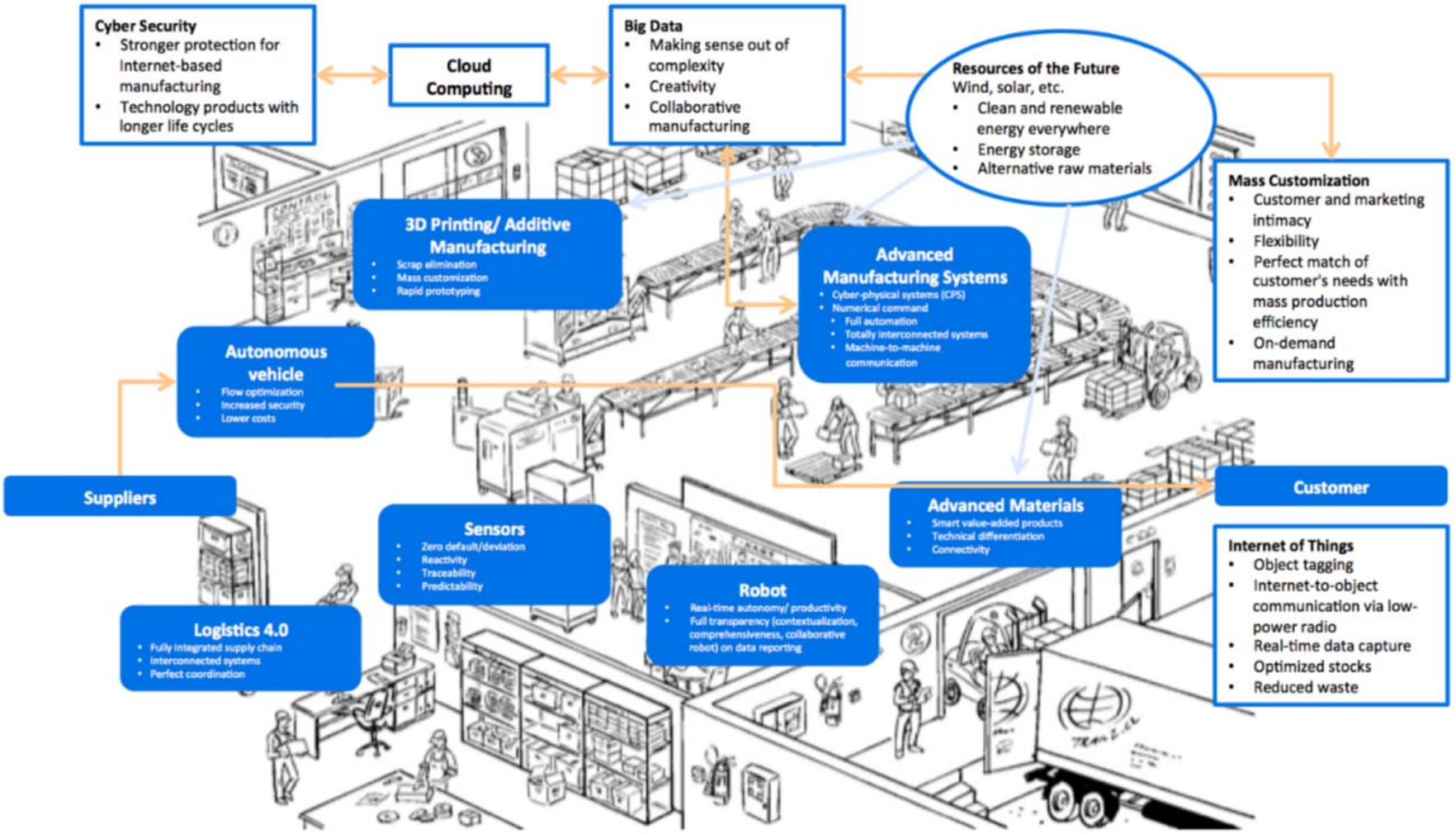
Smart Factories and Secure Cloud Storage Services: Vulnerabilities and Countermeasures

Chia-Mu Yu
Yuan Ze University

IFIP WG 10.4 Meeting
25 Jun 2016

Smart Factory

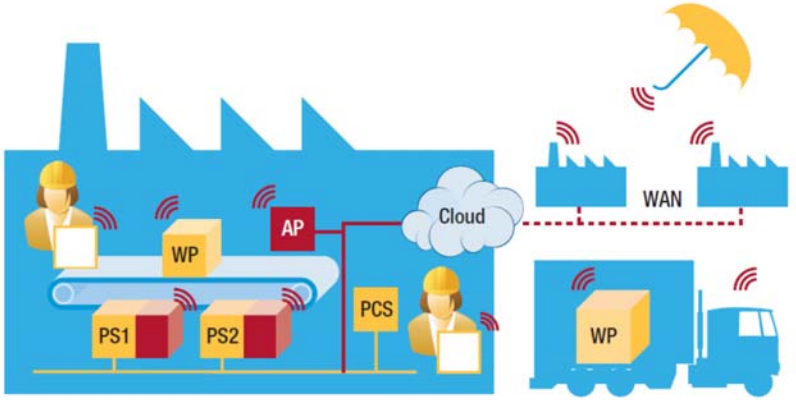


https://www.dcvision.cz/industry40/?___store=en&___from_store=cs

Two Cloud Security
One IoT Security

will be covered in this presentation

Motivating Scenario



https://e2e.ti.com/blogs_/b/thinkinnovate/archive/2015/09/15/smart-factories-of-the-future-enabling-technologies-for-industry-4-0

Motivating Scenario

- Restriction
  - My school does not have enough storage
  - Factory owner does not want to release data
  - Factory owner does not have enough storage

- OK… let's try Dropbox (cloud storage)

- Factory owner: Snowden said it's unsafe…
  Factory owner: I want my data encrypted and efficiency

Cloud Storage Security

Ah… kind of motivated by smart factory

Cloud Storage Providers

Cloud Storage Providers

Google Drive

Plans:

15 GB	100 GB	1 TB
FREE!	$1.99/month	$9.99/month
Current plan	Choose	

Dropbox Basic

Free

2 GB of space

Great Space Race!

The Great Space Race has ended! You can see the final results below!

Global Leaderboard

	SCHOOL	NUMBER OF SPACE RACERS	TOTAL POINTS
1	National University of Singapore	20,532	45,090 points
2	National Taiwan University	16,645	40,292 points
3	Politecnico di Milano	14,425	33,841 points
4	Nanyang Technological University	14,983	33,731 points

Cloud Storage Providers

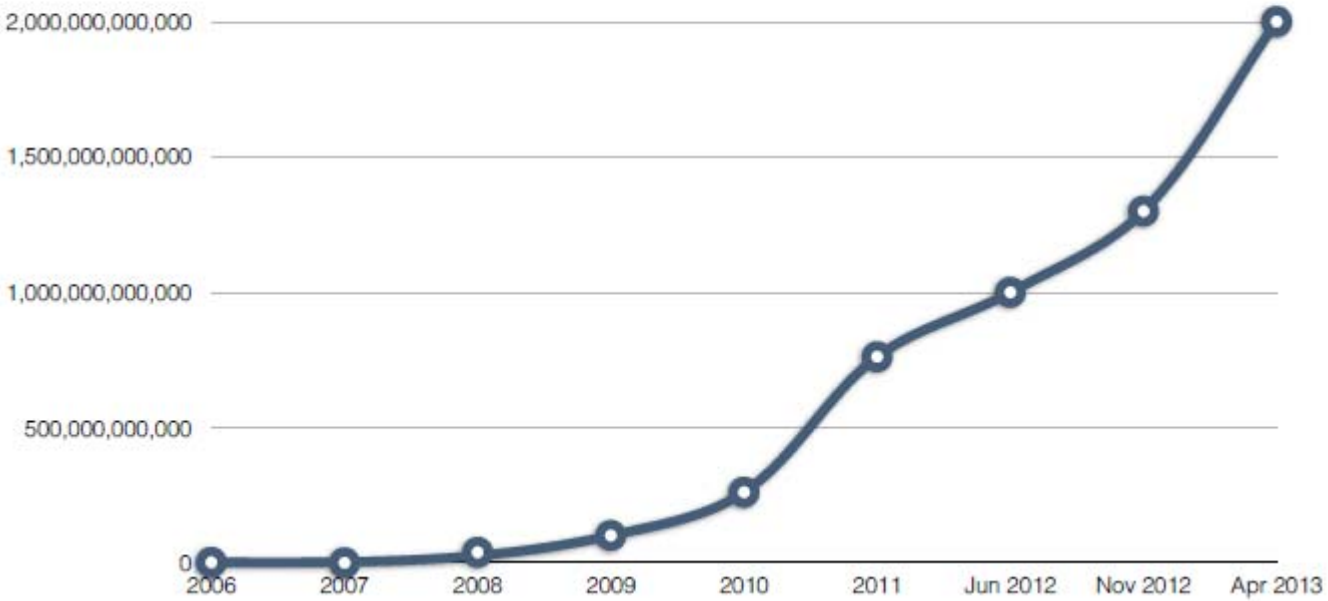
Individual Amazon S3 objects can range in size from **1 byte** to **5 terabytes**. The largest object that can be uploaded in a single PUT is **5 gigabytes**. For objects larger than **100 megabytes**, customers should consider using the Multipart Upload capability.

