Cloud Cryptography

Seny Kamara
Cryptography Group
Microsoft Research

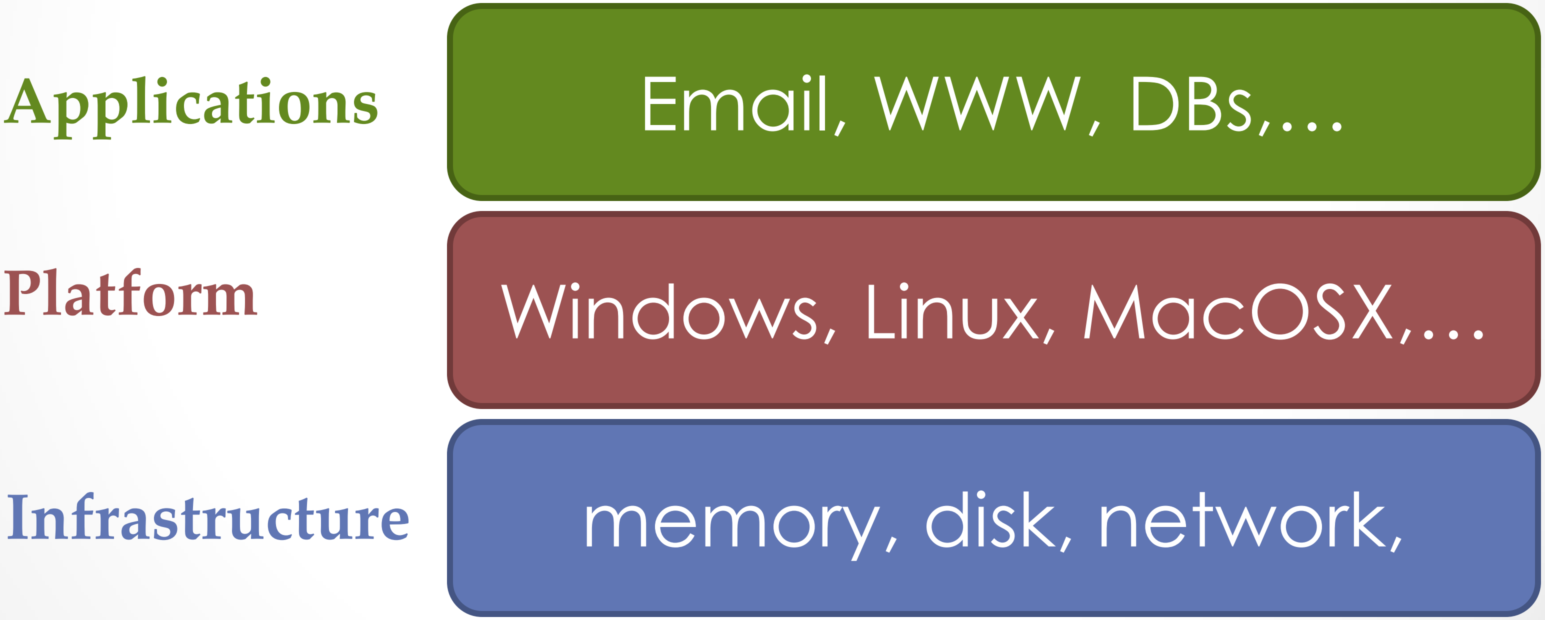
Outline

- Cloud Architecture
  - What is cloud computing?
- Cloud Ecosystem
  - Who provides and who consumes cloud services?
- Cloud Cryptography
  - What are the security concerns & how can cryptography help?

Computing as a Service

- Computing is a vital resource
  - Enterprises, governments, scientists, consumers, …
- Computing is manageable at small scales…
  - e.g., PCs, laptops, smart phones
- …but becomes hard to manage at large scales
  - build and manage infrastructure, schedule backups, hardware maintenance, software maintenance, security, trained workforce, …
- Why not outsource it?

