



# Crypto

**Frederik Armknecht**  
University of Mannheim

March 19, 2015

Troopers, Heidelberg

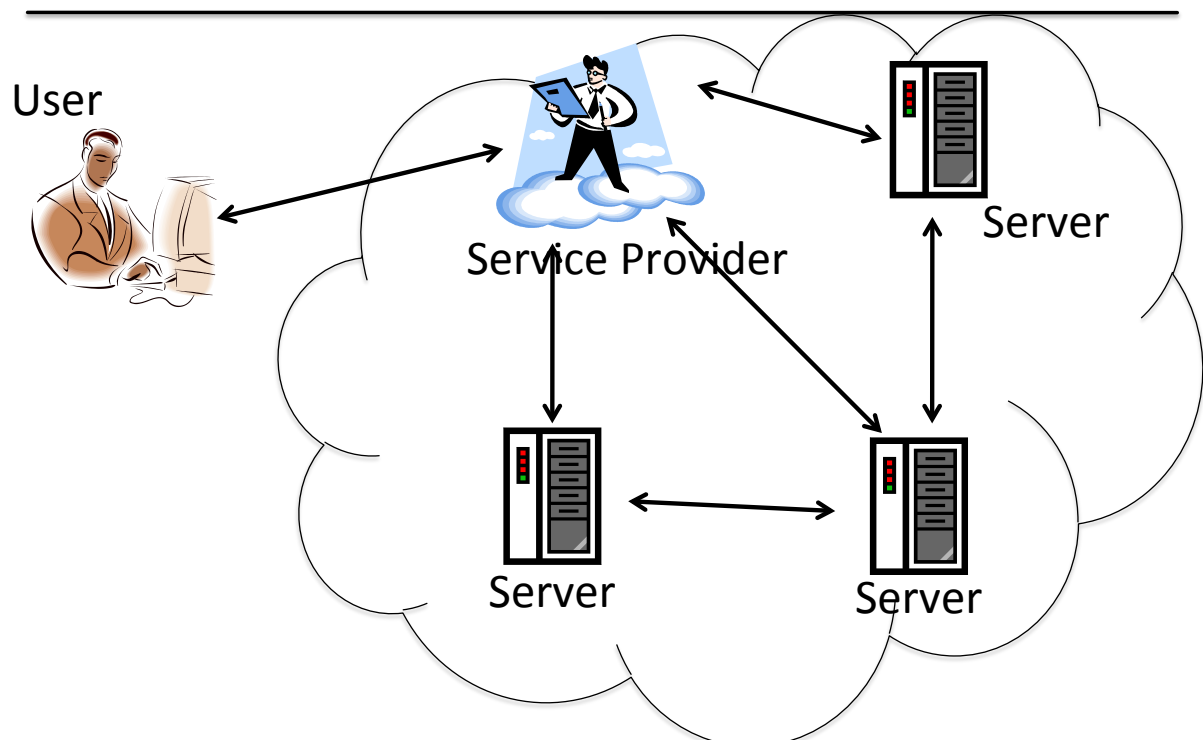
## Agenda

---

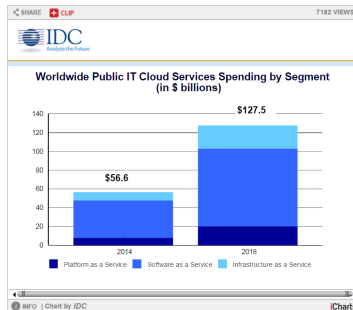
- **Motivation**
- **Data Breach**
- **Data Loss**
- **Conclusion**

# Motivation

## Cloud Computing



# Importance of Cloud Computing



- \$56.6 billion in 2014
- Will grow to more than \$127 billion in 2018
- 5-year compound annual growth rate of 22.8% (about 6 times the rate of growth for the overall IT market)



- 90% of US companies use some form of cloud computing

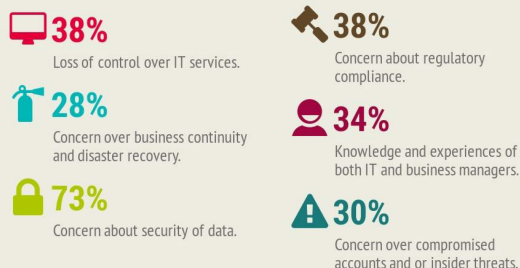
# Challenge: Cloud Security



The Cloud Security Alliance (CSA) is a not-for-profit organization with a mission to promote the use of best practices for providing security assurance within Cloud Computing, and to provide education on the uses of Cloud Computing to help secure all other forms of computing. The Cloud Security Alliance is led by a broad coalition of industry practitioners, corporations, associations and other key stakeholders.



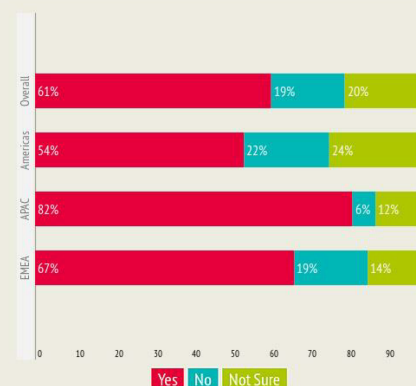
## Top challenges holding back Cloud projects.



© 2015 Cloud Security Alliance - All Rights Reserved.

10

## Is security of data residing in the cloud an executive or board-level concern?



# Top Risks



Top Threats Working Group

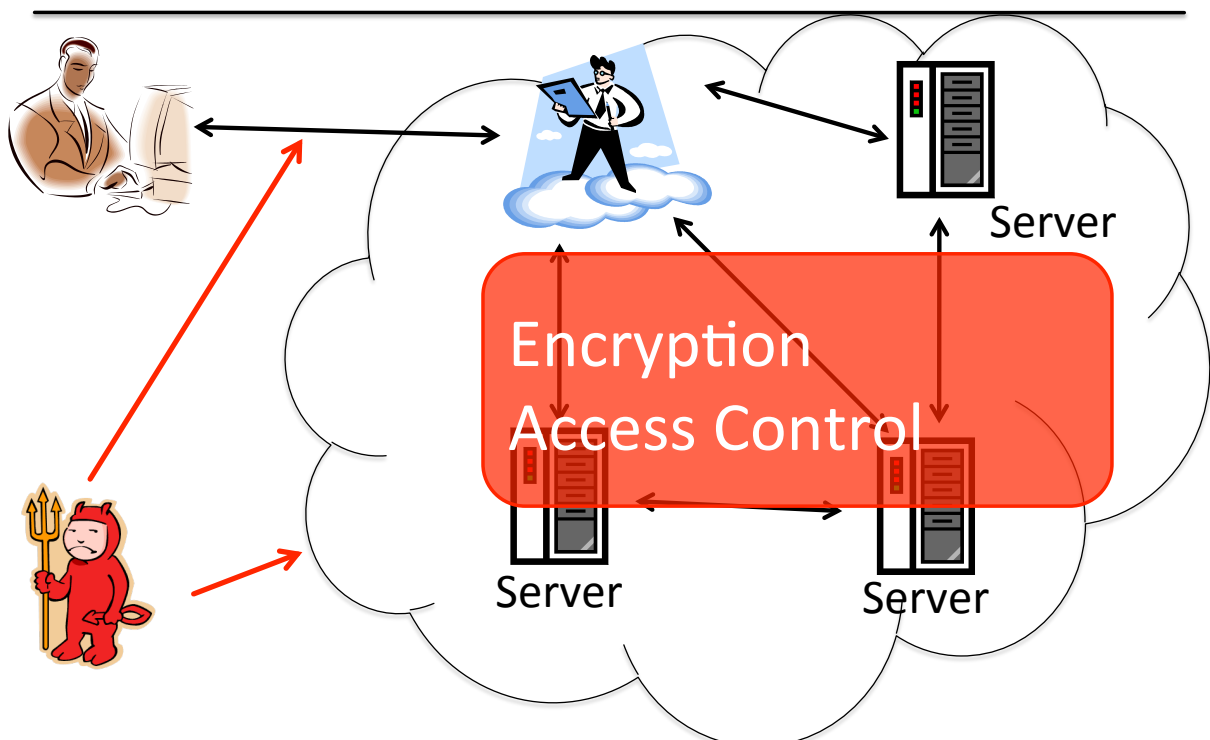
## The Notorious Nine

### Cloud Computing Top Threats in 2013

To identify the top threats, CSA conducted a survey of industry experts to compile professional opinion on the greatest vulnerabilities within cloud computing. The Top Threats working group used these survey results alongside their expertise to craft the final 2013 report. The survey methodology validated that the threat listing reflects the most current concerns of the industry. In this most recent edition of this report, experts identified the following nine critical threats to cloud security (ranked in order of severity):

