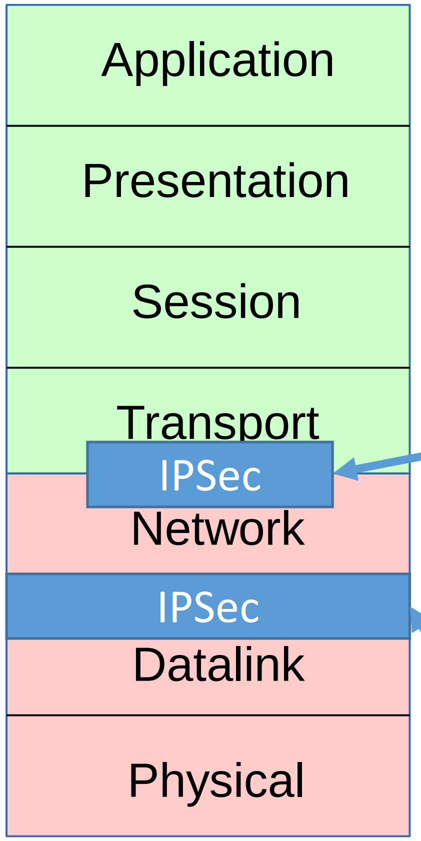
Last week – IPSec

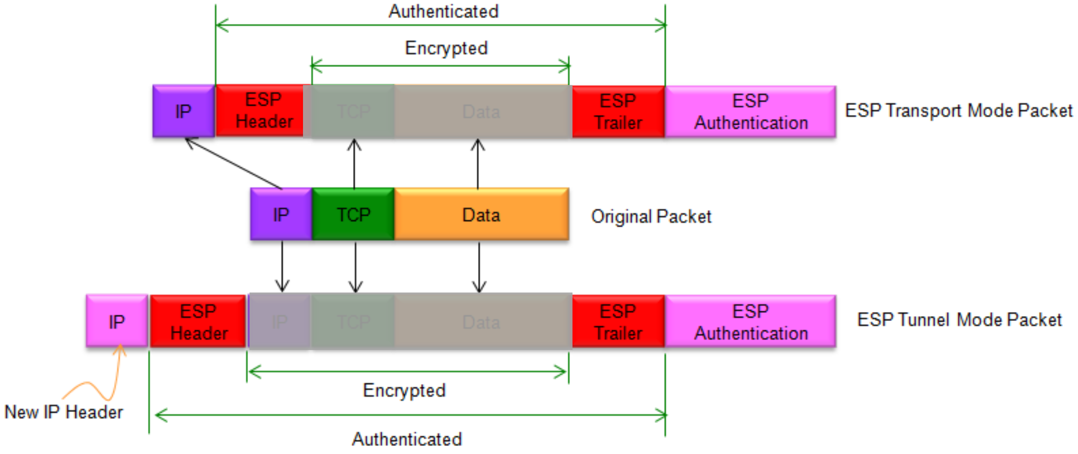


Open Systems
Initiative
(OSI) Model '94

Cryptographic protection at the IP level

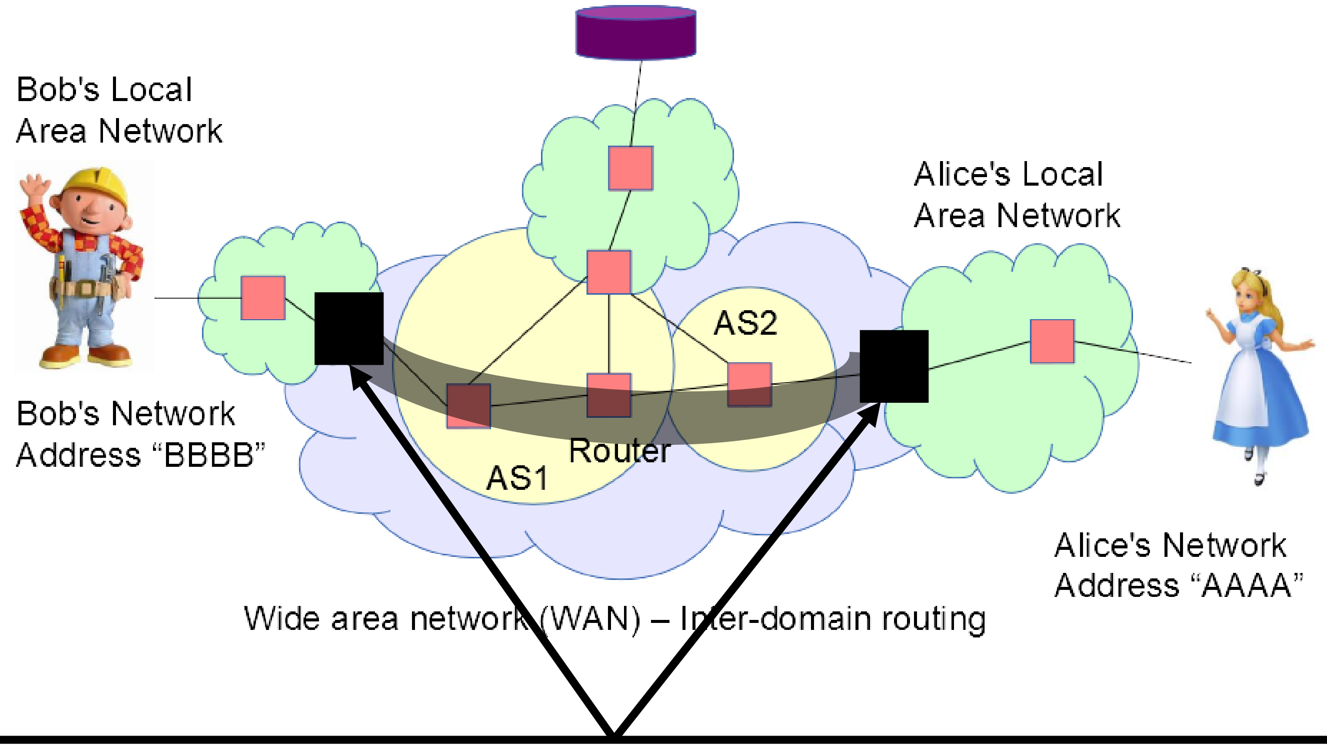
- Key exchange based on public key cryptography or shared symmetric keys
- **Authentication Header (AH)**: authentication & integrity (HMAC), protection from replay attacks (sequence number)
- **Encapsulating Security Payload (ESP)**: confidentiality

IPSec in **TRANSPORT MODE**, **encrypts payload** but keeps the headers.



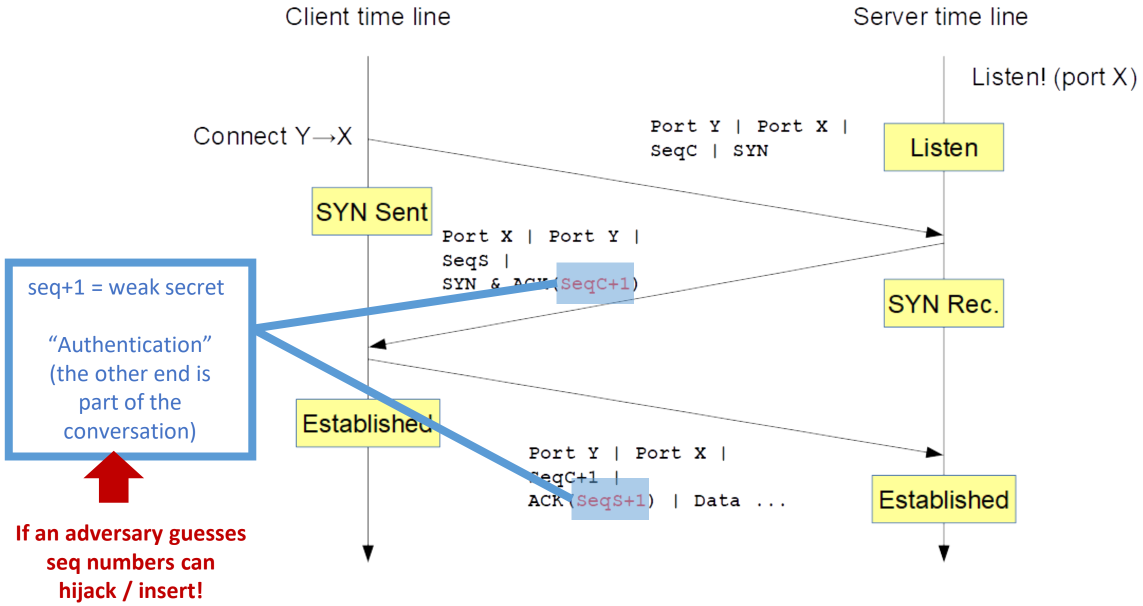
IPSec in **TUNNEL MODE**, **encrypts payload** and **the headers**.

Last week – VPN



- Builds on IPSec in tunnel mode
  - Looks like one single network (Bob routes to Alice as if it was a LAN)
  - Inside VPN “tunnel” fully protected packets: confidentiality, authentication, integrity, reply

Last week – TCP Hijack


Client time line
Server time line
Listen! (port X)
Connect Y→X
Port Y | Port X | SeqC | SYN
Listen
SYN Sent
Port X | Port Y | SeqS | SYN & ACK(SeqC+1)
SYN Rec.
Established
Port Y | Port X | SeqC+1 | ACK(SeqS+1) | Data ...
Established

seq+1 = weak secret

“Authentication” (the other end is part of the conversation)

If an adversary guesses seq numbers can hijack / insert!

