DP5: PIR for Privacy-preserving Presence

Ian Goldberg George Danezis

joint work with Nikita Borisov

Cryptography, Security, and Privacy Research Lab
University of Waterloo

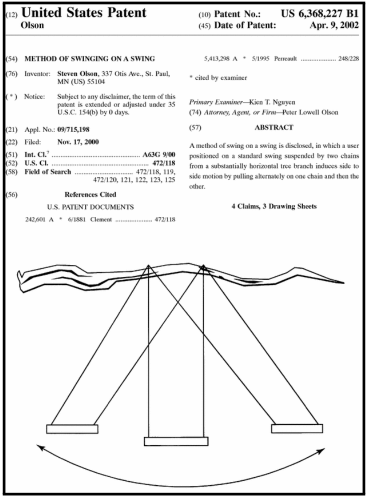
University College London

31C3, 29 December 2014

Private information retrieval

Private information retrieval

Private information retrieval



(12) **United States Patent**
Olson

(10) **Patent No.: US 6,368,227 B1**
(45) **Date of Patent: Apr. 9, 2002**

(54) **METHOD OF SWINGING ON A SWING**

(76) Inventor: **Steven Olson**, 337 Otis Ave., St. Paul, MN (US) 55104

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **09/715,198**

(22) Filed: **Nov. 17, 2000**

(51) **Int. Cl.**[7] **A63G 9/00**
(52) **U.S. Cl.** **472/118**
(58) **Field of Search** 472/118, 119, 472/120, 121, 122, 123, 125

(56) **References Cited**

U.S. PATENT DOCUMENTS

242,601 A * 6/1881 Clement 472/118
5,413,298 A * 5/1995 Perreault 248/228

* cited by examiner

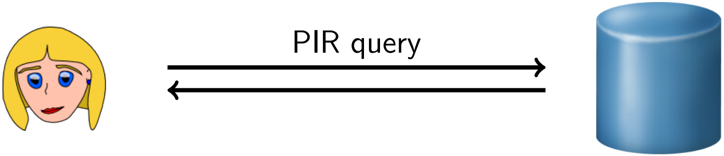
Primary Examiner—Kien T. Nguyen
(74) *Attorney, Agent, or Firm*—Peter Lowell Olson

(57) **ABSTRACT**

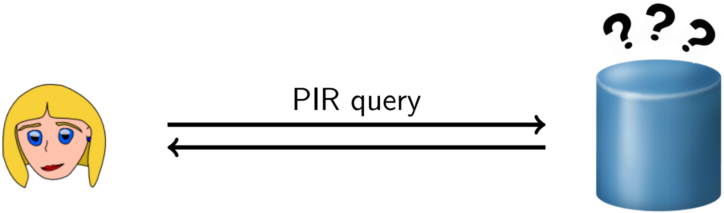
A method of swing on a swing is disclosed, in which a user positioned on a standard swing suspended by two chains from a substantially horizontal tree branch induces side to side motion by pulling alternately on one chain and then the other.

4 Claims, 3 Drawing Sheets

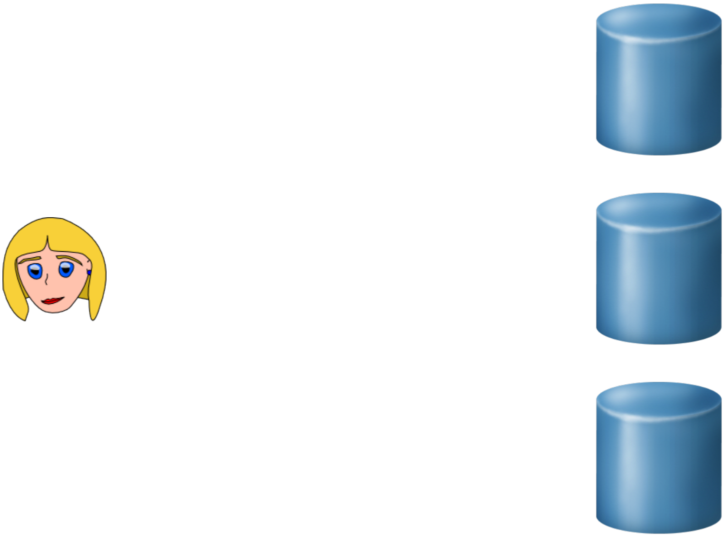
Private information retrieval



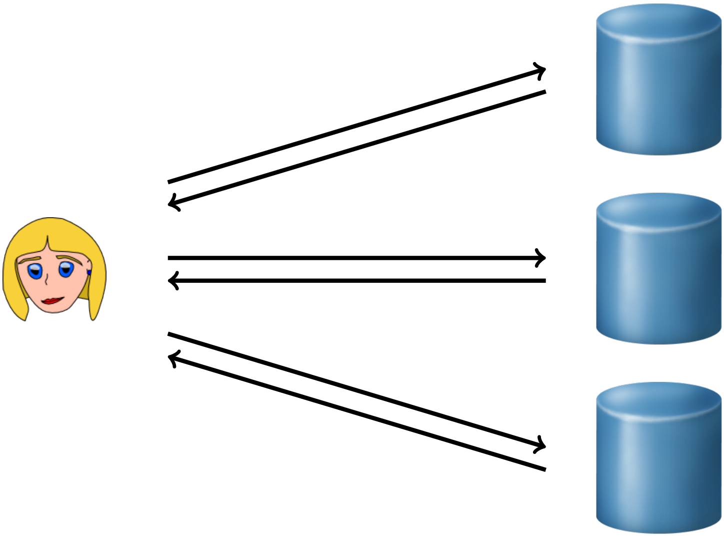
Private information retrieval



IT-PIR



IT-PIR



A simple PIR protocol

$$D = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 0 & 1 & \dots & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & \dots & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & \dots & 1 \\ 1 & 0 & 1 & 0 & 1 & 1 & 0 & \dots & 0 \\ & & & & \vdots & & & \ddots & \vdots \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

A simple PIR protocol

$$D = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 0 & 1 & \dots & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & \dots & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & \dots & 1 \\ 1 & 0 & 1 & 0 & 1 & 1 & 0 & \dots & 0 \\ & & & & \vdots & & & \ddots & \vdots \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

- If $\mathbf{e}_i = [0\ 0\ 1\ 0\ \dots\ 0]$, then $\mathbf{e}_i \cdot D = \text{Block } i$
- $\mathbf{v}_1 \cdot D + \mathbf{v}_2 \cdot D + \cdots + \mathbf{v}_\ell \cdot D = (\mathbf{v}_1 + \mathbf{v}_2 + \cdots + \mathbf{v}_\ell) \cdot D$

A simple PIR protocol

$$D = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 0 & 1 & \dots & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & \dots & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & \dots & 1 \\ 1 & 0 & 1 & 0 & 1 & 1 & 0 & \dots & 0 \\ & & & & \vdots & & & \ddots & \vdots \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

