Anonymity

ECE 422 / CS 461

University of Illinois

Goals

- Understand anonymity and how it relates to other security properties
- Present arguments for and against anonymity online
- Understand how anonymity tools work and how they can be attacked
- Understand how web tracking works and how it can be avoided

Anonymity

- Concealing one’s identity

- Contrast with confidentiality
  - Confidentiality is about contents (**what** was said)
  - Anonymity is about identities (**who** said it and **to whom**)

- Confidentiality **does NOT** mean no information leakage
  - What can be leaked? To whom?

Metadata – Data About Data

- **Who** are the parties communicating?
- **What** was their means of communication?
- **Where** are they? (network or geographic location)
- **When**, how long, and how often did they communicate?
- **How** much data was shared?

“We kill people based on metadata.”

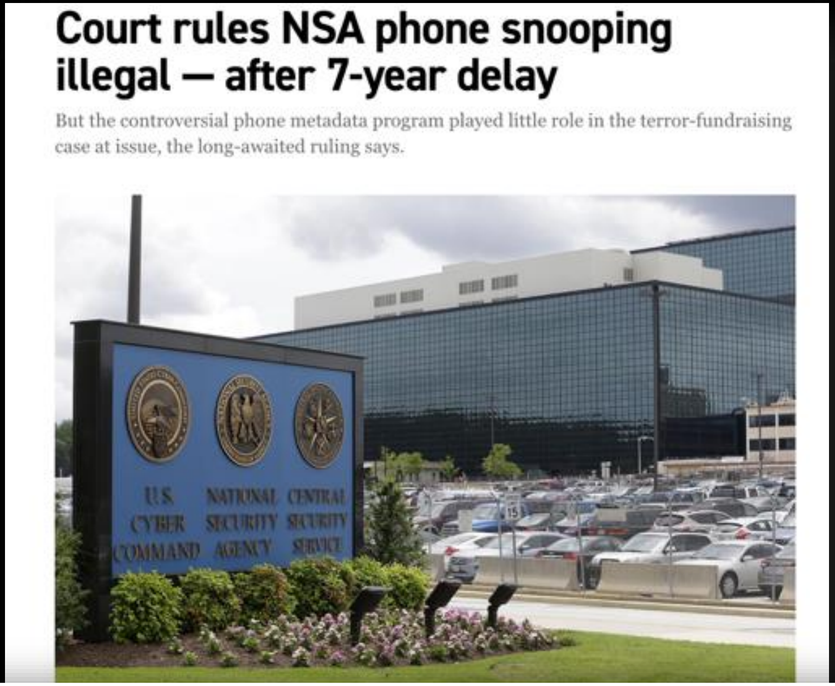
- Michael Hayden, Former Director of NSA and CIA

Is online anonymity a good or bad thing?

Arguments For and Against

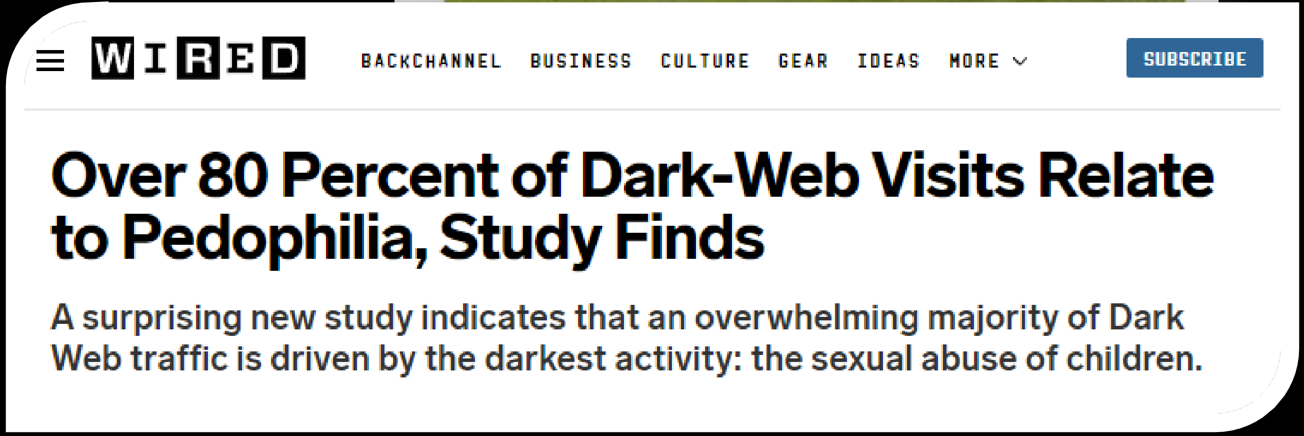
- For:
  - Civil liberties: Freedom from surveillance
  - Protect whistleblowers
  - Prevent user profiling and discrimination

- Against:
  - Illegal and criminal activities
  - Misinformation
  - Toxicity



Court rules NSA phone snooping illegal — after 7-year delay

But the controversial phone metadata program played little role in the terror-fundraising case at issue, the long-awaited ruling says.



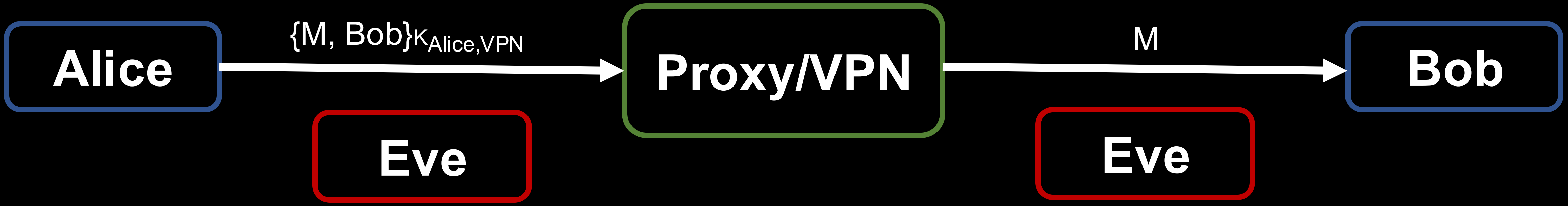
Over 80 Percent of Dark-Web Visits Relate to Pedophilia, Study Finds

A surprising new study indicates that an overwhelming majority of Dark Web traffic is driven by the darkest activity: the sexual abuse of children.

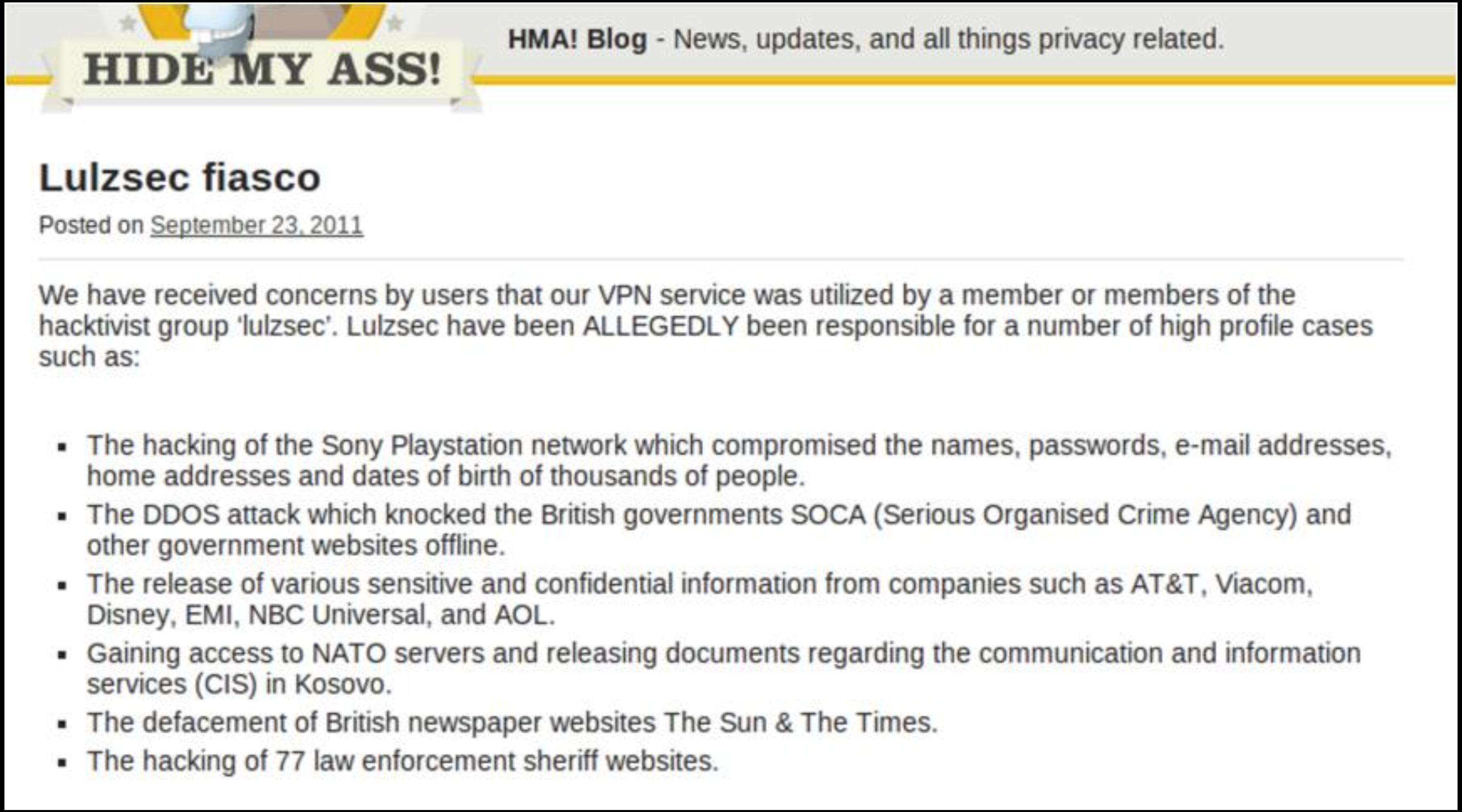
How can anonymity be achieved?