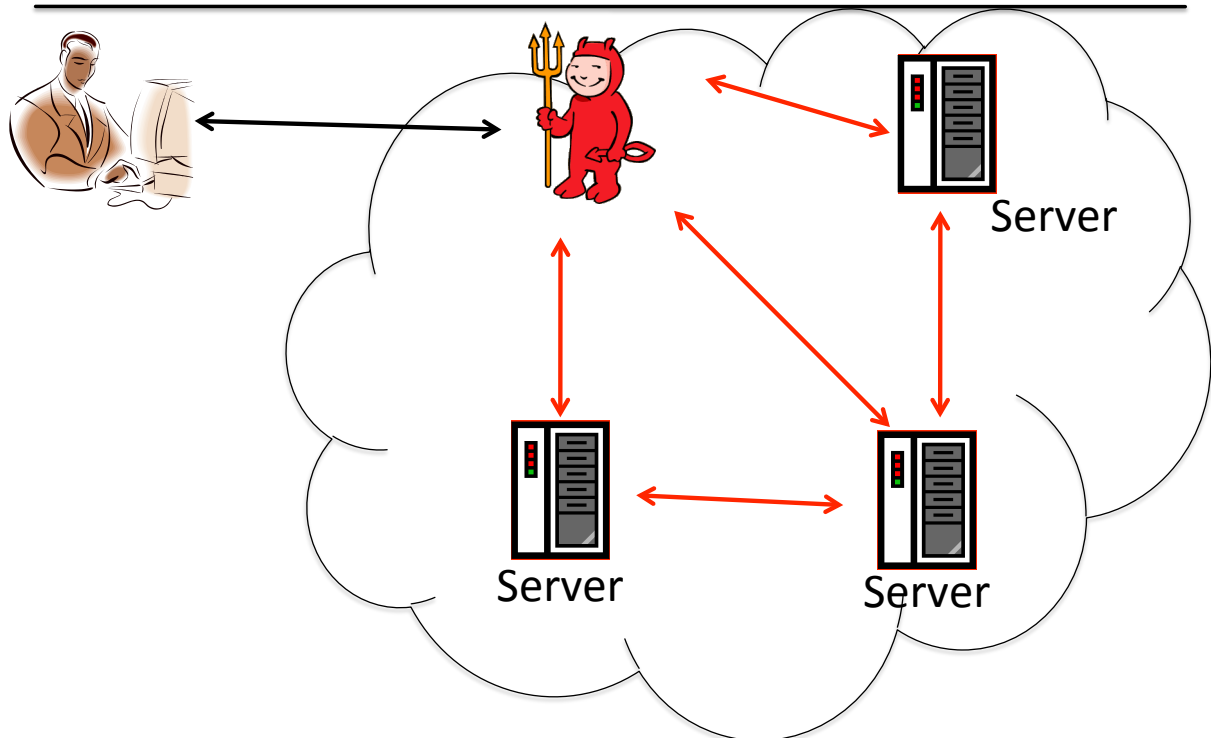
1. Data Breaches
2. Data Loss
3. Account or Service Traffic Hijacking
4. Insecure Interfaces and APIs
5. Denial of Service
6. Malicious Insiders
7. Abuse of Cloud Services
8. Insufficient Due Diligence
9. Shared Technology Vulnerabilities

## Outsider Attacker





# Insider Attacker?

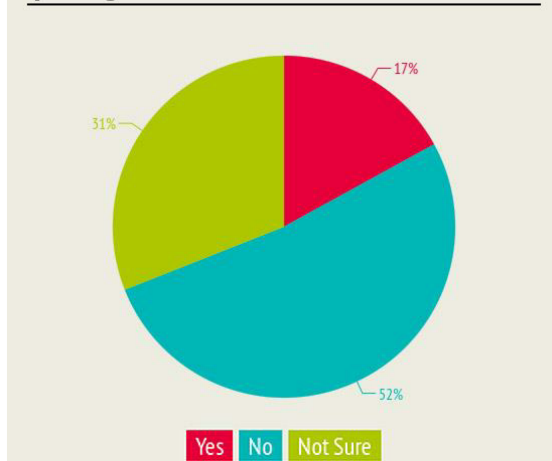


Frederik Armknecht

9

CLOUD ADOPTION PRACTICES & PRIORITIES SURVEY REPORT January 2015

**Has your organization experienced an insider threat incident in the last year, such as an employee downloading sensitive data before quitting?**



## State of Security

Companies are increasingly under attack as criminal organizations and state-sponsored groups attempt to steal sensitive data. Not surprisingly, IT professionals see the top security issues facing their organizations as malware (63 percent), advanced persistent threats (53 percent), compromised accounts (43 percent), and insider threats (42 percent). Although companies are focused on external threats, 17 percent reported a known insider threat incident in the last 12 months, such as an employee downloading sensitive data before quitting. Troublingly, 31 percent were not sure if such an incident occurred. This uncertainty should raise some concern about whether companies have the right resources to identify and stop these types of threats.

**More software vulnerabilities have been uncovered in 2014 than any other year on record.**

© 2015 Cloud Security Alliance - All Rights Reserved.

Frederik Armknecht

10

# This Talk

---

- If service provider is mistrusted (or careless), traditional cryptographic methods cannot be used anymore
- Aim of this talk: Discuss novel cryptographic methods that may help to protect
  - Principles
  - Advantages/Disadvantages
  - Current state
- Focus: cryptographic building blocks, not comprehensive solutions



## *Data Breach*

# Most Significant Risk

CLOUD SECURITY ALLIANCE The Notorious Nine: Cloud Computing Top Threats in 2013

## 1.0 Top Threat: Data Breaches

It's every CIO's worst nightmare: the organization's sensitive internal data falls into the hands of their competitors. While this scenario has kept executives awake at night long before the advent of computing, cloud computing introduces significant new avenues of attack. In November 2012, researchers from the University of North Carolina, the University of Wisconsin and RSA Corporation released a paper describing how a virtual machine could use side channel timing information to extract private cryptographic keys being used in other virtual machines on the same physical server. However, in many cases an attacker wouldn't even need to go to such lengths. If a multitenant cloud service database is not properly designed, a flaw in one client's application could allow an attacker access not only to that client's data, but every other client's data as well.

### SERVICE MODEL

IaaS

PaaS

SaaS

### RISK MATRIX



**If nobody is trusted,  
data should be intrinsically protected**

# Concerns Are Justified

## Latest Incidents

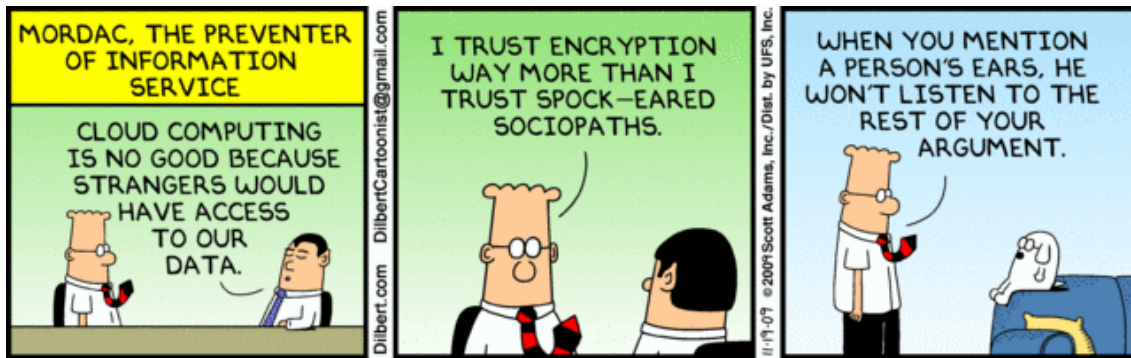
RECORDS	DATE	ORGANIZATIONS
22,867	2015-03-12	Google, eNom Inc.
5,514	2015-03-10	Blue Cross Blue Shield of Michigan
4,697	2015-03-10	Texas A&M University
?	2015-03-09	Unknown Organization, NEXTEP SYSTEMS, Zoupl
141	2015-03-09	Grillin' Wood
14	2015-03-09	Playdowns.c
?	2015-03-05	Sportklinik B
?	2015-03-04	Mandarin Or
7,945	2015-03-03	Neofriends
?	2015-03-02	Natural Groc

