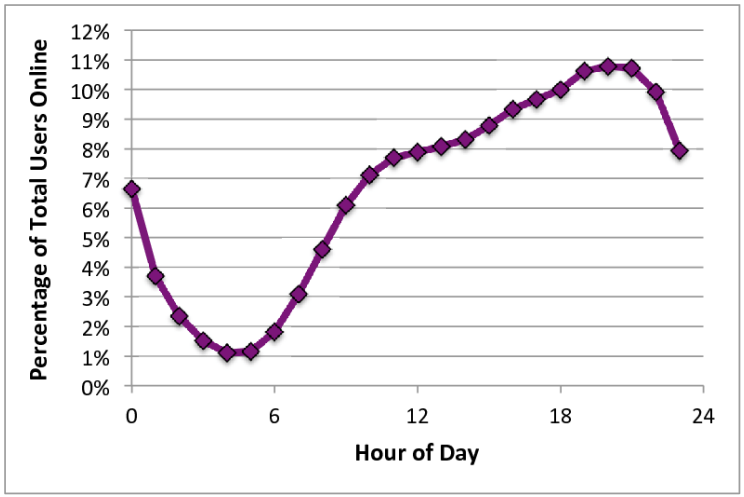
- Alice in Repressistan uses Tor to post onblog server hosted in Repressistan
- State ISP controls *both* entry and exit hops
- Fingerprint & correlate traffic to **deanonymize**



Traffic Analysis: Example 2

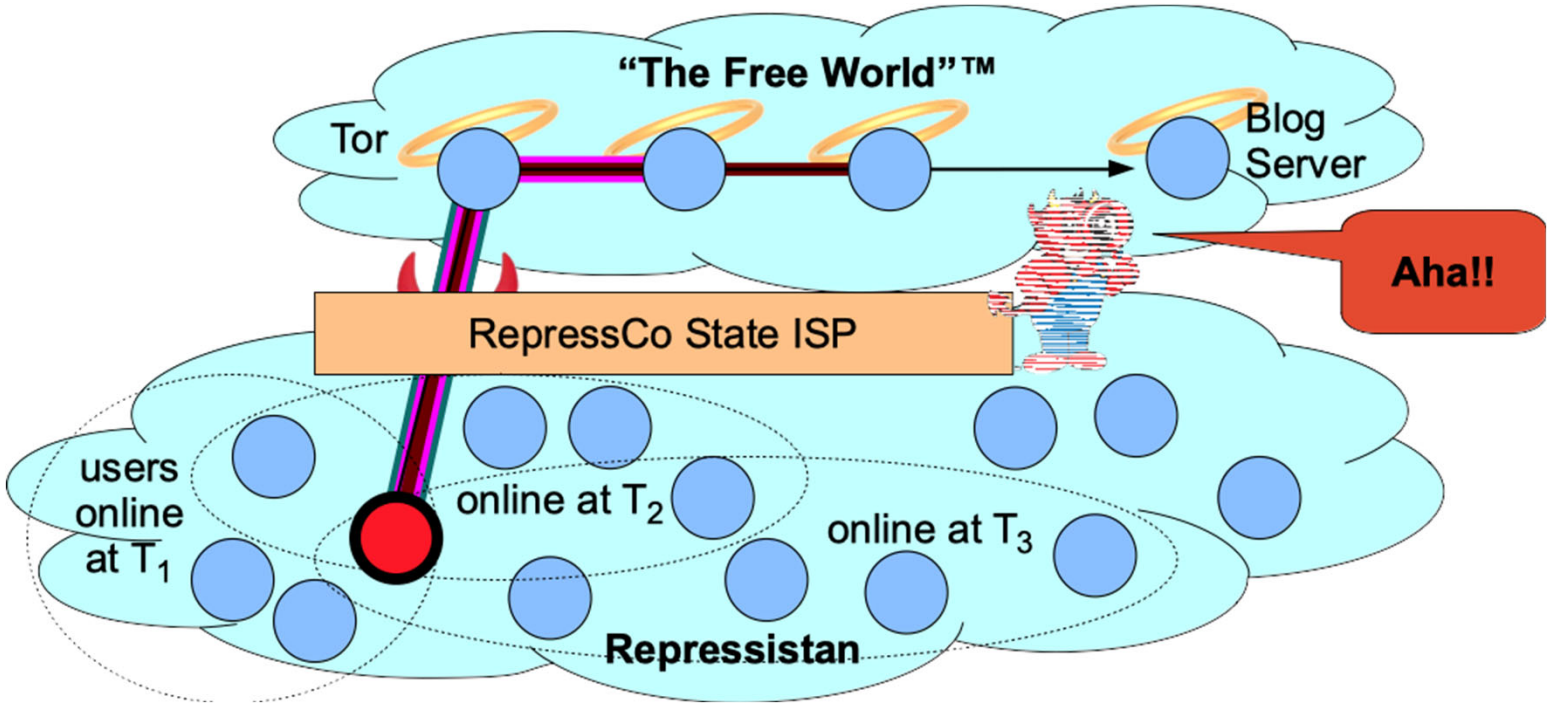
- Bob in Repressistan posts via Tor to blog hosted in “The Free World”™
- Tor Metrics: 50,000 users/day connect from Repressistan
  - Good anonymity, right?
- But ISP logs tell police when users are online; blog post has timestamp
  - How many users are online ***at same time Bob posts?***
  - ~5,000 at 7PM?~500 at 5AM?





