Segurança na Nuvem

da Confidencialidade à Disponibilidade dos Dados

Miguel Correia

Trabalho conjunto com Alysson Bessani, Francisco Rocha,
P. Sousa, B. Quaresma, F. André, S. Abreu

Jornadas Técnicas de Computação em Nuvem
ESTG-IPL - Leiria, Março 2012

INSTITUTO SUPERIOR TÉCNICO

FCT Fundação para a Ciência e a Tecnologia

inesc id lisboa

Cloud computing in a nutshell

- Computing as a utility
- Public cloud – cloud service provider (CSP) different from the cloud user (typ. a company)
- Pay-per-use / pay-as-you-go
- Resource pooling / multi-tenancy
- Elasticity
- Large-scale datacenters

Microsoft's Chicago datacenter

Cloud computing service models

- What is the service provided by the cloud?
- Infrastructure as a Service (IaaS): virtual machines, storage (e.g., Amazon EC2, Amazon S3)
- Platform as a Service (PaaS): programming and execution (e.g., Google AppEngine, Force.com, Windows Azure)
- Software as a Service (SaaS): mostly web applications (e.g., Yahoo! Mail, Google Docs, Facebook,…)

Security of what from whom?

- Victim is not the user or cloud; cloud is attack tool
  - User (bad) uses the cloud (good) to attack others
  - SPAM, DDoS, hosting malicious data, botnet C&C
- Victim is the cloud
  - User or someone else (bad) attack the cloud (good)
- Victim is the user
  - Cloud insider or another user (bad) attacks the user (good)

Security in the cloud
(from the cloud user viewpoint)

- Security is a key aspect of cloud computing
  - Factor that favors and prevents adoption
  - That’s how it should be!
- Challenges
  - The system is no longer in the user premises
  - The infrastructure is shared with other users
  - The access is made through the internet
- The three classical security attributes can be jeopardized: confidentiality, integrity, availability

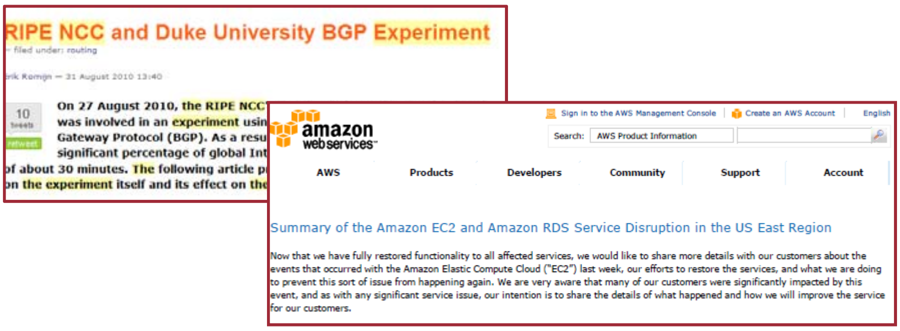
Outline

- Security and dependability threats in the cloud
- Stealing data in the cloud
- Approach 1: improve the IaaS cloud infrastructure
- Approach 2: build a storage cloud-of-clouds
- Conclusions

SECURITY AND DEPENDABILITY THREATS IN THE CLOUD

Unavailability

- Problems in the Internet are relatively frequent
  - Congestion, problems with routers / switches / links,…
  - Routing problems (Cisco bug + RIPE NCC test Aug.’10)
- Problems at the cloud (e.g., Amazon EC2 outage May’11)



RIPE NCC and Duke University BGP Experiment

– filed under: routing

rik Romijn — 31 August 2010 13:40

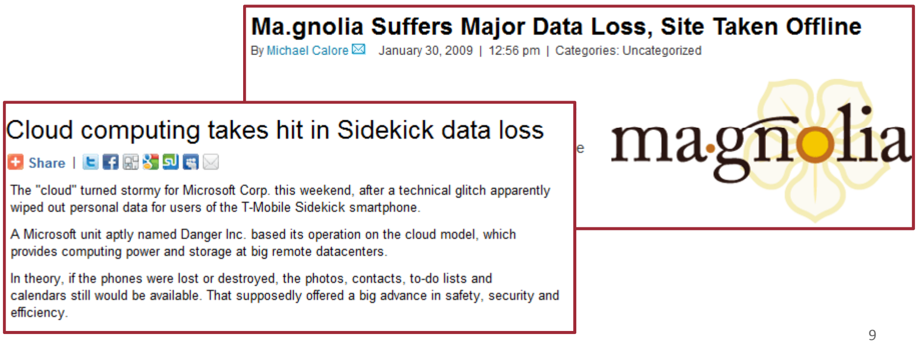
On 27 August 2010, the RIPE NCC
was involved in an experiment usin
Gateway Protocol (BGP). As a resu
significant percentage of global Int
of about 30 minutes. The following article p
on the experiment itself and its effect on the

Summary of the Amazon EC2 and Amazon RDS Service Disruption in the US East Region

Now that we have fully restored functionality to all affected services, we would like to share more details with our customers about the events that occurred with the Amazon Elastic Compute Cloud ("EC2") last week, our efforts to restore the services, and what we are doing to prevent this sort of issue from happening again. We are very aware that many of our customers were significantly impacted by this event, and as with any significant service issue, our intention is to share the details of what happened and how we will improve the service for our customers.

Loss and corruption of data

- Can happen in the cloud as anywhere else
  - Ma.gnolia lost all users’ data, half TB (Feb.’09)
  - Danger Inc. / Sideckick lost contacts, notes, photos, etc. of its clients; took days to recover (Oct.’09)



Ma.gnolia Suffers Major Data Loss, Site Taken Offline

By Michael Calore | January 30, 2009 | 12:56 pm | Categories: Uncategorized

Cloud computing takes hit in Sidekick data loss

The "cloud" turned stormy for Microsoft Corp. this weekend, after a technical glitch apparently wiped out personal data for users of the T-Mobile Sidekick smartphone.

A Microsoft unit aptly named Danger Inc. based its operation on the cloud model, which provides computing power and storage at big remote datacenters.

In theory, if the phones were lost or destroyed, the photos, contacts, to-do lists and calendars still would be available. That supposedly offered a big advance in safety, security and efficiency.

Attacks through management interface

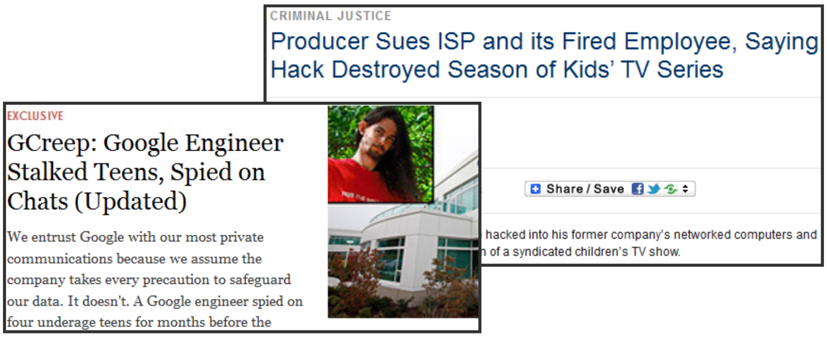
- Cloud users have access to management interfaces
  - Operations: control/monitor virtual machines, users,…
  - Interfaces: web console, web services, REST
- Personification attacks through the interface
  - The usual culprits: CSRF, SQL injection, XSS, XML Signature Wrapping (recently found possible in EC2 and Eucalyptus)
- Phishing / social engineering attacks to obtain authentication credentials

Attacks between VMs

- In IaaS, VMs of several users usually share the same physical machine – co-residence
- Attack in two steps
  - Attacker instantiates several VMs until co-residence with the victim is achieved
  - The attacker's VM attacks the victim, e.g., using a vulnerability in the hypervisor or using shared resources to obtain confidential information

Malicious insider and confidentiality

- The data is in the cloud and the malicious insider is a real problem
  - CyberLynk (March'09) and Google (early'10) events



CRIMINAL JUSTICE

Producer Sues ISP and its Fired Employee, Saying Hack Destroyed Season of Kids' TV Series

Share / Save

hacked into his former company's networked computers and

n of a syndicated children's TV show.