**DATALOSSdb**  
open security foundation

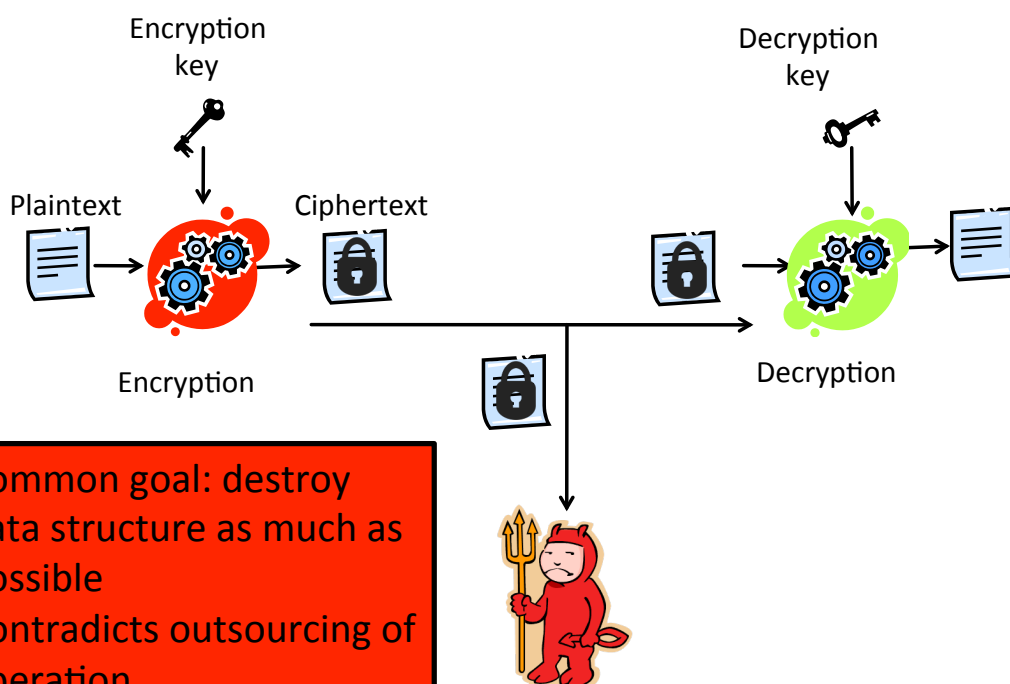
## Largest Incidents

RECORDS	DATE	ORGANIZATIONS
220,000,000	2014-08-22	Unknown Organization
152,000,000	2013-10-03	Adobe Systems, Inc.
150,000,000	2012-03-17	Shanghai Roadway D&B Marketing Services Co. Ltd
145,000,000	2014-05-21	eBay Inc.
140,000,000	2013-06-08	Unknown Organization
130,000,000	2009-01-20	Heartland Payment Systems, Tower Federal Credit Union, Beverly National Bank, North Middlesex Savings Bank, Golden Chick
110,000,000	2013-12-18	Target Brands, Inc., Fazio Mechanical Services, Inc.
109,000,000	2014-09-02	Home Depot, Unknown Organization
104,000,000	2014-01-20	Korea Credit Bureau, NH Nonghyup Card, Lotte Card, KB Kookmin Card
94,000,000	2007-01-17	TJX Companies Inc.

# What Dilbert says...

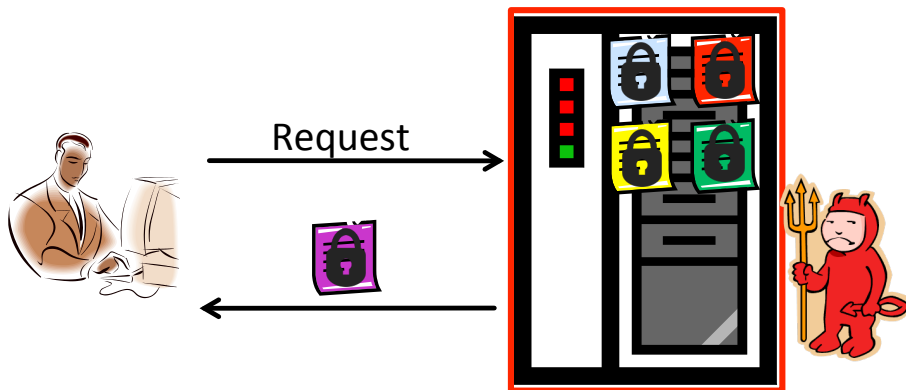


# Encryption



## What If ...

- Store data encrypted
- Requests are operated on the encrypted data
- Service provider returns correct result **WITHOUT** knowing content of the data and/or of result



## Focus Today

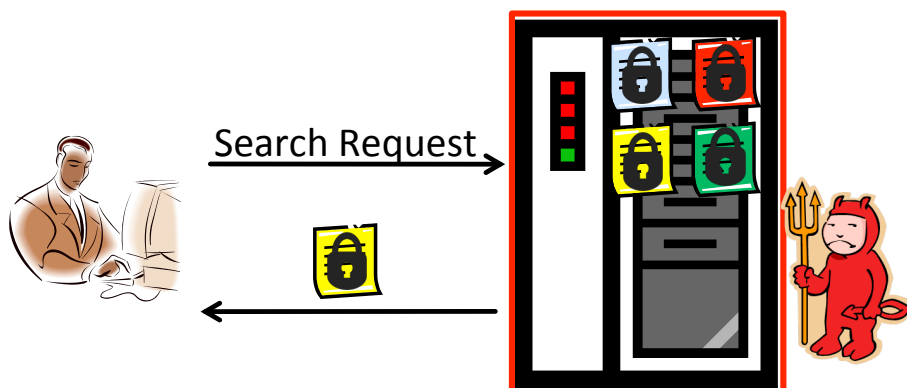
- **Searchable Encryption**
  - Allows to search on encrypted data
- **Homomorphic Encryption**
  - Allows to compute on encrypted data

# Encryption

## Searchable Encryption

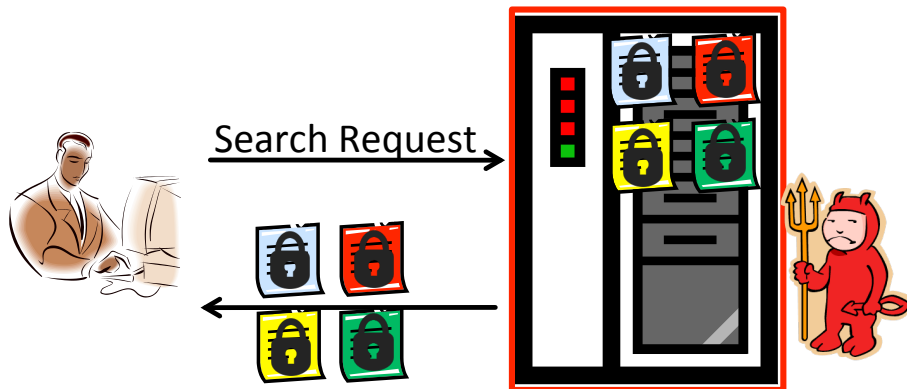
## Scenario

- Encrypted data outsourced
- User has search request



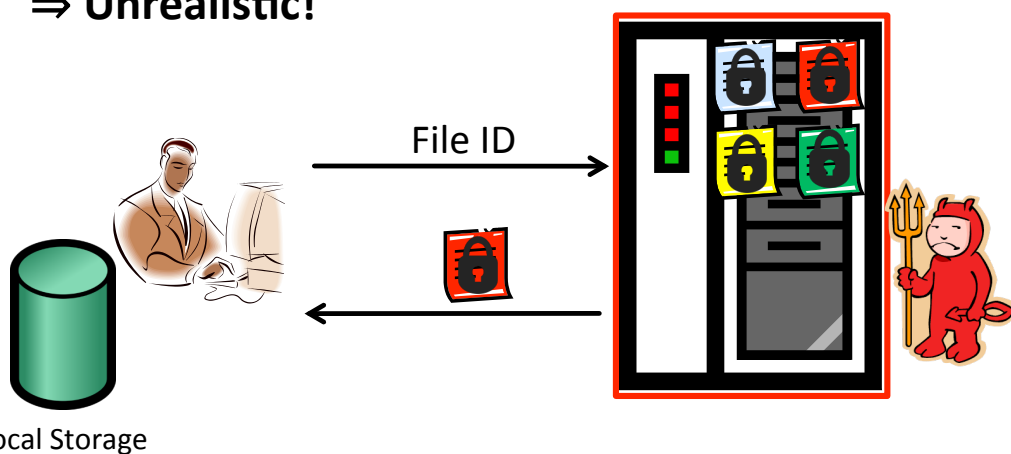
## Simple Solution 1

- Return complete encrypted data base
  - High communication and communication effort
- ⇒ Unrealistic!



## Simple Solution 2

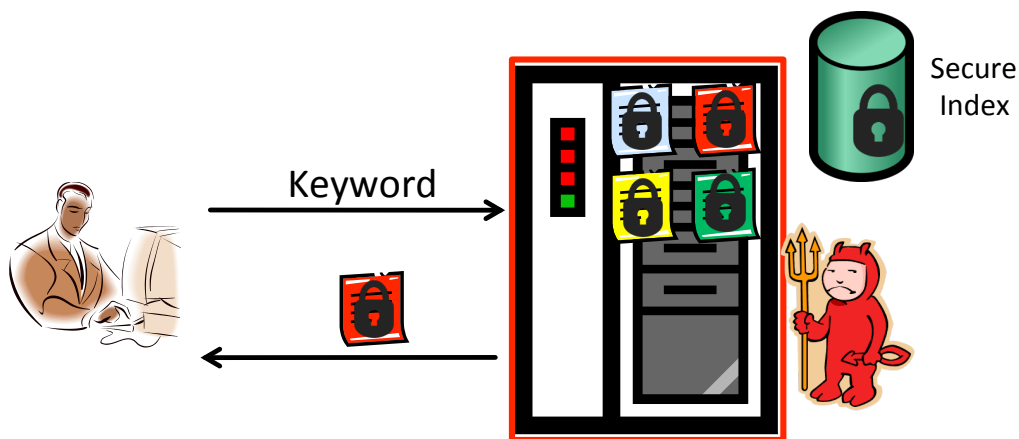
- Locally store information on outsourced data
  - Send file ID only
- ⇒ Large local storage
- ⇒ Unrealistic!





