Everyone is na

Privacy on peer-to-peer systems

Hello there! Thanks a lot for taking your time to listen to this talk. I'm Gonçalo Pestana a systems developer and research enthusiast working on privacy in P2P networks and many other things.

Today we will discuss about how P2P systems play together with user privacy, what kind of data may be disclosed when using P2P based networks and what to do about it.

My research and work - and what we're talking about today - focus on the tension between decentralisation and peer to peer technology and privacy.

While I do believe that P2P technology and systems have a lot of potential in terms of scalability and enabling people and communities to build services that really matter, I also believe that privacy online is the most important digital right we can ask for, and as we know, it has been under.

Privacy on peer-to-peer (P2P) systems

context

a threat model (why is this even important?)

This talk is rather unstructured and please feel free too interrupt at any point. I'd rather turn this next 30m into a discussion.

We'll start by giving some context and definitions to make sure we're all on the same page and then define what is the threat model and why is this important conversation to have.

I'll then demonstrate a few examples of how P2P systems being used by thousands of people are currently exposing user's privacy

And we'll wrap up by briefly discussing about some ideas on how to mitigate some of the problems discussed throughout the talk

everyone is naked ..

.. and putting some clothes on

Everyone is naked @gpestana

Privacy on peer-to-peer (P2P) systems in a nu

Scalability, availability

New paradigm where people and communities c
interoperate and collaborate without the need for
stewardship

Collaboration —> centralised systems disclose u
behaviour and social graph to one centralised ent

naive P2P systems potentially disclose that information to everyone in the network

But let's start from the end.

P2P systems differ from a client-server interaction mo by turning every entity in the system both a client and server. In these systems, network peers are both consumers and providers and collaborate between ea other to reach a common goal.

P2P systems are extremely interesting solutions for building systems that do not rely on central authorities and to provide scalability and resilience in the client-server model.

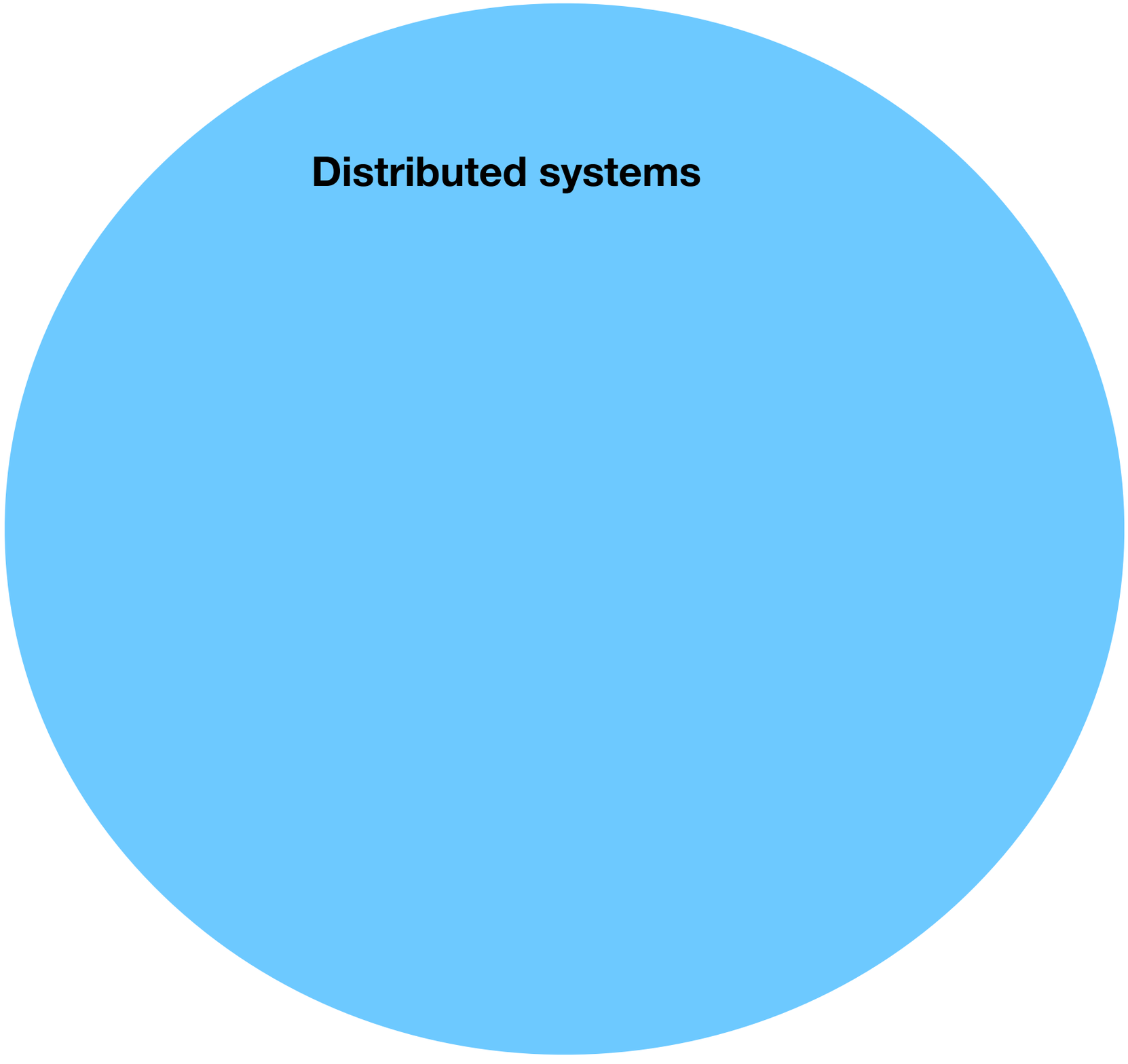
P2P technology has helped centralised systems to sc up and become more resilient and also it has open the door for the decentralised web, where services are maintained without a central authority and users collaborate with each other to reach a common goal.

However powerful, P2P systems can easily become a privacy threat. The main reason for that is collaboratio peers need to communicate what they want to each other. When designed and implemented naively. P2P networks can disclose a lot of information about user' behaviour, social graph, etc.

