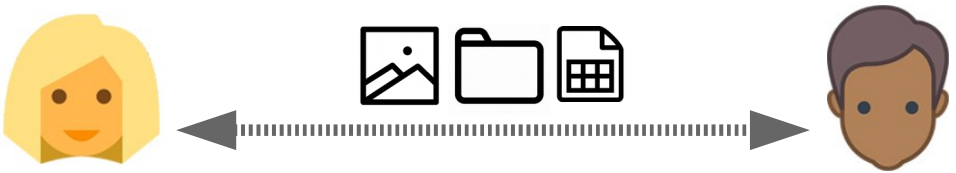
Stefan Contiu – PhD Defense

13 Nov. 2019

Applied Cryptographic Access Control for Untrusted Cloud Storage

LABORATOIRE BORDELAIS DE RECHERCHE EN INFORMATIQUE LaBRI

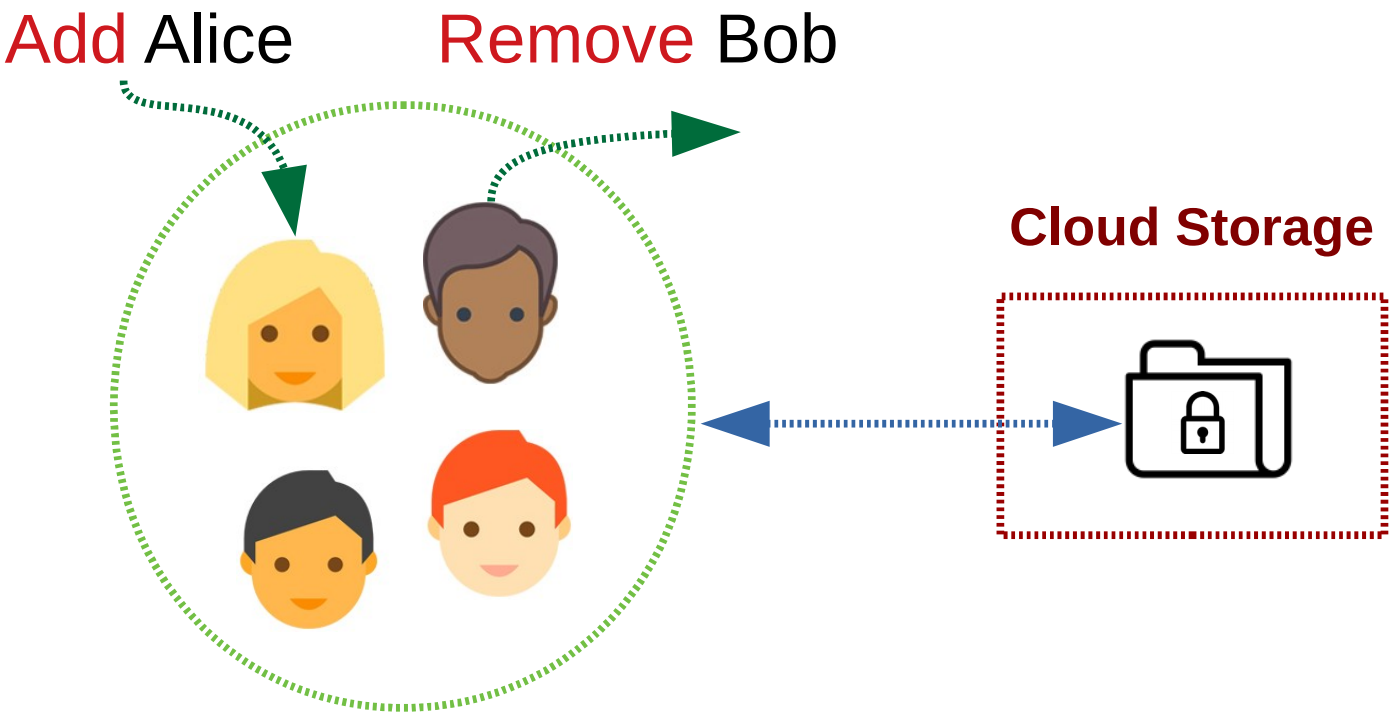
Sharing Data over Clouds



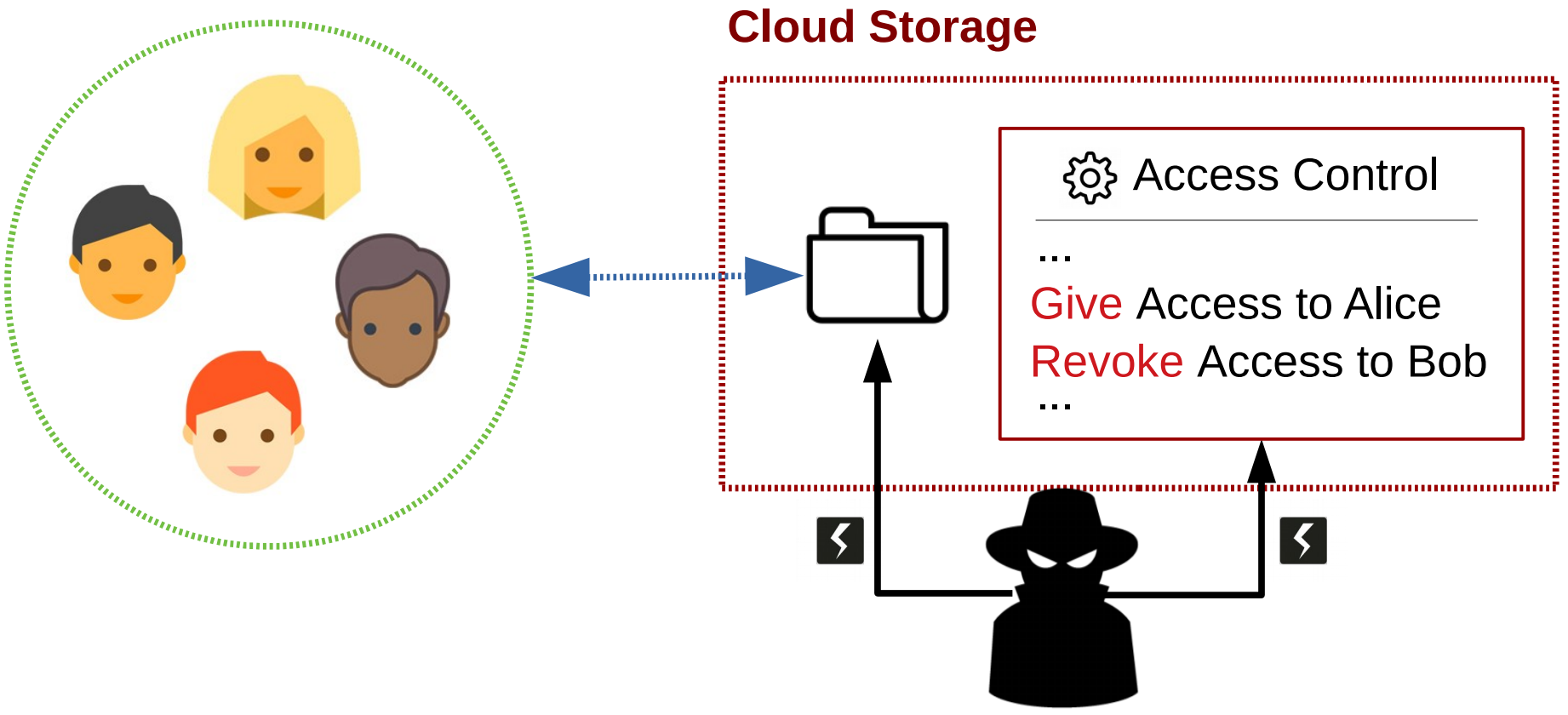
Data sharing is *easier* and *cheaper* than ever before.

Dropbox : 500 M active users.

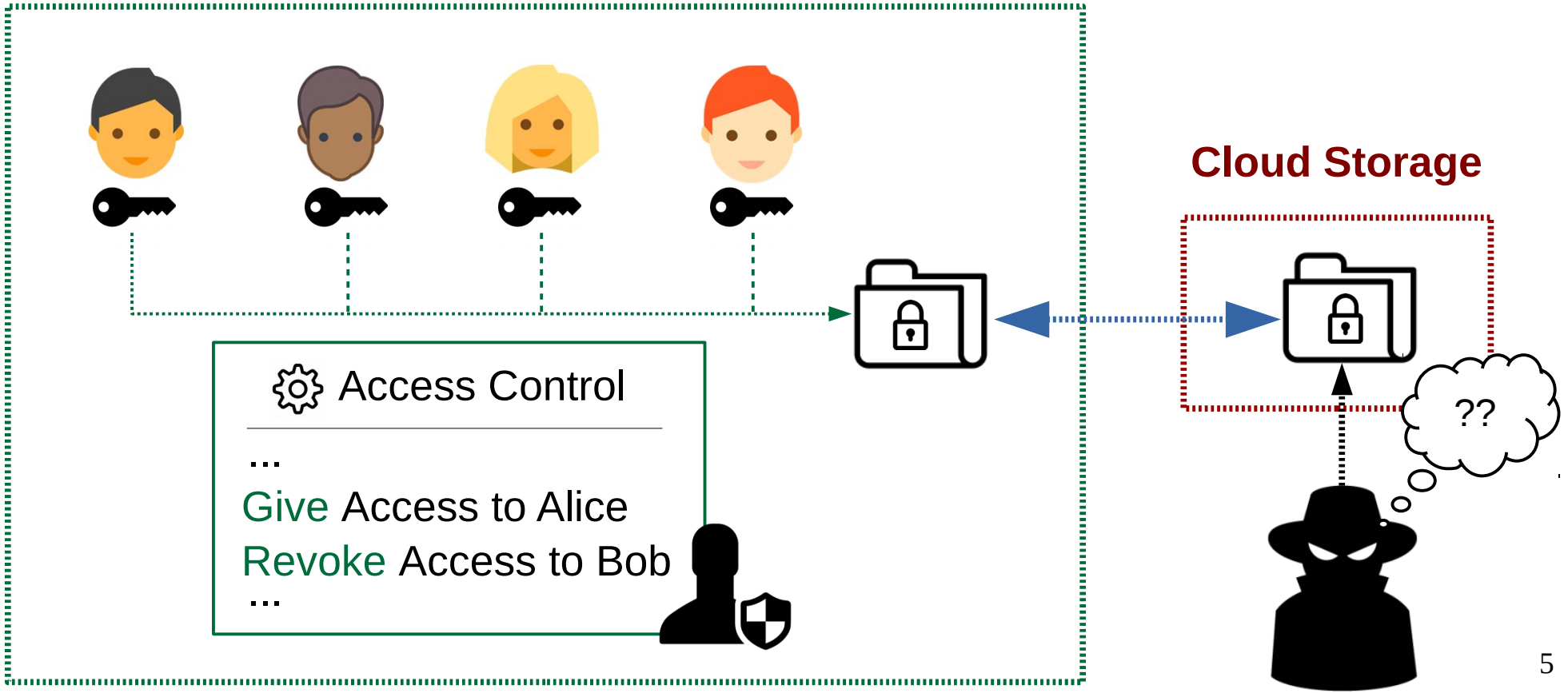
Group Access Control (GAC)



Cloud performs GAC



End-to-end Encryption (E2EE)



GAC for End-to-end Encryption

Access Control

Give Access to Alice
Revoke Access to Bob

Enforced Cryptographically by end-users.