Computing Architecture

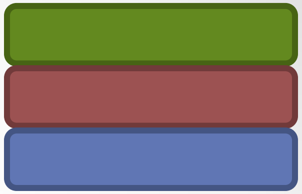


Applications — Email, WWW, DBs,…

Platform — Windows, Linux, MacOSX,…

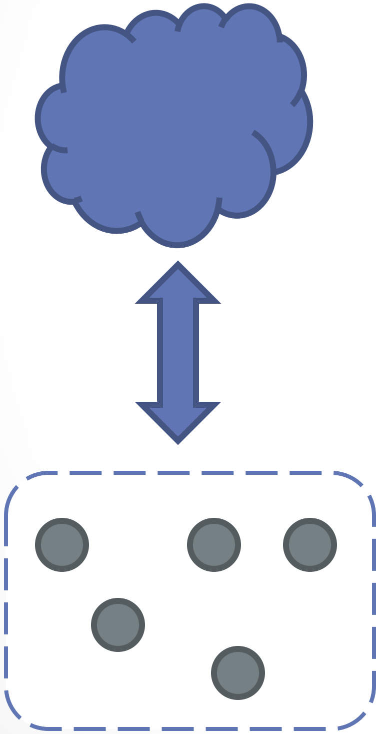
Infrastructure — memory, disk, network,

Cloud Services

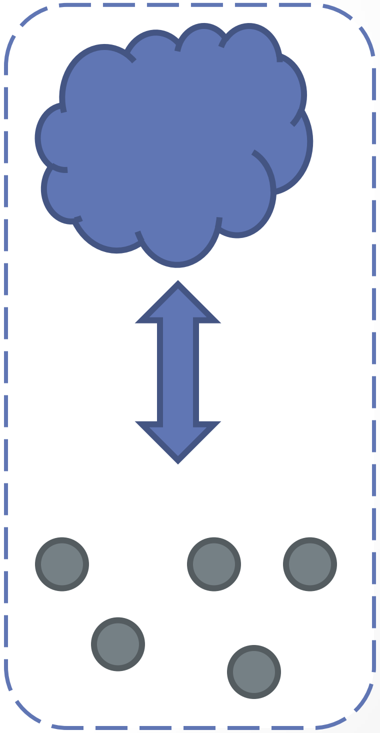


- Infrastructure as a service
  - Service: customer can store data in the cloud
  - Customer: enterprise, developers
  - e.g., MS Azure storage, Amazon S3
- Platform as a service
  - Service: customer can run its apps in the cloud
  - Customer: developers
  - e.g., MS Azure, Amazon EC2, Google AppEngine,
- Software as a service
  - Service: customer makes use of app in the cloud
  - Customer: consumers & enterprise
  - e.g., web-based email, Flickr, delicious, Facebook, Office Web, Google Docs, …

Cloud Deployment Models



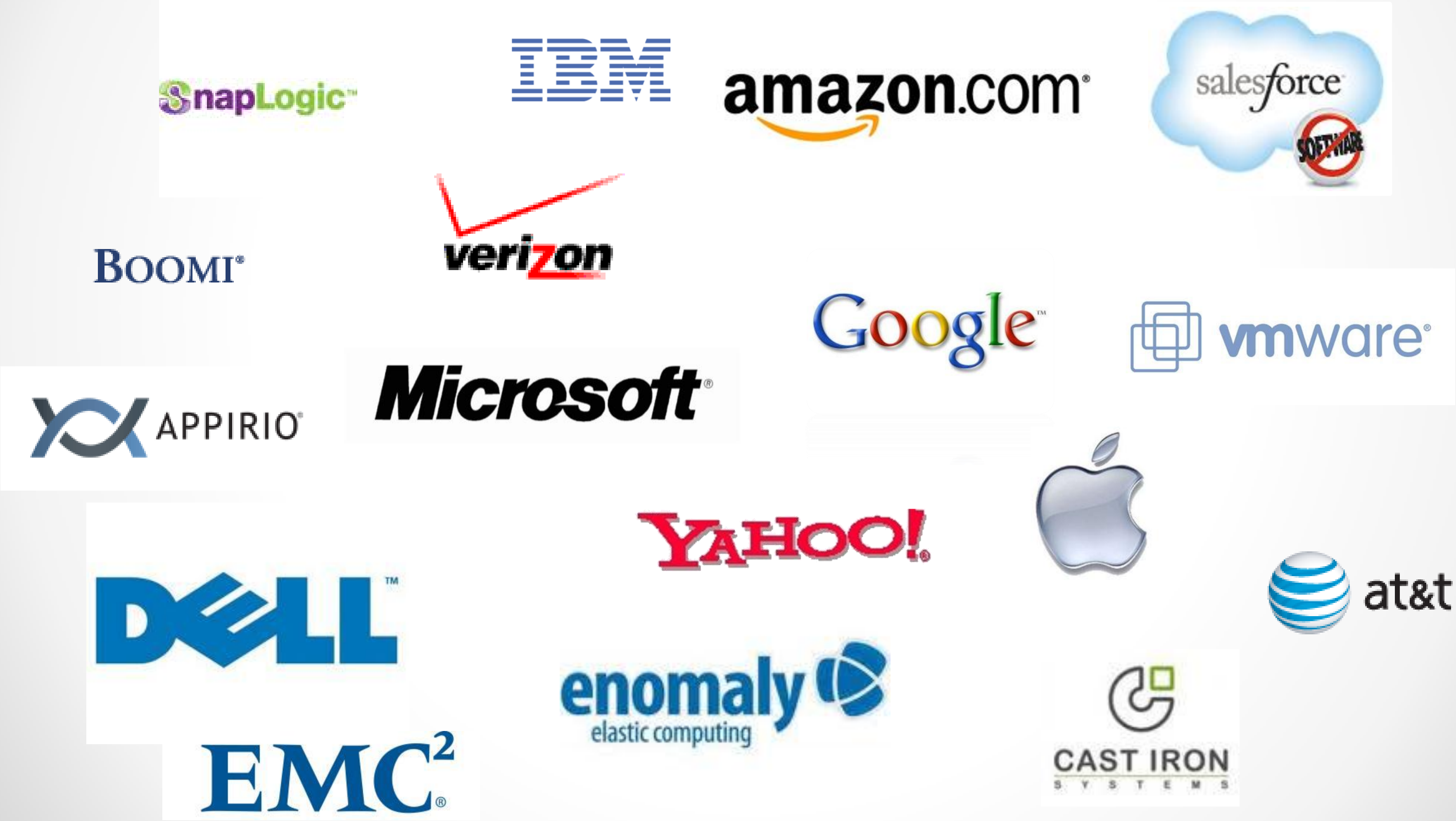
Public



Private

The Cloud Ecosystem

Who Provides Cloud Services?