Virtual Private Network (VPN)

- A proxy (intermediary) that relays traffic
  - Alice sends to proxy: message and the destination (both encrypted)
  - Proxy decrypts and forwards message to destination
  - Bob does not learn that M is from Alice
  - Eve (passive eavesdropper) does not learn that Alice is talking to Bob



Proxy and VPN



HMA! Blog - News, updates, and all things privacy related.

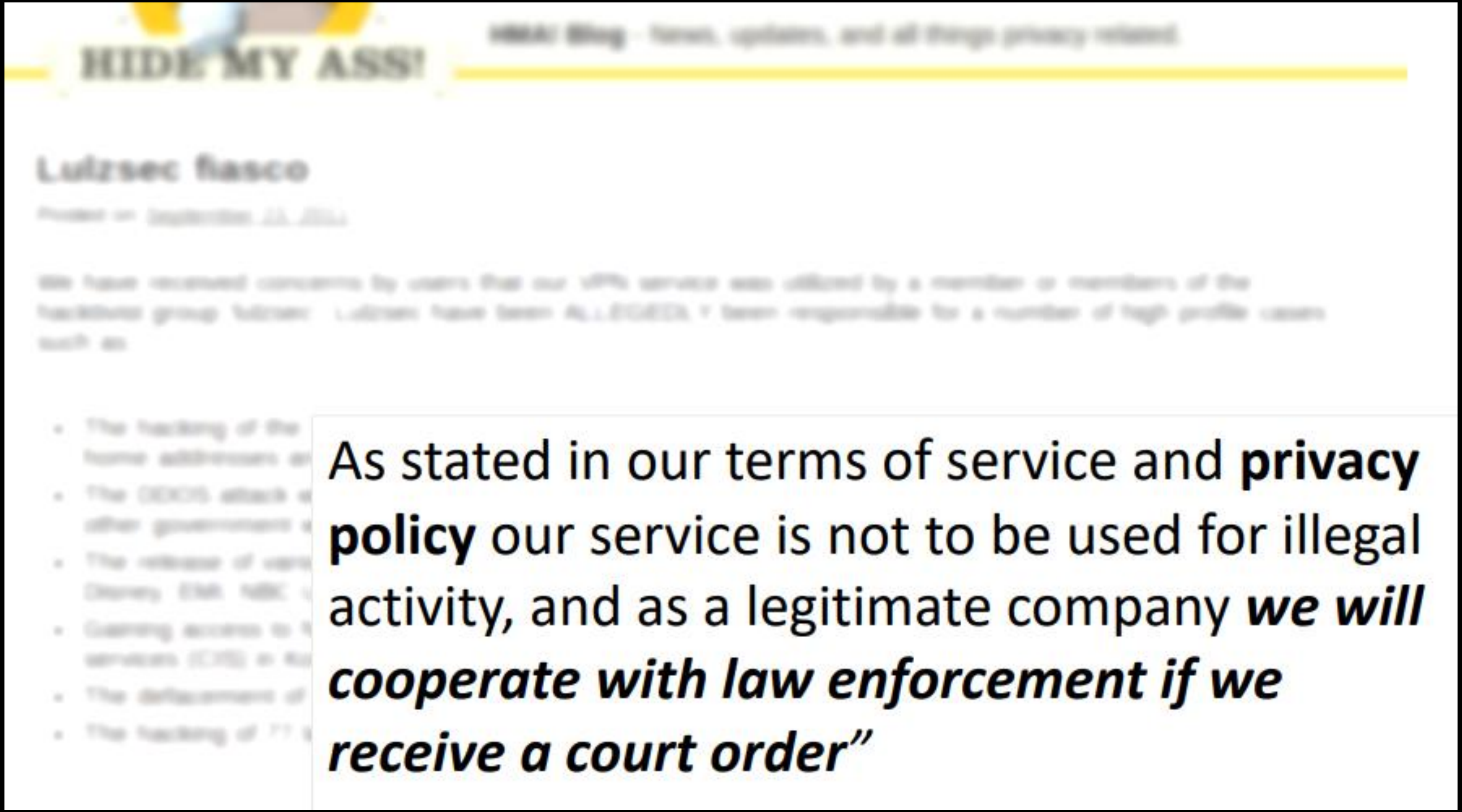
Lulzsec fiasco

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.

Proxy and VPN



HMA! Blog - News, updates, and all things privacy related.

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:

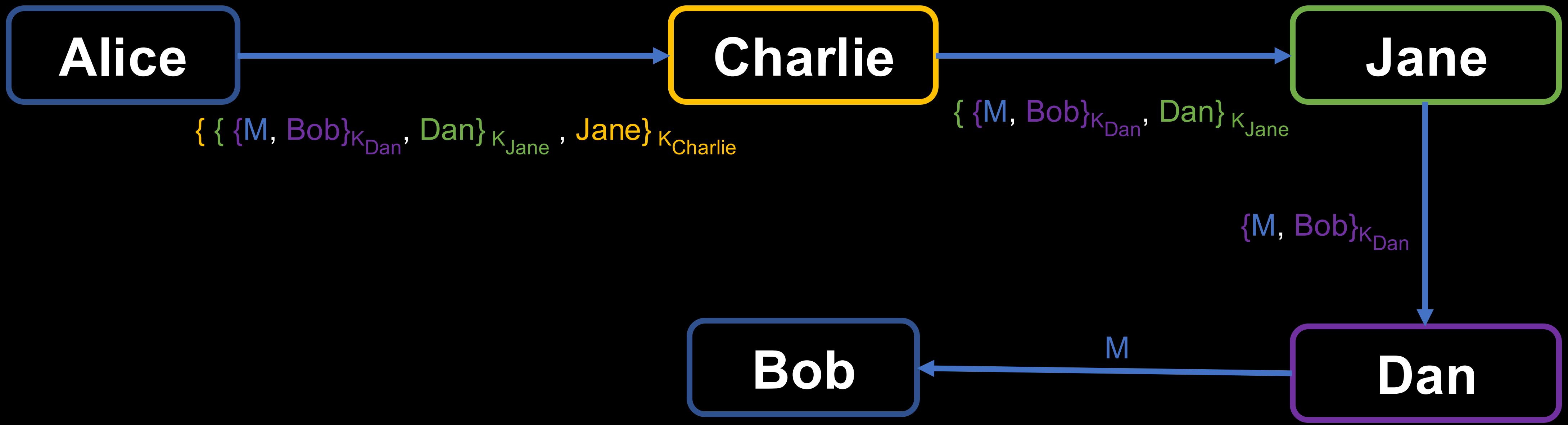
As stated in our terms of service and **privacy policy** our service is not to be used for illegal activity, and as a legitimate company ***we will cooperate with law enforcement if we receive a court order***"

Better Approach: Tor

- Stands for “The Onion Router”
- Idea: Multiple hops of proxies so no single hop knows everything
- Works at the transport layer, allows a user to make TCP connections (usually to websites) without revealing its IP address