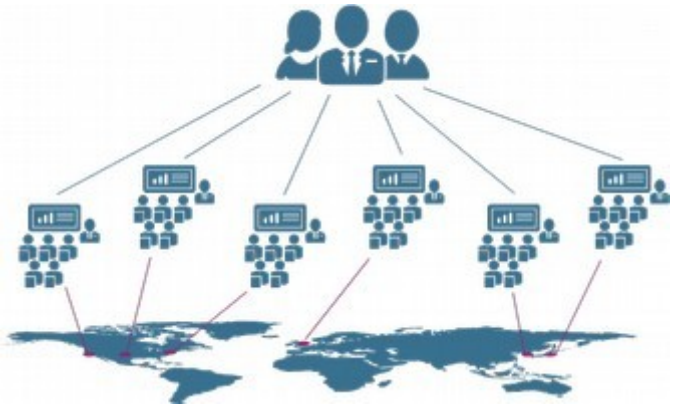
POSSIBLE WITH **parsec.cloud** BY SCILLE

E2EE storage & sharing

usability

open source

But, what about large scale ?



Large organization has :

- large user base,
- dynamic workloads,
- large data volume.

What are the GAC requirments ?

Concrete scenario : United Nations

1. Confidentiality for Large Groups

United Nations

Peacekeeping Operations

> 100K users

UNHCR
The UN Refugee Agency

WFP
World Food Programme

Troops

Police

Field Support

Field Support

Field Support

Personnel

Budget

Logistics

IT & Comm.

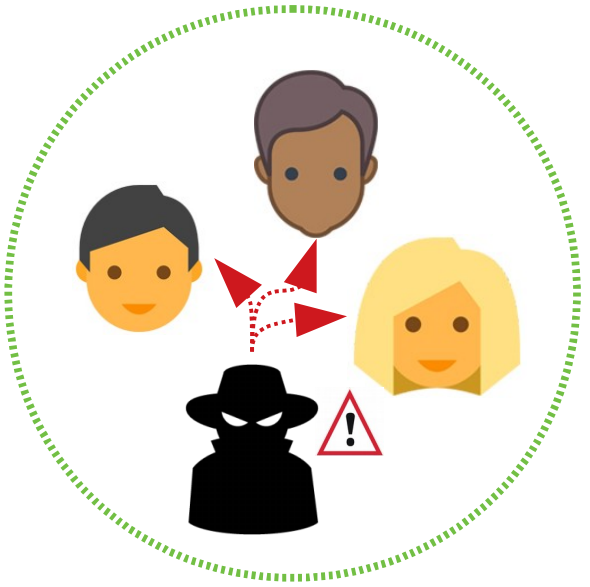
2. Anonymity inside Large Groups

United Nations

Peacekeeping Operations

Security Clearances

External Contractors



3. Revocation of Large Data-sets

United Nations

Peacekeeping Operations

Imagery Intelligence

Revoked User

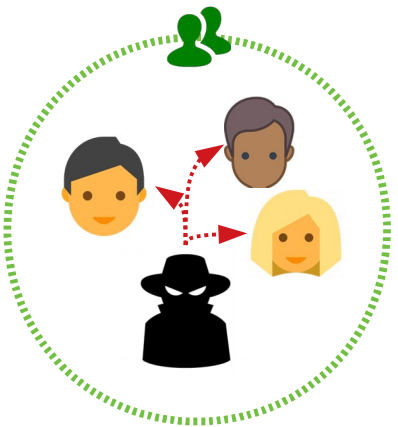
Petabyte(s)

Does E2EE Group Access Control work well?

Confidentiality of **Large** Groups

Not efficient.

Anonymity inside **Large** Groups

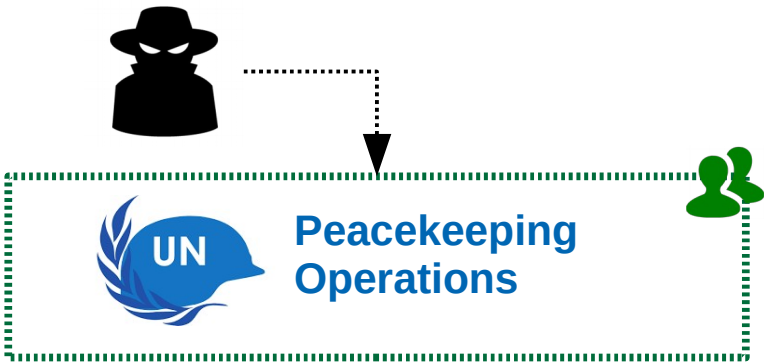


Not efficient.

Revocation of **Large** Data

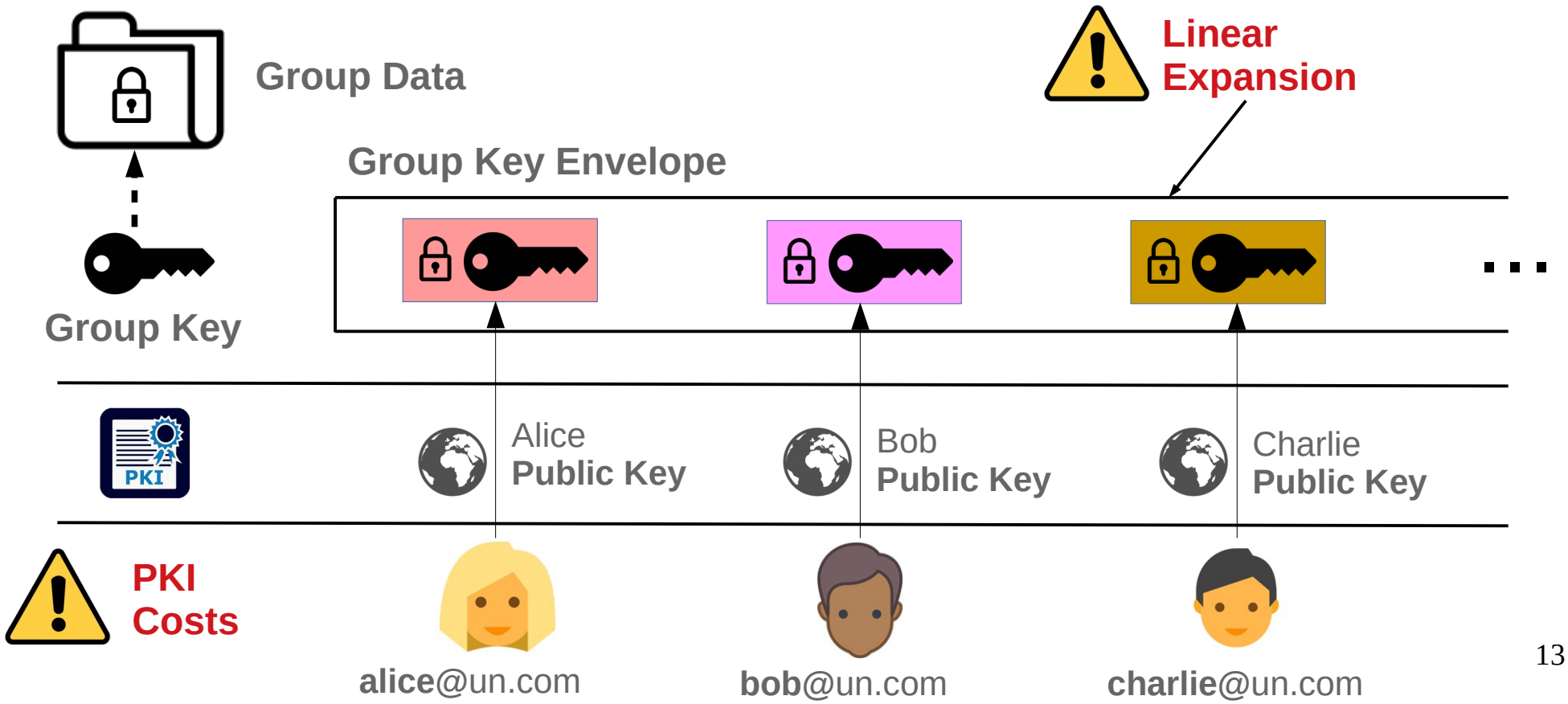
Not efficient.