Who: a man in the middle adversary (MITM)

- can observe communication
- can intercept and inject packets

What:

1- Wait for TCP session to be established

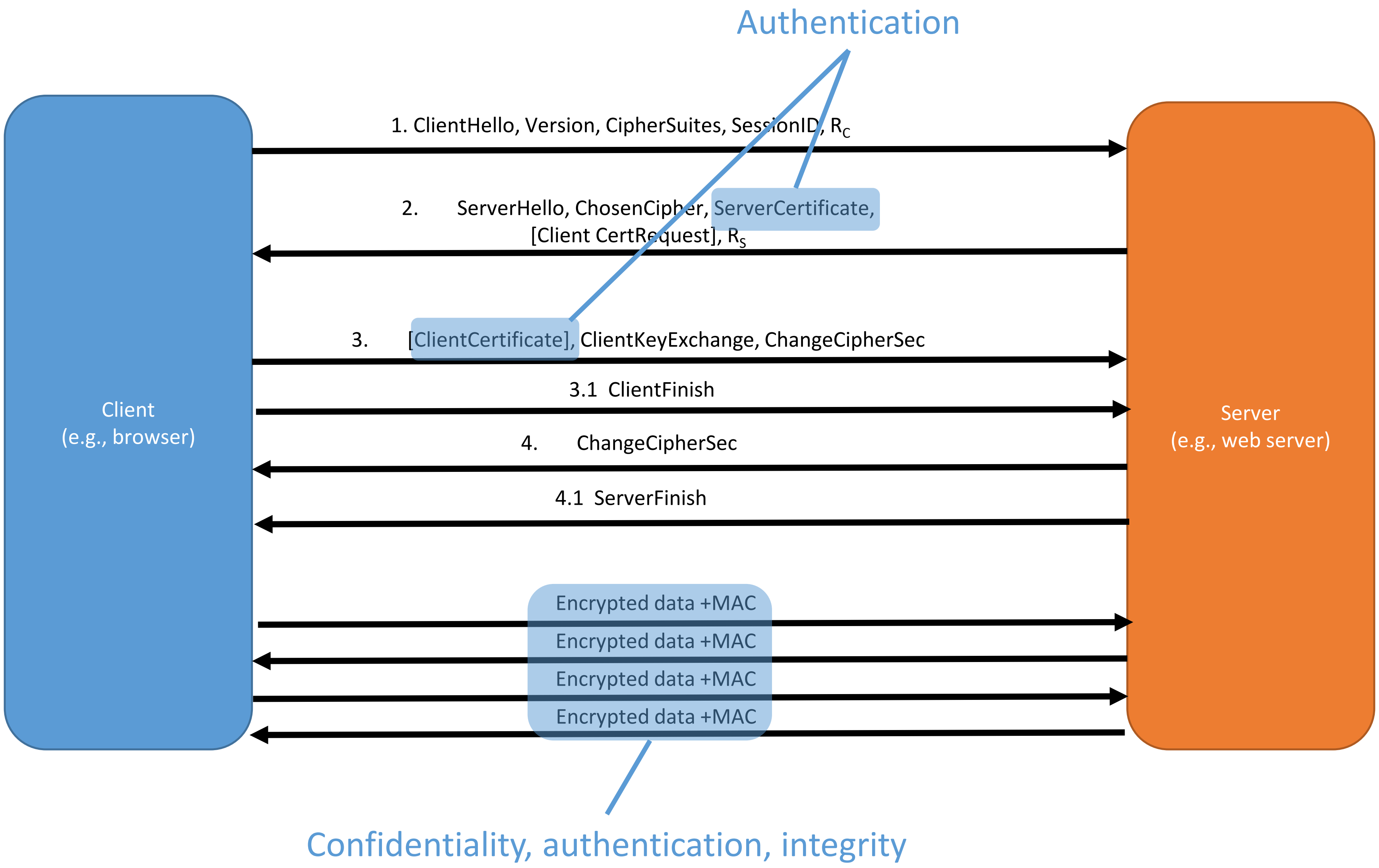
2- Wait for authentication phase to be over

3- Only then use knowledge of sequence numbers to take over the session and inject malicious traffic.

4- Use malicious traffic to execute commands, …

5- The genuine connection gets cancelled

Last week: TLS



Authentication

Client
(e.g., browser)

Server
(e.g., web server)

1. ClientHello, Version, CipherSuites, SessionID, R_C

2. ServerHello, ChosenCipher, ServerCertificate, [Client CertRequest], R_S

3. [ClientCertificate], ClientKeyExchange, ChangeCipherSec

3.1 ClientFinish

4. ChangeCipherSec

4.1 ServerFinish

Encrypted data +MAC

Encrypted data +MAC

Encrypted data +MAC

Encrypted data +MAC

Confidentiality, authentication, integrity

Last week - Denial of Service

Goal: prevent legitimate users from accessing a service

Option A - Crash victim: exploit software flaws to make it stop

Option B – Exhaust victim's resources

- Network: Bandwidth
- Host
  - Kernel: TCP connection state tables, etc.
  - Application: CPU, memory, etc.

Last week – Non cryptographic defenses

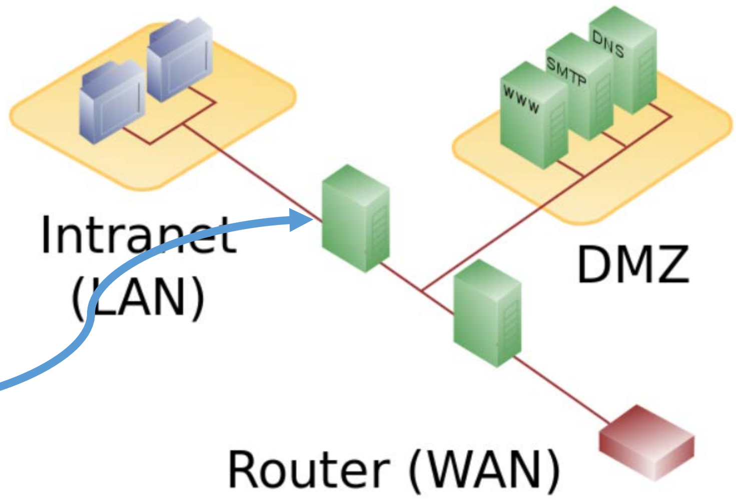
NAT: (translation from public to local IPs) no access to internal IPs

Firewall: filter flows according to a policy

Stateless vs stateful (can remember properties of the flow)

Headers vs content (Deep packet inspection)

Demilitarized zone (DMZ)



bastion host

Computer Security (COM-301)
Privacy

Carmela Troncoso

SPRING Lab

carmela.troncoso@epfl.ch

Some slides/ideas adapted from: George Danezis, Bart Preneel, Claudia Diaz, Seda Guerses

Goal of this lecture

Understanding:

- There are different conceptions of privacy depending on the adversary model

- Depending on the adversary model one relies of different Privacy Enhancing Technologies: different protection degree

- Privacy requires to protect information beyond content: The need to protect meta-data

The context: Availability of data

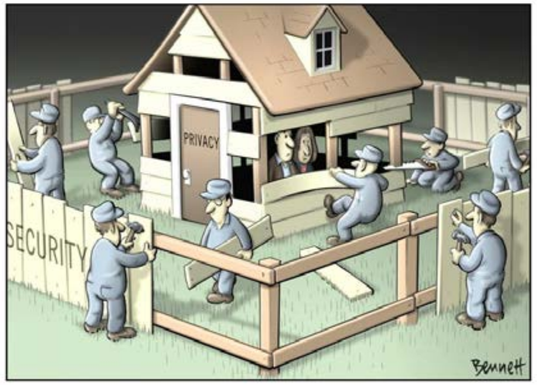
Intelligent data-based applications

- Recommendation systems
  - Movies (Netflix)
  - Products (Amazon)
  - Friends (Social networks)
  - Music (Spotify, iTunes)
- Location based services
  - Friend finders
  - Maps
  - Points of interest
- Health monitoring
- Children/Elderly trackers
- Smart metering
- Intelligent buildings

Individual applications are legitimate

Together they become a cheap
SURVEILLANCE INFRASTRUCTURE

We need privacy!



But what about security!!?!?!?!

Privacy IS a security property

For individuals

protection against crime / identity theft, control over one’s information, protection against profiling and manipulation.

For companies

protection of trade secrets, business strategy, internal operations, access to patents

For governments / military

protection of national secrets, confidentiality of law enforcement investigations, diplomatic activities, political negotiations

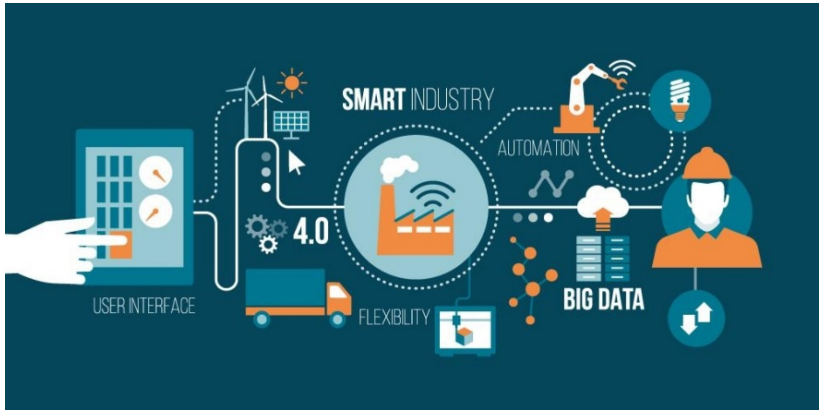
Privacy **IS** a security property

Denying privacy to some is denying privacy to all!!

INFRASTRUCTURE IS SHA

Individuals, Industry, and Governments use the same applications

amazon
web services



Directly
(Cloud-based services, Industry 4.0, Blockchain)

Indirectly
(employers are users)

and Privacy is important for society

Daniel Solove, Prof. of Law

"Part of what makes a society a good place in which to live is the extent to which it allows people freedom from the intrusiveness of others. **A society without privacy protection would be suffocation**"

Not so much Orwell's "Big Brother" as Kafka's "The Trial":

"...a bureaucracy with inscrutable purposes that uses people's information to make important decisions about them, yet denies the people the ability to participate in how their information is used"

"The problems captured by the Kafka metaphor are of a different sort than the problems caused by surveillance. They often do not result in inhibition or chilling. Instead, they are problems of information processing—the storage, use, or analysis of data—rather than information collection."