# Searchable Symmetric Encryption

- User outsources encrypted data AND secure index
- Search request: Create search token, receive all fitting ciphertexts



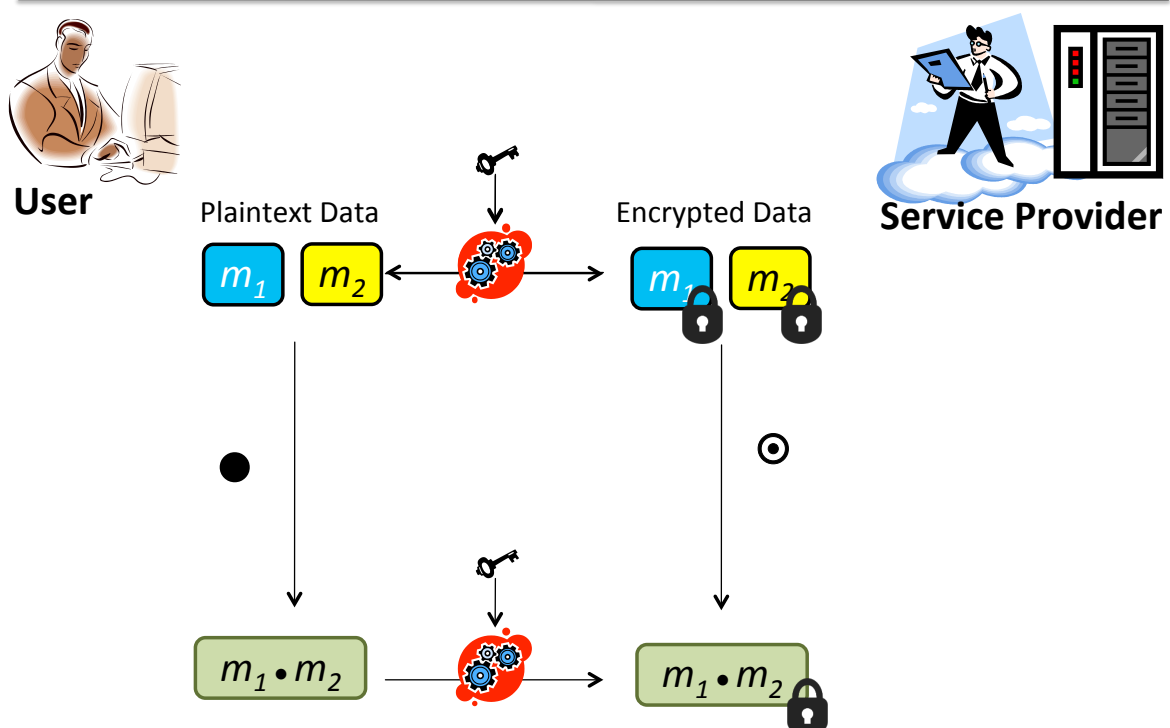
Scheme	Dynamism	Security	Search Time	Index Size
Song et al. [2000]	Static	CPA	$n/p$	N/A
Goh [2003]	Dynamic	CKA1	$n/p$	$n$
Chang and Mitzenmacher [2005]	Static	CKA1	$n/p$	$m \cdot n$
Curtmola et al. (SSE-1) [2006]	Static	CKA1	$r$	$m + n$
Curtmola et al. (SSE-1) [2006]	Static	CKA2	$r$	$m \cdot n$
Van Liesdonk et al. [2010]	Dynamic	CKA2	$n$	$m \cdot n$
Chase and Kamara [2010]	Static	CKA2	$r$	$m \cdot n$
Kurosawa and Ohtaki [2012]	Static	UC	$n$	$m \cdot n$
Kamara et al. [2012]	Dynamic	CKA2	$r$	$m + n$
Kamara and Papamanthou [2013]	Dynamic	CKA2	$r/p \cdot \log(n)$	$m \cdot n$
Yavuz and Guajardo [2015]	Dynamic	CKA2	$m/p$	$m \cdot n$

- $n$  = # outsourced data files
- $m$  = #keywords
- $r$  = #documents containing keyword








# Encryption

## Homomorphic Encryption

## Homomorphic Encryption



# Example: RSA (1978)

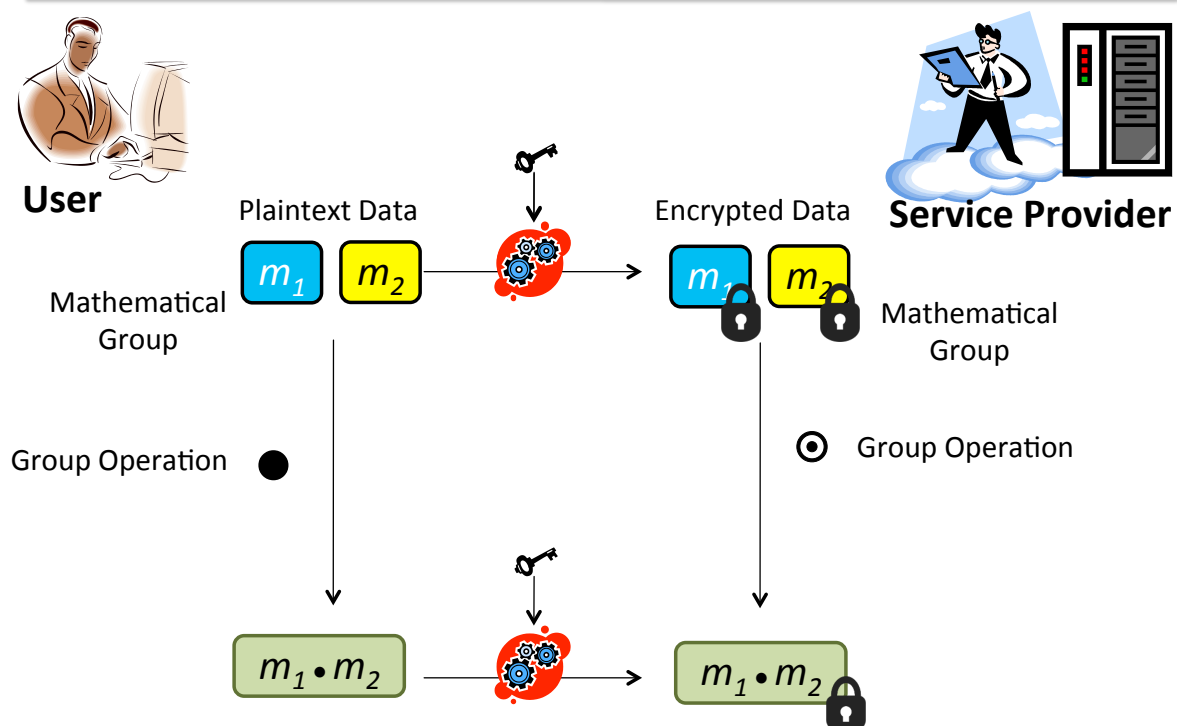
-  **Parameters:**  $N = p \cdot q$  with  $p, q$  large primes (approx. 1000 bits)
-  **Plaintext space:**  $Z_N (= \{0, \dots, N-1\} \text{ modulo } N)$
-  **Ciphertext:**  $Z_N (= \{0, \dots, N-1\} \text{ modulo } N)$
-  **Encryption Key:**  $e \in Z_N$  with  $\gcd(e, (p-1)(q-1)) = 1$
-  **Decryption key:**  $d \in Z_N$  with  $e \cdot d \bmod ((p-1) \cdot (q-1)) = 1$
-  **Encryption of  $m$ :**  $c := m^e \bmod N$
-  **Decryption of  $c$ :**  $c^d \bmod N = m$



Homomorphism:  $(m_1)^e \cdot (m_2)^e = (m_1 \cdot m_2)^e$

$$m_1 \cdot m_2 = m_1 \cdot m_2$$

# Group-Homomorphic Encryption



# Complete Characterization

Armknrecht, Katzenbeisser, Peter; DCC 2013

UNIVERSITÄT  
MANNHEIM

TECHNISCHE  
UNIVERSITÄT  
DARMSTADT

**Definition 3** (GIFT *scheme*) GIFT is a public key encryption scheme  $\mathcal{E}_G = (G, E, D)$  with

Key generation:  $G$  takes a security parameter  $\lambda$  as input and outputs a tuple  $(pk, sk)$  where  $pk$  is the public key that contains descriptions of