Confidentiality of **Large** Groups

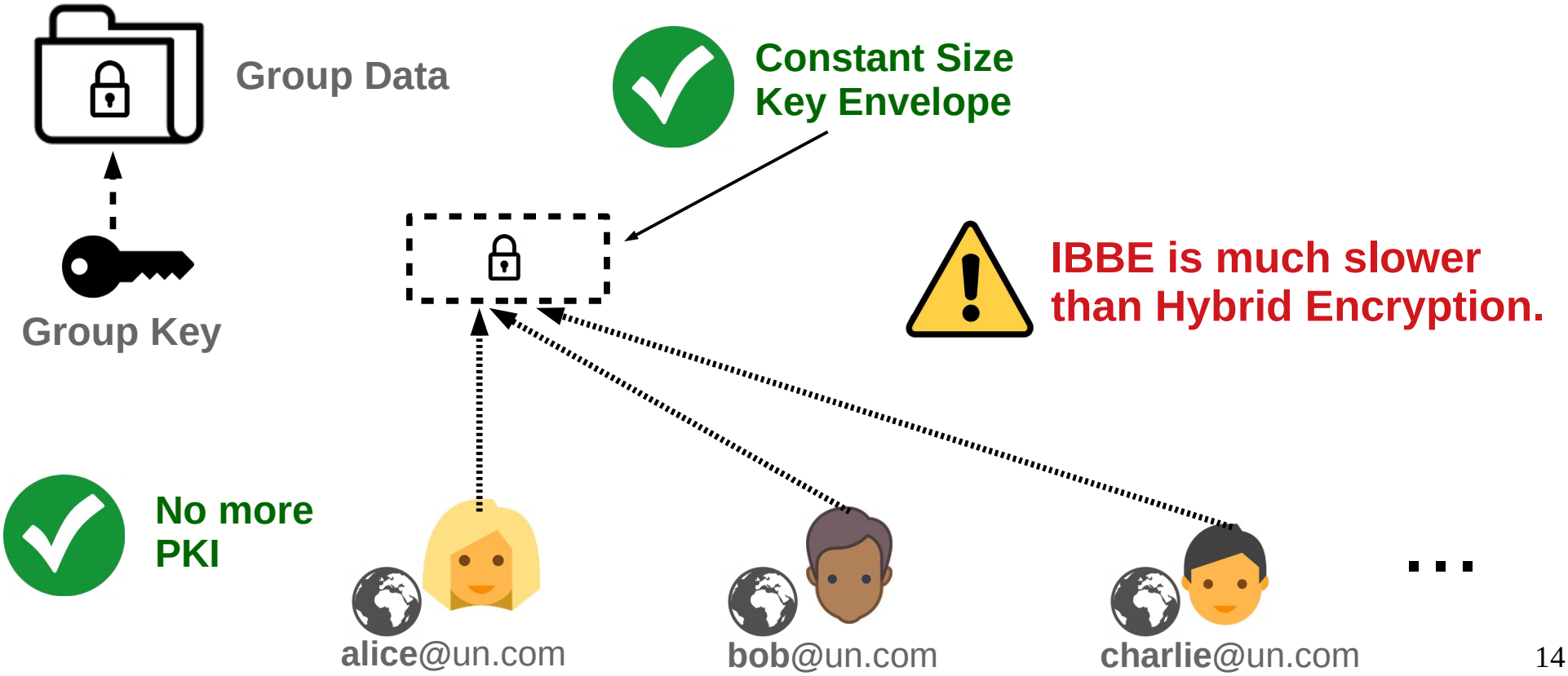


Not efficient.

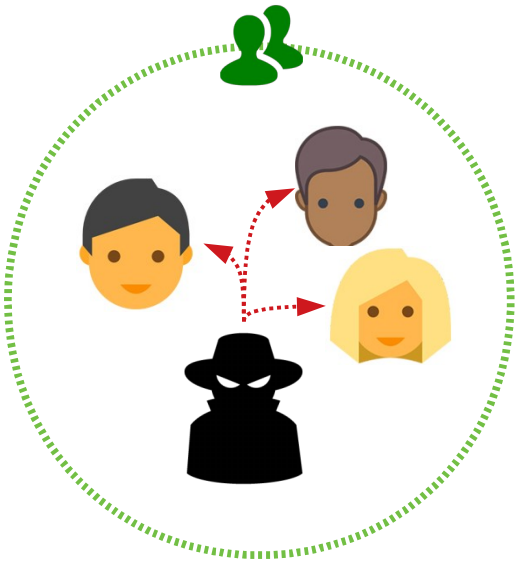
Hybrid Encryption (HE)



Identity Based Broadcast Encryption (IBBE)



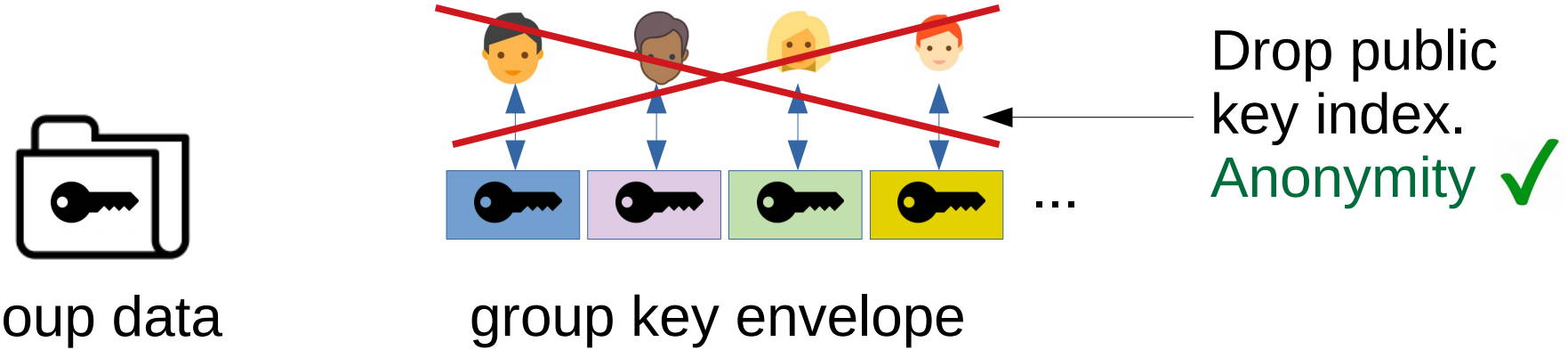
Anonymity inside **Large** Groups



⚠ Not efficient.

Pretty Good Privacy (PGP)

- `hidden-recipient` mode :



Without an index, decryption : *N/2* trials to find the key.

Anonymous Broadcast Encrypt.

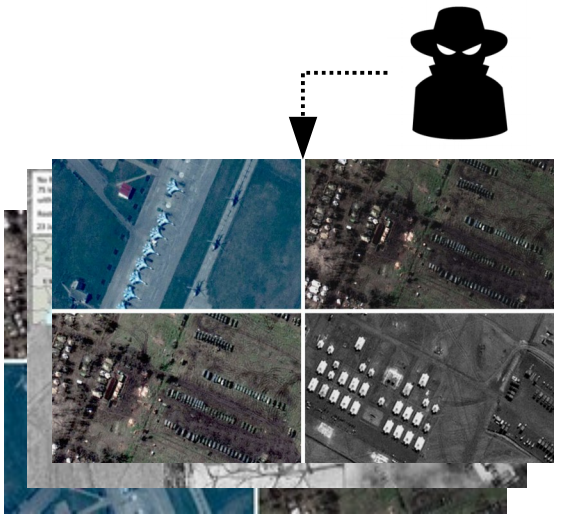
- Uses PGP method : “drop the public key index”
- Sign each time constructing envelope (IND-CCA)

Impractical for large user bases : 330 users/s

Confidentiality of
Large Groups

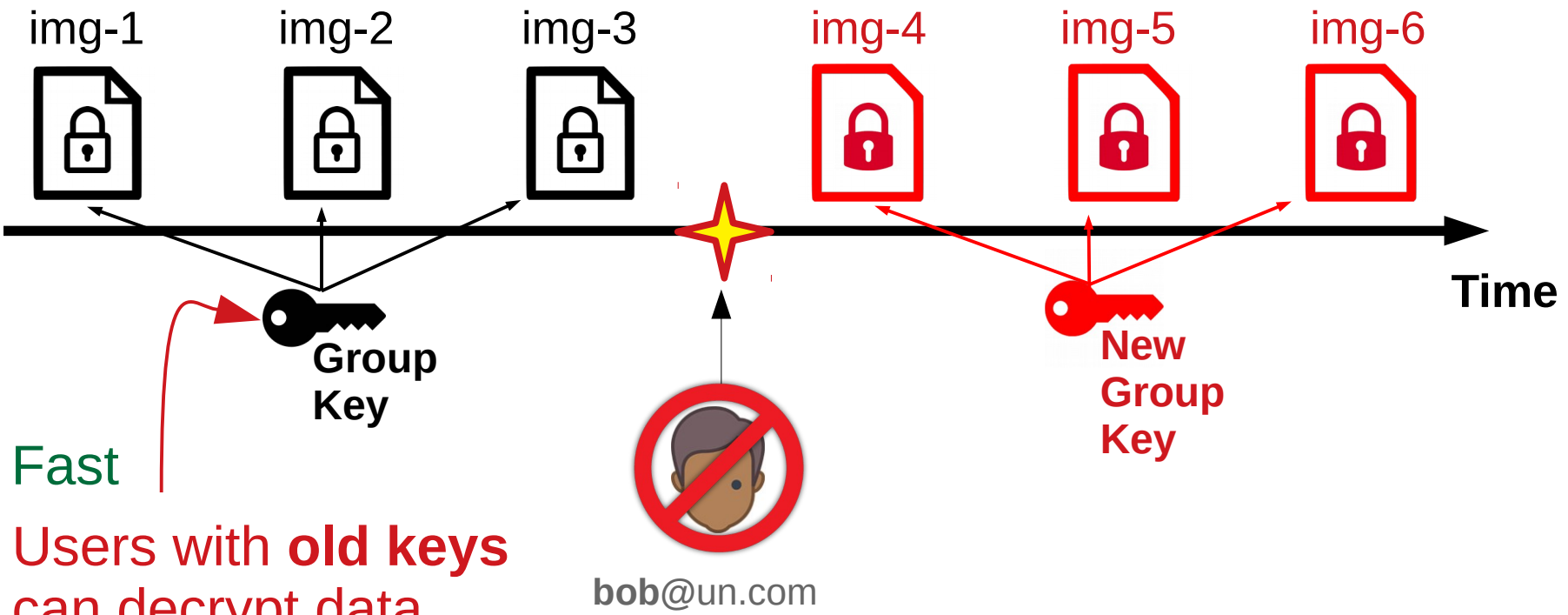
Anonymity inside
Large Groups

Revocation of **Large** Data



Not efficient.

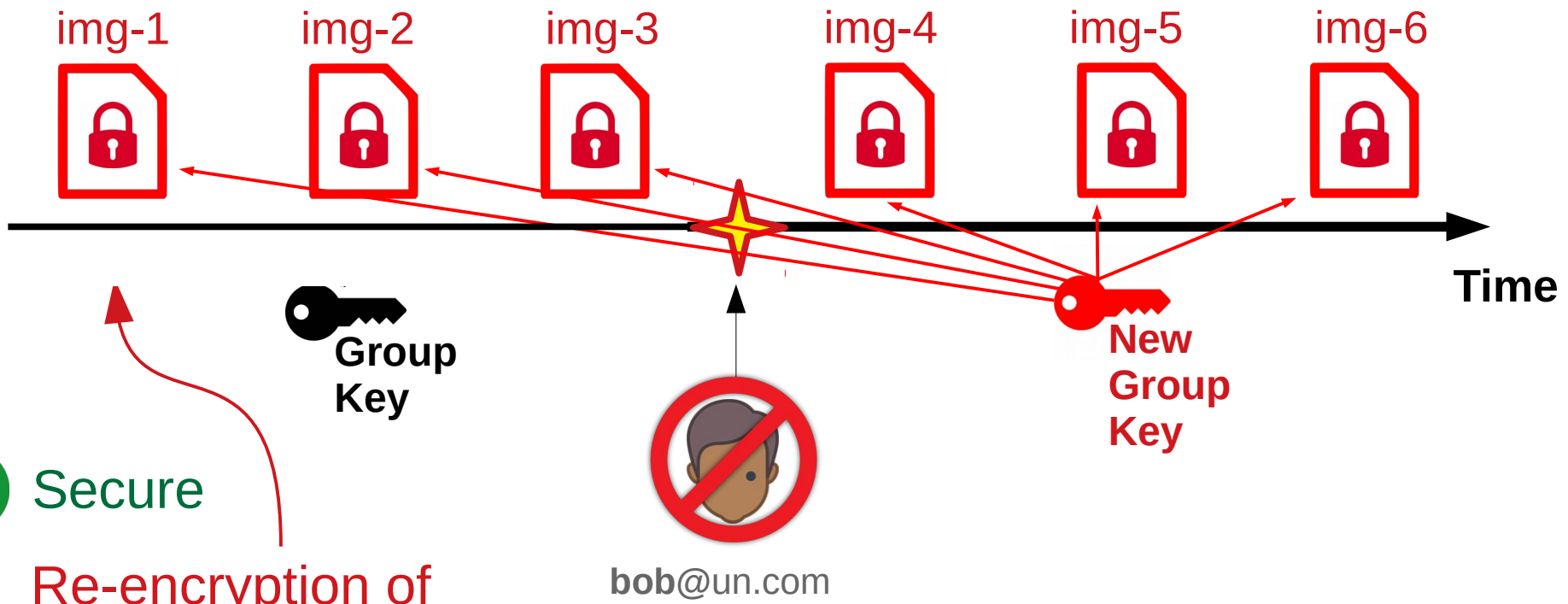
Lazy Revocation



Fast

Users with **old keys** can decrypt data after revoked.

