Enforcing Privacy in Online Services

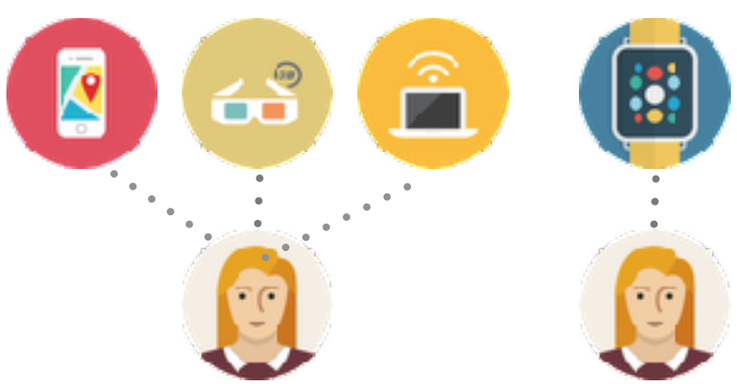
Sonia Ben Mokhtar
Directrice de Recherches CNRS

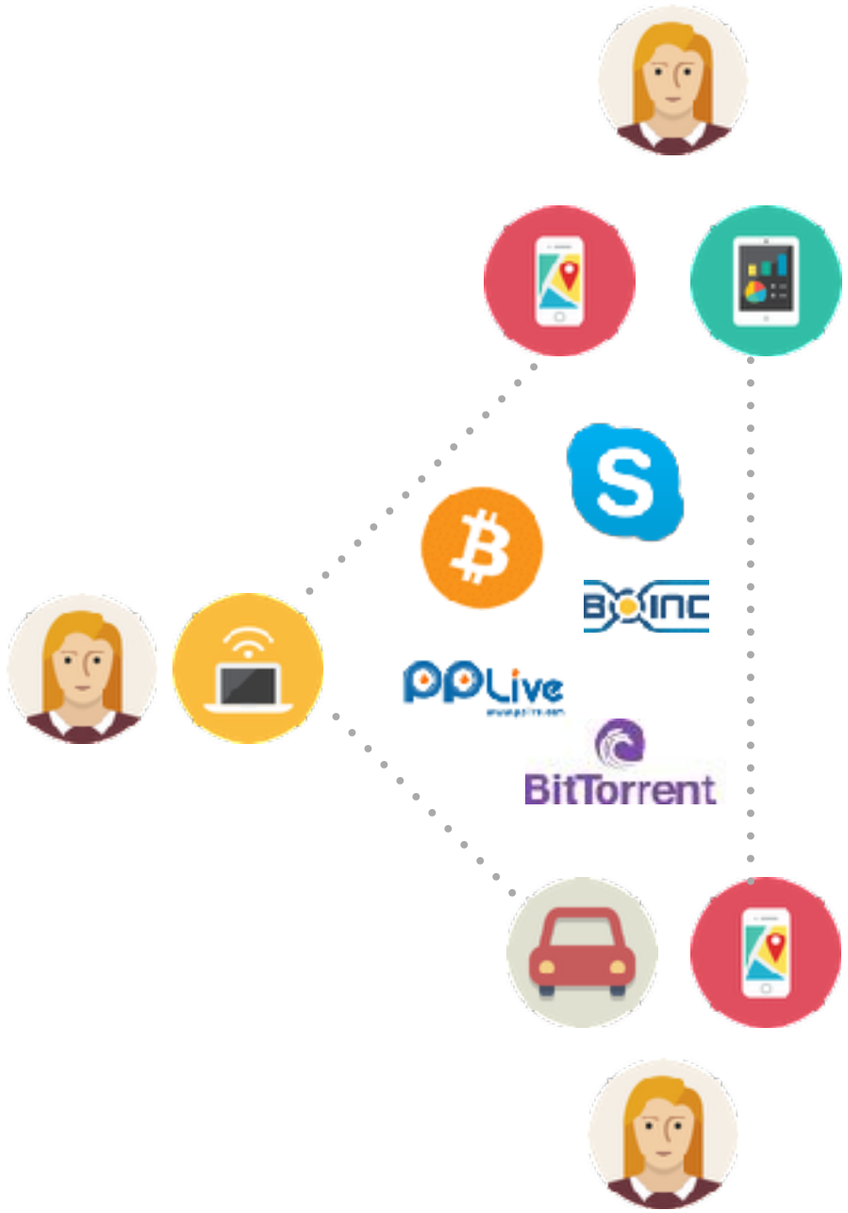
Paris P2P Festival
09/01/2020

Who am I?

- Head of the distributed systems & information retrieval group @LIRIS lab, Lyon, France
  - liris.cnrs.fr/drim
- Research topics
  - Distributed systems
    - Fault-tolerance
    - Performance
    - *Privacy*

Today's Distributed Systems





An example: Web Search

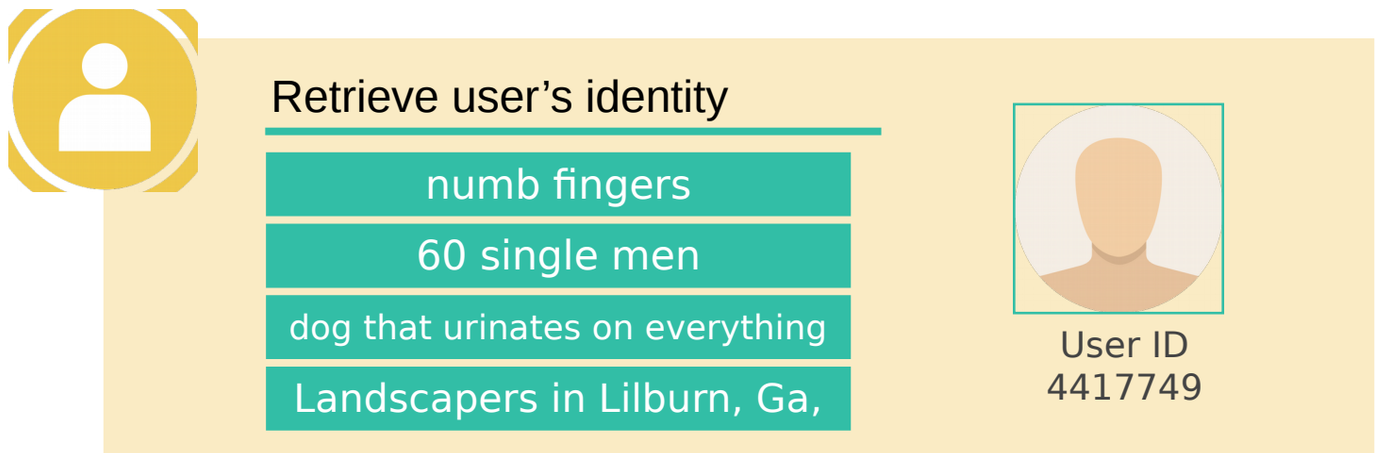
Every day, millions of users are querying **SEARCH ENGINES**

> We also use this information [*that we collect from all of our services*] to offer you tailored content – like giving you more **relevant search results** and **ads**.

http://www.google.com/policies/privacy/

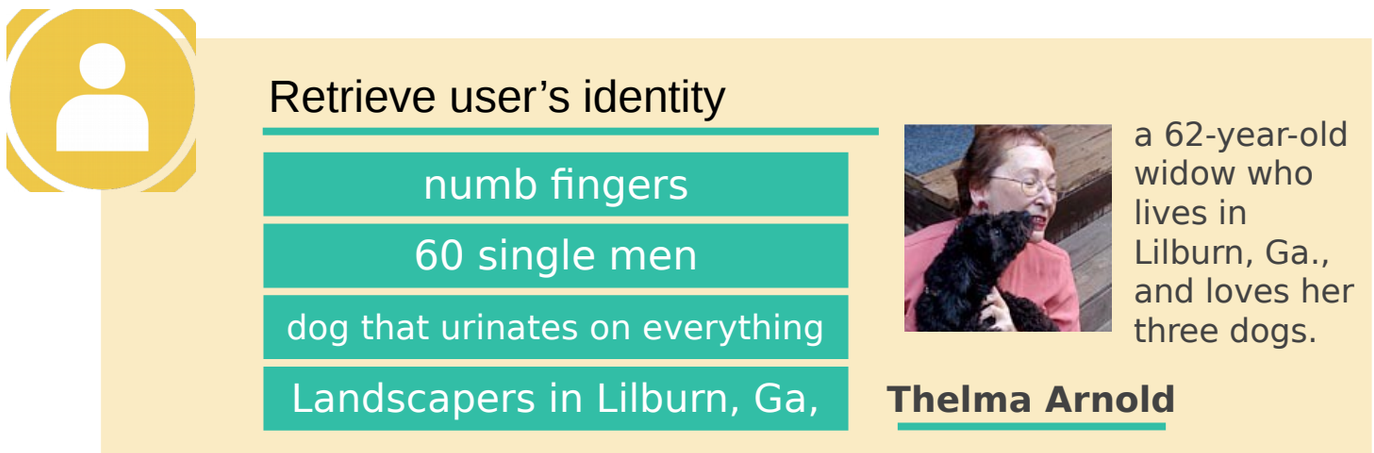
USER PROFILES

Web Search: Privacy Threats



Barbaro, Michael, Tom Zeller, and Saul Hansell. "A face is exposed for AOL searcher no. 4417749." New York Times 9.2008 (2006): 8For.

Web Search: Privacy Threats



Barbaro, Michael, Tom Zeller, and Saul Hansell. "A face is exposed for AOL searcher no. 4417749." New York Times 9.2008 (2006): 8For.

Web Search: Privacy Threats

Retrieve user's identity

numb fingers

60 single men

dog that urinates on everything

Landscapers in Lilburn, Ga,

a 62-year-old widow who lives in Lilburn, Ga., and loves her three dogs.

Thelma Arnold

Barbaro, Michael, Tom Zeller, and Saul Hansell. "A face is exposed for AOL searcher no. 4417749." New York Times 9.2008 (2006): 8For.