EXCLUSIVE

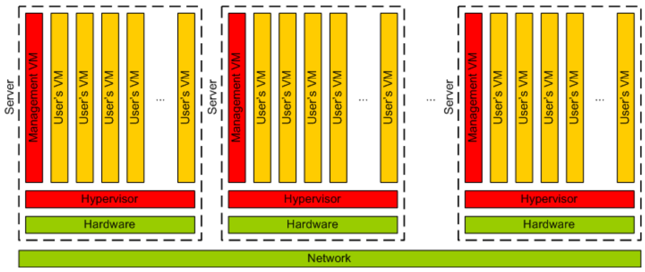
GCreep: Google Engineer Stalked Teens, Spied on Chats (Updated)

We entrust Google with our most private communications because we assume the company takes every precaution to safeguard our data. It doesn't. A Google engineer spied on four underage teens for months before the

STEALING DATA IN THE CLOUD

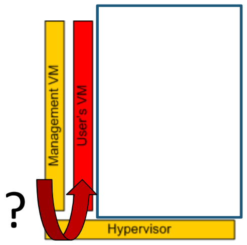
Infrastructure as a Service (VM)

- Service provided is the execution of Virtual Machines (VMs) – full software stack, including OS
- Servers run an Hypervisor (or VMM) that supports the execution of several VMs



Experimental environment

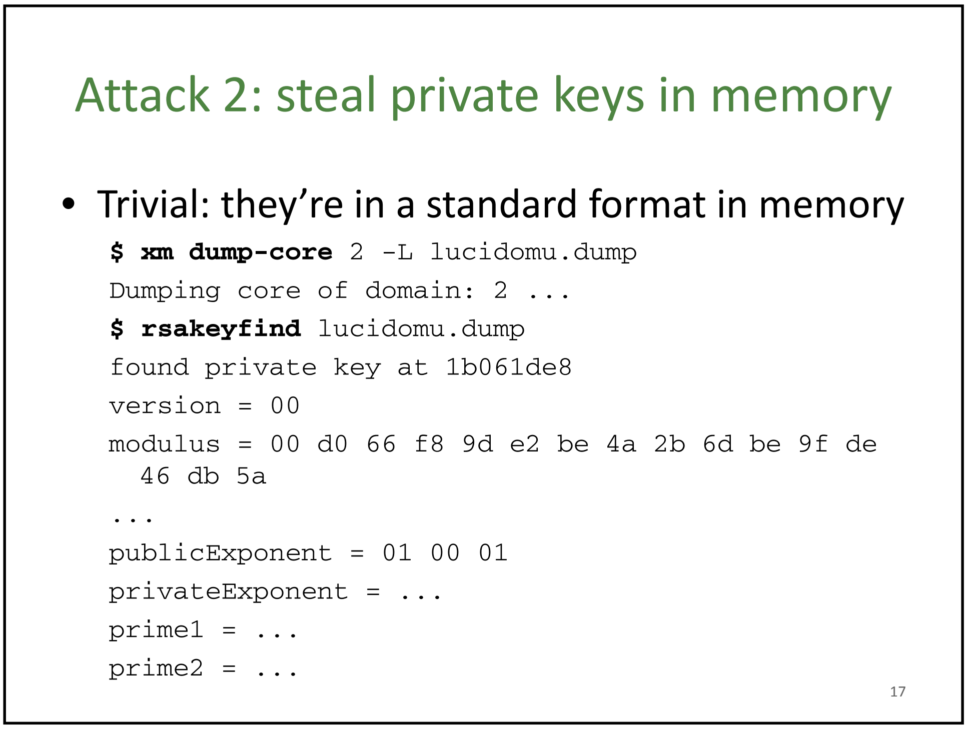
- We played the role of a malicious insider with access to the management VM
- The “cloud” was just a single machine
  - Hypervisor was Xen
  - Management VM was Xen Dom 0 with Linux
  - 1 user VM (victim) with Linux and an Apache server
  - Malicious insider had login in Dom 0



Attack 1: steal passwords in memory

- Trivial: take mem snapshot, look for passwords

```
$ xm dump-core 2 -L lucidomu.dump
Dumping core of domain: 2 ...
$ cat lucidomu.dump | strings | grep loginpwd
loginpwd
loginpwd
$ cat lucidomu.dump | strings | grep
  apachersapwd
apachersapwd
apachersapwd
apachersapwd
```

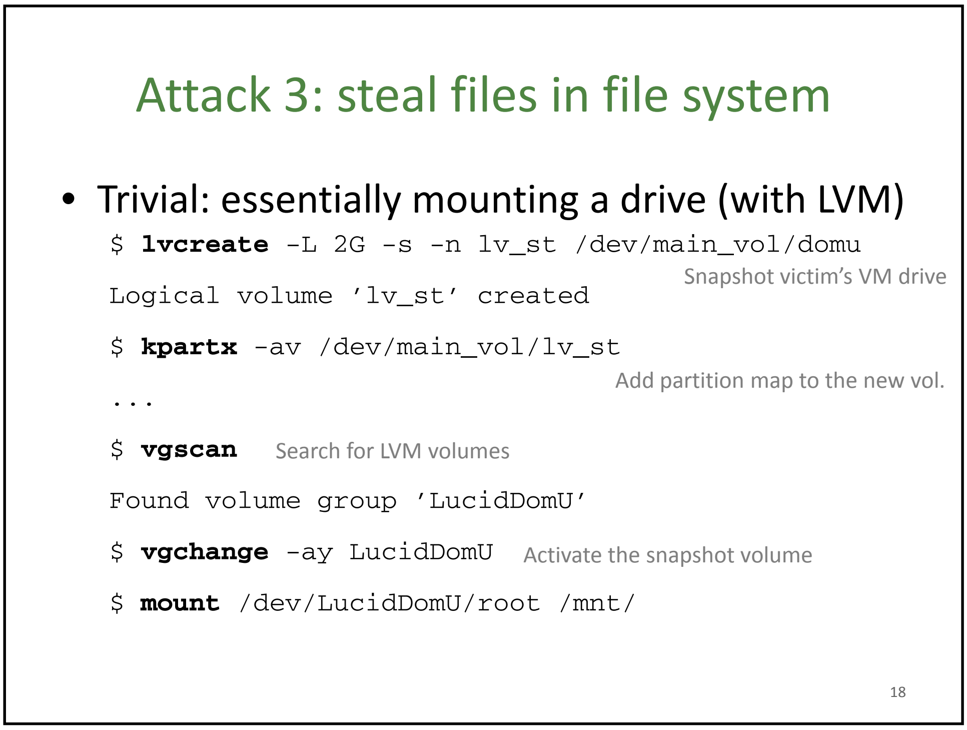


Attack 2: steal private keys in memory

- Trivial: they're in a standard format in memory

```
$ xm dump-core 2 -L lucidomu.dump
Dumping core of domain: 2 ...
$ rsakeyfind lucidomu.dump
found private key at 1b061de8
version = 00
modulus = 00 d0 66 f8 9d e2 be 4a 2b 6d be 9f de
  46 db 5a
...
publicExponent = 01 00 01
privateExponent = ...
prime1 = ...
prime2 = ...
```

