EURECOM
Sophia Antipolis

Security and Privacy for Cloud Computing

Refik Molva

Cloud Computing

- **Outsourcing**
  - Storage
  - Computation

- **High availability**
- **No maintenance**
- **Decreased Costs**
- **Elasticity & Flexibility**

Security & Privacy Challenges

Outsourcing

- Potentially untrusted Service Provider

- Data storage and computations

⇒ New requirements (PoR, verifiability, . . .)

⇒ Crypto schemes dealing with untrusted partners

- PIR
- Secure multi-party computation
- Computing with encrypted functions
- Verifiability: proof of data possession, proof of execution

Security & Privacy Challenges

Large scale

- Data

- Computations

⇒ **Severely asymmetric scenarios**

- Customer (verifier) << Service Provider (prover)
- “Quantum leap”: classical schemes don’t work, need for new approaches
- Example: integrity – customer cannot even keep a hash value per data split

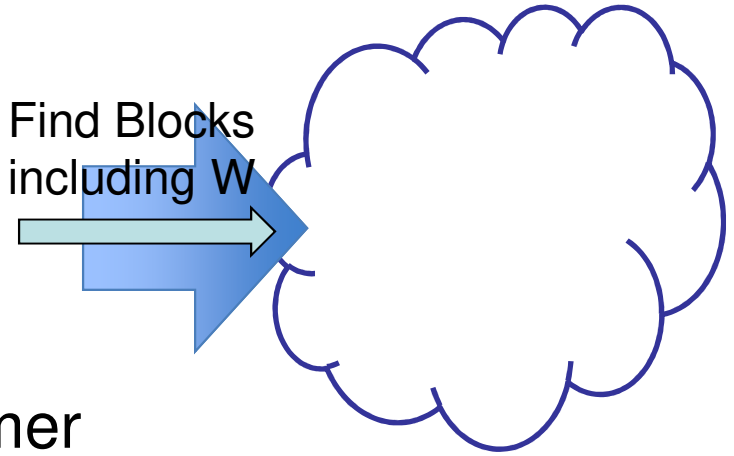
⇒ **Joint Crypto & Cloud schemes**

Security & Privacy Solutions

- **Privacy**
  - Privacy preserving word search
  - Multi-user searchable encryption
- **Integrity**
  - Proof of Retrievability
  - Message-locked PoR
- **Verifiability**
  - Verifiable computation
  - Proof composition

Privacy preserving word search

Outsourced Backup Service



- several years' corporate data
- regularly stored in the Cloud
- Privacy ⇨ Encryption by the customer
- Query: only a small portion needs to be restored
- How to find it without downloading the entire DB?

Requirement for a new solution

- to **search words** in an **encrypted DB**
- with **privacy**

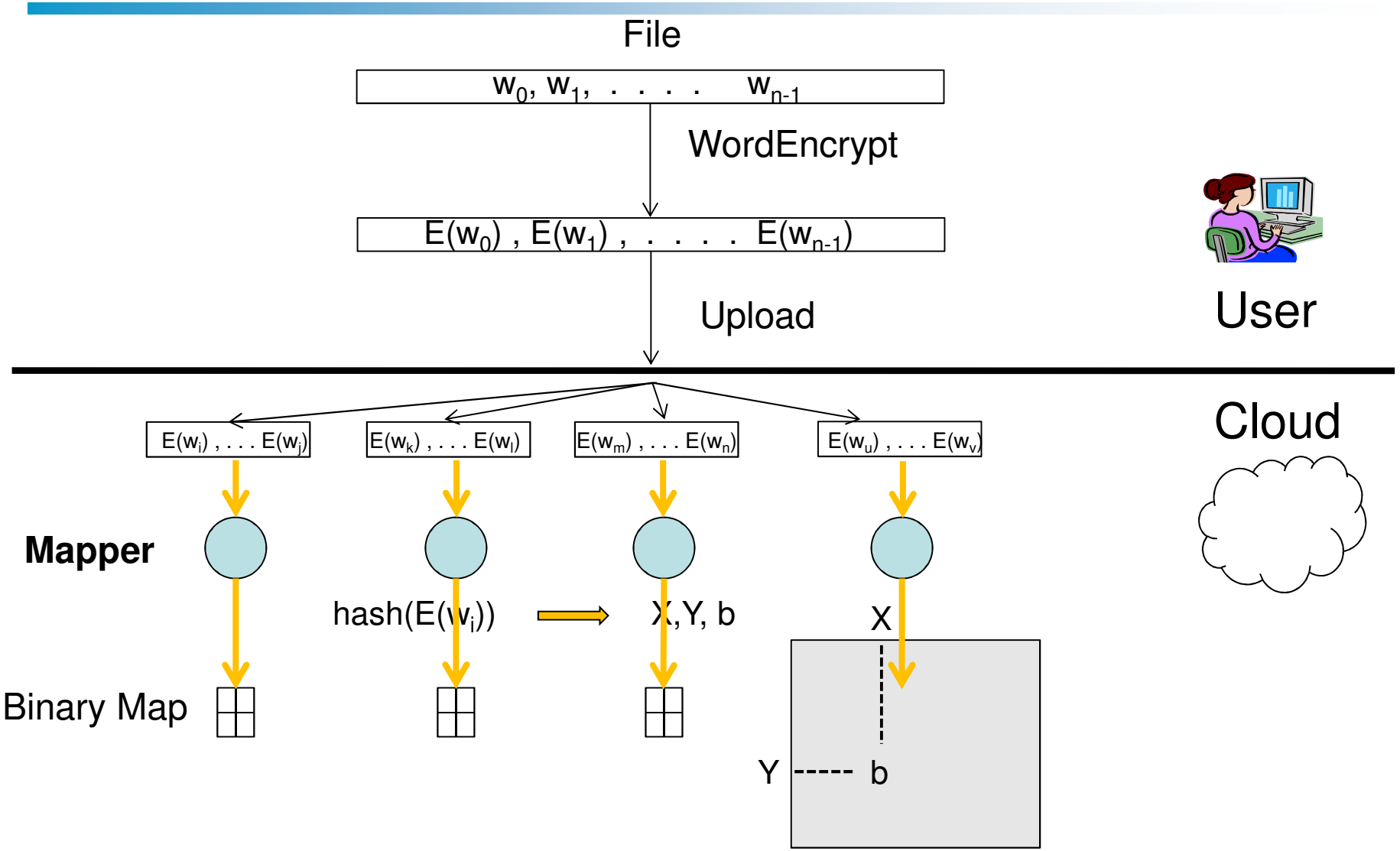
Privacy preserving word search

- **Existing solutions not scalable**
  - Encrypted keyword search algorithms
  - Private information retrieval (PIR)