Infer extra information

Age

Zip Code

Religion

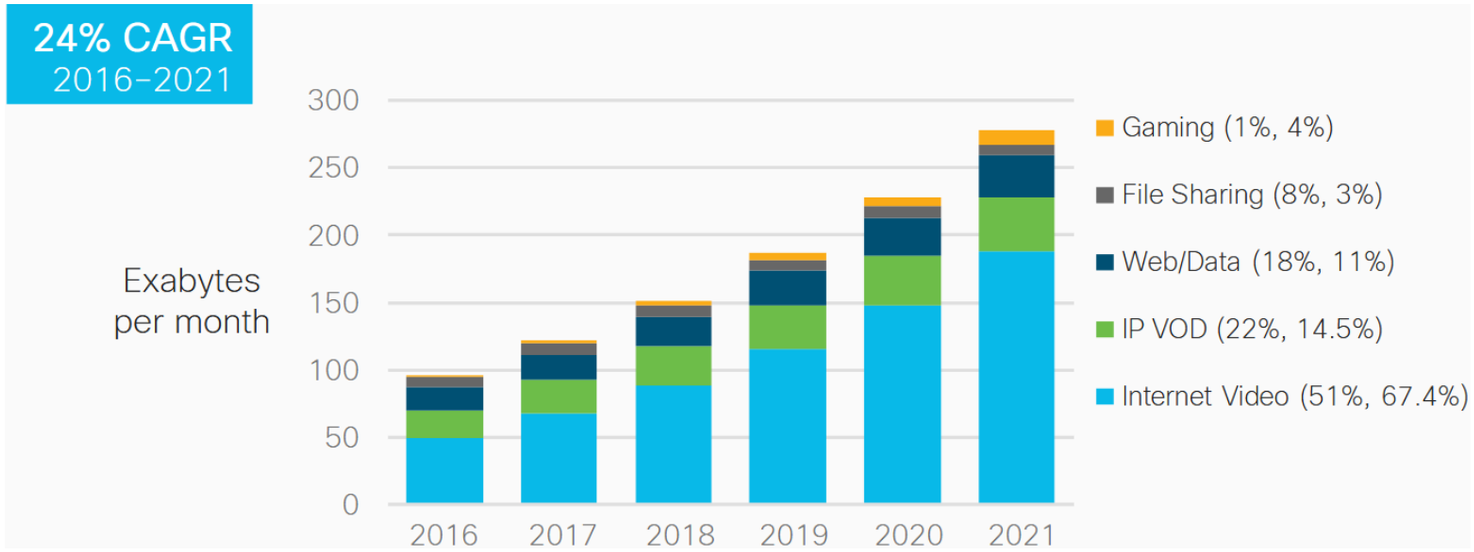
Gender

Diseases

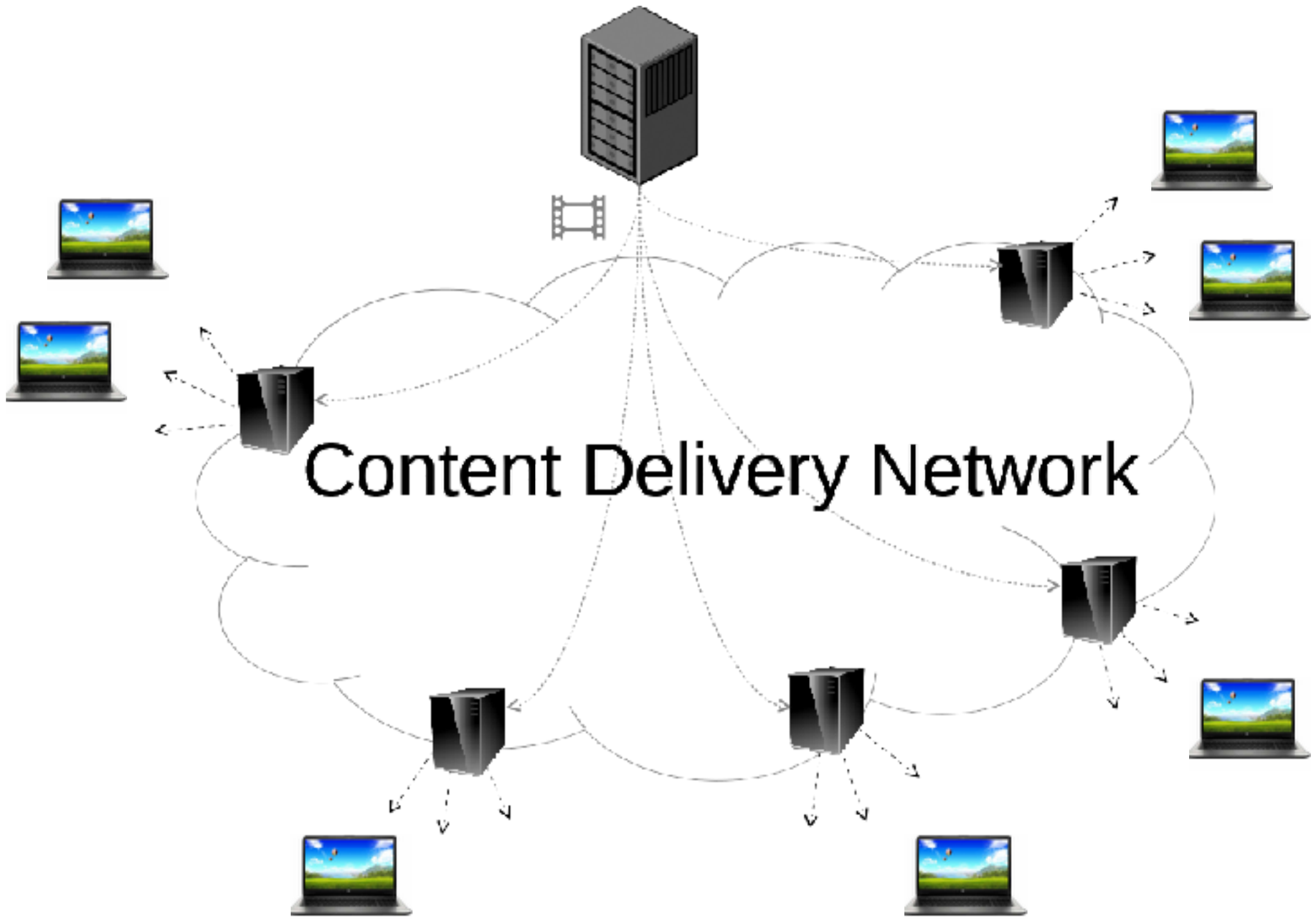
Interests

Jones, Rosie, et al. "I know what you did last summer: query logs and user privacy." Proceedings of the sixteenth ACM conference on Conference on information and knowledge management. ACM, 2007.