"...not only frustrate the individual by creating a sense of helplessness and powerlessness, but they also affect social structure by altering the kind of relationships people have with the institutions that make important decisions about their lives."

and Privacy is important for society

Daniel Solove,
Prof. of Law

“Part of what makes a society a good place in which to live is the extent to which it allows people freedom from the intrusiveness of others. **A society without privacy protection would be suffocation**”

Not so much Orwell’s “Big Brother” as Kafka’s “The Trial”:

> “...a bureaucracy with inscrutable purposes that uses people’s information to make important decisions about them, yet denies the people the ability to participate in how their information is used”
>
> “The problems captured by the Kafka metaphor are of a different sort than the problems caused by surveillance. They often do not result in inhibition or chilling. Instead, they are problems of information processing—the storage, use, or analysis of data—rather than information collection.”
>
> “...not only frustrate the individual by creating a sense of helplessness and powerlessness, but they also affect social structure by altering the kind of relationships people have with the institutions that make important decisions about their lives.”

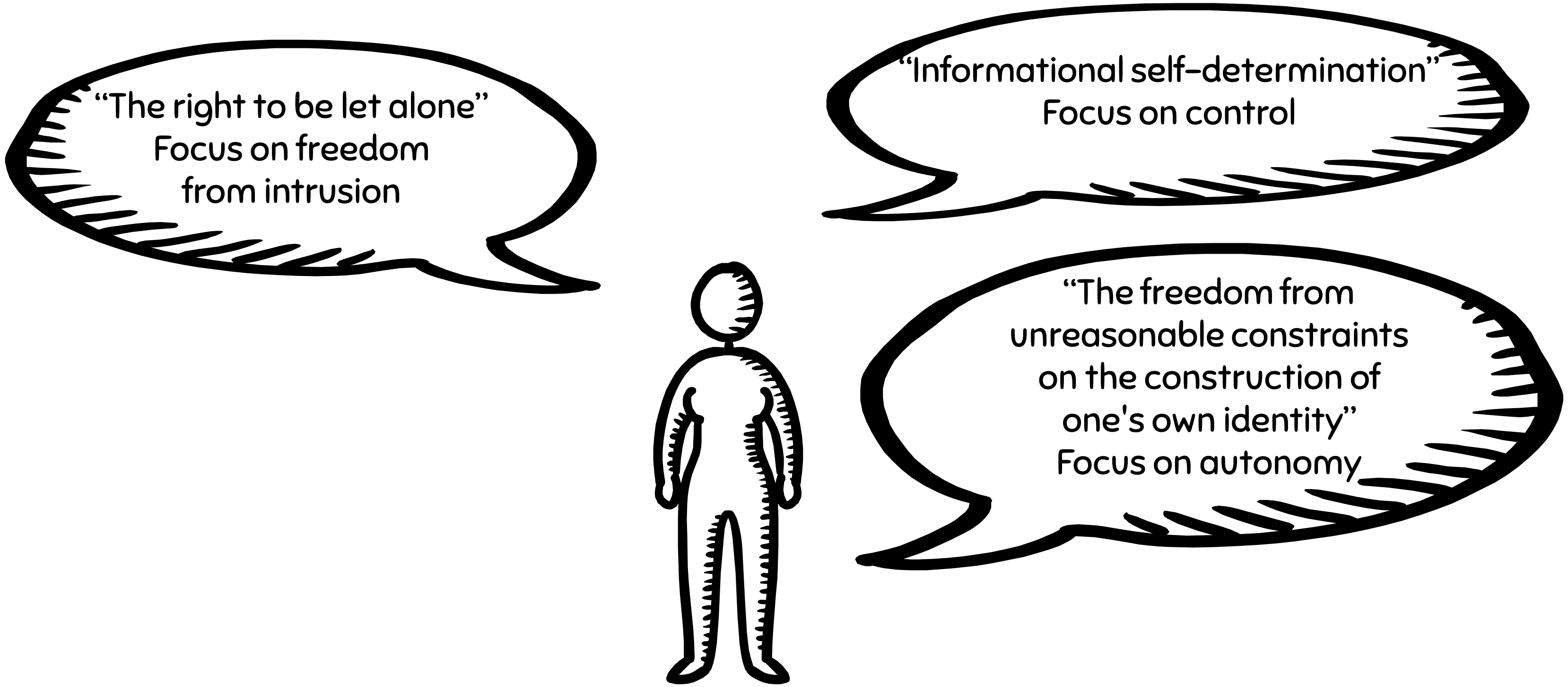
MAY BECOME

ONE RING TO RULE THEM ALL

What is privacy

Abstract and subjective concept, hard to define

Dependent on cultural issues, study discipline, stakeholder, context



"The right to be let alone"
Focus on freedom
from intrusion

"Informational self-determination"
Focus on control

"The freedom from
unreasonable constraints
on the construction of
one's own identity"
Focus on autonomy

What is privacy in Privacy Enhancing Technologies

PETs

3 different types of PETs depending on ...

the concerns they address

their goals

their challenges and limitations

Not these PETs!!!

Gürses, Seda, and Claudia Diaz. "Two tales of privacy in online social networks." IEEE Security & Privacy 11.3 (2013): 29-37.
Diaz, Claudia, and Seda Gürses. "Understanding the landscape of privacy technologies." Information Security Summit (2012): 58-63.
Danezis, George, and Seda Gürses. "A critical review of 10 years of privacy technology." Surveillance cultures: a global surveillance society (2010): 1-16.

1 - Social Privacy

Concerns - The privacy problem is defined by **Users**

Technology brings problems

“My parents discovered I'm gay”

“My boss knows I am looking for other job”

“My friends saw my naked pictures”

Goals - Do not surprise the user

Two main approaches

Support decision making

Help identifying actions impact

Limitations

Only protects from other users: **trusted service provider**!

Limited by user’s capability to understand policies

Based on user expectations – What if the expectations are null?

Common Industry approach
Make users comfortable

2 - Institutional Privacy

GDPR

The General Data Protection Regulation

CONCERNS - The privacy problem is defined by Legislation

Data **should not** be collected without user consent or processed for illegitimate uses

Data should be secured: correct, integrity, deletion

GOALS – Compliance with data protection principles

informed consent

purpose limitation

data minimization

subject access rights

Preserving the security of data

Auditability and accountability

LIMITATIONS

Never questions collection – assumes it is necessary

Trusted service provider! No technical measures to protect data from them

Limits misuse, but not collection (based on consent)

Limited scope (personal data != all data)

3 – Anti-surveillance Privacy

CONCERNS - The privacy problem is defined by **Security Experts**

Data is disclosed **by default** through the ICT infrastructure: **the adversary is anybody**

Concerned about: censorship, surveillance, freedom of speech,…

GOALS – Minimize

Default disclosure of personal information to anyone - both explicit and implicit!

Minimize the need to trust others

LIMITATIONS

Privacy-preserving designs are narrow – difficult to create "general purpose privacy"

Usability problems both for developers and users

how the @$%&#$Ŷ& do I program this?