Onion Routing

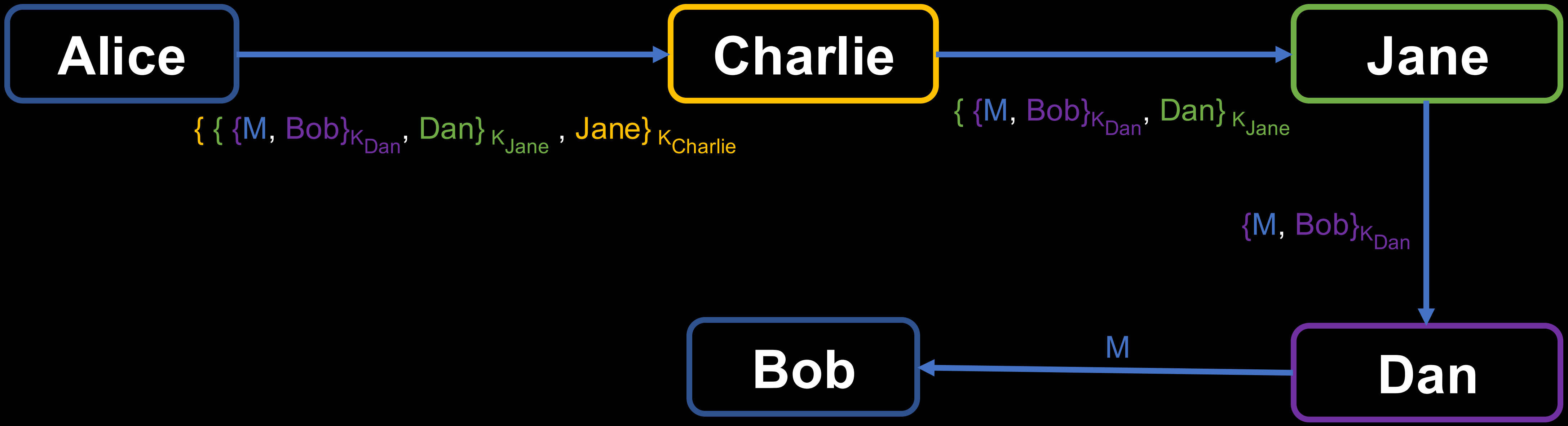
- Alice encrypts in layers, denoted by { }
- Each relays decrypts its layer using key shared only with Alice to see the next hop



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

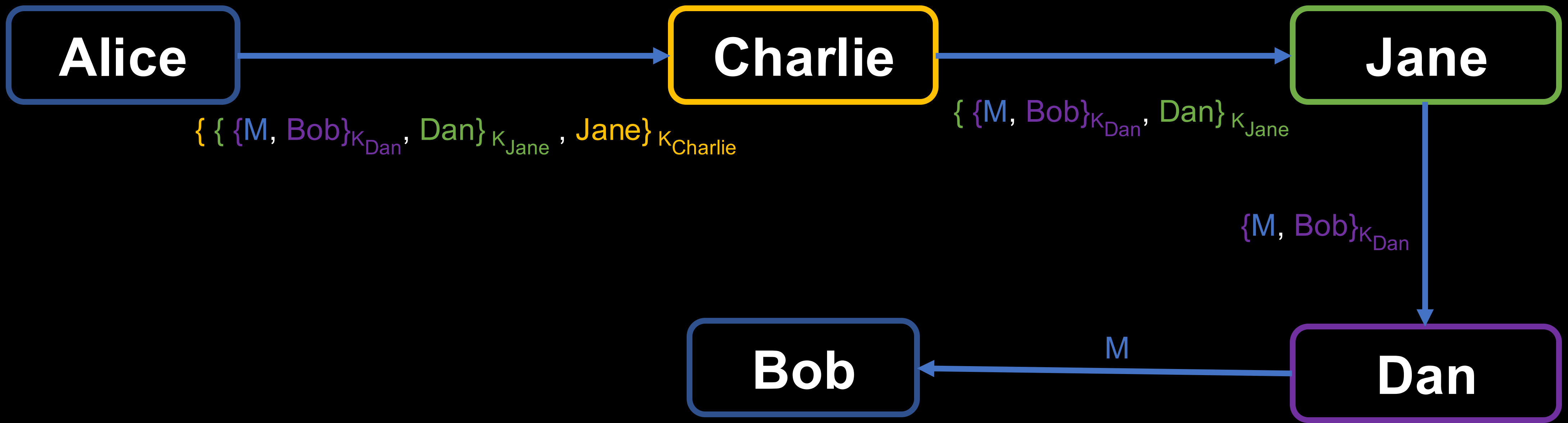
Onion Routing

- Every node knows previous and next hops
- No single node knows both source (Alice) and dest (Bob)


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

Onion Routing

- Entry (“guard”) node Charlie knows: Alice is using Tor, next hop is Jane. Does not know destination (Bob).



Alice

Charlie

Jane

Dan

Bob

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

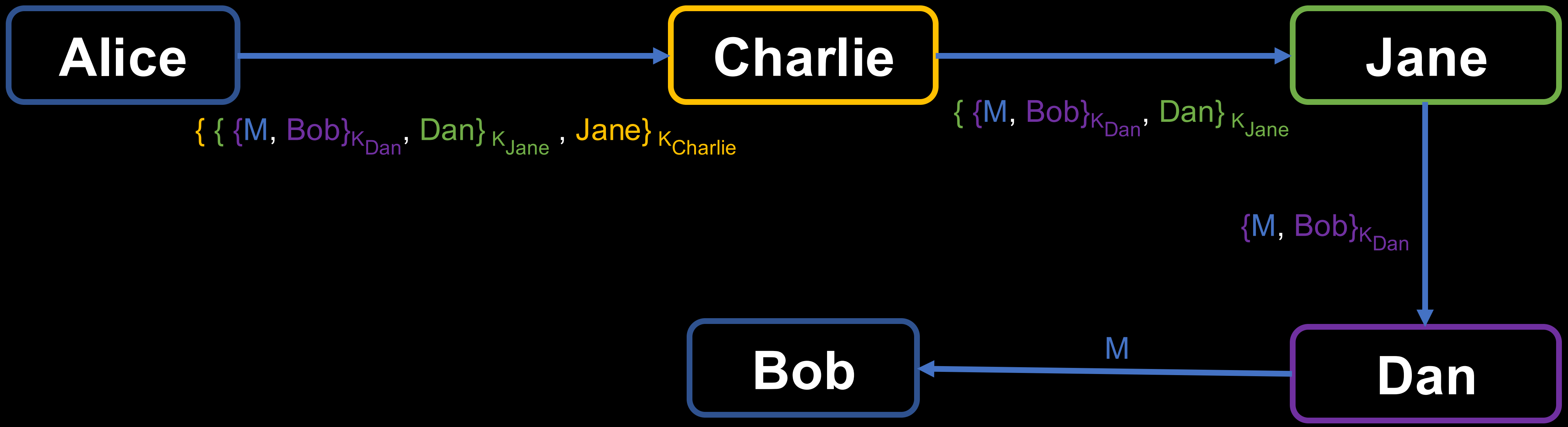
$\{\{M, Bob\}_{K_{Dan}}, Dan\}_{K_{Jane}}$

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

M

Onion Routing

- Middle ("transit") node Jane knows all 3 intermediate nodes



Alice

Charlie

Jane

Dan

Bob

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

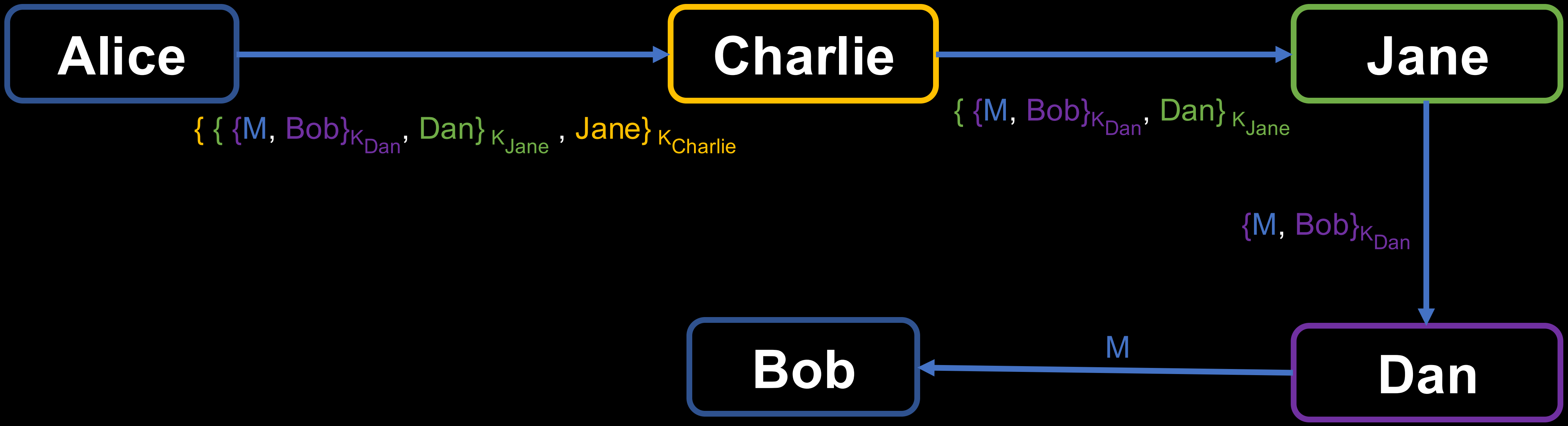
$\{\{M, Bob\}_{K_{Dan}}, Dan\}_{K_{Jane}}$