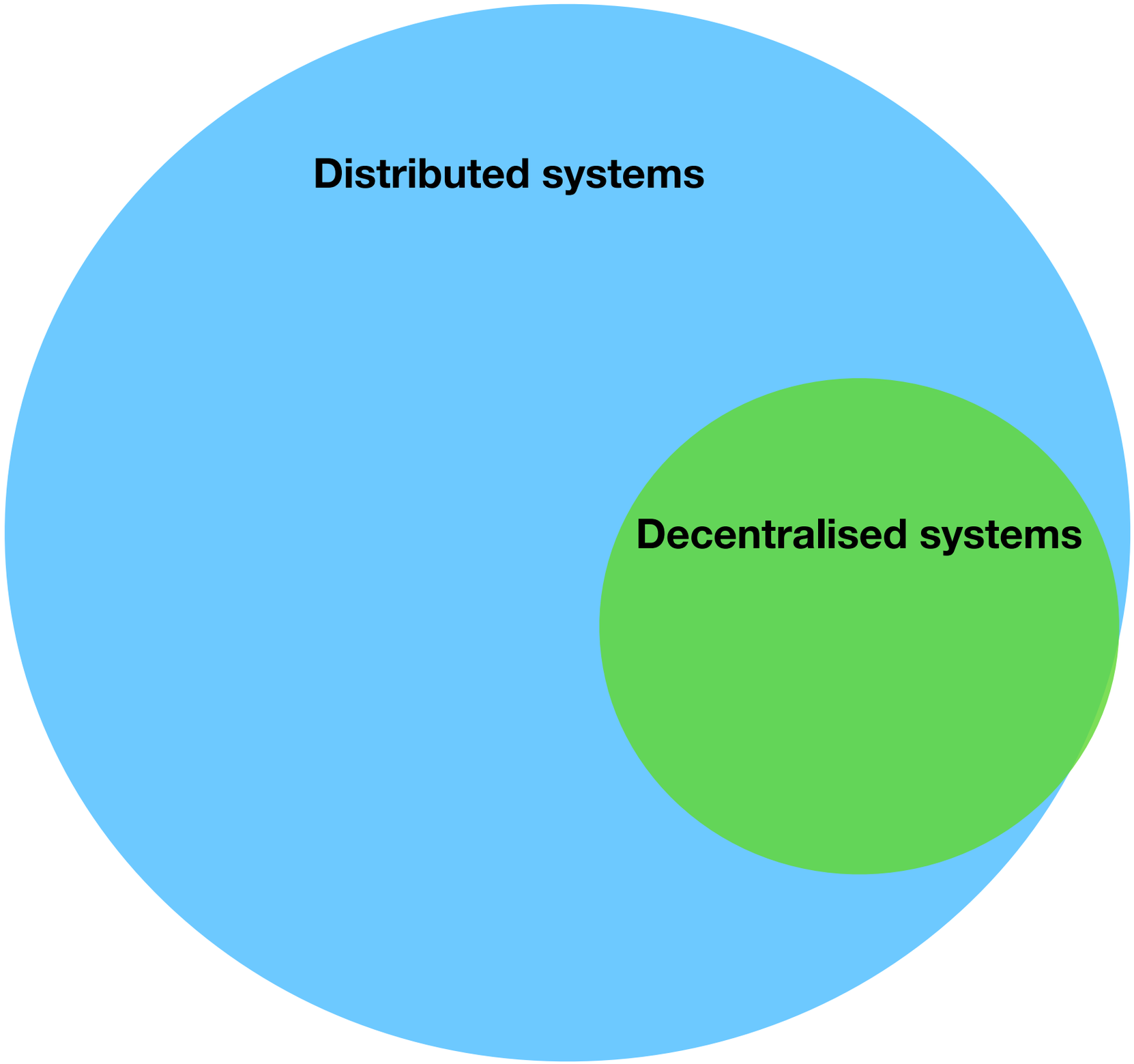
It is possible the problems of privacy in P2P networks but it comes at a cost of complexity and performance

——

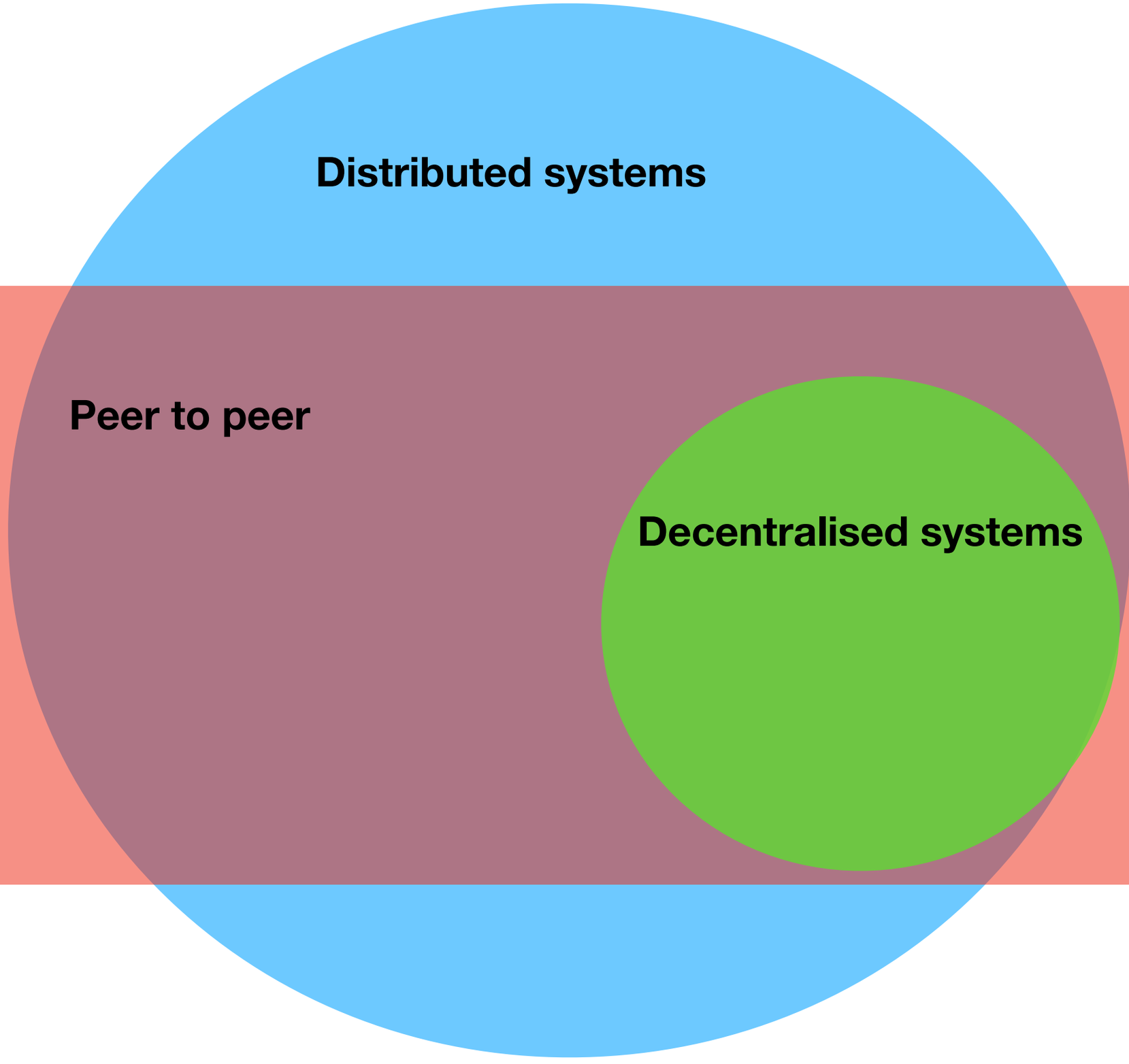
If you are in a rush, you can leave now. Hopefully this