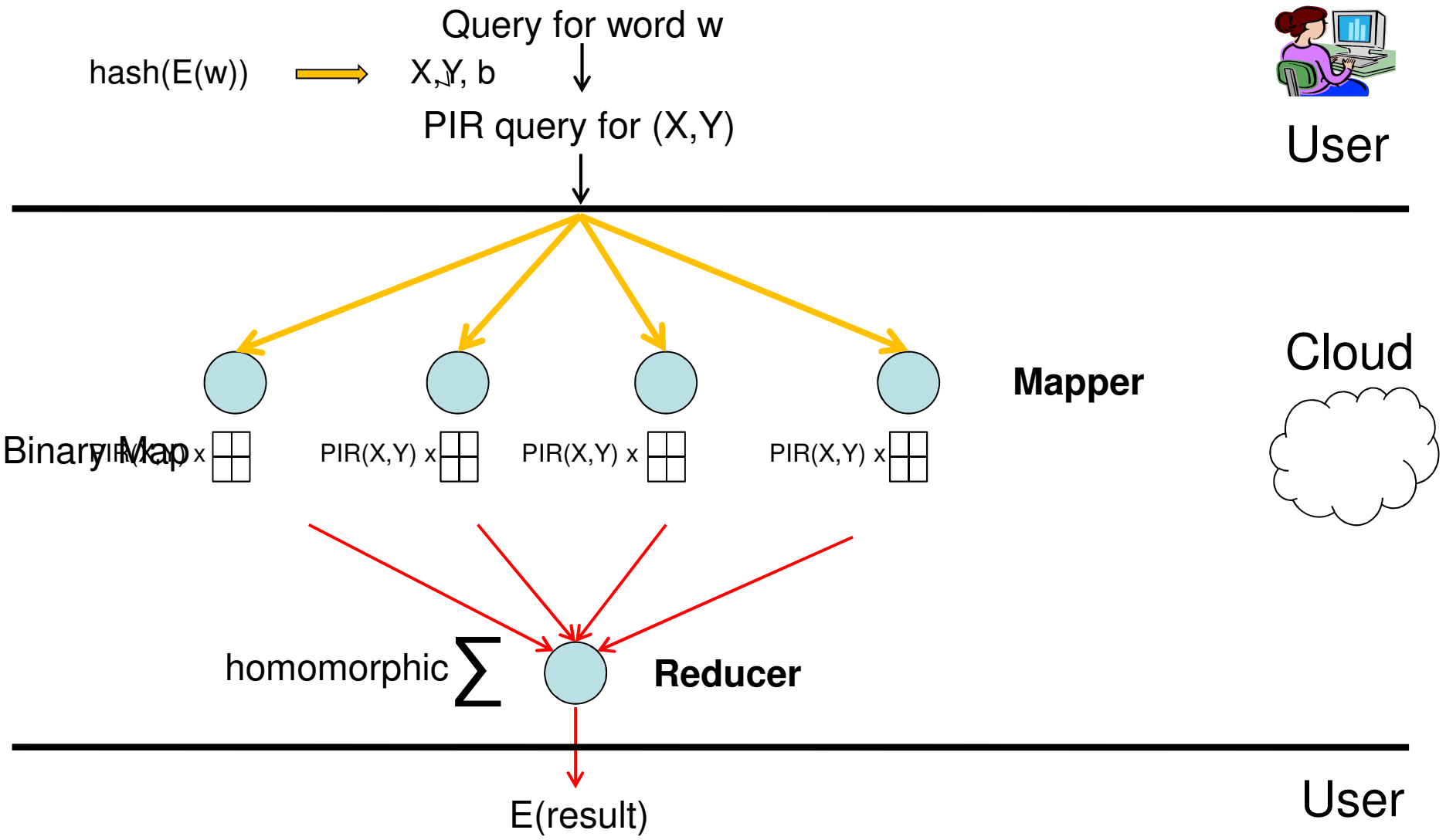
- **PRISM: Privacy preserving search in MapReduce**
  - Data and query privacy
  - Idea: PIR on intermediate data maps
  - Advantage: parallelism with MapReduce

PRISM - Upload

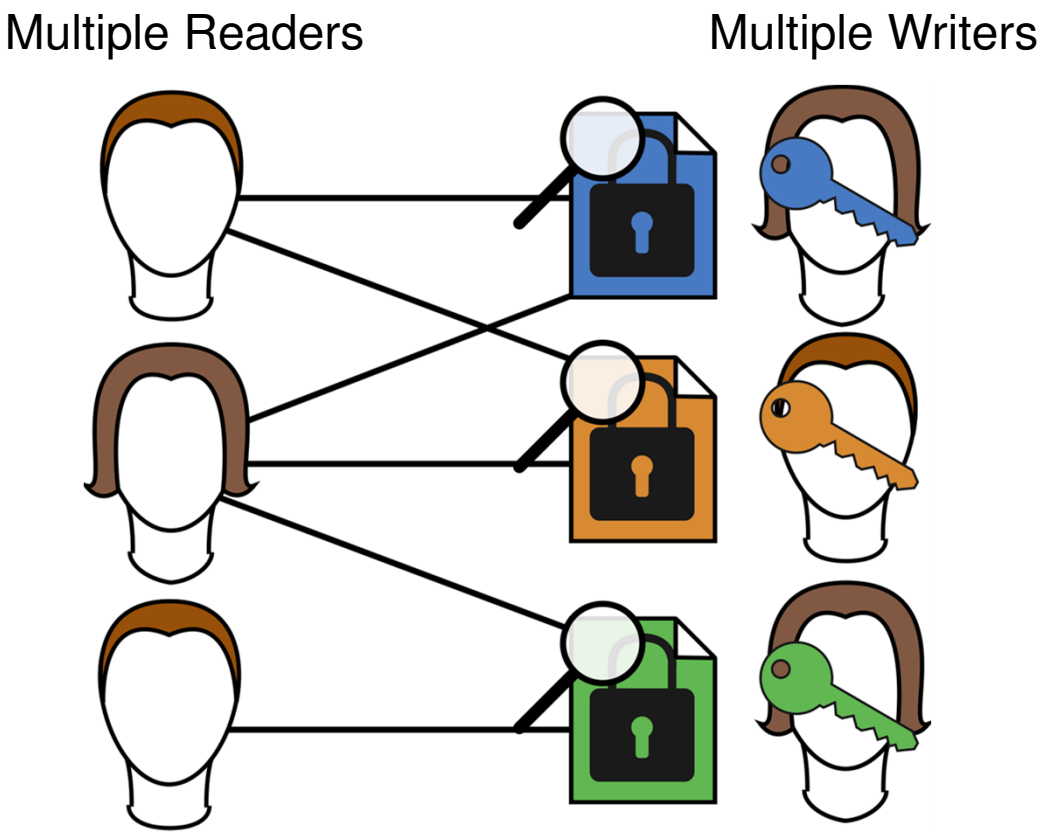
[PETS'12]



PRISM – Word Search

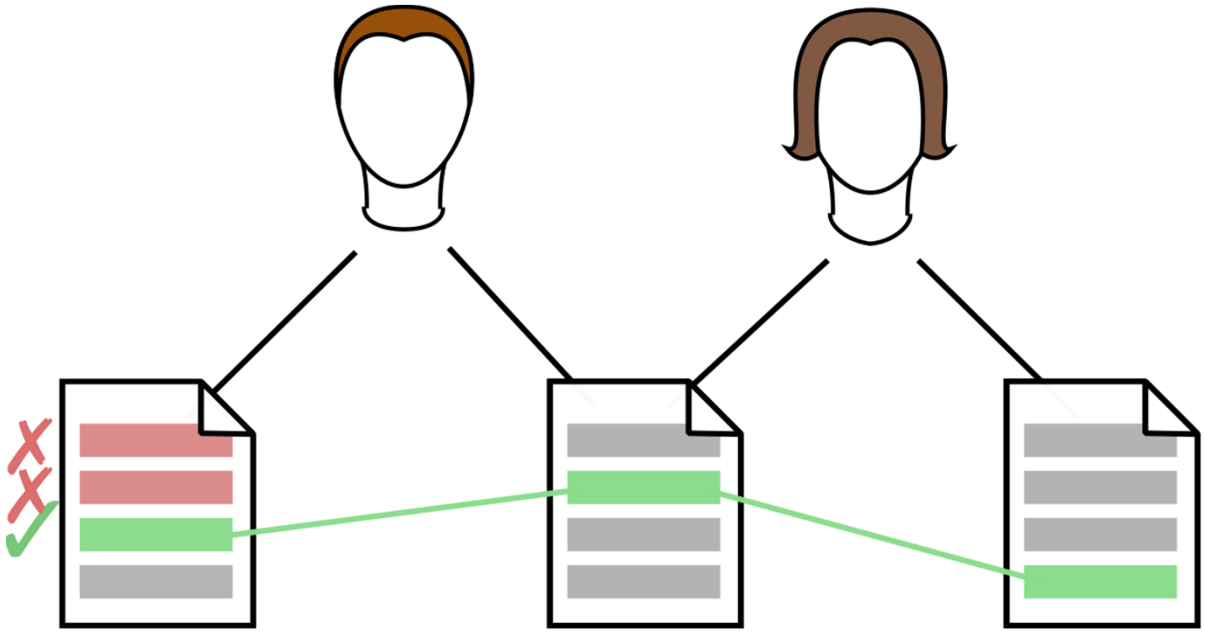


Multi-user Searchable Encryption (MUSE)



