

# Lecture 26: Anonymity

COMP 411, Fall 2022

Victoria Manfredi



**Acknowledgements:** materials adapted from Computer Networking: A Top Down Approach 7<sup>th</sup> edition: ©1996-2016, J.F Kurose and K.W. Ross, All Rights Reserved as well as from Avi Kak's lecture 12 slides at <https://engineering.purdue.edu/kak/compsec/>

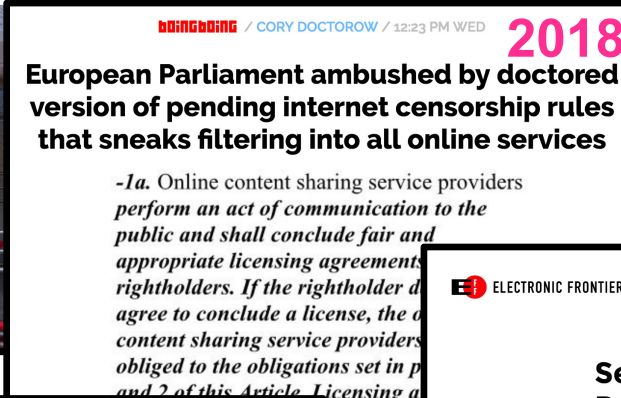
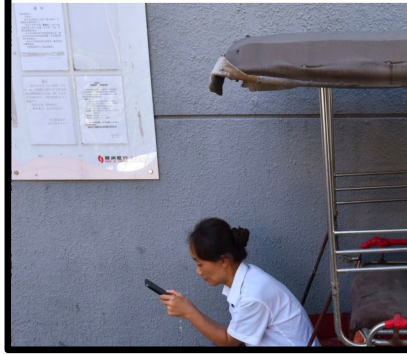
# Today

1. Anonymous communication
  - Tor
2. What have we covered?

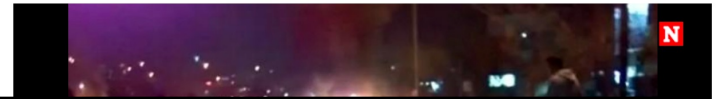
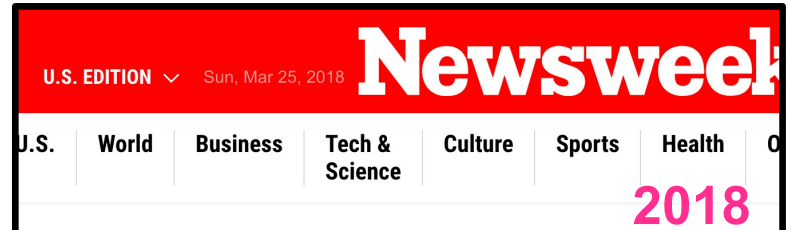
# Anonymity and Internet Censorship

## **OVERVIEW**

# Censorship in the real world



-1a. Online content sharing service providers perform an act of communication to the public and shall conclude fair and appropriate licensing agreements with rightholders. If the rightholder does not agree to conclude a license, the content sharing service providers are obliged to the obligations set in paragraph 2 of this Article. Licensing agreements shall be concluded on fair and reasonable terms between the content sharing service providers and the rightholders. The content sharing service providers shall be liable for the content shared on their platform in line with the licensing agreements. The content sharing service providers do not have a right of representation.



# China's state-of-the-art censorship

## The Great Firewall of China

- searches in China give ``alternate'' results for certain words
- terminates connections if packets contain certain words
- ... plus much more!

Man in China sentenced to five years' jail for running VPN

2017

As part of an internet 'cleanup', Wu Xiangyang was also fined an amount equal to his profits since starting service in 2013



GREAT FIRE GreatFire.org  
@GreatFireChina

2018

Following

The authorities temporarily censored the letter "N" on social media in China as Chinese netizens were trying to calculate how long Xi Jinping might stay in power.

$$\begin{pmatrix} N_{t+l_1} \\ N_{t+l_2} \\ N_{t+l_3} \end{pmatrix} = \begin{pmatrix} F_1 & F_2 & F_3 \\ S_1 & 0 & 0 \\ 0 & S_2 & 0 \end{pmatrix} \begin{pmatrix} N_{t_1} \\ N_{t_2} \\ N_{t_3} \end{pmatrix}.$$

1:25 AM - 27 Feb 2018



VPNs banned in China and Russia

# Why aren't TLS and IPsec enough?

## Packet headers are plaintext

- src and dst IP addresses visible to everyone
- that's how Internet routing is able to function ...

## Traffic analysis attacks

- obtain info about original src of packets and their ultimate dst

## Even IPsec is vulnerable

- packet sniffer at any point before packets get to encapsulator used for Tunnel Mode knows both pkt src and dst

# Anonymous communication

## **TOR**

# Acknowledgements

## Most of this section based on

- Vanbever slides for “Anonymity on Quicksand: Using BGP to compromise Tor”
  - [https://nsg.ee.ethz.ch/fileadmin/user\\_upload/publications/nsg\\_vanbever\\_bgp\\_tor\\_hotnets\\_slides\\_2014\\_slides.pdf](https://nsg.ee.ethz.ch/fileadmin/user_upload/publications/nsg_vanbever_bgp_tor_hotnets_slides_2014_slides.pdf)
- Avi Kak lecture 20 slides
  - <https://engineering.purdue.edu/kak/compsec>
- <https://www.torproject.org/about/overview>

# The Onion Router (Tor)

## Goals

- enable user to access blocked sites
- hide who is talking to whom from adversary

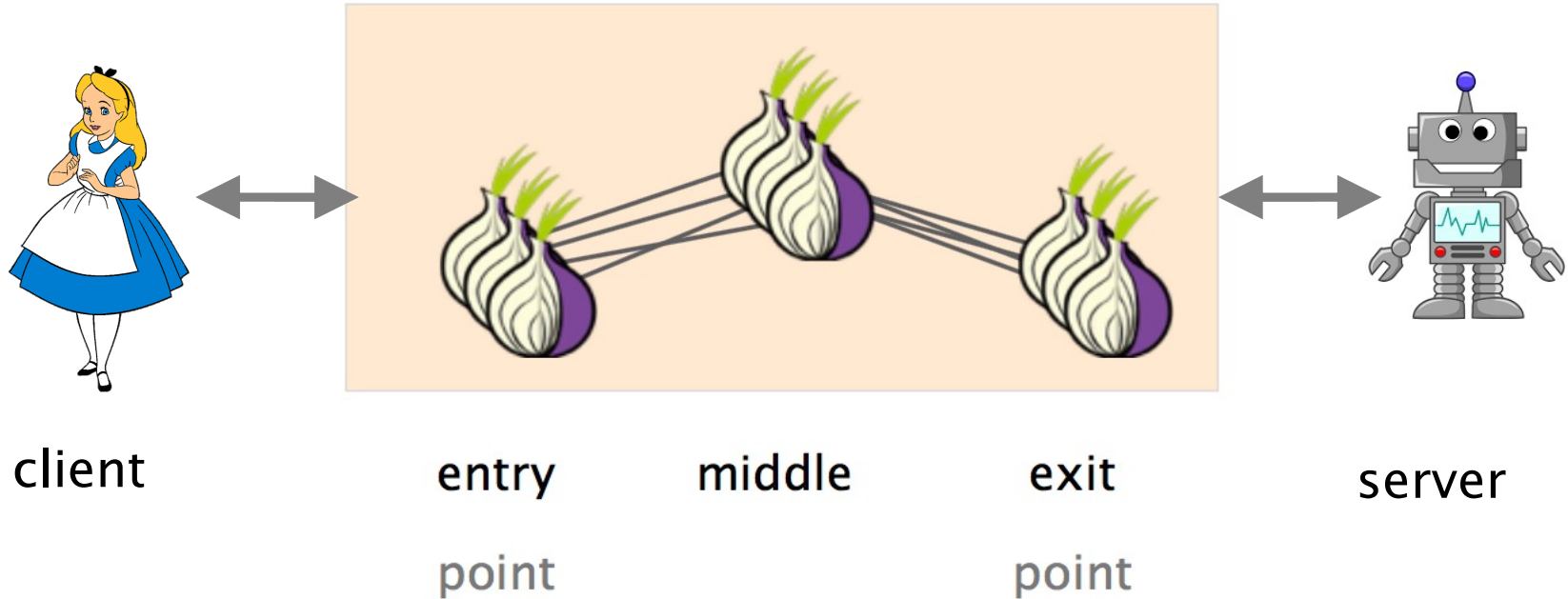
## Uses onion routing to provide anonymity