Resources in different network locations, orchestration through message passing



No central truth,
control, authority



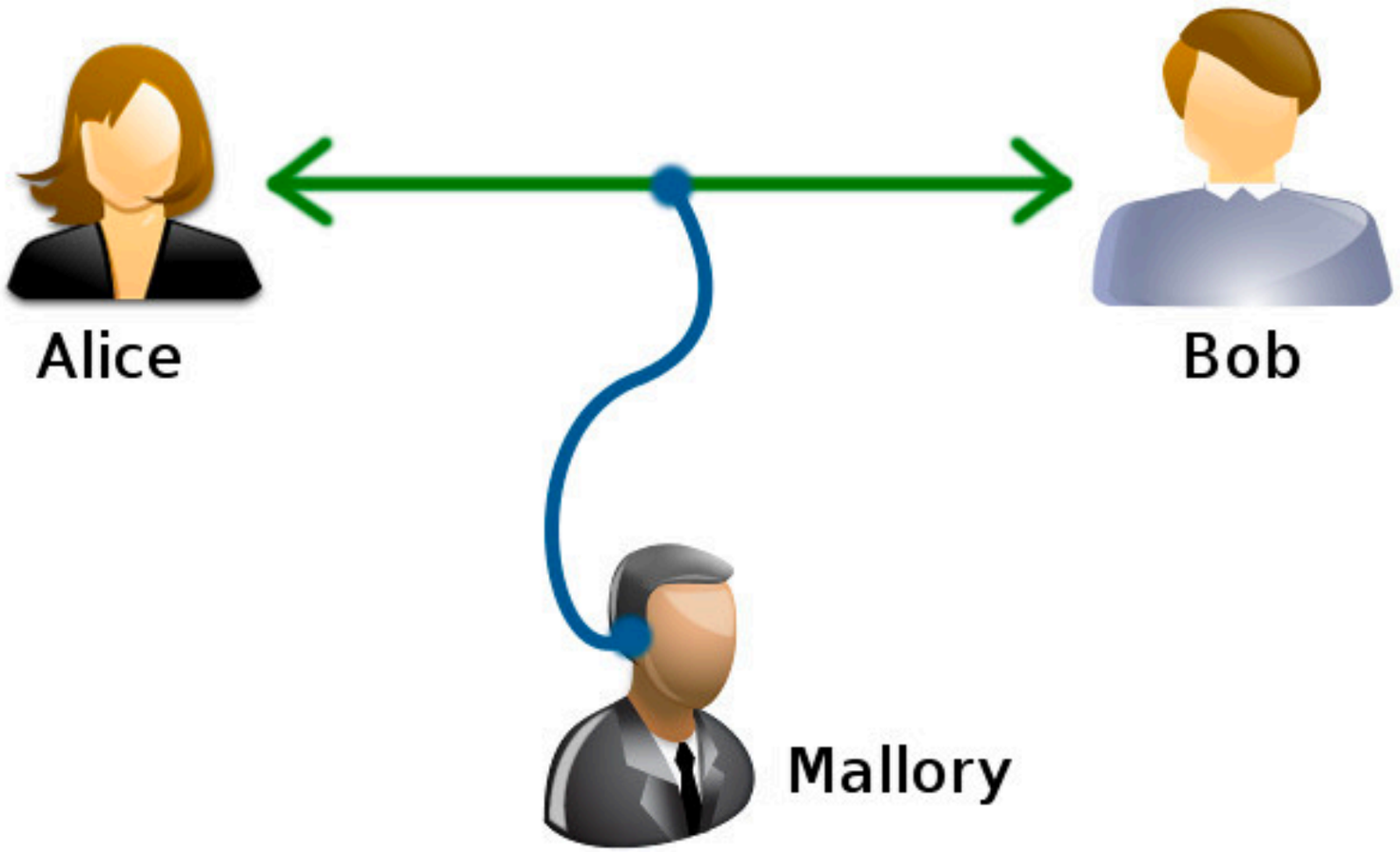
Decentralised and hybrid (independent nodes and central auth)

Metadata data that describes and gives information about other data. Any type of info used to **infer** user behaviour

Encryption message/data encoding

Metadata data that describes and gives information about other data. Any type of info used to **infer** user behaviour

Encryption message/data encoding



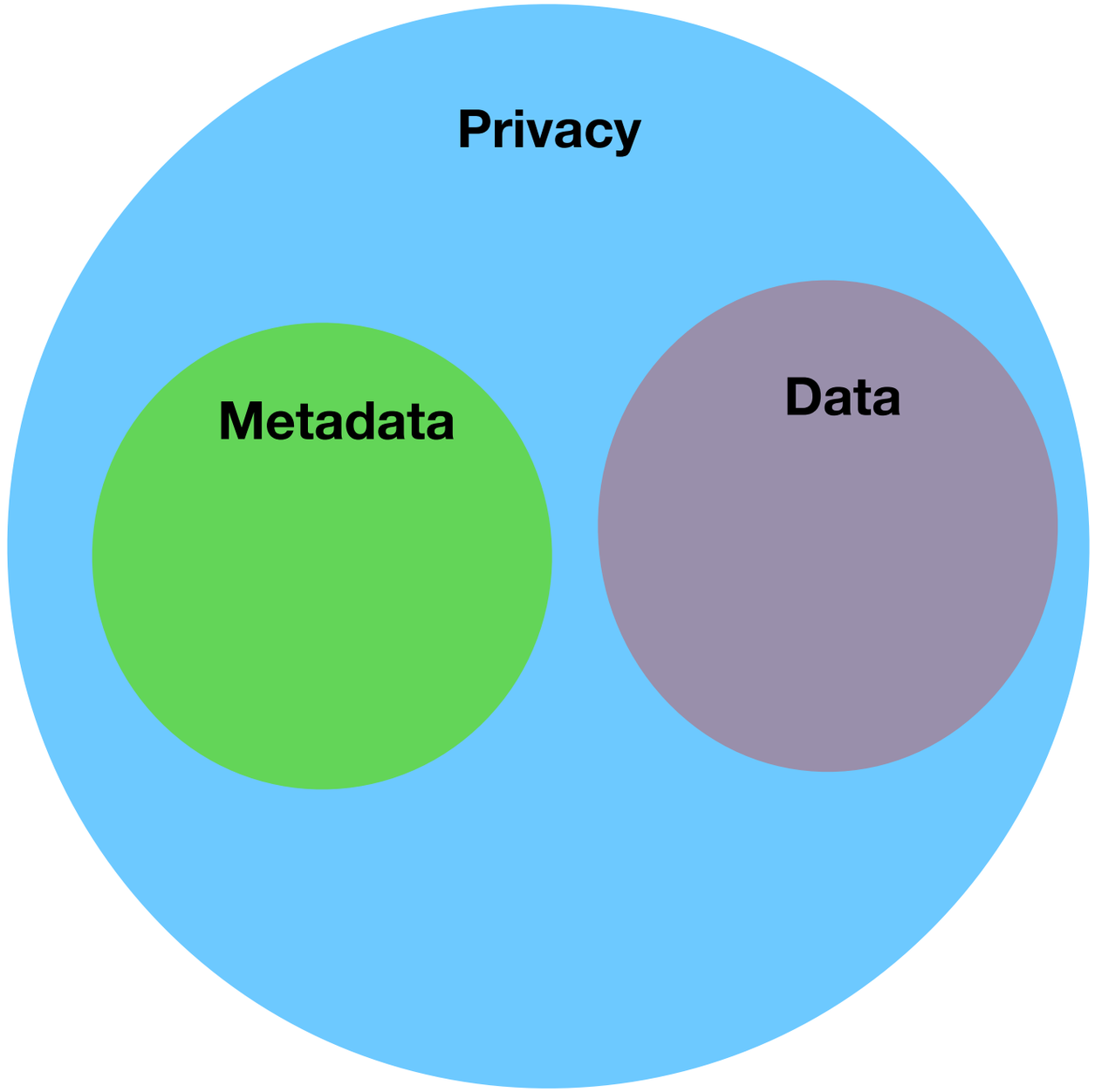
Everyone is naked @gpestana

Metadata data that describes and gives information about other data.
Any type of info used to **infer** user behaviour

Encryption message/data encoding

Privacy measures leaks of

metadata and information



“We kill people based on metadata"

Metadata is often overlooked when we talk about privacy, but the fact is that metadata can reveal an awful amount of information about a person: what content someone has accessed, when, who did this person contacted and when.