Cloud Infrastructure Providers

- Provide access to infrastructure
  - e.g., Amazon, Microsoft, Google, IBM, EMC, Equinix, AT&T, Verizon
- Characteristics
  - Requires very large investments
    - build data centers
    - acquire expertise
    - provide physical security
    - energy consumption
    - …
  - Large (often) publicly traded companies
  - Have a reputation to uphold

Cloud Service Companies

- Provide cloud-based applications
  - e.g., Salesforce, GoGrid, NetSuite
- Characteristics
  - Requires small investment
    - developers
    - Platform/infrastructure services from larger cloud providers
  - Startups (often) privately held

Who Consumes Cloud Services?

- Consumers
  - e.g., Facebook (500+ M), Web-based email (840 M), Flickr, Dropbox, …
- Enterprise
  - E.g., Amazon EC2/S3, MS Azure, Google AppEngine, Google Apps
- Governments
  - 120,000 US Dept. of Agriculture employees will move to MS cloud services
  - 17,000 Gen. Serv. Admin. Employees will move to Google cloud services
- Local Governments
  - 100,000 NYC emplyees will move to MS cloud services
  - 34,000 L.A. emplyees will move to Google cloud services

Cloud Cryptography

Concerns

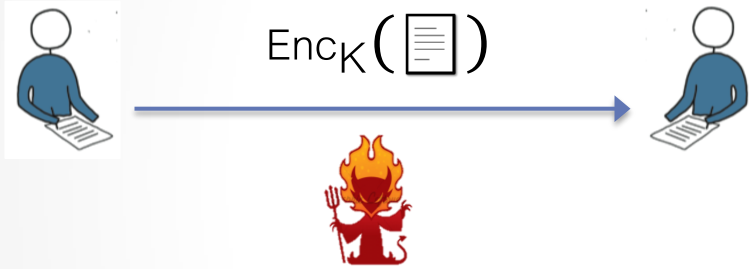
- Outsider security
  - Can other tenants, hackers, competitors access my data?
- Insider security
  - Can the cloud operator (and its employees) access my data?
- Intellectual property
  - Can outsiders or insiders see my code and algorithms?
- Compliance
  - Can I remain compliant if I move to the cloud?
- Availability
  - Can I access my data or service at all times?

Modern Cryptography

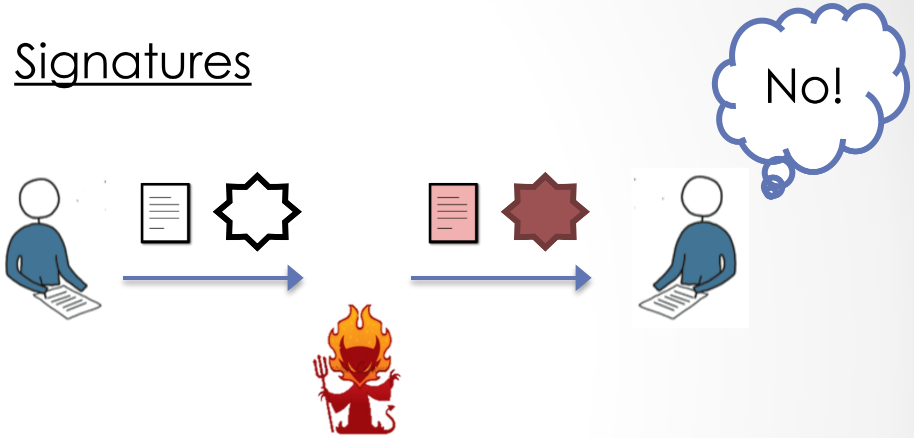
- Primitives
  - e.g., encryption, digital signatures, hash functions, pseudo-random generators, …
- Protocols
  - e.g., key agreement, zero-knowledge proofs, multi-party computation
- Security definitions
  - Formal definition of what it means to be secure
- “Proofs” of security
  - Proof that primitive/protocol meets security definition
  - Unconditional security (e.g., one-time pad)
  - Conditional security (e.g., RSA, El Gamal,…)
- Leads to very strong security guarantees
  - e.g., digital signatures are widely accepted in court
  - SHA-2, AES, ECC are certified for government use by NIST & NSA