17



Attack 3: steal files in file system

- Trivial: essentially mounting a drive (with LVM)

```
$ lvcreate -L 2G -s -n lv_st /dev/main_vol/domu
                                    Snapshot victim's VM drive
Logical volume 'lv_st' created
$ kpartx -av /dev/main_vol/lv_st
                                Add partition map to the new vol.
...
$ vgscan    Search for LVM volumes
Found volume group 'LucidDomU'
$ vgchange -ay LucidDomU   Activate the snapshot volume
$ mount /dev/LucidDomU/root /mnt/
```

18

Current solutions?

- From “Cloud Computing Roundtable” (Nov/Dec 2010)
  - 5 directors/senior staff from: Google, Microsoft, Cisco, Amazon, Cloud Security Alliance
- “We have very strict procedures in place for when our employees are allowed to [physically] access the machines the customer data resides on.”
  - Excellent, but the attacks we saw can be done remotely
- “We keep track of every action that they take on those machines, and we log all that information for later audits”
  - Excellent, but detecting in later audits is usually too late

“there’re some things that will never go into [our cloud], for example, our SAP back end”

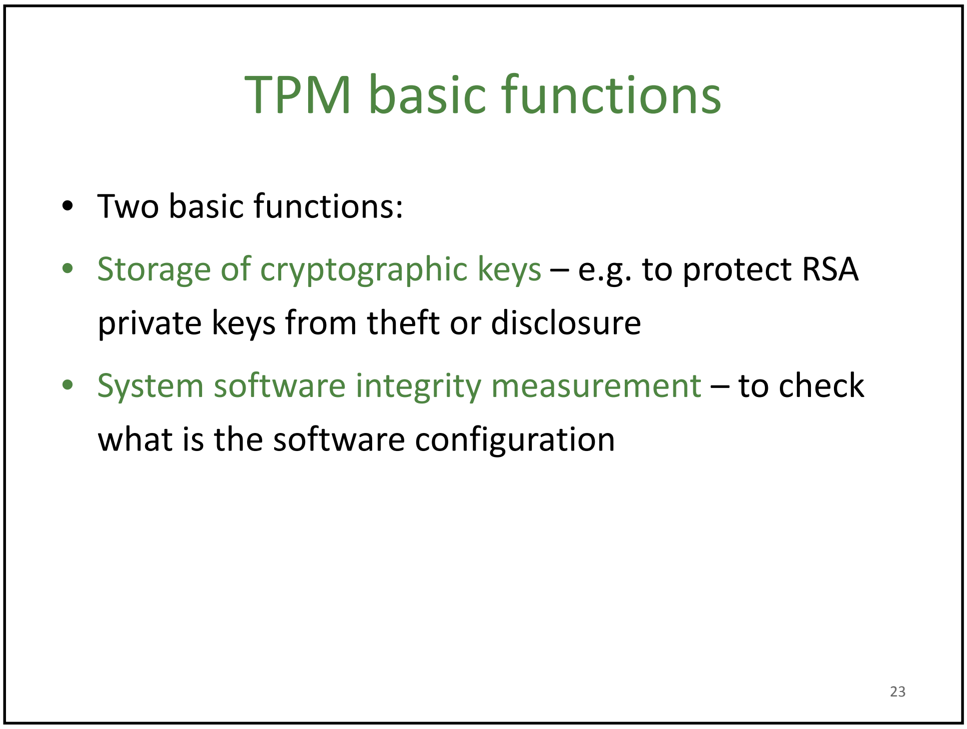
Cryptography?

- Obvious solution: simply encrypt the data
- But what is data in IaaS?
  - User files, web pages, databases, program variables, etc.
  - Is it possible to modify applications to handle encrypted data? An application server (Tomcat, JBoss,…)?
  - Where do we store the encryption keys safely?
- Some applications manipulate data
  - Arithmetic w/encrypted data: fully homomorphic encrypt.
  - Slow, doesn’t work if data encrypted with different keys, application server has also to be modified

APPROACH 1: IMPROVE THE IAAS CLOUD INFRASTRUCTURE

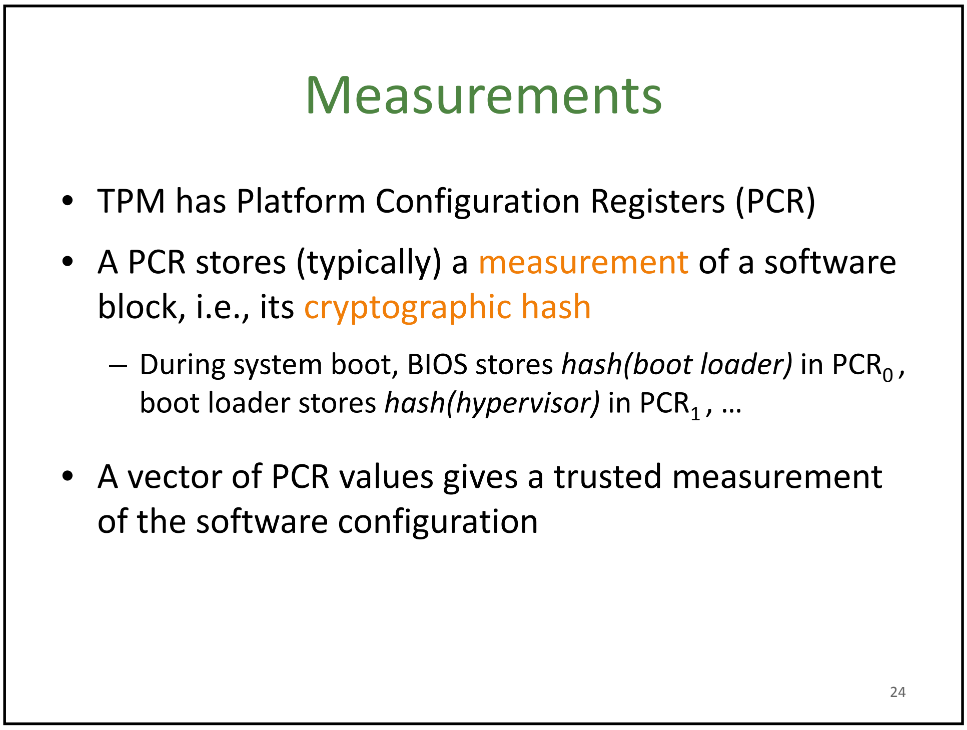
Key idea

- To prove to the cloud user that its data is in a server with a “good” software configuration (e.g., in which the management VM has no snapshot function)
- Do this using the Trusted Platform Module (TPM), a security chip designed by the Trusted Computing Group, now shipping with common PC hardware



TPM basic functions

- Two basic functions:
- Storage of cryptographic keys – e.g. to protect RSA private keys from theft or disclosure
- System software integrity measurement – to check what is the software configuration