Active Revocation



Secure

Re-encryption of 1 PB takes days...

Instrument for Efficiency

Availability of Trusted Execution Environments

- Isolate code and data inside an *enclave*.
- Provide execution confidentiality and results integrity.

Administrators are equipped with TEE.

Intel SGX as TEE

Widespread adoption in the research world.

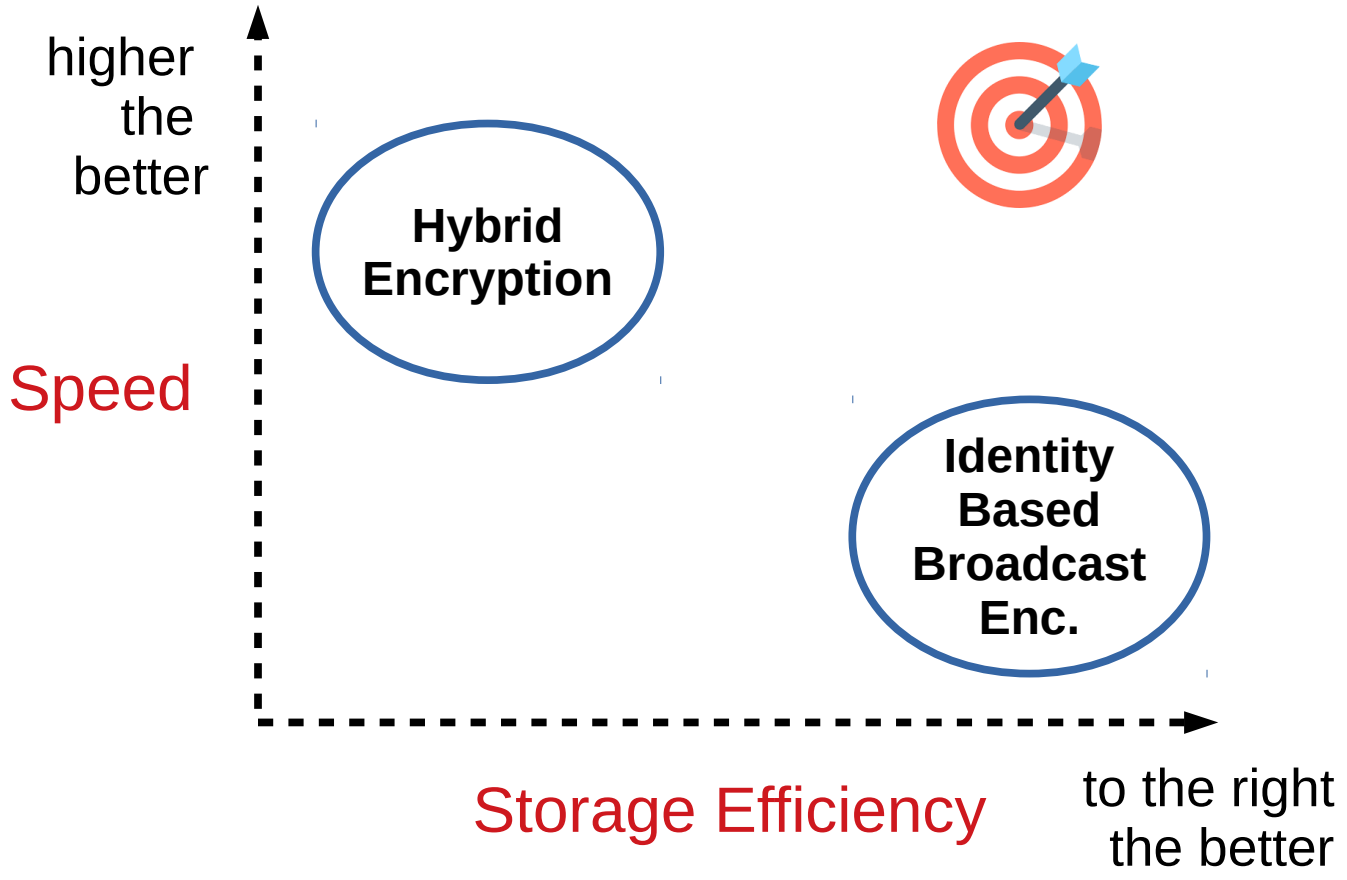
Can persist data outside enclave by sealing.

Can be attested before running.

Limitations : memory, context switch.

Confidentiality of **Large** Groups

HE vs. IBBE

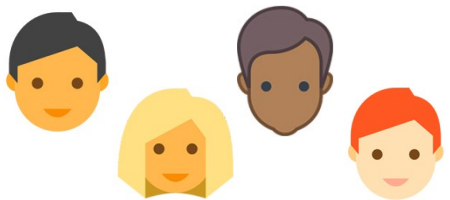

higher
the
better
Speed
Hybrid
Encryption
Identity
Based
Broadcast
Enc.
Storage Efficiency
to the right
the better

Zoom in IBBE

IBBE : O(1) storage $O(n^2)$ computation

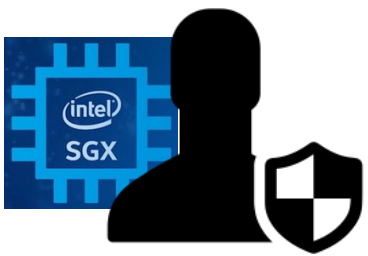
Q : Who runs access control changes?

“Traditional” IBBE :



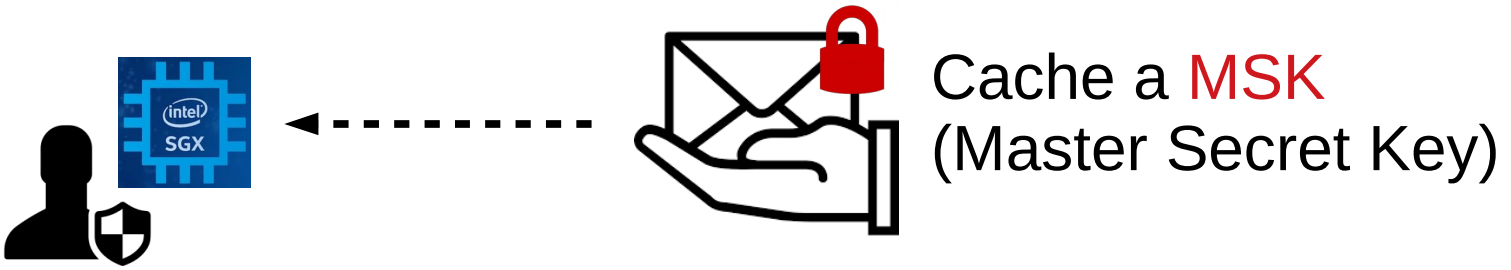
anybody

Our context :



administrator w/ TEE

Running IBBE in SGX

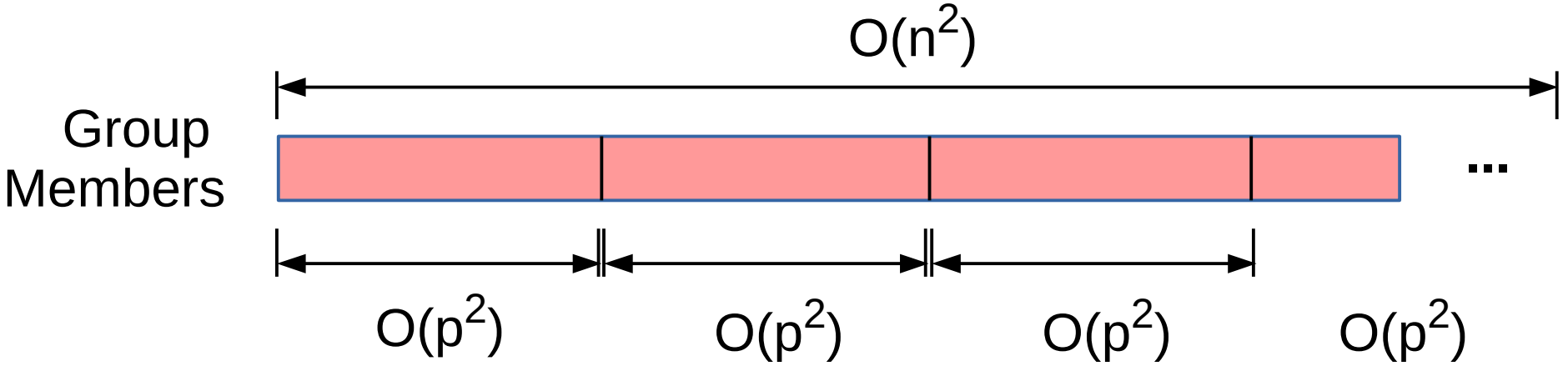


- Use MSK for Administrator Operations
- Computational Cost : $O(n^2) \rightarrow O(n)$
- Users do not have TEE