performance hit

unintuitive technologies

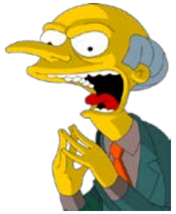
Industry lacks incentives

The adversary is anyone and VERY powerful

Intelligence agencies

Your Parents

Your Children



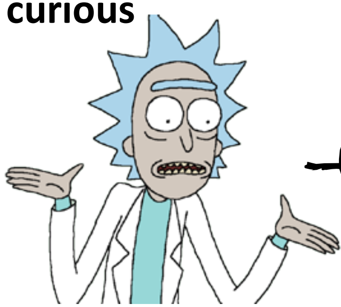
The Boss

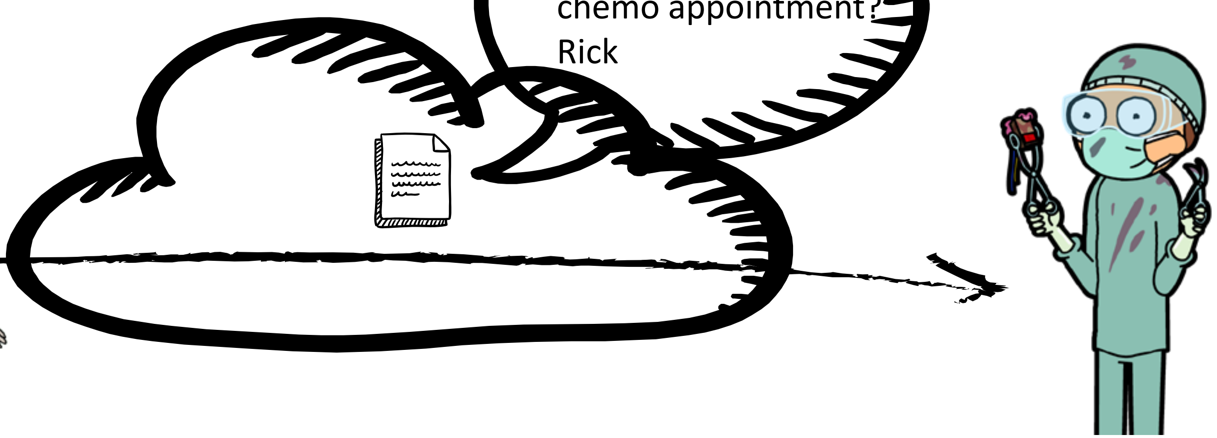
Your Roomates

Anybody curious

ISPs

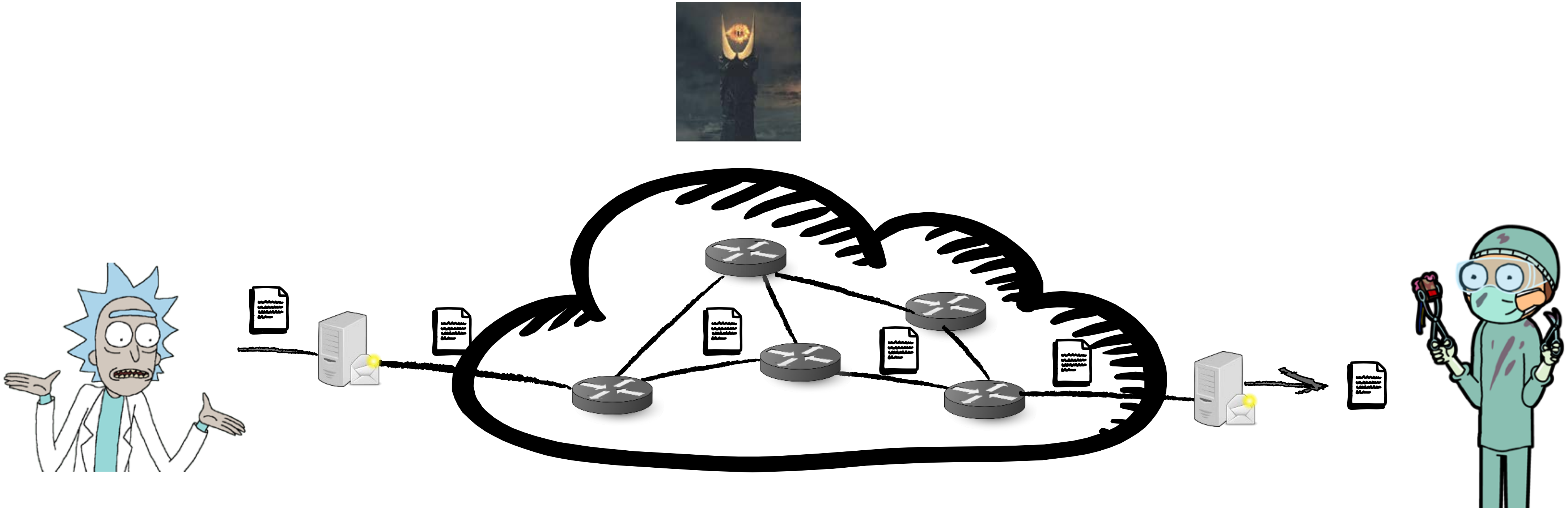
SysAdmins





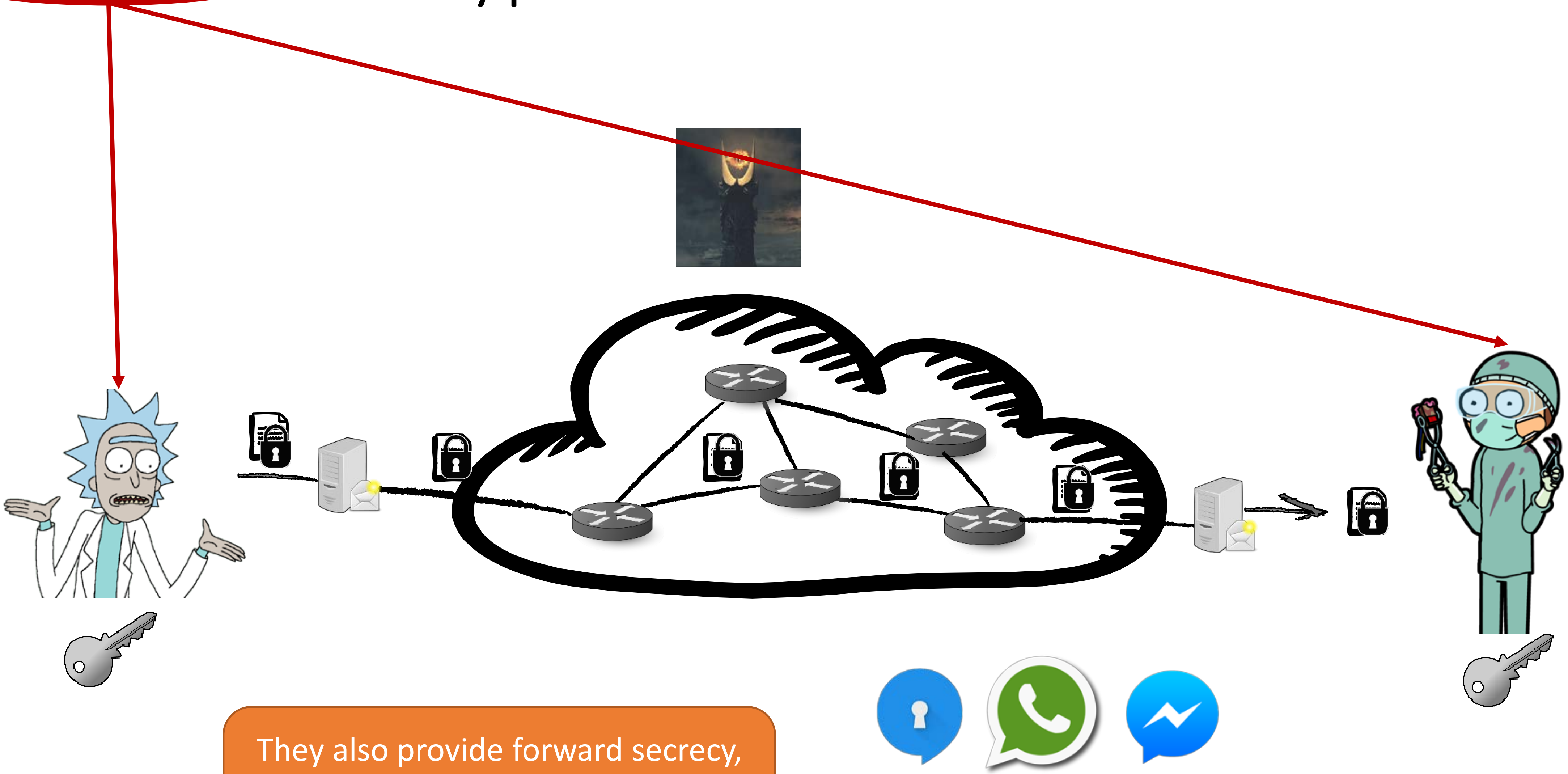
End to End Encryption

What is an End?



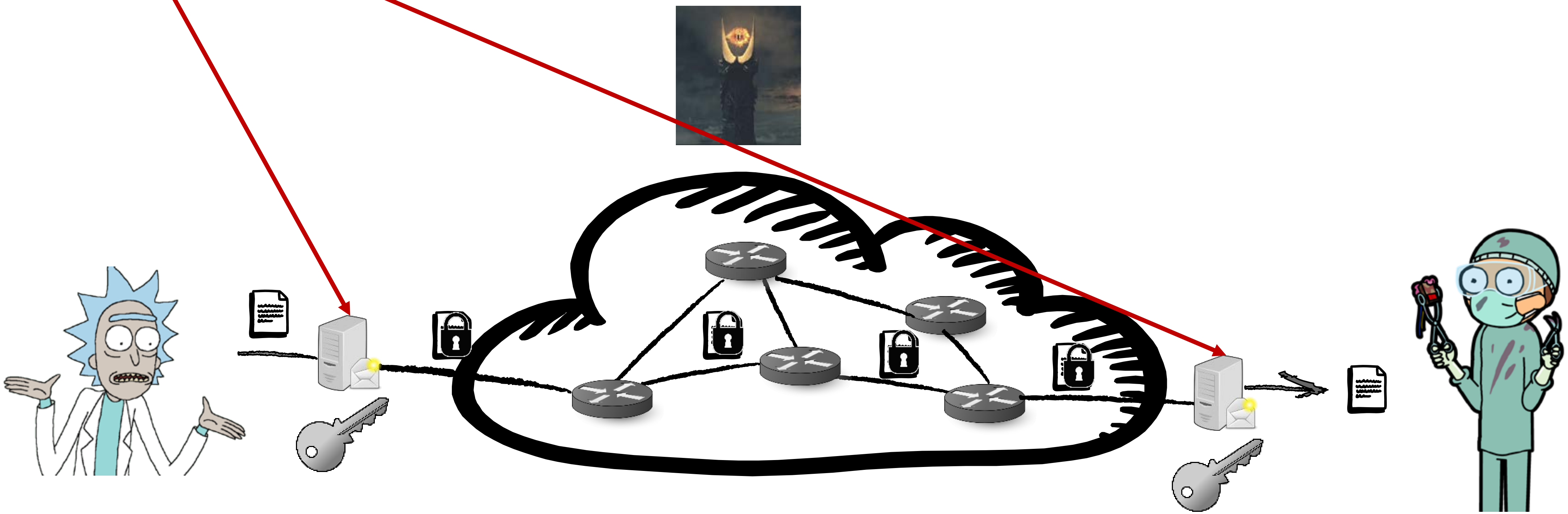
Cryptography → Confidentiality!
(and integrity and authenticity)