- by Roger Dingledine, Nick Mathewson, Paul Syverson

# Tor bounces traffic around network of relays

Relays are end-hosts on Internet

Tor network is overlay on Internet

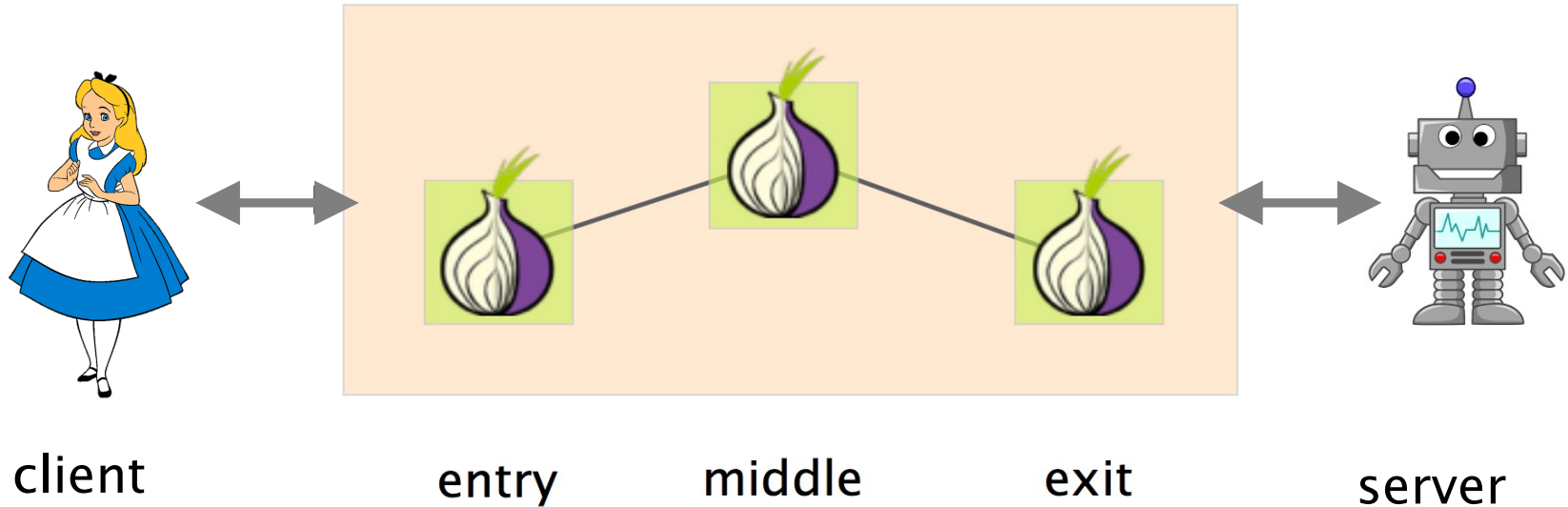


How?

# Client selects 3 relays, 1 of each type

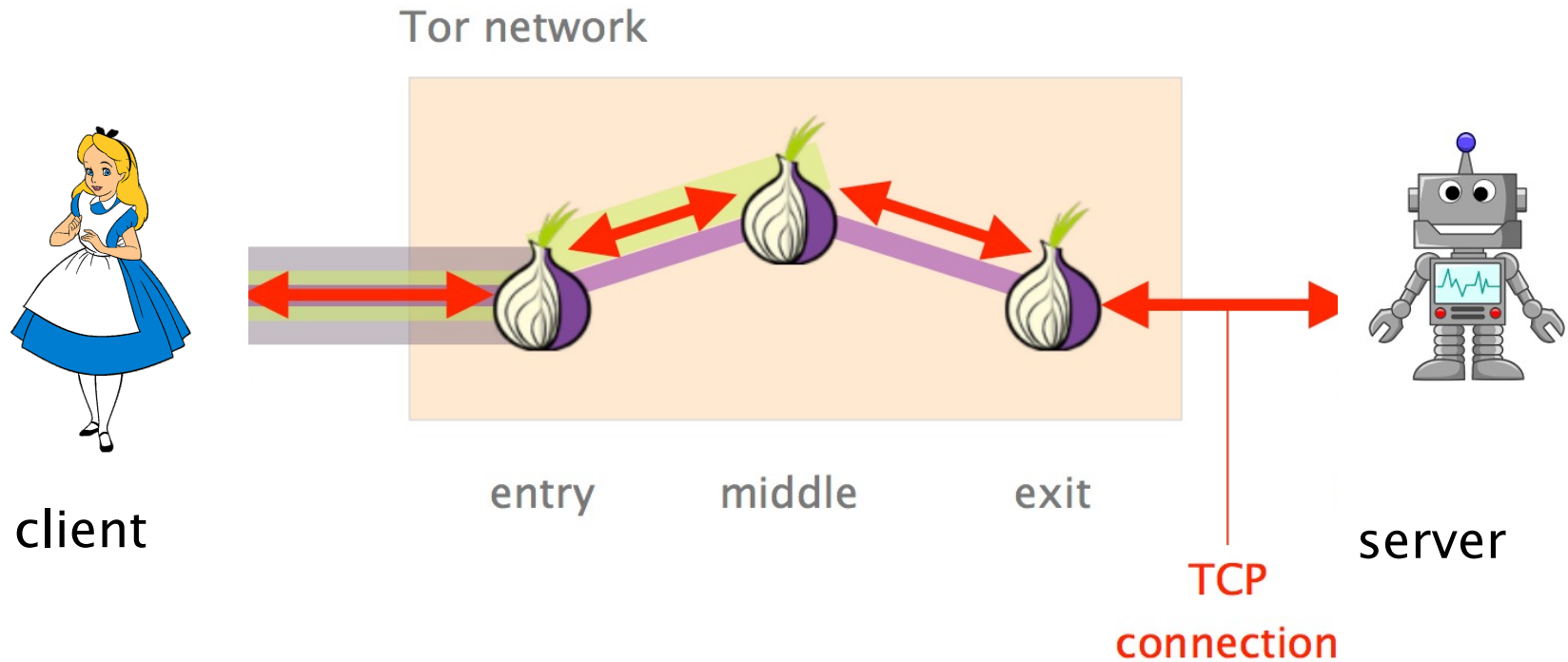
Relays are end-hosts on Internet

Tor network **is overlay on Internet**



Client's onion proxy queries Tor directory for IP addresses of onion routers in Tor overlay, chooses 3