Running IBBE by Users

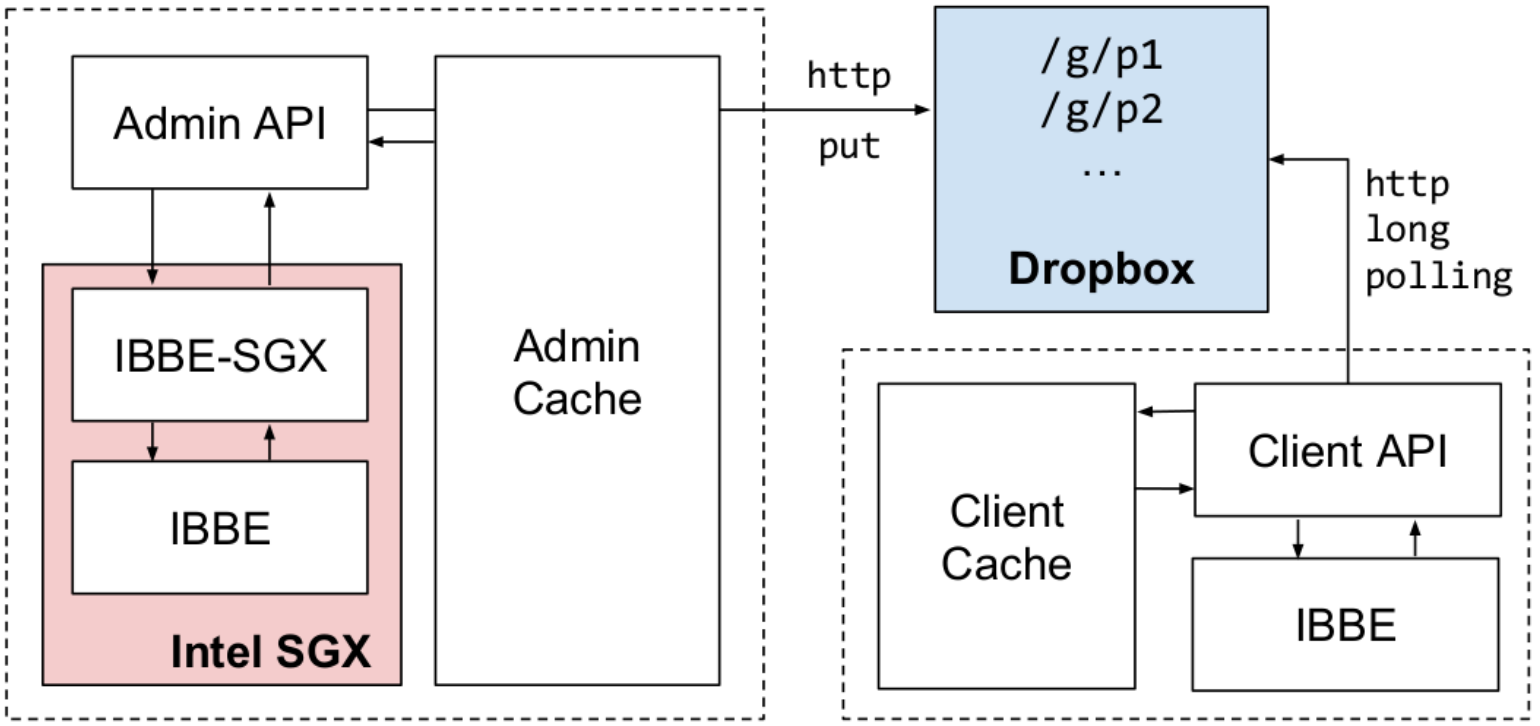
IBBE User Side (no SGX) : $O(n^2)$

Split the Group into Partitions :

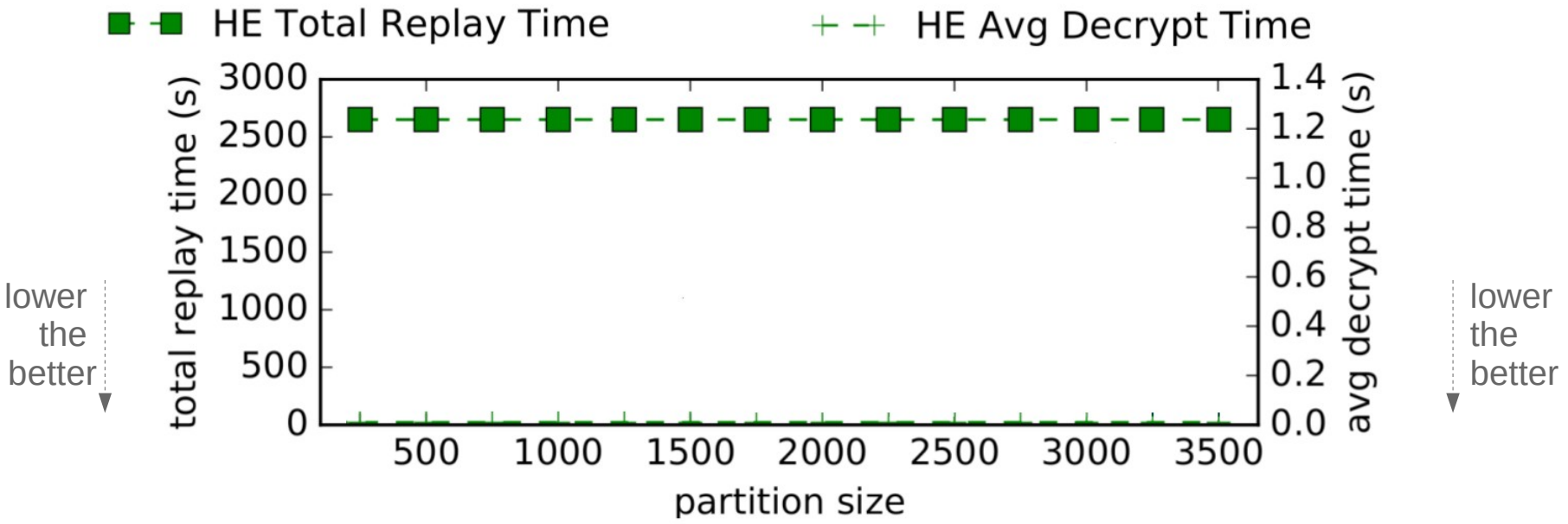


...where p << n.

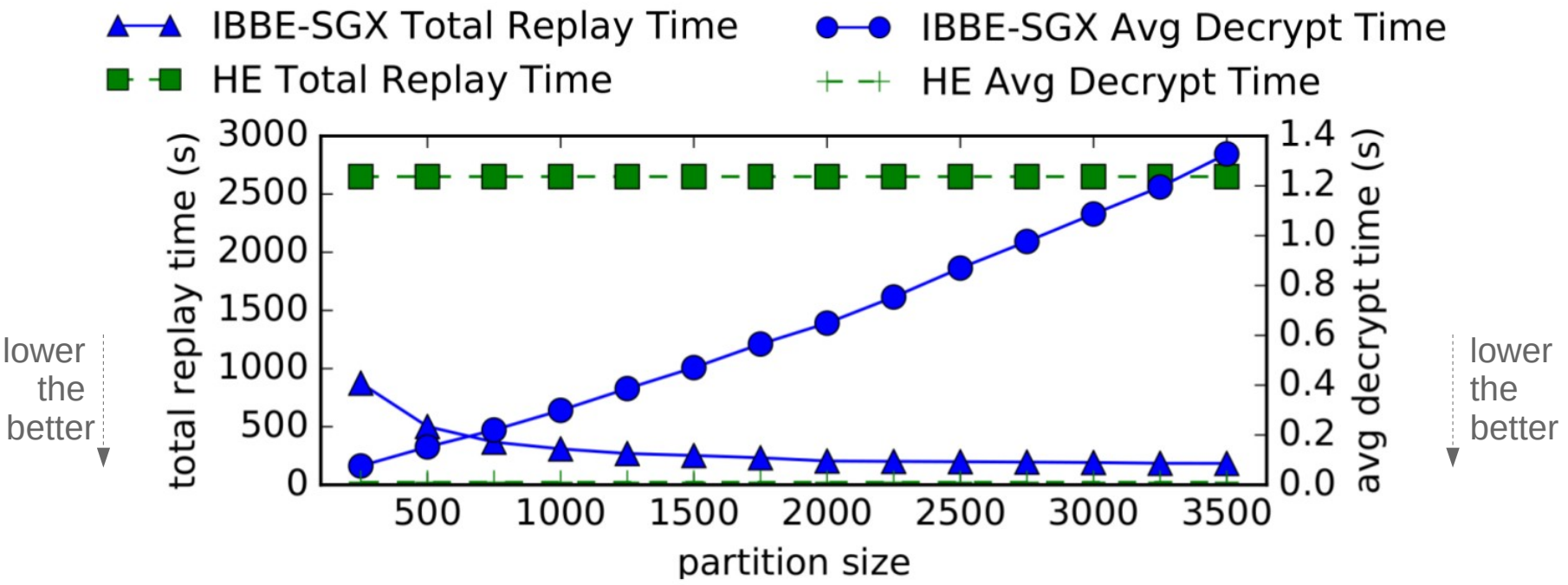
System Big Picture



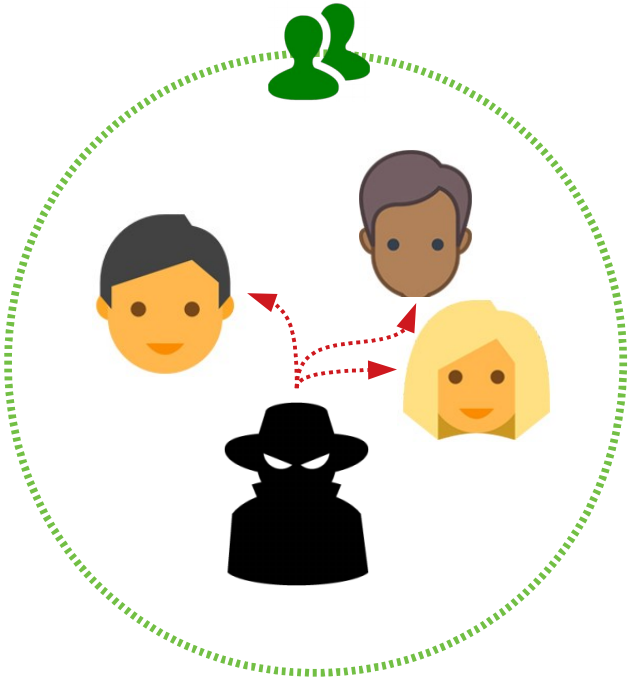
Real Trace Replay



Real Trace Replay

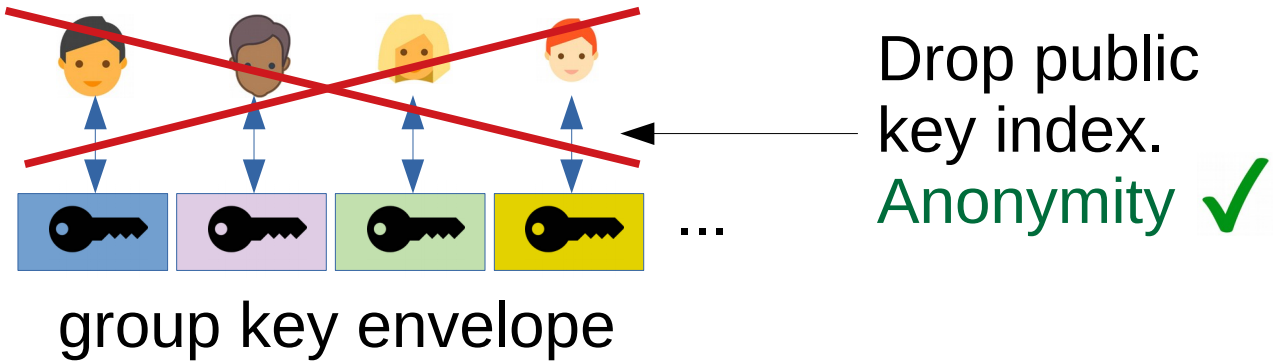


Anonymity inside **Large** Groups



Recall state-of-the-art

- IBBE-SGX does not support anonymity :
  - Operations and partitioning require identities.
- Anonymous Broadcast Encryption :