amazon web services / big data / cloud computing

Amazon S3 goes exponer
now stores 2 trillion objects

Amazon S3 FAQs - Amazon Web Services
aws.amazon.com/s3/faqs/

by Derrick Harris APR. 18, 2013 - 7:56 AM PDT

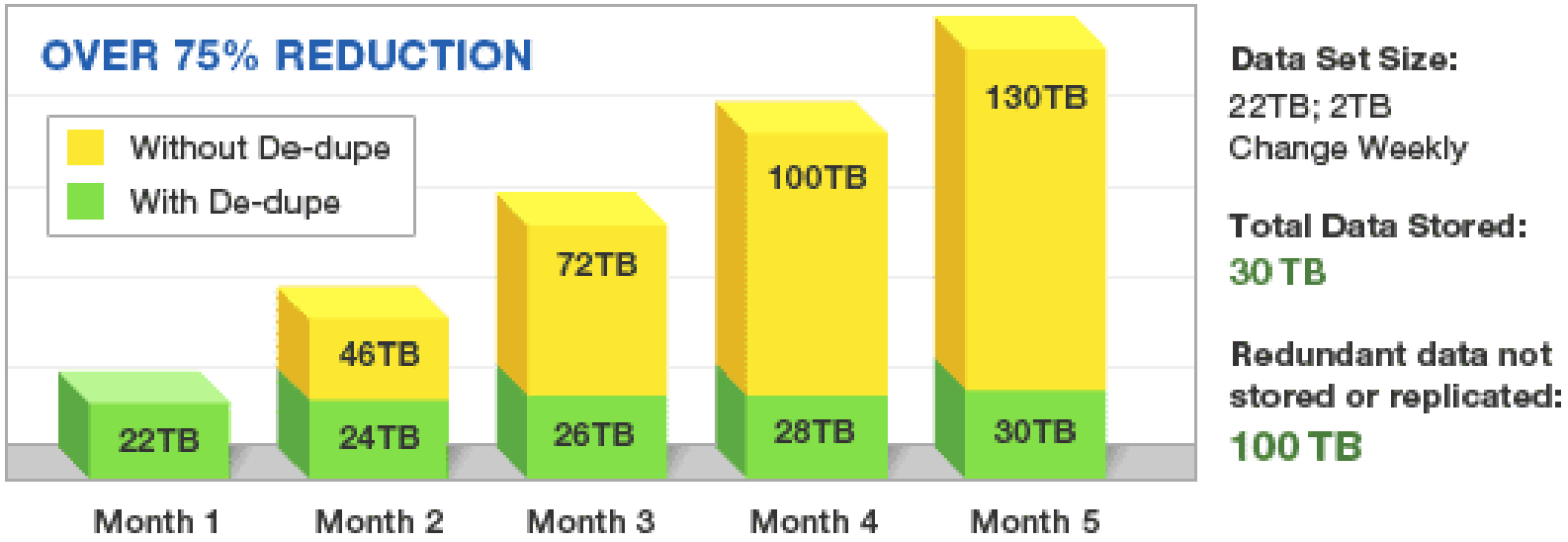


Data Deduplication

- People keeps uploading stuffs to cloud

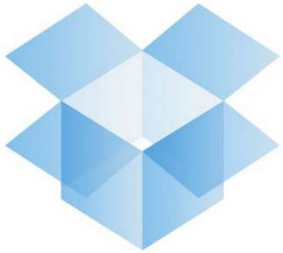
Data Deduplication

- Data **de***duplication*
  - A way of **avoiding storing the same file twice**



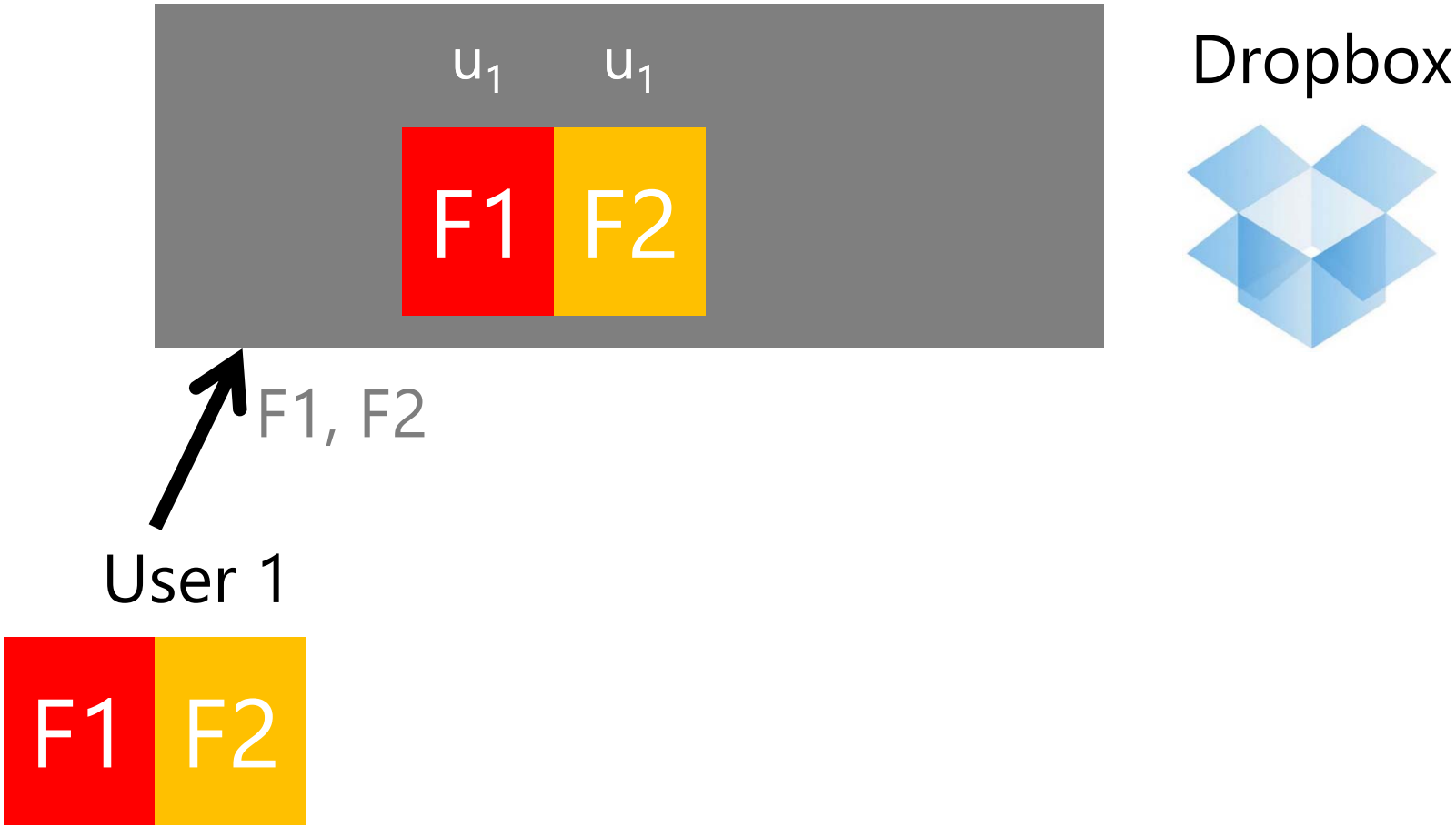
Cross-User Server-Side Data Deduplication

Dropbox

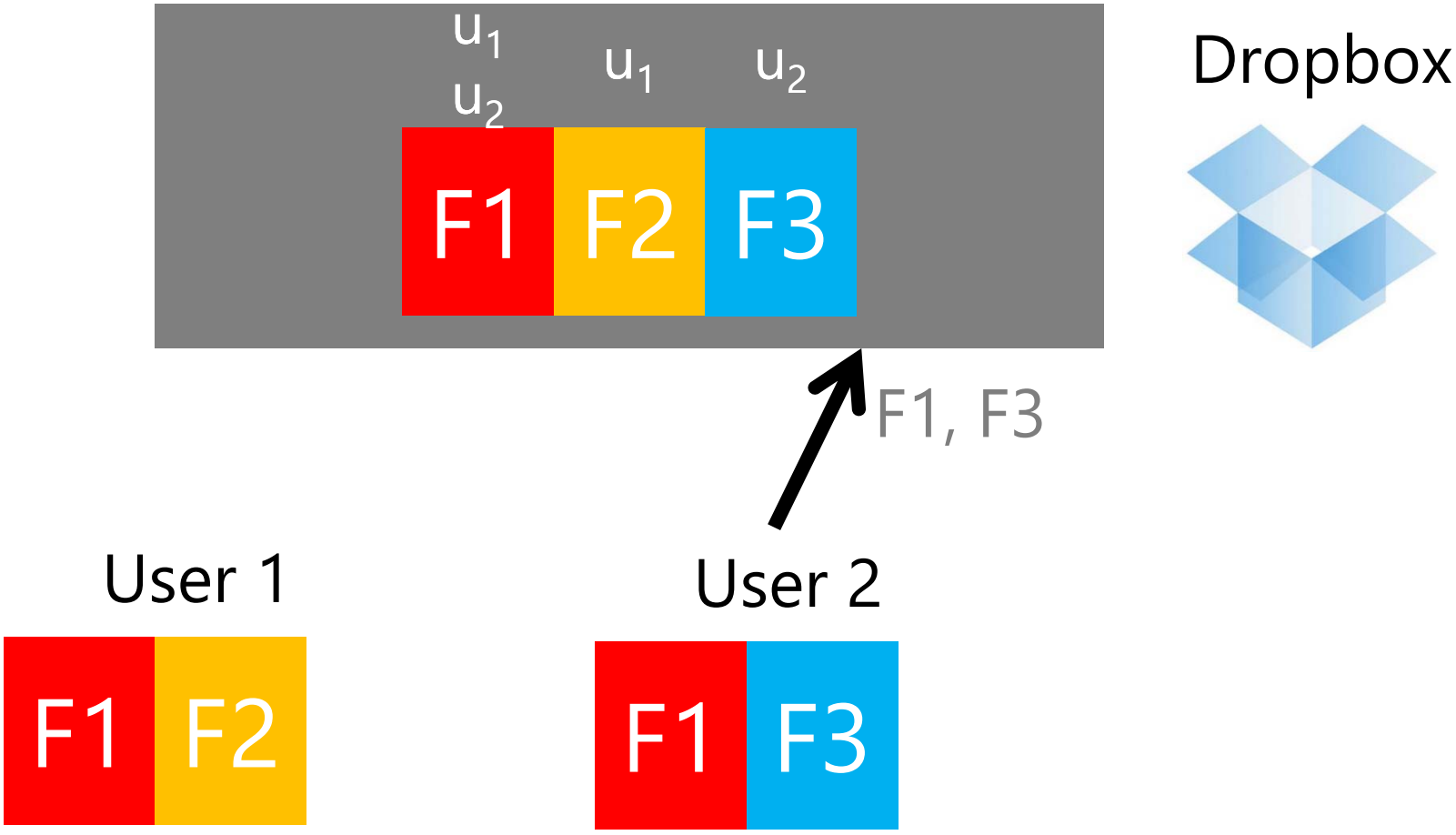


User 1

Cross-User Server-Side Data Deduplication

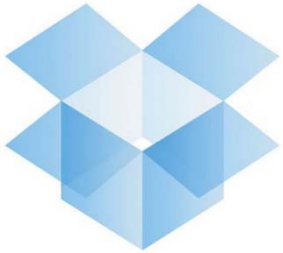


Cross-User Server-Side Data Deduplication



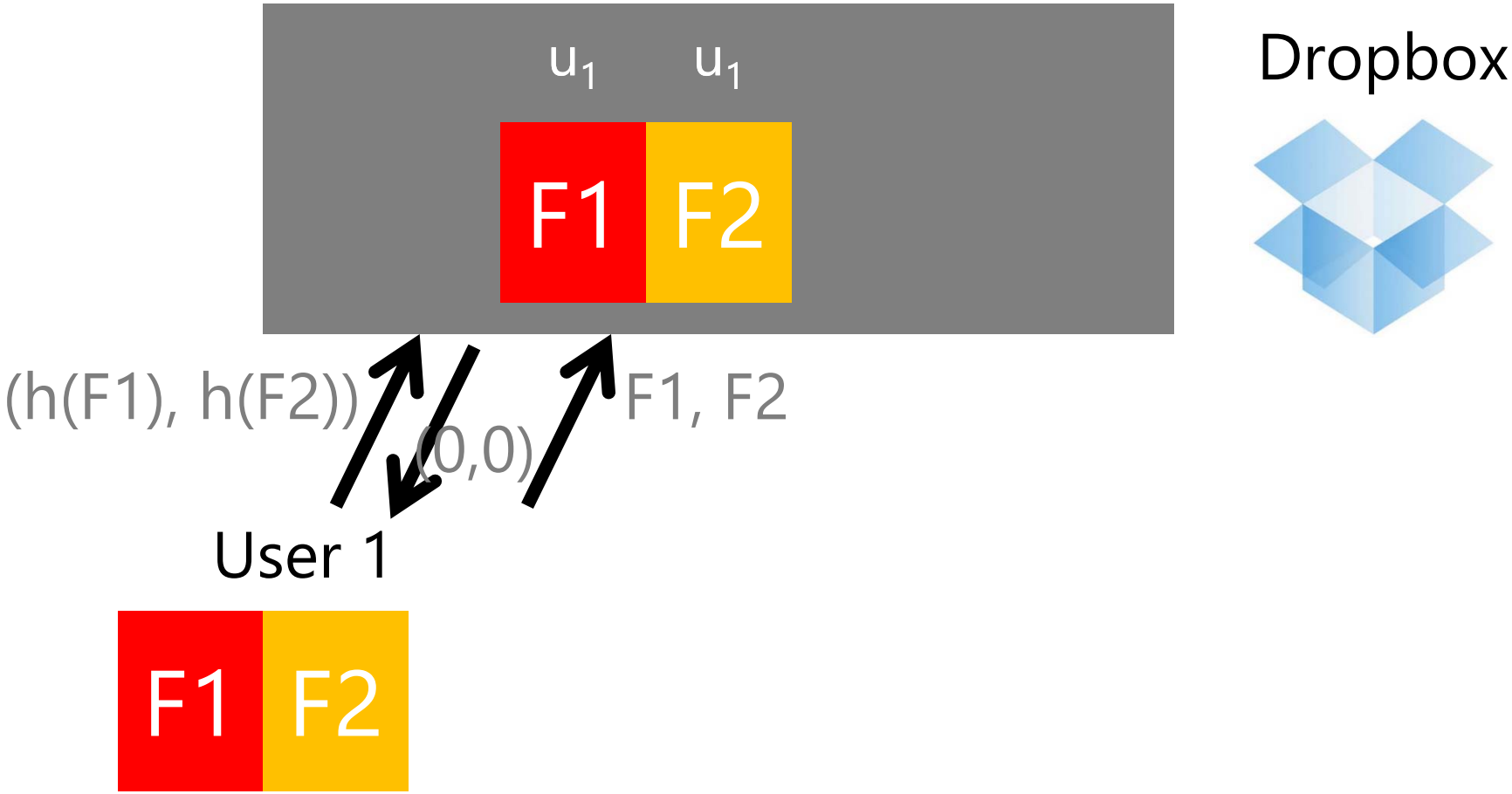
Cross-User Client-Side Data Deduplication