Modern Cryptography

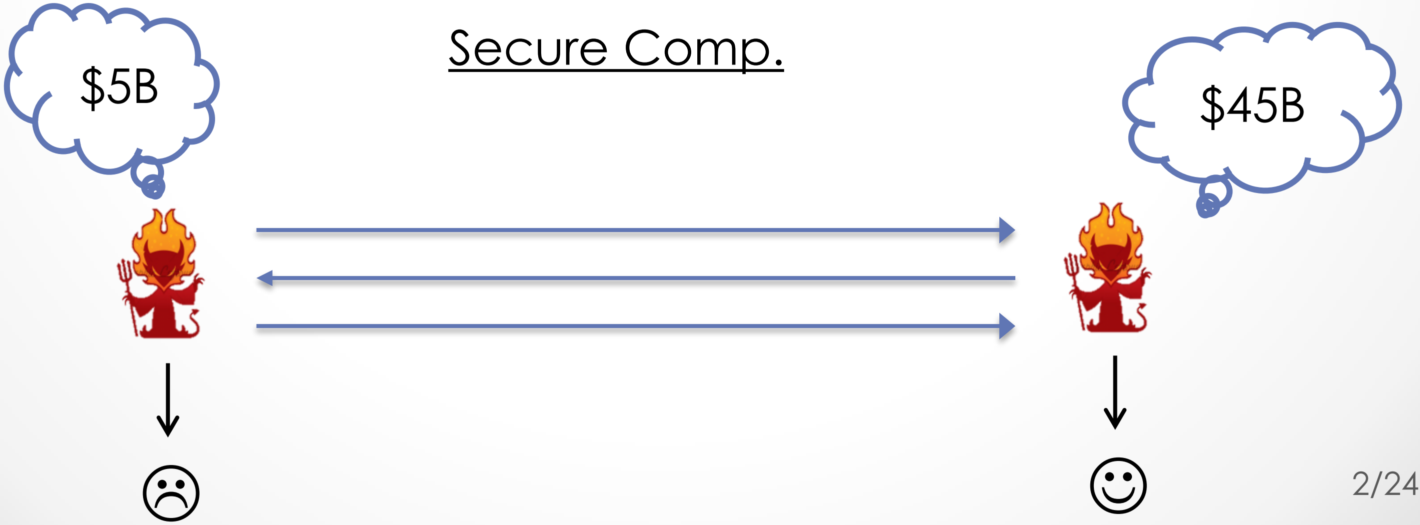
Encryption



Signatures



Secure Comp.



Cloud Cryptography

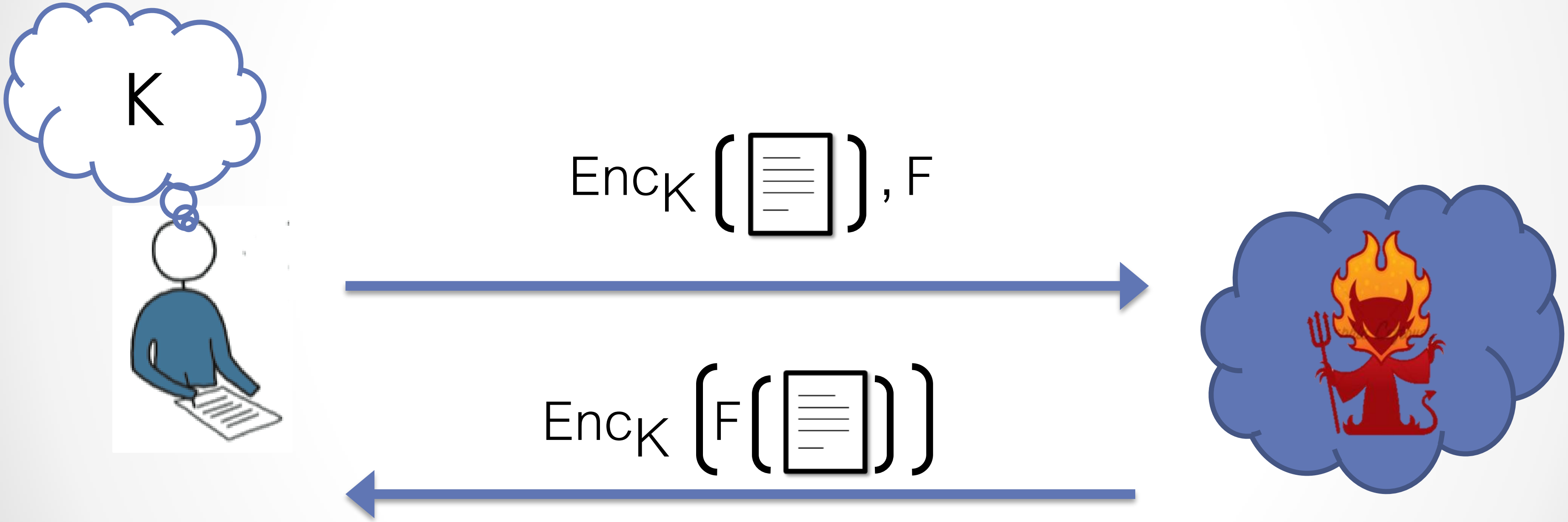
- Current crypto tools are inappropriate for the cloud
  - Due to assumptions about how tools will be used
  - Results in efficiency loss & insecurity
- New tools
  - Homomorphic encryption
  - Searchable/Structured encryption
  - Proofs of storage
  - Server-aided secure computation

[…,G09,…]

Homomorphic Encryption

- Encryption that supports comp. on encrypted data
  - *Fully* homomorphic [G09, DGHV10]
  - Partially homomorphic [SYY99, BGN05, IP07,GHV10a,GHV10b,KR11]
- Guarantees that
  - Cloud never sees plaintext/message
- Pros
  - FHE is general-purpose
  - Partial & parallel HE can be efficient
- Cons
  - FHE is inefficient (but improvements are being made rapidly)