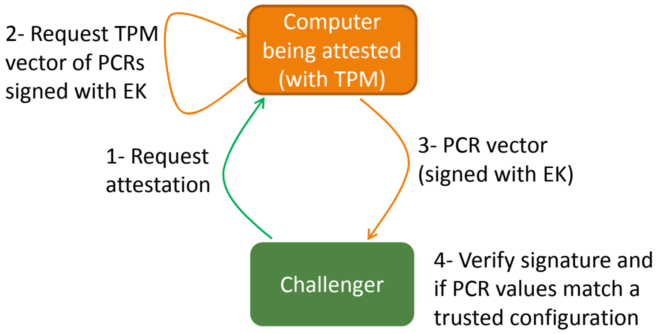


Measurements

- TPM has Platform Configuration Registers (PCR)
- A PCR stores (typically) a measurement of a software block, i.e., its cryptographic hash
  - During system boot, BIOS stores *hash(boot loader)* in PCR_0 , boot loader stores *hash(hypervisor)* in PCR_1 , …
- A vector of PCR values gives a trusted measurement of the software configuration

Remote attestation

- Computer gives to a challenger a measurement of the software configuration, i.e., a vector of PCR values
  - Challenger has the Endorsement Key Certificate, signed by the TPM vendor (means it's a real TPM)



Computer being attested (with TPM)

Challenger

1- Request attestation

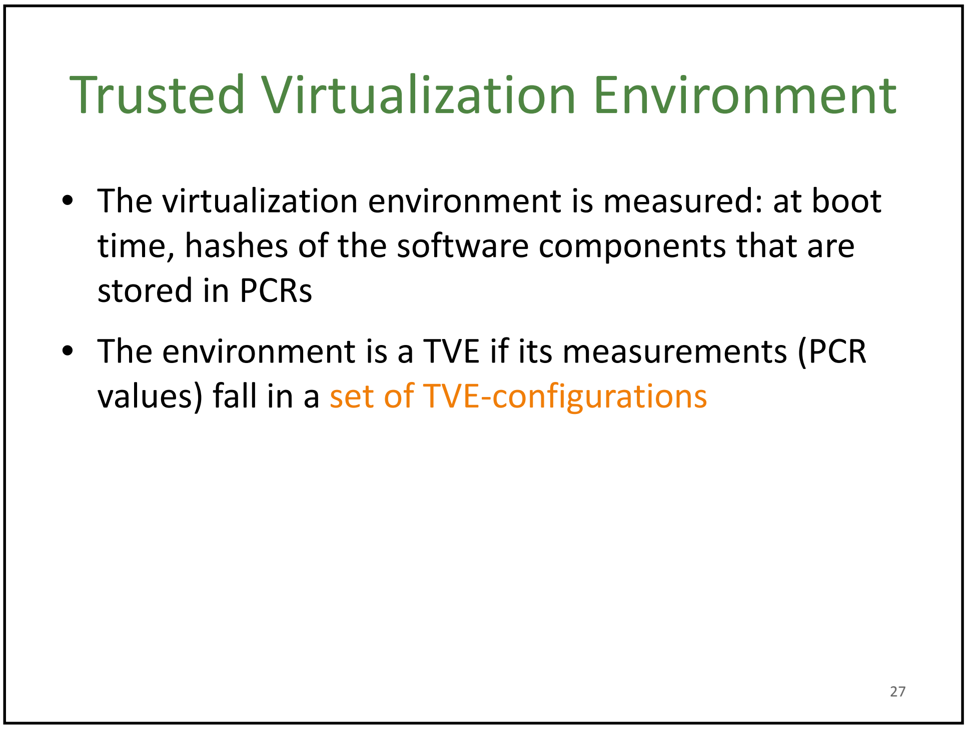
2- Request TPM vector of PCRs signed with EK

3- PCR vector (signed with EK)

4- Verify signature and if PCR values match a trusted configuration

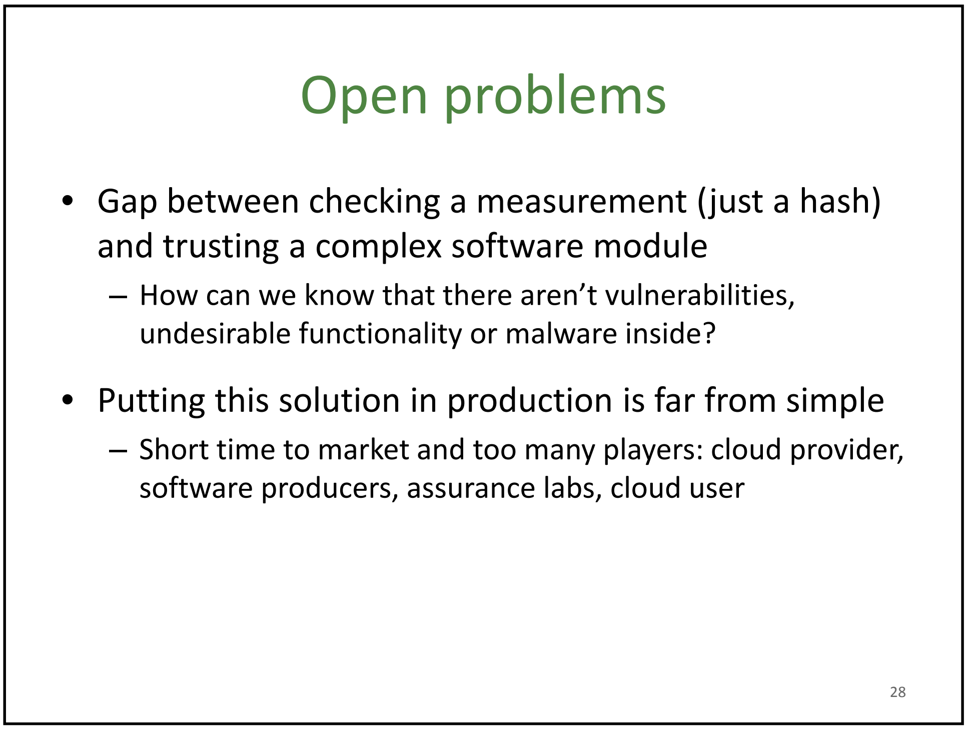
Approach overview

- Servers run a Trusted Virtualization Environment (TVE), formed by hypervisor + management VM that the user trusts
- TVE does not provide dangerous operations to administrators: snapshot, volume mount
- TVE provides only trusted versions of certain operations: launch, migrate, backup, terminate VMs
- VMs enter and leave a TVE encrypted
- Users do remote attestation of TVEs/operations to be sure that their VMs are either in a TVE or encrypted



Trusted Virtualization Environment

- The virtualization environment is measured: at boot time, hashes of the software components that are stored in PCRs
- The environment is a TVE if its measurements (PCR values) fall in a set of TVE-configurations