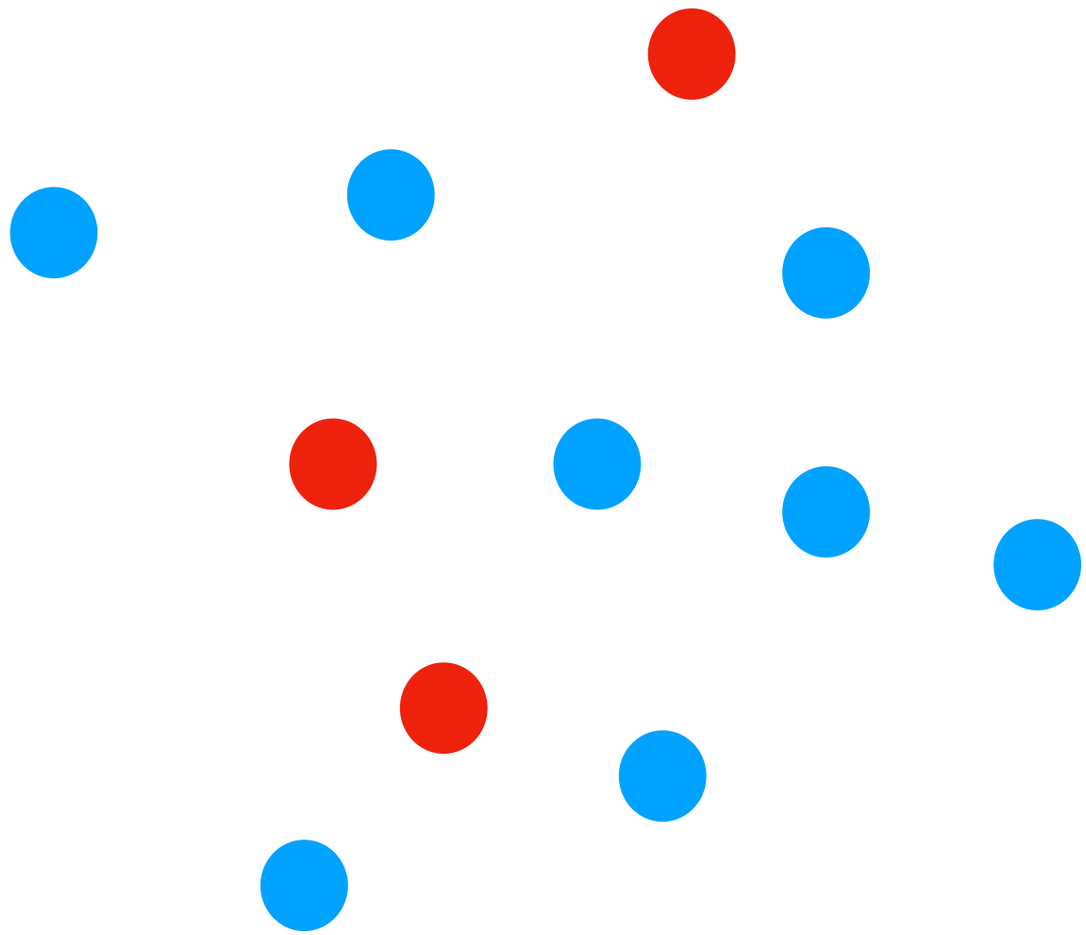
Even though the content of messages is encrypted, just by observing interactions and patterns we can create a rather accurate profile of a person.

Nowadays, we know that metadata is being used as a weapon, for example.

And in P2P systems the metadata attack surface increases considerably. Let’s see why and some concrete and tangible cases.

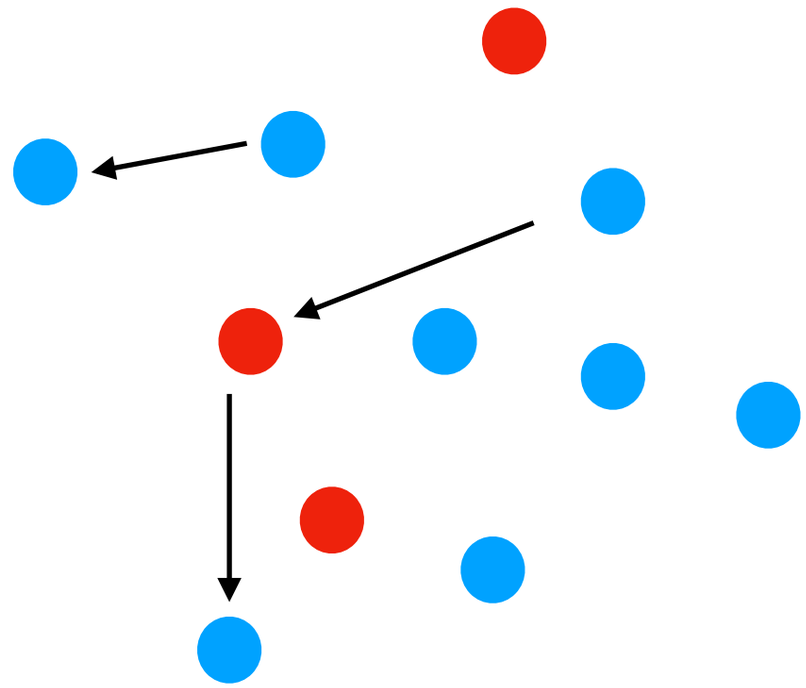
Threat model

"bring your own node” network (with or without central auth)