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

M

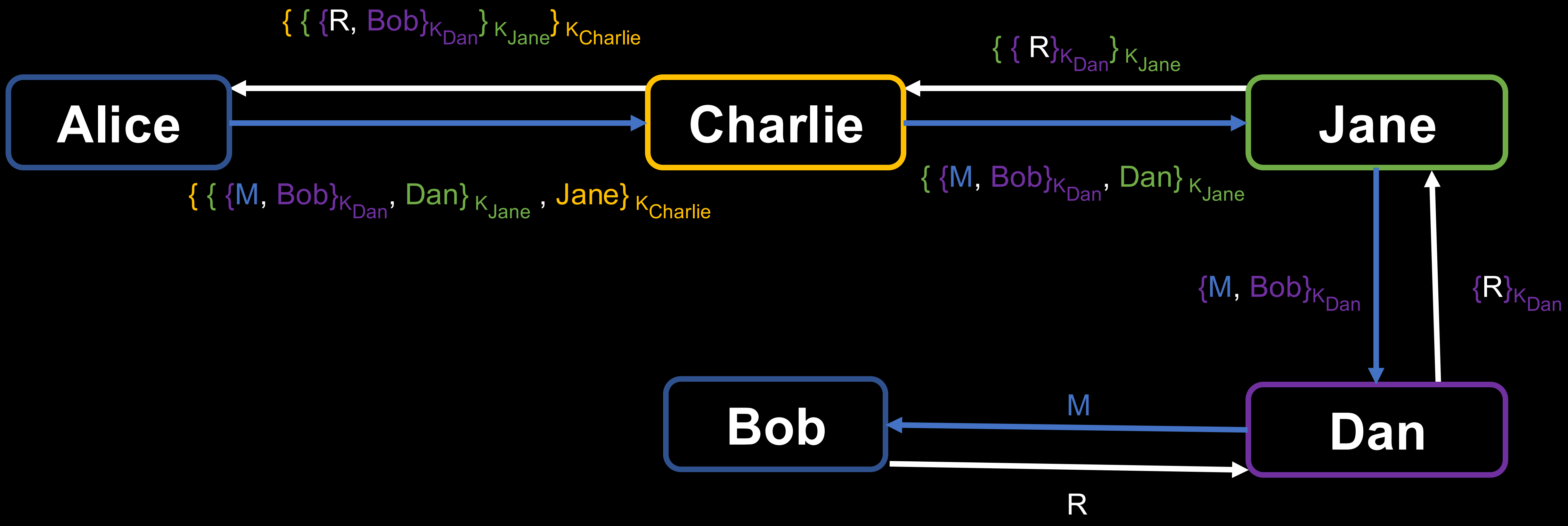
Onion Routing

- Exit node Dan knows: Some Tor user is connecting to Bob, previous hop is Charlie. Does not know source (Alice).


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

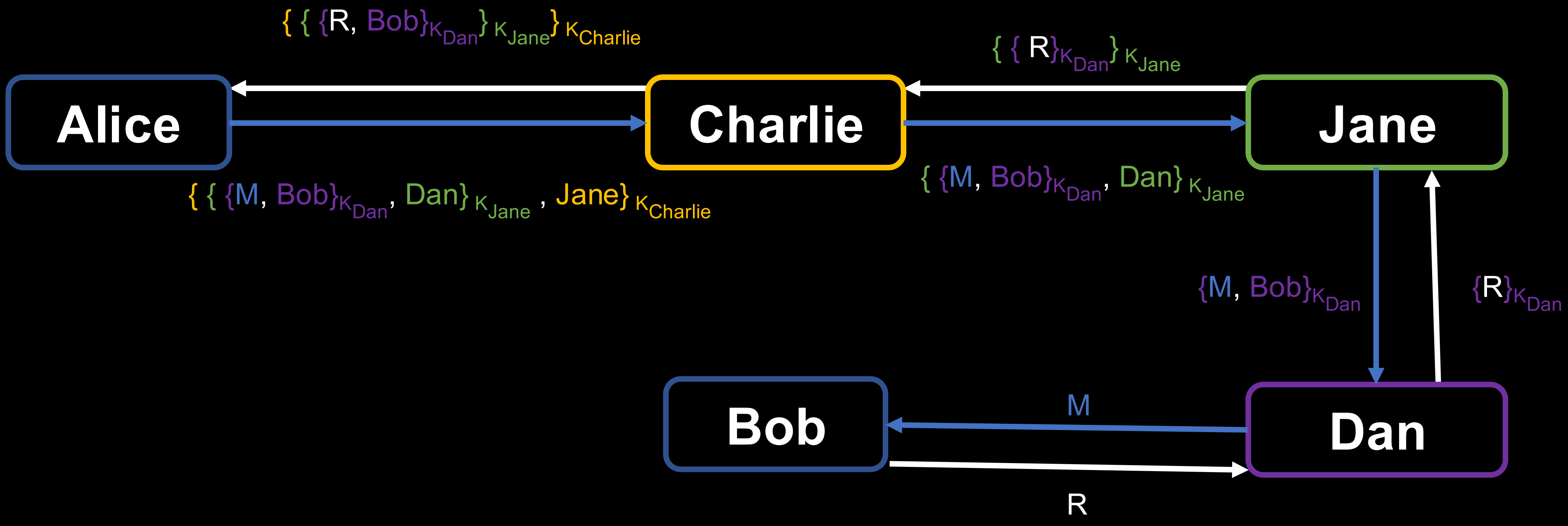
Onion Routing Responses

- Bob sends response R. Each relay adds an encryption layer using their own key (shared with only Alice)


$\{\{\{R, Bob\}_{K_{Dan}}\}_{K_{Jane}}\}_{K_{Charlie}}$
$\{\{R\}_{K_{Dan}}\}_{K_{Jane}}$
Alice
Charlie
Jane
$\{\{\{M, Bob\}_{K_{Dan}}, Dan\}_{K_{Jane}}, Jane\}_{K_{Charlie}}$
$\{\{M, Bob\}_{K_{Dan}}, Dan\}_{K_{Jane}}$
$\{M, Bob\}_{K_{Dan}}$
$\{R\}_{K_{Dan}}$
M
Bob
Dan
R

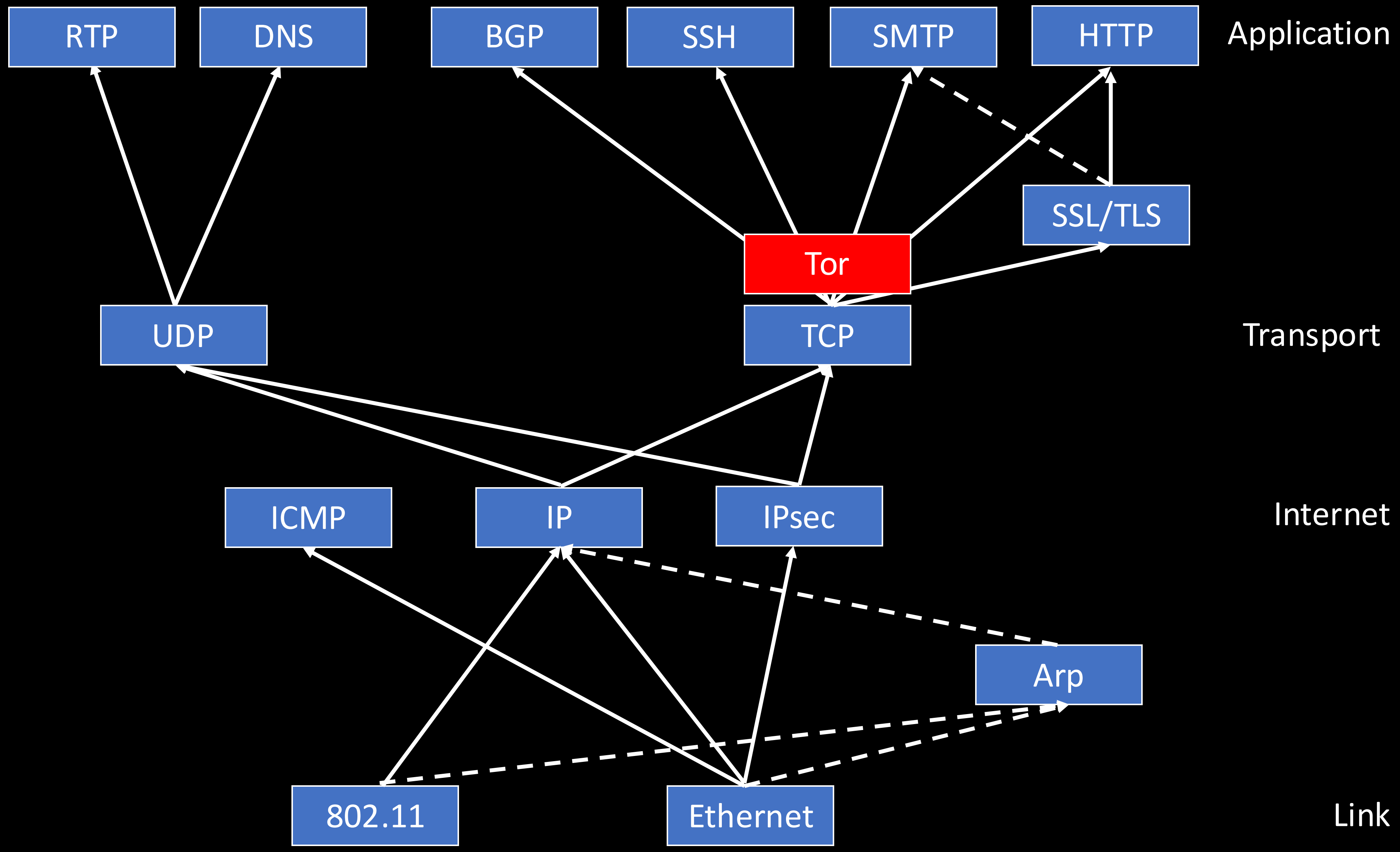
Onion Routing

- Note that Tor does NOT provide encryption between the exit node and the destination. (User can opt to encrypt herself.)