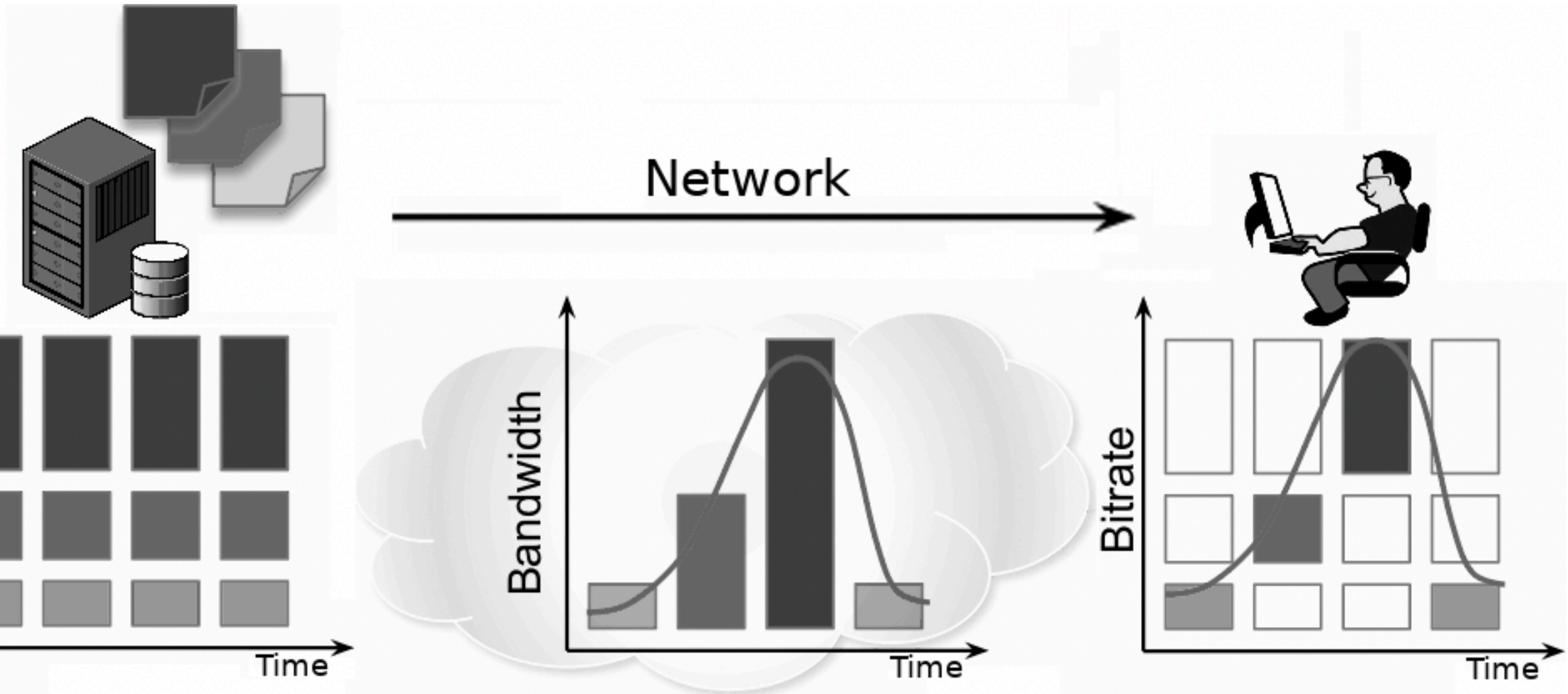
Another Example: Video Streaming



Streaming over CDNs

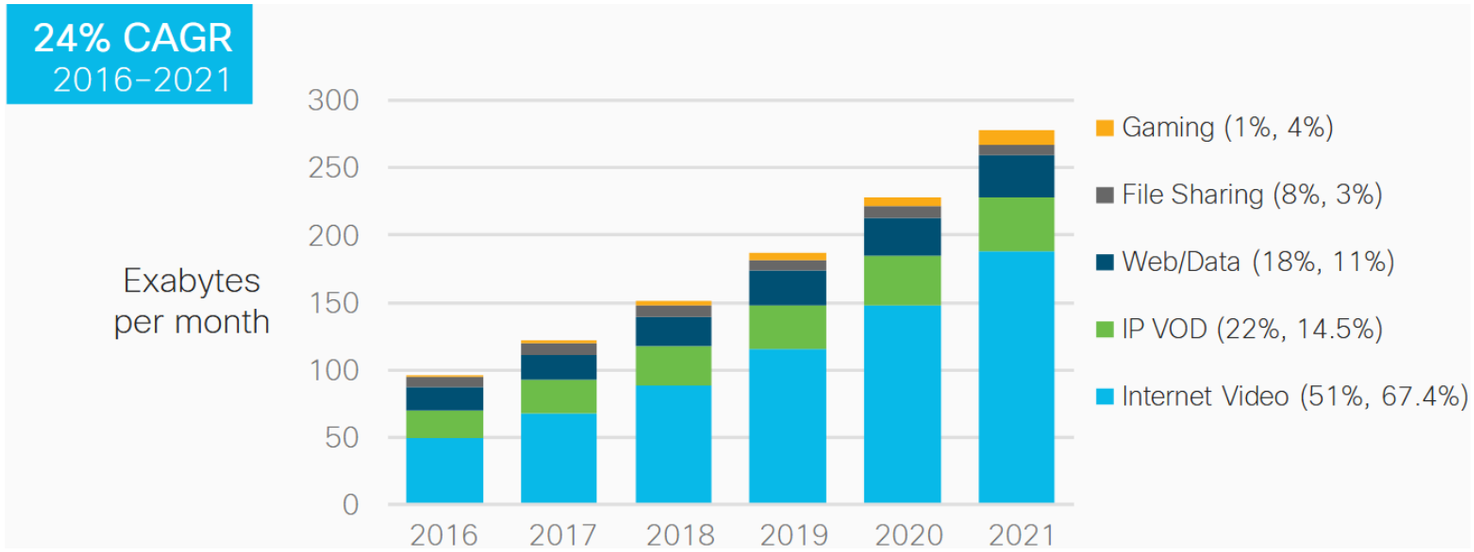


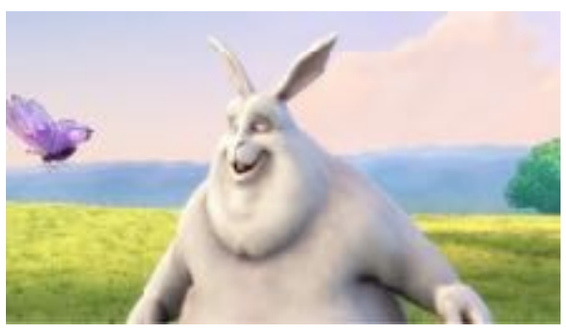
Streaming with DASH

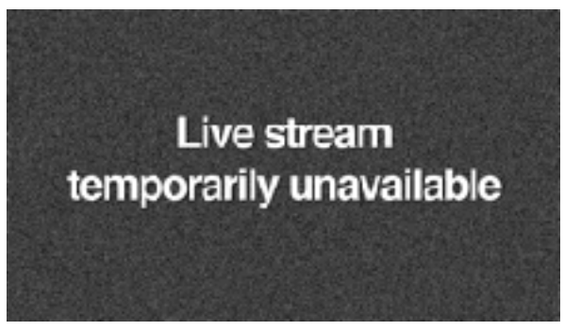


Still…

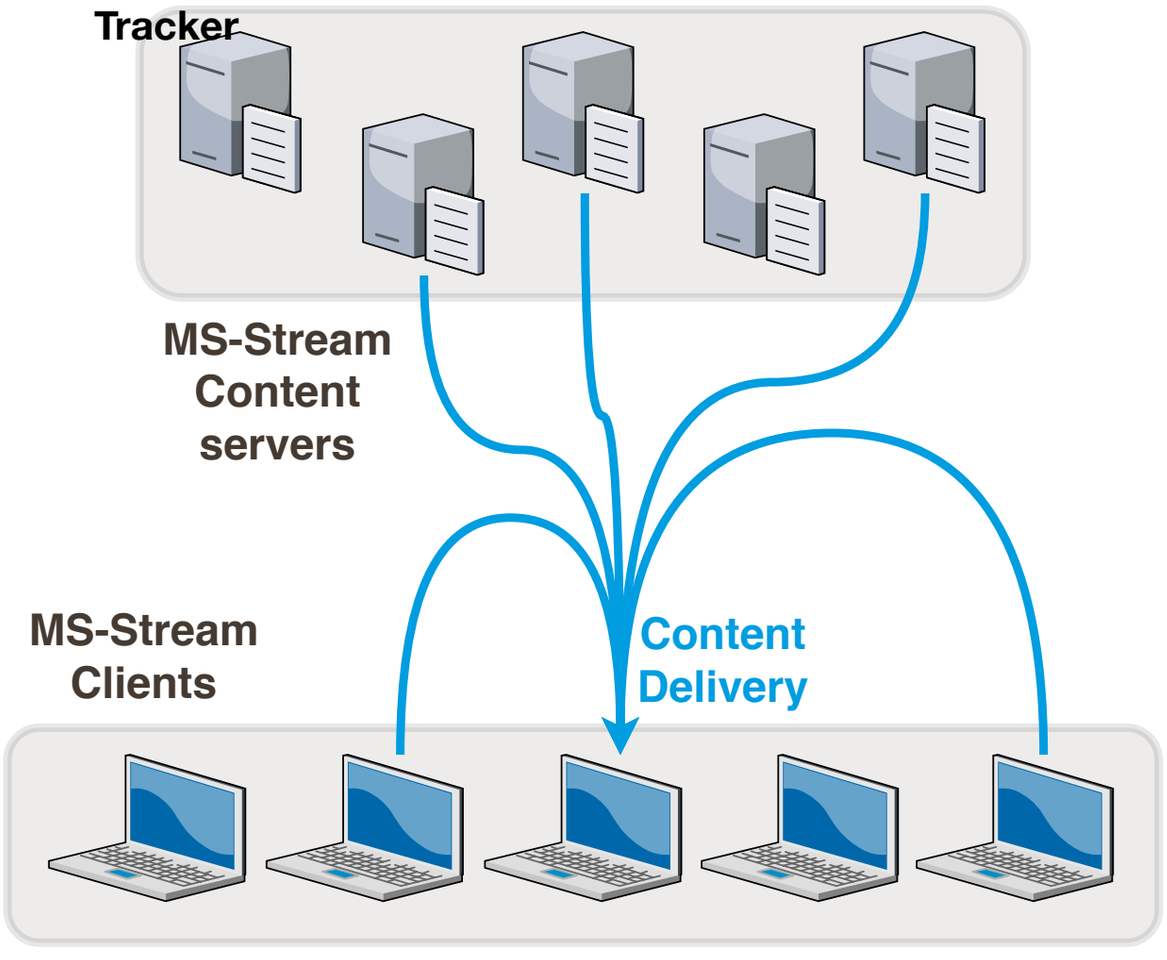

NETFLIX
YouTube
twitch

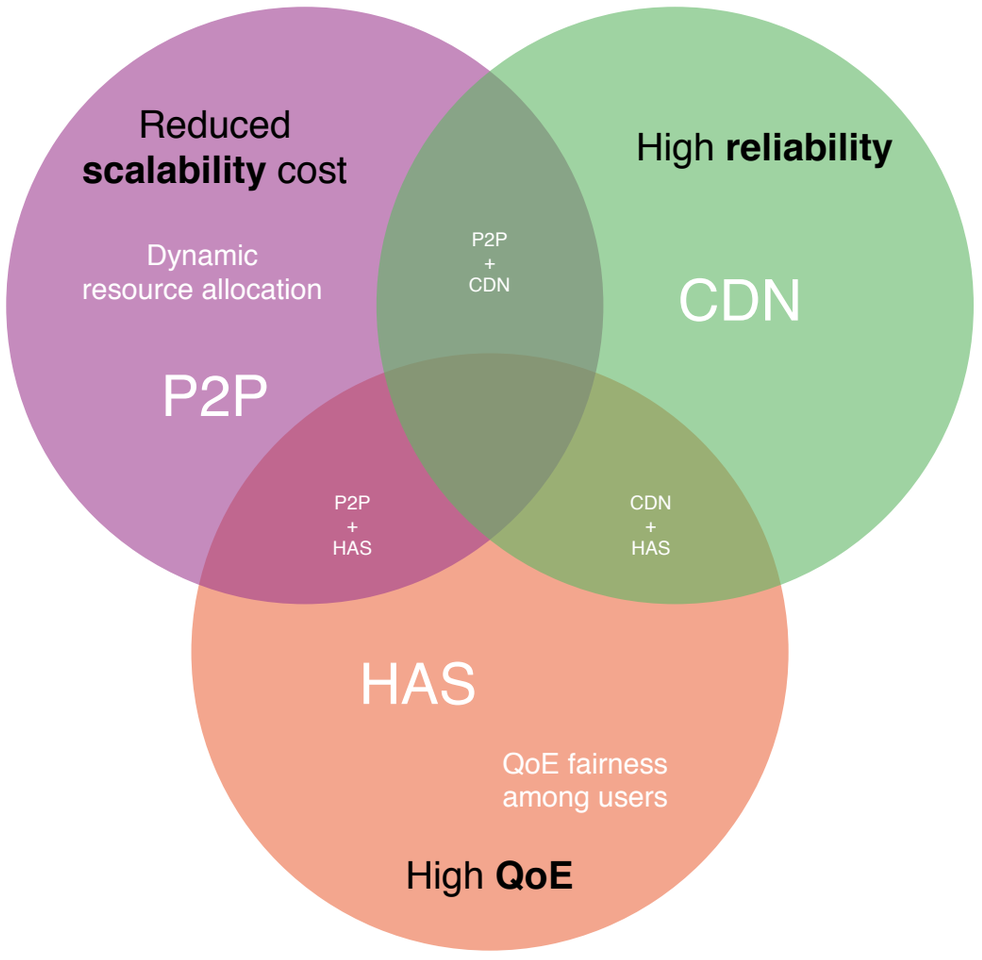

24% CAGR
2016–2021
Exabytes
per month
300
250
200
150
100
50
0
2016
2017
2018
2019
2020
2021
Gaming (1%, 4%)
File Sharing (8%, 3%)
Web/Data (18%, 11%)
IP VOD (22%, 14.5%)
Internet Video (51%, 67.4%)