SotA - Access pattern leakage *[PETS'17]*

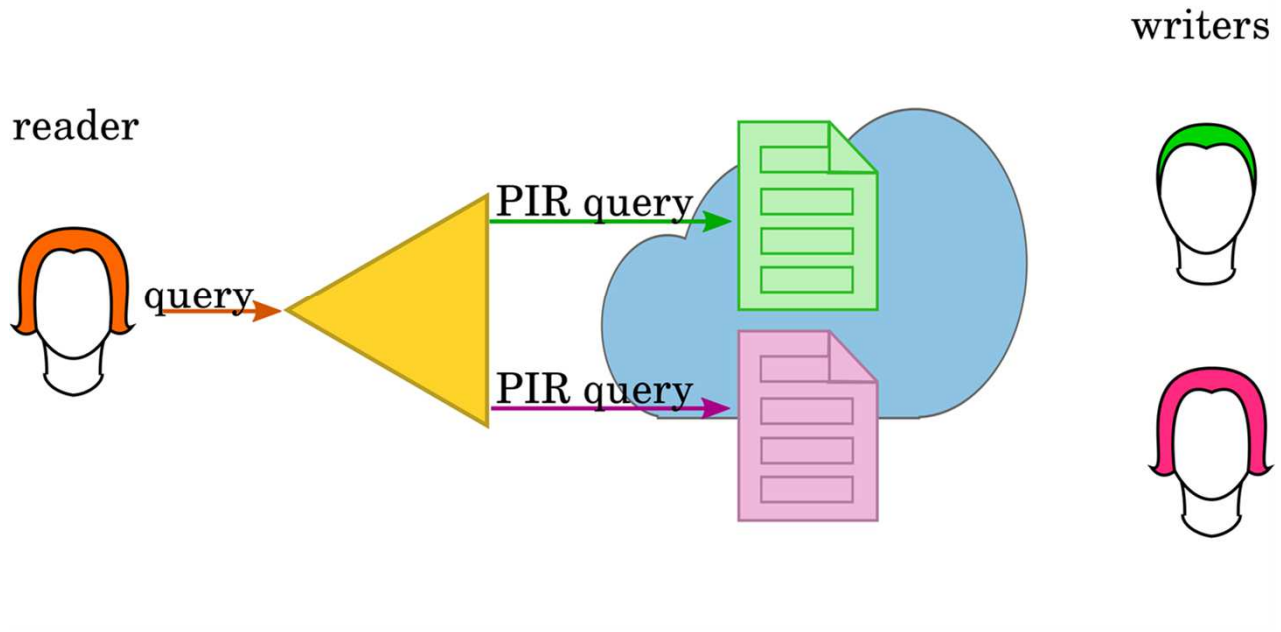
- **Iterative Testing**



- *Each encrypted keyword is tested separately in all documents*
- *Similarities between documents & position of the keyword revealed*

Collusion (CSP, User) ⇒ Privacy Breach

MUSE: Multi-User Searchable Encryption
[ISC'15]



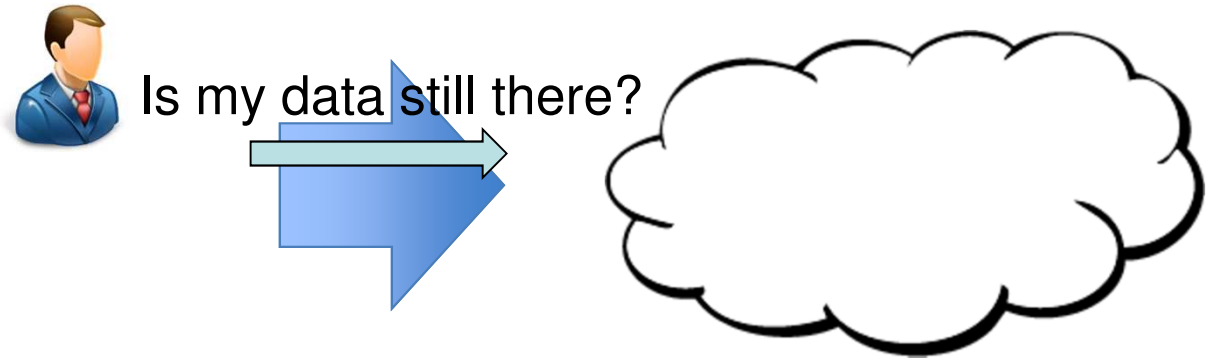
No collusion between Proxy and CSP

Cloud Security Research

- **Privacy**
  - Privacy preserving word search
  - Multi-user searchable encryption
- **Integrity**
  - Proof of Retrievability
  - Message-locked PoR
- **Verifiability**
  - Verifiable computation
  - Proof composition

Proof of Retrievability

- **Motivating scenario: outsourced storage**



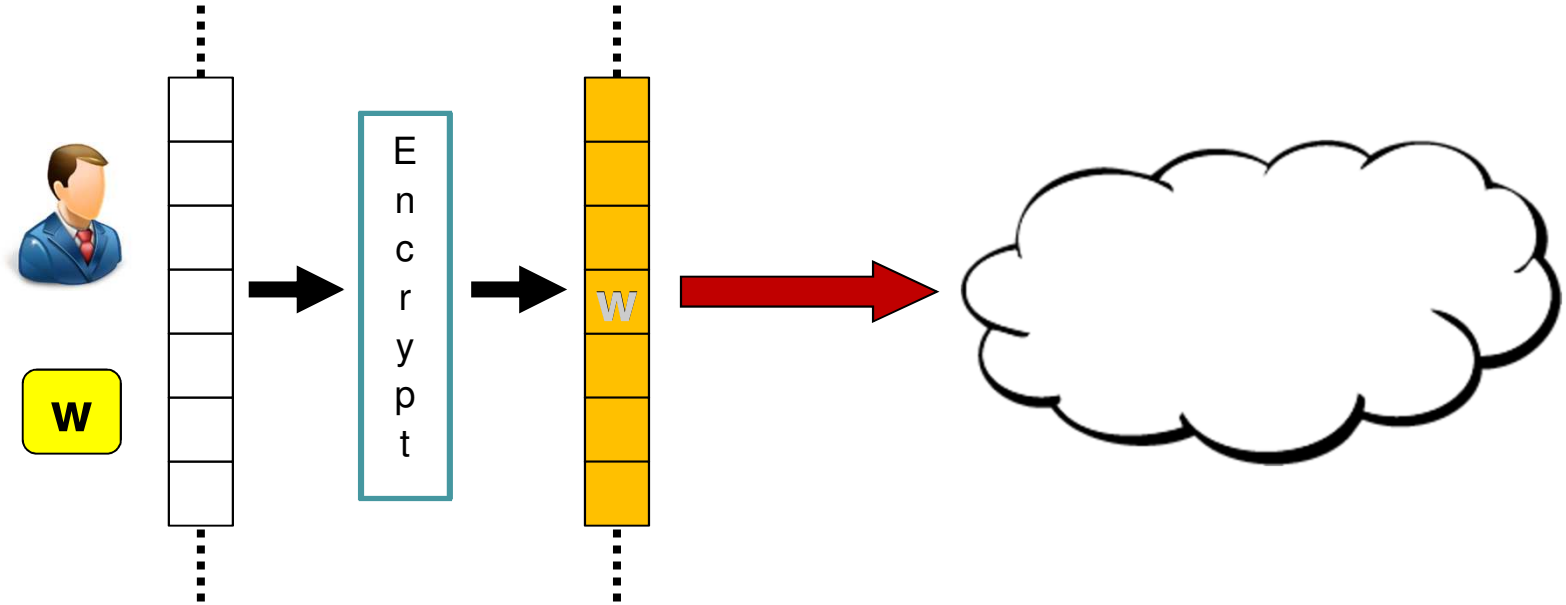
- **Requirements**
  - Integrity check by Client
  - No data stored at Client
  - No bulk data transfer