- a non-trivial group  $\mathcal{P}$  of plaintexts and a non-trivial group  $\widehat{\mathcal{C}}$  of ciphertexts together with a non-trivial subgroup  $\mathcal{C} \leq \widehat{\mathcal{C}}$  that will act as the set of encryptions
- a non-trivial, proper normal subgroup  $\mathcal{N}$  of  $\mathcal{C}$  such that  $|\mathcal{C}/\mathcal{N}| = |\mathcal{P}|$
- an efficient isomorphism  $\varphi : \mathcal{P} \rightarrow \mathcal{R}$  where  $\mathcal{R} \subseteq \mathcal{C}$  (not necessarily a subgroup but certainly a group, cf. Remark 1) is a system of representatives of  $\mathcal{C}/\mathcal{N}$ ,

and  $sk$  is the secret key that contains

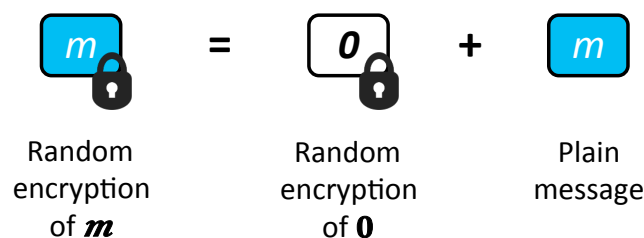
- an efficient description of  $\varphi^{-1} \circ \nu$  with the epimorphism  $\nu : \mathcal{C} \rightarrow \mathcal{R}$  such that  $\nu(c)$  is the unique representative  $r \in \mathcal{R}$  with  $c = r \cdot n$  for some  $n \in \mathcal{N}$ .
- an efficient function  $\delta : \widehat{\mathcal{C}} \rightarrow \{0, 1\}$  such that  $\delta(c) = 1 \iff c \in \mathcal{C}$ .

Encryption:  $E$  takes the public key  $pk$  and a message  $m \in \mathcal{P}$  as input and outputs the ciphertext  $c := \varphi(m) \cdot n \in \mathcal{C}$  where  $n \leftarrow \mathcal{N}$ .

Decryption:  $D$  takes the secret key  $sk$  and a ciphertext  $c \in \widehat{\mathcal{C}}$  as input. If  $\delta(c) = 0$ , it outputs  $\perp$ , otherwise it outputs the plaintext  $\varphi^{-1}(\nu(c)) \in \mathcal{P}$ .

## Consequences

- **Encryption format:**



- **Application of the framework**

- Security analysis
- New designs

- **Group-case well understood**

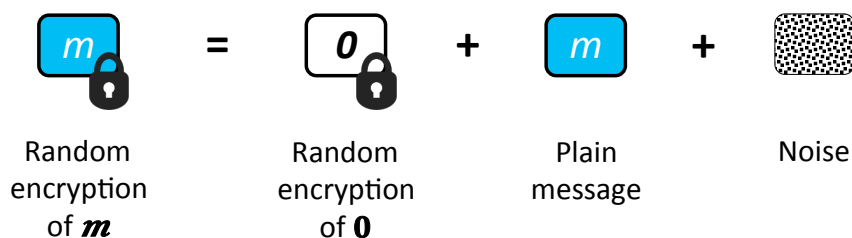
- **Beyond groups?**

# Gentry's Breakthrough Result (2009)



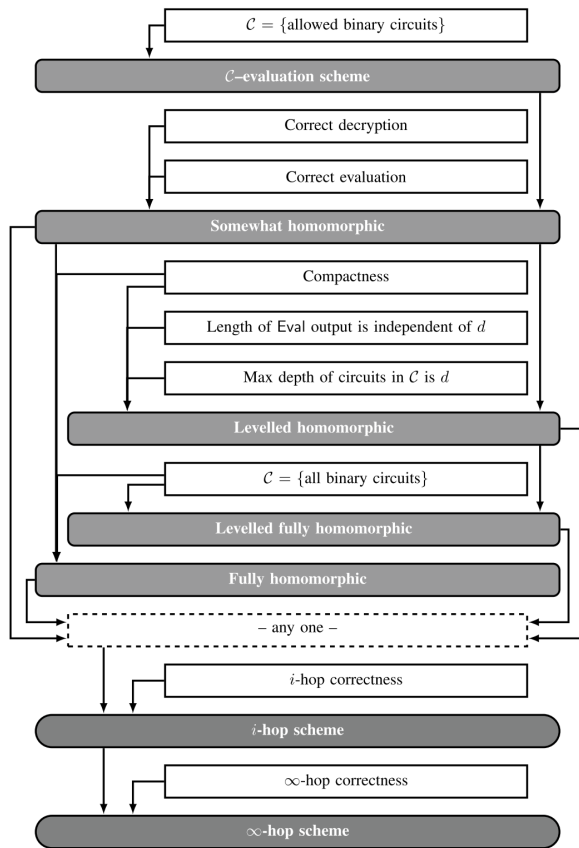
## High Level View

- **Encryption format:**



- **Noise  $\Rightarrow$  higher security**
- **Combinations of ciphertexts increase noise**
- **Main challenge: dealing with noise**
- **Also: theory more involved**

# Definition Jungle



# State of the Art

Scheme	Underlying Problems	Asymptotic Runtime	Concrete Instantiation Runtime
Gentry: A Fully Homomorphic Encryption Scheme [18]	BDDP & SSSP	$\mathcal{O}(\lambda^6 \log(\lambda))$ per gate	-
van Dijk, Gentry, Halevi, Vaikuntanathan: FHE over the Integers [35]	AGCD & SSSP	$\mathcal{O}(\lambda^{10})$	-
Coron, Naccache, Tibouchi: Public Key Compression and Modulus Switsching for FHE over the Integers [13]	DAGCD & SSSP	-	Recryption (a step that takes place after every addition/multiplication) takes about 11 minutes.
Brakerski, Vaikuntanathan: Efficient FHE from (standard) LWE [9]	DLWE	$\tilde{\mathcal{O}}(\lambda^{2^C})$ where $C$ is a very large parameter that ensures bootstrappability.	-
Brakerski, Vaikuntanathan: FHE from Ring-LWE and Security for Key Dependent Messages [10]	PLWE	-	-
Brakerski, Gentry, Vaikuntanathan: FHE without Bootstrapping [8]	RLWE	Per-gate computation overhead $\tilde{\mathcal{O}}(\lambda \cdot d^3)$ (where $d$ is the depth of the circuit) without bootstrapping, $\tilde{\mathcal{O}}(\lambda^2)$ with bootstrapping.	In [21]: 36 hours for an AES encryption on a supercomputer
Smart, Vercauteren: FHE with Relatively Small Key and Ciphertext Sizes [34]	PCP & SSSP	-	Key generation took several hours even for small parameters which do not deliver a fully homomorphic scheme, for larger parameters the keys could not be generated
Rohloff, Cousins: A Scalable Implementation of Fully Homomorphic Encryption Built on NTRU [32]	SVP & RLWE	-	Recryption at 275 seconds on 20 cores with 64-bit security
Halevi, Shoup: Bootstrapping for HELib [27]	RLWE	-	Vectors of 1024 elements from $\text{GF}(2^{16})$ was recrypted in 5.5 minutes at security level $\approx 76$ , single CPU core.

## DARPA spends \$20 million on homomorphic encryption

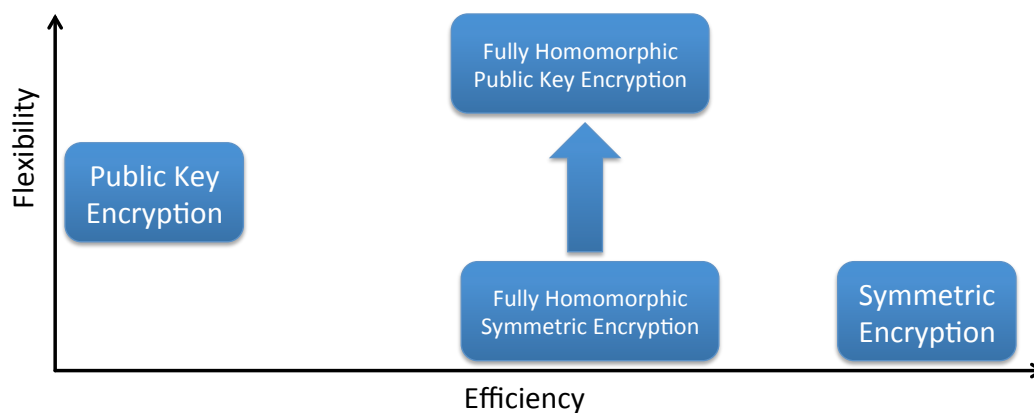
Written by Alex Armstrong

Tuesday, 19 April 2011 09:33

The US military research agency has awarded almost \$5 million to speed the performance of an algorithm that could make cloud computing secure.