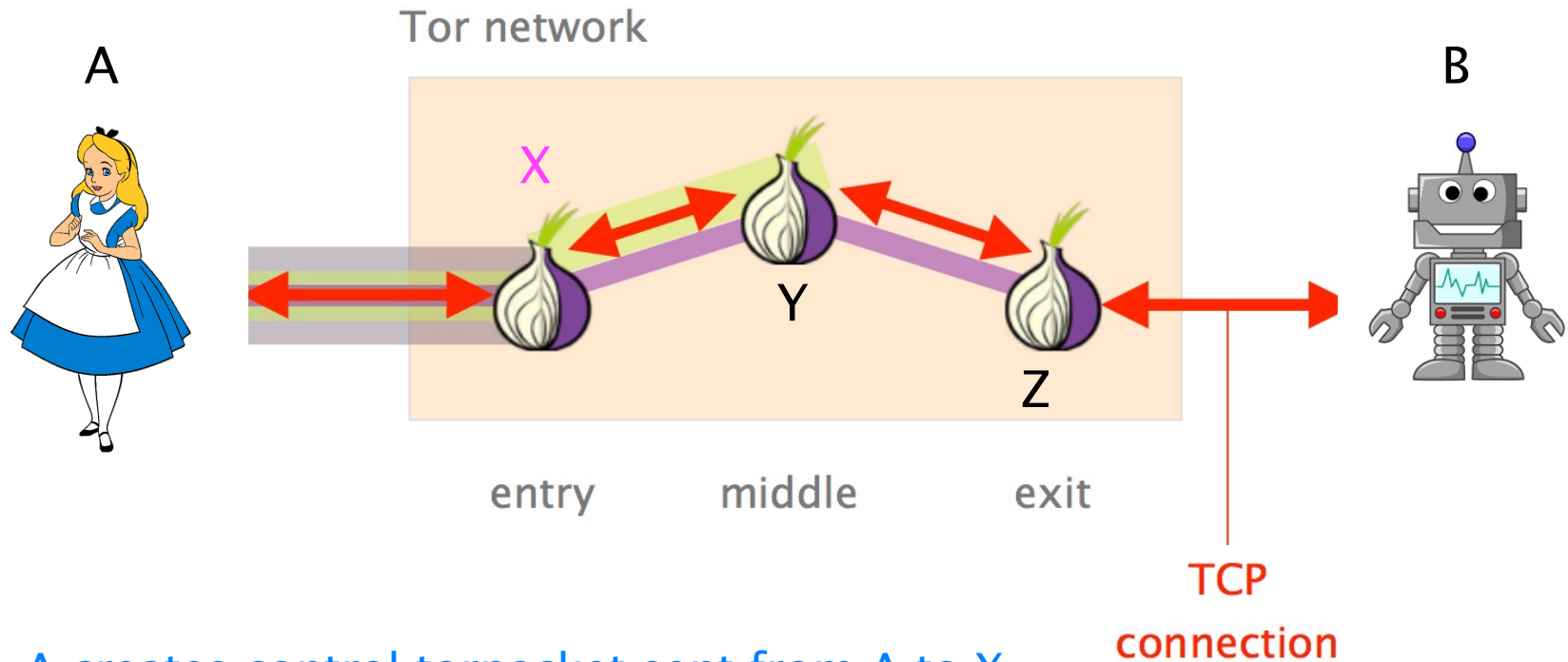
# Client incrementally builds encrypted circuit



Every Onion router has

1. (static) public RSA key known to client's onion proxy
2. Diffie-Hellman (DH) key  $K$  created between client's onion proxy and each selected onion router

# How A extends circuit to X



A creates control torpacket sent from A to X

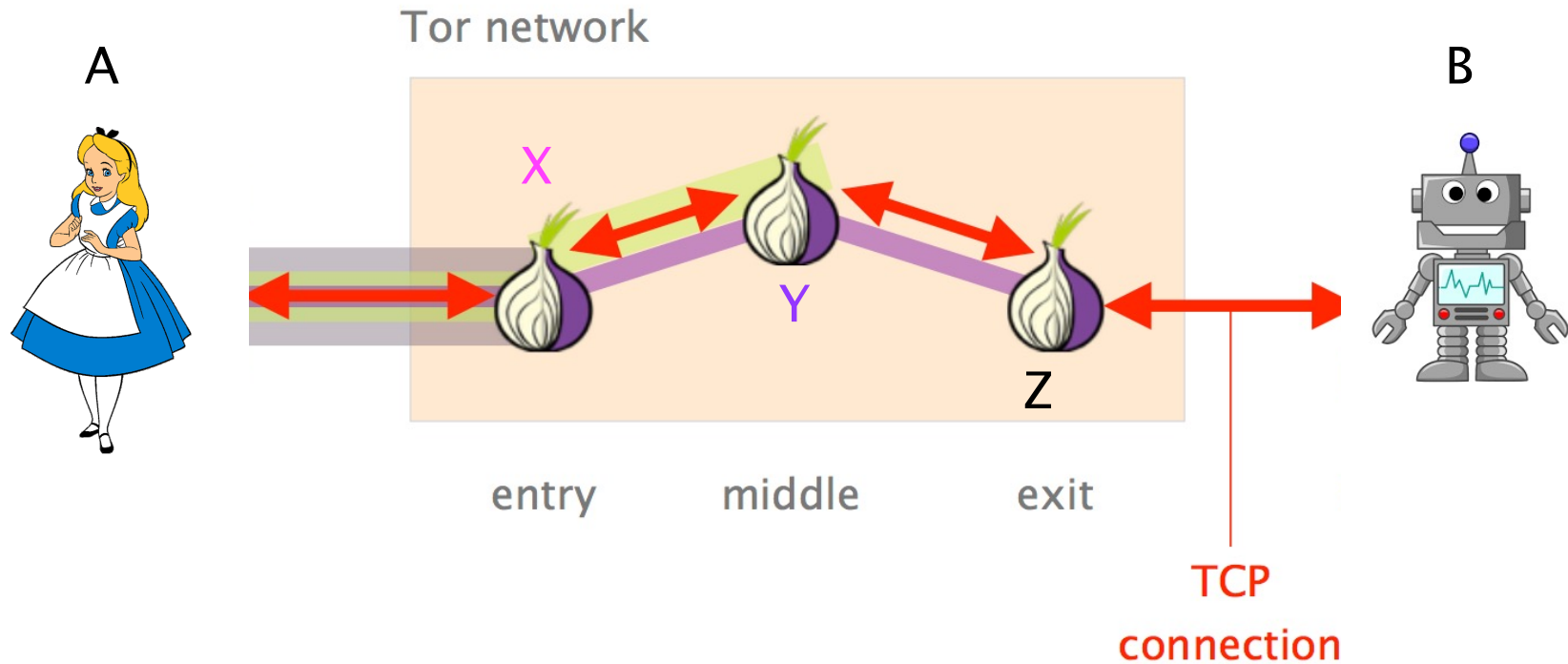
- data field contains A's DH key  $K_{A \rightarrow X}$  encrypted with X's RSA public key