Everyone is naked @gpestana

Threat model



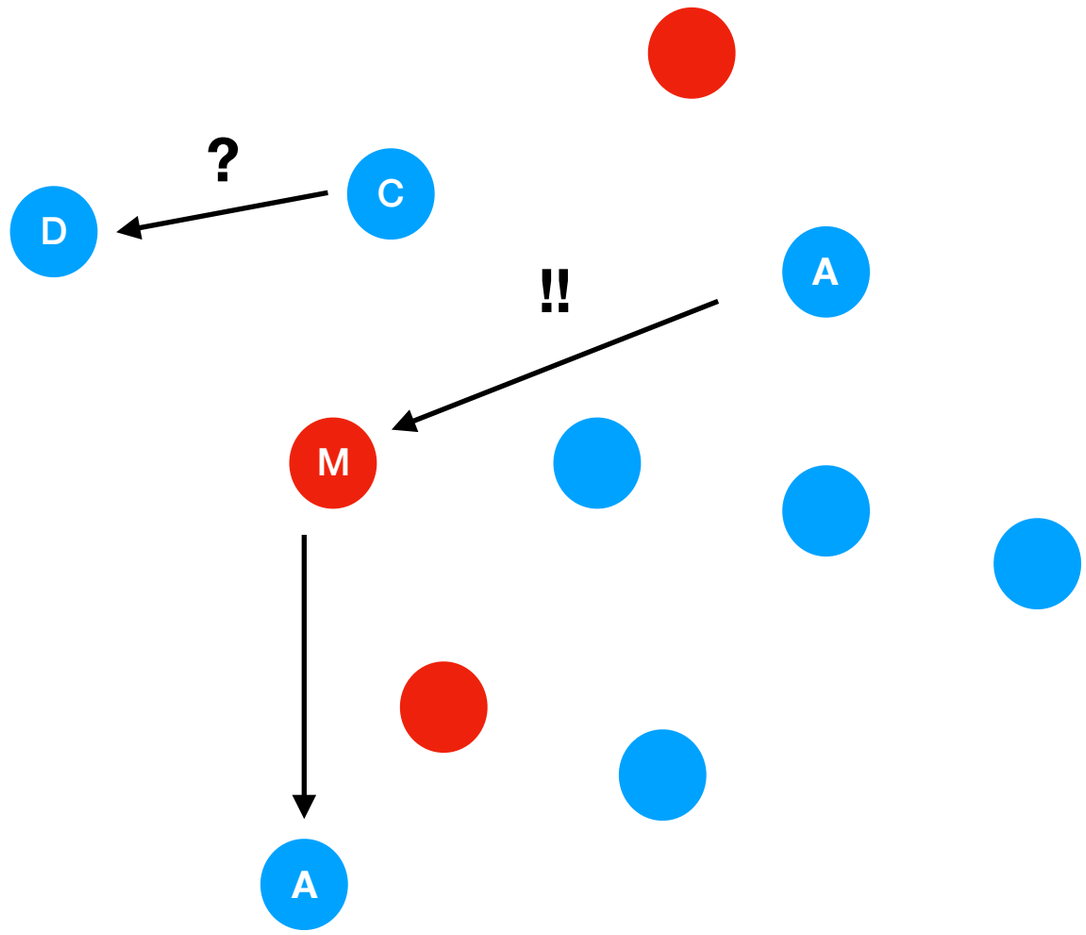
Local, passive adversary controlling up to 20% of the network

- should not learn about others' **behaviour** and **interests**

- should not learn other's **social graphs**

Threat model

"bring your own node” network (with or without central auth)

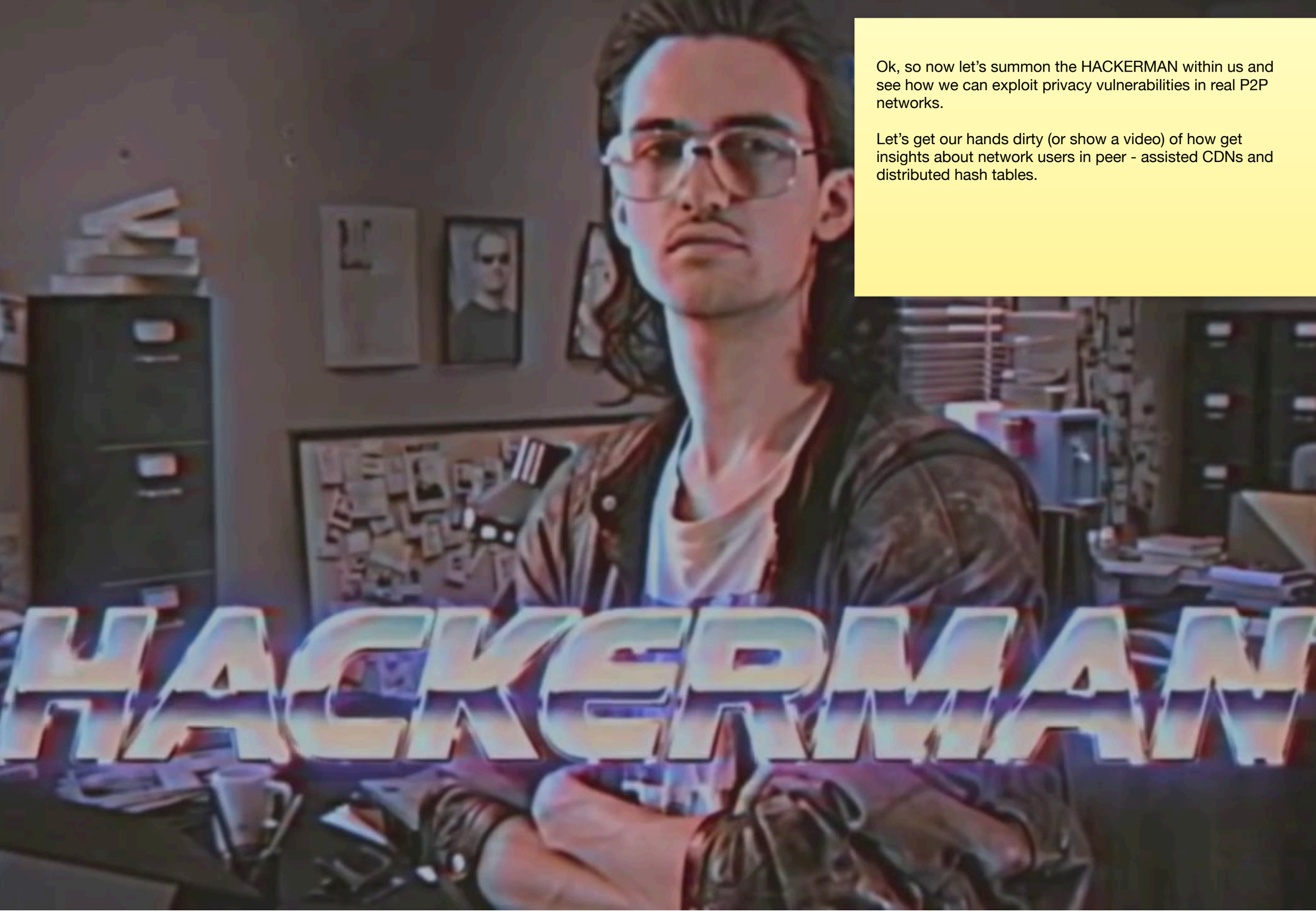


Everyone is naked @gpestana

Threat model

Collaboration and **reduced view** of the network requires information passing (and most of the times leakage) between peers