⇨ **Proof of Retrievability (POR)**

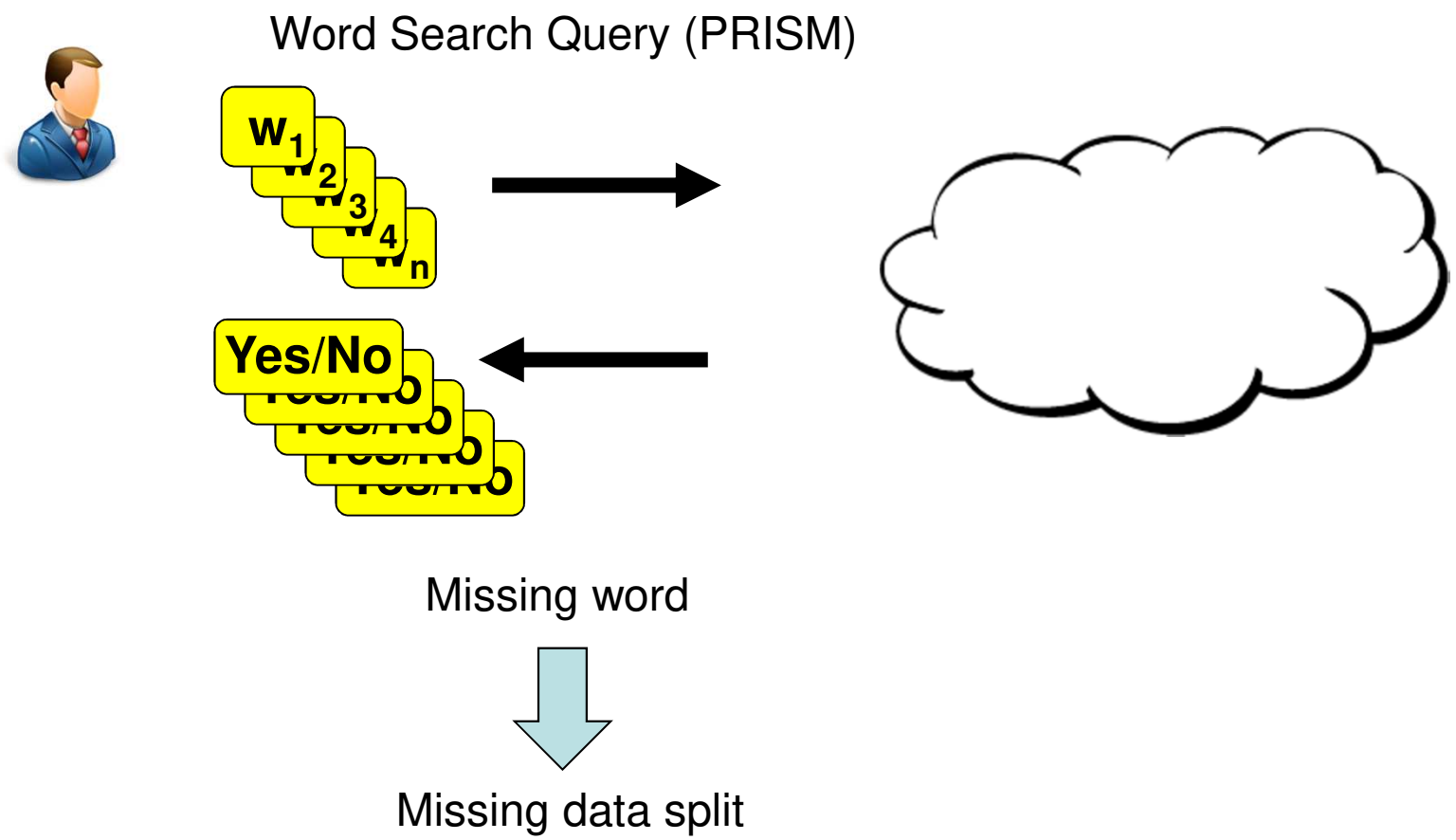
Proof of Retrievability – Related Work

- **Related work** *[Deswarte et. al, Filho et. al, ...]*
  - Deterministic
    - ☞ Verification of the entire data ⇒ costly
  - Probabilistic *[Ateniese et. al, Shacham et.al, Juels et al, ...]*
    - ☞ Tags for each block + random verification ⇒ cost of homomorphic ops
    - ☞ randomly located sentinels => limited # of verifications
- **StealthGuard** *[ESORICS'14]*
  - privacy preserving search of watchdogs
  - Unbounded # of queries

Proof of Retrievability - StealthGuard



Proof of Retrievability – StealthGuard

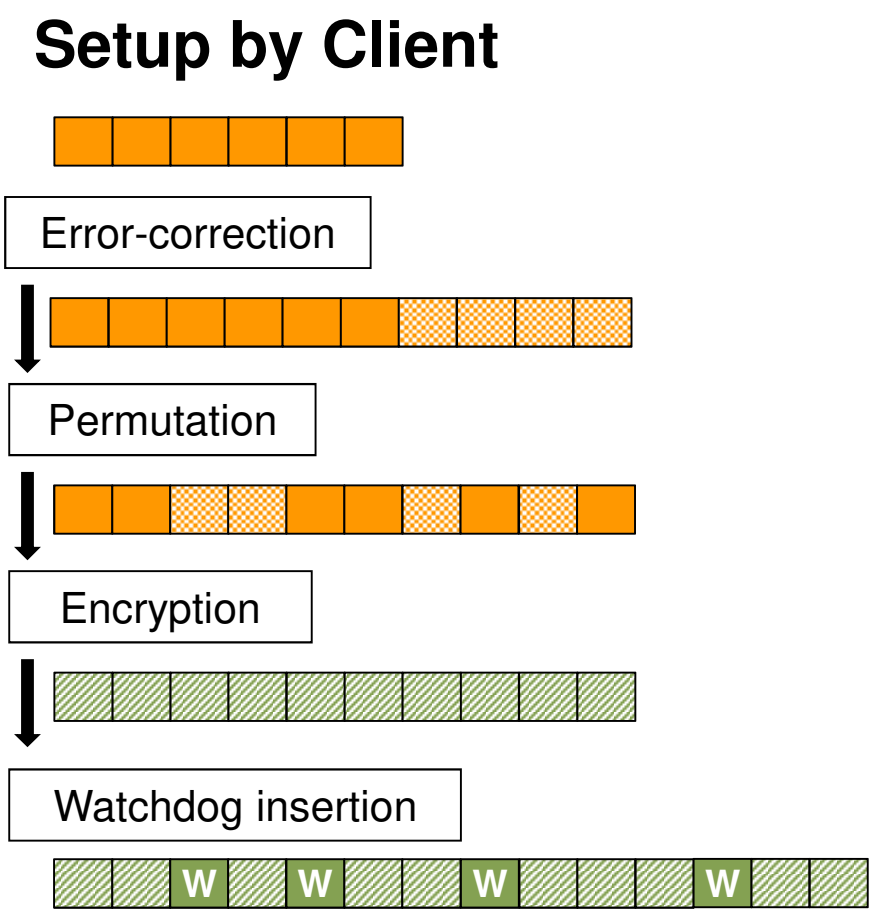


How many watchdogs to check?