X responds back to A with control torpacket

- data field contains X's DH key  $K_{X \rightarrow A}$

Now both A and X can calculate secret session key  $S_{AX}$  for their link

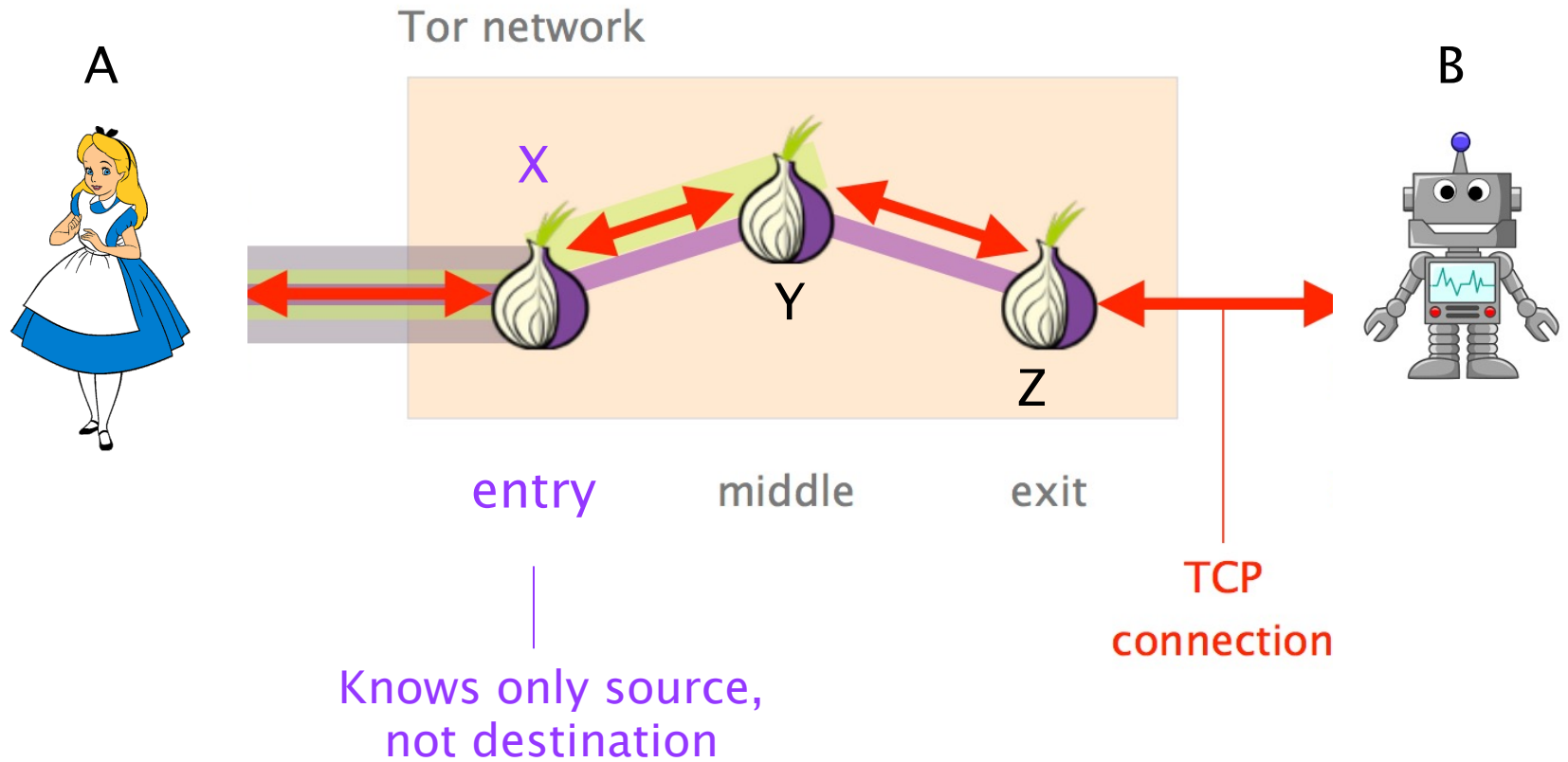
# How A extends circuit to Y



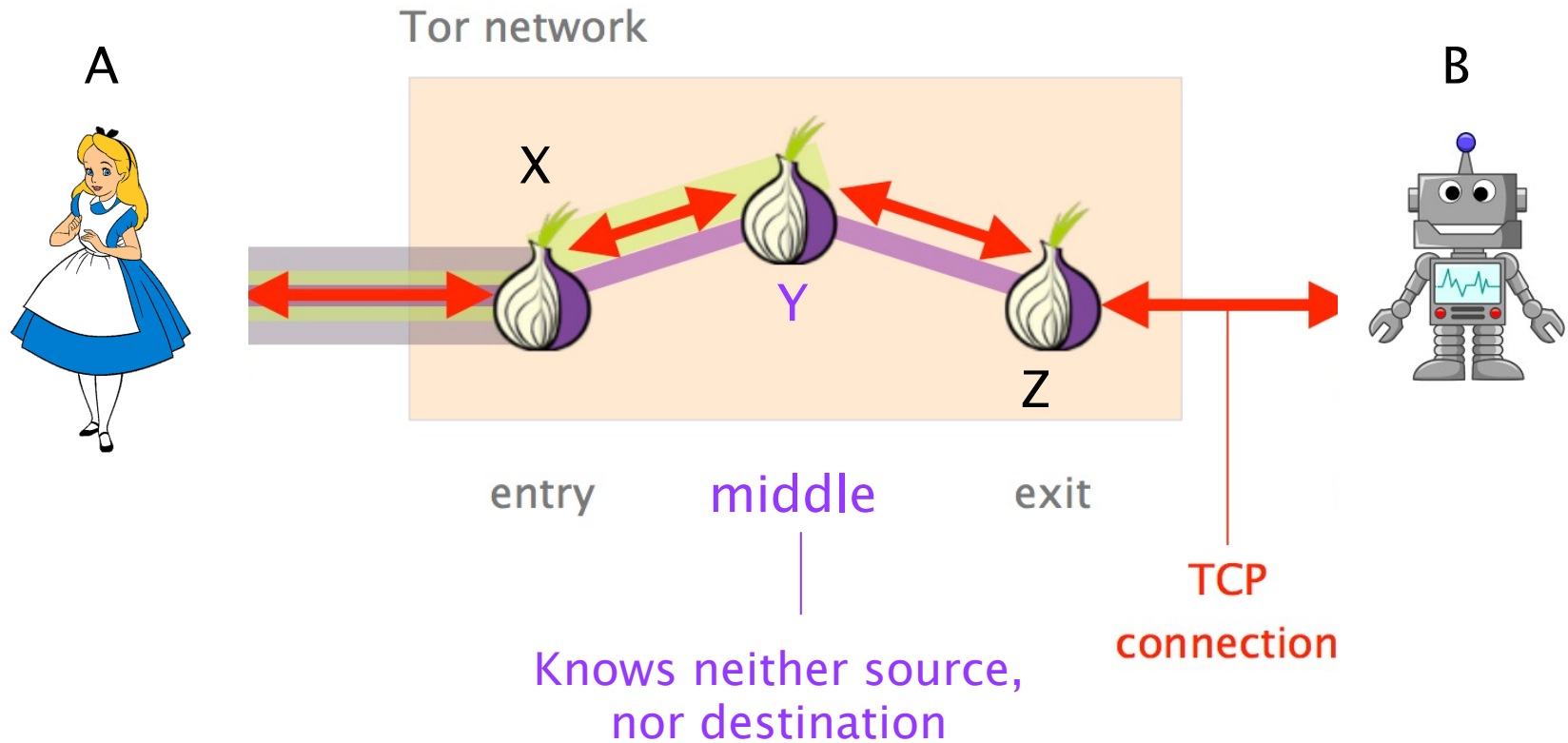
## A sends X relay extend torpacket to extend circuit