Everyone is naked @gpestana

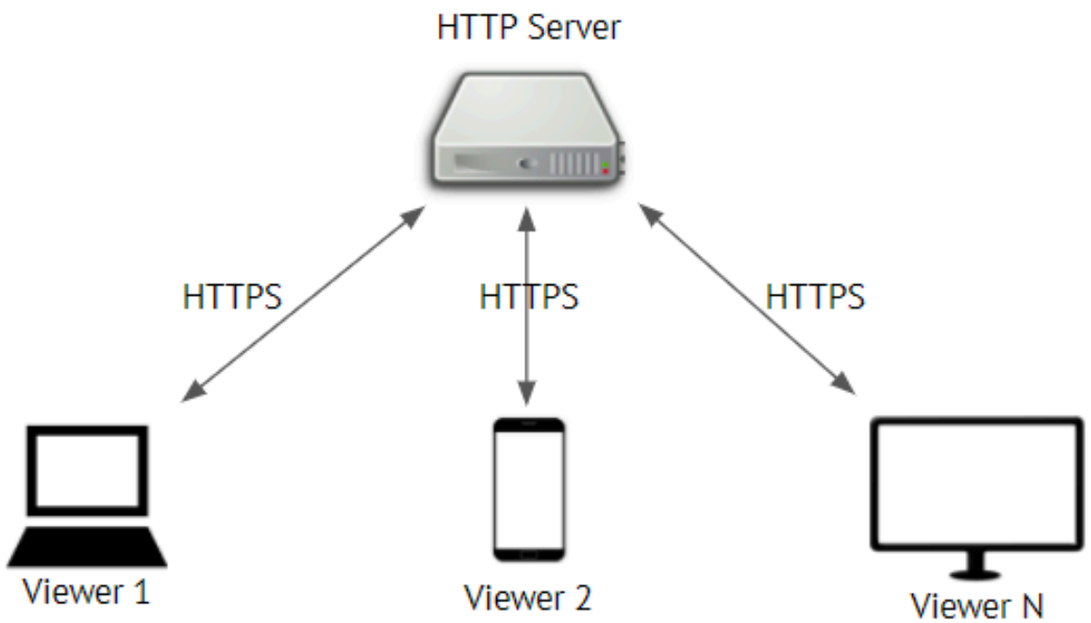


Ok, so now let’s summon the HACKERMAN within us and see how we can exploit privacy vulnerabilities in real P2P networks.

Let’s get our hands dirty (or show a video) of how get insights about network users in peer - assisted CDNs and distributed hash tables.

Peer-assisted Content Distribution Network (pCD

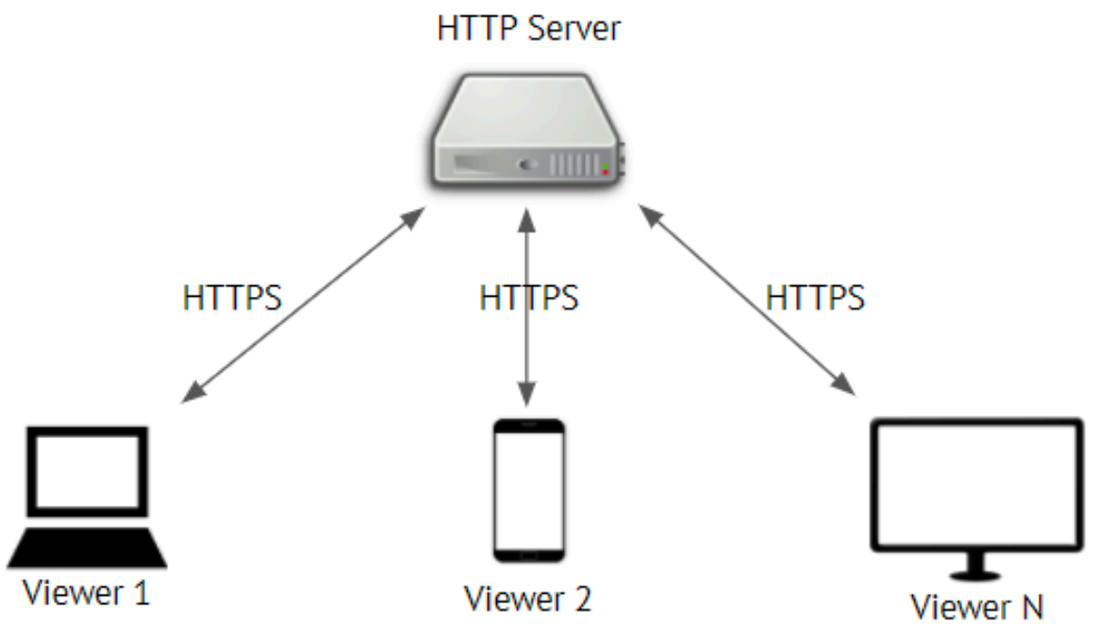
Let's start with the case of a peer assisted content distributed network

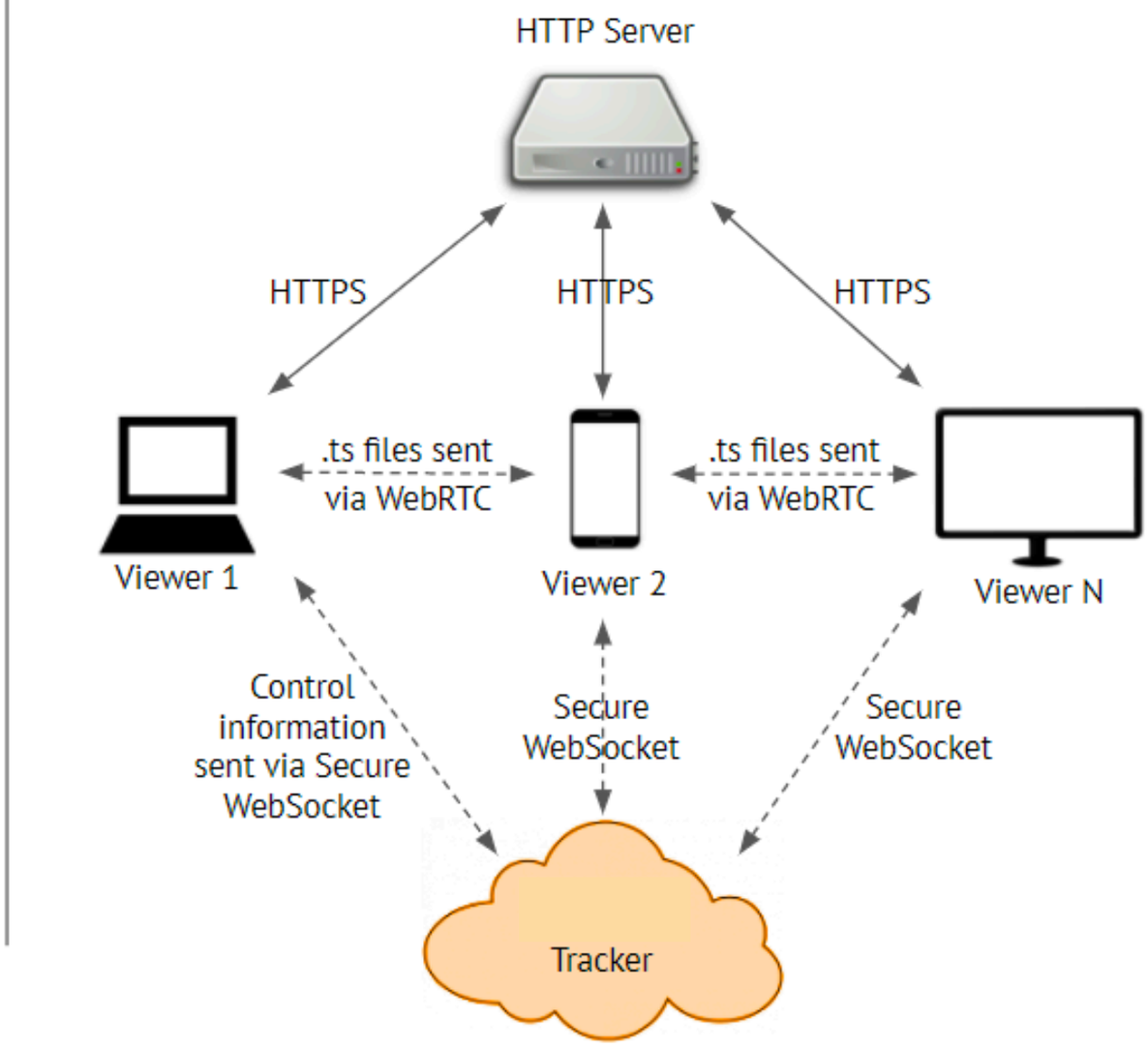


Everyone is naked @gpestana

Peer-assisted Content Distribution Network (pCD

Let's start with the case of a peer assisted content distributed network





Everyone is naked @gpestana

Peer-assisted Content Distribution Network (pCDN)