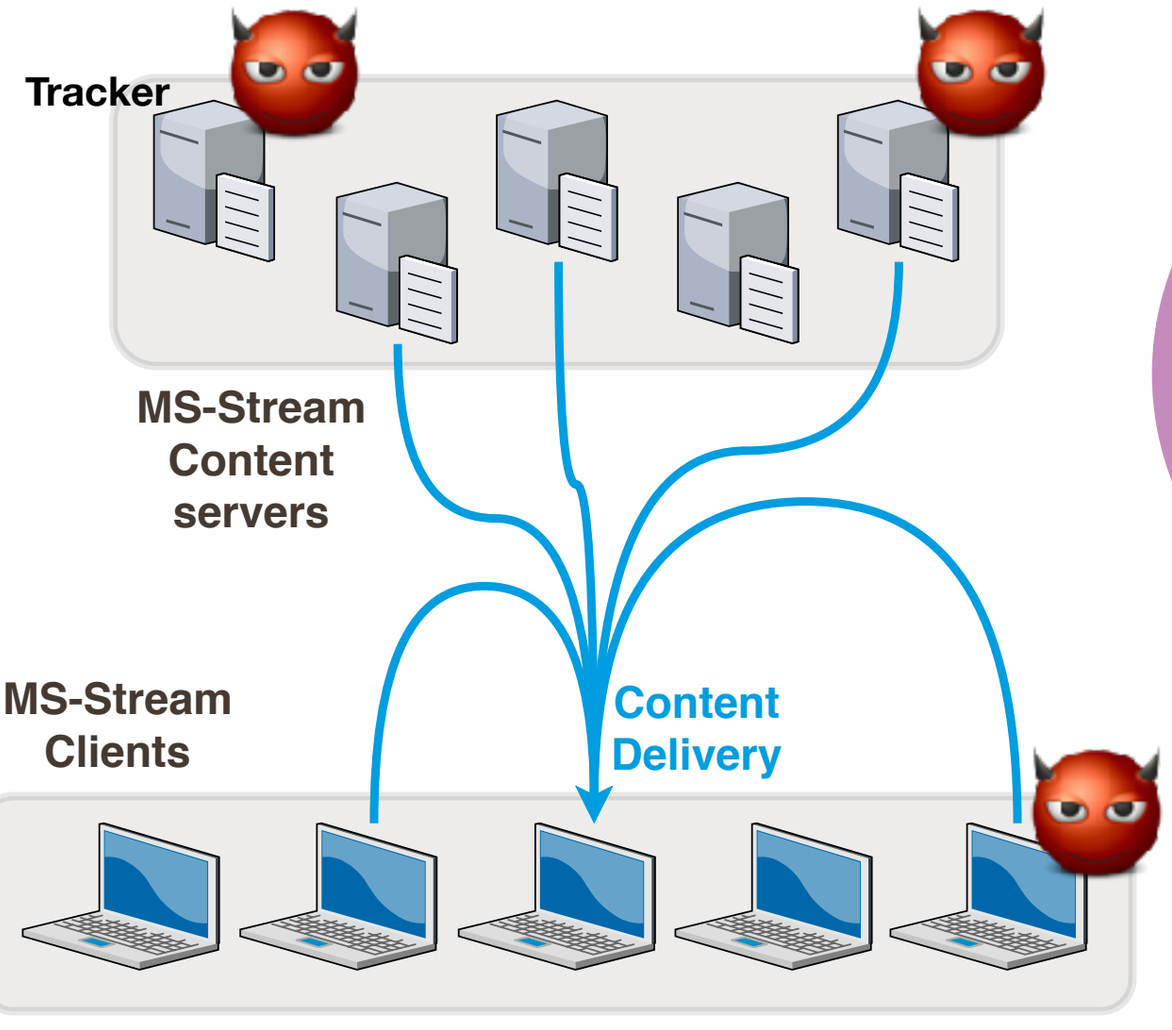

Live stream
temporarily unavailable

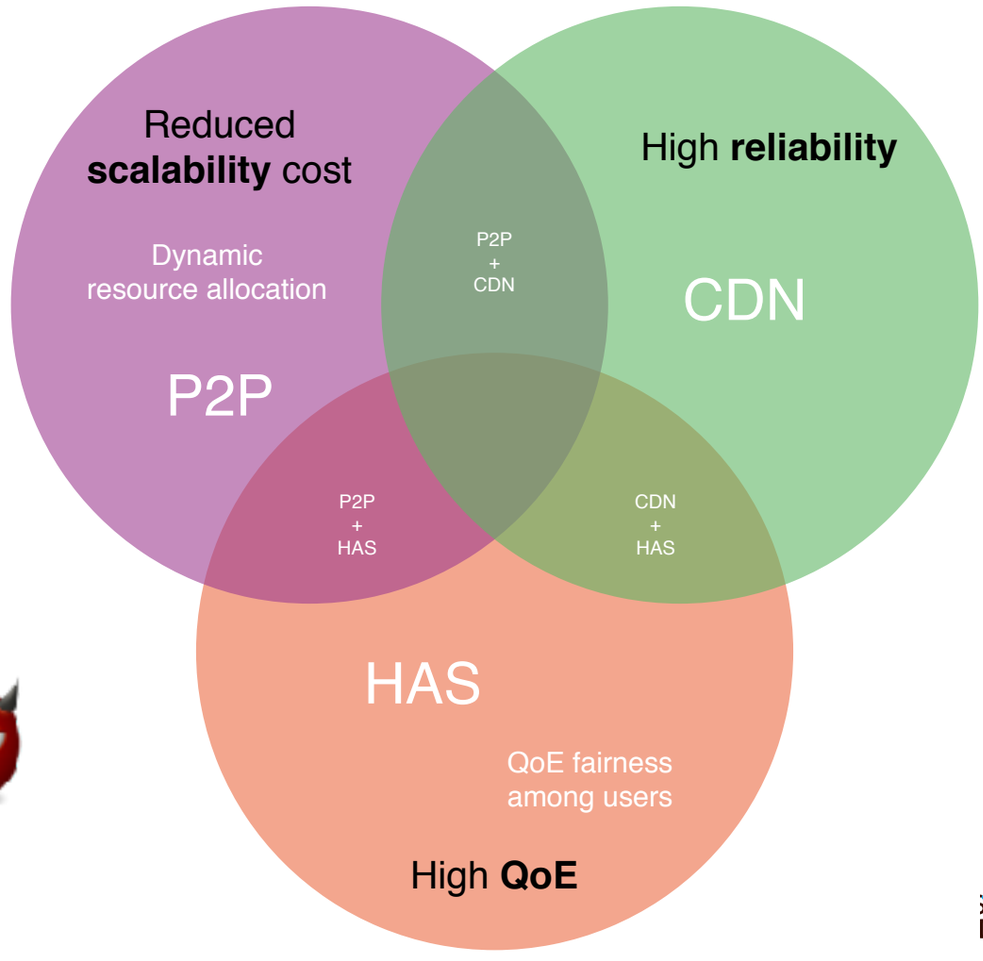
Multi-Source Streaming





Privacy issues



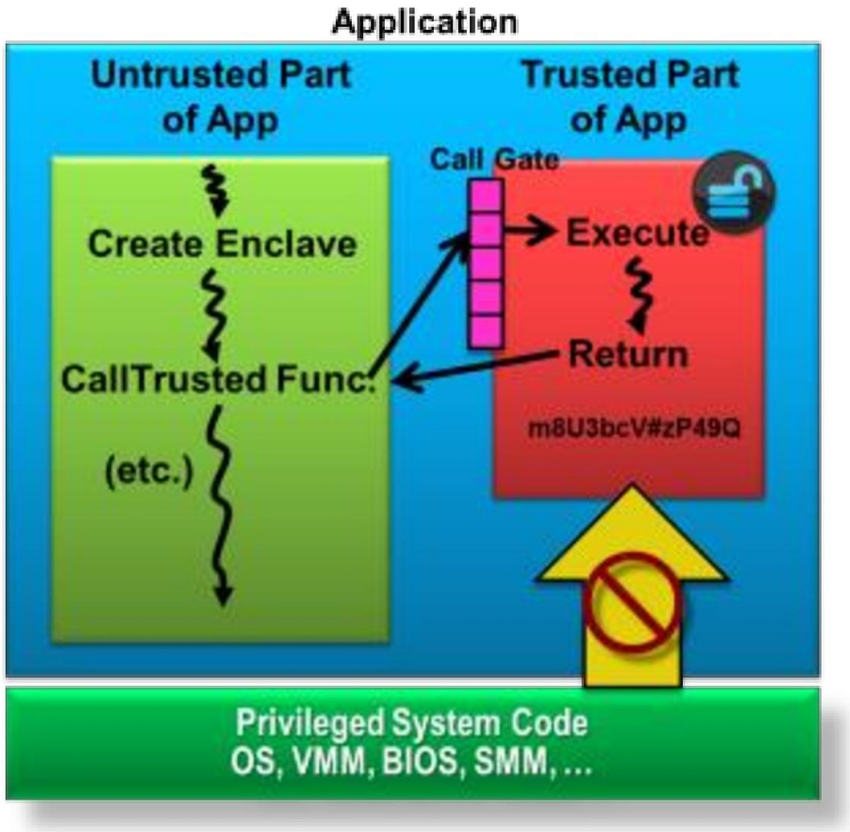


Outline

- Introduction
- Motivation/Privacy Threats
- **Enforcing Privacy in Online Services**
  - Using Intel SGX processors and P2P
    - Decentralized Proxy Service for Web Search
    - Edge-assisted Video Streaming
- Conclusion & Perspectives

SGX: Software Guard Extensions

- New instruction set since Intel Skylake processors (2015)
- Provides a protected environment called *enclave*
  - Memory encryption, integrity and freshness
  - Not even the OS or hypervisor are able to inspect
  - Suitable for using in hostile environments (cloud)



- Limitations:
  - Memory usage is limited to 128 MB per CPU

Outline

- Introduction
- Motivation/Privacy Threats
- Enforcing Privacy in Online Services
  - **Using Intel SGX processors and P2P**
    - Decentralized Proxy Service for Web Search
    - Edge-assisted Video Streaming
- Conclusion & Perspectives

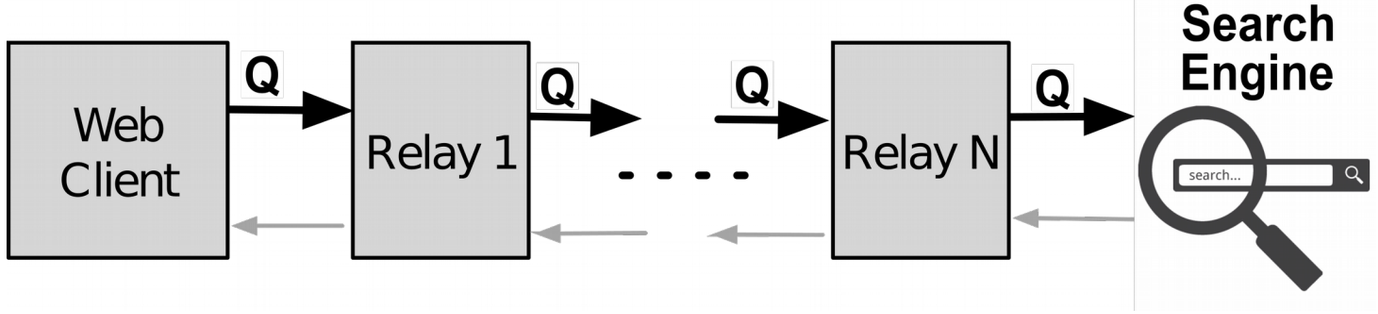
Private Web search

- How can users protect their privacy from curious search engines?

1. Hiding identities (IP Address)
2. Making queries and user’s interests indistinguishable

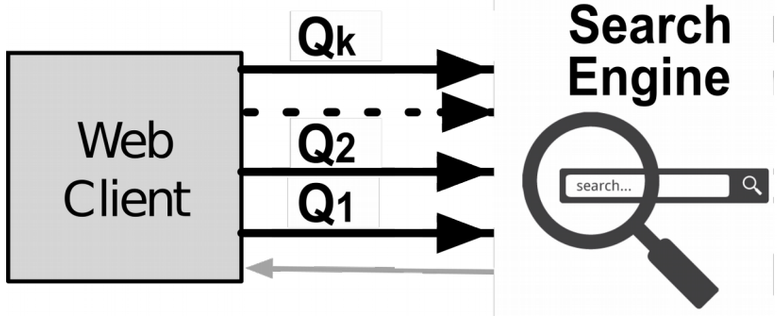
Unlinkability

Unlinkability between user and query (Tor)

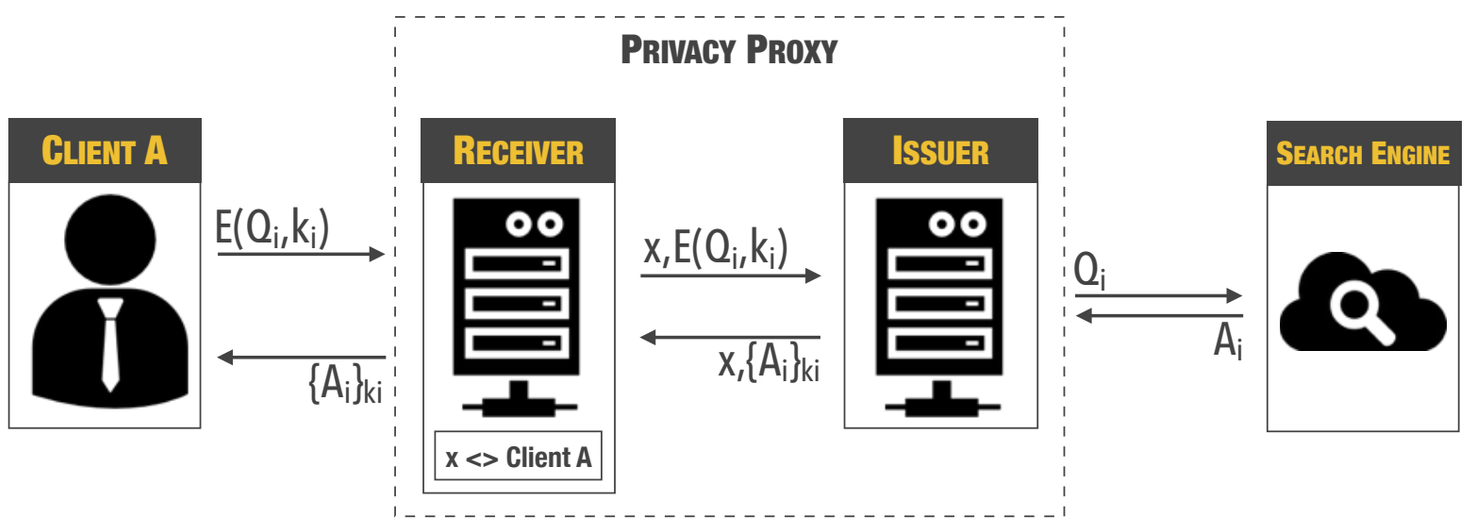


Indistinguishability

Indistinguishability between real and fake queries (TrackMeNot)



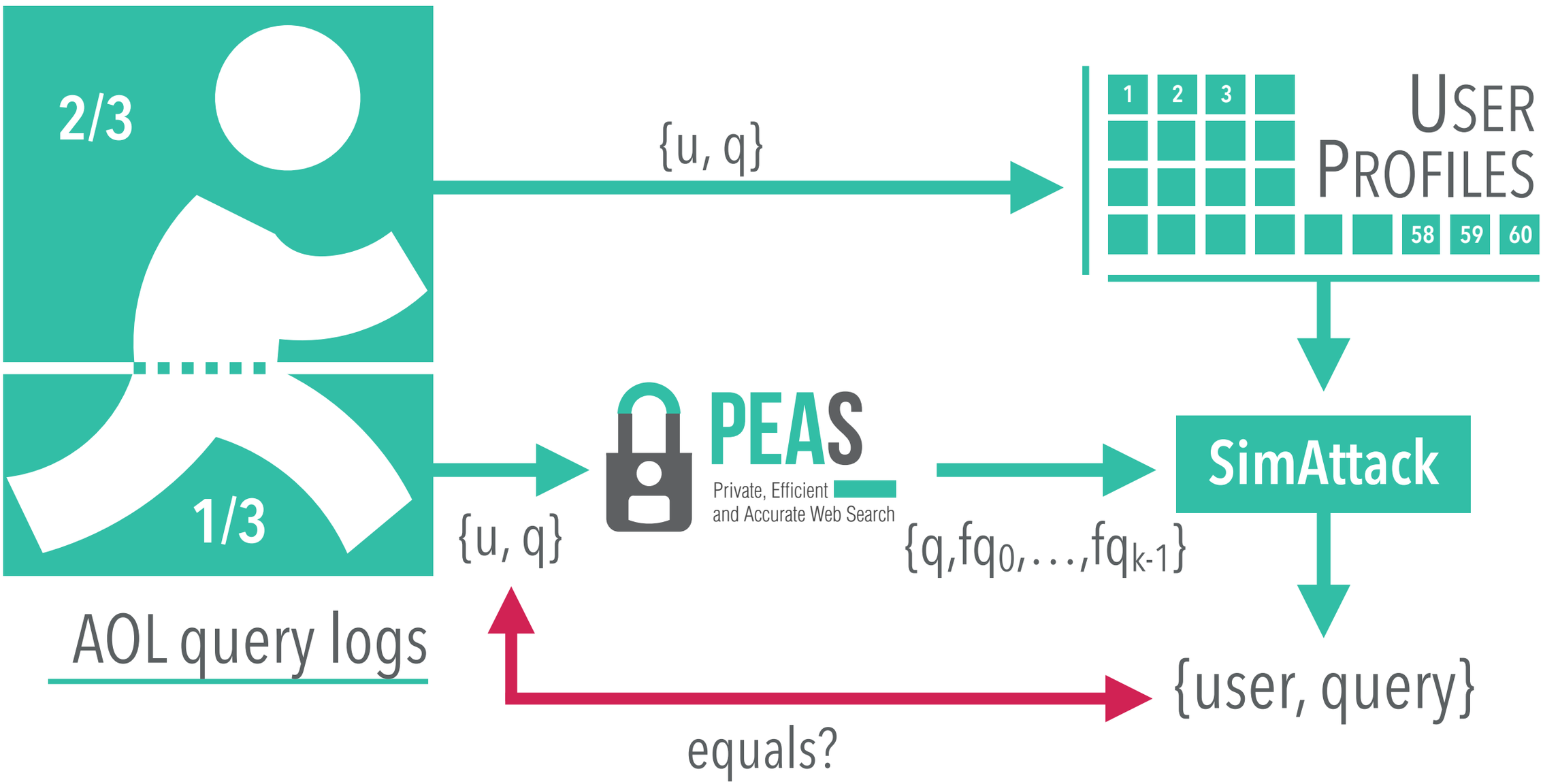
Unlink. + Indisting.: PEAS



$E(m)$	RSA encryption of message m with the public key of the issuer
$\{m\}_i$	AES encryption of message m with key K_i
Q_i	i-th query of user U
K_i	AES encryption key associated with query Q_i
A_i	Answer to query Q_i
X	An anonymous identifier