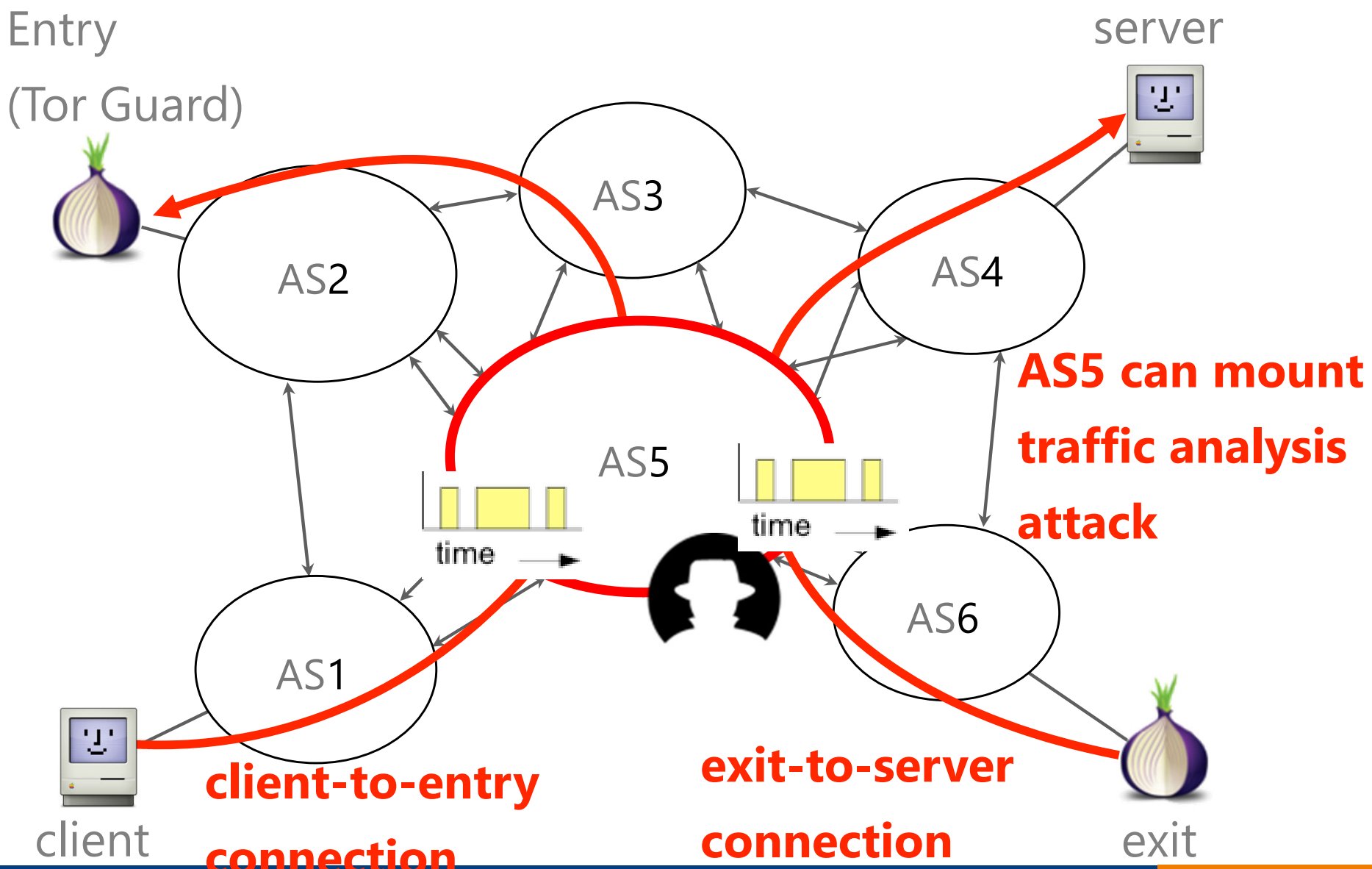
Intersection Attack: Example

- Bob signs posts with pseudonym “AnoniBob”
  - Posts 3 signed messages at times T1, T2, T3
  - Police find sets of users online each time, **intersect**