Now the Defense Advanced Research Projects Agency (DARPA), an agency of the United States Department of Defense responsible for the development of new technology for use by the military, has awarded \$4.9 million to research contractor, [Galois Inc](#), to turn the algorithm into something practical. This is part of a larger project funded to the tune of \$20 million called Programming Computation on Encrypted Data or PROCEED (presumably the term homomorphic is too technical). The goal of the project is to speed up the algorithm by a factor of 10 million - which is clearly not an easy optimisation factor to achieve.

## What can we expect?





## Personal Opinion

---

- Unlikely to see efficient fully-homomorphic encryption
- Counter-question: do we need fully-homomorphism in practice?
  - Examples exist where a scheme with less functionalities would be sufficient
  - Adapted homomorphic encryption schemes
- Potential for specific use cases

## *Data Loss*

# Second-Most Significant Risk

CLOUD SECURITY ALLIANCE The Notorious Nine: Cloud Computing Top Threats in 2013

## 2.0 Top Threat: Data Loss

For both consumers and businesses, the prospect of permanently losing one's data is terrifying. Just ask Mat Honan, writer for Wired magazine: in the summer of 2012, attackers broke into Mat's Apple, Gmail and Twitter accounts. They then used that access to erase all of his personal data in those accounts, including all of the baby pictures Mat had taken of his 18-month-old daughter.

Of course, data stored in the cloud can be lost due to reasons other than malicious attackers. Any accidental deletion by the cloud service provider, or worse, a physical catastrophe such as a fire or earthquake, could lead to the permanent loss of customers' data unless the provider takes adequate measures to backup data. Furthermore, the burden of avoiding data loss does not fall solely on the provider's shoulders. If a customer encrypts his or her data before uploading it to the cloud, but loses the encryption key, the data will be lost as well.

### SERVICE MODEL

IaaS

PaaS

SaaS

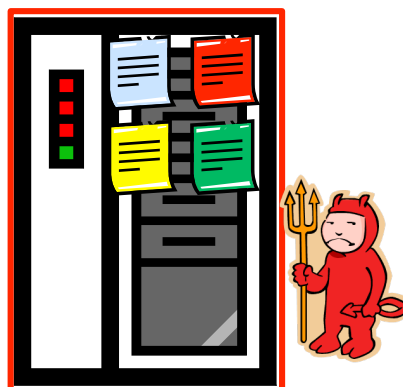
### RISK MATRIX



### RISK ANALYSIS

## Scenario

- (Possibly encrypted) data outsourced
- User worries: data loss



# Amazon S3

## Reduced Redundancy Storage (RRS)

### Q: What is RRS?

Reduced Redundancy Storage (RRS) is a new storage option within Amazon S3 that enables customers to reduce their costs by storing non-critical, reproducible data at lower levels of redundancy than Amazon S3's standard storage. RRS provides a lower cost, less durable, highly available storage option that is designed to sustain the loss of data in a single facility.

### Q: Why would I choose to use RRS?

RRS is ideal for **non-critical or reproducible data**. For example, RRS is a **cost-effective** solution for sharing media content that is durably stored elsewhere. RRS also makes sense if you are storing thumbnails and other resized images that can be easily reproduced from an original image.

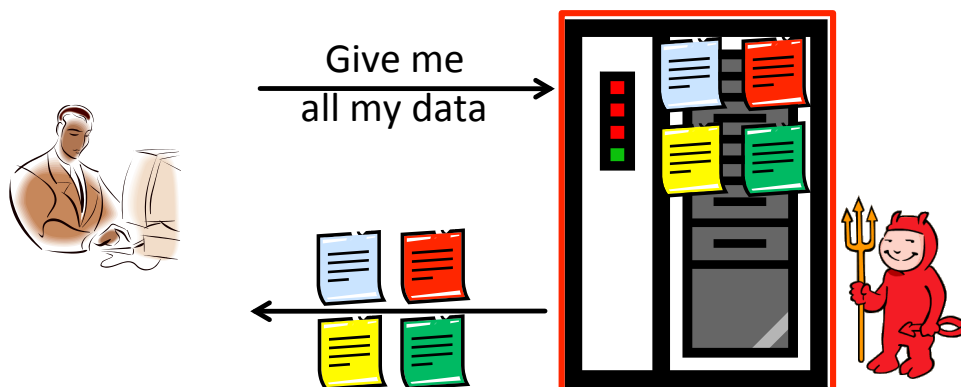
### Q: What is the durability of Amazon S3 when using RRS?

RRS is designed to provide 99.99% durability of objects over a given year. This durability level corresponds to an average annual expected loss of 0.01% of objects. For example, if you store 10,000 objects using the RRS option, you can on average expect to incur an annual loss of a single object (i.e. 0.01% of 10,000 objects). This annual loss represents an expected average and does not guarantee the loss of 0.01% of objects in a given year.

The RRS option stores objects on multiple devices across multiple facilities, providing 400 times the durability of a typical disk drive, but does not replicate objects as many times as standard Amazon S3 storage, and thus is even more cost effective. In addition, RRS is designed to sustain the loss of data in a single facility.

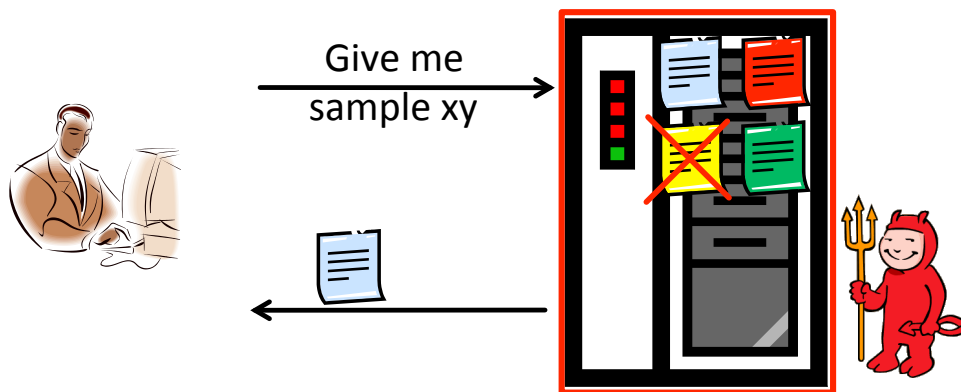
## Stupid “Solution” 1

- Download complete data from time to time
- Inefficient
- How to check?



## Stupid “Solution” 2

- Ask for random samples from time to time
- Inefficient
- How to check?
- Security

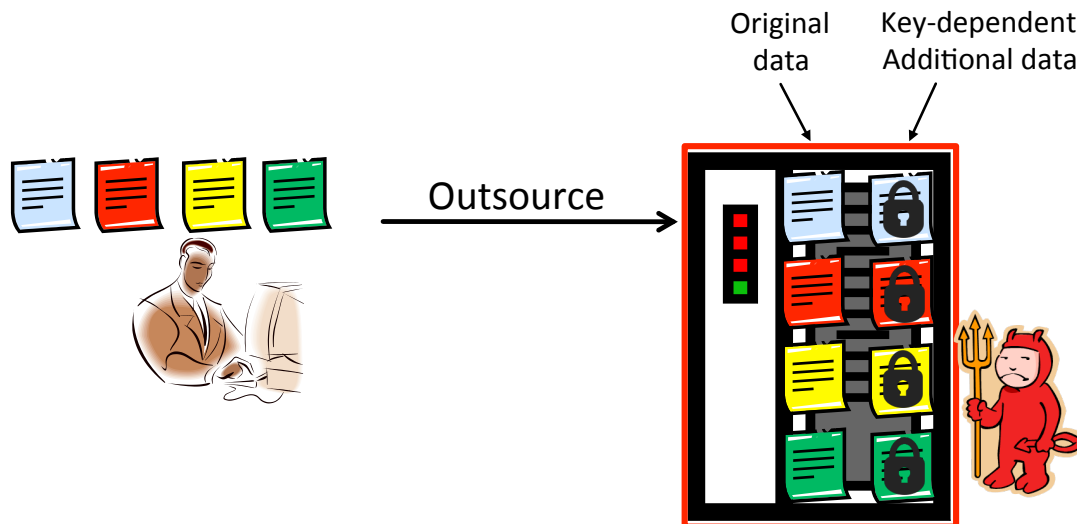


## Proof of Retrievability (POR)

- Cryptographic protocol between user and service provider
- Goal: Ensure that outsourced file is still intact and extractable
- Attacker model: service provider rational (aims to reduce costs)
- Of high academic interest, e.g.,
  - Ateniese et al. (CCS 2007)
  - Juels et al. (CCS 2007)
  - Shacham, Waters (Asiacrypt 2008)
  - Bowers et al. (CCS 2009)
  - Dodis et al. (TCC 2009)
  - Erway et al. (CCS 2009)
  - Bowers et al. (CCS 2011)
  - Shi et al. (CCS 2013)
  - Armknecht et al. (CCS 2014)