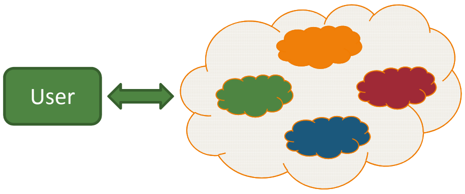
Open problems

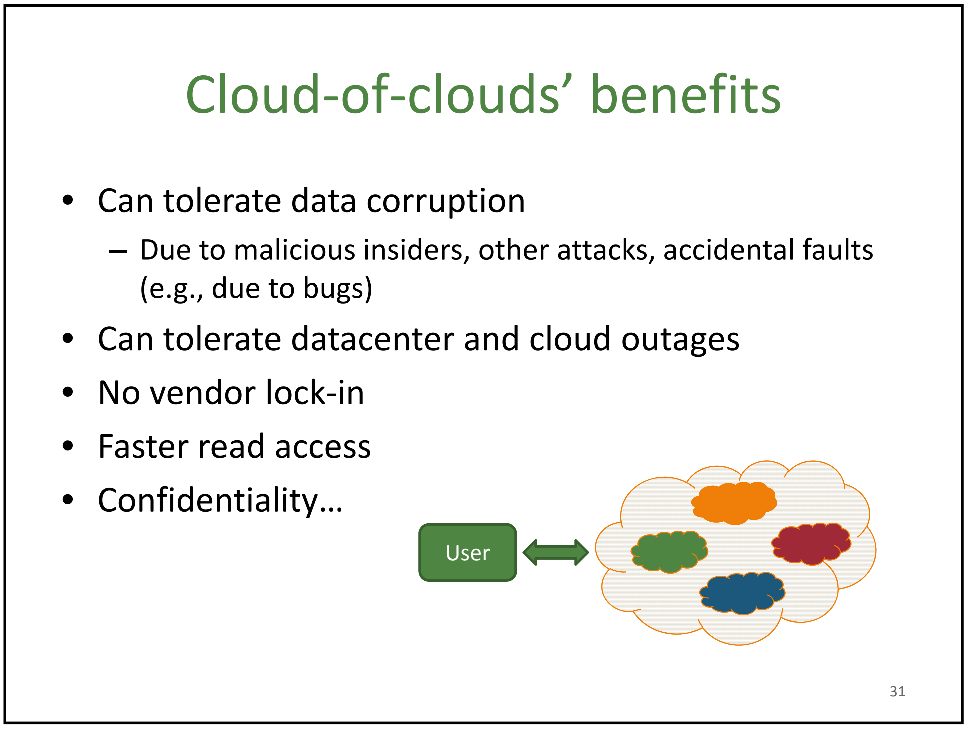
- Gap between checking a measurement (just a hash) and trusting a complex software module
  - How can we know that there aren’t vulnerabilities, undesirable functionality or malware inside?
- Putting this solution in production is far from simple
  - Short time to market and too many players: cloud provider, software producers, assurance labs, cloud user

APPROACH 2: BUILD A STORAGE CLOUD-OF-CLOUDS

Securing the cloud

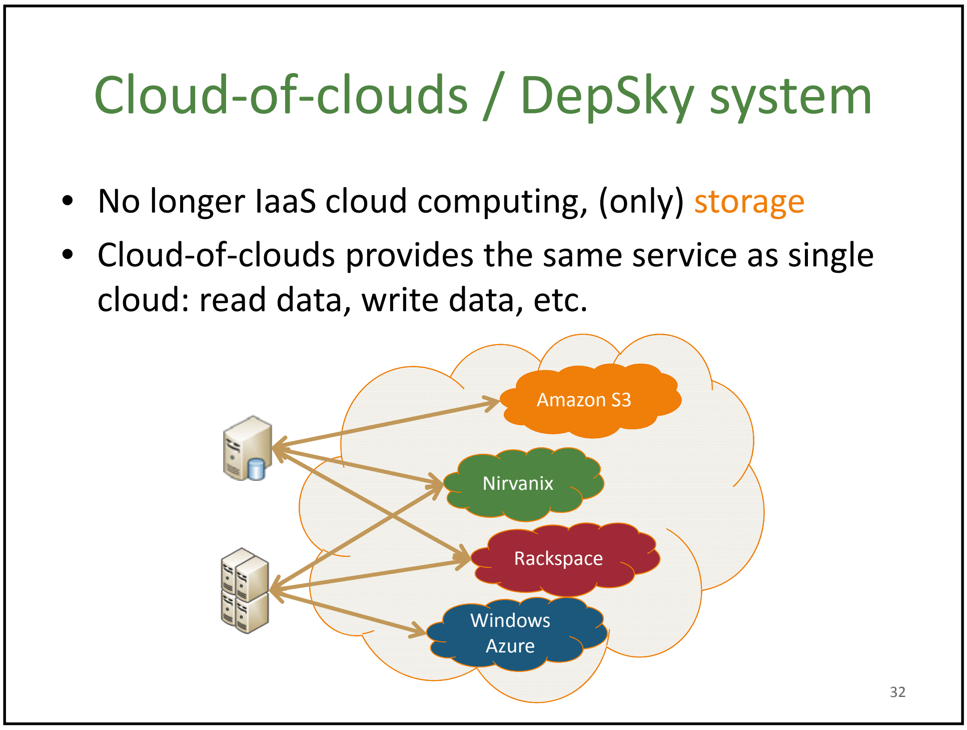
- 1st approach: improve the cloud infrastructure with trusted computing ✓
- 2nd approach: build a (virtual) cloud-of-clouds based on a few clouds
- First can be implemented by providers, second by users





Cloud-of-clouds’ benefits

- Can tolerate data corruption
  - Due to malicious insiders, other attacks, accidental faults (e.g., due to bugs)
- Can tolerate datacenter and cloud outages
- No vendor lock-in
- Faster read access
- Confidentiality…



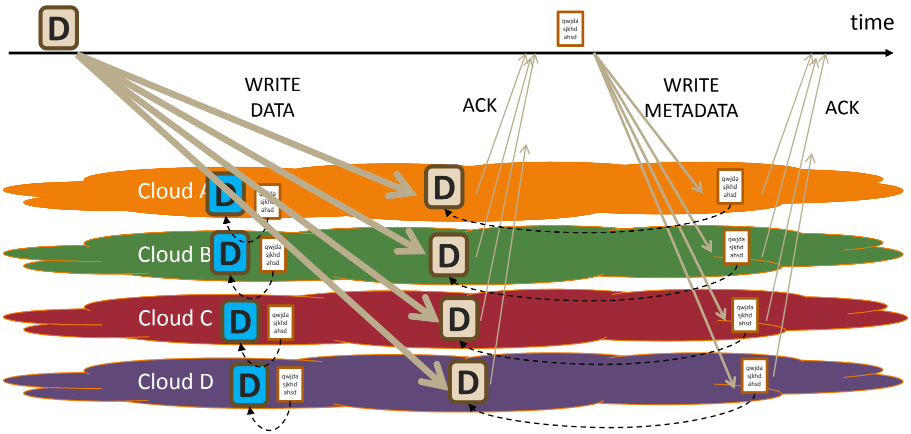
Cloud-of-clouds / DepSky system

- No longer IaaS cloud computing, (only) storage
- Cloud-of-clouds provides the same service as single cloud: read data, write data, etc.