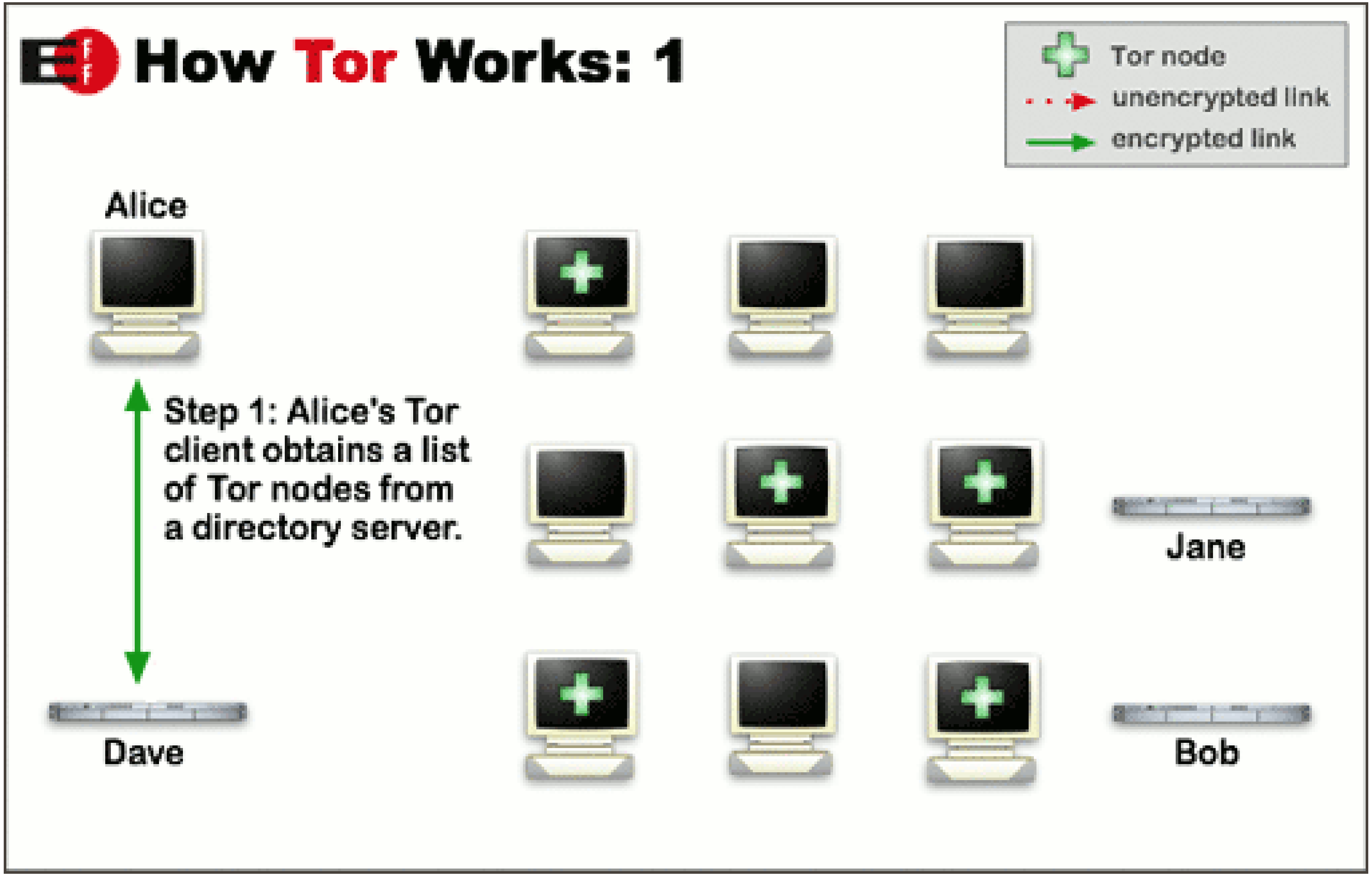
$\{\{\{R, Bob\}_{K_{Dan}}\}_{K_{Jane}}\}_{K_{Charlie}}$
$\{\{R\}_{K_{Dan}}\}_{K_{Jane}}$
Alice
Charlie
Jane
$\{\{\{M, Bob\}_{K_{Dan}}, Dan\}_{K_{Jane}}, Jane\}_{K_{Charlie}}$
$\{\{M, Bob\}_{K_{Dan}}, Dan\}_{K_{Jane}}$
$\{M, Bob\}_{K_{Dan}}$
$\{R\}_{K_{Dan}}$
Bob
M
Dan
R

Tor in Network Layers



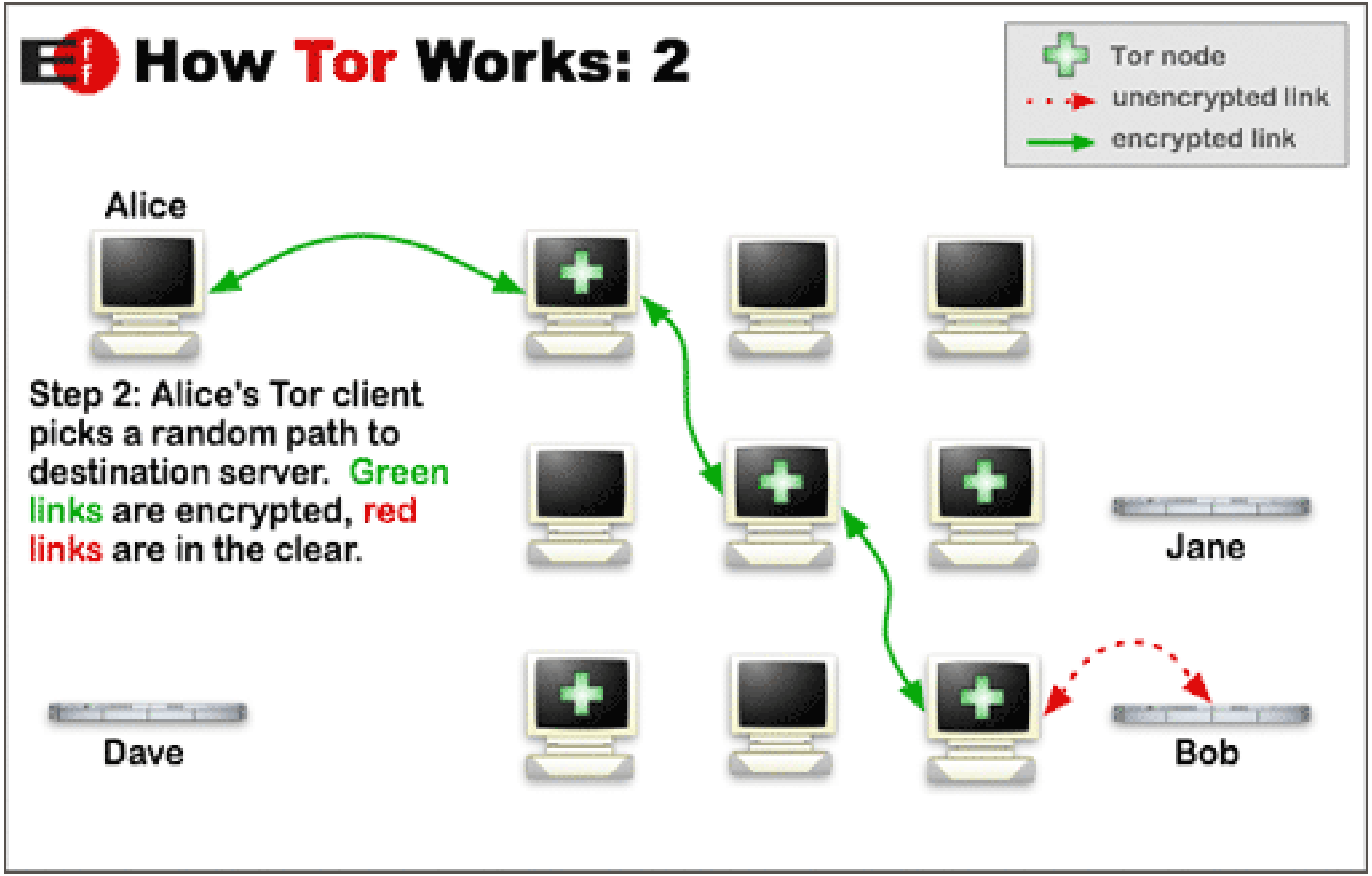
Tor Network

- Made up of volunteer-run nodes (currently around 6,300 worldwide)
- Node listings are available in directory servers (currently 9, run by Tor project)
- A user randomly picks relay nodes for Tor connections. By default, 3 nodes are selected (Tradeoffs?)