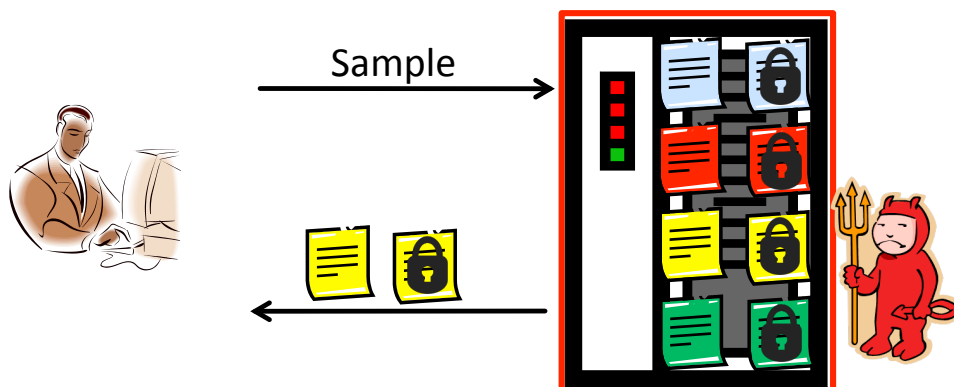
# POR - Setup

- User preprocesses data, using some key
- Additional data used in verification step (next slide)



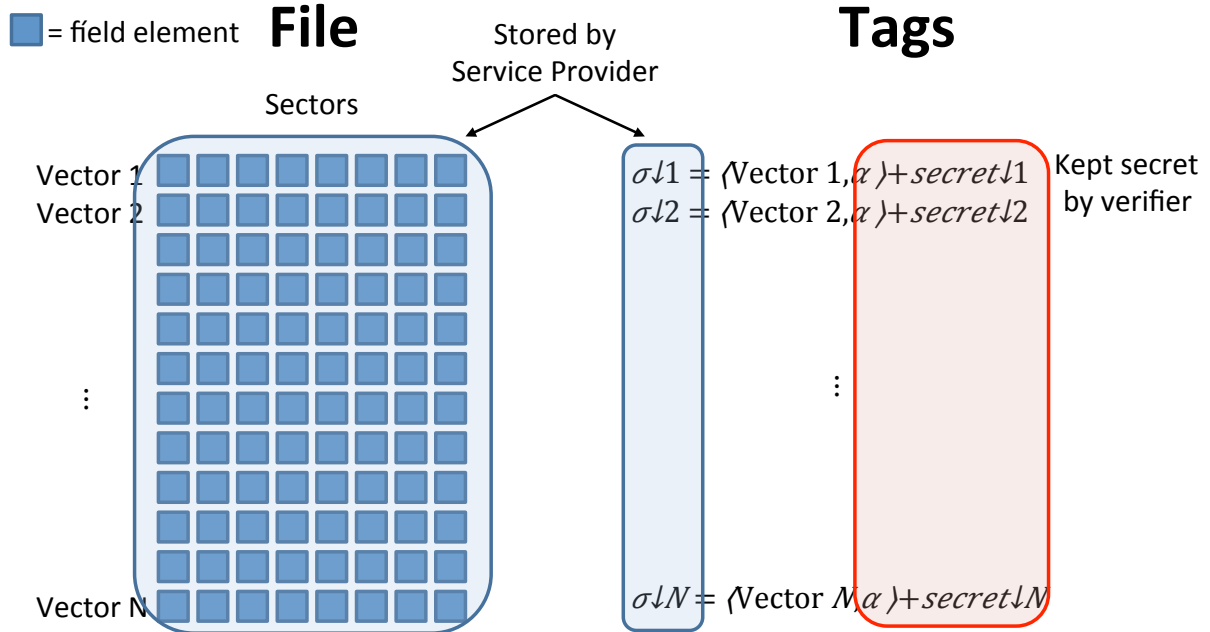
# POR - Verification

- From time to time, user requests for a sample
- Sample includes original data and additional key-dependent data
- Check data for completeness and correctness
- Original data can be reconstructed from samples



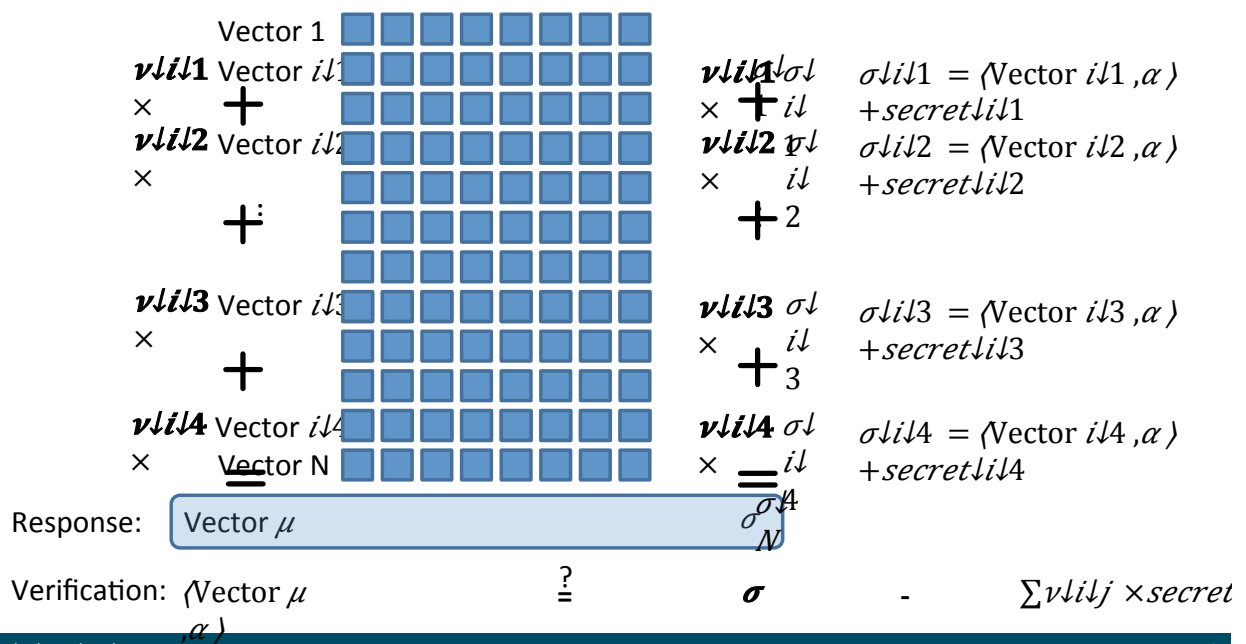
# SW-POR - Setup

H. Shacham, B. Waters (Asiacrypt'08)



# SW-POR - Verify

Challenge: Set of indices and values  $Q = \{(i, v_i)\} \mid i \in I$



# Motivation

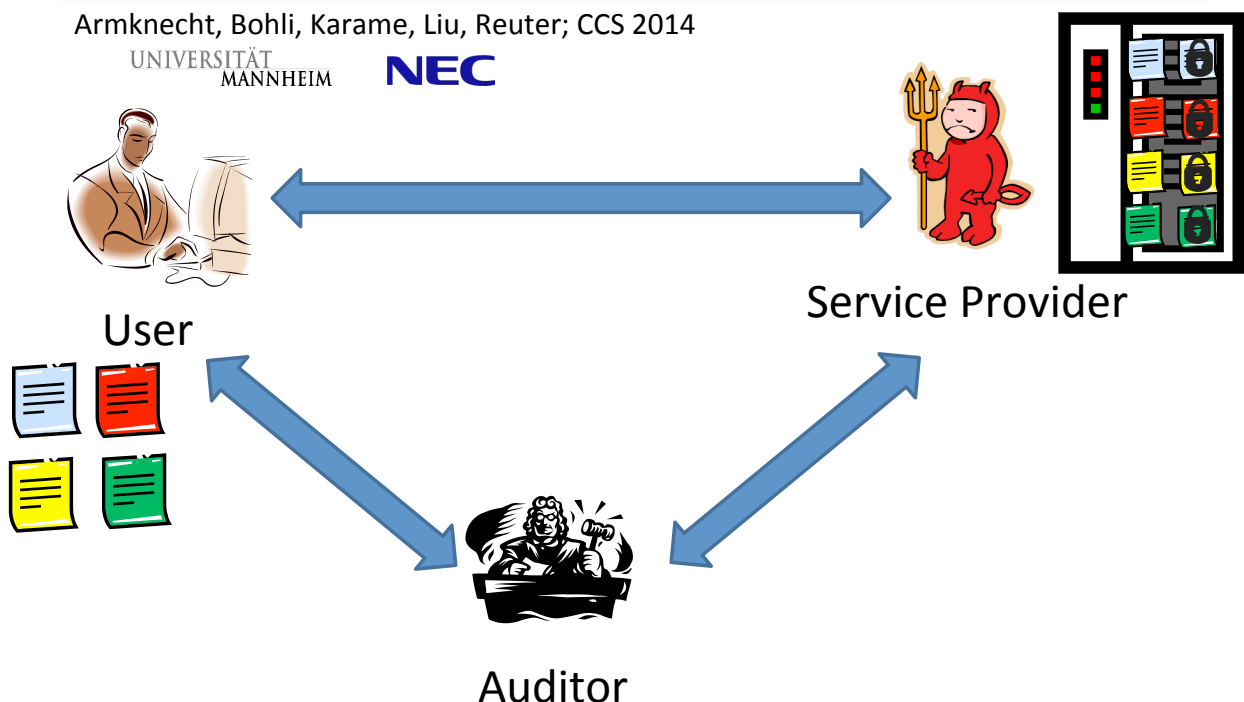
- Several POR (and related Proof-of-Data-Possession (PDP) do exit
- Main drawback for practical application: requires regular involvement of user
- Although Storage as a Service is of high practical relevance, POR hardly used
- Need to solve: POR as a Service