Dropbox

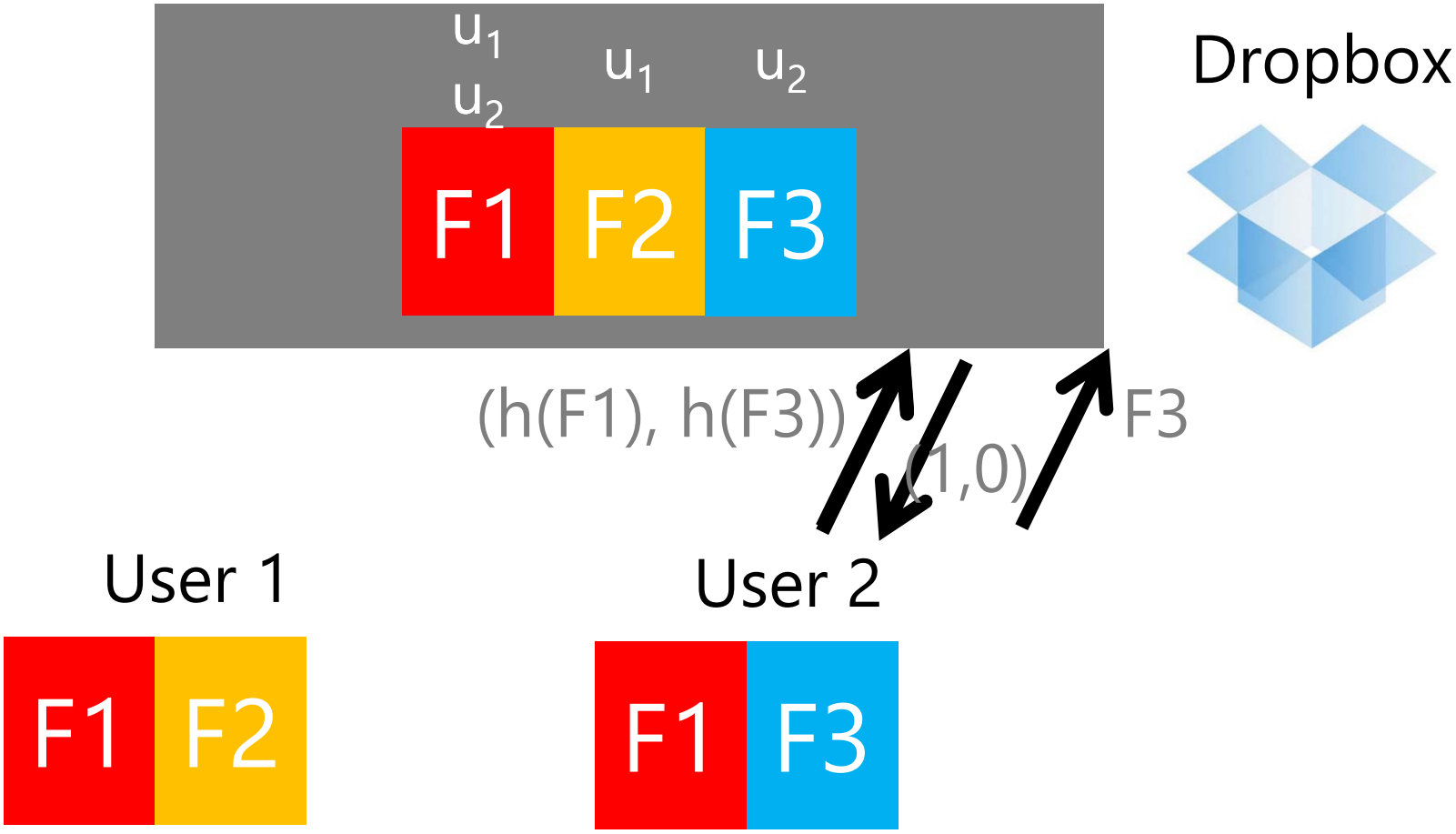


User 1

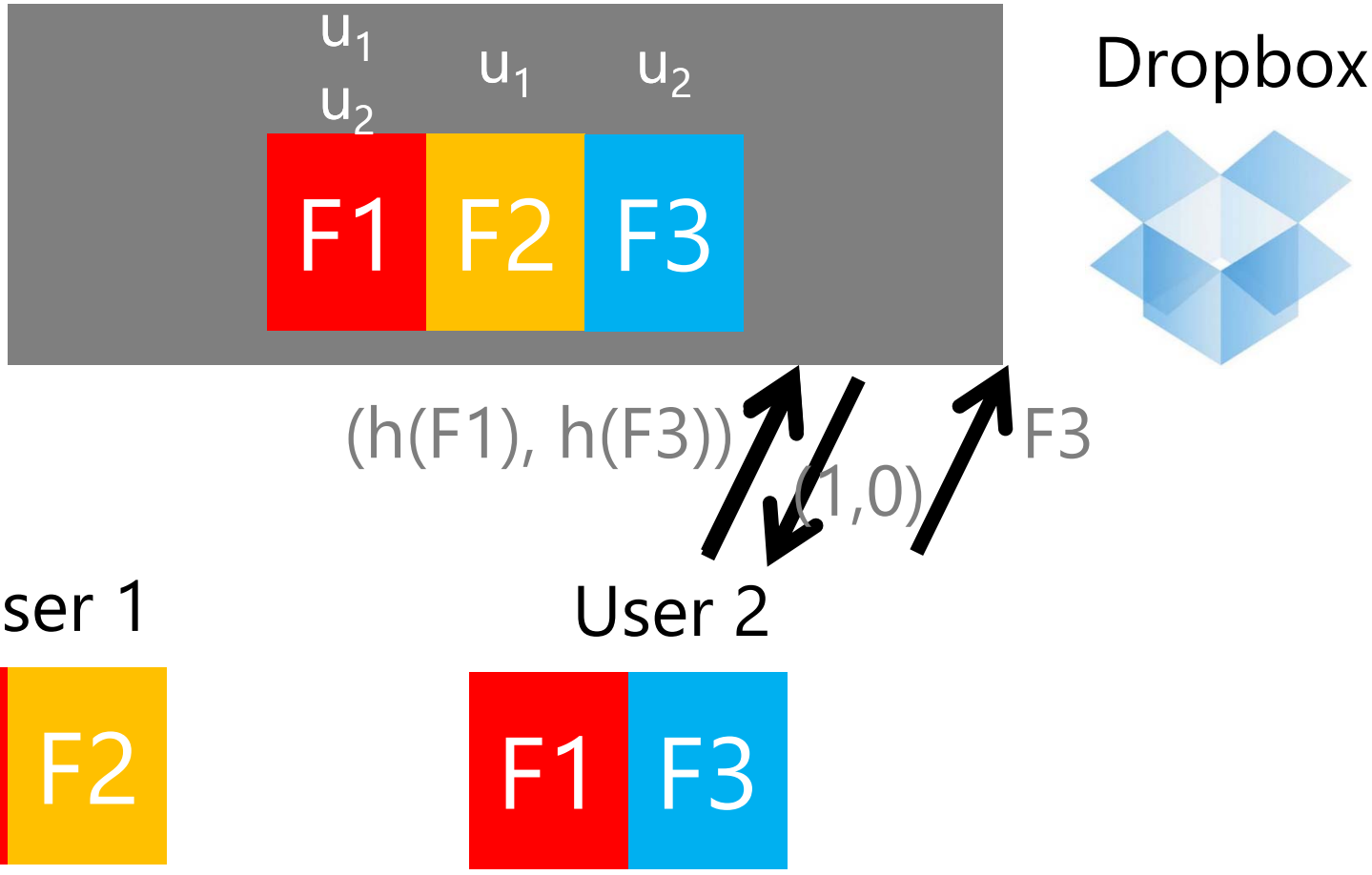
Cross-User Client-Side Data Deduplication



Cross-User Client-Side Data Deduplication



Cross-User Client-Side Data Deduplication

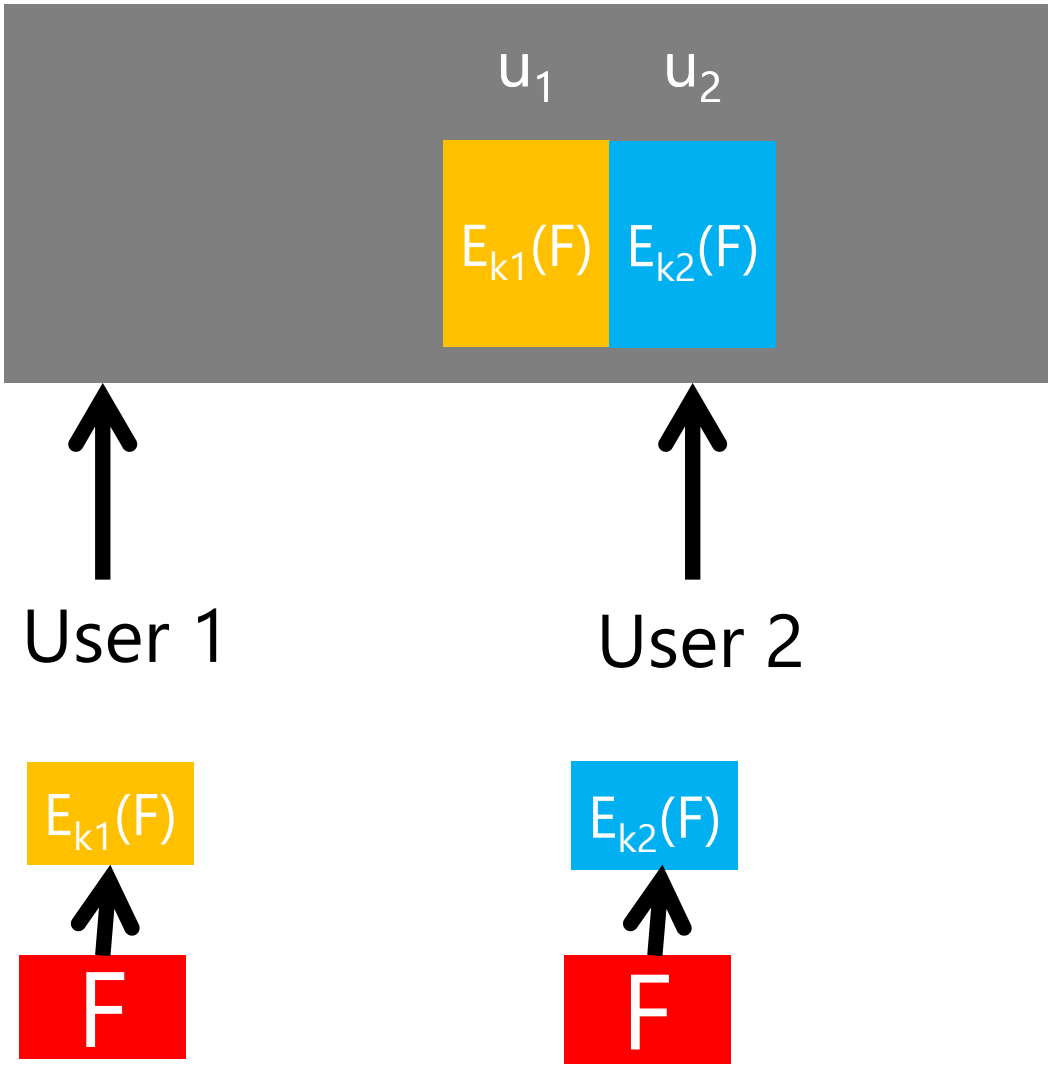


Save the transmission of one F1

Secure Deduplication

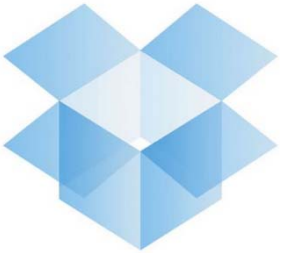
- **Data could be sensitive**
  - Data need to be encrypted before uploaded
  - However, totally destroys deduplication capability

Encryption Meets Deduplication



Convergent Encryption

Dropbox



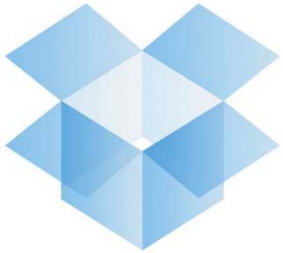
User 1

F

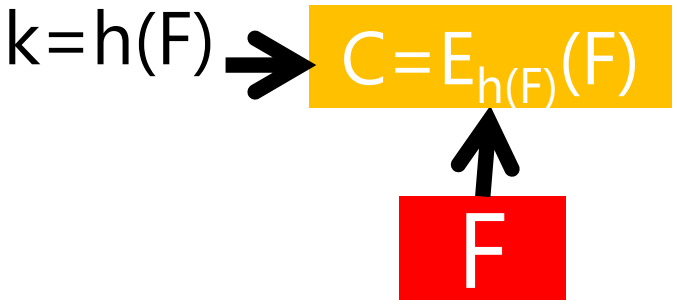
Reclaiming space from duplicate files in a serverless distributed file system, *IEEE ICDCS*, 2001