End to End Encryption



They also provide forward secrecy, by using ephemeral keys.
How: advanced crypto

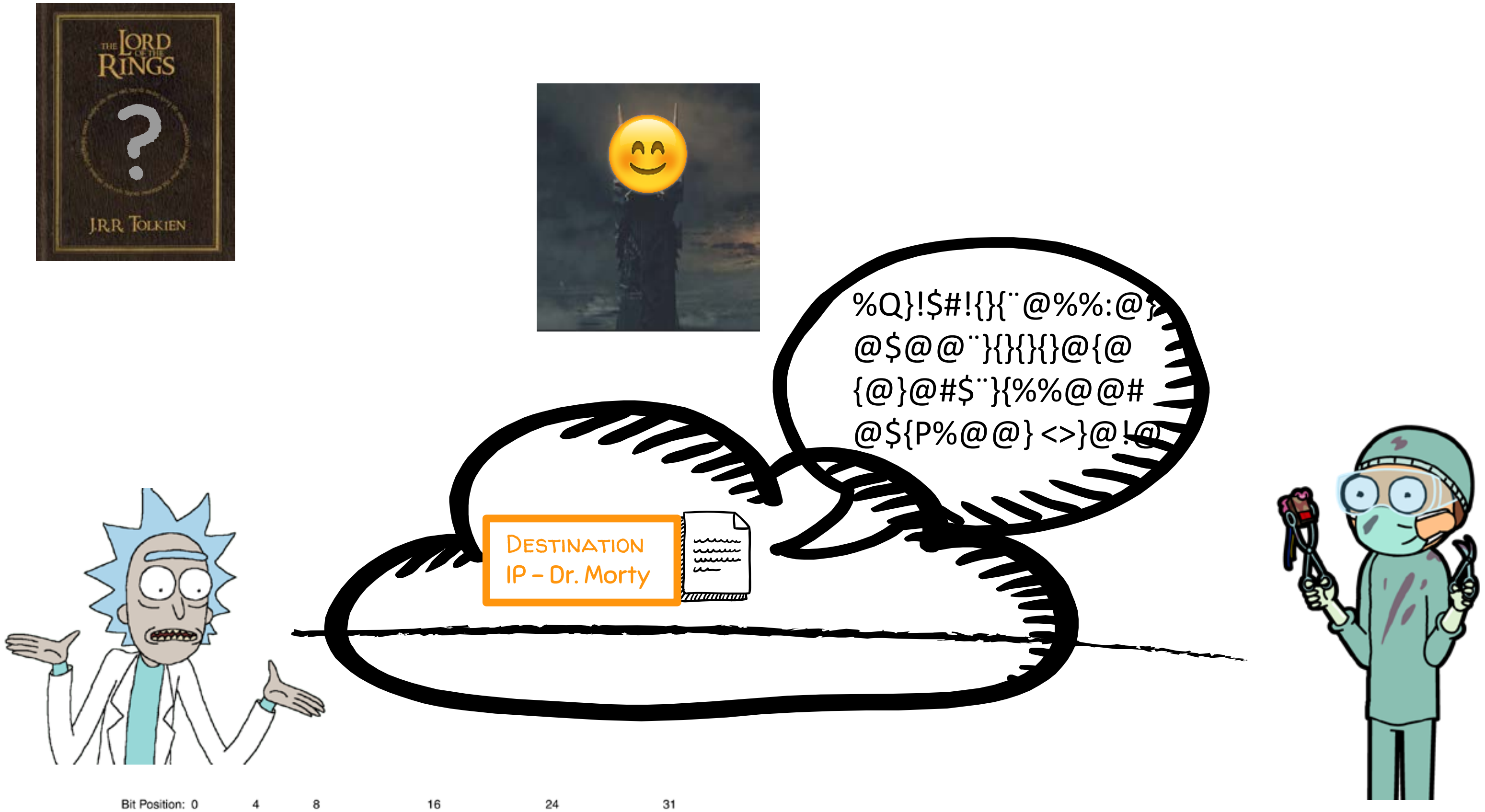
End to End Encryption

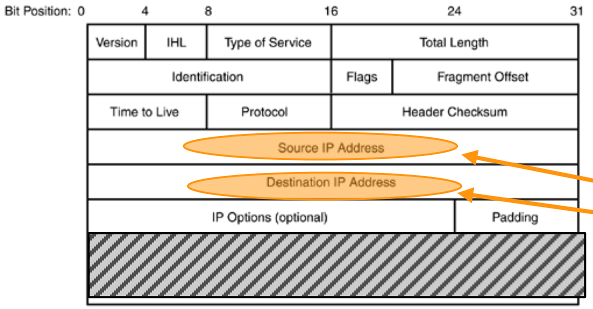


Hotmail | Outlook

But we can encrypt! What is the problem?

The problem is Traffic Analysis





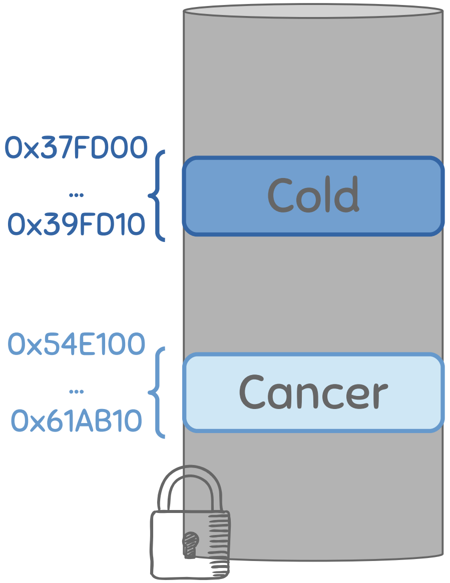
Bit Position: 0 4 8 16 24 31

Version	IHL	Type of Service	Total Length	
Identification			Flags	Fragment Offset
Time to Live		Protocol	Header Checksum	
Source IP Address				
Destination IP Address				
IP Options (optional)				Padding

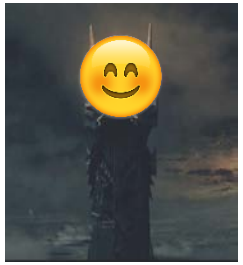
IPv4 Header (RFC 791, 1981)

Same for Ethernet, TCP, SMTP, IRC, HTTP, ...

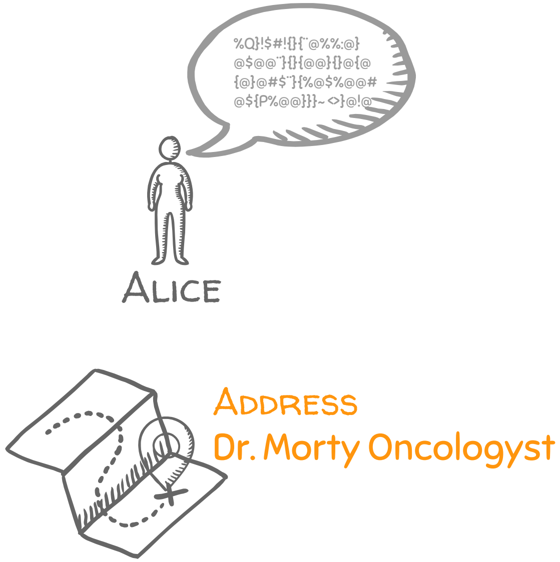
Other metadata is also sensitive!!



The address where data is stored may reveal information about the content.
Example: medical database with patients with mild and severe diseases in different locations



Implicit data is as important as explicit data!



The address where an action happens may reveal information about the action / user.
Example: sending a message from an Oncologist clinic reveals information about the sender

Traffic WHAT?

Wikipedia: traffic analysis is the process of intercepting and examining messages in order to deduce information from patterns in communication

Traffic WHAT?

Wikipedia: traffic analysis is the process of intercepting and examining messages in order to deduce information from patterns in communication

Making use of “just” traffic data of a communication (aka metadata) to extract information (as opposed to analyzing content or perform cryptanalysis)

Identities of communicating parties

Timing, frequency, duration

Location

Volume

Device

Traffic WHAT?

Wikipedia: traffic analysis is the process of intercepting and examining messages in order to deduce information from patterns in communication

Making use of “just” traffic data of a communication (aka metadata) to extract information (as opposed to analyzing content or perform cryptanalysis)

Identities of communicating parties

Timing, frequency, duration

Location

Volume

Device

Military Roots

M. Herman: “These non-textual techniques can establish **targets' locations**, order-of-battle and **movement**. Even when messages are not being deciphered, traffic analysis of the target's Command, Control, Communications and intelligence system and its patterns of behavior provides indications of his **intentions** and **states of mind**”

WWI: British troops finding German boats.

WWII: assessing size of German Air Force, fingerprinting of transmitters or operators (localization of troops).

Nowadays

Diffie&Landau: ”Traffic analysis, not cryptanalysis, is the backbone of communications intelligence”

Stewart Baker (NSA): “metadata **absolutely tells you everything about somebody’s life**. If you have enough metadata, you don’t really need content.”

Tempora, MUSCULAR → XkeyScore, PRISM

Herman, Michael. Intelligence power in peace and war. Cambridge University Press, 1996.
Diffie, Whitfield, and Susan Landau. Privacy on the line: The politics of wiretapping and encryption. MIT press, 2010.
http://www.theguardian.com/world/interactive/2013/nov/01/snowden-nsa-files-surveillance-revelations-decoded

We need to protect the communication layer!
Why a**nonymous communications?**

If you are a cyber-criminal!

DRM infringement, hacker, spammer, terrorist, etc.

But, also if you are:

- Journalist
- Whistleblower
- Human rights activist
- Business executive
- Military/intelligence personnel
- Abuse victims

And normal people??

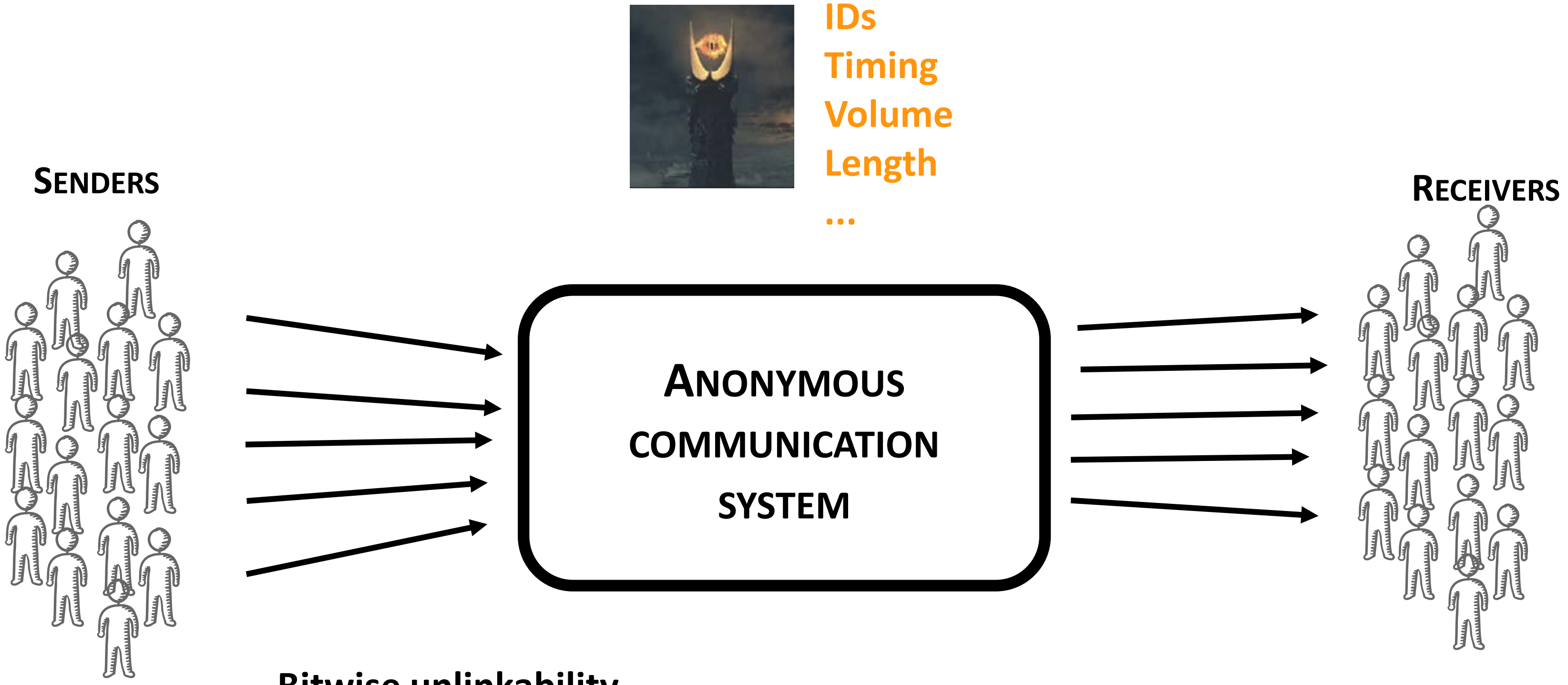
- Avoid tracking by advertising companies
- Protect sensitive personal information from businesses, like insurance companies, banks, etc.
- Express unpopular or controversial opinions
- Have a dual life
  - A professor who is a pro in LoL!
- Try uncommon things