How Tor Works: 1

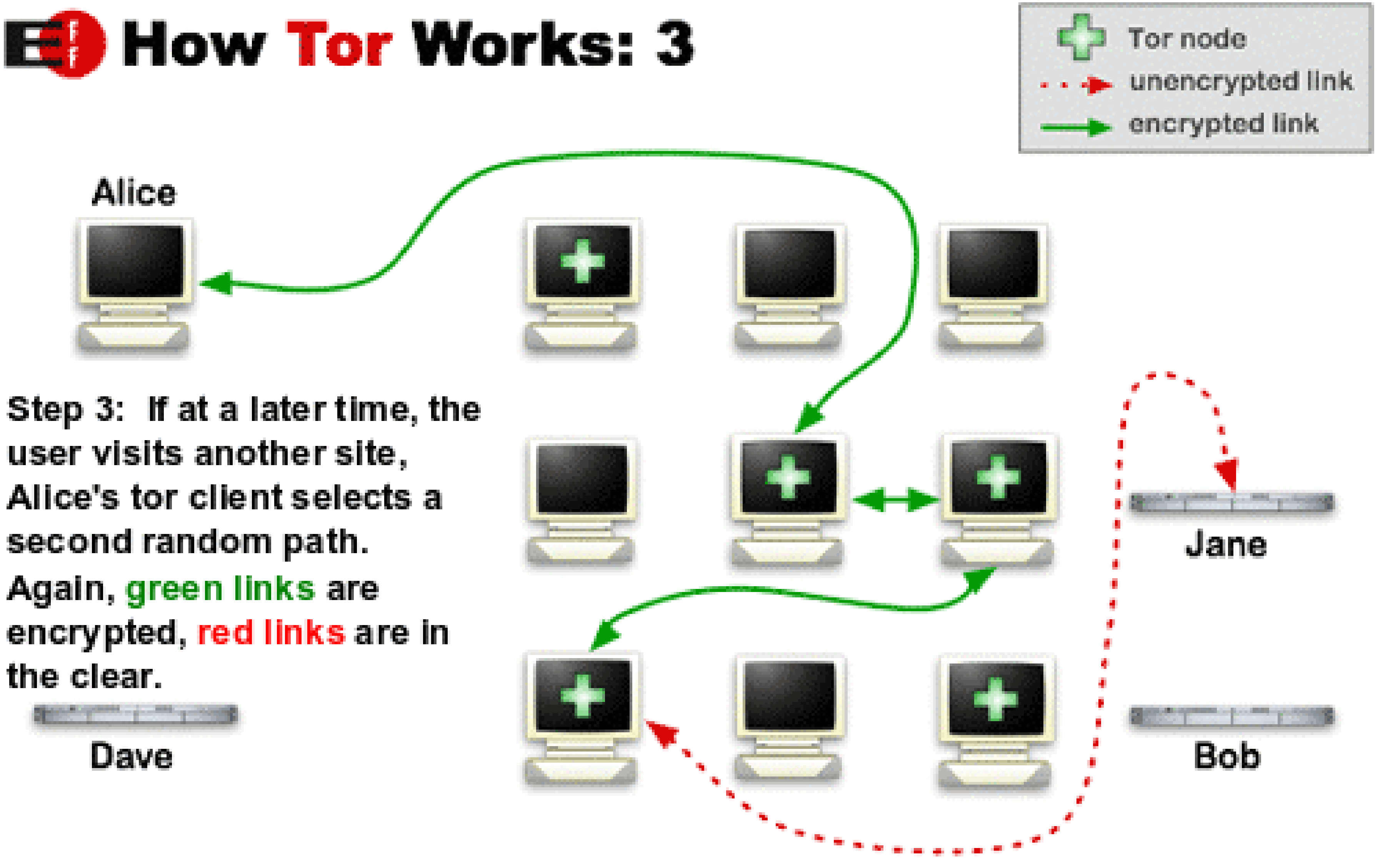
Step 1: Alice's Tor client obtains a list of Tor nodes from a directory server.



How Tor Works: 2

Step 2: Alice's Tor client picks a random path to destination server. Green links are encrypted, red links are in the clear.

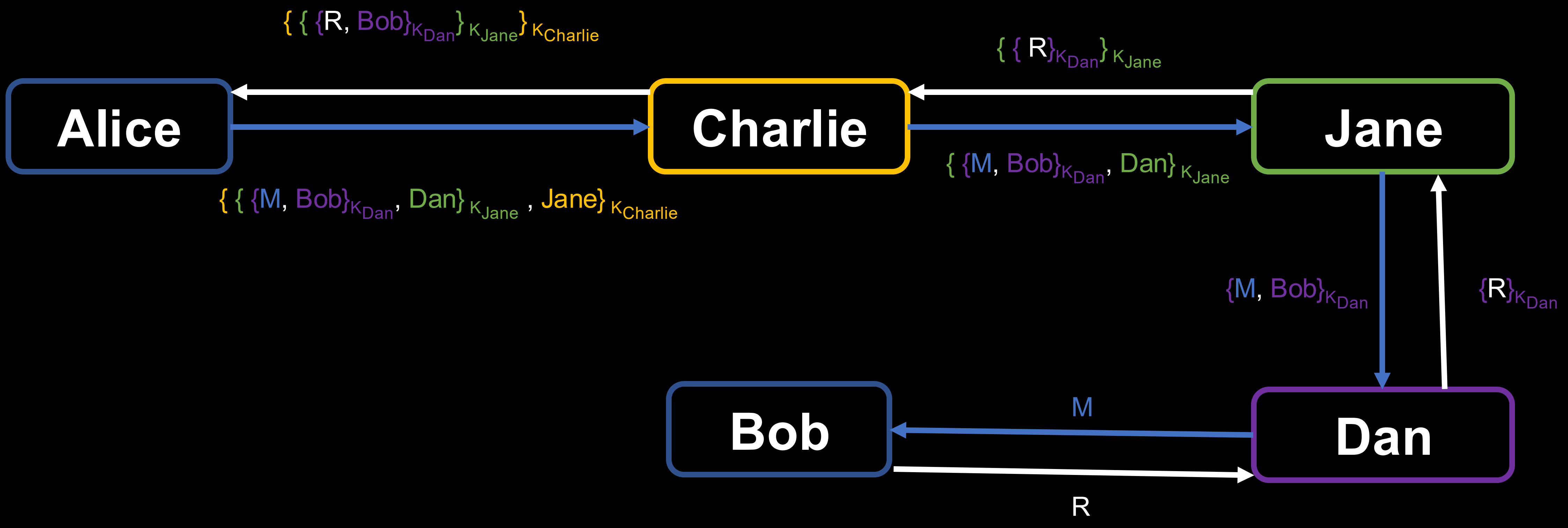
How Tor Works: 3



Step 3: If at a later time, the user visits another site, Alice's tor client selects a second random path. Again, green links are encrypted, red links are in the clear.

Tor Anonymity

- Claim: If *at least one* of the relay nodes is honest, it is hard to link Alice to Bob…. Issues? Attacks?



$\{\{\{R, Bob\}_{K_{Dan}}\}_{K_{Jane}}\}_{K_{Charlie}}$
$\{\{R\}_{K_{Dan}}\}_{K_{Jane}}$
Alice
Charlie
Jane
$\{\{\{M, Bob\}_{K_{Dan}}, Dan\}_{K_{Jane}}, Jane\}_{K_{Charlie}}$
$\{\{M, Bob\}_{K_{Dan}}, Dan\}_{K_{Jane}}$
$\{M, Bob\}_{K_{Dan}}$
$\{R\}_{K_{Dan}}$
M
Bob
Dan
R