Homomorphic Encryption



K

$\mathrm{Enc}_K(\quad), F$

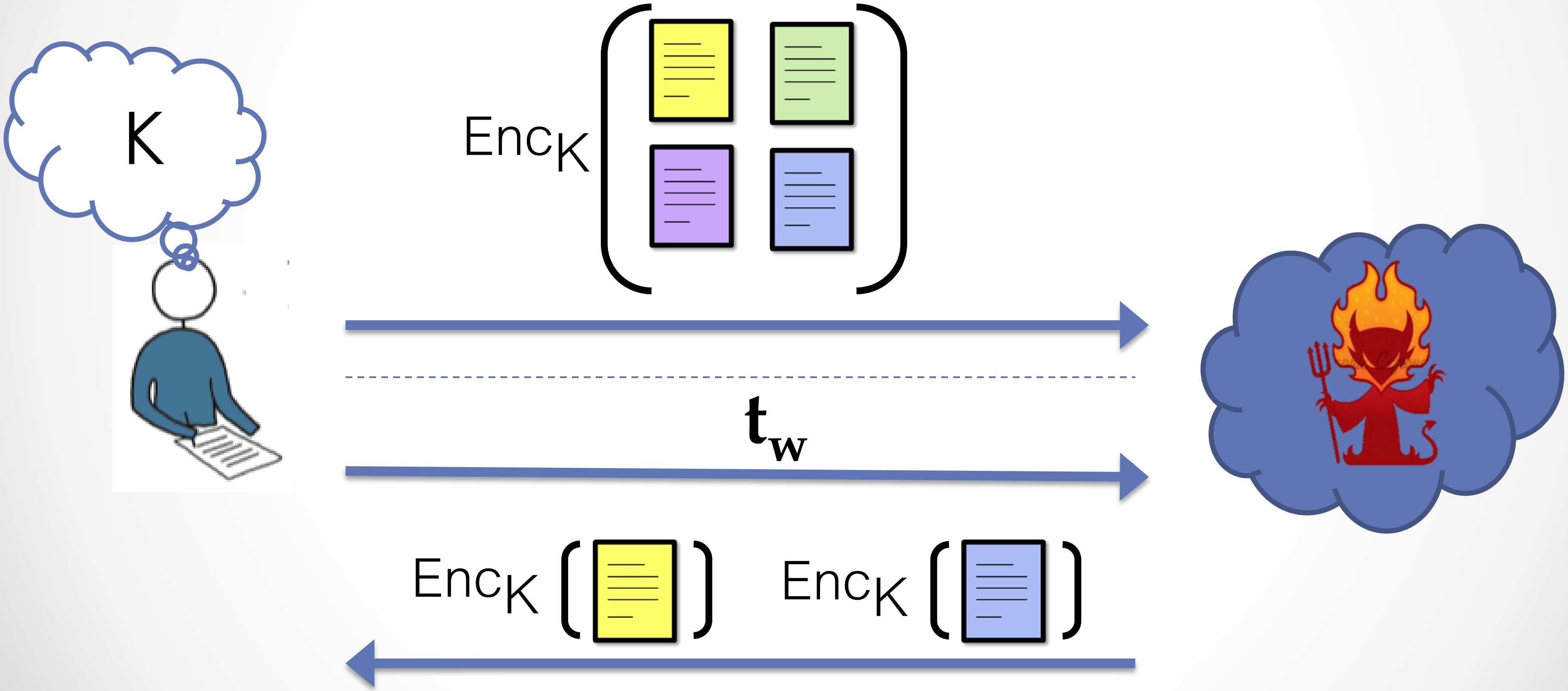
$\mathrm{Enc}_K(F(\quad))$

[SWP01]

Searchable Encryption

- Encryption that supports search on encrypted text
  - Symmetric key [SWP01,Goh03,CM05,CGKO06]
  - Public key [BDOP06, BKOS07,…]
- Guarantees that
  - Cloud never sees documents
  - Cloud never sees search keywords
- Pros
  - Symmetric variant is very efficient!
- Cons
  - Reveals access and search patterns
  - [GO96] shows how to hide this but it is expensive