Everyone is naked @gpestana

Peer-assisted Content Distribution Network (pCDN)

Privacy verdict It’s a big mess™

Everyone is naked @gpestana

Distributed Hash Tables (DHT)

Goal decentralised key-value store

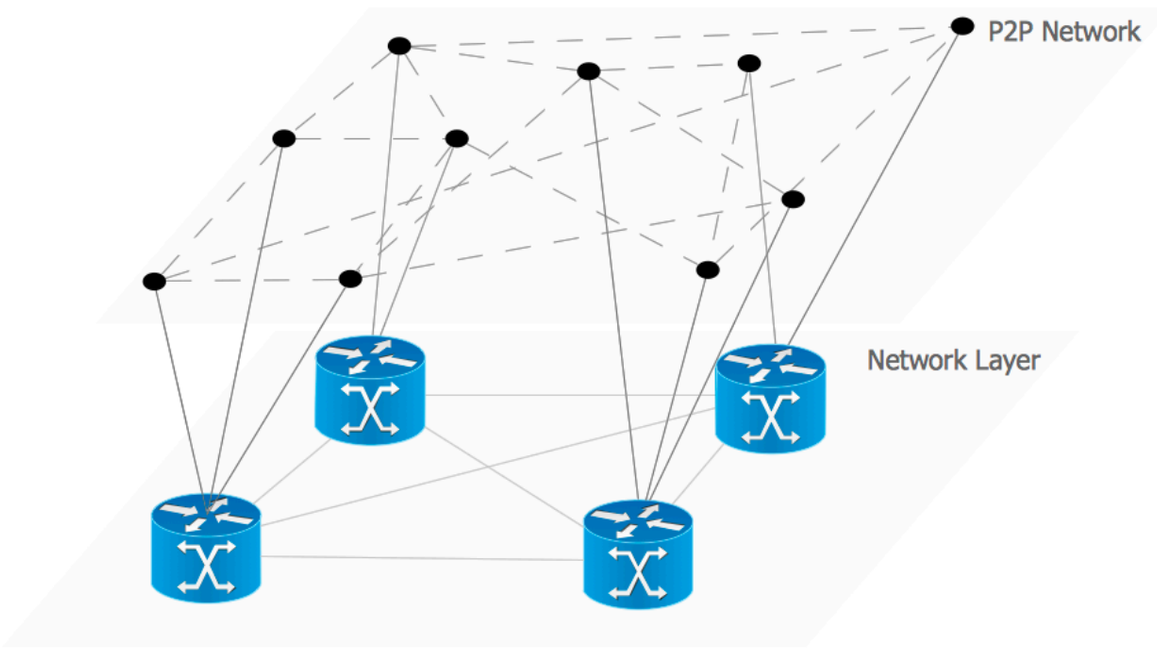
Nodes ID and resources ID share domain (node with unique ID 0010 is responsible for storing resource with iD 0010)

Nodes maintain a finger table with other peers info (reduced view of the network)

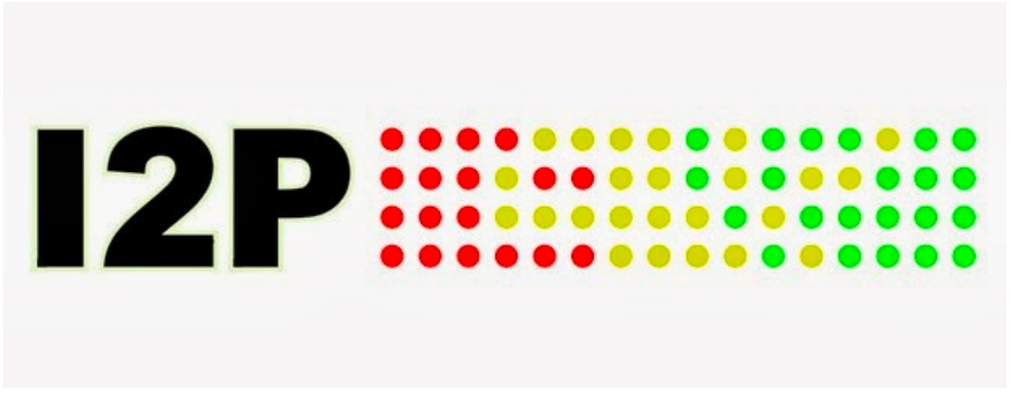
API

get(resource_id)

store(resource)



Distributed Hash Tables (DHT)



Distributed Hash Tables (DHT)