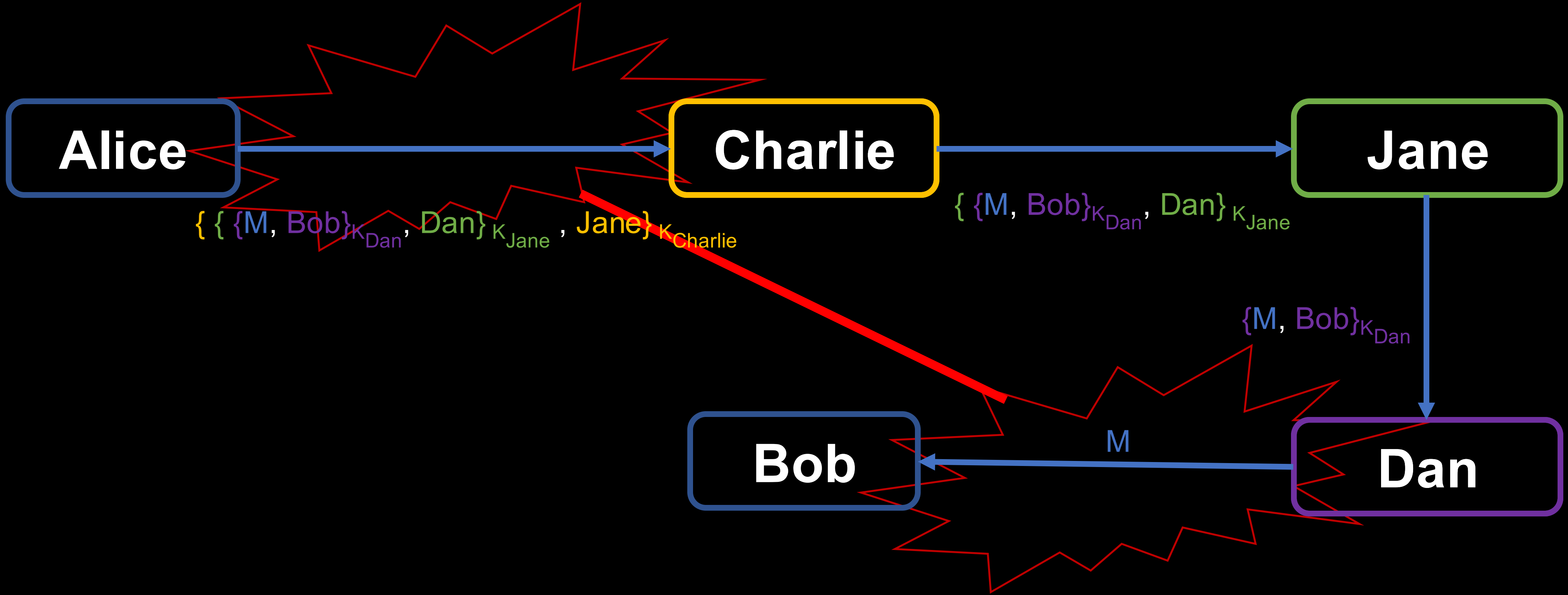
Adversary-Controlled Nodes

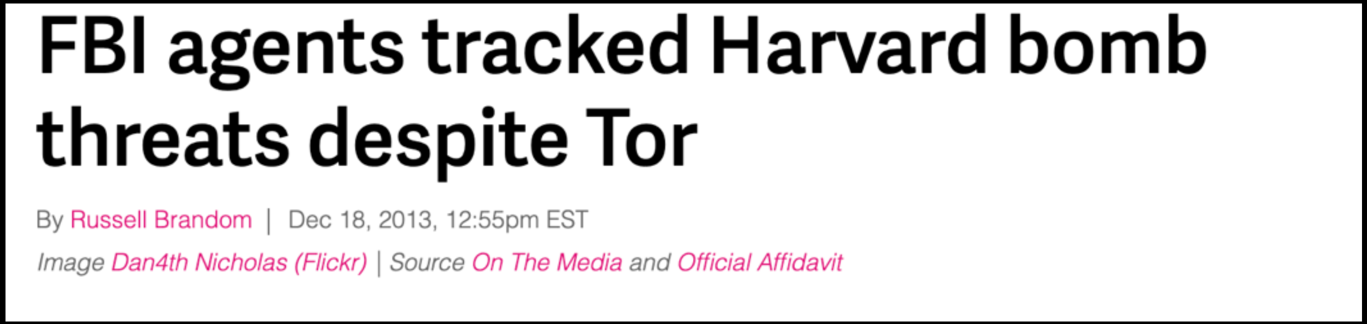
- Adversary volunteers many nodes and hopes that a user picks a path consisting of all its nodes

- Adversary tries to compromise directory server

Side Channel Correlation

- Observing first & last connections allows an attacker to conduct linking using side channels like packet length and timing


Alice
Charlie
Jane
Bob
Dan
{ { {M, Bob}$_{K_{Dan}}$, Dan}$_{K_{Jane}}$, Jane}$_{K_{Charlie}}$
{ {M, Bob}$_{K_{Dan}}$, Dan}$_{K_{Jane}}$
{M, Bob}$_{K_{Dan}}$
M

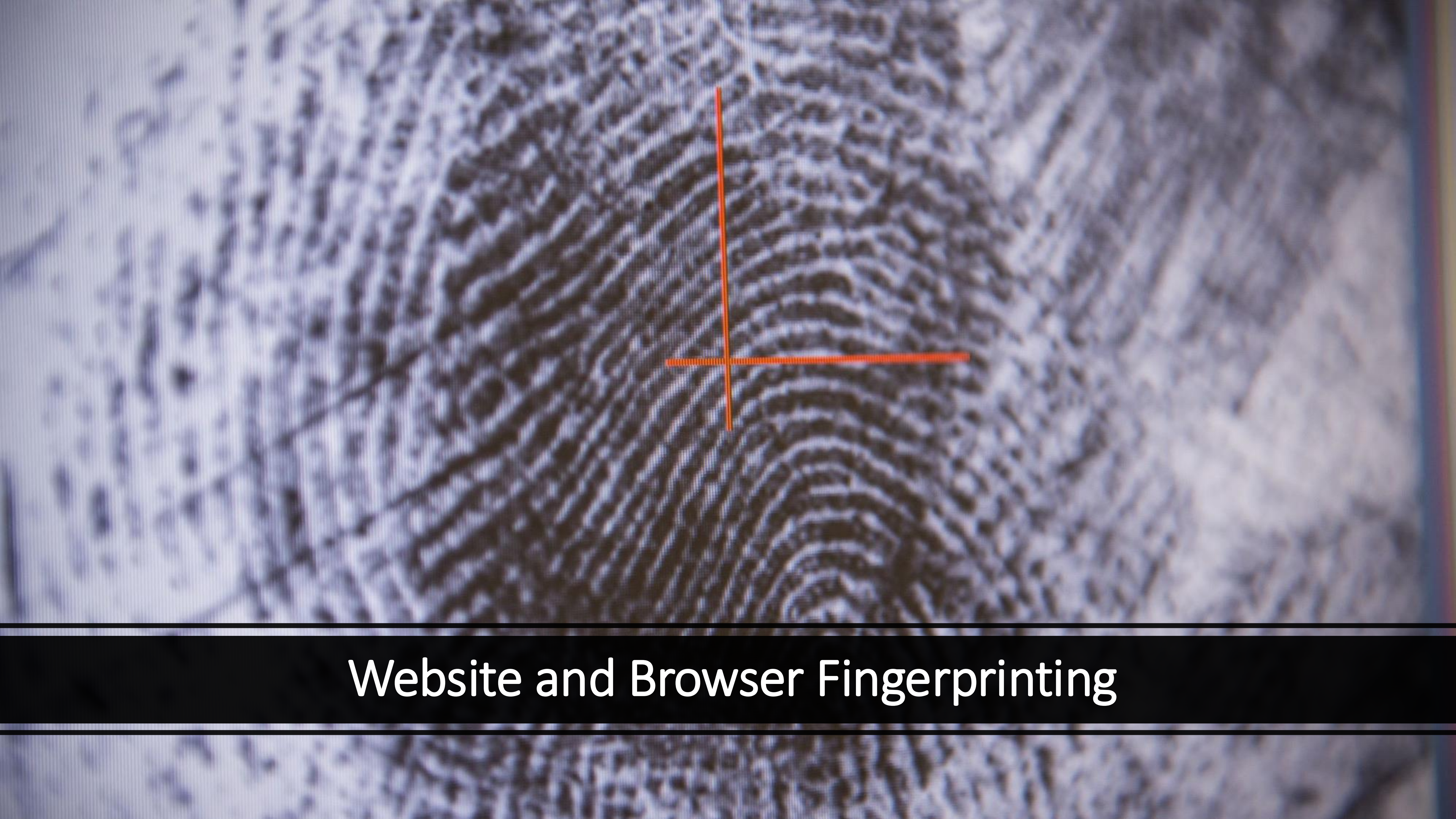


FBI agents tracked Harvard bomb threats despite Tor

By Russell Brandom | Dec 18, 2013, 12:55pm EST

Image Dan4th Nicholas (Flickr) | Source On The Media and Official Affidavit

“Anonymity loves company”