- data field includes
  - DH key  $K_{A \rightarrow Y}$  for new terminal node on path, Y, encrypted with Y's RSA public key to prevent X from seeing
  - identity of new node Y
- everything encrypted with session key  $S_{AX}$

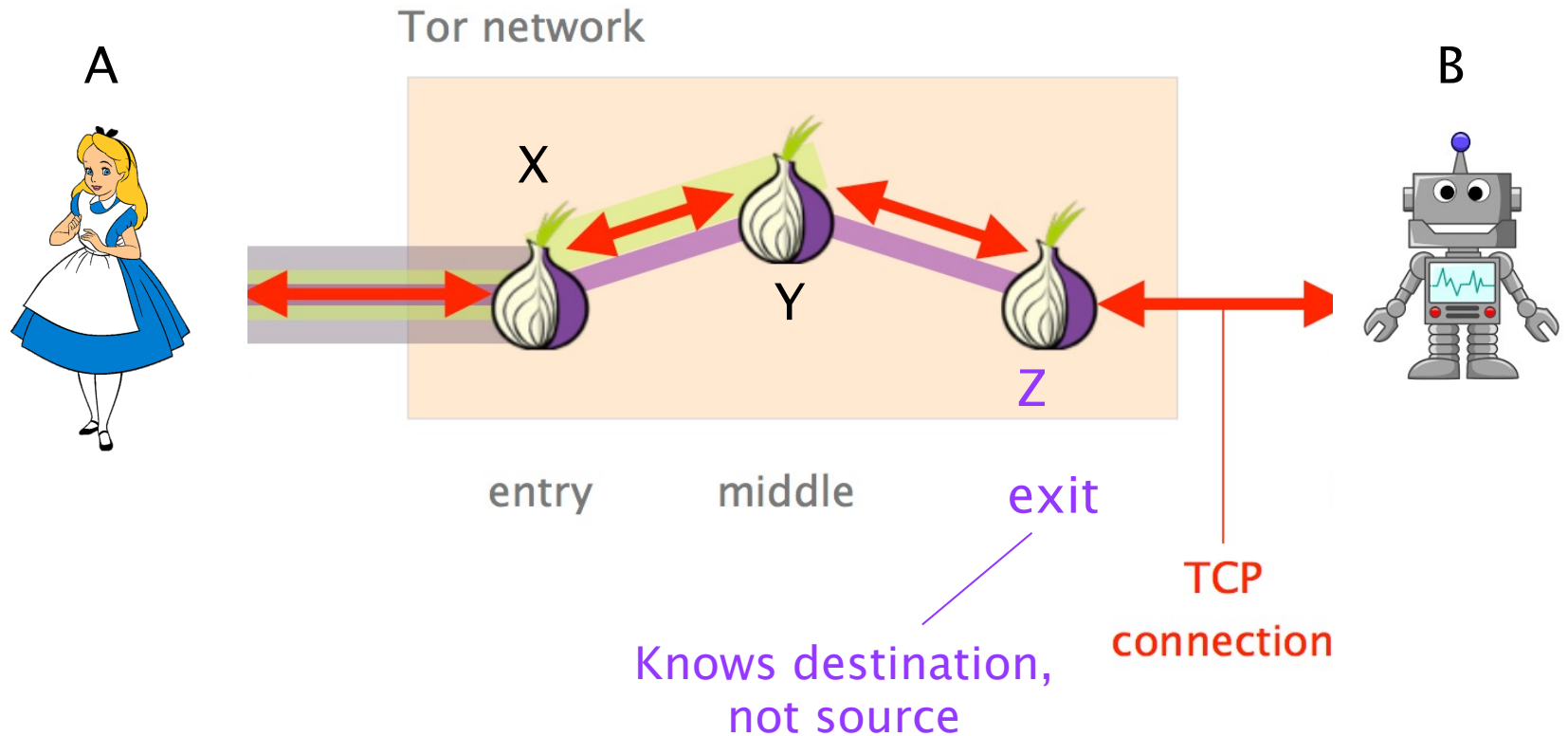