- If $\mathbf{e}_i = [0\ 0\ 1\ 0\ \dots\ 0]$, then $\mathbf{e}_i \cdot D =$ Block i
- $\mathbf{v}_1 \cdot D + \mathbf{v}_2 \cdot D + \cdots + \mathbf{v}_\ell \cdot D = (\mathbf{v}_1 + \mathbf{v}_2 + \cdots + \mathbf{v}_\ell) \cdot D$

A simple PIR protocol

$$D = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 0 & 1 & \dots & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & \dots & 0 \\ \color{red}1 & \color{red}1 & \color{red}0 & \color{red}0 & \color{red}1 & \color{red}1 & \color{red}0 & \color{red}\dots & \color{red}1 \\ 1 & 0 & & 0 & 1 & 1 & 0 & \dots & 0 \\ & & & & \vdots & & & \ddots & \vdots \\ 0 & 1 & 1 & & & 0 & 0 & & 1 \end{bmatrix}$$

Previous work: variable-sized records

- If $\mathbf{e}_i$
- $\mathbf{v}_1 \cdot D + \mathbf{v}_2 \cdot D +$

A simple PIR protocol

$$\begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 0 & 1 & \dots & 0 \\ & & & & & & & \vdots & \vdots \\ 0 & & 1 & 1 & 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

Previous work: lookups by keyword or SQL

- If $\mathbf{e}_i = [0\ 0\ 1\ 0\ \dots\ 0]$, then $\mathbf{e}_i \cdot D =$ Block i
- $\mathbf{v}_1 \cdot D + \mathbf{v}_2 \cdot D + \cdots + \mathbf{v}_\ell \cdot D = (\mathbf{v}_1 + \mathbf{v}_2 + \cdots + \mathbf{v}_\ell) \cdot D$

A simple PIR protocol

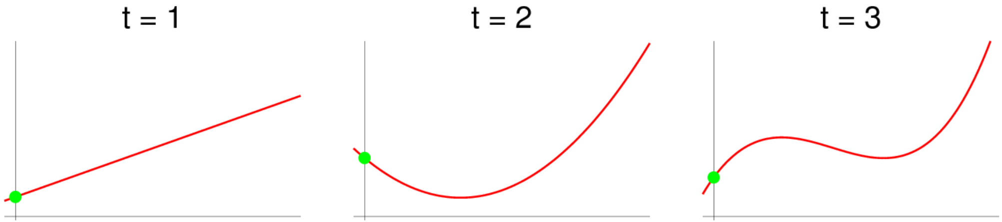
$$D = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 0 & 1 & \dots & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & \dots & 0 \\ \color{red}1 & \color{red}1 & \color{red}0 & \color{red}0 & \color{red}1 & \color{red}1 & \color{red}0 & \color{red}\dots & \color{red}1 \\ 1 & 0 & 1 & 0 & 1 & 1 & 0 & & 0 \\ 0 & 1 & ? & ? & 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

Previous work: robustness

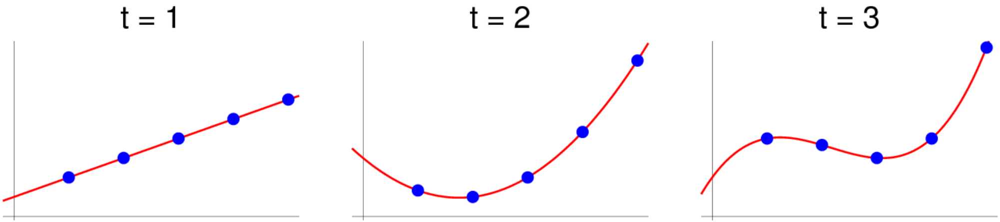
- If $\mathbf{e}_i = [0\ 0\ {\color{red}1}\ 0\ \dots\ 0]$, th $\mathbf{e}_i \cdot D =$ Block i
- $\mathbf{v}_1 \cdot D + \mathbf{v}_2 \cdot D + \cdots + \mathbf{v}_\ell \cdot D = (\mathbf{v}_1 + \mathbf{v}_2 + \cdots + \mathbf{v}_\ell) \cdot D$