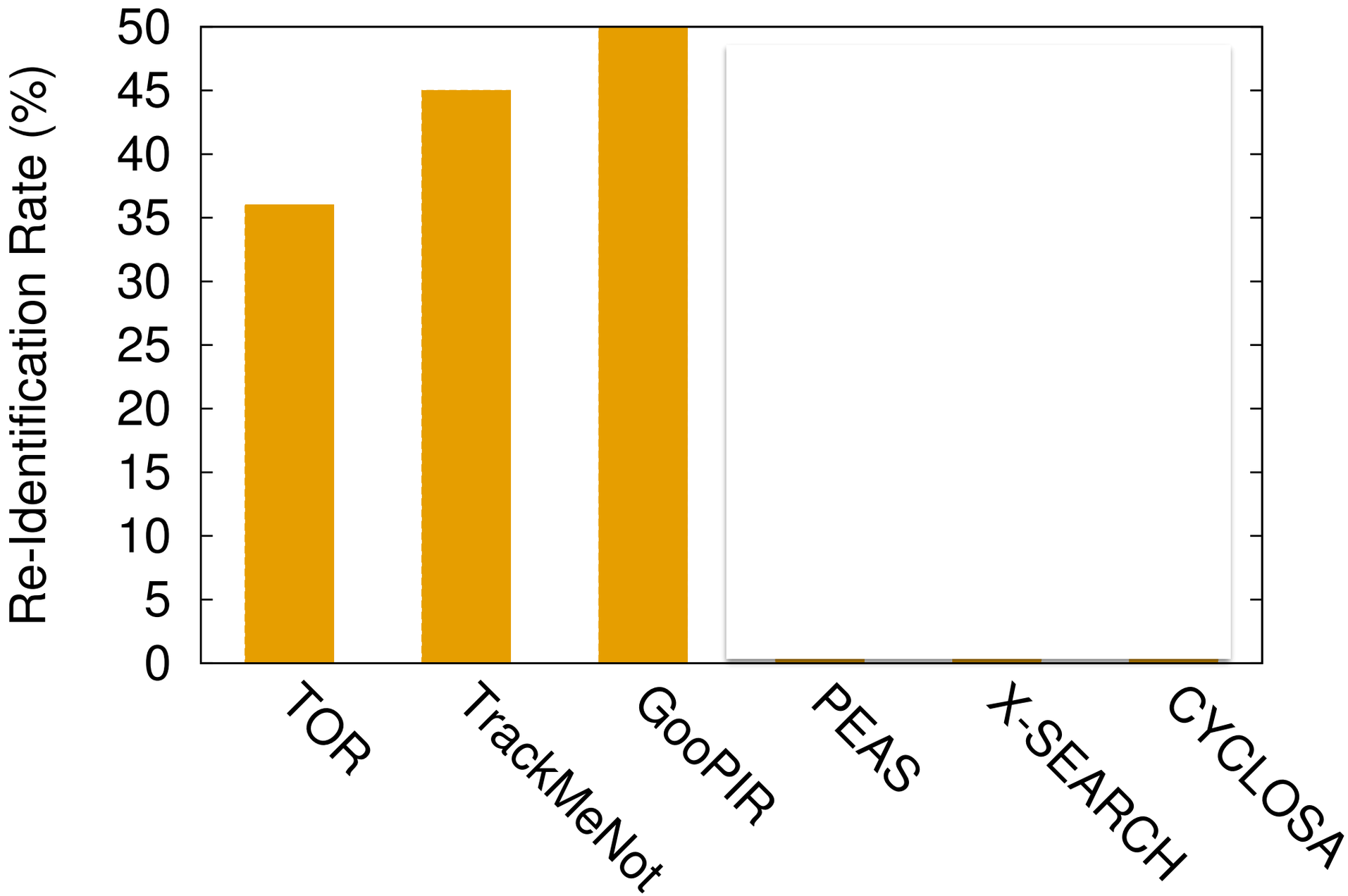
PEAS: Private, Efficient and Accurate Web Search. IEEE TrustCom'15. 2015.

Measuring Privacy

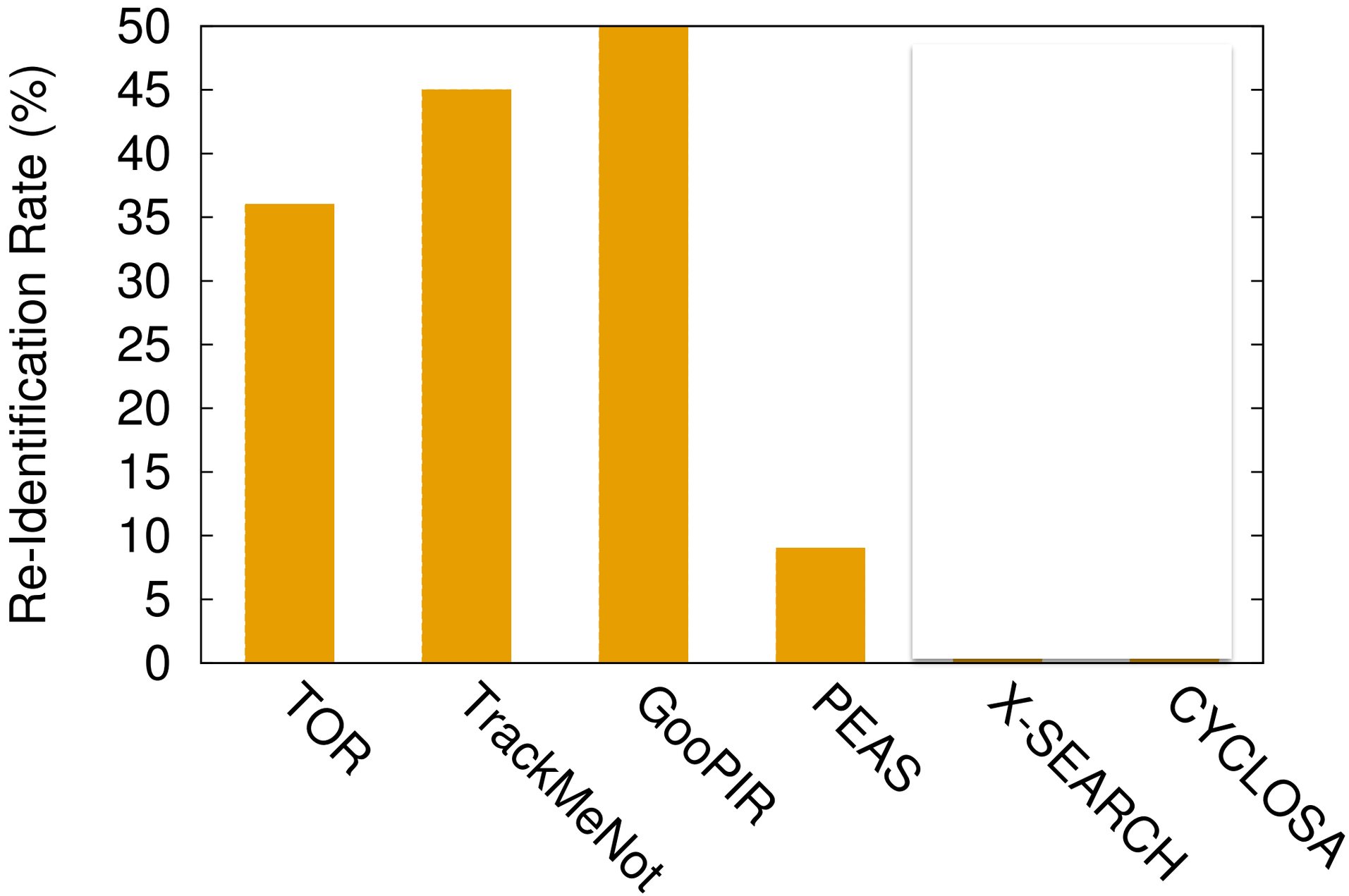


SimAttack: private web search under fire. Journal of Internet Services and Applications 7(1): 2:1-2:17 (2016).

Measuring privacy



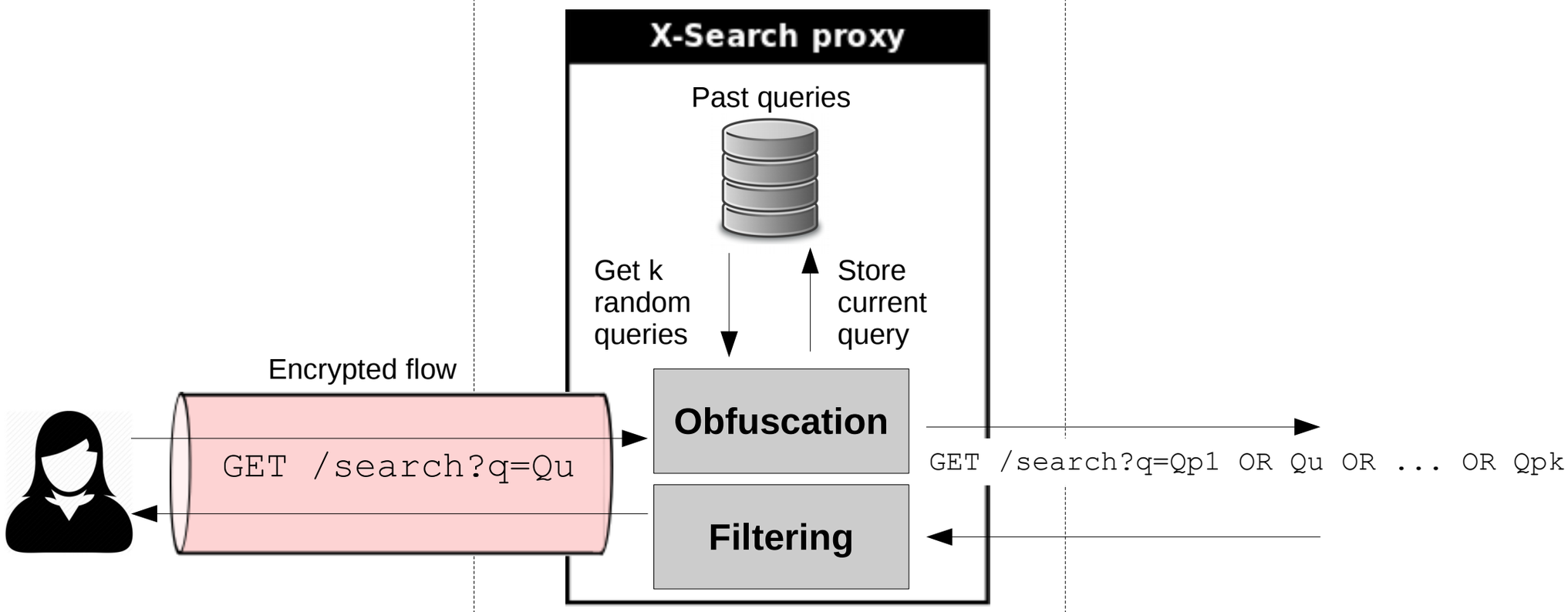
Measuring privacy



PEAS limitations

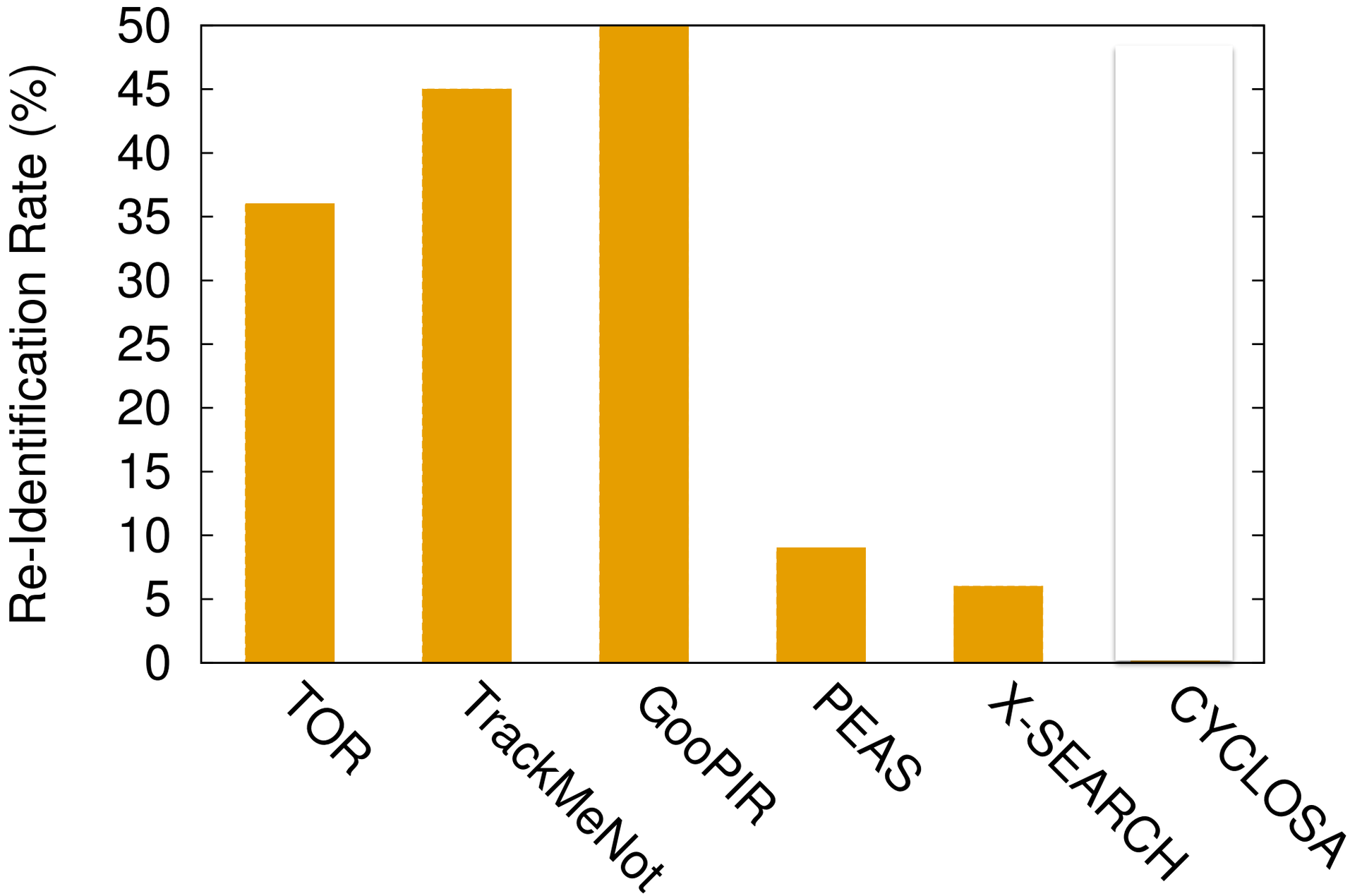
- Weak adversarial model
  - Relies on two non colluding servers
- Quality of fake queries
- Scalability