…

-It feels good to have some privacy!

https://www.eff.org/deeplinks/2013/10/online-anonymity-not-only-trolls-and-political-dissidents
http://geekfeminism.wikia.com/wiki/Who_is_harmed_by_a_%22Real_Names%22_policy%3F

Anonymous communications – Abstract model



IDs
Timing
Volume
Length
...

SENDERS

ANONYMOUS COMMUNICATION SYSTEM

RECEIVERS

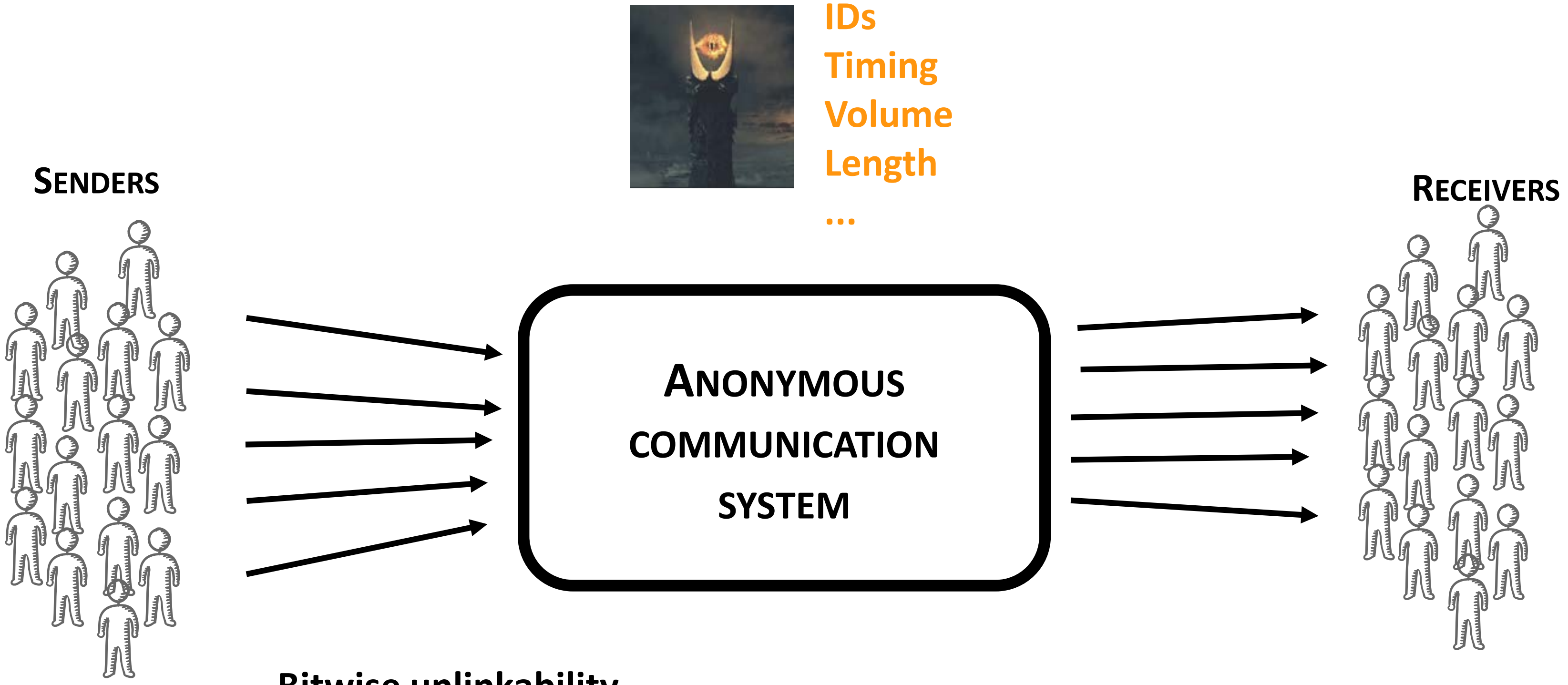
Bitwise unlinkability

Use cryptography to make inputs and outputs to the anonymous communication systems appearance (bits) different

(re)packetizing + (re)schedule

Destroy patterns (traffic analysis resistance)

Anonymous communications – Abstract model



IDs

Timing

Volume

Length

...

SENDERS

ANONYMOUS COMMUNICATION SYSTEM

RECEIVERS

Bitwise unlinkability

Use cryptography to make inputs and outputs to the anonymous communication systems appearance (bits) different

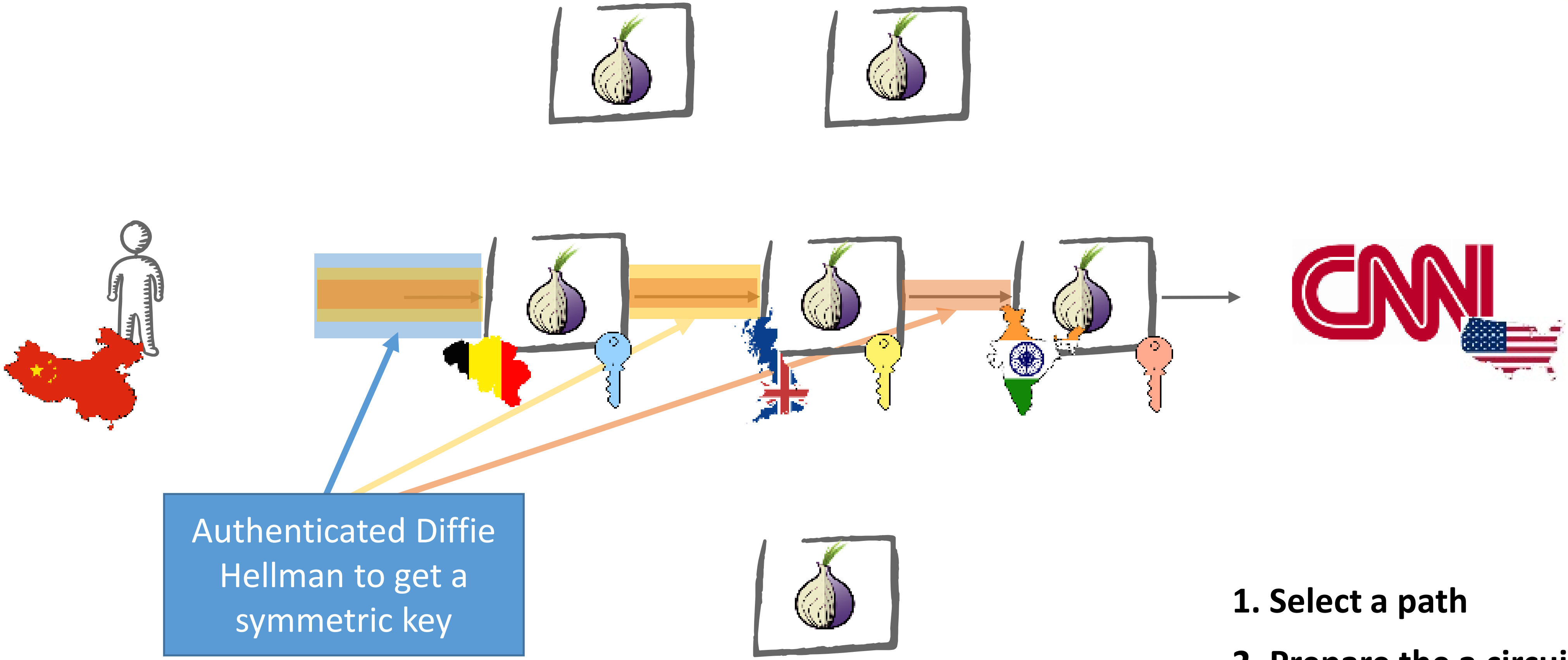
(re)packetizing + (re)schedule + (re)routing

Destroy patterns (traffic analysis resistance)