# Each relay has only local knowledge



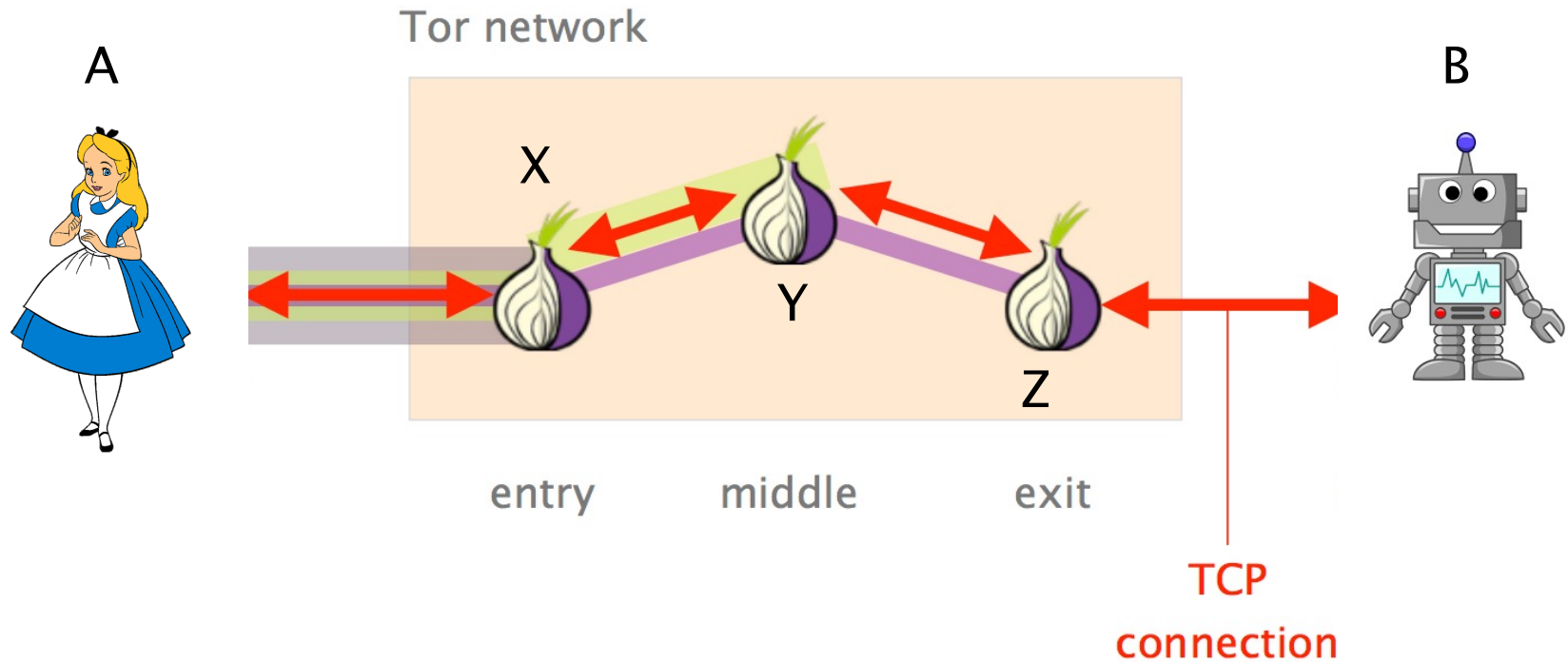
# Each relay has only local knowledge



# Each relay has only local knowledge



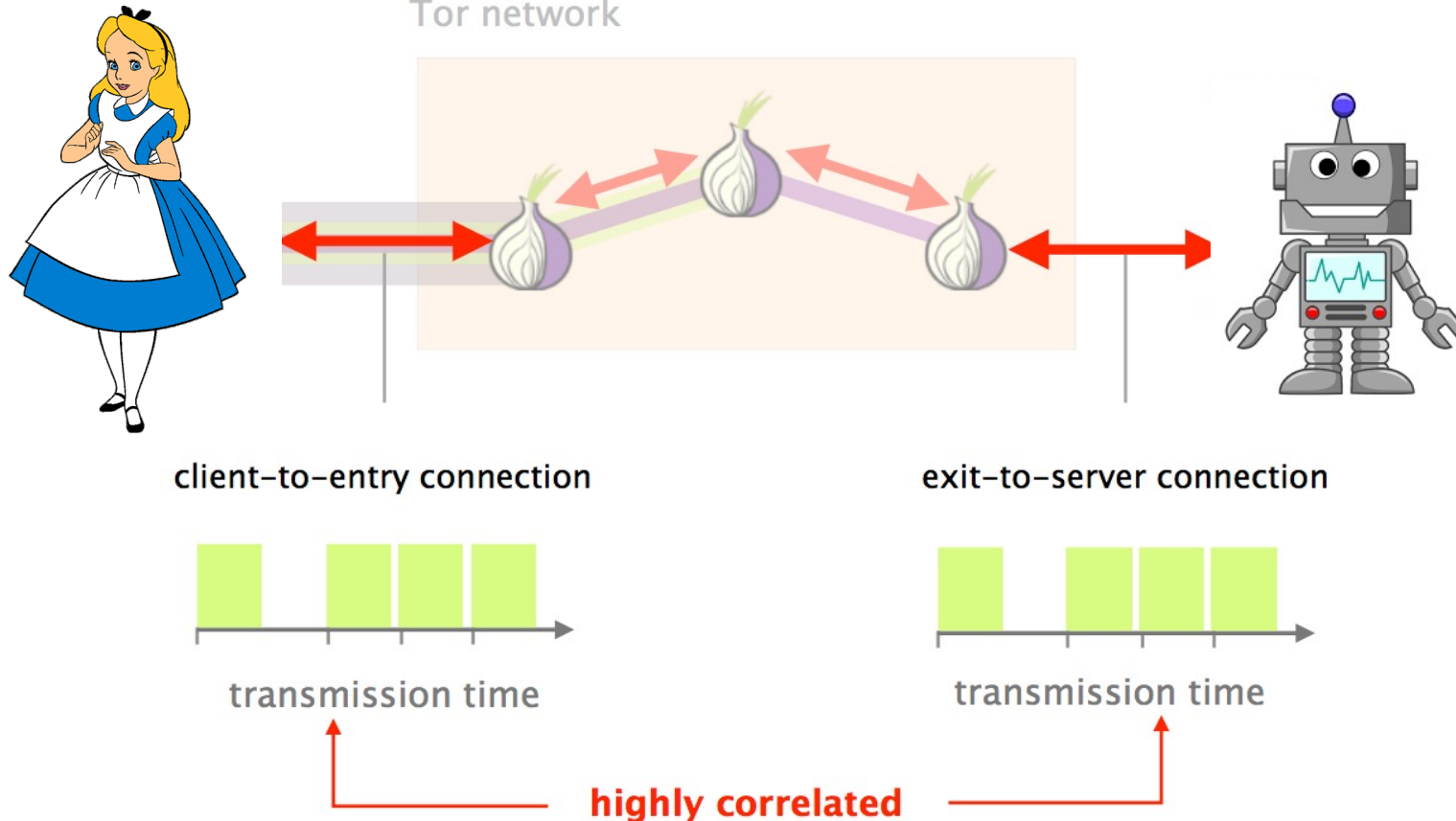
# Anonymous communication



Anonymous communication takes place by forwarding across consecutive tunnels running over TLS underneath