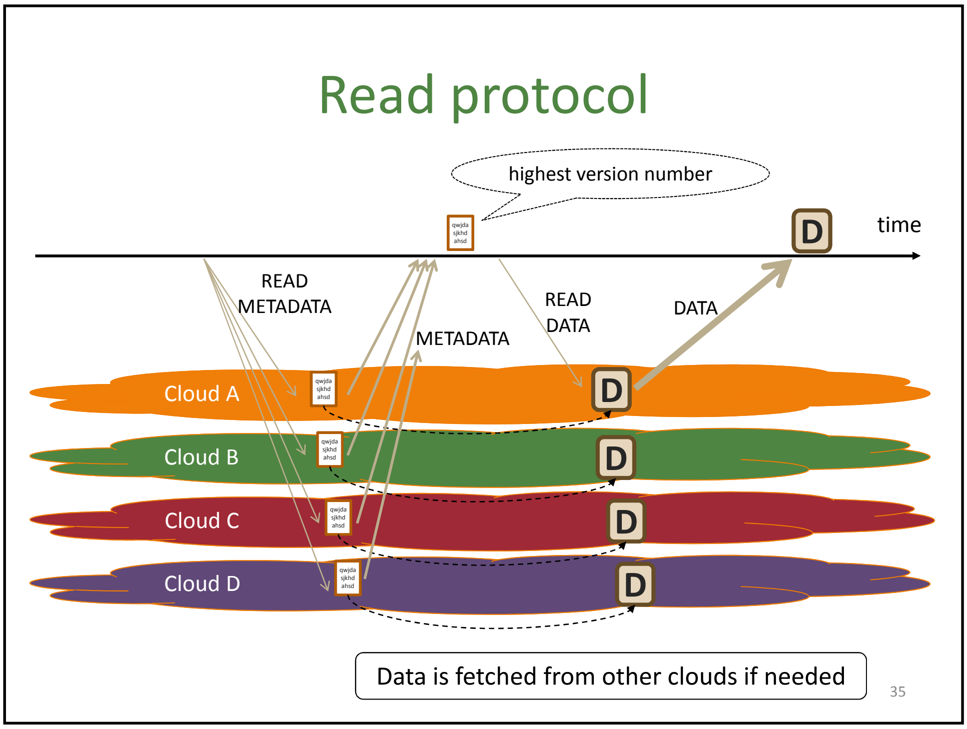
DepSky design principles

- 1. No trust on individual cloud providers
  - Distributed trust is obtained by using several clouds
- 2. Use storage clouds as they are
  - No server-side code in the replication protocols
- 3. Data is updatable
  - Byzantine quorum replication protocols for consistency

Write protocol



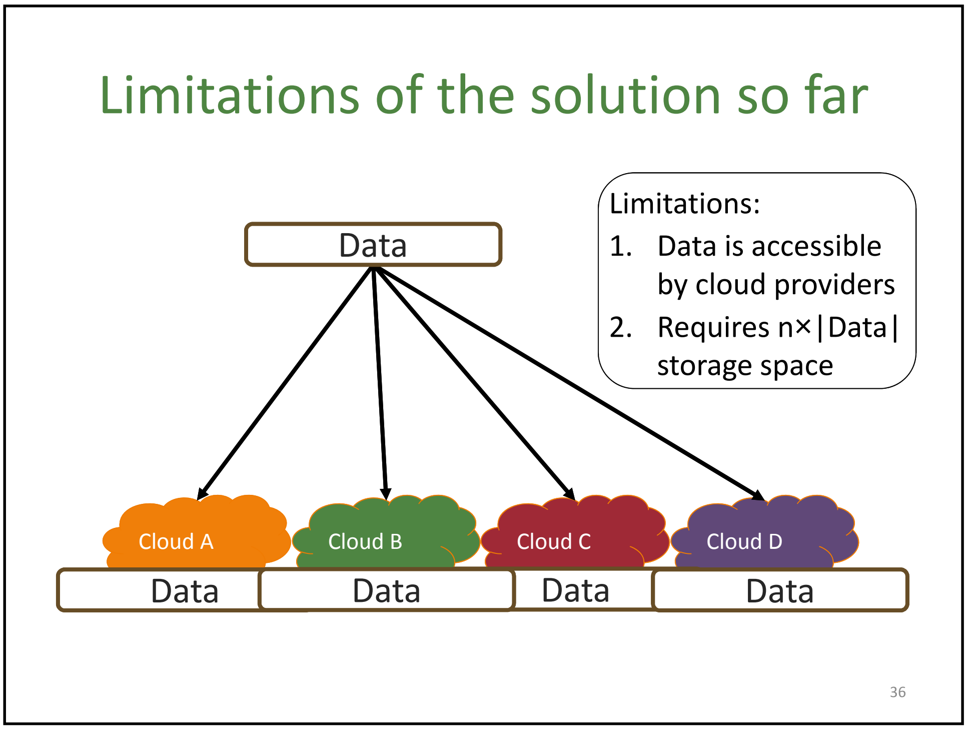
Assume a version of the file is already stored



Read protocol

Data is fetched from other clouds if needed

35

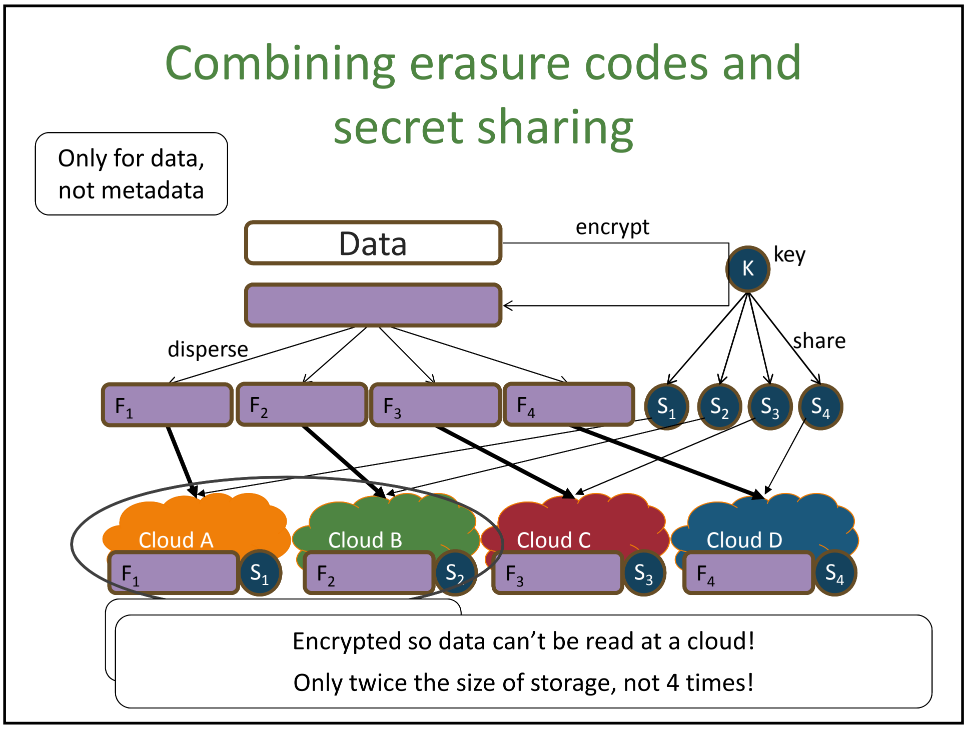


Limitations of the solution so far

Limitations:

1. Data is accessible by cloud providers
2. Requires n×|Data| storage space

36

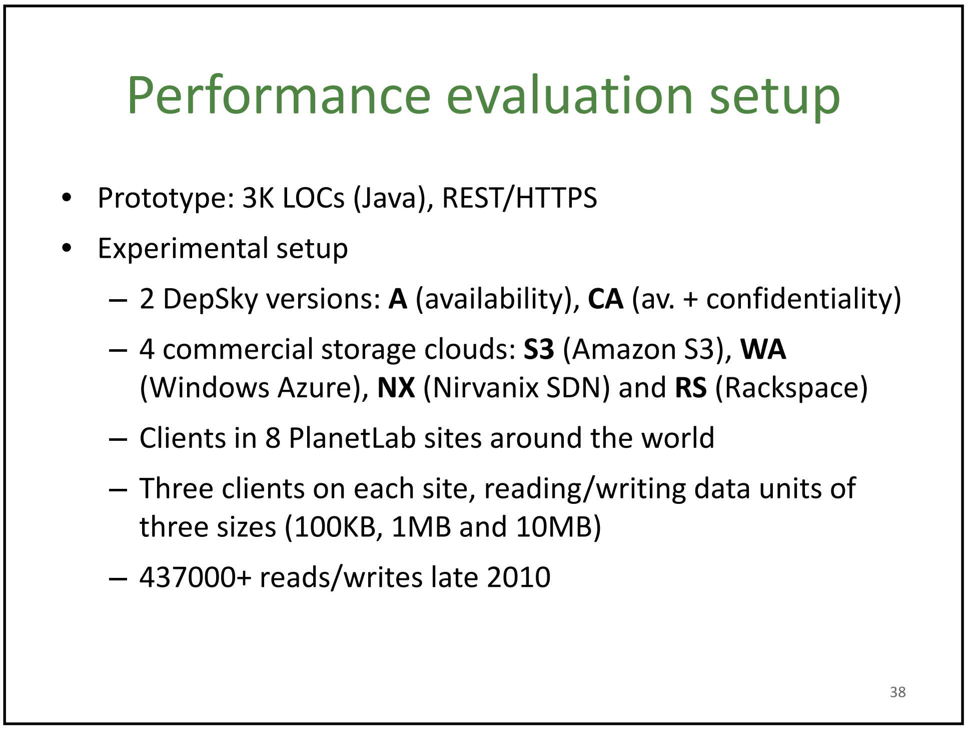


Combining erasure codes and secret sharing

Only for data, not metadata

Encrypted so data can't be read at a cloud!

Only twice the size of storage, not 4 times!

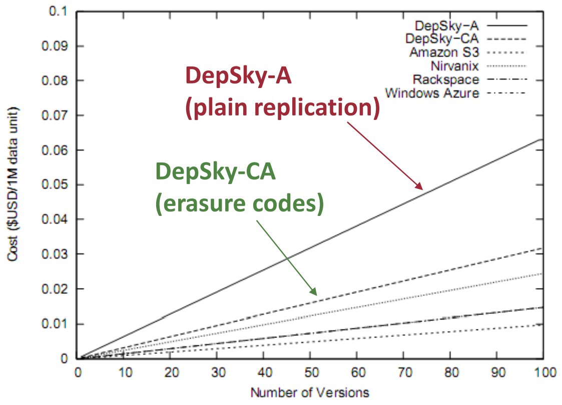


Performance evaluation setup

- Prototype: 3K LOCs (Java), REST/HTTPS
- Experimental setup
  - 2 DepSky versions: **A** (availability), **CA** (av. + confidentiality)
  - 4 commercial storage clouds: **S3** (Amazon S3), **WA** (Windows Azure), **NX** (Nirvanix SDN) and **RS** (Rackspace)
  - Clients in 8 PlanetLab sites around the world
  - Three clients on each site, reading/writing data units of three sizes (100KB, 1MB and 10MB)
  - 437000+ reads/writes late 2010

38

DepSky storage costs ($)



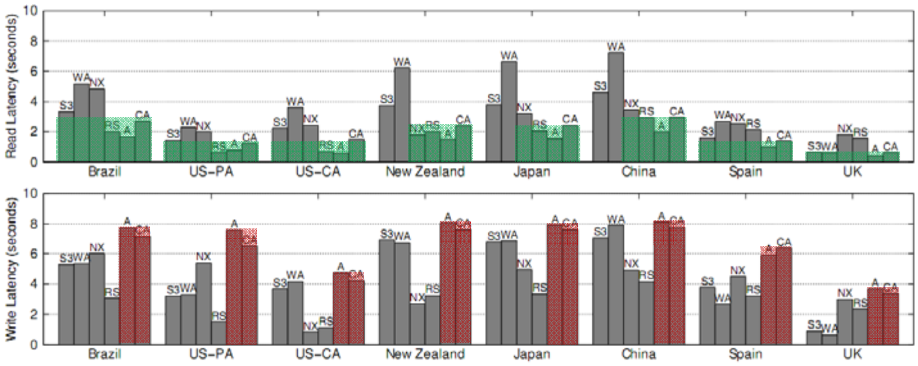
DepSky-A (plain replication)

DepSky-CA (erasure codes)

DepSky-CA storage cost (1M DU) ≈ 2×(average cloud cost)

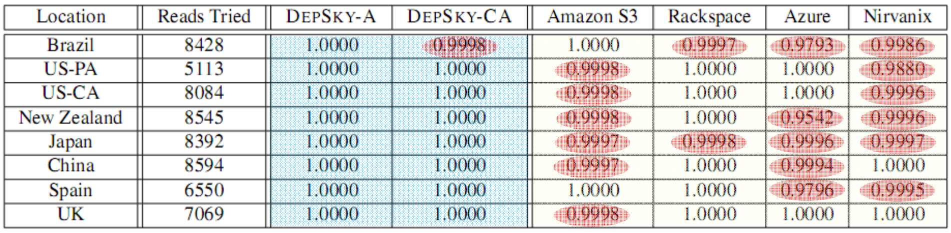
DepSky latency (100KB DU)