## Outsourced POR (OPOR)

Armknecht, Bohli, Karame, Liu, Reuter; CCS 2014

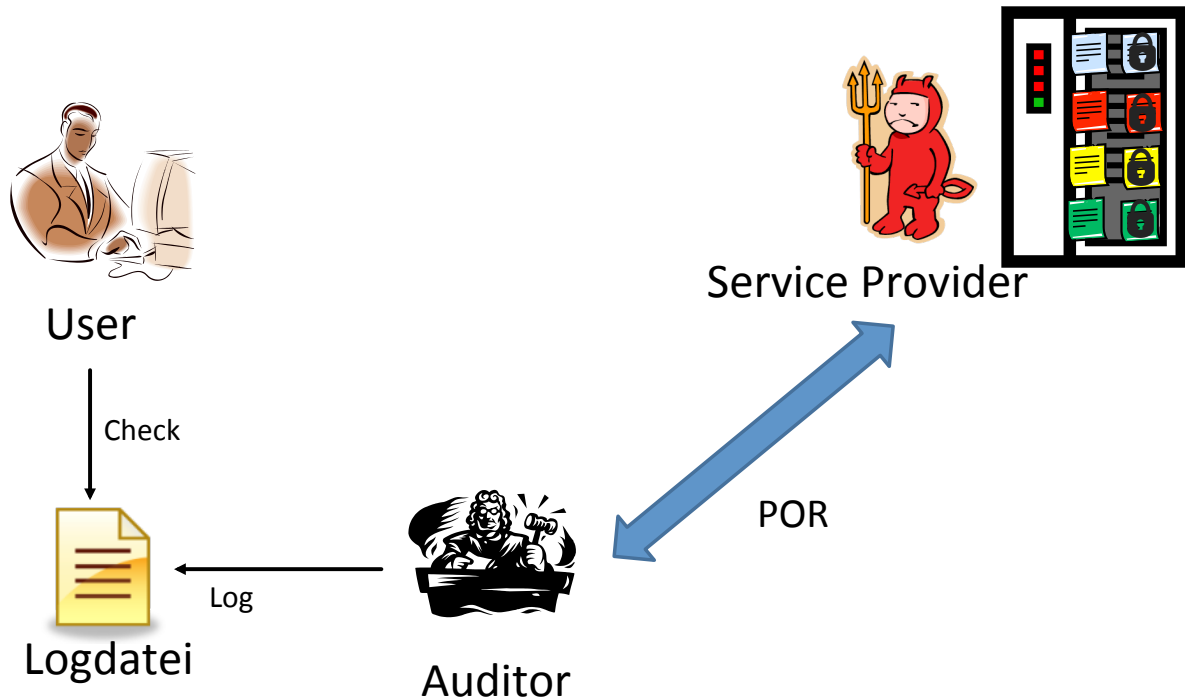
UNIVERSITÄT  
MANNHEIM

NEC

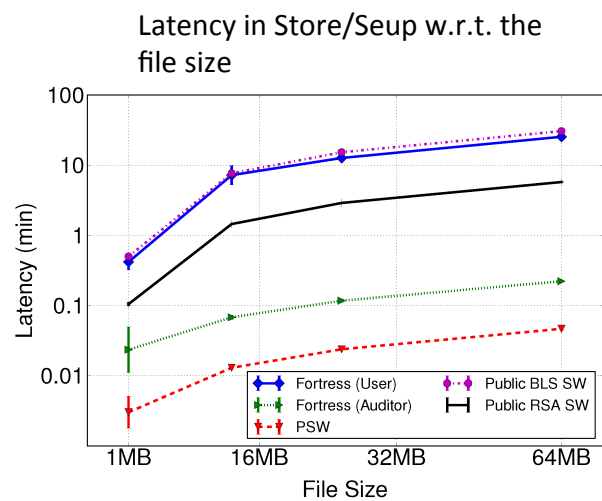
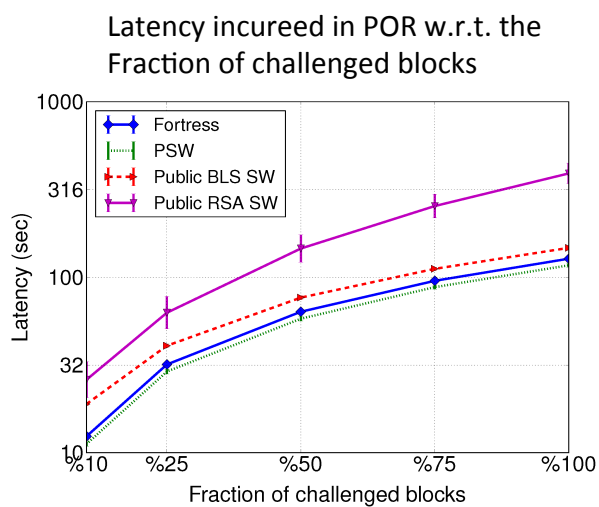




# Outsourced POR (OPOR)



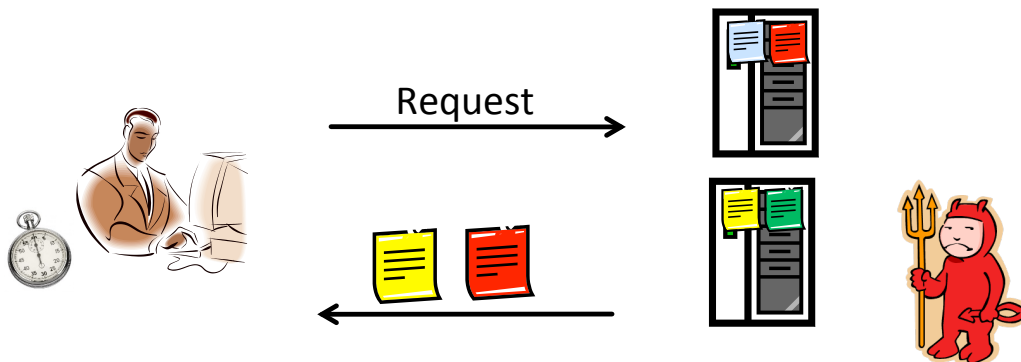
## Fortress - Performance



# Proof of Redundancy

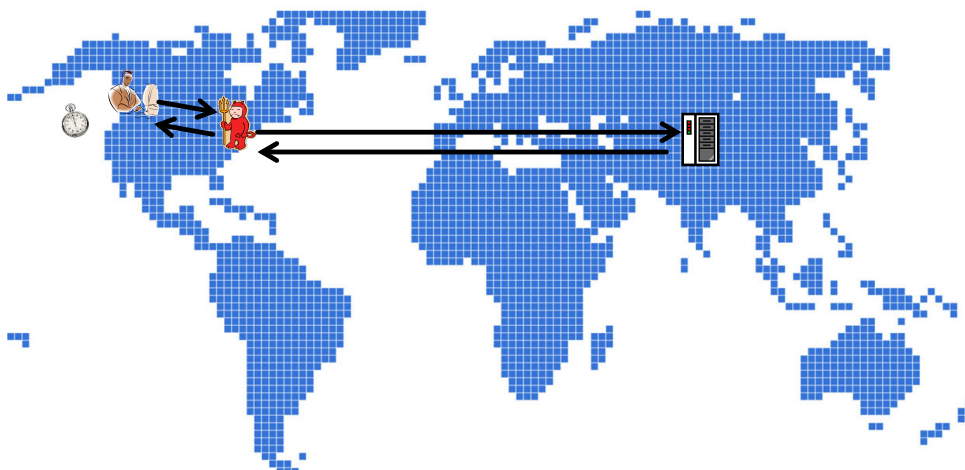
Bowers et al. [CCS 2011]

- **POR are possibly “too late”**
- **Better: check if storage is sufficiently fault tolerant, i.e., distributed over different hard drives**
- **Idea: Make requests for different hard drives, check response time**



# Proof of Location

- **Also important: WHERE is the data stored**
- **Similar idea: measure response time**



# ***Conclusions***

## **Summary**

---

- **Security concerns with respect to cloud computing**
- **Traditional mechanisms obviously not sufficient**
- **New cryptographic techniques may help**

# Thank You!

---