# Tor assumes adversary observes at 1 location

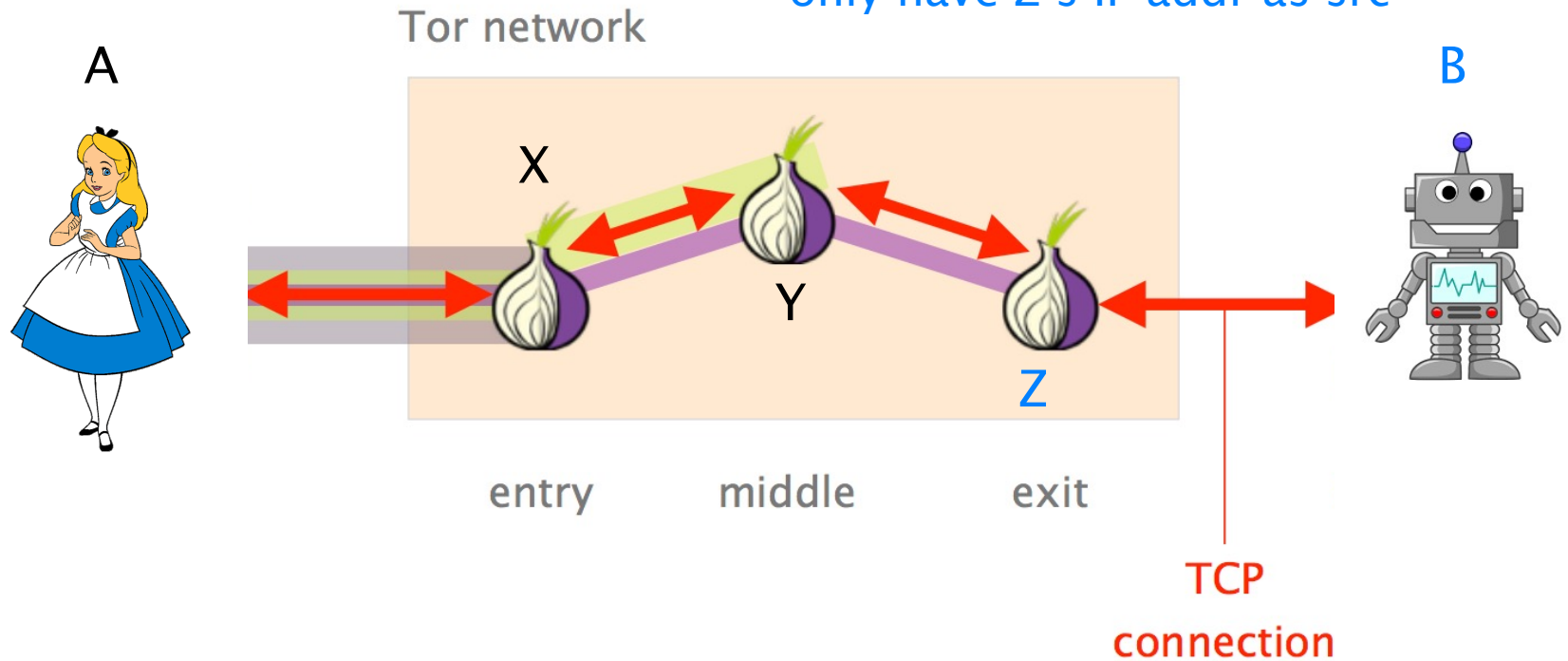
If adversary observes at 2 locations, can break Tor  
because relays don't reorder packets



What kind of adversary might be able to  
observe traffic at 2 locations?

# Can Z see IP address of A?

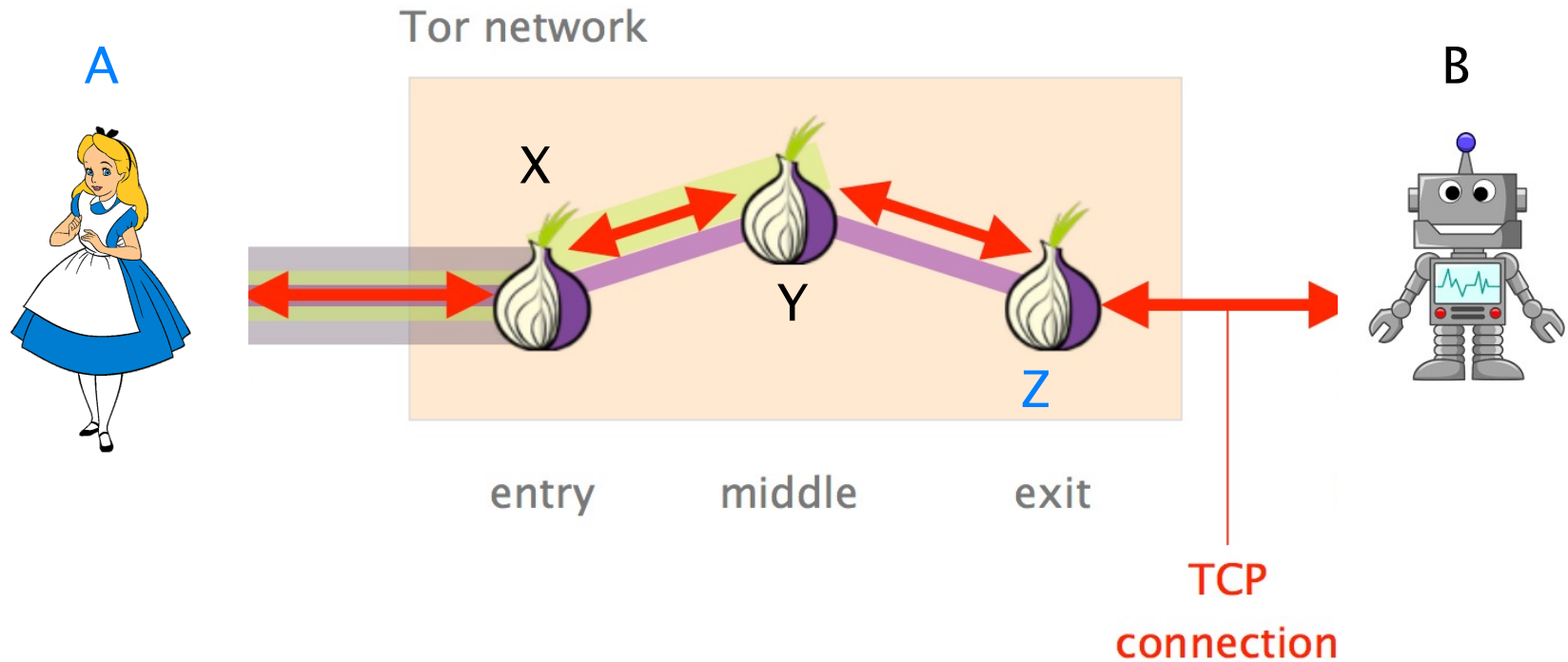
Should not be possible ... packets from Z to B  
only have Z's IP addr as src



## “One Bad Apple Spoils the Bunch”

- by Stevens Le Blond, Pere Manils, Abdelberi Chaabane, Mohamed Ali Kaafar, Claude Castelluccia, Arnaud Legout, and Walid Dabbous
- Takes advantage of peculiarities of BitTorrent protocol to **reveal src IP addr of 10,000 hosts using Tor** for BitTorrent downloads during 23 day period in 2011

# Can Z see data sent to/from A/B?



Not if node A is trying to reach HTTPS site:  
end-to-end encryption of pkt payload

# What else does adversary see?

That user is connecting to Tor onion router

- just using a proxy can attract unwanted attention

What can adversary do?

- drop pkts with dst IP addr associated with Tor onion routers

**What have we covered?**  
**... SINCE THE MIDTERM**

# Transport layer

- Congestion control
- Flow control
- Seq #s and ACKs

... not tested on midterm, so will definitely ask about!

# Network layer

- Router functions
- Internet Protocol
- Addressing
- Link-state vs. distance vector routing
- Intra-domain routing protocols
  - OSPF
- Inter-domain routing protocols
  - BGP

... spent multiple weeks on, so will get multiple questions

# Link layer

- ARP
- Ethernet and frames
- Switches (vs. Routers)

... spent 1 class on. Will definitely get question, but only so much that I can ask about link layer ...