Everyone is naked @gpestana

Everyone is naked @gpestana

Everyone’s new clothes

Attacker learned who is interested in a specific content

Attacker learned personal information based on content cached by network peers

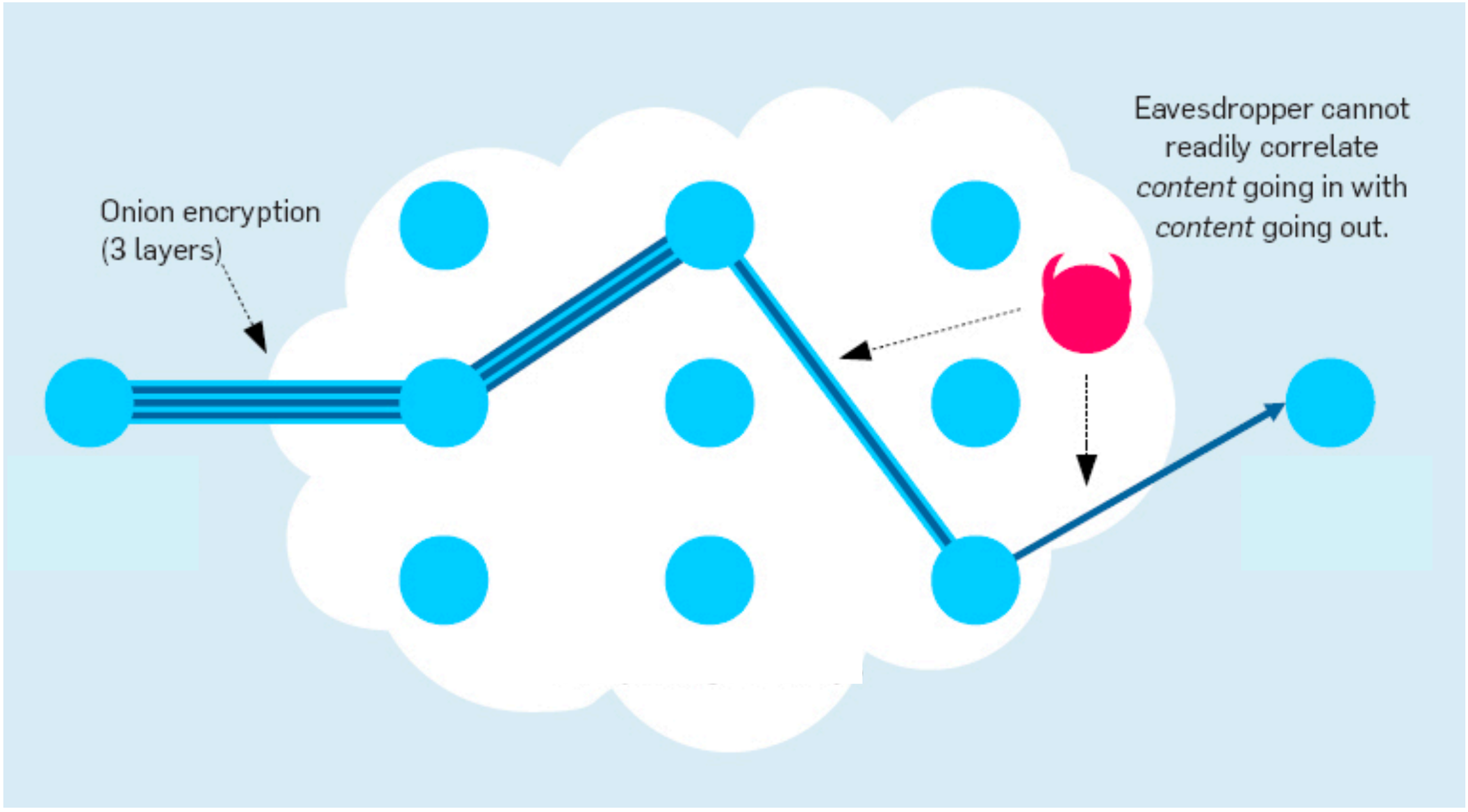
Location based, targeted attacks, global/local attacks

Everyone is naked @gpestana

Everyone is naked @gpestana

Everyone is naked @gpestana

Onion routing



Onion routing

https://github.com/hashmatter/p3lib

p3lib-sphinx, IPFS and onion routing demo

Everyone is naked @gpestana

Many other (possible, not tested, complex) solutions

Mixnets

Friend to Friend routing (**F2F**)

Private Information Retrieval (**PIR**) and Multi Party Computation (**MPC**)

Zero knowledge and **applied crypto** (e.g. randomised requests and responses)

Many other solutions

Added complexity

Encryption is free, privacy is not!

Not enough research and engineering

Mixnets

Friend to Friend routing (**F2F**)

Private Information Retrieval (**PIR**) and Multi Party Computation (**MPC**)

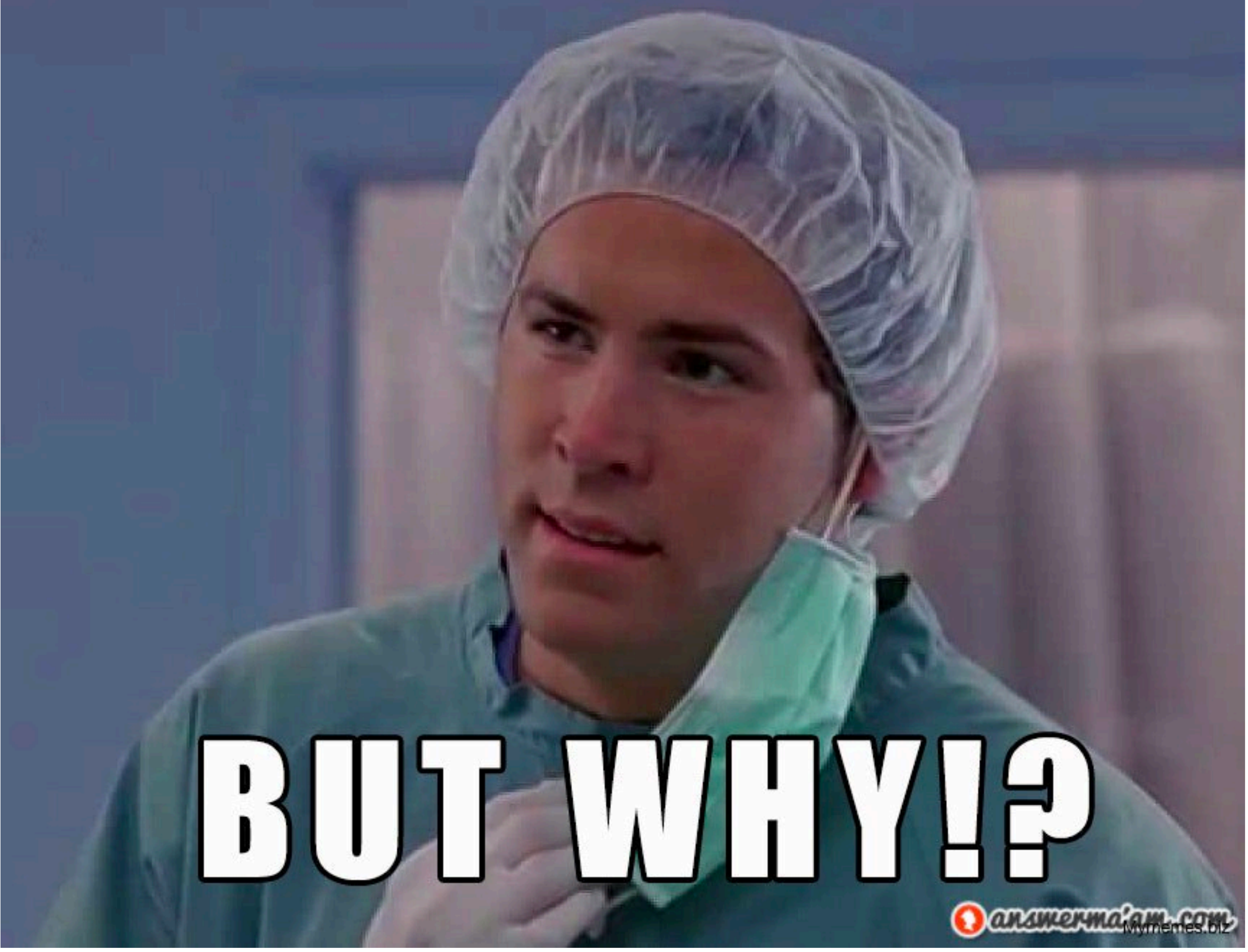
Zero knowledge and **applied crypto** (e.g. randomised requests and responses)

Putting some clothes on

How to create and maintain entropy/randomness for **plausible deniability** (I requested this resource but it does not mean I want it)

How to successfully collaborate without revealing full information (I got a request from a user but cannot get any info about its objective)

Everyone is naked @gpestana



Everyone is naked @gpestana

Everyone is naked

Privacy on peer-to-peer systems

hashmatter.com

github.com/gpestana/p2psec

gpestana.com

@gpestana