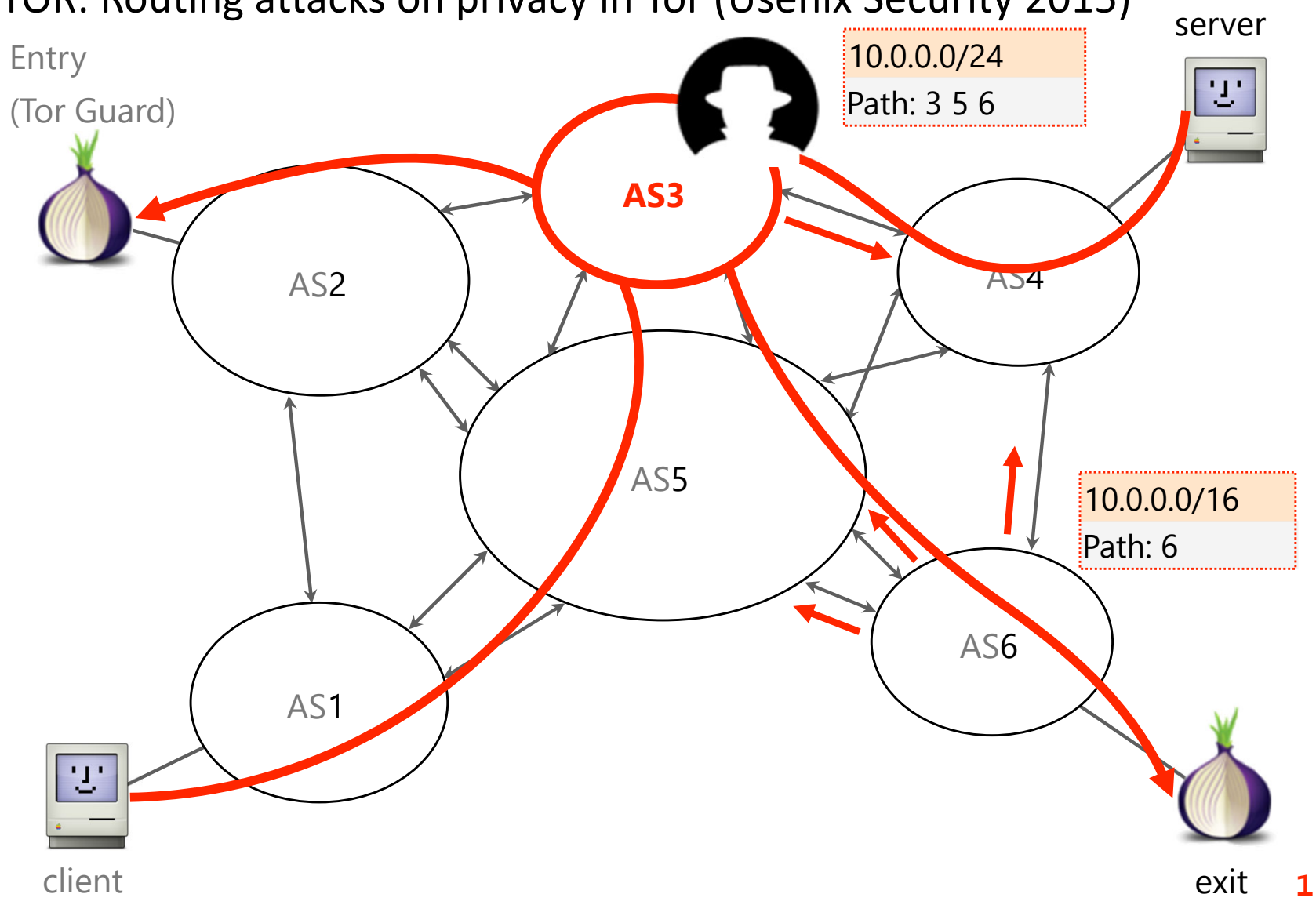
Traffic correlation attacks by *ISPs*

RAPTOR: Routing attacks on privacy in Tor (Usenix Security 2015)



Traffic correlation attacks by ISPs by ***hijacking prefixes***

RAPTOR: Routing attacks on privacy in Tor (Usenix Security 2015)



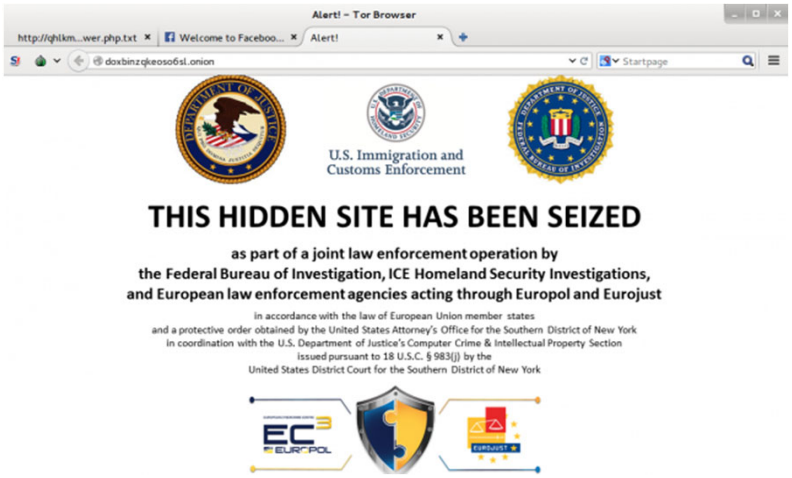
Other attacks: Malicious Exit Nodes

- Attackers can deploy Tor onion routers that allows 'exits'
- Exit nodes do not know clients; yet, they see the destination
  - If HTTP used, see the contents as well.
  - Application layer protocol may reveal information about the source (e.g., BitTorrent)

- Honeypot research in 2015
  - Set up a fake Bitcoin wallet website with no TLS
  - Logins through all 1,400 Tor exit nodes
  - In one month, 16 exit nodes tried to login!

Other Attacks: Operation Onymous

- An international law enforcement operation targeting darknet markets and other hidden services operating on the Tor network.
- Joint effort by Federal Bureau of Investigation (FBI) and the European Union Intelligence Agency Europol.
- Hundreds of sites, including Silk Road 2.0, were shut down.
- How did they do?
  - Not announced.
  - Research done by CMU might be replicated.



So, is Tor broken?

- Non goals:
  - Side channels (e.g., timing, packet size)
  - Traffic analysis
  - Directory servers
  - Malicious ORs
  - Not hiding who is connected to the network
  - Lack of anonymity set
  - …

Summary

- Protocols for anonymous communication
  - High-latency
    - Chaum Mixes as a building block
  - Low-latency
    - Onion Routing and Tor
- Knowing the accurate anonymity guarantee against what attacker capability is important!

NEXT WEEK:
ANTI-CENSORSHIP SYSTEM

Two papers to read

- [Paper 1] Parrot is dead (IEEE S&P 2012)
  - How anti-censorship systems easily fail in practice

- [Paper 2] Telex (Usenix Security 2011)
  - How to fool censorship system with the help of friendly ISPs