Convergent Encryption

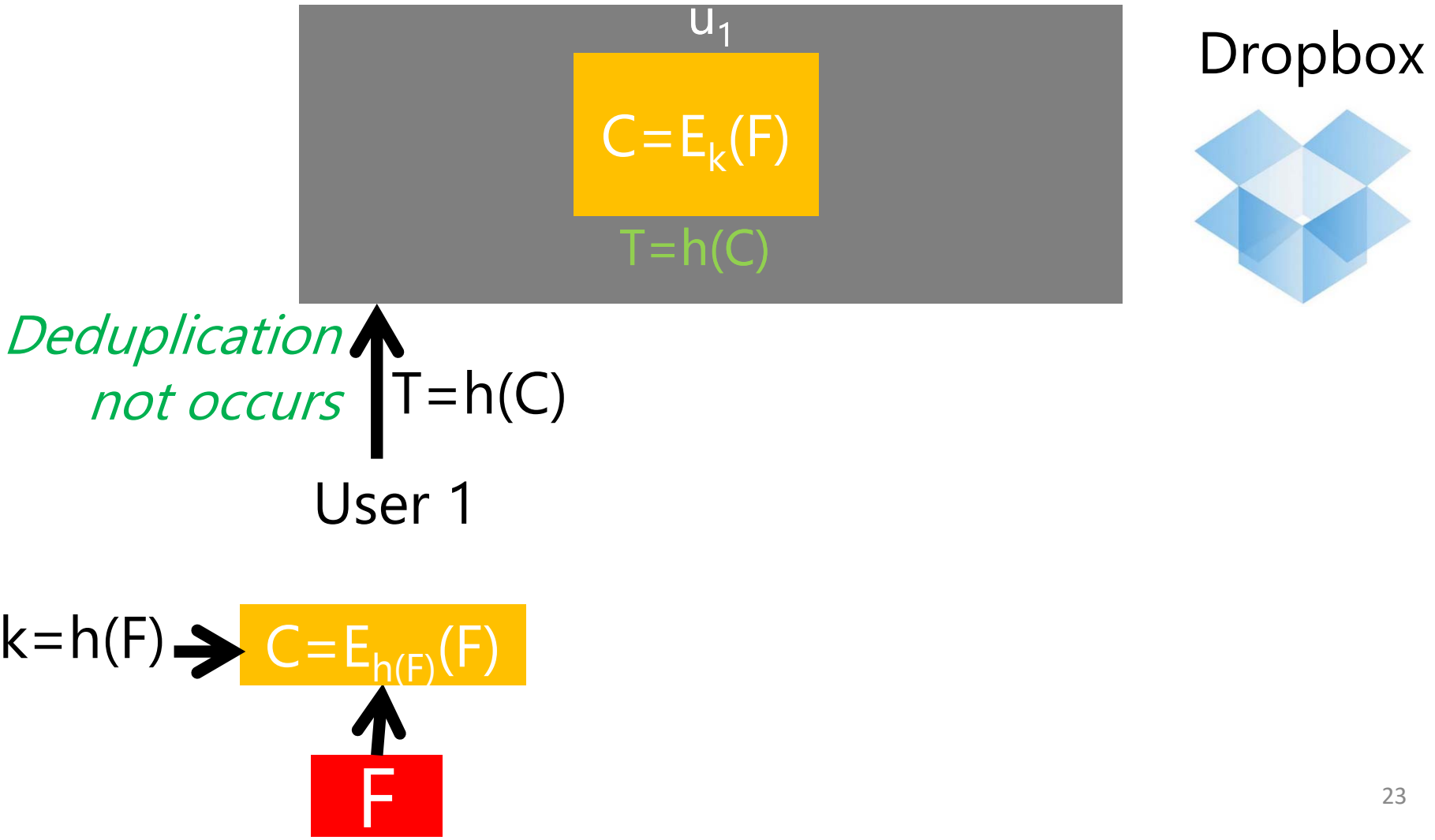
Dropbox



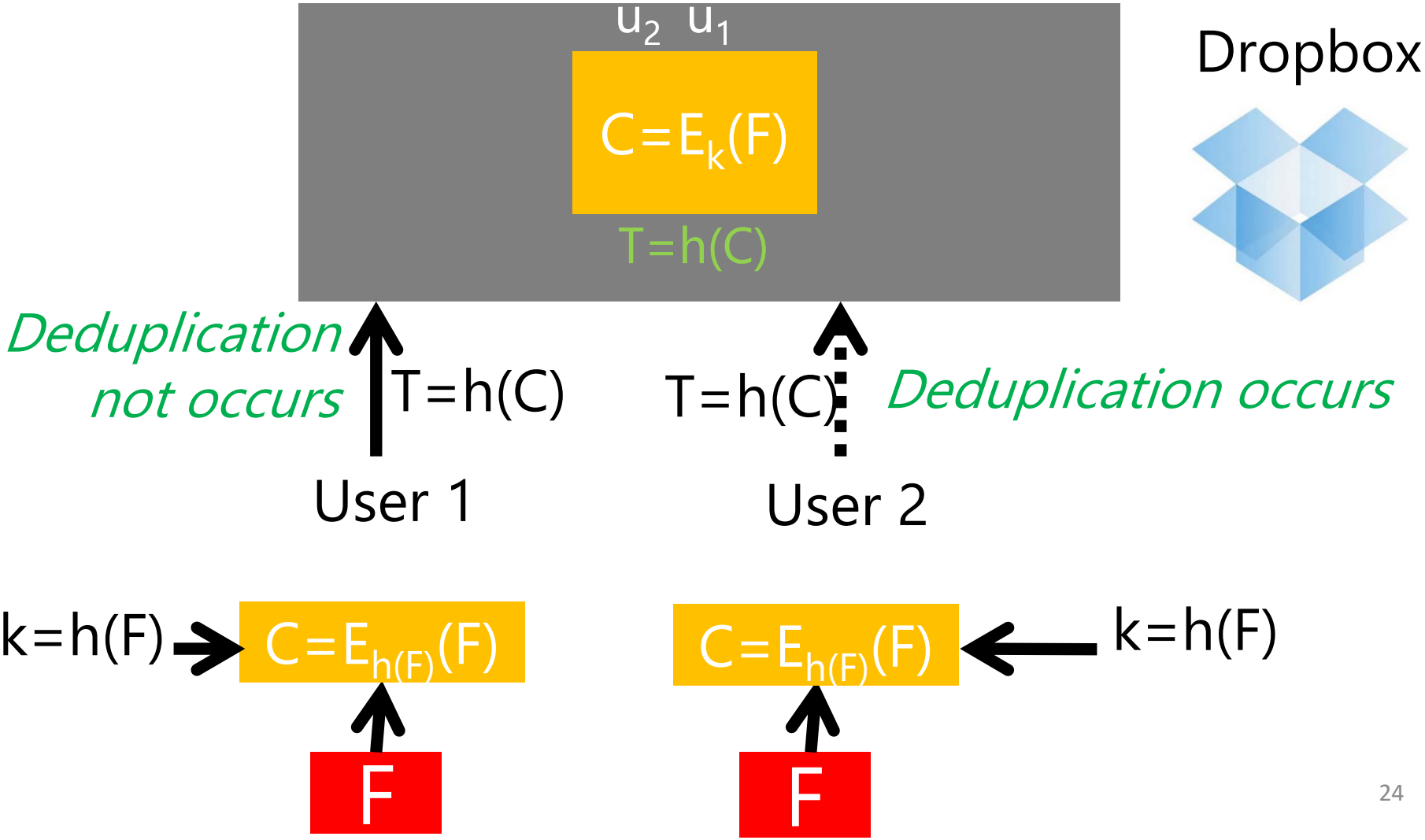
User 1



Convergent Encryption



Convergent Encryption



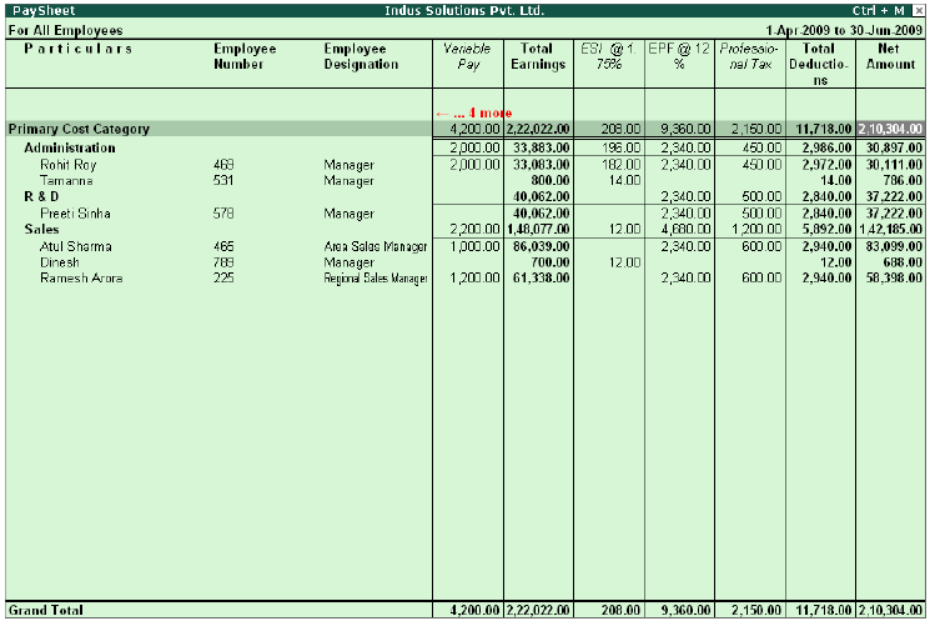
Try Every Possible Patterns!

- Convergent Encryption (CE)
  - Good for both data deduplication and privacy

- The weakness
  - **Brute force attack**

Weakness

- File predictability
  - In real life, file content is usually predictable
    - Pay sheet example
    - Chef's secret sauce
    - Engineer's parameter
    - etc



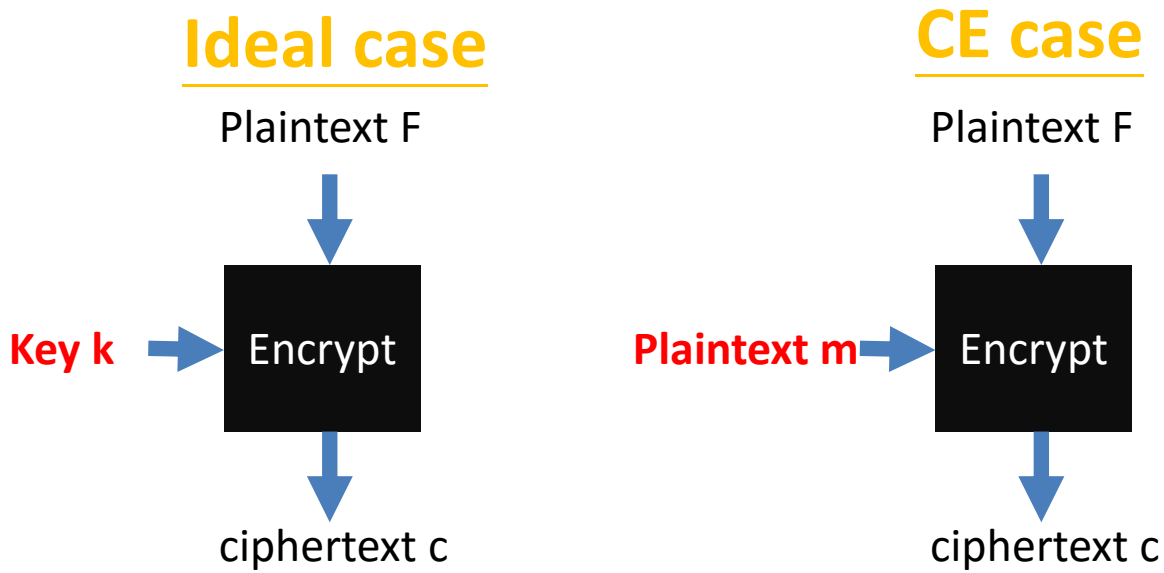
PaySheet — Indus Solutions Pvt. Ltd. — Ctrl + M

For All Employees — 1-Apr-2009 to 30-Jun-2009

Particulars	Employee Number	Employee Designation	Variable Pay	Total Earnings	ESI @ 1.75%	EPF @ 12 %	Professional Tax	Total Deductions	Net Amount
			← ... 1 more						
Primary Cost Category			4,200.00	2,22,022.00	208.00	9,360.00	2,150.00	11,718.00	2,10,304.00
Administration			2,000.00	33,883.00	196.00	2,340.00	450.00	2,986.00	30,897.00
Rohit Roy	468	Manager	2,000.00	33,083.00	182.00	2,340.00	450.00	2,972.00	30,111.00
Tamanna	531	Manager		800.00	14.00			14.00	786.00
R & D				40,062.00		2,340.00	500.00	2,840.00	37,222.00
Preeti Sinha	578	Manager		40,062.00		2,340.00	500.00	2,840.00	37,222.00
Sales			2,200.00	1,46,077.00	12.00	4,680.00	1,200.00	5,892.00	1,42,185.00
Atul Sharma	465	Area Sales Manager	1,000.00	86,039.00		2,340.00	600.00	2,940.00	83,099.00
Dinesh	789	Manager		700.00	12.00			12.00	688.00
Ramesh Arora	225	Regional Sales Manager	1,200.00	61,338.00		2,340.00	600.00	2,940.00	58,398.00
Grand Total			**4,200.00**	**2,22,022.00**	**208.00**	**9,360.00**	**2,150.00**	**11,718.00**	**2,10,304.00**

Weakness

- Brute force attack
  - MLE is weaker than conventional use of AES
    - Reason is that ***CE is keyless***



Our Requirements

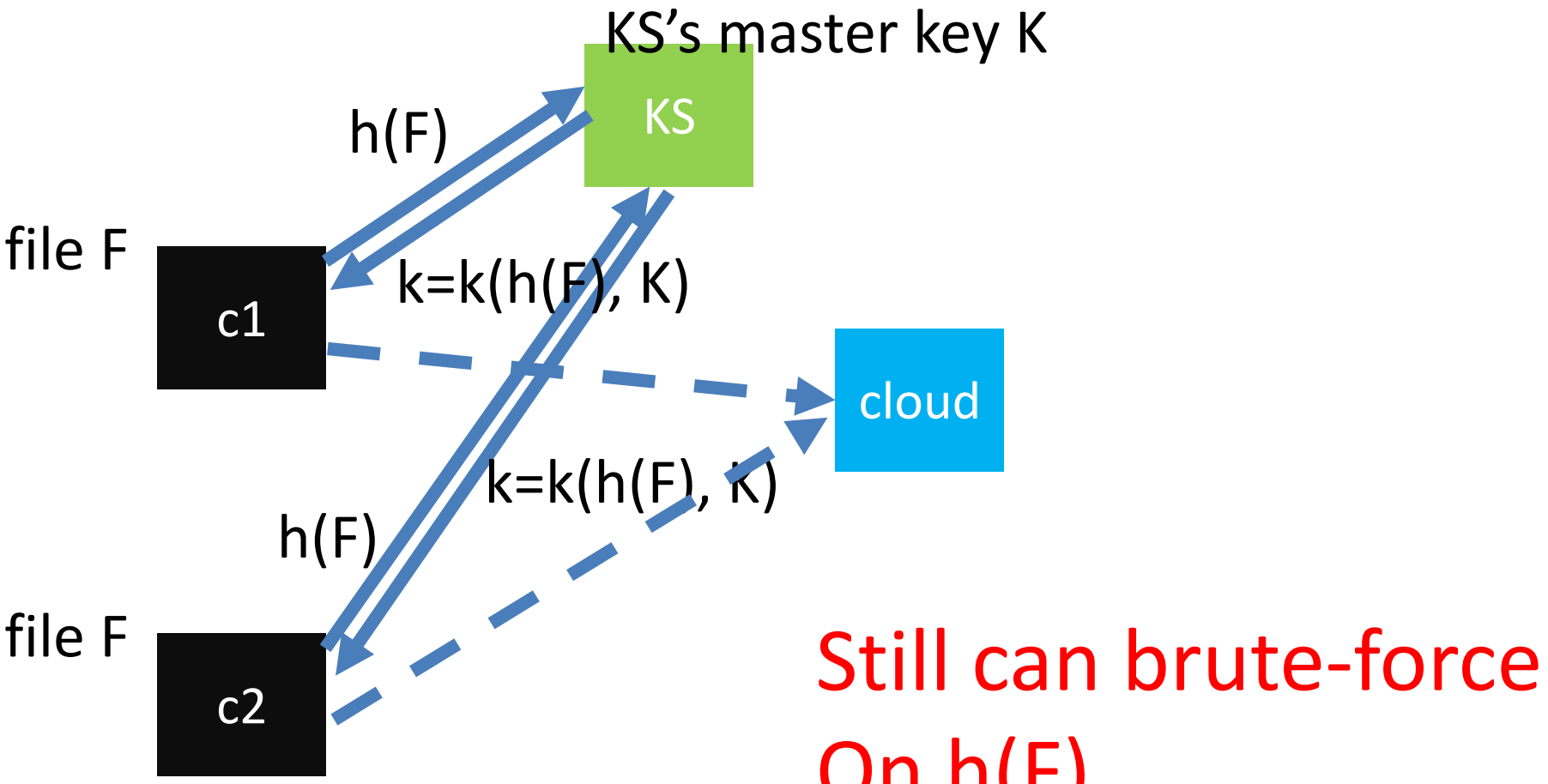
- Data deduplication
- Computation efficiency
- Brute-force resiliency

DupLESS

- How to overcome weakness?
  - A new secret
  - Idea is to deploy an additional **key server** (**KS**) that is responsible for generating keys for encryption purpose

DupLESS server-aided encryption for deduplicated storage, *USENIX Security*, 2013