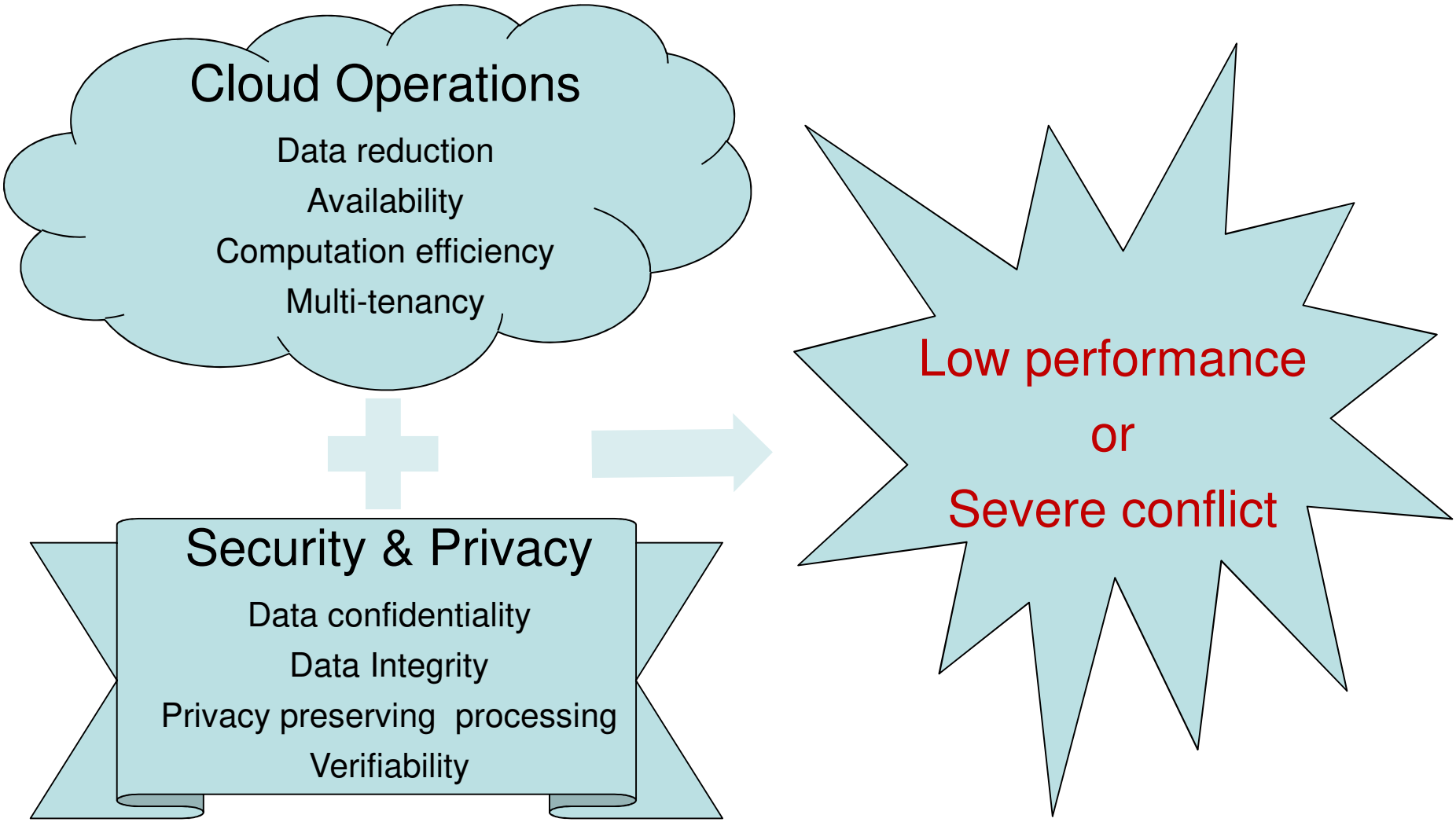
or how to detect lack of retrievability?

Adversary Model:
Bernoulli processes ρ_Adv
$\Rightarrow$ **Error-correction** $\rho_n = f(\tau, \rho_{ECC}, m)$

Retrievability: $\rho_{Adv} \leq \rho_n$
File not retrievable: $\rho_{Adv} > \rho_n$
Detection: $\gamma = f(\tau, \rho_n)$

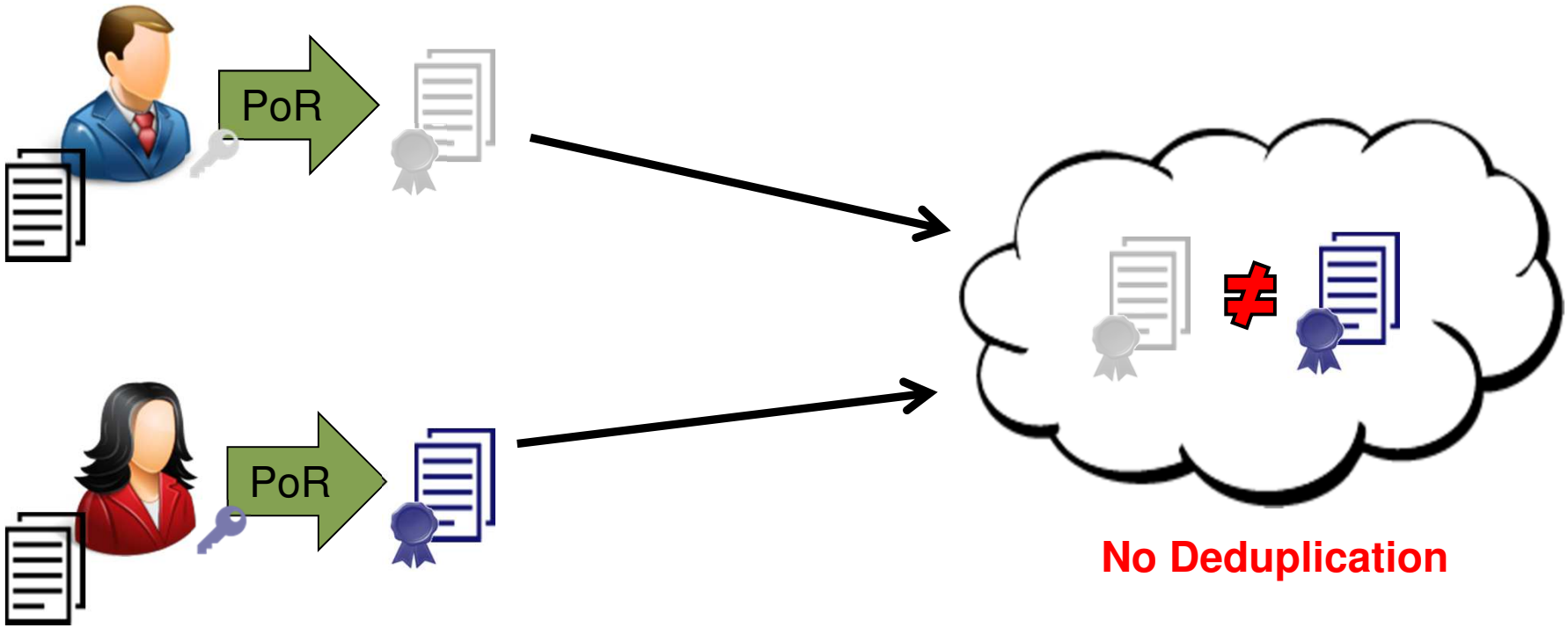


The Integration Problem



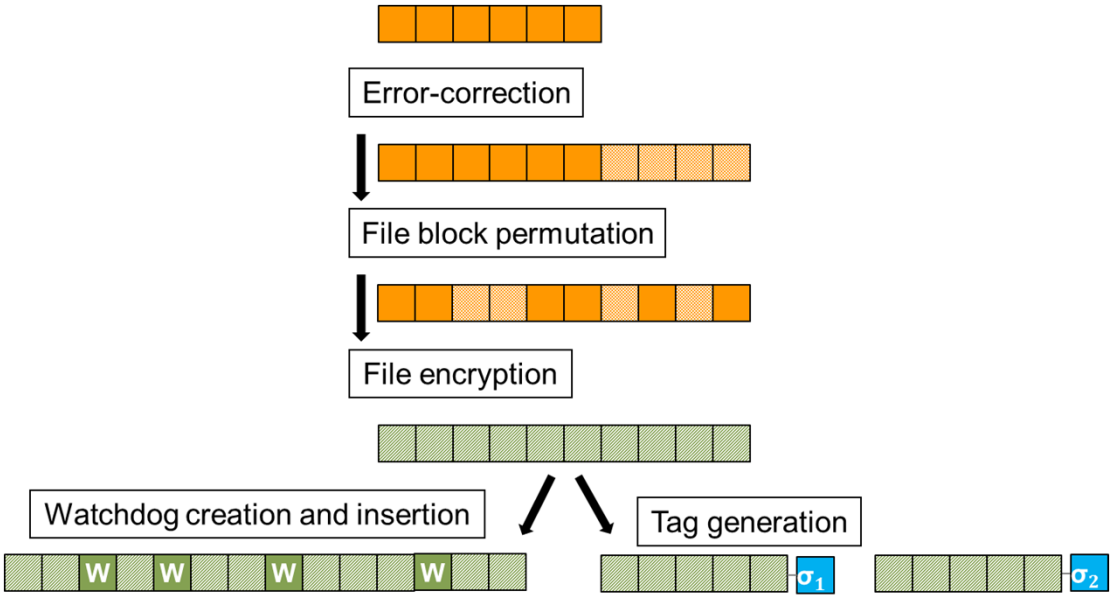
Conflict between PoR & deduplication

- **PoR → User specific encoding**
- **Deduplication → Keep a unique copy in storage**