Shamir secret sharing

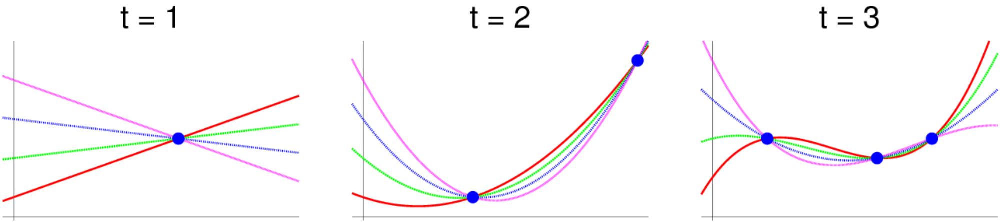
Shamir secret sharing



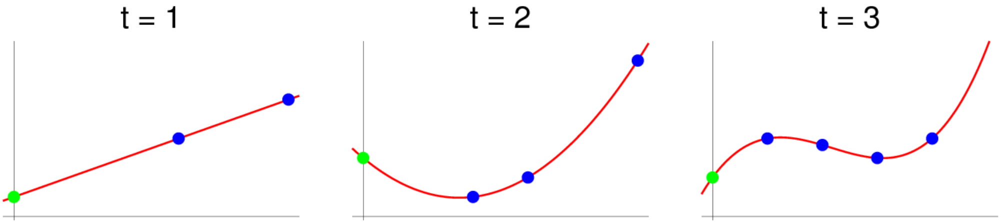
Shamir secret sharing



Shamir secret sharing

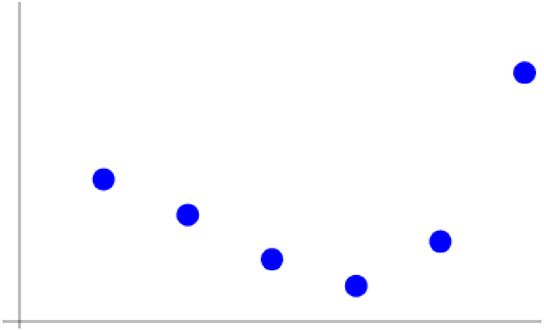


Shamir secret sharing

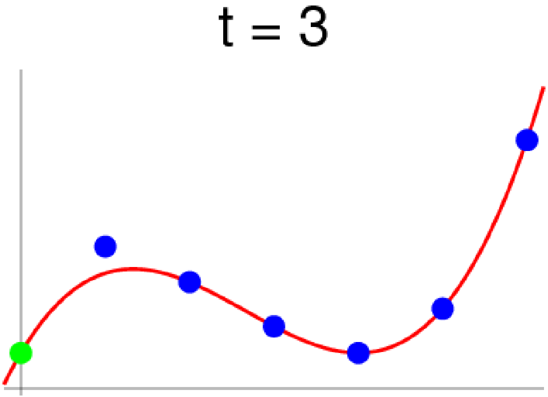


Error correction

$t = 3$



Error correction



Error correction

Error correction

How to try it out

Percy++ open-source library

git://git-crysp.uwaterloo.ca/percy

http://percy.sourceforge.net/

Social applications

Social applications

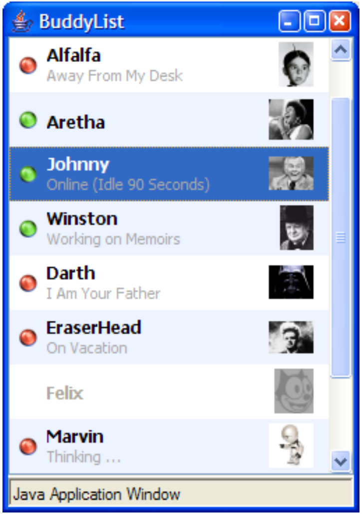
Social applications

Social applications

Social applications

Online presence

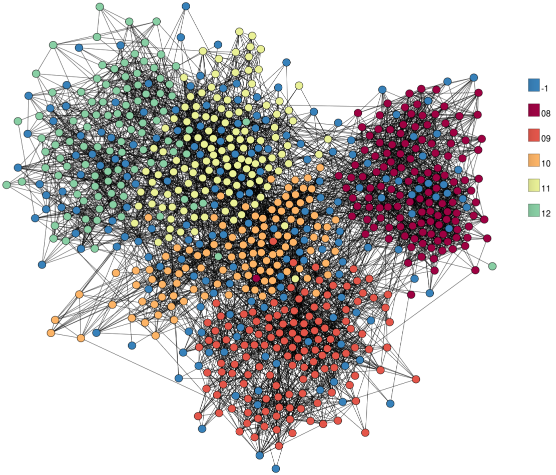
Online presence



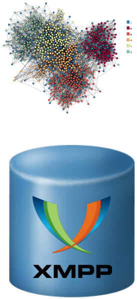
How it typically works

How it typically works

How it typically works

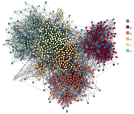


How it typically works

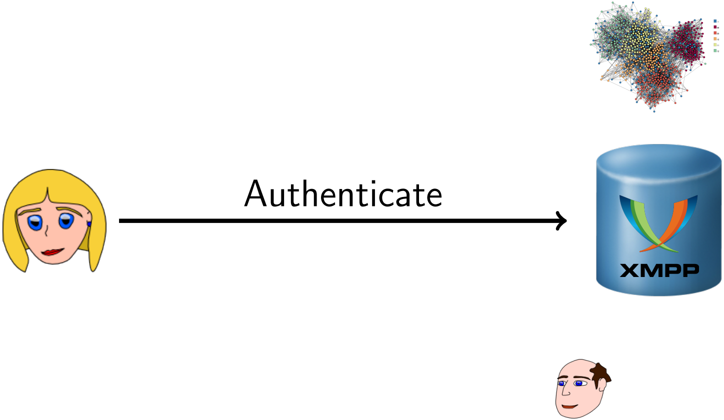


How it typically works

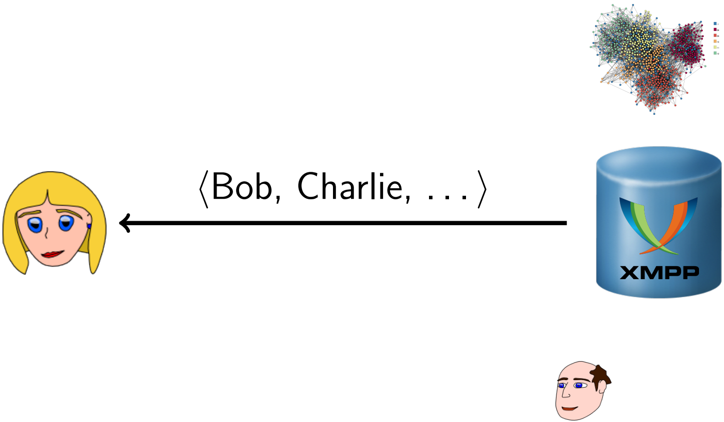
How it typically works



How it typically works

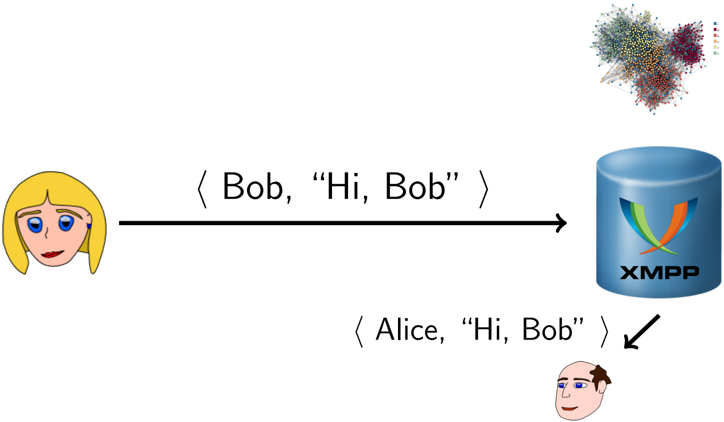


How it typically works

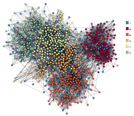


How it typically works

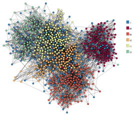
How it typically works



The problem



The problem



The problem

CONTROVERSIES

NSA Collects Online Address Books and Buddy Lists

The agency captures contacts when they're transmitted across global servers, dodging domestic requirements mandating prior authorization for data collection inside the U.S.

By Courtney Subramanian @cmsub | Oct. 14, 2013 | 3 Comments

Share | Like 55 | Tweet 77 | +1 2 | Share 11 | Pin it | Read Later

Patrick Semansky / AP

This June 6, 213 file photo shows the sign outside the National Security Agency (NSA) campus in Fort Meade, Md.

Senior intelligence officers and leaked documents from National Security Agency whistleblower Edward Snowden reveal that the NSA is amassing millions of contacts via online address books and instant-messaging buddy lists.