Instrument for efficiency

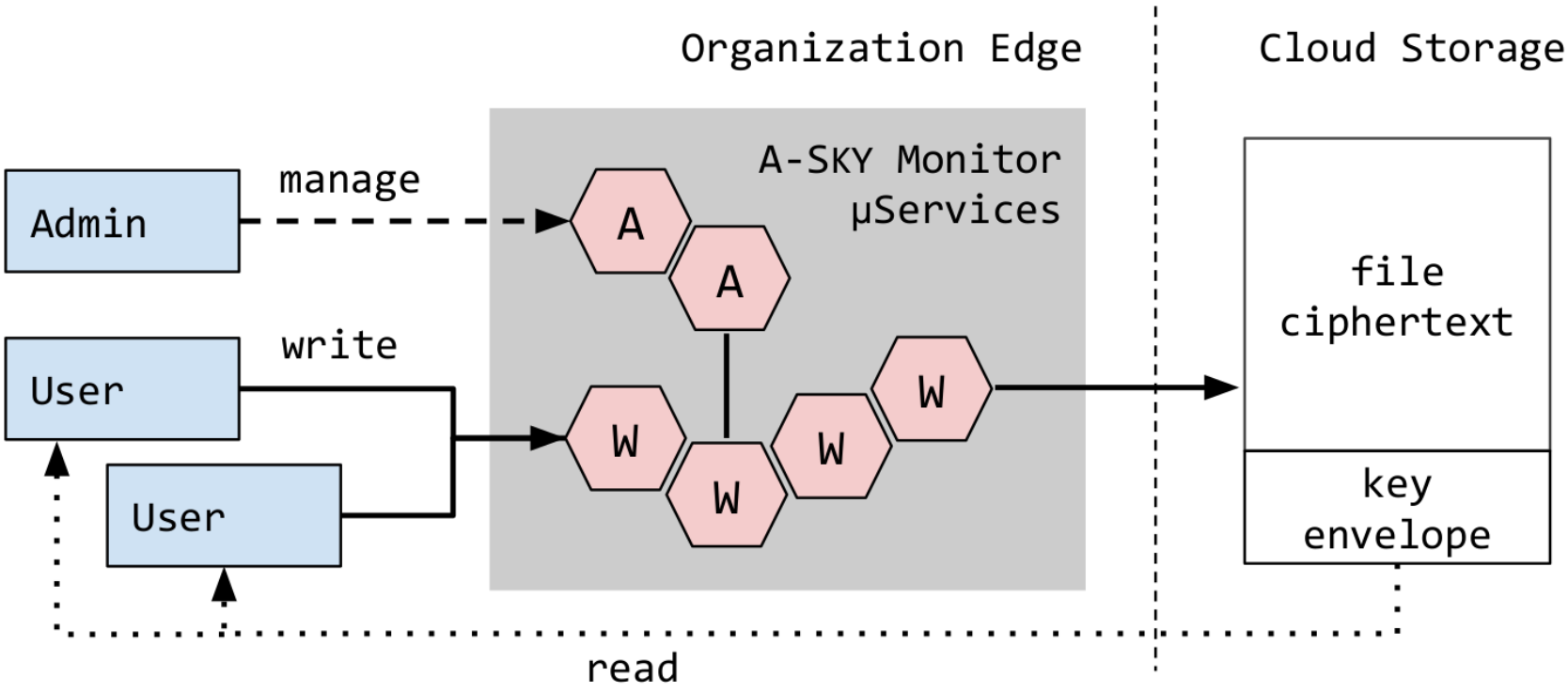
- Trusted Execution Environments (TEE)

Intel Software Guard Extensions

⚠ Performance Limitations ⚠

1. Proxy all *writes* through an SGX enclave but not *reads*.
2. Elastically *scale* depending on load.

A-SKY : Solution Overview

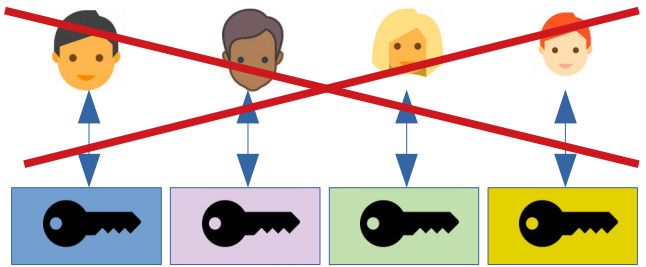


A AccessControl

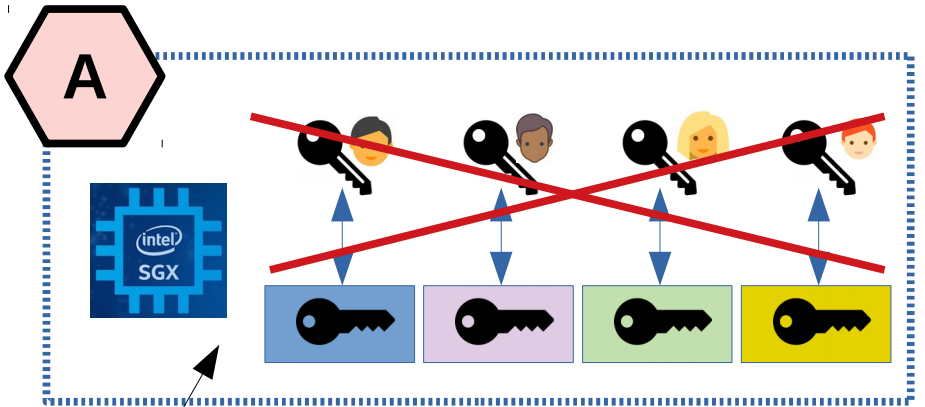
W WriterProxy

Efficiency Gain for Key Envelope

Traditional ANOBE
uses Public Key Encryption



A-SKY uses
Symmetric Key Encryption



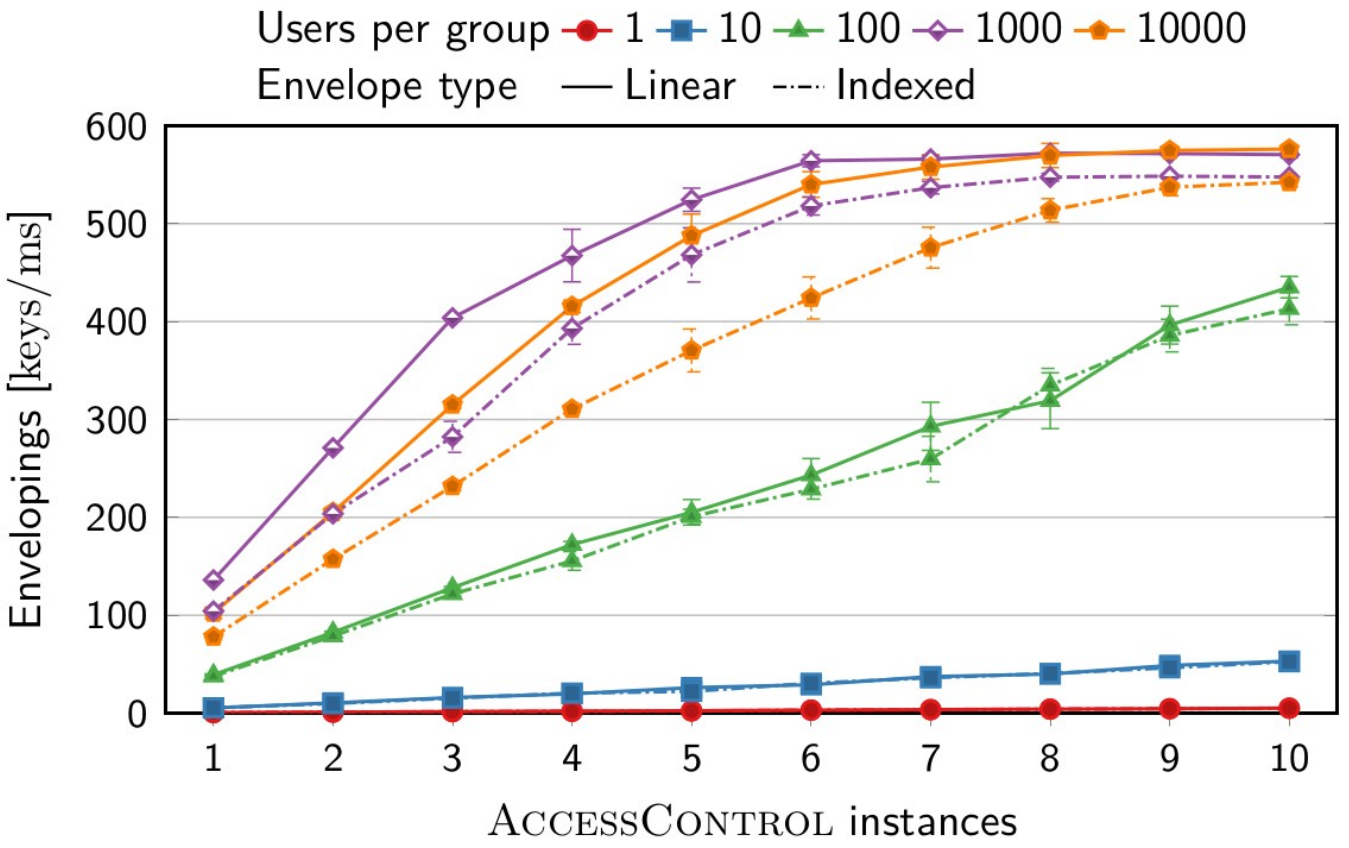
✓ Fast & Small Ciphertexts

Evaluation : ANOBE vs A-SKY

	Enveloping $[\|\mathcal{G}\|/s]$	De-enveloping $[\|\mathcal{G}\|/s]$
ANOBE	3.3×10^2	5×10^3
A-SKY	1.9×10^6	2.5×10^6
Faster by	**3.7 OoM**	**2.6 OoM**