Message-locked PoR - Idea

- **PoR setup (Tags and Watchdogs)**



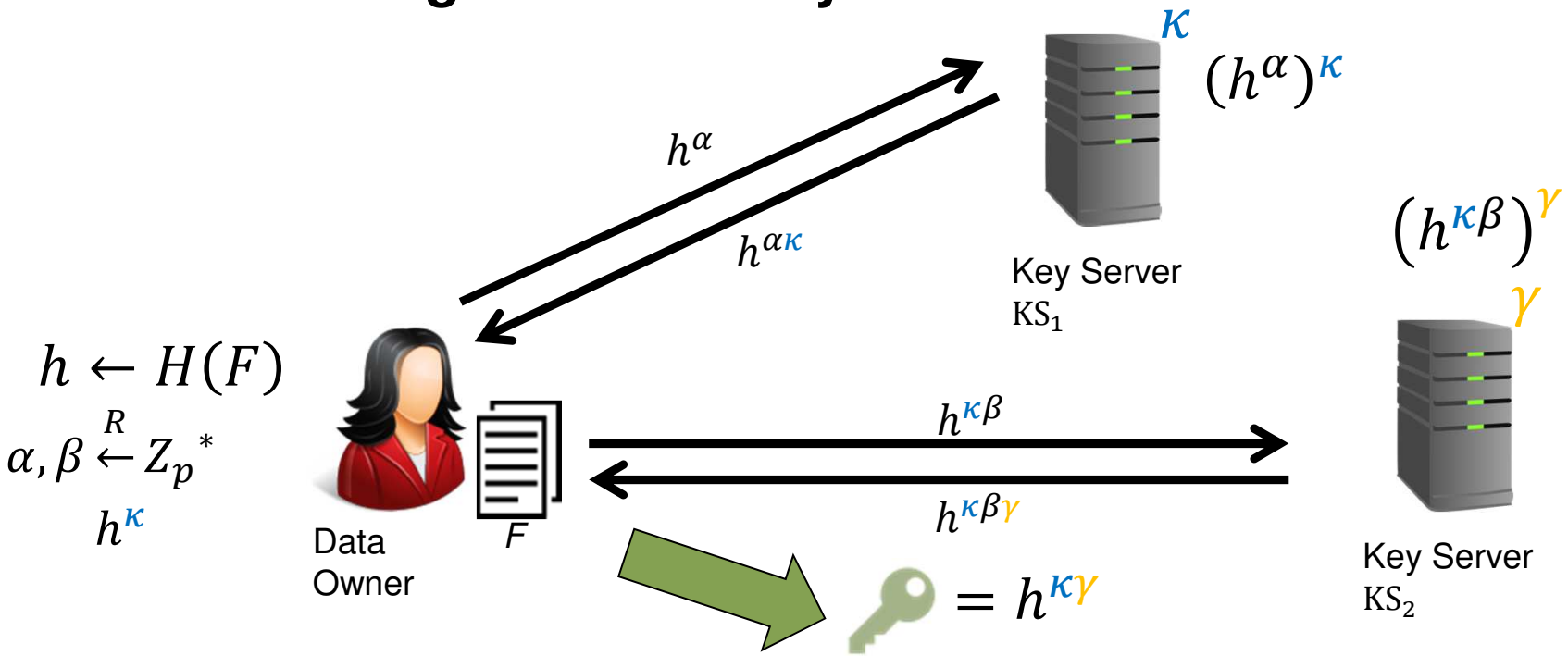
- **PoR can be represented by $P(F, K)$**
- **Derive K from file content**

Convergent Encryption ($K = H(F)$) suffers from dictionary attacks

$\Rightarrow$ **Secure Message-Locked Key Generation**

Message-locked PoR *[CCSW'16]*

- **Secure Message-Locked Key Generation**

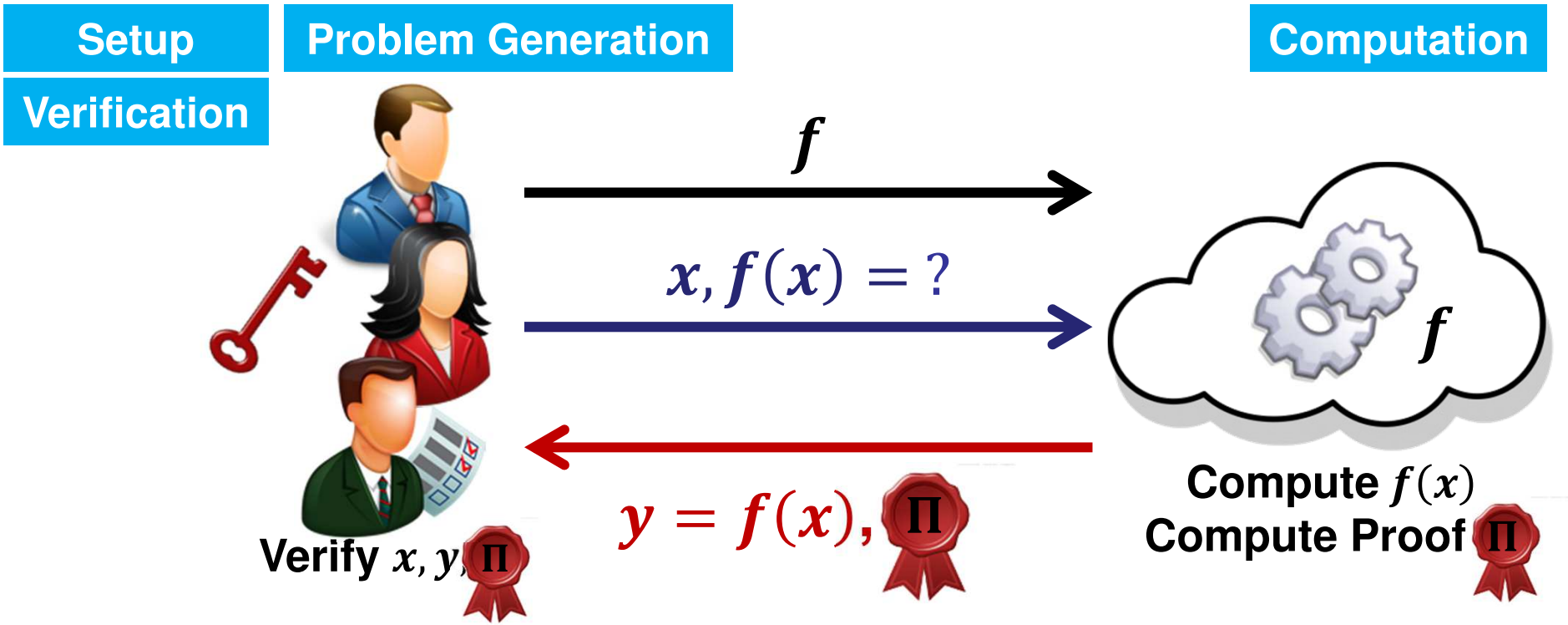


- **Message-locked PoR = PoR using 🔑 in $P(F, K)$**
  - StealthGuard – watchdogs
  - Private Compact PoR - tags [Shacham et al 2008]

Cloud Security Research

- **Privacy**
  - Privacy preserving word search
  - De-duplication on encrypted data
- **Integrity**
  - Proof of Retrievability
- **Verifiability**
  - Verifiable computation
  - Proof composition