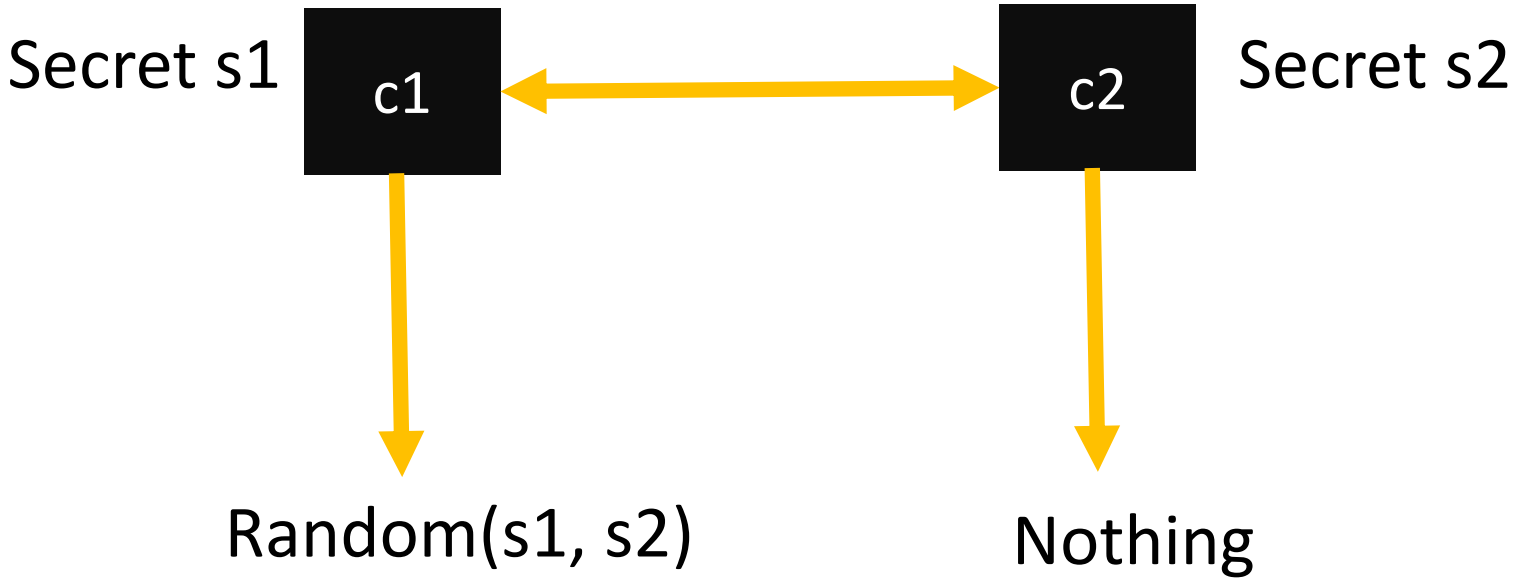
Naïve Implementation of DupLESS



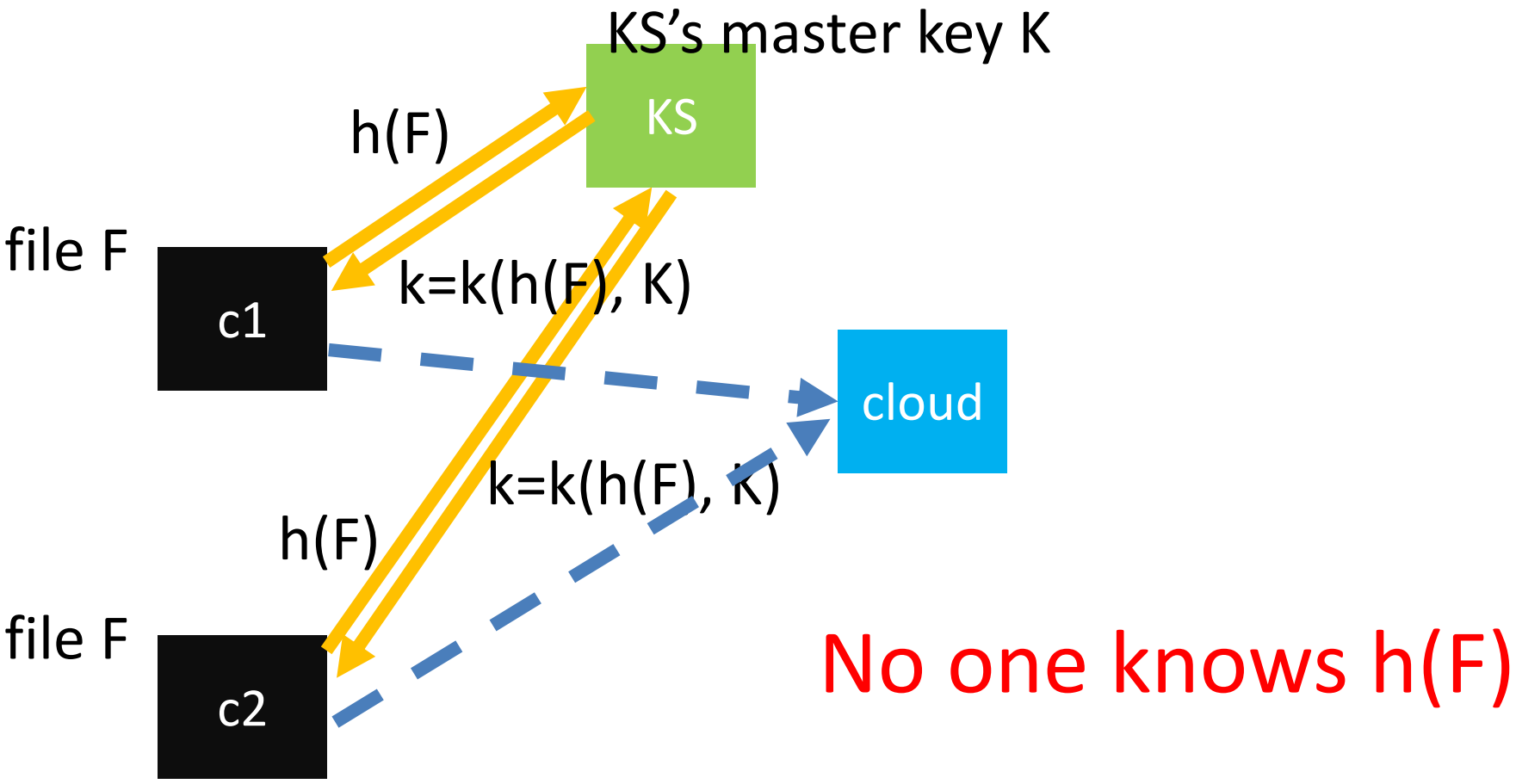
Still can brute-force
On h(F)

OPRF

- **Oblivious pseudorandom functions**
- **Kind of blind signature**

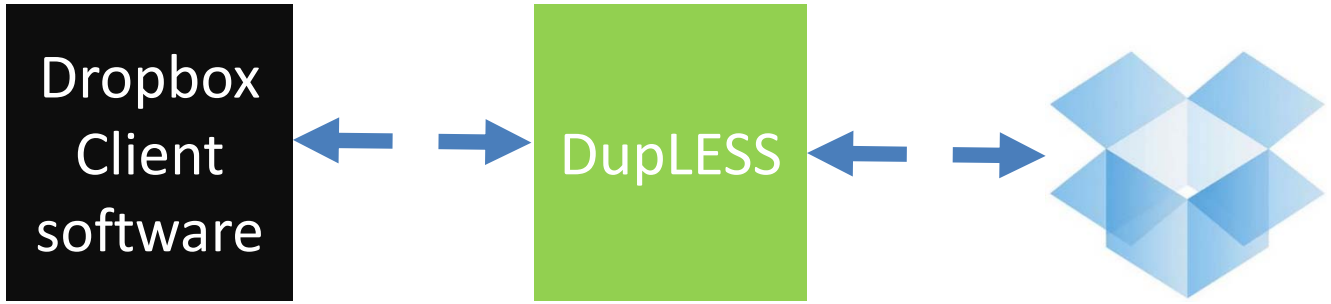


DupLESS


KS's master key K
KS
h(F)
file F
c1
k=k(h(F), K)
cloud
k=k(h(F), K)
h(F)
file F
c2
No one knows h(F)

DupLESS

- DupLESS does not need to modify cloud
  - It can be an additional software layer



- Factory owner: I'm happy

DupLESS

- DupLESS seems to have no weakness
  - No
  - **It has no practical use!**
  - **Who will be in charge of key server?**

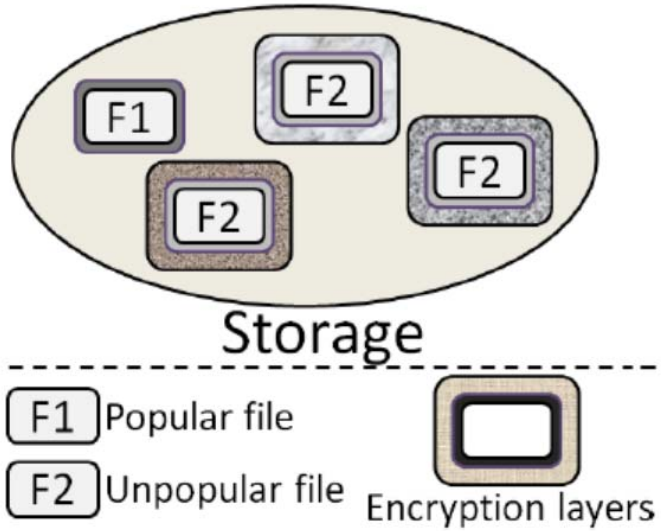
SecDep

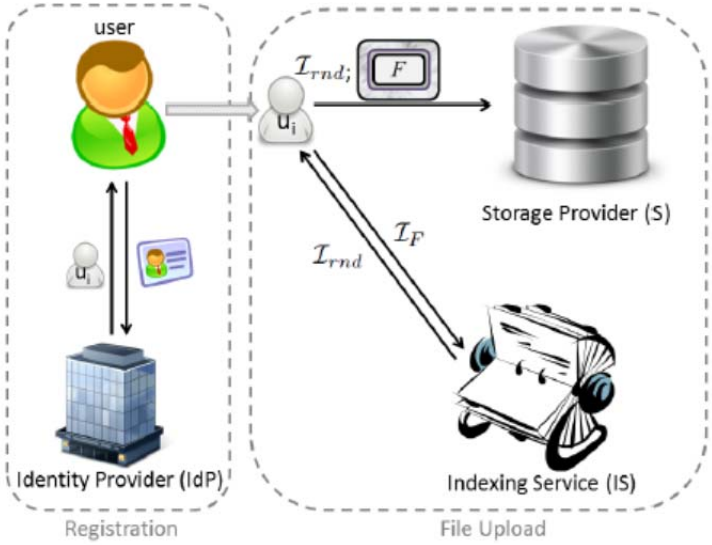
- DupLESS client always talks to KS, would inefficient in chunk level
  - Upload a file
    - Talk to KS in file level, to get file-level key and check dedup status in cloud
    - If not deduped, talk to KS again in chunk level, to get chunk-level key
- Maintains keys (file/chunk level) by client itself is cumbersome
  - Multiple KS
  - Distribute secret shares of key to KSs

Secdep: a user-aware efficient fine-grained secure deduplication scheme with multi-level key management, *MSST*, 2015

Threshold CE

- Dedup according to file popularity
- Each file is encrypted in two layers; the first is, the second is threshold CE





A secure data deduplication scheme for cloud storage, *FC*, 2014

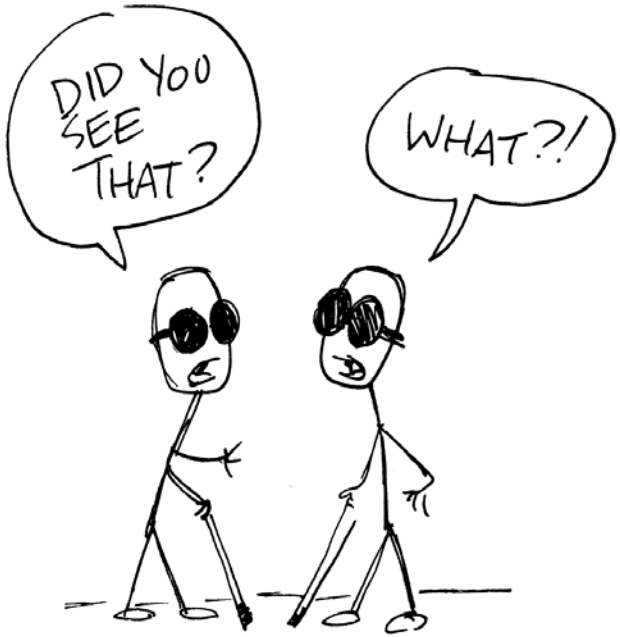
No KS Solution?

- Where the difficulty from?
  - Cannot send h(f)
    - Brute force attack for low-entropy file f
  - Cannot send E(f)
    - No bandwidth saving
  - Cannot communicate with additional trusted server and communicate via trusted channel
    - Awful assumption

Secure deduplication of encrypted data without additional independent servers, *ACM CCS*, 2015

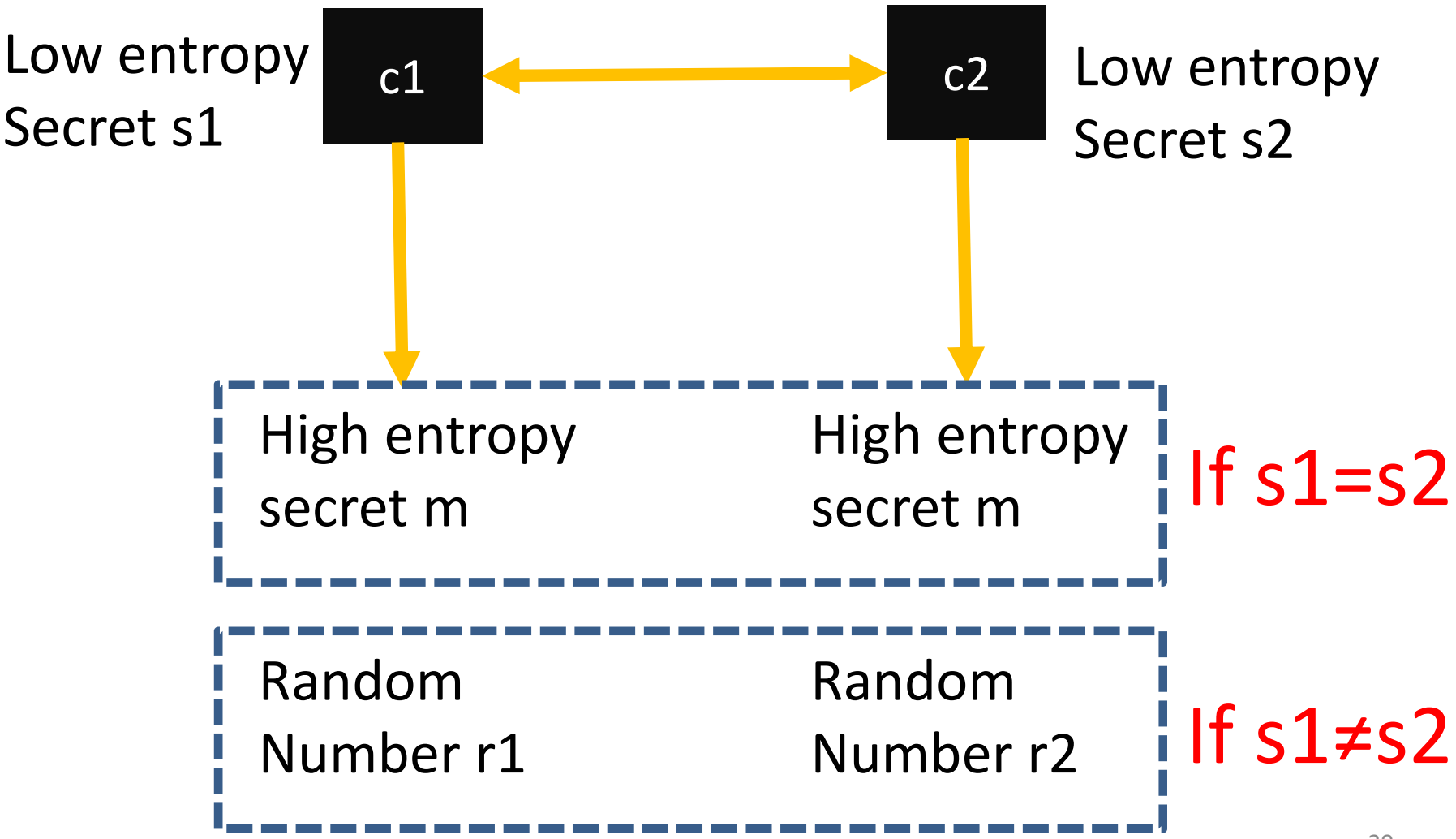
PAKE

- Password Authenticated Key Exchange
- Enable users to establish a common key based on their low entropy password only