X-Search



X-Search: Revisiting Private Web Search using Intel SGX. Middleware 2017.

Measuring privacy

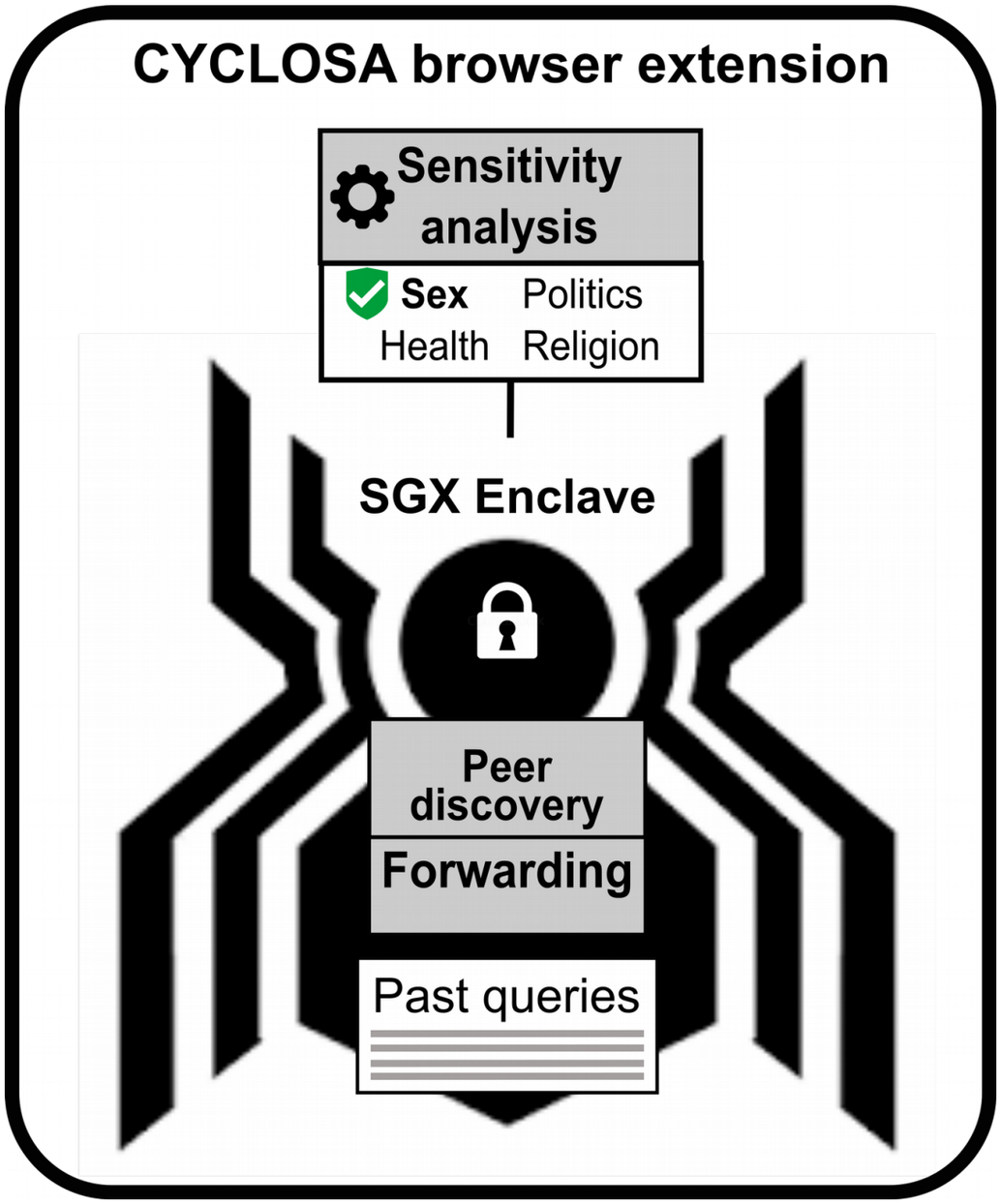


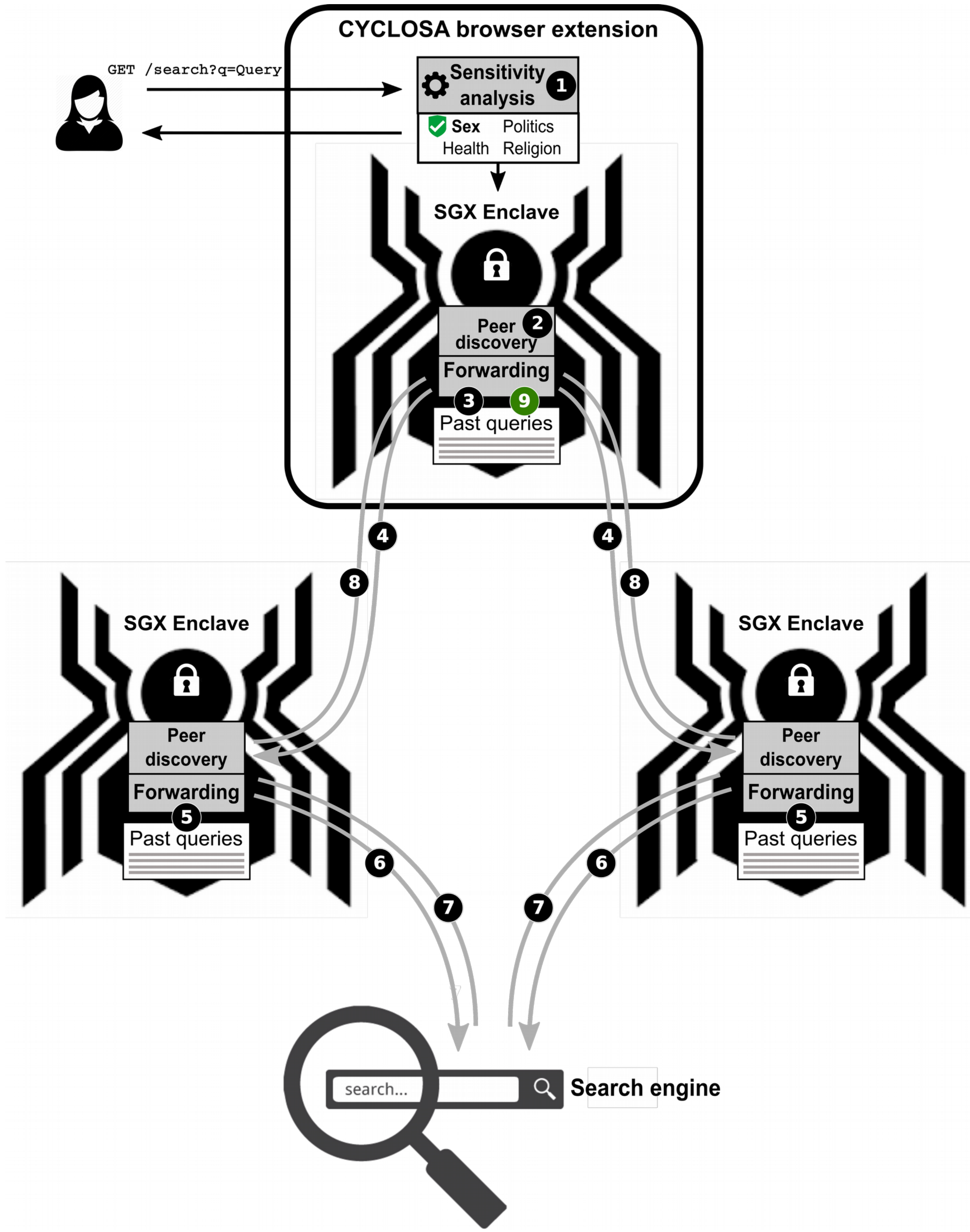
X-Search Limitations

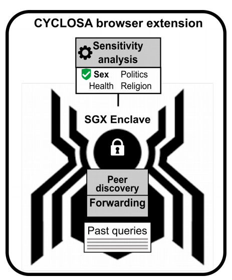
- Scalability
- Query limitation wrt search engine
- Accuracy

Cyclosa Architecture

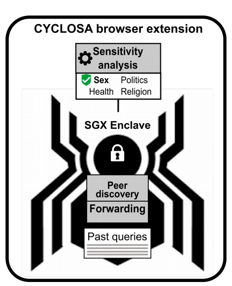
- Every node in the system acts as a proxy node for others
- Use Intel SGX
- Built as a browser extension
- Considers query sensitivity



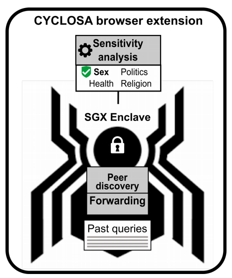

CYCLOSA browser extension
GET /search?q=Query
Sensitivity analysis
1
Sex
Politics
Health
Religion
SGX Enclave
Peer discovery
2
Forwarding
3
9
Past queries
4
4
8
8
SGX Enclave
Peer discovery
Forwarding
5
Past queries
SGX Enclave
Peer discovery
Forwarding
5
Past queries
6
7
7
6
search...
Search engine