Verifiable Computation



R1: Cost(Verify) ≪ Cost(Compute)

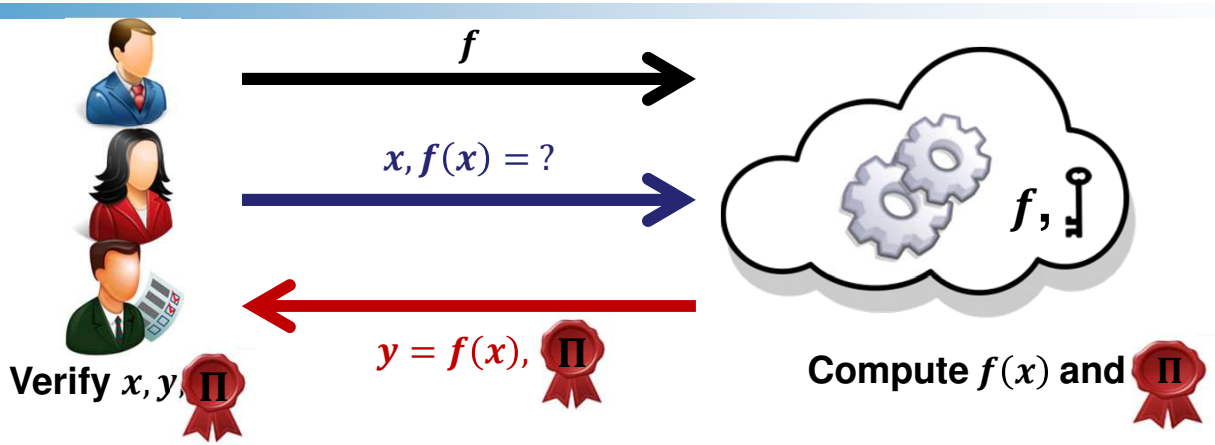
R2: Public delegatability *[Parno et al. 2012]*

Anyone can submit a computation request

R3: Public verifiability *[Parno et al. 2012]*

Anyone can verify a computation result

Verifiability for 3 Operations

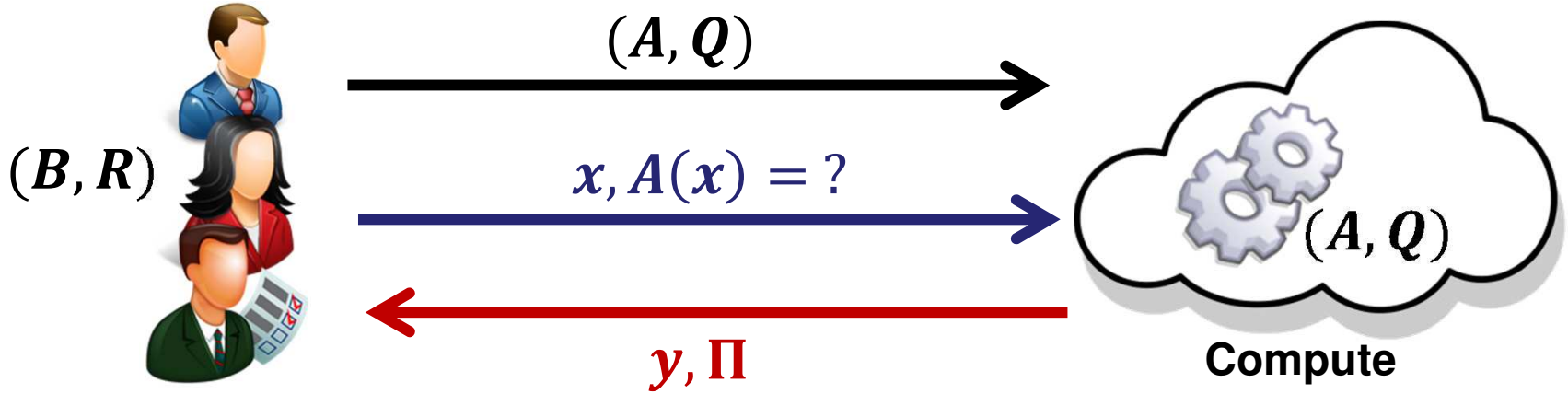


	High-Degree Polynomial Evaluation	Large Matrix Multiplication	Conjunctive Keyword Search
f	$A(X) = \sum_{i=0}^{d} a_i X^i \in \mathbb{F}_p[X]$	$M.\vec{x}$ with $\mathbf{M} = [M_{ij}] \in \mathbb{F}_p^{n \times m}$	Search(.)
x	$x \in \mathbb{F}_p$	$\vec{x} = (x_1, x_2, \ldots, x_m)^\top \in \mathbb{F}_p^m$	Keywords $\mathbb{W} = \{\omega_1, \omega_2, \ldots, \omega_n\}$
y	$y = A(x) \in \mathbb{F}_p$	$\vec{y} = (y_1, y_2, \ldots, y_n)^\top = M\vec{x} \in \mathbb{F}_p^n$	ID of files F_i such that $\mathbb{W} \subset F_i$

Verifiable Polynomial Evaluation – Idea

Euclidean Division of Polynomials

$$A = QB + R$$



Verify

$$y = \Pi\, B(x) + R(x)\ ?$$

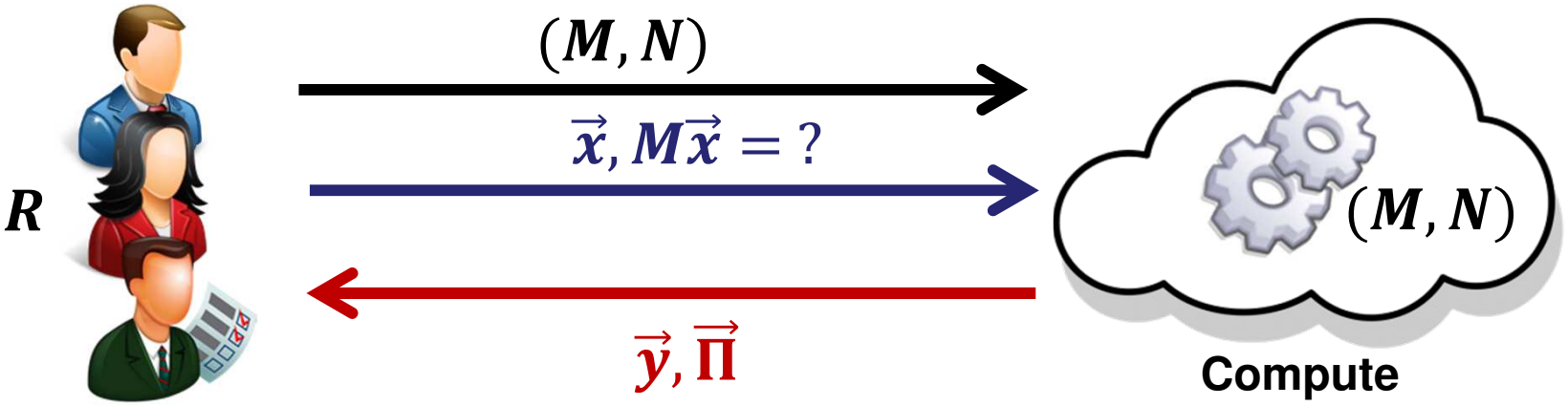
B, R **small degree**

Verifiable Matrix Multiplication – Idea

Auxiliary Matrices

$$N = \delta M + R$$

R *pseudo-random*



Verify

$$\vec{\Pi} = \delta\vec{y} + R\vec{x}\ ?$$

Projection $\vec{\lambda} \cdot \vec{\Pi} = \vec{\lambda} \cdot \delta\vec{y} + \vec{\lambda} \cdot (R\vec{x})$

Cloud Security Research

- **Privacy**
  - Privacy preserving word search
  - De-duplication on encrypted data
- **Integrity**
  - Proof of Retrievability
- **Verifiability**
  - Verifiable computation
  - Proof composition