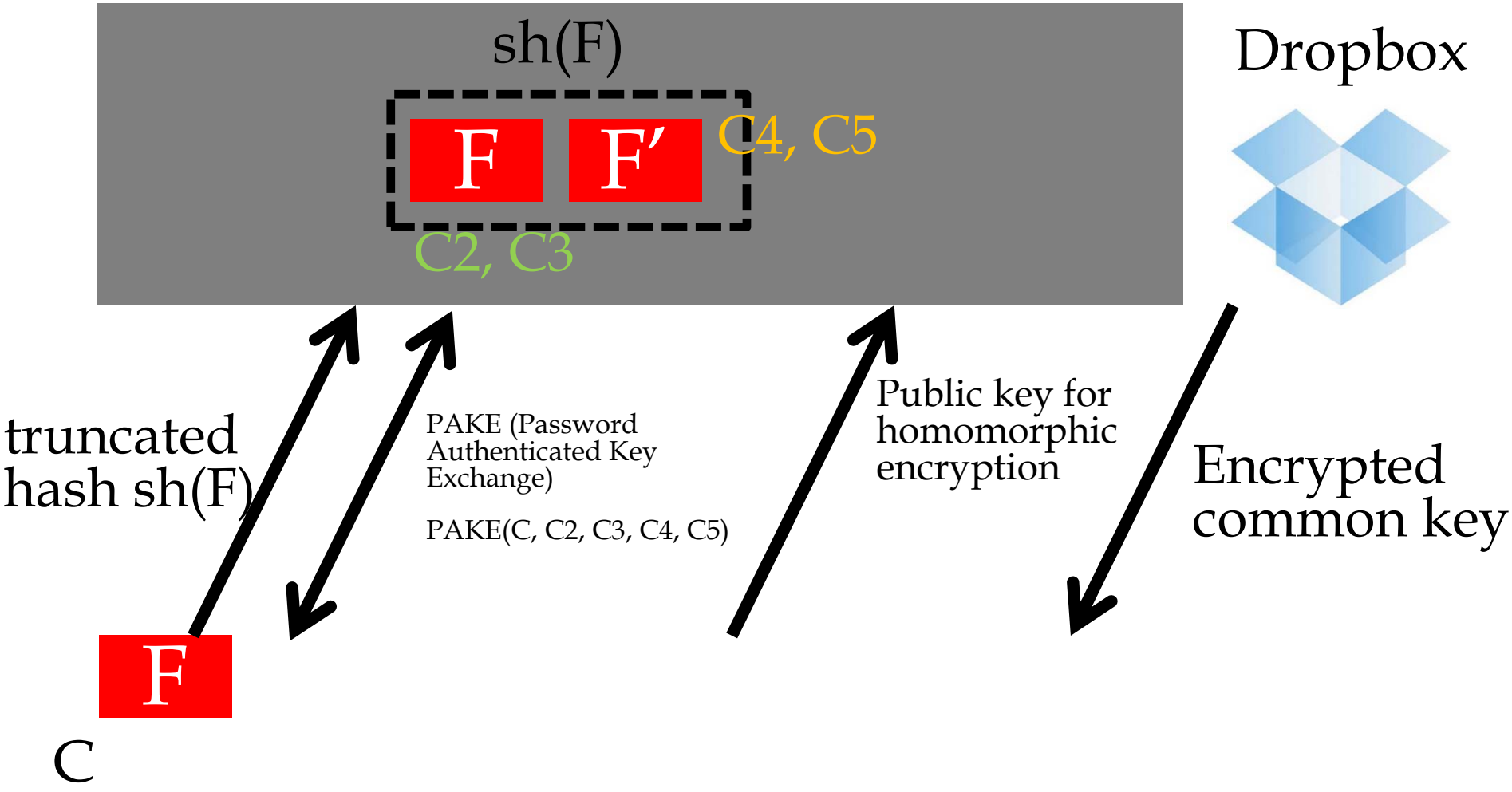


Encrypted key exchange: password-based protocols secure against dictionary attacks, *IEEE S&P*, 1992

PAKE


Low entropy
Secret s1
c1
c2
Low entropy
Secret s2
High entropy
secret m
High entropy
secret m
If s1=s2
Random
Number r1
Random
Number r2
If s1≠s2

PAKE-based Solution



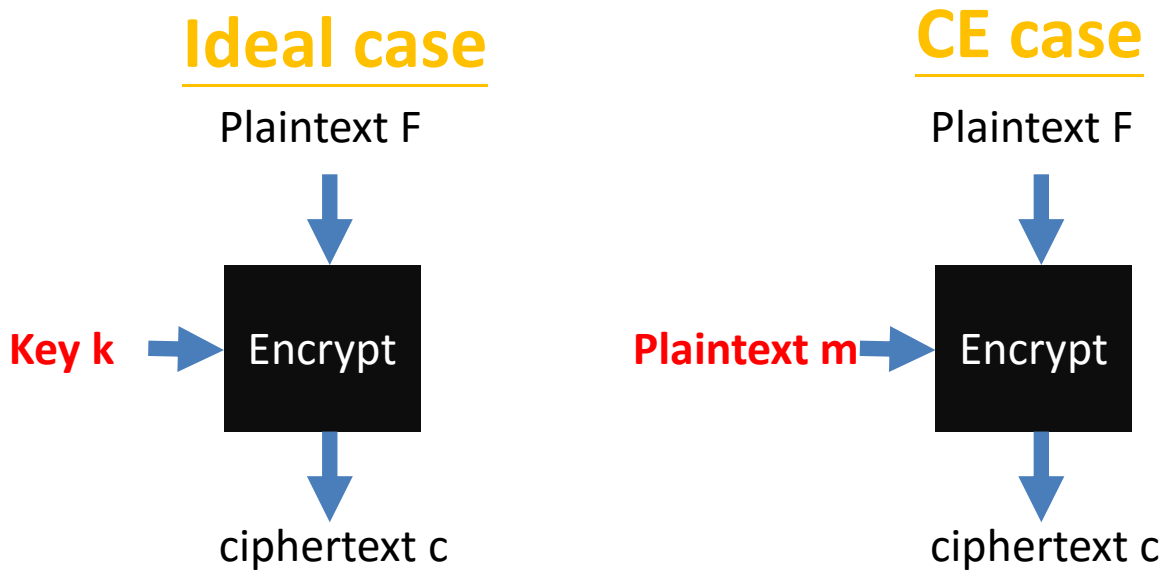
Secure deduplication of encrypted data without additional independent servers, *ACM CCS*, 2015

PAKE-based Solution

- Two heavyweight weapons
  - PAKE
  - Homomorphic encryption
- Have significant theoretical contribution but still no practical impact

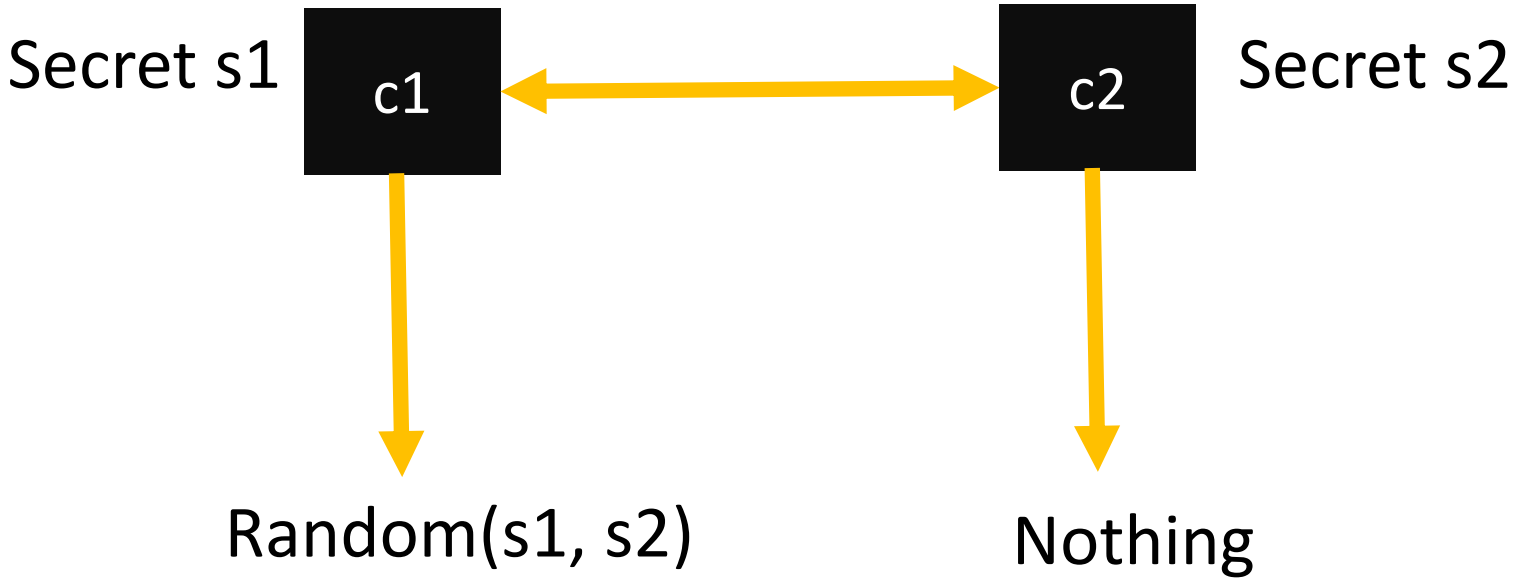
Rethinking PAKE-based Solution

- In fact, KS is still there; **everyone can be KS**
  - Essentially, we need an additional secret for brute-force attack