Searchable (Symm.) Encryption

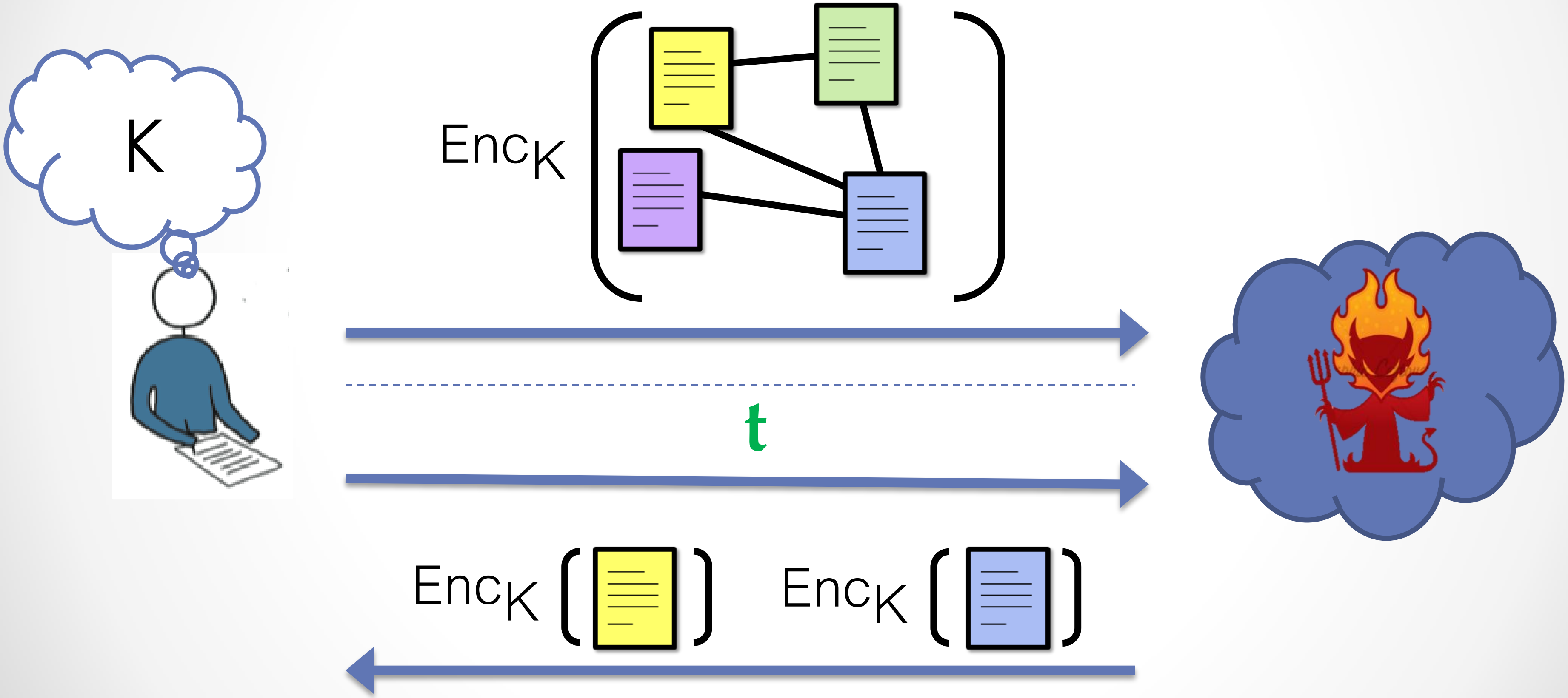


[CK10]

Structured Encryption

- Encryption that supports queries on encrypted data
  - Query over encrypted graphs [CK10]
  - Query over encrypted web graphs [CK10]
- Guarantees that
  - Cloud never sees data
  - Cloud never sees queries
- Pros
  - Symmetric variant is very efficient!
- Cons
  - Reveals access and search patterns