Proof Composition Problem

Verifiability of general purpose programs

- **Efficient methods for handling sequence of operations Pinocchio** *[Parno et al]*
- **Efficient schemes for a single very complex operation**
- **No technique achieving both purposes**

$$
\begin{aligned}
&\textit{program } P(x)\\
&\qquad a \coloneqq A(x)\\
&\qquad b \coloneqq B(a)\\
&\qquad \vdots\\
&\qquad z \coloneqq Z(y)
\end{aligned}
$$

Example:

$$
\begin{aligned}
&\textit{program } NN2(x)\\
&\qquad a \coloneqq M_1.x\\
&\qquad b \coloneqq relu(a)\\
&\qquad c \coloneqq M_2.b
\end{aligned}
$$

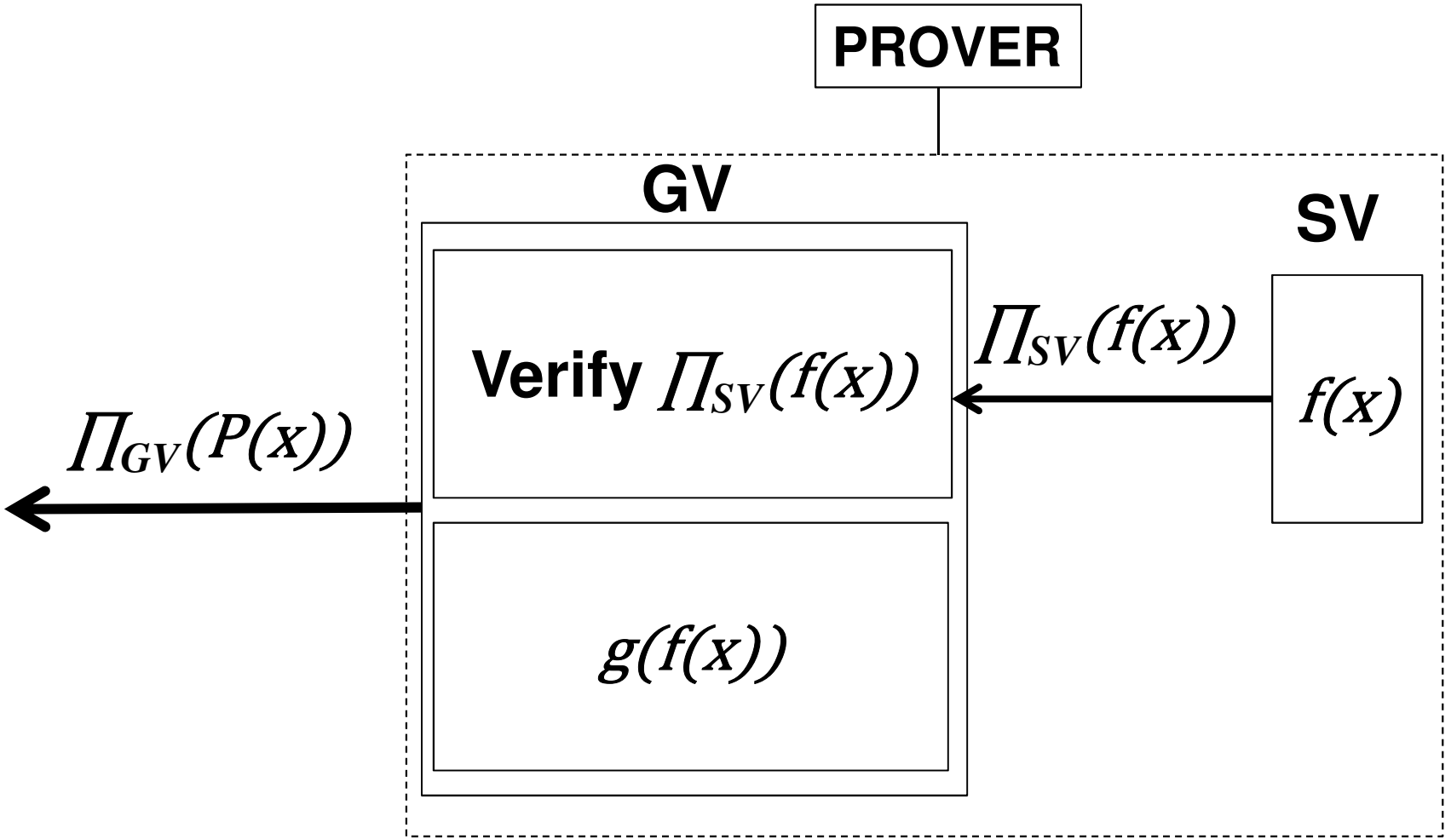
Proof Composition - Problem

$$program\ P(x)$$
$$a := f(x)$$
$$b := g(a)$$

- f : high complexity
- g : low complexity
- GV : verifiability for a sequence of operations (Pinocchio)
- SV : verifiability for a complex function (product of very large matrices)

Proof Composition - Idea

Outsourced proof generation



Conclusion

- **Privacy**
- **Integrity**
- **Verifiability**

Outsourcing

Big Data

Outlook

- **Efficient & practical**
- **Integration - S&P with Cloud, DB**
- **“New topics”**
  - Secure deletion
  - Proof of reliability
  - Verifiability / location, physical memory

Acknowledgments

- **EU Projects**

- **Collaborators**
  - Julien Keuffer, Iraklis Leontiadis, Pasquale Puzio, Cédric Van Rompay, Dimitrios Vasilopoulos
  - Monir Azraoui, Erik Blass, Roberto Di Pietro, Kaoutar Elkhiyaoui, Melek Önen

Papers

- *A leakage-abuse attack against multi-user searchable encryption,* C. Van Rompay, R. Molva, M. Önen, PETS 2017
- *Reconciling security and functional requirements in multi-tenant clouds,* G. Karame, M. Neugschwandtner, M. Önen, H. Ritzdorf, SCC 2017
- *Message-locked proofs of retrievability with secure deduplication,* D. Vasilopoulos, M. Önen, Melek; K. Elkhiyaoui, R. Molva, CCSW 2016
- *Efficient Techniques for Publicly Verifiable Delegation of Computation,* M. Azraoui, K. Elkhiyaoui, M. Önen, R. Molva, ASIACCS 2016
- *PUDA- Privacy and Unforgeability for Data Aggregation,* I. Leontiadis, K. Elkhiyaoui, M. Önen, R. Molva, in CANS 2015
- *Publicly verifiable conjunctive keyword search in outsourcing databases,* M. Azraoui, K. Elkhiyaoui, M. Önen, R. Molva, SPC 2015,
- *Multi-user searchable encryption in the cloud,* C. Van Rompay, R. Molva, M. Önen, ISC'15
- *PerfectDedup,* P. Puzio, R. Molva, M. Önen, S. Loureiro, DPM 2015
- *StealthGuard: Proofs of Retrievability with hidden watchdogs,* M. Azraoui, K. Elkhiyaoui, R. Molva, M Önen, ESORICS 2014

Papers (cont'd)

- *Privacy preserving delegated word search,* K. Elkhiyaoui, M. Önen, R. Molva, SECRYPT 2014
- *A-PPL: An accountability policy language, M. Azraoui*, K. Elkhiyaoui, M. Önen, K. Bernsmed, A. Santana de Oliveira, J. Sendor, DPM 2014
- *Private and dynamic time-series data aggregation with trust relaxation*, I. Leontiadis, K. Elkhiyaoui, R. Molva, CANS 2014
- *Privacy preserving statistics in the smart grid*, I. Leontiadis, R. Molva, M. Önen, DASEC 2014
- *ClouDedup:Secure deduplication with encrypted data for cloud storage*, P. Puzio, R. Molva, M. Önen, S. Loureiro, CLOUDCOM 2013
- *Privacy preserving delegated word-search in the cloud*, K. Elkhiyaoui, M. Önen, R. Molva, TCLOUDS 2013
- *PRISM- Privacy preserving Search in Map Reduce,* E.-O. Blass, R. Di Pietro, R. Molva, M. Önen, PETS 2012

EURECOM
Sophia Antipolis

THANK YOU