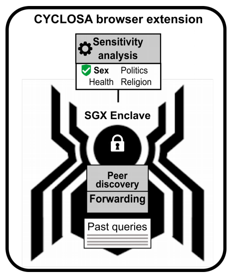
20



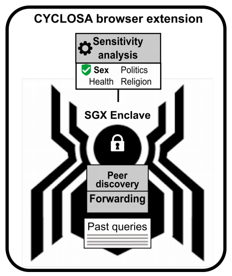
20



20

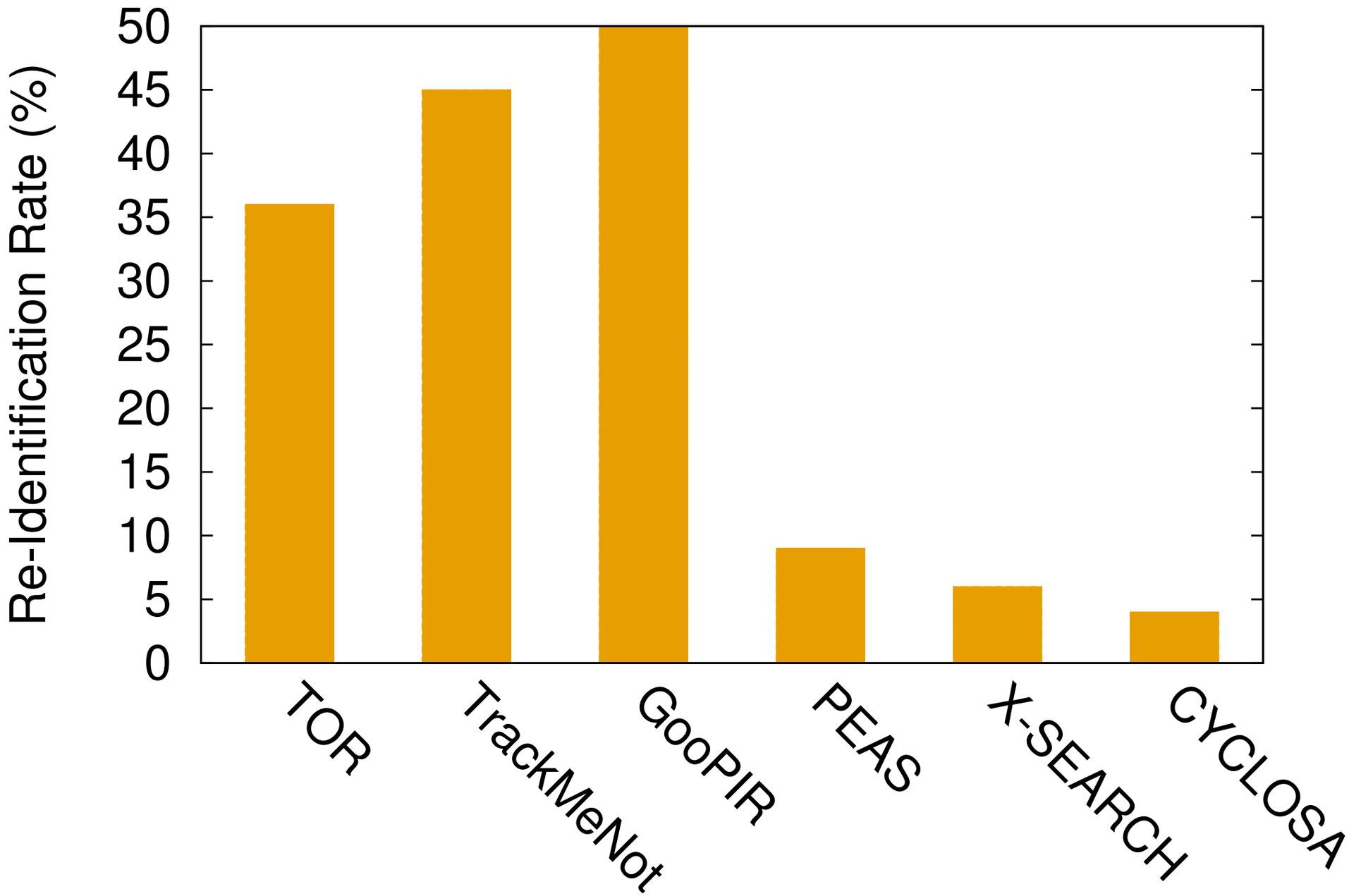


20



20

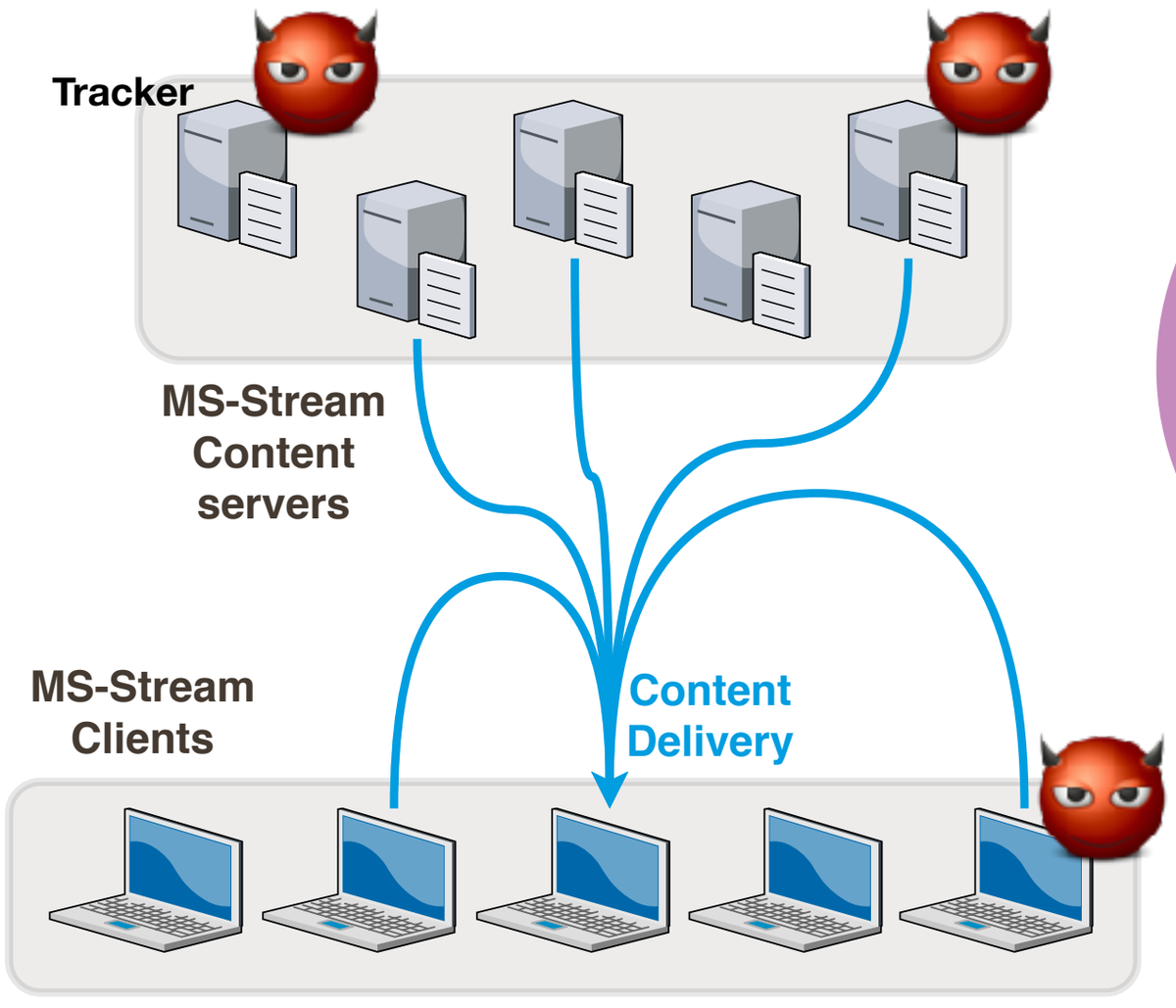
Measuring privacy

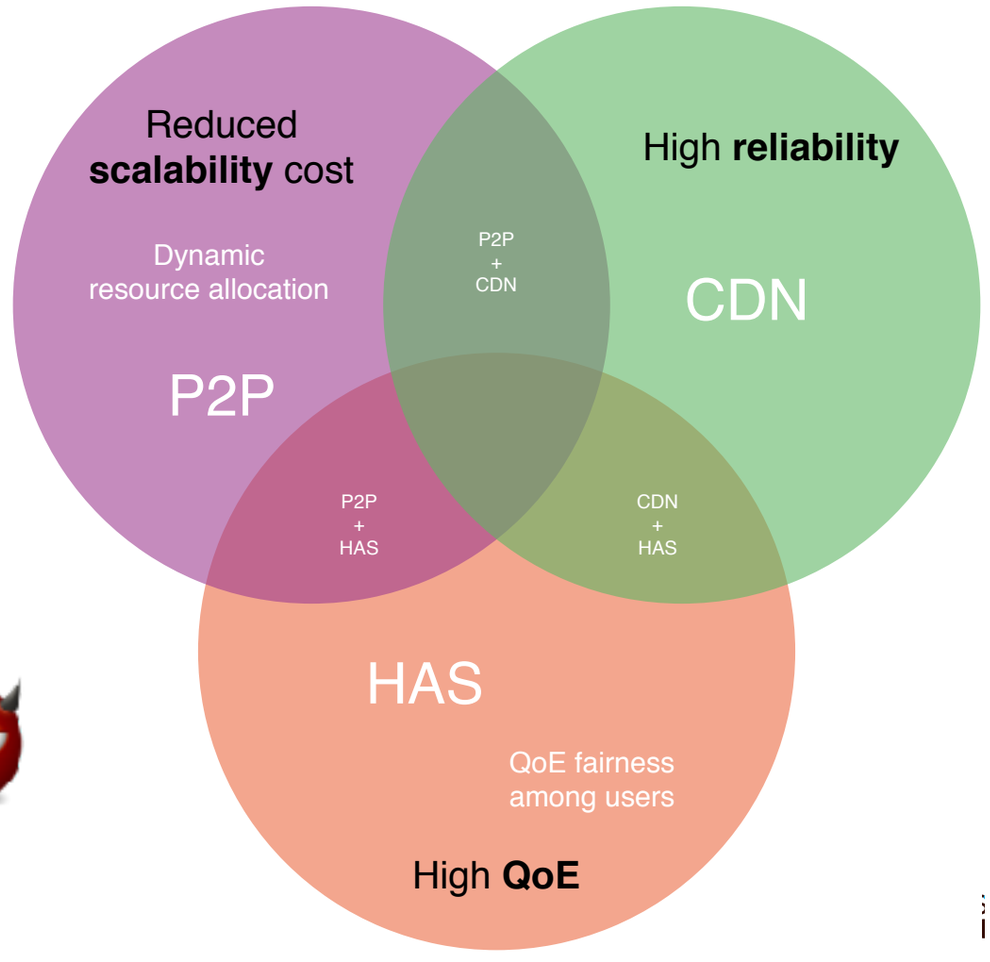


Outline

- Introduction
- Motivation/Privacy Threats
- Enforcing Privacy in Online Services
  - Using Intel SGX processors and P2P
    - Decentralized Proxy Service for Web Search
    - **Edge-assisted Video Streaming**
- Conclusion & Perspectives