# Security

- Confidentiality
  - symmetric encryption
  - public key encryption
- Authentication
- Message integrity
- TLS
- IPsec
- Tor

... spent multiple weeks on, so will get multiple questions

# What will I definitely not ask you about?

- How to derive keys for public-key cryptography
- ipsec
- Tor

# What should you definitely review?

## TCP finite state machine

- when/why state transitions occur
- what information is important to keep track of

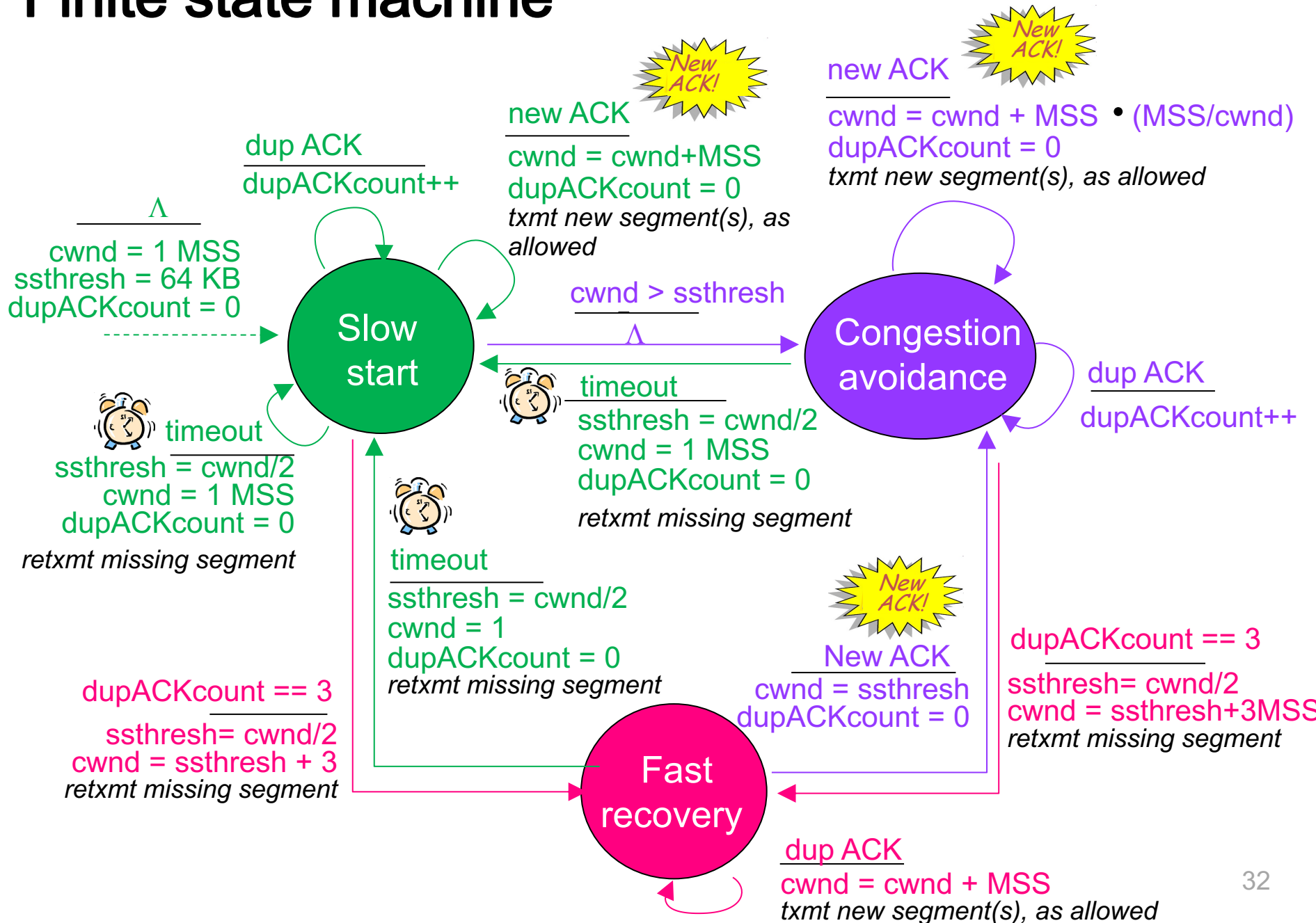
## A day in the life of a web request

- can you fit all protocols we've covered into this?
  - including TLS, BGP, OSPF, ...
  - what lower layer protocols do upper layer protocols run over?
  - what protocols need to execute depending on what info is available?
- think about number of RTTs incurred too

## Network and link layer addressing

- when, why both?
- how is each used?

# Finite state machine



# Final **OVERVIEW**

# Final overview

In our normal classroom on Sat., Dec. 15 from 9a-12p

- closed book, closed notes
- cumulative but covers primarily material in lectures 13 to 25

## 5-6 questions

- transport layer short questions
- network and link layer short questions
- security short questions
- given network, answer questions about it
- design a secure communication protocol
- challenge question

# Problem 1 to 3 – 10 points each

Similar to review questions in book, should only need to write a few sentences to answer

## Problem 1: transport layer short questions

- ~3 in total

## Problem 2: network and link layer short questions

- ~5 in total

## Problem 3: security short questions

- ~3 in total

# Problem 4 – 10 points

## Given network answer questions

- what protocols are used and where?
  - TLS, ARP, DNS, BGP, ...
- how is addressing done?
  - network-layer
  - link-layer
- what if NAT is in use?
- ...

# Problem 5 – 10 points

Given scenario, design secure data transfer protocol

- e.g., suppose that Alice has a data file,  $d$ , that Bob needs.
  1. Alice and Bob want to make sure that if anyone intercepts the file during its transmission, then they cannot understand its content.
  2. Bob also wants to know whether or not whatever is transmitted from Alice to Bob has not been corrupted or altered in transit, and
  3. that the file was sent by Alice. Bob will only need to convince himself of that, no one else
  4. Bob and Alice are computationally limited, so their goal is to transfer the file while meeting criteria 1-4 above, but at the same time, be computationally efficient.

You may assume:

- Symmetric key. Alice and Bob share a secret symmetric key that no one else knows, and Bob and Alice both know that no one else knows it.
- Public keys. There is a public key infrastructure available (e.g., a CA that has Bob and Alice's public keys, and that the public key of the CA is known to Bob and Alice).

# Problem 6 – 5 points

Something to challenge you ...

- haven't yet decided if/what this question will be

**More questions**  
**TEST YOURSELF**

# True or False?

Each network adapter has a unique MAC address

- do switches use these MAC addresses when forwarding frames?

For Ethernet, if a network adapter determines that a frame it has just received is addressed to a different adapter

- it discards the frame without sending an error message to the network layer
- it discards the frame and sends an error message to the network layer
- it delivers the frame to the network layer, and lets the network layer decide what to do
- it sends a NACK (not acknowledged frame) to the sending host

# True or False?

In a distance-vector routing algorithm, each node has a map of the entire network and determines the shortest path from itself to all other nodes in the network.

The network portion of an IP address is the same for all the hosts on the same IP network.

# ... choose one

The link-state algorithm has the following properties:

- it requires the source node to know the costs between every pair of adjacent nodes in the graph
- it determines the shortest path from the source node to all other nodes
- after the kth iteration, the least-cost paths are known to k nodes
- all of the above

The ARP protocol

- runs on top of TCP
- runs on top of UDP
- runs directly on top of IP
- none of the above

**... choose one**

In routing among ASs, which of the following issues dominates:

- geographical distance between ASs
- policy
- number of ASs traversed
- current congestion levels in the ASs

Every autonomous system must use the same intra-autonomous system (domain) routing algorithm.

- True
- False