Email | Print | Share | Comment

Follow @TIMEPolitics

The program, under NSA's Special Source Operations branch, collects more than 250 million contacts in its database per year. A single day's data found that the agency accumulated 444,743 email address books from Yahoo, 105,068 from Hotmail, 82,857 from

“We kill people based on metadata”

General Michael Hayden, former Director of NSA

http://www.youtube.com/watch?v=UdQiz0Vavmc

Want: private presence

Presence features

Threat model

Security goals

Want: private presence

Presence features

Threat model

Security goals

- Friend registration
- Presence registration
- Presence status query
- Friend suspension / revocation

Want: private presence

Presence features

Threat model

Security goals

- Global passive adversary
- Dishonest users
- Secure end hosts
- Threshold of honest infrastructure servers
- Can't break strong crypto

Want: private presence

Presence features

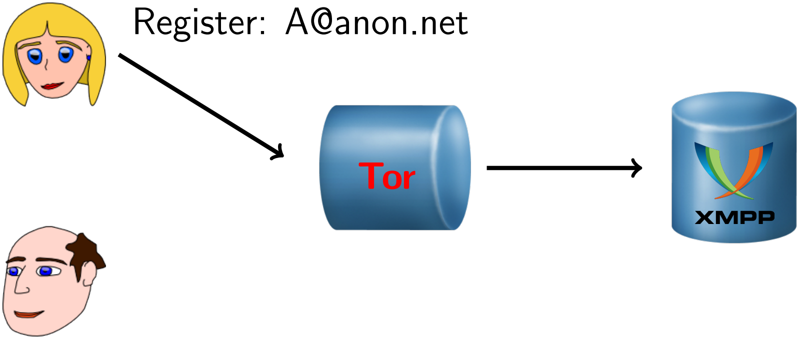
Threat model

Security goals

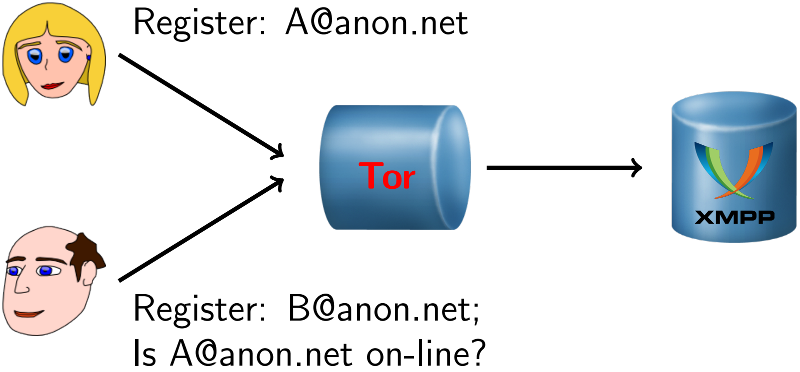
- Privacy of social network
- Privacy, integrity of presence and auxiliary data
- Unlinkability
- Suspension / revocation indistinguishable from offline
- Forward and backward secrecy
- Auditability

Why not 'just use Tor'?

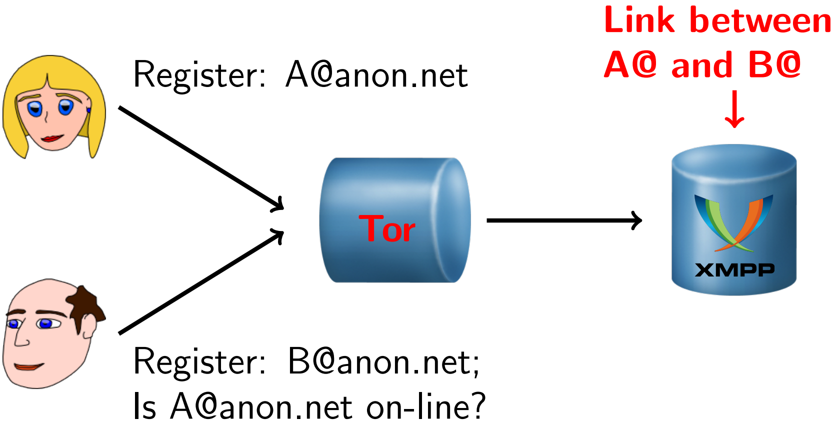
Why not 'just use Tor'?



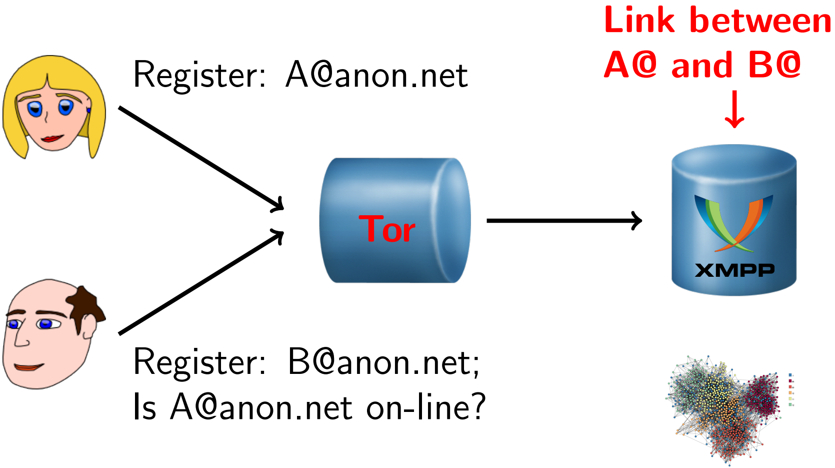
Why not 'just use Tor'?



Why not 'just use Tor'?



Why not ‘just use Tor’?



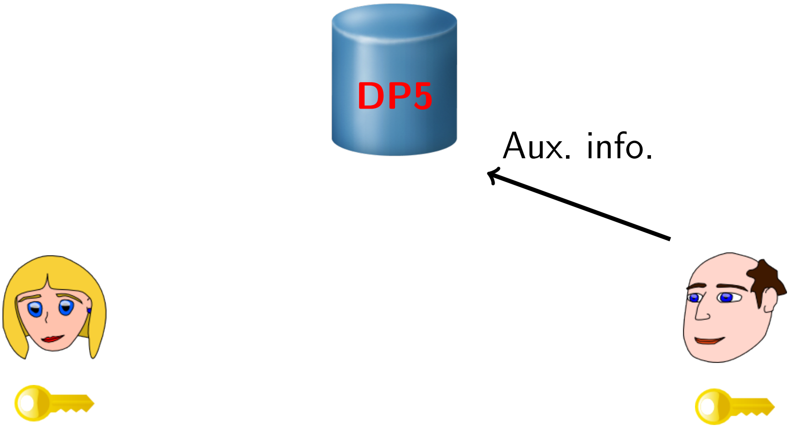
Why not 'just use Tor'?

- 'Anonymous social graph' isomorphic to real social graph $\rightarrow$ anonymization attacks.
    - Easy to de-anonymize using side graphs (Remember Netflix!)
- Pile-up the tricks?
    - Do not register B@ – can still link all friends to a pseudonym.
    - Use a separate circuit per since single friend? $\rightarrow$ Millions of circuits.
    - ...
- DP5 aims: do not require an anonymous channel; do not leak any social graph!