Structured Encryption



K

Enc_K

t

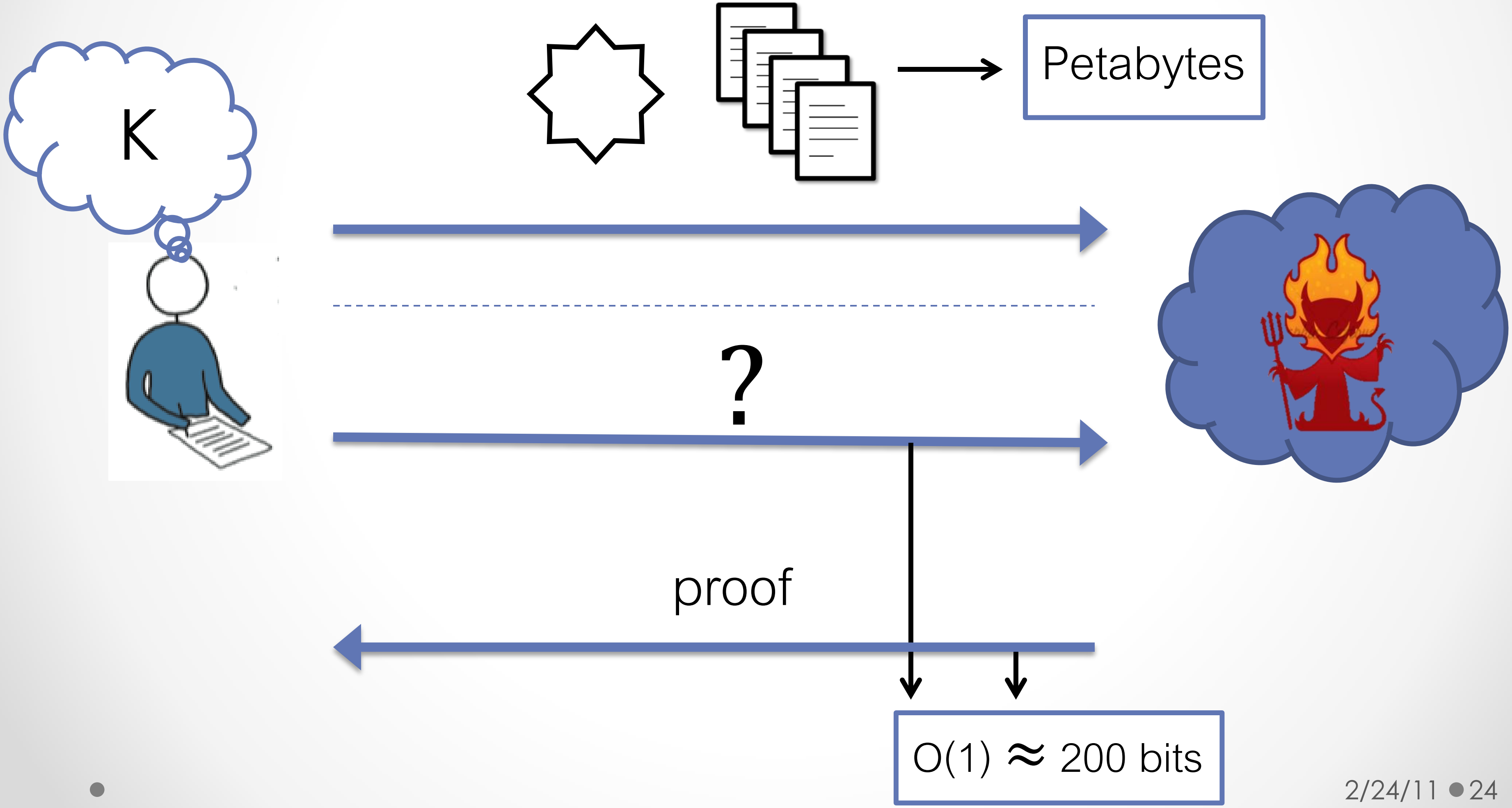
Enc_K Enc_K

[JK07, ABC+07]

Proofs of Storage

- Tamper detection without knowing original file
  - Symmetric-key [JK07, SW08, DVW10]
  - Public-key [ABC+07, SW08, AKK10]
- Guarantees that
  - Cloud will be caught if it tampers with data
- Pros
  - Symmetric variant is efficient!
  - Verification does not require copy of original data
- Cons
  - --

Proofs of Storage



K

Petabytes

?

proof

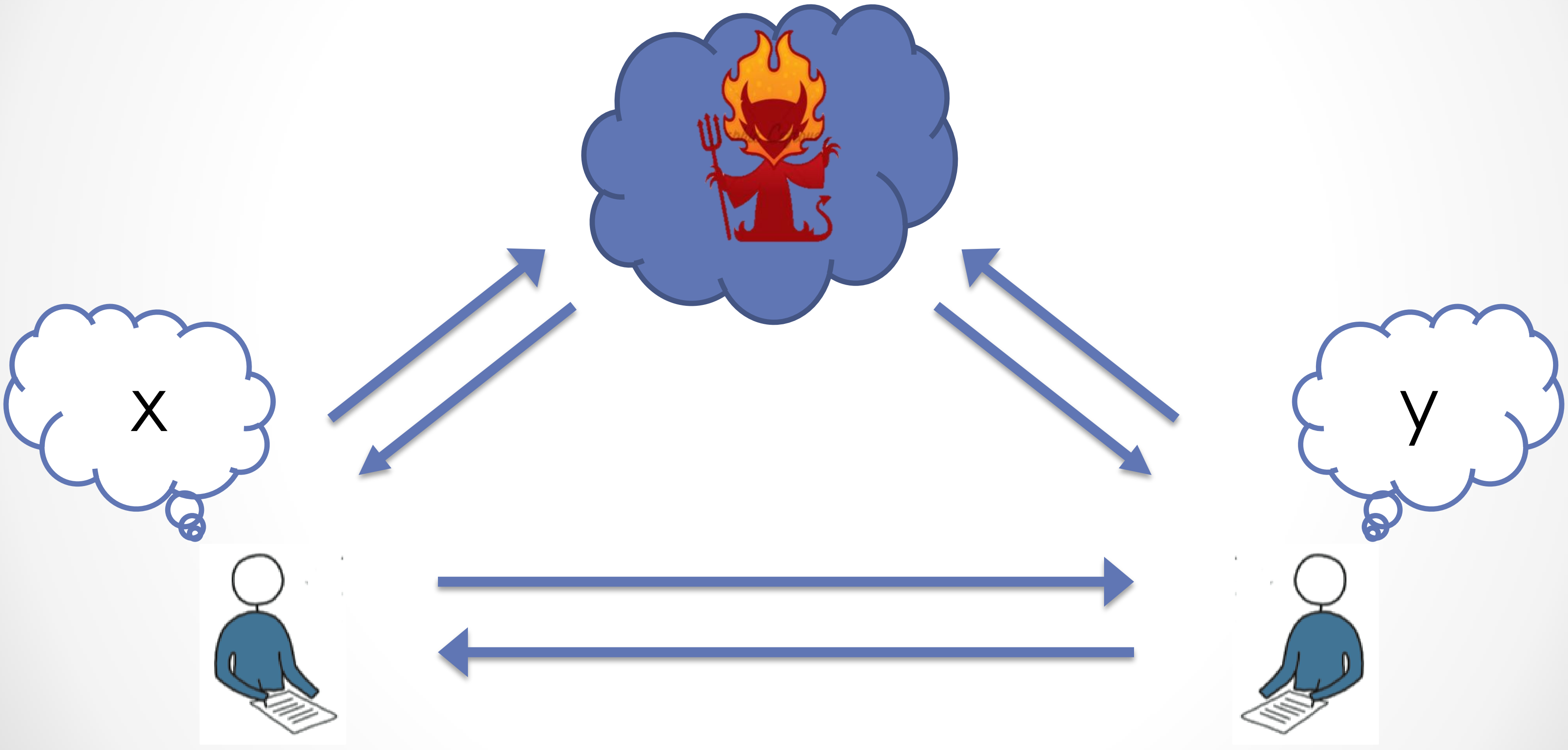
O(1) ≈ 200 bits

[KMR11]

Server-Aided Secure Comp.