Website and Browser Fingerprinting

Website Fingerprinting

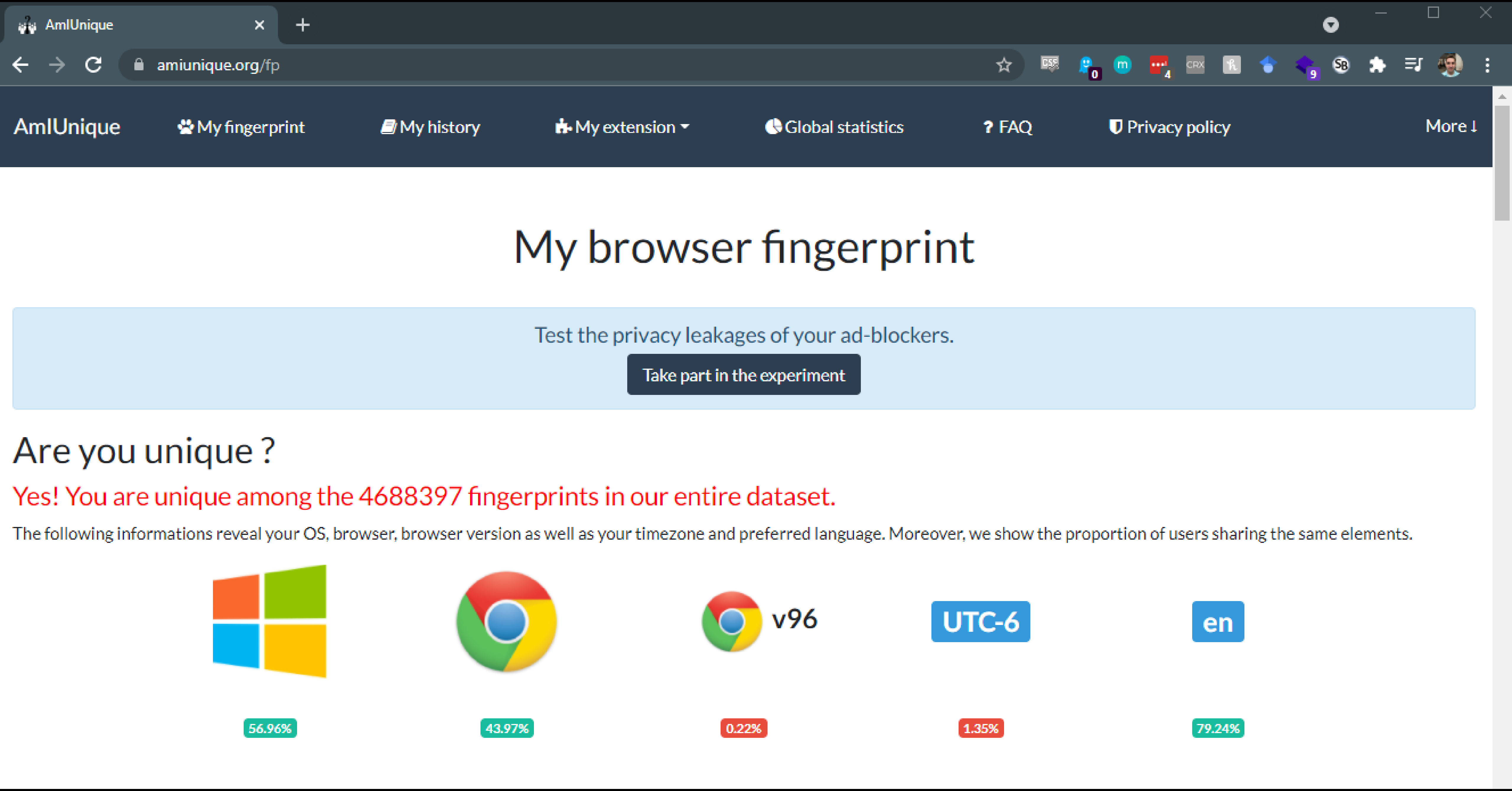
- If we load a website many times, we expect the amount of data transmitted and the timing of that data to be similar

- Machine learning makes these attacks more potent – can factor in packet size, time between packets, number of packets

Browser/User Fingerprinting

- Fundamental challenge: Websites can run code (JavaScript) in the browser (on the computer) of the user

- What information can they gather?



My browser fingerprint

Test the privacy leakages of your ad-blockers.

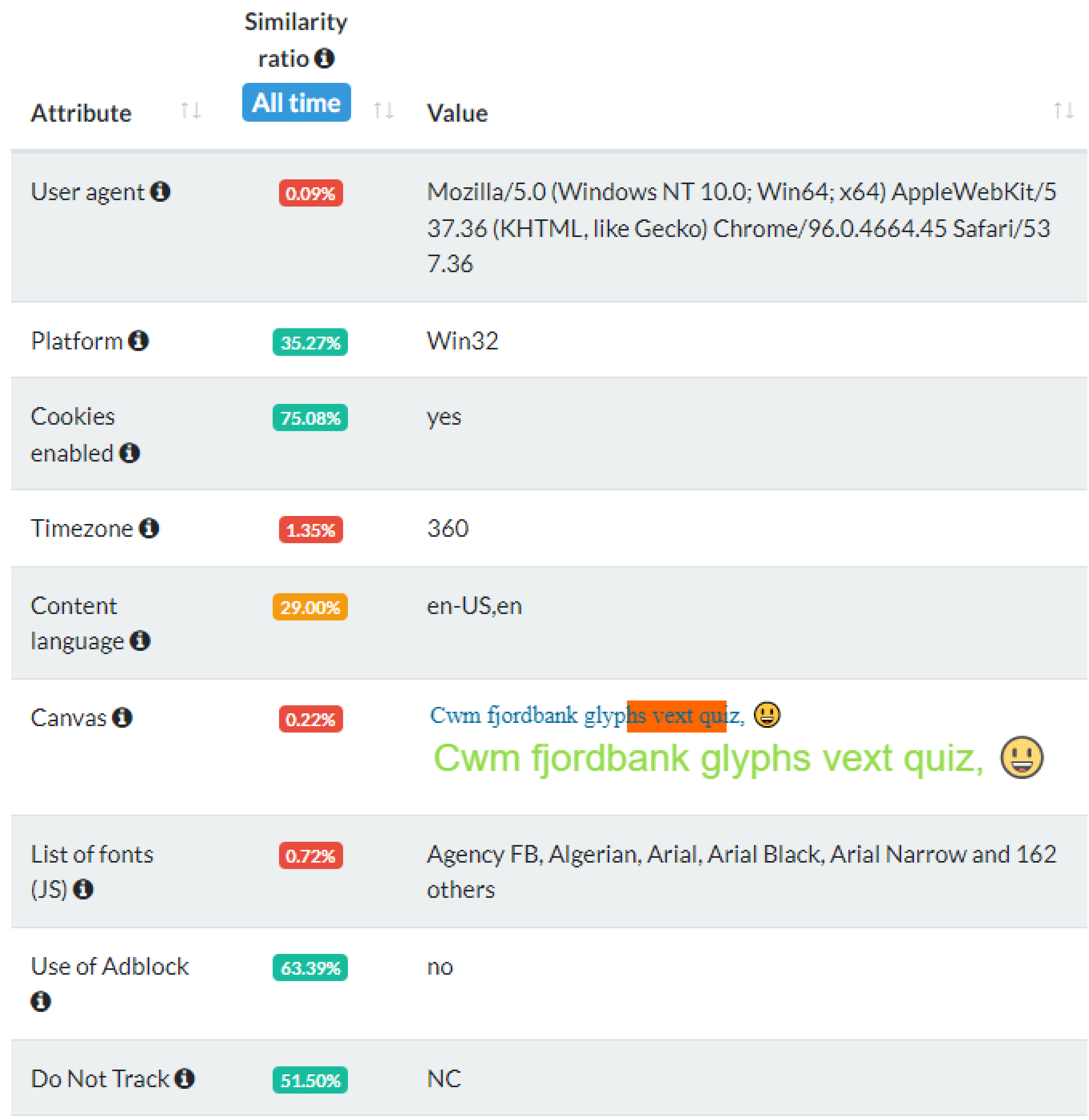
Take part in the experiment

Are you unique ?

Yes! You are unique among the 4688397 fingerprints in our entire dataset.

The following informations reveal your OS, browser, browser version as well as your timezone and preferred language. Moreover, we show the proportion of users sharing the same elements.

56.96% 43.97% 0.22% 1.35% 79.24%



Attribute ↑↓	Similarity ratio ℹ All time ↑↓	Value ↑↓
User agent ℹ	0.09%	Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/96.0.4664.45 Safari/537.36
Platform ℹ	35.27%	Win32
Cookies enabled ℹ	75.08%	yes
Timezone ℹ	1.35%	360
Content language ℹ	29.00%	en-US,en
Canvas ℹ	0.22%	Cwm fjordbank glyphs vext quiz, 😃 Cwm fjordbank glyphs vext quiz, 😃
List of fonts (JS) ℹ	0.72%	Agency FB, Algerian, Arial, Arial Black, Arial Narrow and 162 others
Use of Adblock ℹ	63.39%	no
Do Not Track ℹ	51.50%	NC

Preventing Browser Fingerprinting

- Block scripts that are known to do it

- Disable JavaScript entirely

- Tor Browser supports these settings … and provide “company”

Summary

Anonymity is about concealing one’s identity

There are arguments for and against

State of the art solutions: Proxies/VPNs and Tor

Achieving anonymity is hard against a sophisticated attacker because of side channels

“Anonymity loves company”