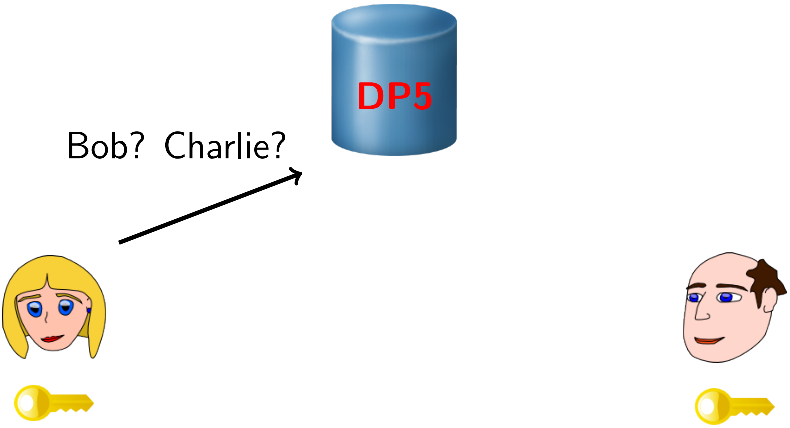
Introducing DP5 (High level idea)

Introducing DP5 (High level idea)

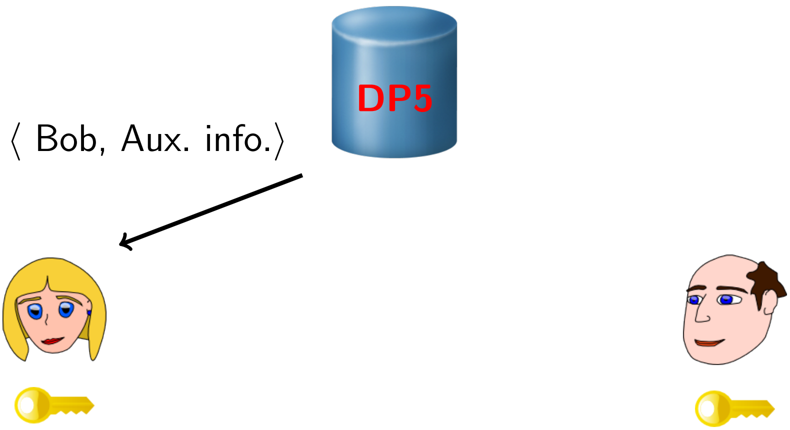
Introducing DP5 (High level idea)



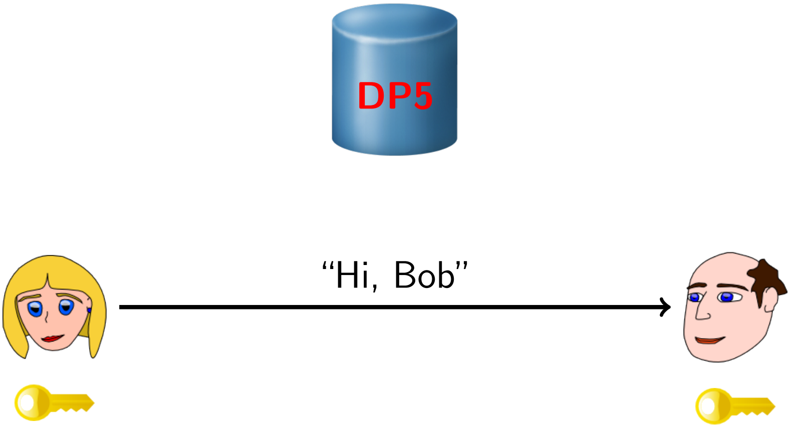
Introducing DP5 (High level idea)



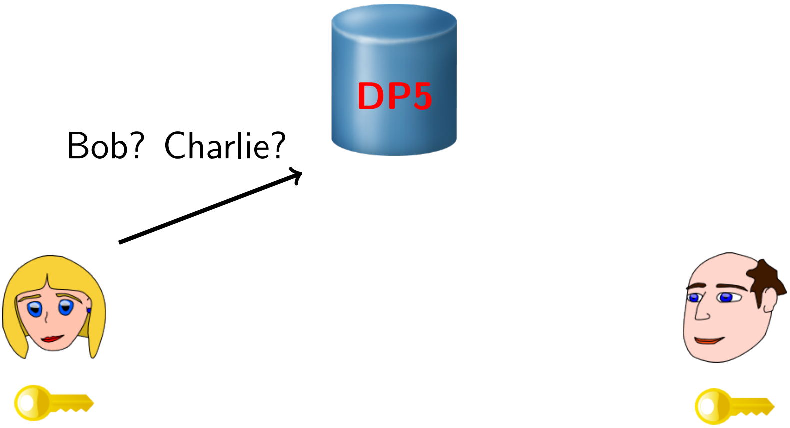
Introducing DP5 (High level idea)



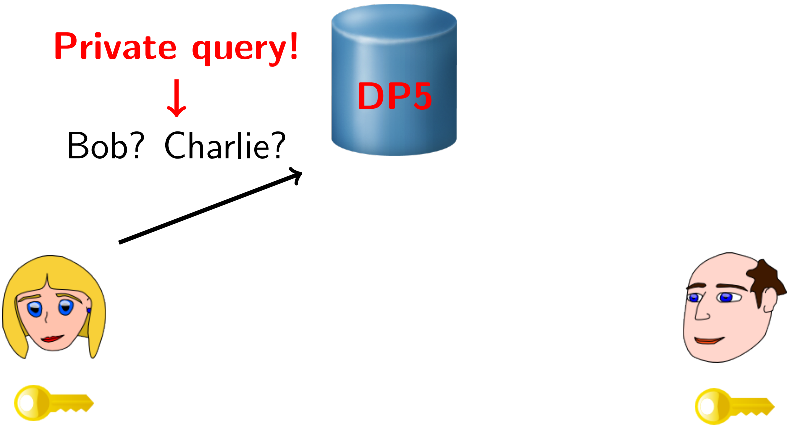
Introducing DP5 (High level idea)