Back to Video Streaming

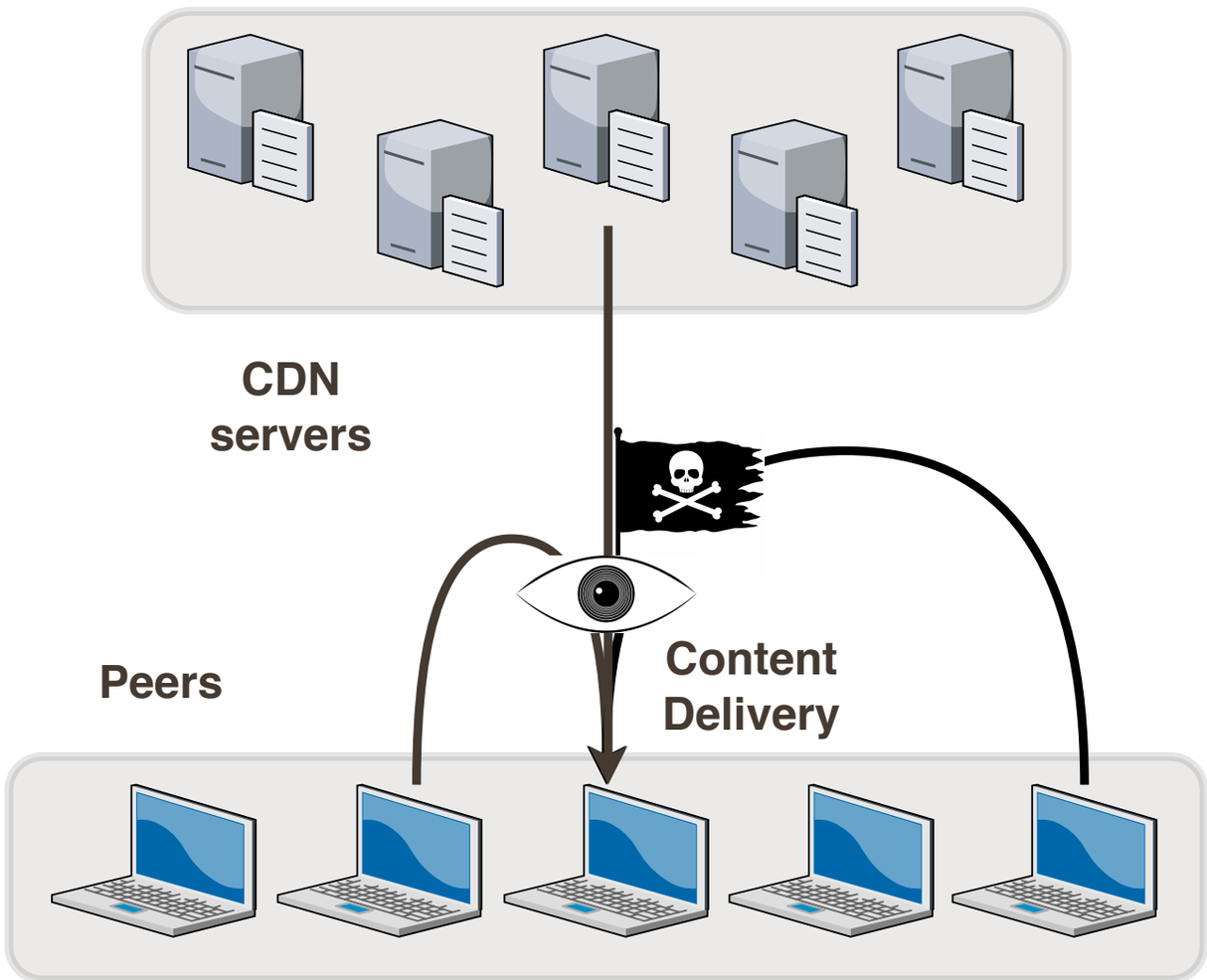


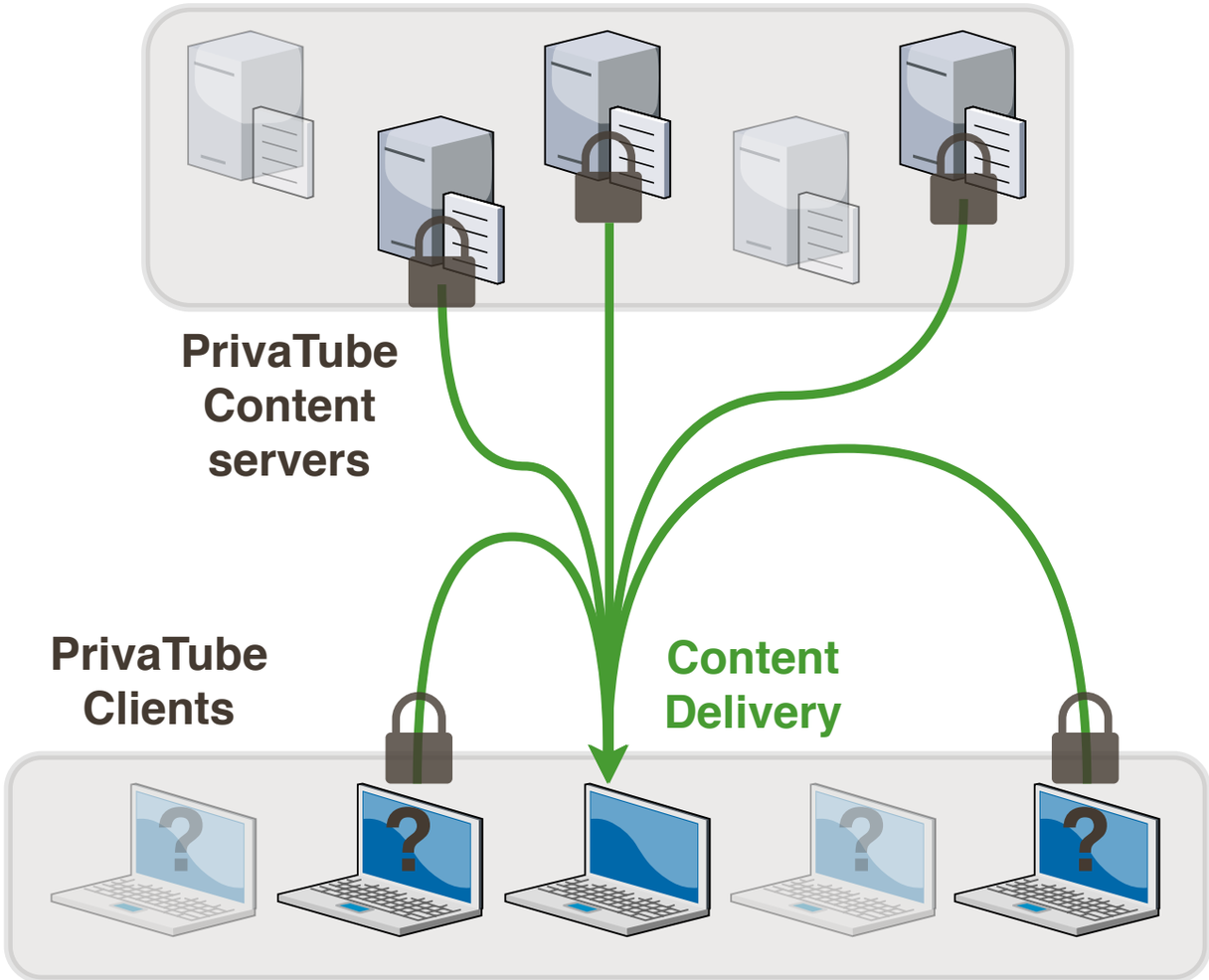


Privacy Objectives

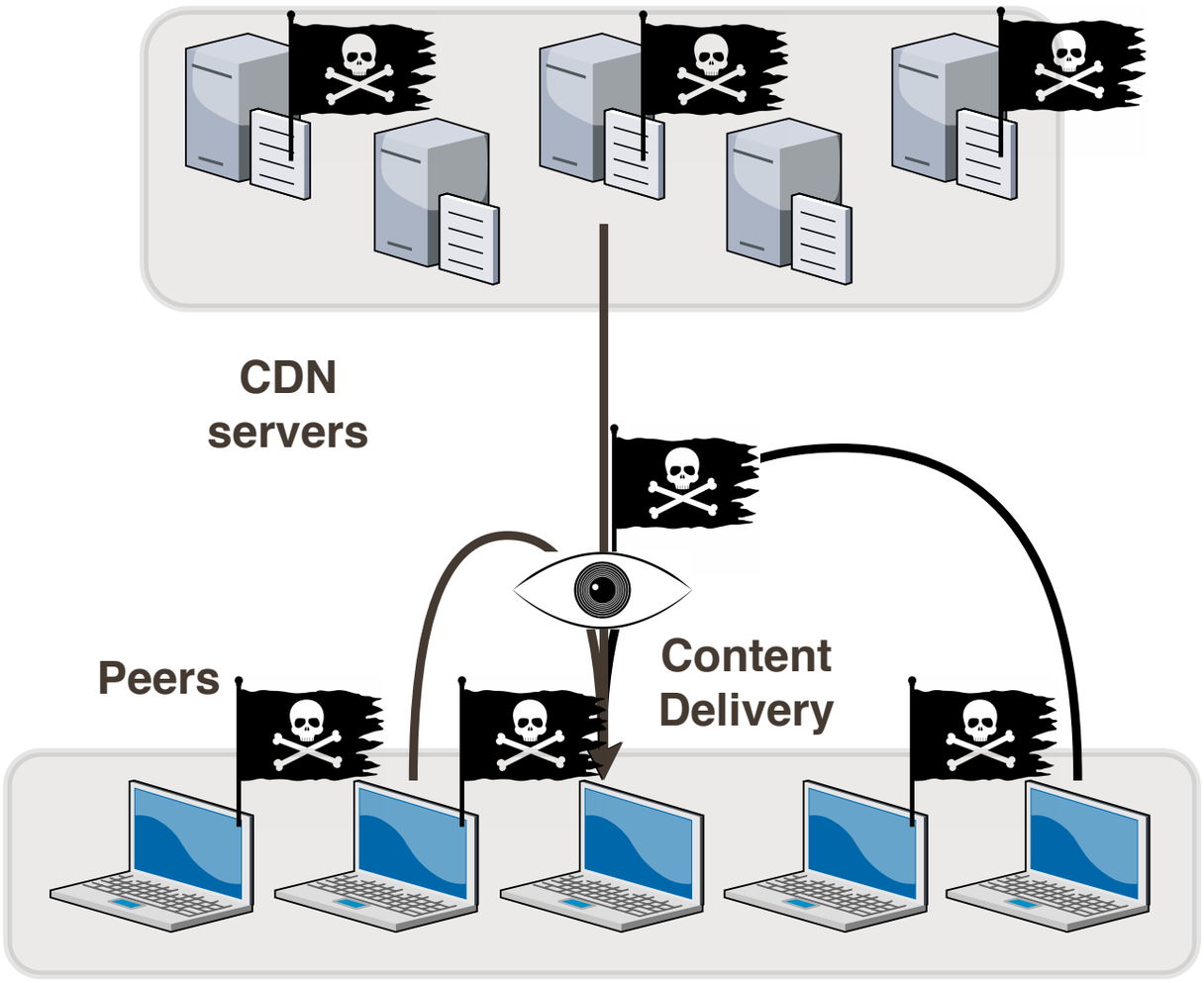
- Allow users to access video streams while enforcing delta-unlinkability
  - The probability of a user ***u*** being interested in a video ***v*** is at most equal to delta
- How?
  - Using TEEs to prevent information leakage (metadata server, the tracker, on the client side)
  - Protecting network traffic
  - Generating fake requests

Using Intel SGX

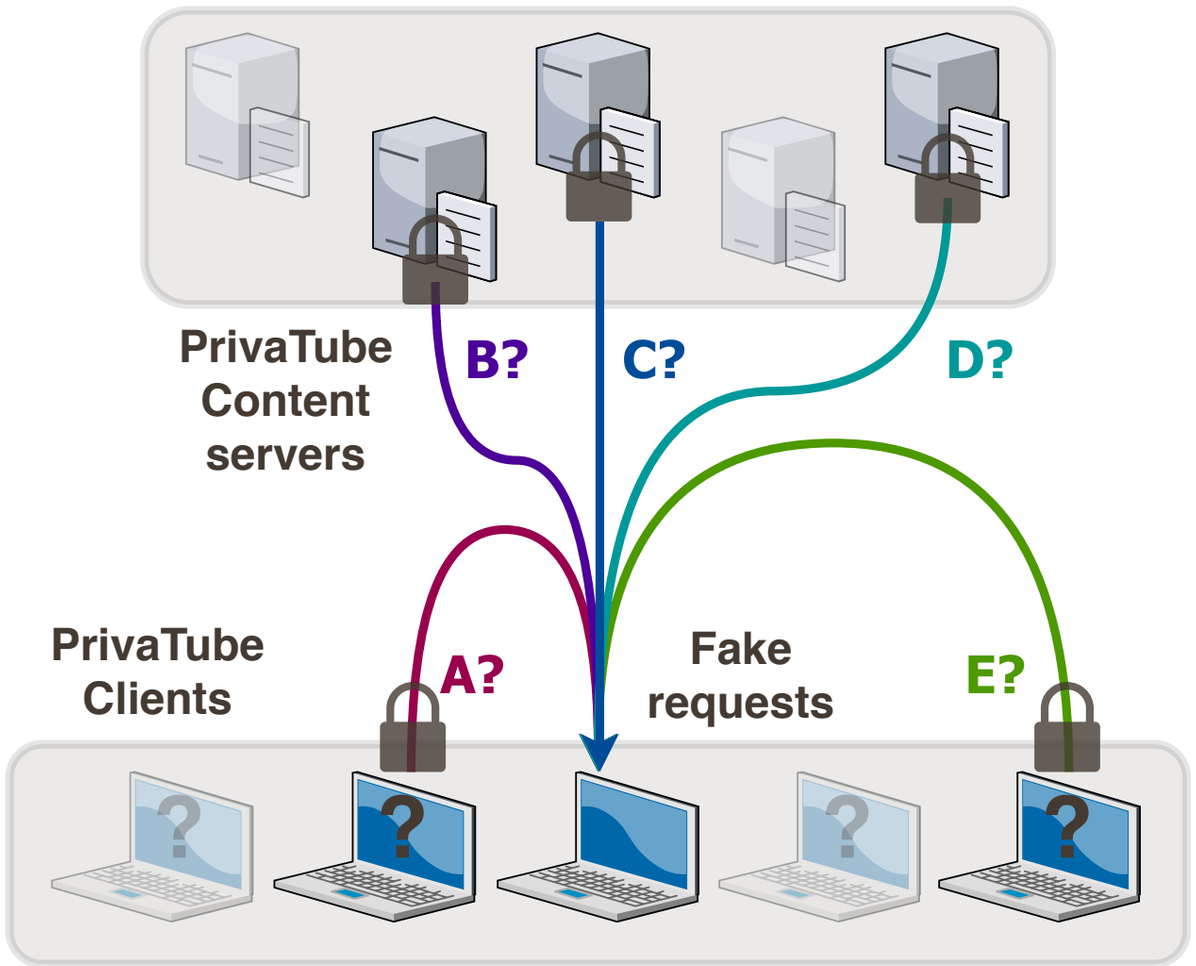




Protecting Against Insider Attacks



Using Fake Requests



Sum up

- Enforcing privacy in online services is important
  - 2 examples **Video Streaming** and **Web Search**
  - Many other examples: Recommender Systems, Location-based services, …
  - P2P and secure hardware can help
  - More info: https://liris.cnrs.fr/drim