- Joint computation w/o revealing inputs
  - (plain) secure computation [Yao82,GMW87,…]
- Guarantees that
  - Parties will not learn each other's inputs
  - Cloud will not learn parties' inputs
- Pros
  - General-purpose (e.g., data mining, voting, negotiations,…)
  - Efficient
- Cons
  - --

Server-Aided Secure Comp.



f(x,y)

Questions?

References

- [MG09]
  - *The NIST Definition of Cloud Computing*
  - http://csrc.nist.gov/groups/SNS/cloud-computing/cloud-def-v15.doc
- [G09]
  - Craig Gentry
  - *Fully Homomorphic Encryption from Ideal Lattices.*
  - ACM Symposium on Theory of Computing, 2009.
- [DGHV10]
  - Martin van Dijk, Craig Gentry, Shai Halevi and Vinod Vaikuntanathan
  - *Fully Homomorphic Encryption Over the Integers*
- [SYY99]
  - Tomas Sanders, Adam Young and Moti Yung
  - *Non-interactive Cryptocomputing for NC1*
  - IEEE Symposium on the Foundations of Computer Science, 1999

References

- [BGN05]
  - Dan Boneh, Eu-Jin Goh and Kobi Nissim
  - *Evaluating 2-DNF Formulas on Ciphertexts*
  - Theory of Cryptography Conference, 2005
- [IP07]
  - Yuval Ishai and Anat Paskin
  - *Evaluating branching programs on encrypted data*
  - Theory of Cryptography Conference, 2007
- [GHV10a]
  - Craig Gentry, Shai Halevi and Vinod Vaikuntanathan
  - *A Simple BGN-style Encryption Scheme from LWE*
  - Advances in Cryptology – Eurocrypt, 2010

References

- [GHV10b]
  - Craig Gentry, Shai Halevi and Vinod Vaikuntanathan
  - *i-hop Homomorphic Encryption Schemes*
  - Advances in Cryptology – CRYPYO, 2010
- [KR11]
  - Seny Kamara and Mariana Raykova
  - *Parallel Homomorphic Encryption*
  - Under submission
- [SWP01]
  - Dawn Song, David Wagner and Adrian Perrig
  - *Practical Techniques for Searches on Encrypted Data*
  - IEEE Security and Privacy Symposium, 2000

References

- [Goh03]
  - Eu-Jin Goh
  - *Secure Indexes*
  - http://eprint.iacr.org/2003/216
- [CM05]
  - Yang-Chen Chang and Michael Mitzenmacher
  - *Privacy preserving keyword searches on remote encrypted data*
  - Conference on Applied Cryptography and Network Security, 2005
- [CGKO06]
  - Reza Curtmola, Juan Garay, Seny Kamara and Rafail Ostrovsky
  - *Symmetric Searchable Encryption: Improved Definitions and Efficient Constructions.*
  - ACM Conference on Computer & Communication Security, 2006

References

- [BDOP04]
  - Dan Boneh, Giovanni di Crescenzo, Rafail Ostrovsky and Giuseppe Persiano
  - *Public-Key Encryption with Keyword Search*
  - Advances in Cryptology – Eurocrypt, 2004
- [BKOS07]
  - Dan Boneh, Eyal Kushilevitz, Rafail Ostrovsky, William E. Skeith III
  - *Public Key Encryption That Allows PIR Queries*
  - Advances in Cryptology – CRYPTO, 2007
- [CK10]
  - Melissa Chase and Seny Kamara
  - *Structured Encryption and Controlled Disclosure*
  - Advances in Cryptology – Asiacrypt, 2010

References

- [JK07]
  - Ari Juels and Burt Kaliski
  - *PORs: Proofs of Retrievability for Large Files*
  - ACM Conference on Computer & Communications Security, 2007
- [ABC+07]
  - Giuseppe Ateniese, Randal C. Burns, Reza Curtmola, Joseph Herring, Lea Kissner, Zachary N. J. Peterson, Dawn Xiaodong Song
  - *Provable Data Possession at Untrusted Stores*
  - ACM Conference on Computer & Communications Security, 2007
- [SW08]
  - Hovav Shacham, Brent Waters
  - *Compact Proofs of Retrievability*
  - Advances in Cryptology – Asiacrypt, 2008

References

- [DVW09]
  - Yevgeniy Dodis, Salil P. Vadhan, Daniel Wichs
  - *Proofs of Retrievability via Hardness Amplification*
  - Theory of Cryptography Conference, 2009
- [AKK10]
  - Giuseppe Ateniese, Seny Kamara and Jonathan Katz
  - *Proofs of Storage from Homomorphic Identification Schemes*
  - Advances in Cryptology – Asiacrypt, 2010
- [KMR11]
  - Seny Kamara, Payman Mohassel and Mariana Raykova
  - *Server-Aided Secure Computation*
  - Under submission