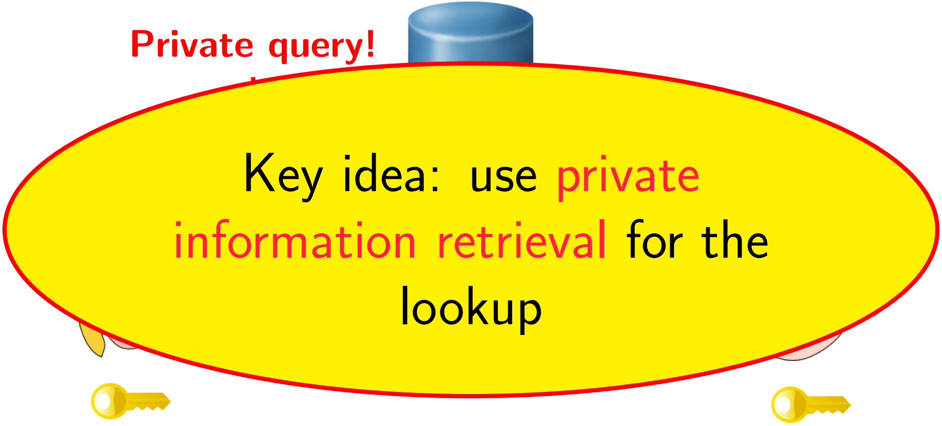


Introducing DP5 (High level idea)



Introducing DP5 (High level idea)

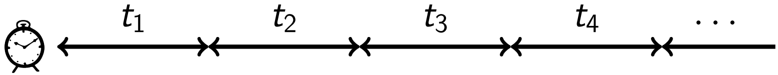




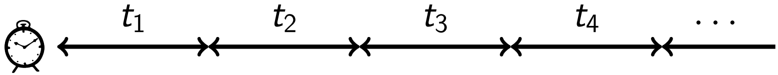
Key idea: use private information retrieval for the lookup

DP5: Strawman version

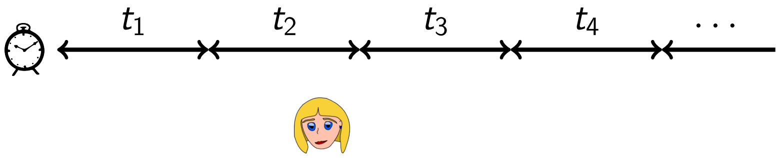
DP5: Strawman version



DP5: Strawman version

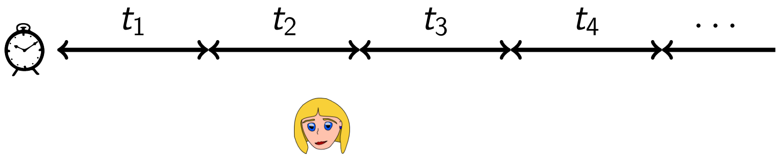


DP5: Strawman version



$$\mathrm{PRF}_{K_{ab}}(t_i)$$

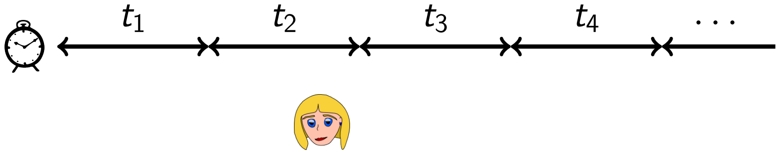
DP5: Strawman version



$\mathrm{PRF}_{K_{ab}}(t_i)$ → K

→ ID

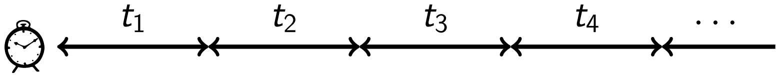
DP5: Strawman version

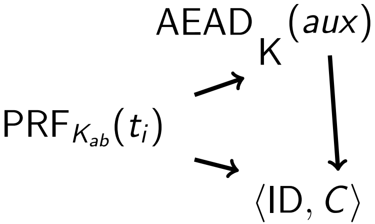


$$\mathrm{AEAD}_{K}(aux)$$

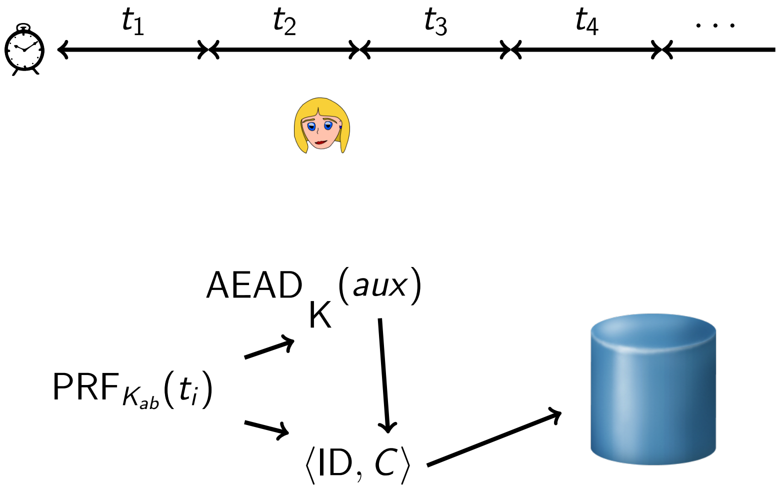
$$\mathrm{PRF}_{K_{ab}}(t_i) \nearrow K \quad \searrow \mathrm{ID}$$