Key Enveloping Performance

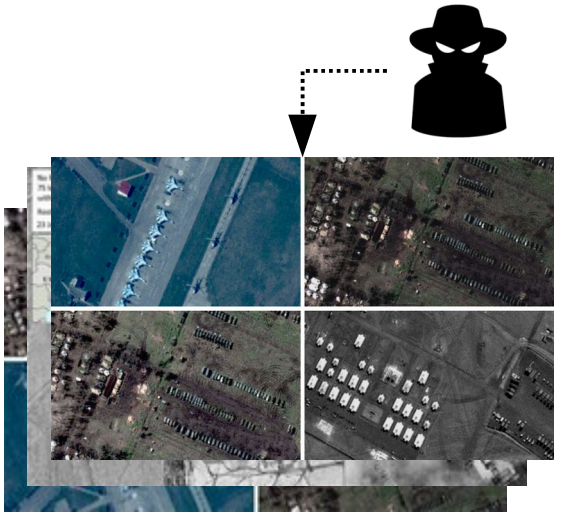
higher the better



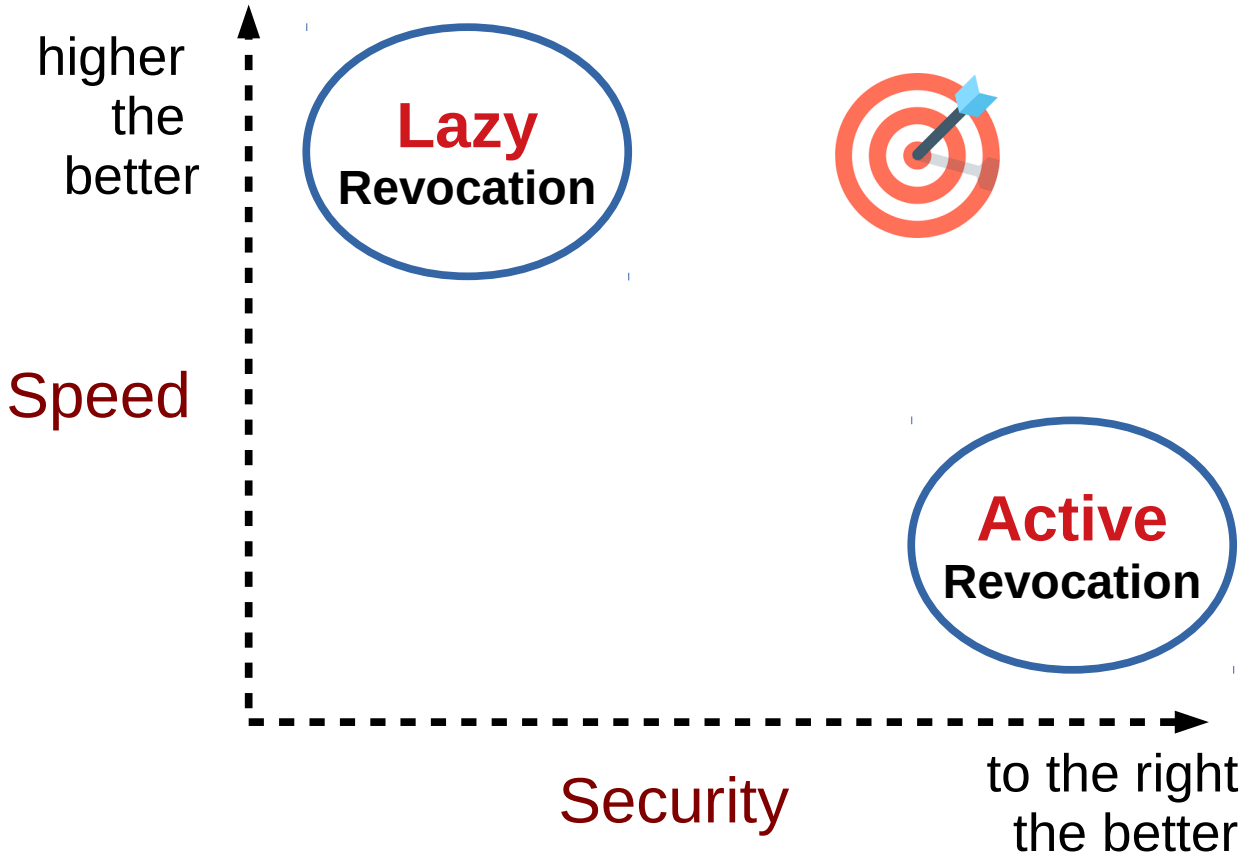
Confidentiality of
Large Groups

Anonymity inside
Large Groups

Revocation of **Large** Data

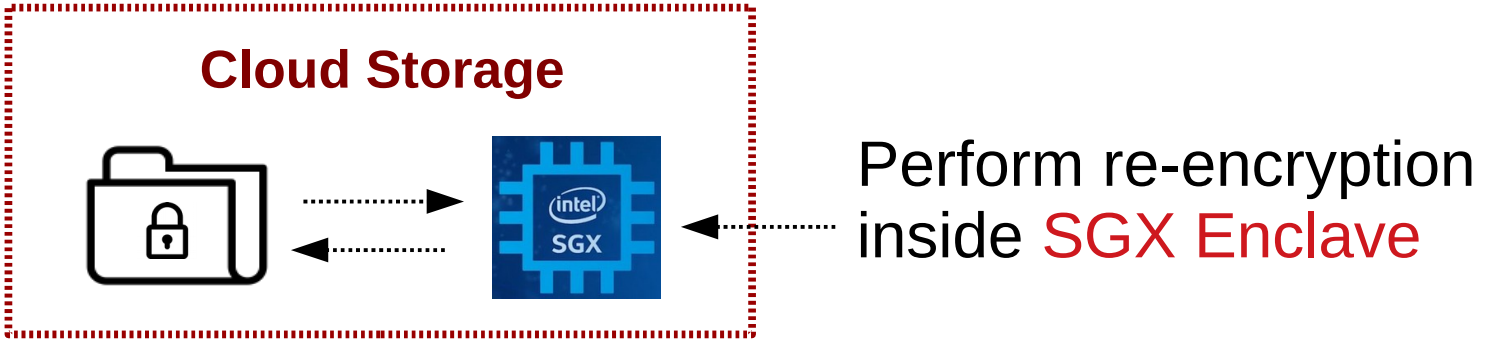


Lazy vs. Active Revocation



Lower I/O of Active Revocation

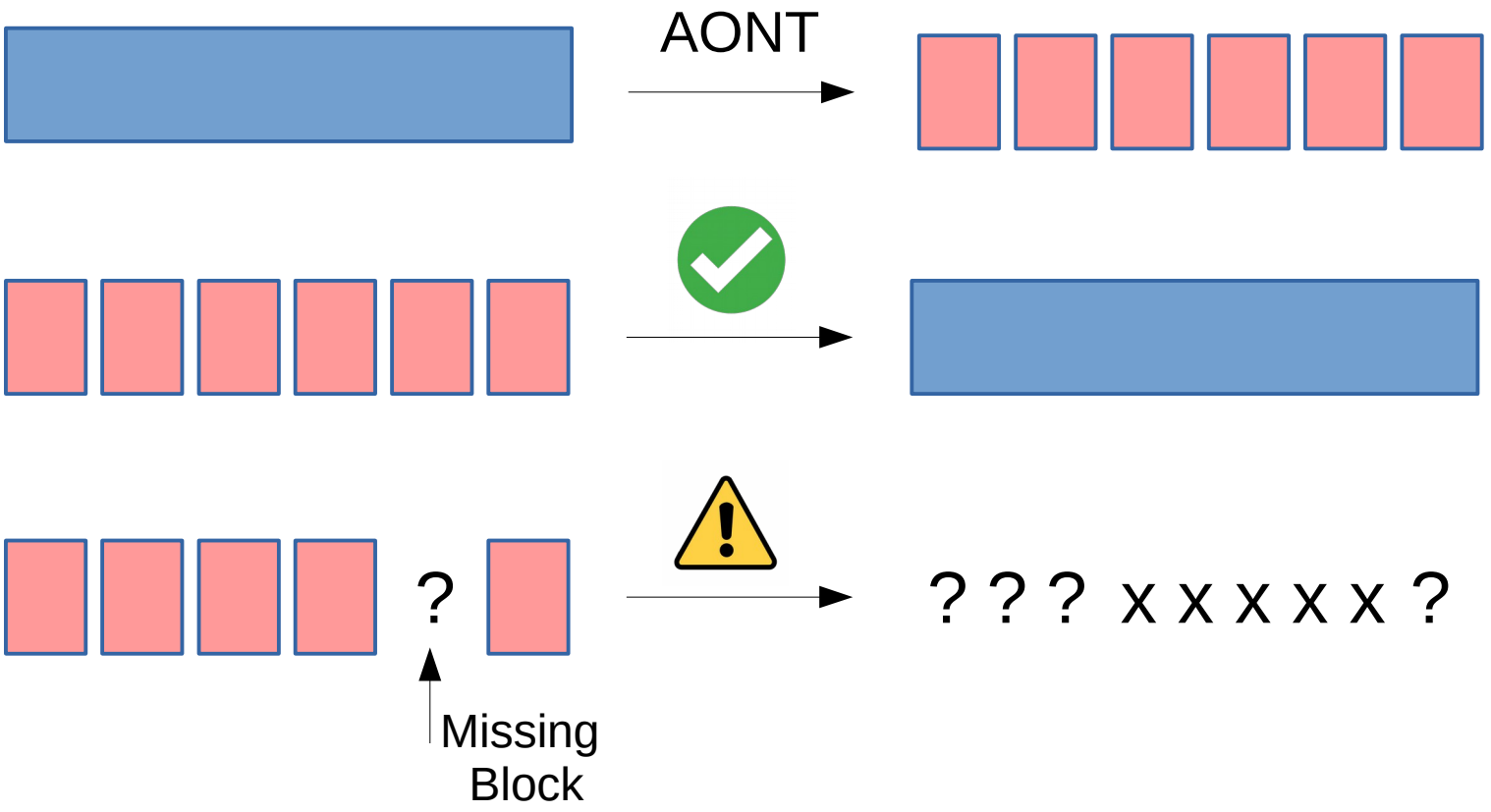
Cloud has slow response time.