Load balancing

Distribute trust

The Tor network – Onion routing

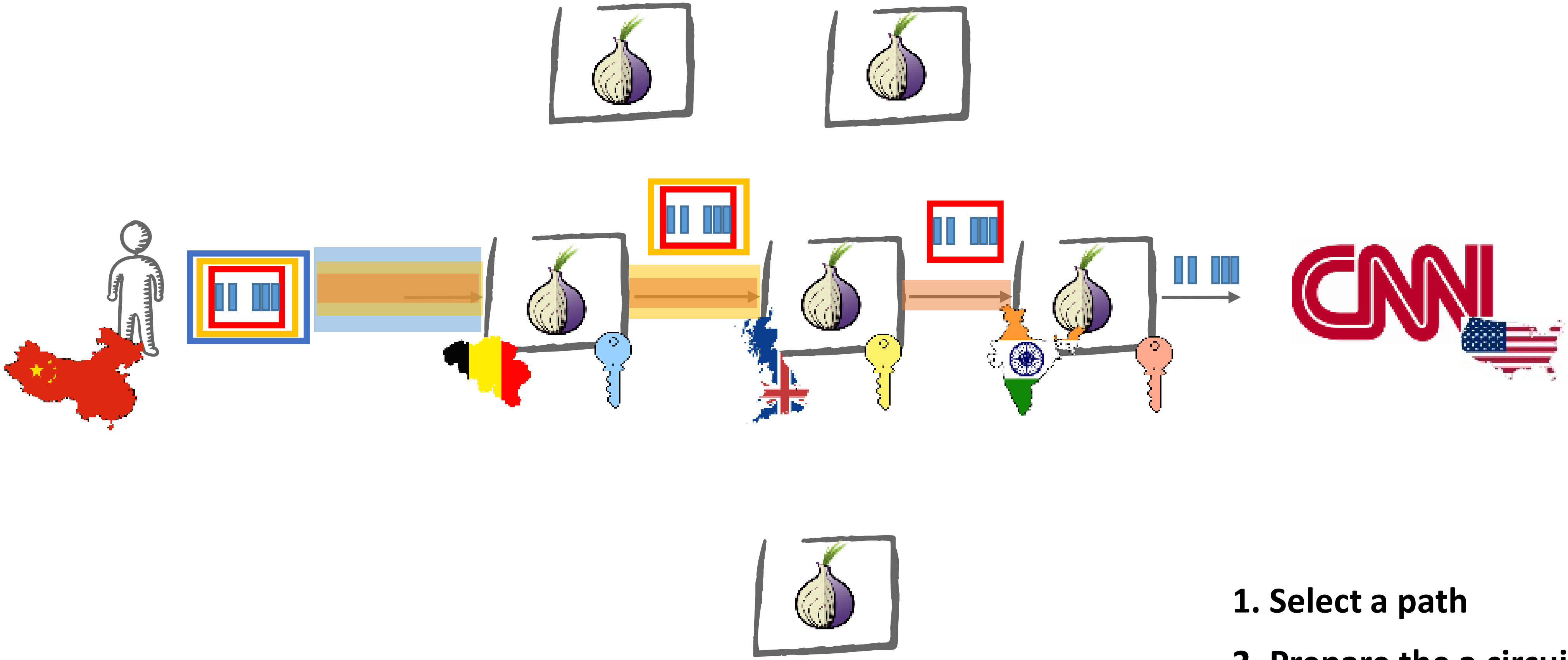


Authenticated Diffie Hellman to get a symmetric key

1. **Select a path**

2. **Prepare the a circuit**

The Tor network – Onion routing

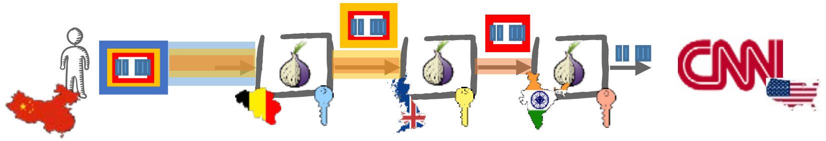


1. Select a path
2. Prepare the a circuit
3. Send stream

Anonymous communications out there

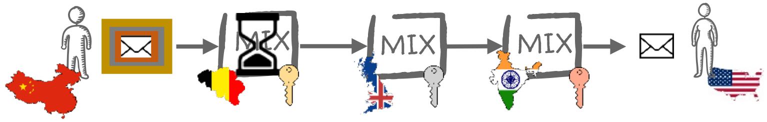
Low Latency

Tor I2P JonDonym



Web browsing, Instant Messaging, streaming

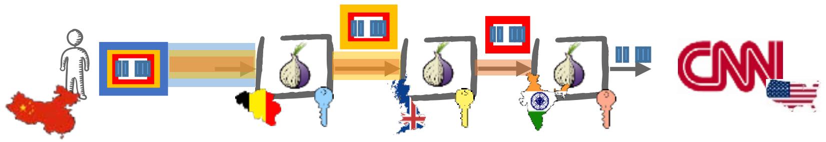
High Latency



Anonymous communications out there

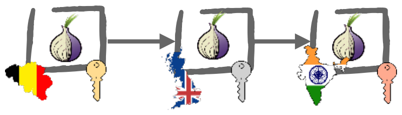
LOW LATENCY

Tor TorProject.org I2P JonDonym

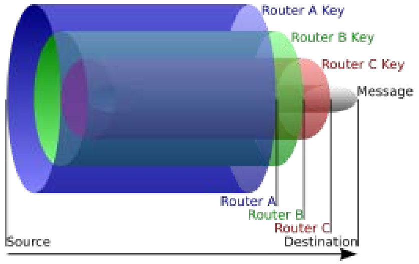


Web browsing, Instant Messaging, streaming

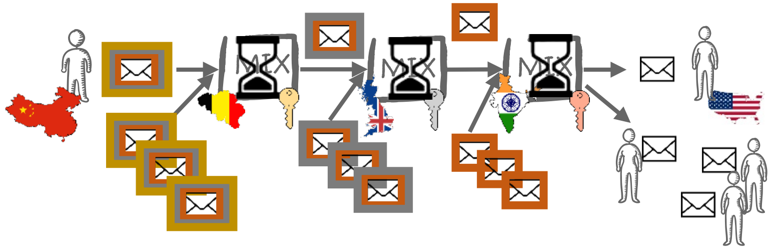
STREAM-based:



fixed for the stream

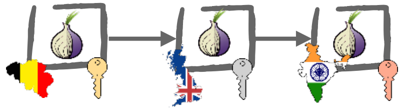


HIGH LATENCY



Email, Voting

MSG-based:

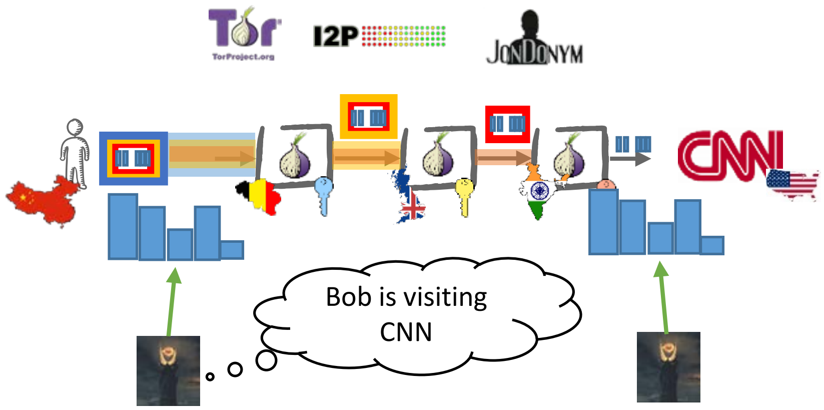


vary every message

One route per message + delays
(slower!)

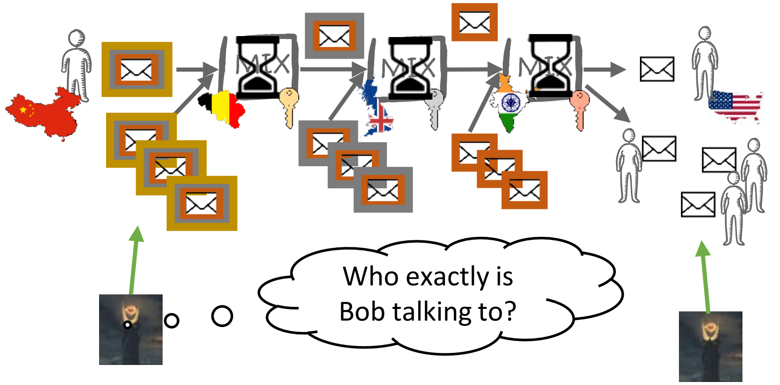
Anonymous communications out there

Low Latency



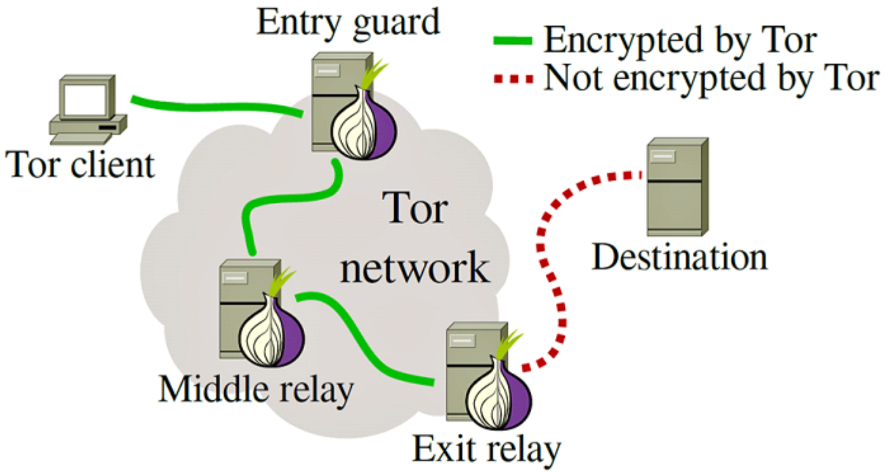
Cannot resist **Global** Adversary
(Tor assumes that the adversary cannot see both edges)

High Latency

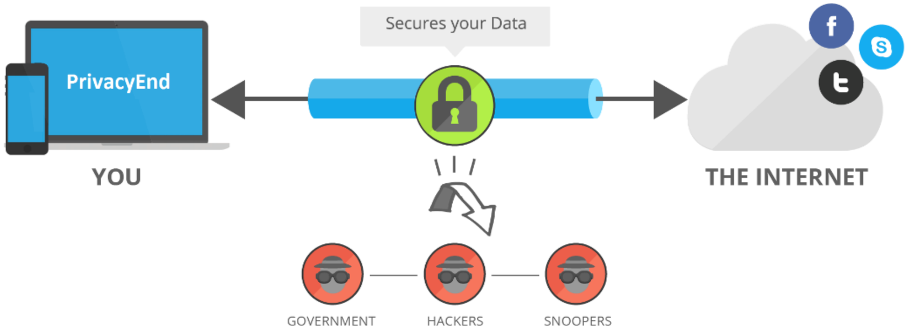


Global Adversary resistance
at the cost of latency
(and long term patterns revealed)