DP5: Strawman version

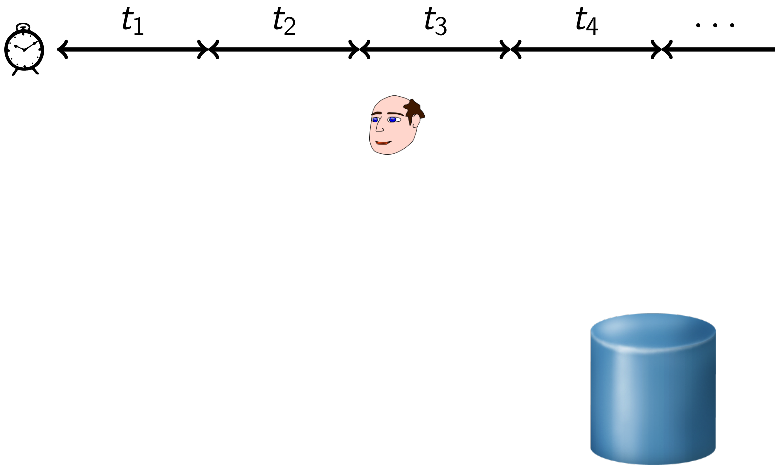




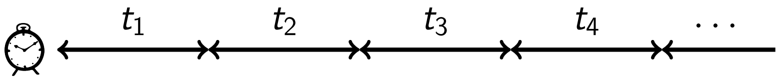
DP5: Strawman version



DP5: Strawman version

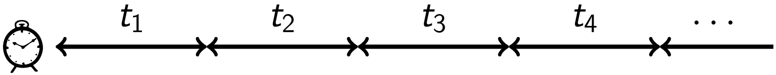


DP5: Strawman version



$\mathrm{PRF}_{K_{ab}}(t_{i-1})$

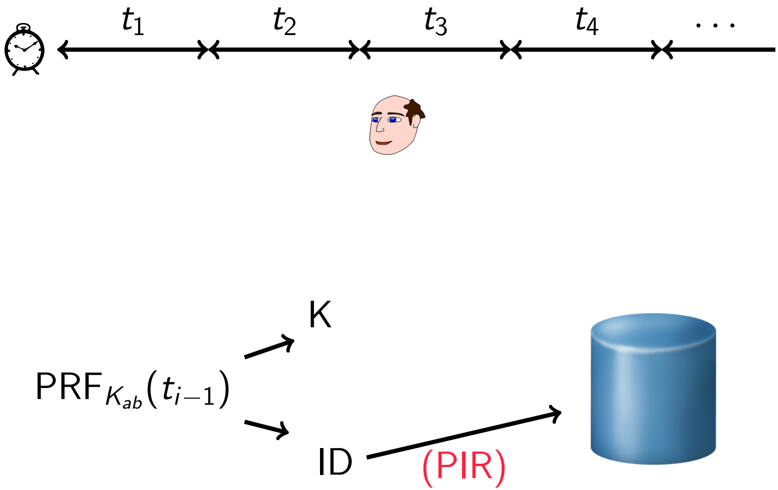
DP5: Strawman version



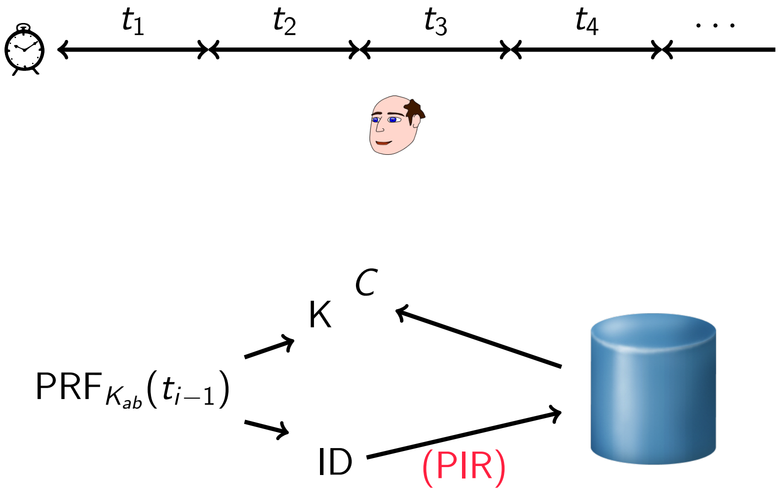
$\mathrm{PRF}_{K_{ab}}(t_{i-1})$ → K

$\mathrm{PRF}_{K_{ab}}(t_{i-1})$ → ID

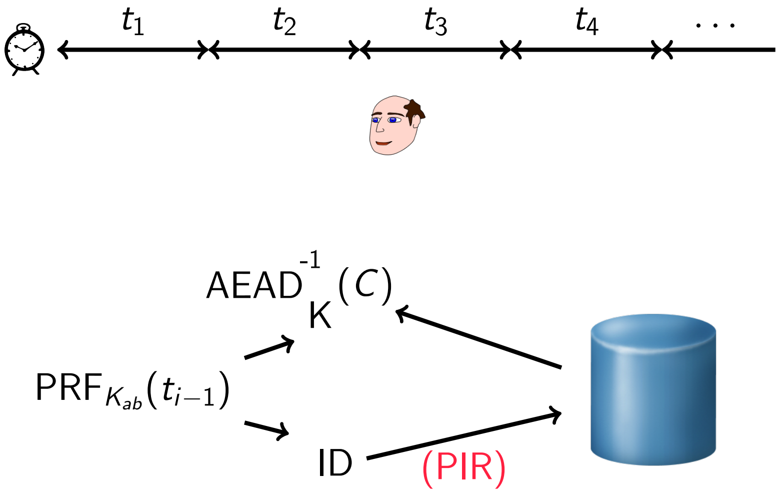
DP5: Strawman version



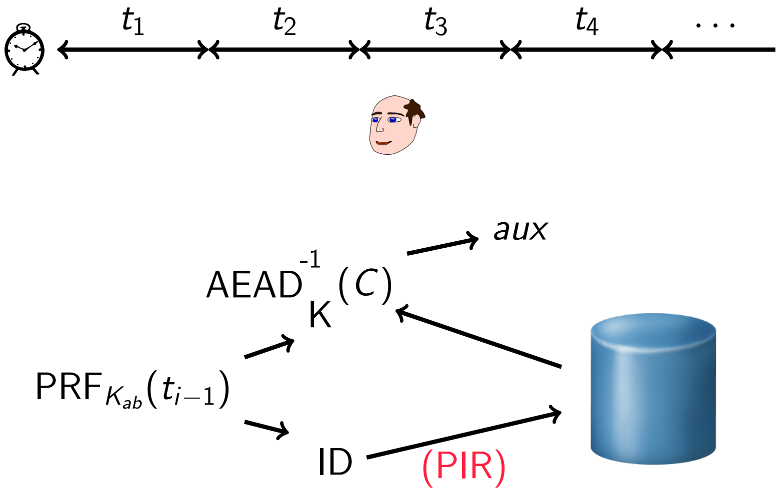
DP5: Strawman version



DP5: Strawman version



DP5: Strawman version



The problem of the large database

The problem of the large database

David Wheeler

The problem of the large database

David Wheele

Any problem in computer science can be solved with another layer of indirection.

The problem of the large database

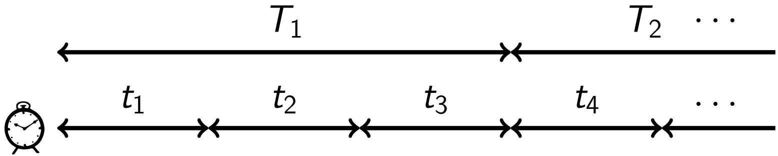
David Wheeler

Any problem in computer science can be solved with another layer of indirection.

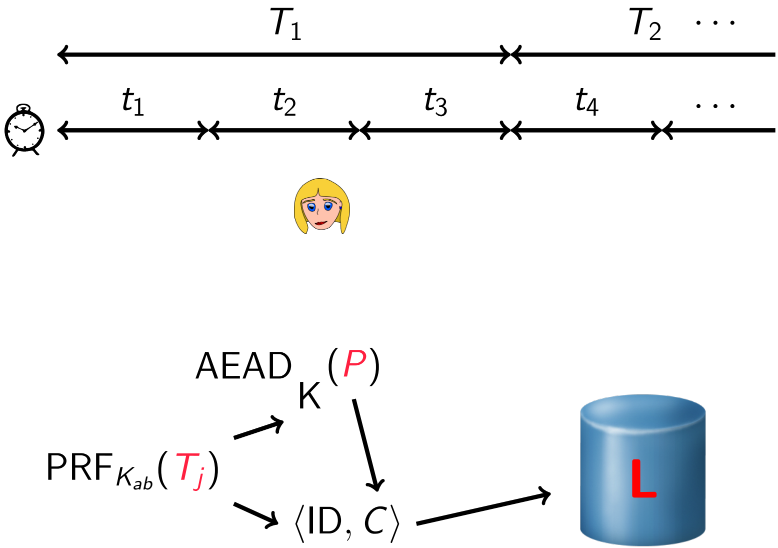
But that will usually create another problem.