DepSky **read** latency is close to the cloud with the **best** latency



DepSky **write** latency is close to the cloud with the **worst** latency

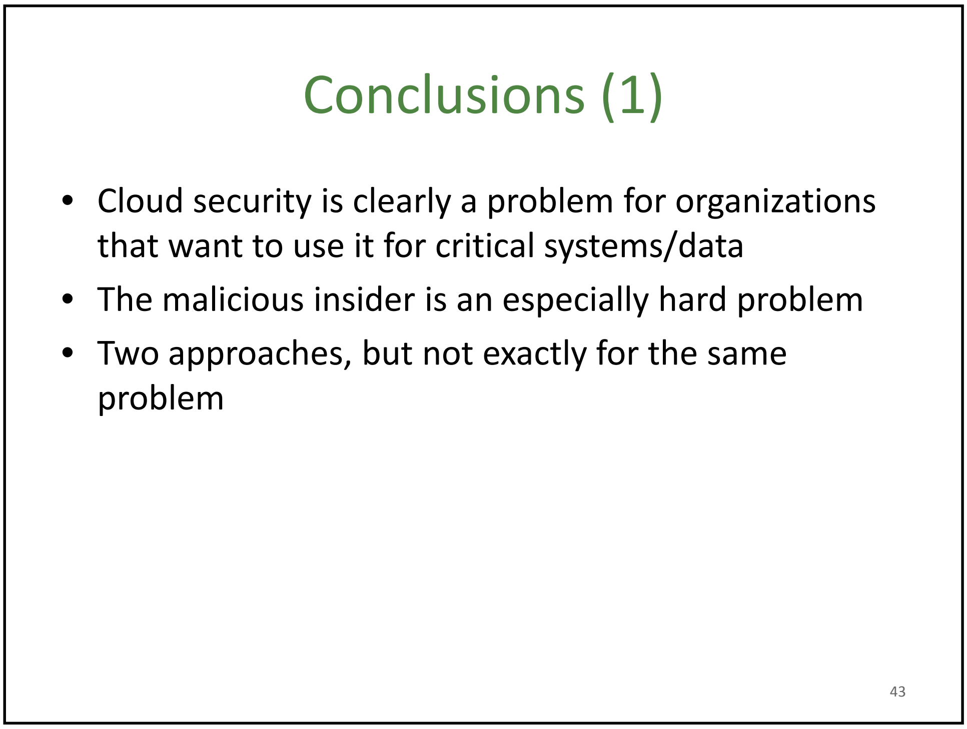
DepSky perceived availability



Location	Reads Tried	DEPSKY-A	DEPSKY-CA	Amazon S3	Rackspace	Azure	Nirvanix
Brazil	8428	1.0000	0.9998	1.0000	0.9997	0.9793	0.9986
US-PA	5113	1.0000	1.0000	0.9998	1.0000	1.0000	0.9880
US-CA	8084	1.0000	1.0000	0.9998	1.0000	1.0000	0.9996
New Zealand	8545	1.0000	1.0000	0.9998	1.0000	0.9542	0.9996
Japan	8392	1.0000	1.0000	0.9997	0.9998	0.9996	0.9997
China	8594	1.0000	1.0000	0.9997	1.0000	0.9994	1.0000
Spain	6550	1.0000	1.0000	1.0000	1.0000	0.9796	0.9995
UK	7069	1.0000	1.0000	0.9998	1.0000	1.0000	1.0000

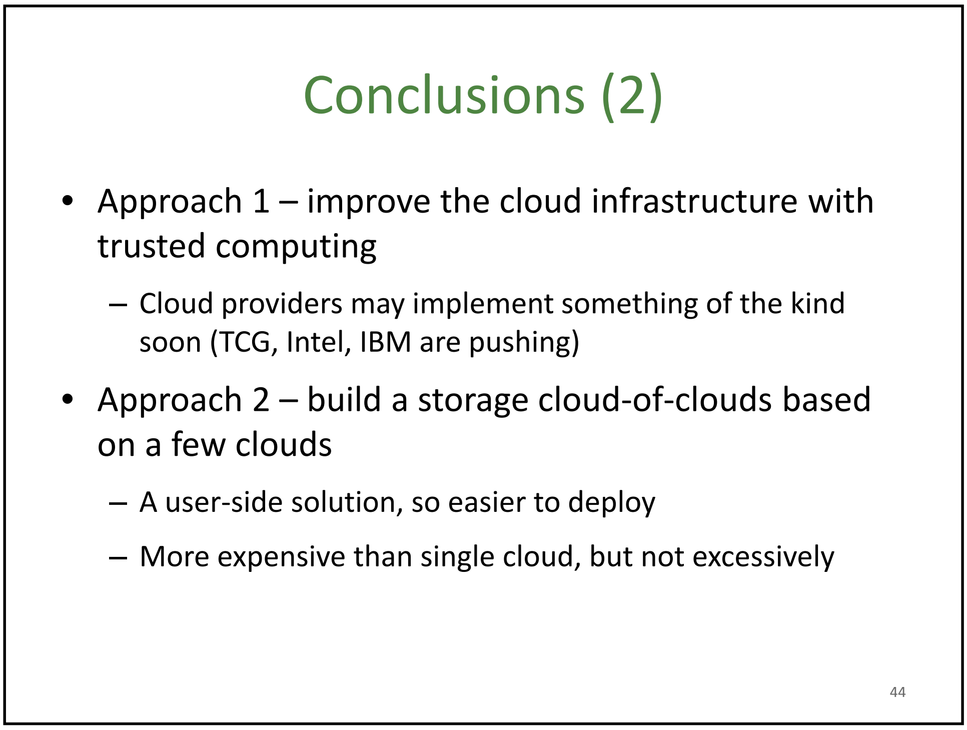
- Apparently, some clouds don’t provide the promised 5 or 6 9’s of availability
- Internet availability plays an important role

CONCLUSIONS



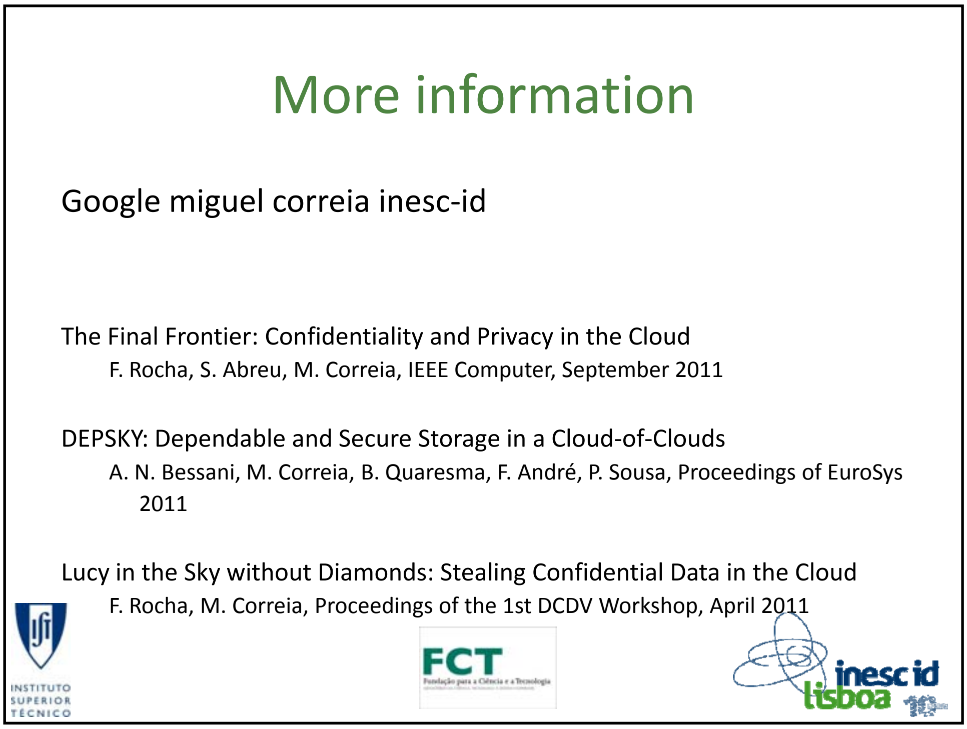
Conclusions (1)

- Cloud security is clearly a problem for organizations that want to use it for critical systems/data
- The malicious insider is an especially hard problem
- Two approaches, but not exactly for the same problem



Conclusions (2)

- Approach 1 – improve the cloud infrastructure with trusted computing
  - Cloud providers may implement something of the kind soon (TCG, Intel, IBM are pushing)
- Approach 2 – build a storage cloud-of-clouds based on a few clouds
  - A user-side solution, so easier to deploy
  - More expensive than single cloud, but not excessively



More information

Google miguel correia inesc-id

The Final Frontier: Confidentiality and Privacy in the Cloud
F. Rocha, S. Abreu, M. Correia, IEEE Computer, September 2011

DEPSKY: Dependable and Secure Storage in a Cloud-of-Clouds
A. N. Bessani, M. Correia, B. Quaresma, F. André, P. Sousa, Proceedings of EuroSys 2011

Lucy in the Sky without Diamonds: Stealing Confidential Data in the Cloud
F. Rocha, M. Correia, Proceedings of the 1st DCDV Workshop, April 2011

INSTITUTO SUPERIOR TÉCNICO

FCT Fundação para a Ciência e a Tecnologia

inesc id lisboa