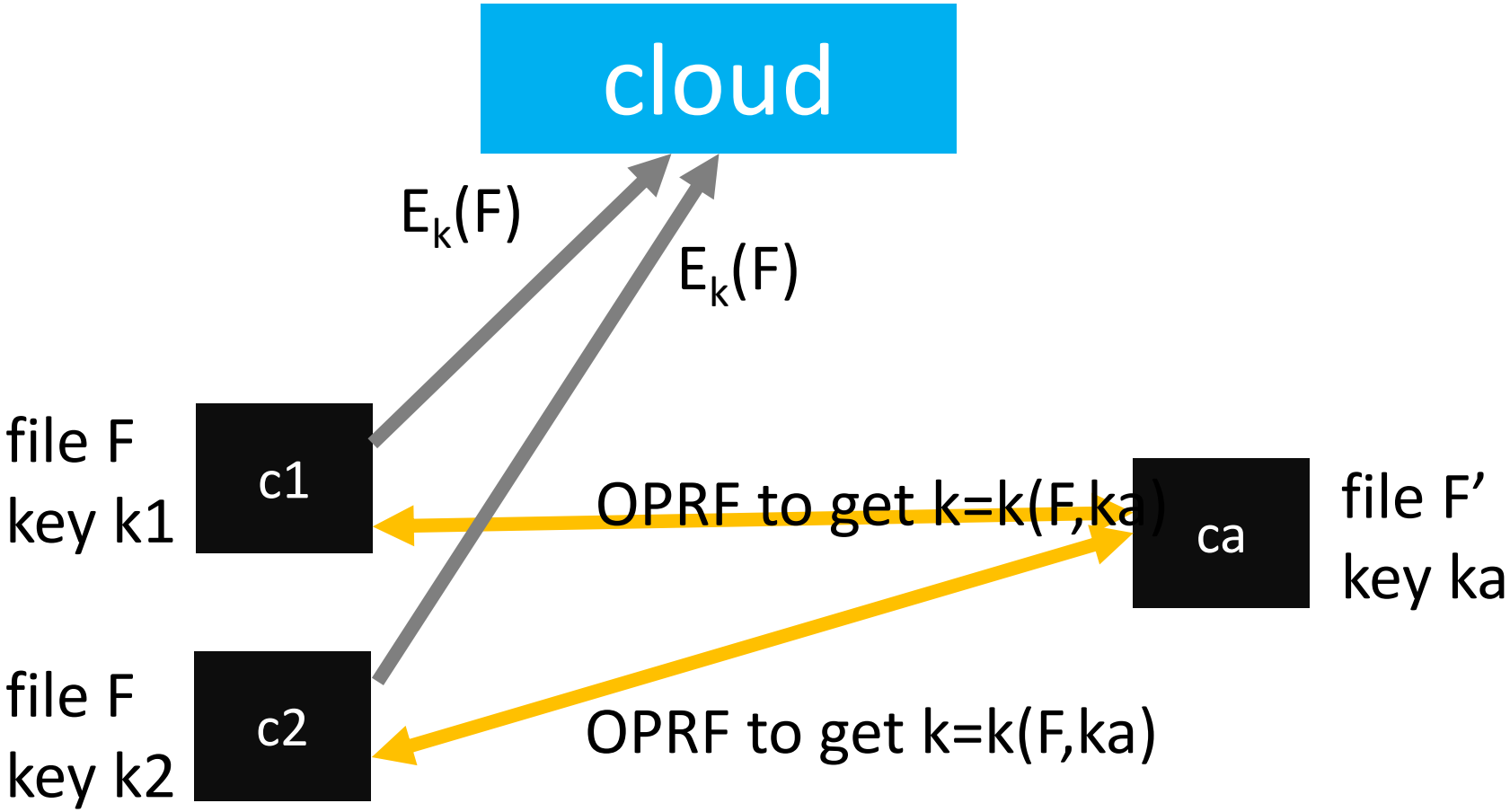


OPRF, again

- **Combine OPRF and the idea that everyone can be KS**

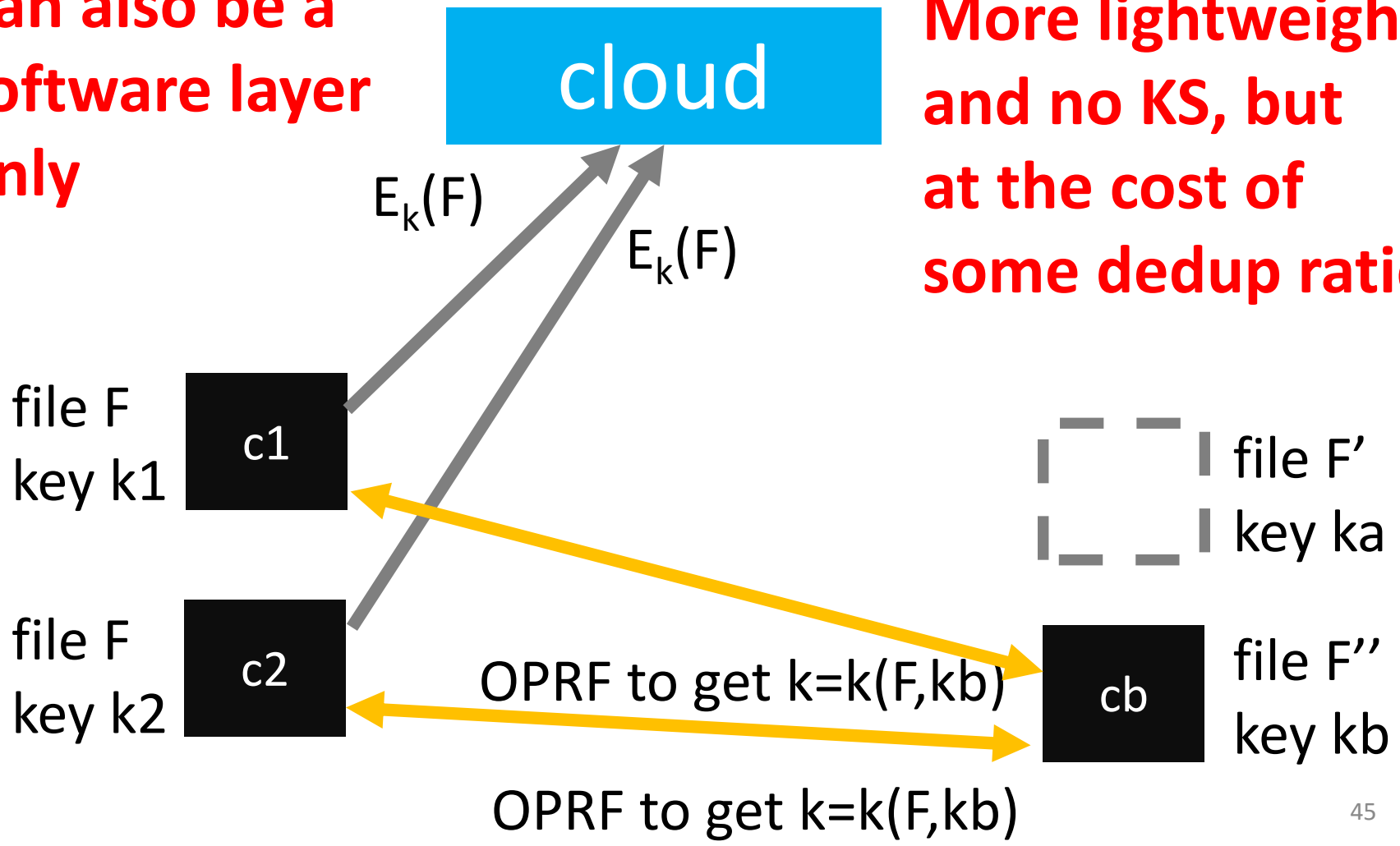


OPRF-based Solution



OPRF-based Solution

Can also be a software layer only

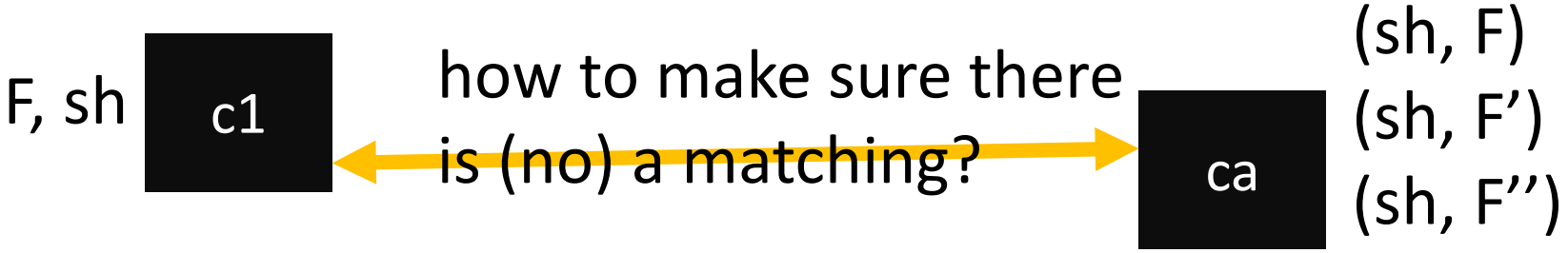


More lightweight and no KS, but at the cost of some dedup ratio

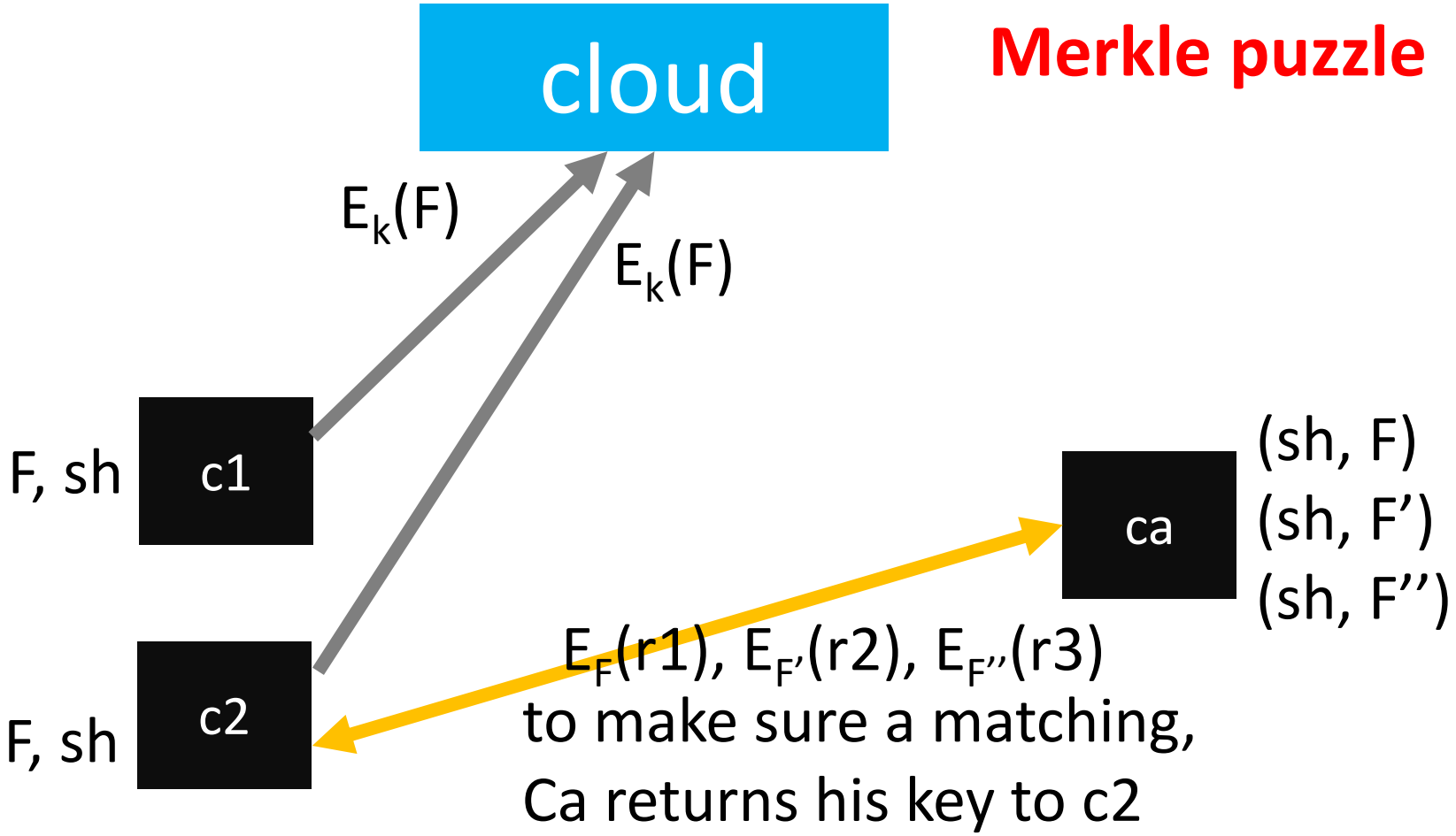
Symmetric Crypo-based Solution

- Should be the best in terms of performance

- Take another route, sh(F)
  - E(F), h(F) are not good, OPRF is heavyweight

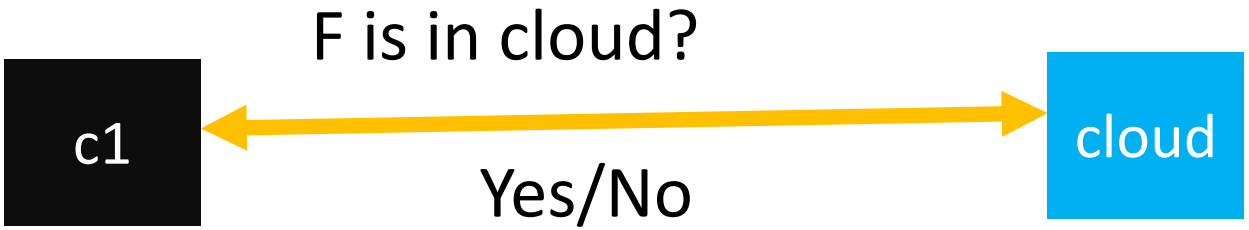


Symmetric Crypo-based Solution



Motivating Scenario

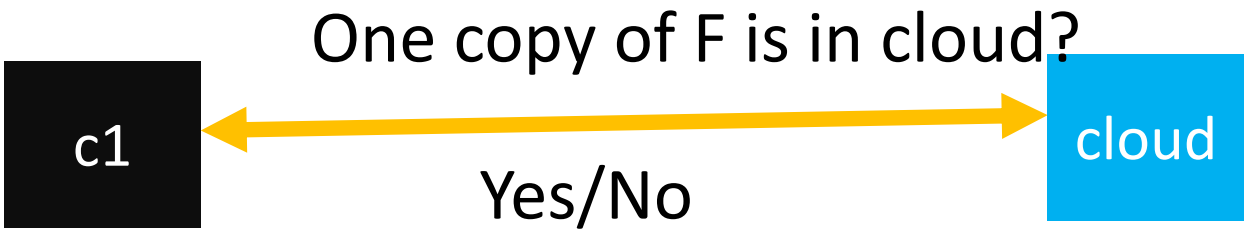
- Factory owner: dedup leaks my secret
- Me: why?
- Factory owner: cloud always returns dedup result!