Two timescales, two databases

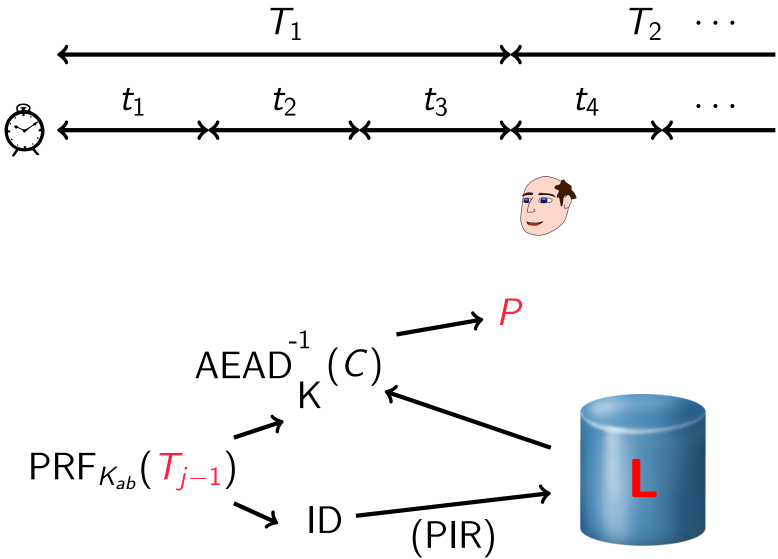
Two timescales, two databases



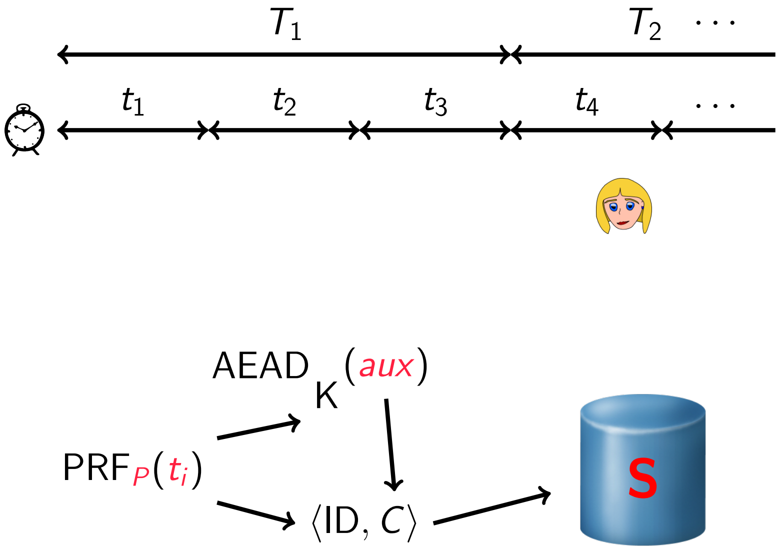
Two timescales, two databases



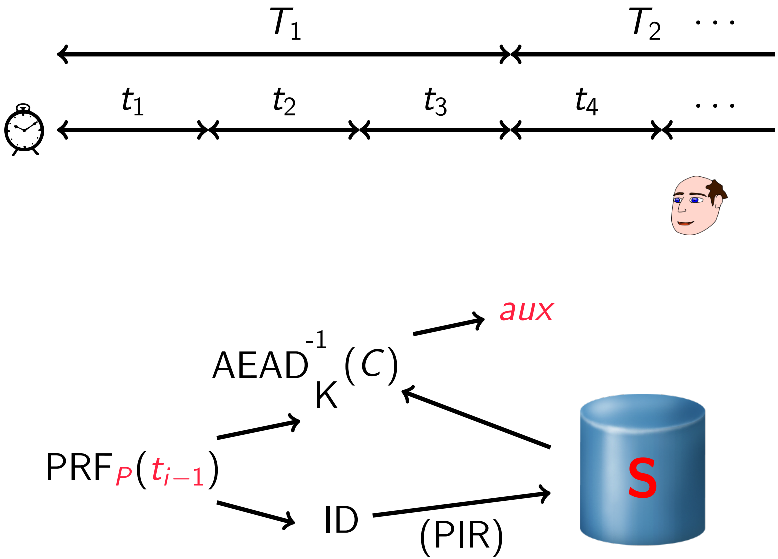
Two timescales, two databases



Two timescales, two databases



Two timescales, two databases



Implementation

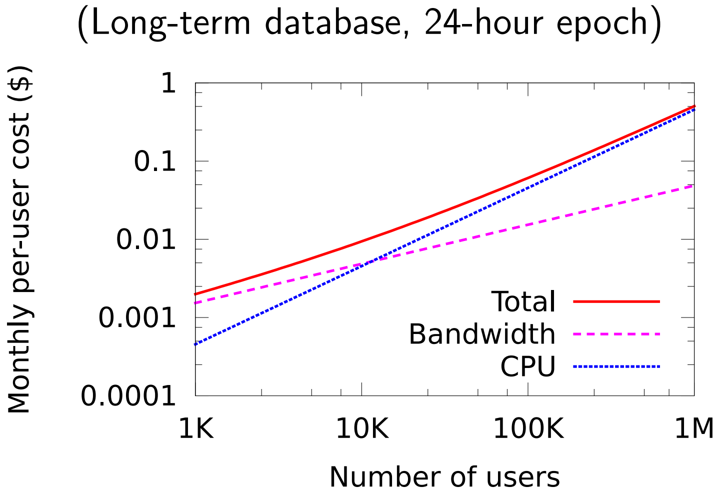
PIR: Percy++ PIR library (C++)

DP5 core: C++, Python Bindings

Networking: Cherrypy framework, Twisted (Python)

Missing: Integration into common chat clients

Cost of running a DP5 PIR server



Takeaways

- Metadata in social communication is being targeted
- Private information retrieval (PIR) allows database lookups without revealing the query to the database servers themselves
- DP5 uses PIR to achieve private presence—people learn when their friends are online (and how to contact them securely) without any server ever learning who is friends with whom

Find out more

- Technical report
  `http://cacr.uwaterloo.ca/techreports/2014/cacr2014-10.pdf`
- Git code repository
  `git://git-crysp.uwaterloo.ca/dp5`