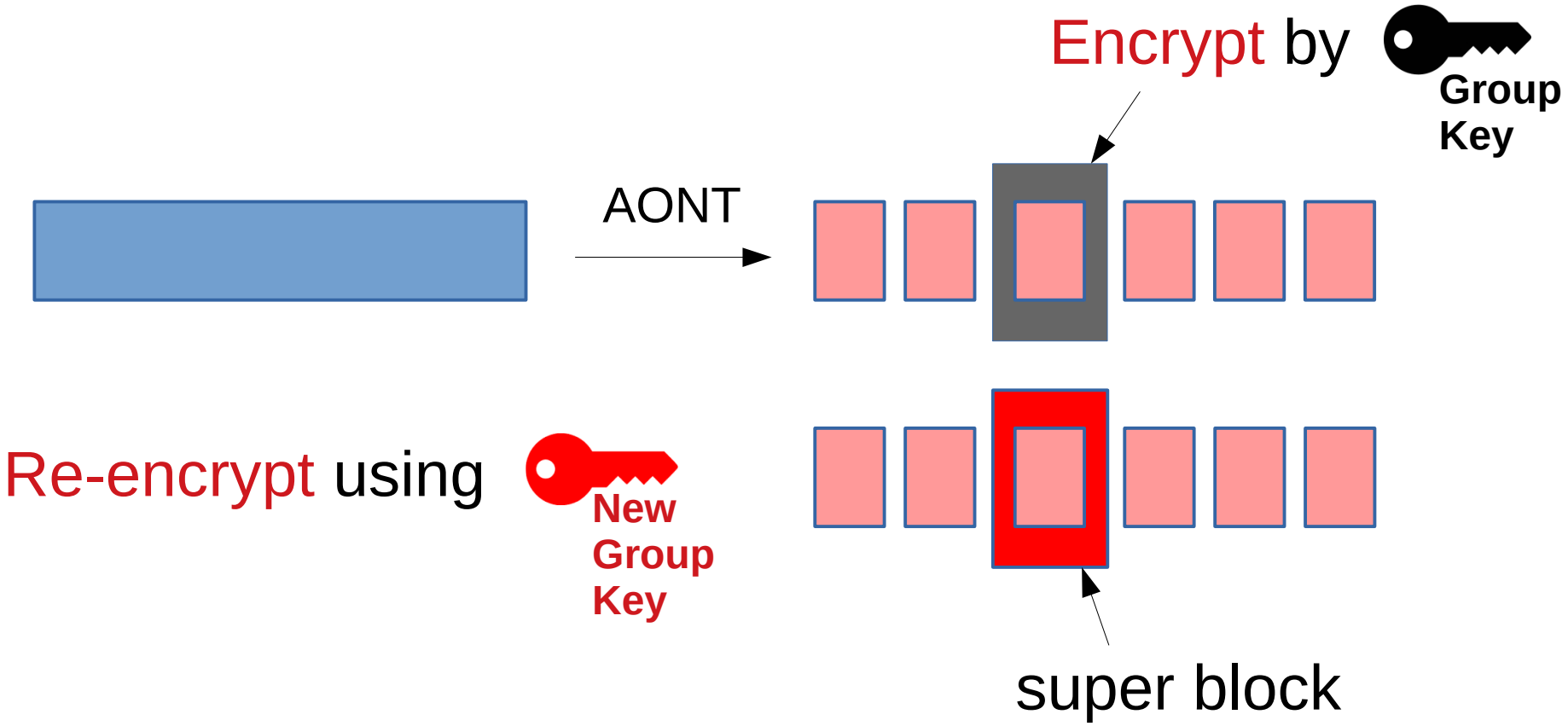
Moving all data to SGX is costly.

Transform the data *s.t.* only parts are re-encrypted.

All Or Nothing Transform

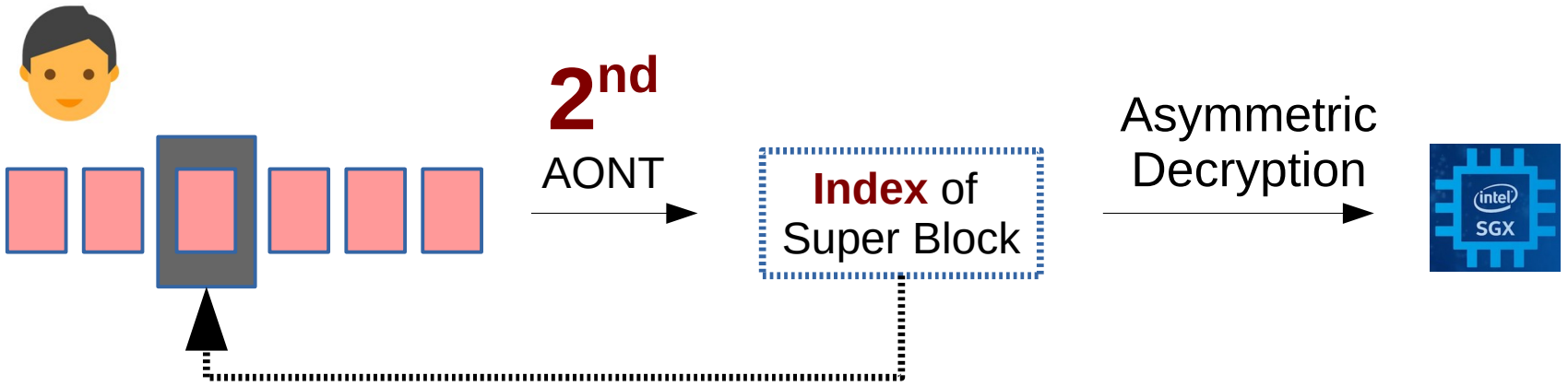


AONT and Super Blocks

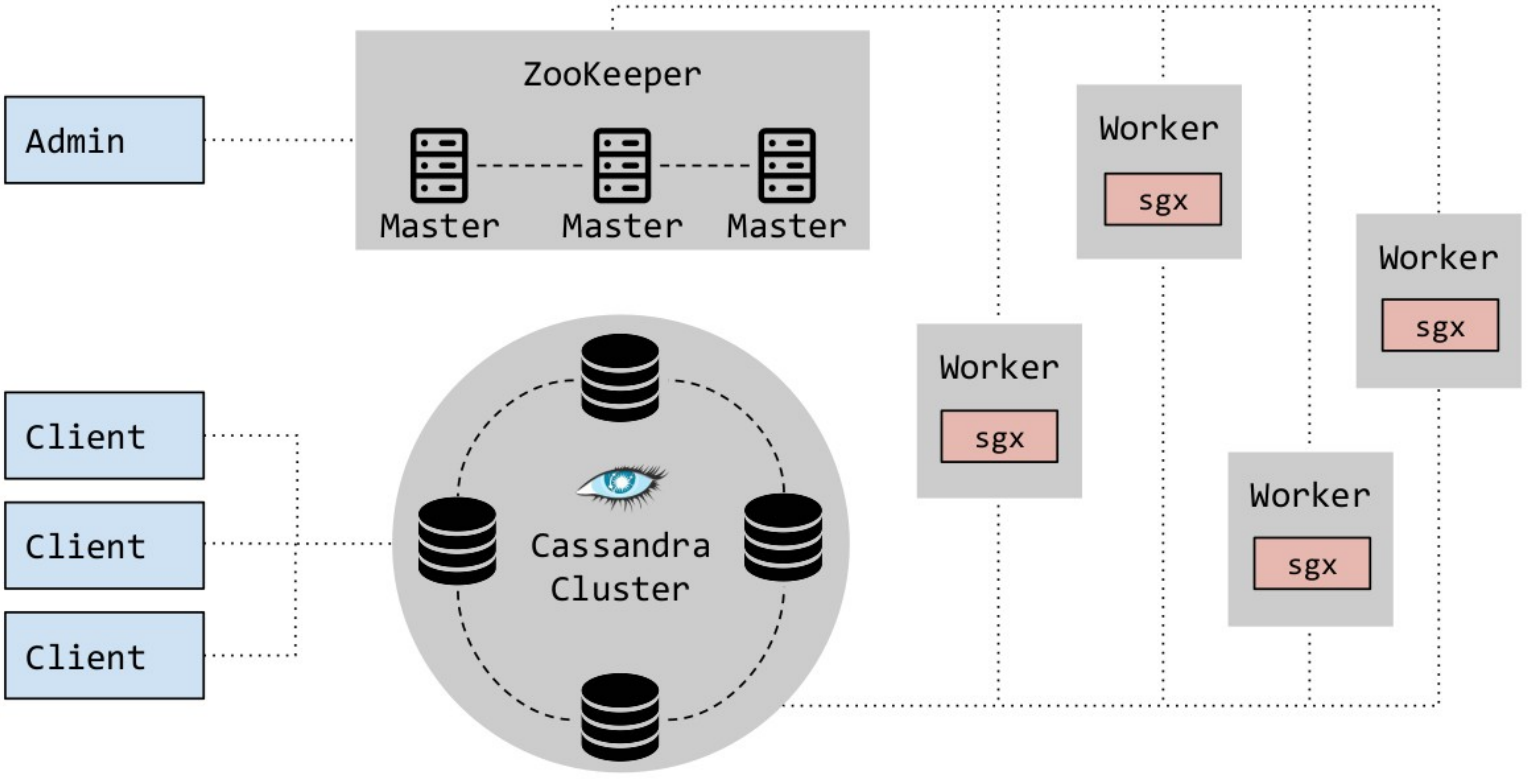


Avoid users getting super blocks

Give index of super block to user,
only if the whole file is downloaded :



R-SKY Implementation



Revocation Benchmark

Images Count	Total Size	Full Re-encryption	R-SKY
1,000	450 GB	42.2 m	7.5 s
10,000	4.3 TB	6.7 h	59 s
100,000	43.3 TB	2.8 d	11.1 m

Future Work

- Asymmetric integration of TEE :
  - ABE, Group Singatures, ZK-Proofs.
- Traceability of access control.
- Decentralized administrative decisions.
- Exploring strengthened threat models.

Industry Transfer

- Research → Industry.
- Chief Operating Officer (COO).
- Parsec v.2 :
  - Large organization deployments.
  - Adoption of TEE.

Conclusion

Context :

E2EE : data is encrypted before stored on cloud.

Group Access Control is performed cryptographically.

Problem :

Group access control is inefficient at large scale.

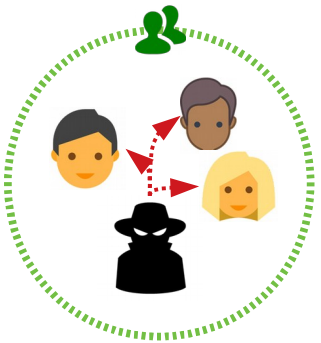
Confidentiality of **Large** Groups

IBBE-SGX

1.2 OoM faster

3 OoM less storage

Anonymity inside **Large** Groups



A-SKY

3 OoM faster

Revocation of **Large** Data

R-SKY

11 min vs. 3 days