Anonymous communications vs. VPN

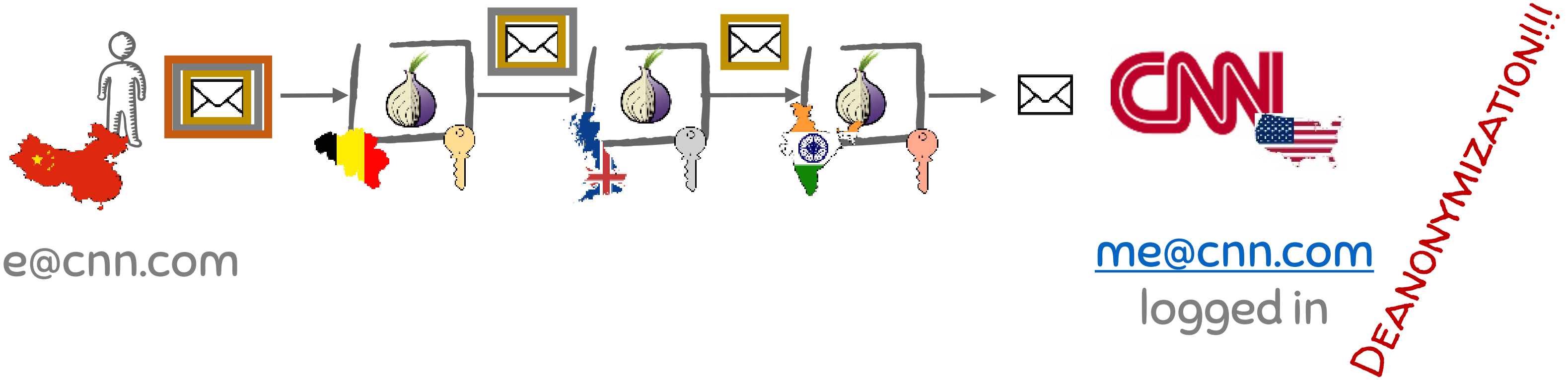


Different trust models!! Who is the adversary?

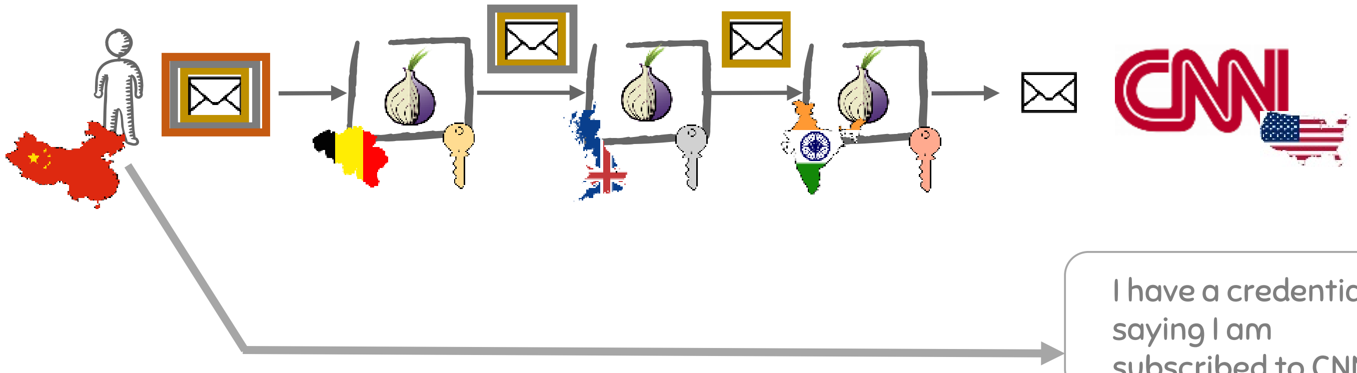


https://www.privacyend.com/difference-between-tor-and-vpn/

Anonymous communications at network layer
what about the application layer?

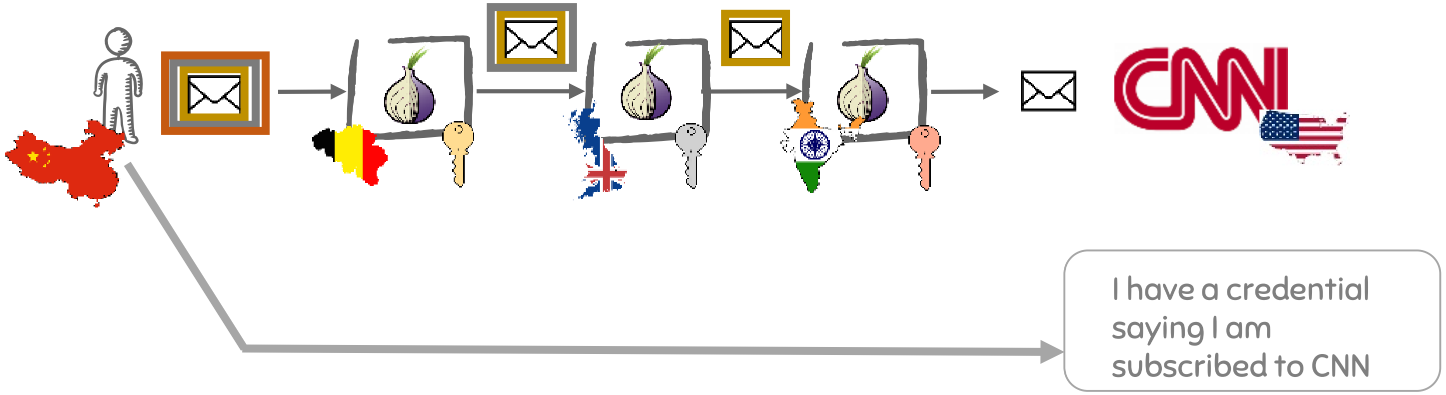


Anonymous communications at network layer what about the application layer?



I have a credential saying I am subscribed to CNN

Anonymous communications at network layer
what about the application layer?



Anonymous credentials
Attribute-based credentials

When used the server **cannot**

- Identify Alice (if her name is not provided)
- Learn anything beyond the info she gives (and what can be inferred)
- Distinguish two users with the same attributes
- Link multiple uses of the same credentials

Public Key Infrastructure
(usual internet authentication)

Signed by a trusted issuer
Certification of attributes
Authentication (secret key)

No data minimization
Users are identifiable
Users can be tracked
(Signature linkable to other contexts where PK is used)

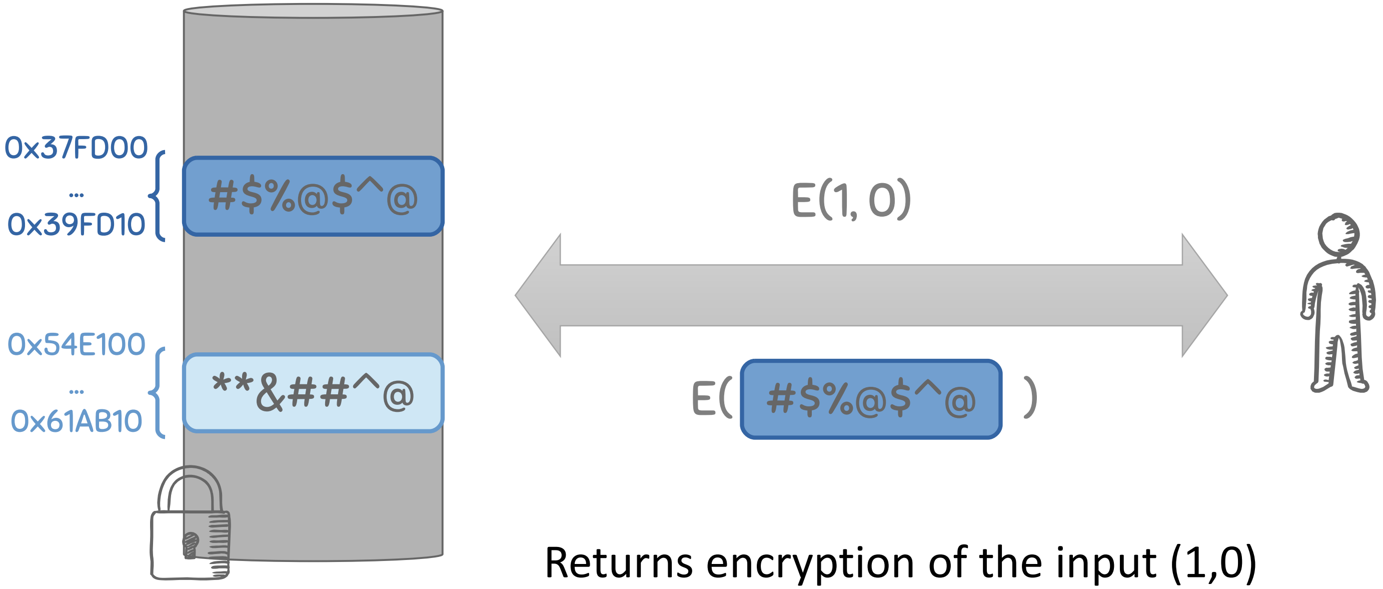
Attribute based credentials

Signed by a trusted issuer
Certification of attributes
Authentication (secret key)

Data minimization
Users are anonymous
Users are unlinkable across contexts

Private Information Retrieval

“is a protocol that allows a user to retrieve an item from a server in possession of a database without revealing which item is retrieved.”



Returns encryption of the input (1,0) multiplied by the rows of the database (using homomorphic encryption)

The multiplication simulates accessing FULL database

Homomorphic, advanced crypto

Examples of other PETs

Private set intersection

a client and a server jointly compute the intersection of their private input sets in a manner that at the end the client learns the intersection and the server learns nothing (one-way PSI) or both learn the intersection (mutual PSI) -- private search

Blind Signatures

a server signs a message produced by a client without learning the content of the message -- eCash

Multiparty computation

parties to jointly compute a function over their inputs while keeping those inputs private –- compute total computations (statistics)

Want more? CS-523: Advanced Privacy Technologies