Threshold = 1

- Original deduplication assumes threshold=1

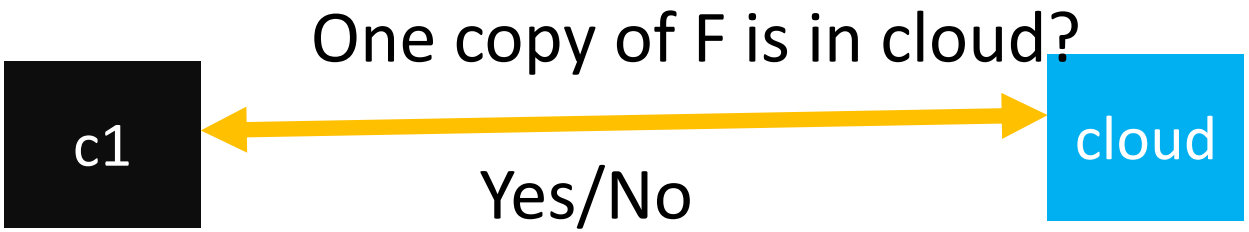
- Easy for attacker to know the file existence status



Random Threshold

- Each file x is associated with a random threshold tx

- tx too large, no dedup
  tx too small, no security



Side channels in cloud services, deduplication in cloud storage, Security and Privacy, 2012

Random Response

- First attempt: randomize the hash response

Chunk existence	Hash response
0	1
1	0/1

- 0-response indicates chunk existence

Random Response

- Second attempt: client uploads two chunks at once

Chunk 1	Chunk 2	response
0	0	2
0	1	1
1	0	1
1	1	1

- Upload **c1⊕c2** to cloud
- Seem to work?
- Fix a chunk not in cloud, infer chunk 2 existence

Random Response

- Third attempt: each result has a time limit

Chunk 1	Chunk 2	response
0	0	2 (t)
0	1	1 (t)
1	0	1 (t)
1	1	1 (t)

- Many accounts query cloud within a short time period

Random Response

- Fourth attempt: client cannot do the query but does not upload the chunk

Chunk 1	Chunk 2	response
0	0	2
0	1	1
1	0	1
1	1	1

- Many accounts will be used by attacker

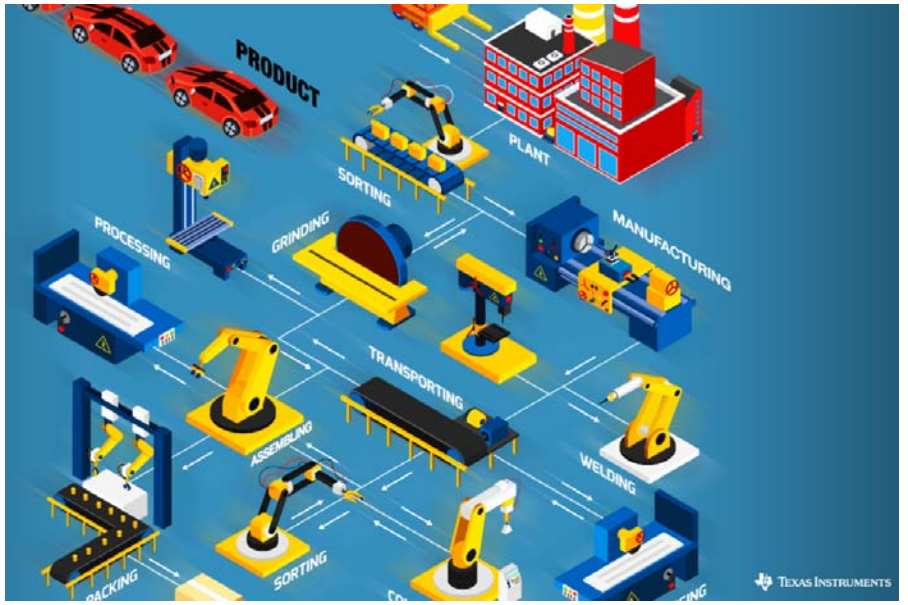
Random Response

- Observation: in any case, at least **c1⊕c2** needs to be uploaded

Chunk 1	Chunk 2	response
0	0	2
0	1	1
1	0	1
1	1	1

- Force client to send the query with the form of (h(c1), h(c2), **c1⊕c2**)

Motivating Scenario



https://e2e.ti.com/blogs_/b/thinkinnovate/archive/2015/09/15/smart-factories-of-the-future-enabling-technologies-for-industry-4-0

Internet of Things
Rule Checking

Augmented Collective Beings

- There are a lot of devices interacting with each other and with users, who are usually not IT professionals.

Cross-device Dependencies

- Explicit dependencies

  If power usage is higher than 50, turn off air conditioner

- Implicit dependencies
  - Via context, like temperature, location, human

  If air conditioner is turned off, temperature increases

Multi-stage Attacks

- Emerging threats via exploiting explicit/implicit dependencies to access higher-value targets

e.g., burglar wishing to break in can first turn off smart plug, which disconnects the air conditioner, which increases the temperature, which then triggers the window to open.

Objective

- Given a bunch of dependency rules, check whether several security and safety constraints are violated

Related work in firewall checking

- [1] checks anomalies that could exist in a single- or multi-firewall environment
- [1]: rules in sequence
- The execution order of firewall rules is fixed with respect to each packet.
  However, every rules operate in parallel in IoT

[1] Conflict Classification and Analysis of Distributed Firewall Policies @ IEEE Journal on Selected Areas in Communications 2005

Related work in SDN

- [2] is a layer between SDN controller and network devices that checks for network-wide invariant violations dynamically as each forwarding rule is inserted.

- The search space in SDN is fixed to the space of IP headers.
  However, in IoT, the search space changes when devices join or leave.

[2] VeriFlow: verifying network-wide invariants in real time @ HotSDN '12 Proceedings of the first workshop on Hot topics in software defined networks

Related works in IoT

- Most works focus on checking the existence of **conflicts**, which means that multiple rules try to use one or more sensors or actuators at the same time, which cause different effects on the environment
  - Conflicts between rules: [3], [4], [5]
  - Conflicts between users: [6]

- May not be applied directly
  - Global constraints may not be converted to rules
  - The conflicts between pairs may be too strict

[3] Policy conflicts in home automation @ Computer Networks: The International Journal of Computer and Telecommunications Networking 2013
[4] DepSys: Dependency Aware Integration of Cyber-Physical Systems for Smart Homes @ ICCPS '14: ACM/IEEE 5th International Conference on Cyber-Physical Systems
[5] An Application Conflict Detection and Resolution System for Smart Homes @ 2015 IEEE/ACM 1st International Workshop on Software Engineering for Smart Cyber-Physical Systems
[6] Conflict detection and resolution in home and building automation systems: a literature review @ Journal of Ambient Intelligence and Humanized Computing October 2014

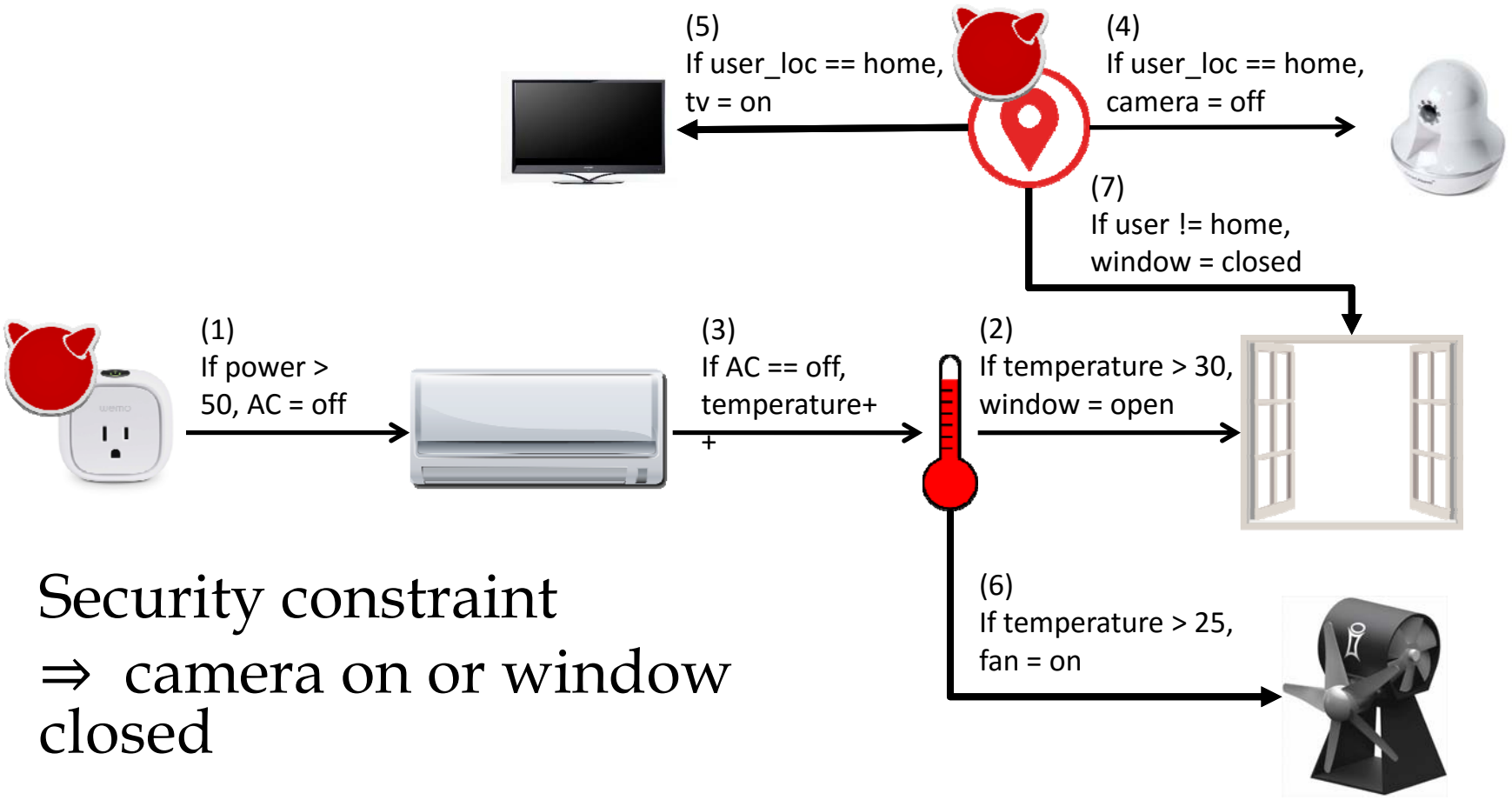
Related works in IoT

- [7] first considers the security challenges of cross-device dependencies in IoT
- [8] is mostly related

  Build a safety-centric programming platform for connected devices in IoT environments. However, the solution they proposed is not fast enough

[7] Handling a trillion (unfixable) flaws on a billion devices: Rethinking network security for the Internet-of-Things @ HotNets-XIV Proceedings of the 14th ACM Workshop on Hot Topics in Networks Article No. 5, 2015
[8] SIFT: Building an Internet of Safe Things @ IPSN '15 Proceedings of the 14th International Conference on Information Processing in Sensor Networks, 2015

Small Dataset



Security constraint

⇒ camera on or window closed

First try

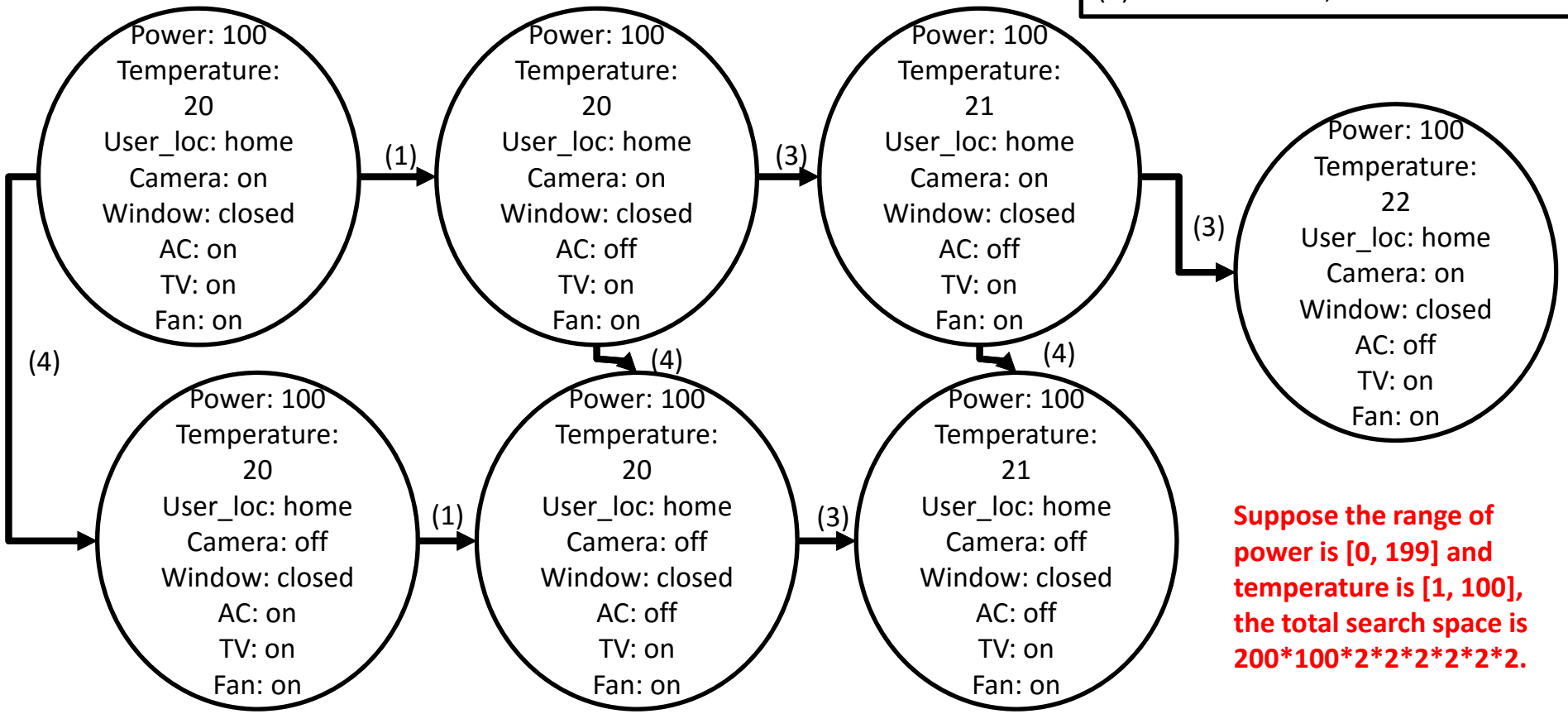
- The “if-this-then-that” clause is similar to the “implication” in logic
  ⇒ try to model the rules in the form of propositional logic

- However, the concept of “state” is absent in simple logic
  ⇒ the situation in which the temperature or the power usage increases cannot be modelled.

- Thus, use state machine to model the rules’ effect on environments

Finite State Machine

(1) If power > 50, AC = off
(2) If temperature > 30, window = open
(3) If AC == off, temperature++
(4) If user_loc == home, camera = off
(5) If user_loc == home, tv = on
(6) If temperature > 25, fan = on
(7) If user != home, window = closed



Suppose the range of power is [0, 199] and temperature is [1, 100], the total search space is 200*100*2*2*2*2*2*2.

Conclusion

- Three security issues
  - Two for cloud
  - One for IoT
  - IoT security is more related to smart factory in a straightforward way


THANK YOU


ANY
QUESTIONS
?


Any comments?
we welcome our customers' comments on our merchandise
and service. Will you take a moment to give us your ideas and suggestions so
that we may continue to provide the best product and customer service possible?
Thank you.