Security In the Cloud

Giovanni Russello
g.russello@auckland.ac.nz

Computing as an Utility

*In **John McCarthy*** 1960 opinioned:

"computation may someday be organized as a ***public utility***."

What is Cloud Computing?

- Appearance of infinite resources on demand
  - No need to plan ahead for load surges
- Outsourcing is more convenient
  - One hour on 1000 servers = 1000 hours in one server
- Flexible Pay-as-you-go model
  - Processing by the hour
- No need for up-front commitments
  - Small and Medium Companies can get very reliable IT infrastructures

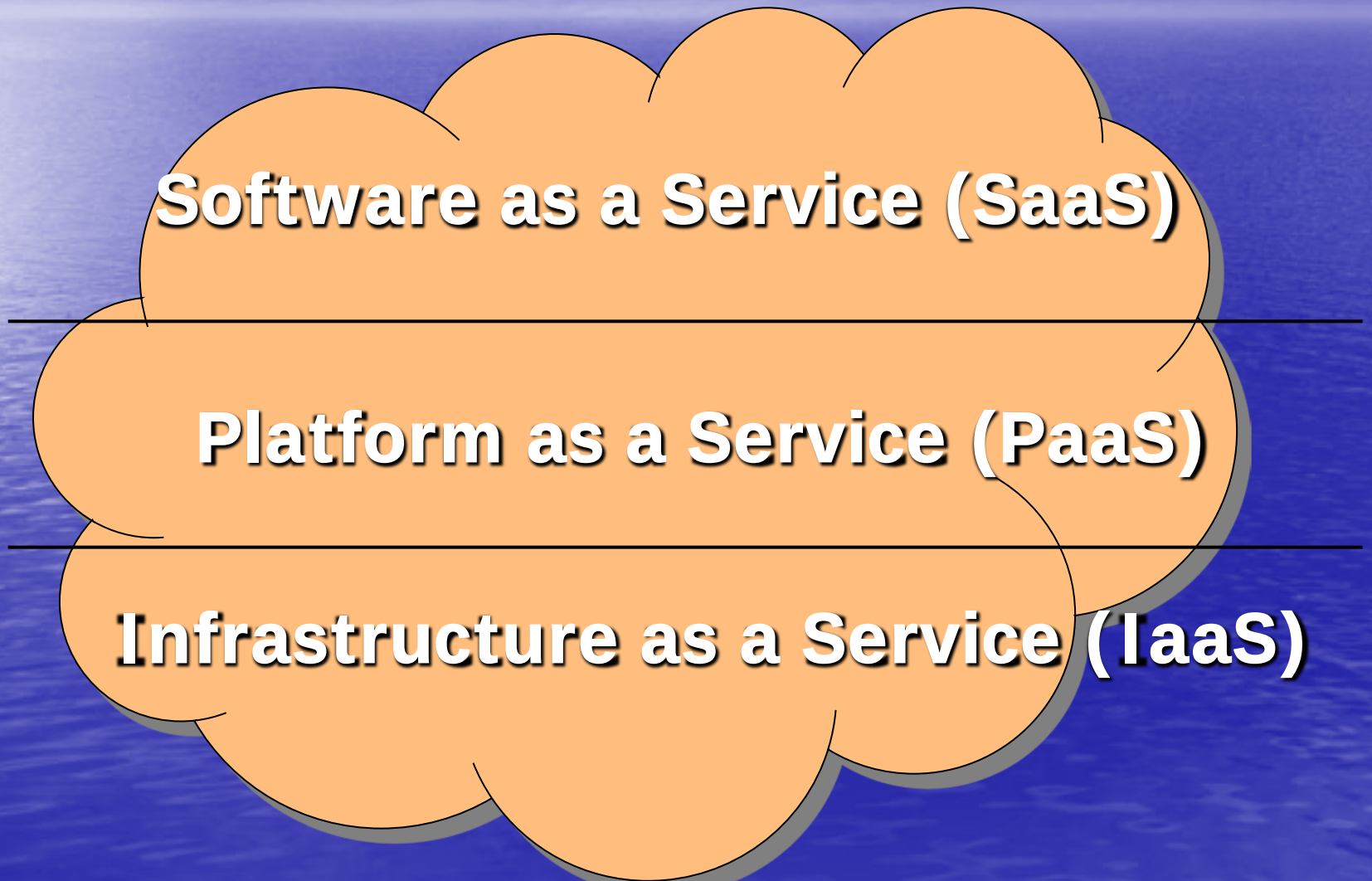
Based on Armbrust et al. - **Above the Clouds: A Berkeley View of Cloud Computing -** Communications April 2010

Some Definitions

Cloud Computing refers to:

- The software offered as Internet services
- The hardware and system software used for providing the services

Cloud Layers



Private vs Public Cloud

- Private cloud refers to internal datacentres of an organisation
- Public cloud refers to datacentres made available to the general public with a pay-as-you-go model

As an analogy

- A semiconductor fabrication line costs over $3 Billions
- Only big players in the market could afford one (Intel, Samsung)
- Then came companies that build chips for others
- Small companies, like nVidia, can capitalise on the chip design without the needs of buying the fab-lines

Who are the big Cloud players?

- Amazon
- Google
- Ebay
- Microsoft

Big datacentres + large-scale software already available

The Cloud Economics

- ***Elasticity***: Shifting the provision risk to the Cloud provider
- ***Pay-as-you-go*** model avoids:
  - Underprovisioning
  - Overprovisioning

Careless Computing?

According to ***Richard Stallman***:

"It's stupidity. It's worst than stupidity"

"I think that marketers like cloud computing because it is devoid of substantive meaning. [] it's an attitude: '*Let any Tom, Dick and Harry* ***hold your data***, let any Tom, Dick and Harry do your computing for you (and ***control it***).' Perhaps the term '***careless computing***' would suit it better."

Storm in the Clouds

Major Security Challenges:

- Availability of Service
- Data Lock-In
- Data Confidentiality

Service Availability

- A cloud computing service by a single provider represents a **Single Point of Failure**
- The provider can go out of business

Data Lock-In

- Cloud Computing API are still proprietary
- Not possible to move from one provider to another

Confidentiality in the Cloud

- Data Confidentiality represents the main obstacle to the adoption of cloud computing
- It is all about trusting valuable data to the cloud
- This data can be strictly regulated (HIPAA, SOX) for auditability

Data Confidentiality Today

- No cloud providers offer data confidentiality as a service
- Amazon Simple Storage Service (S3)
  - "Data stored within Amazon S3 is not encrypted at rest by AWS. However, users can **encrypt their data before it is uploaded** to Amazon S3 " [http://aws.amazon.com/articles/1697?_encoding=UTF8&jiveRedirect=1]

What are the Threats

- User-to-user threat
- User-to-infrastructure threat
- Provider-to-user threat

Protection Mechanisms

The main security mechanism in today cloud is virtualisation

This is effective for user-to-user and user-to-infrastructure threats.

Virtualisation Shortcomings

- However, not all virtualisation software is bug free and
- It is possible to user Cartography to map on which physical server an instance is running

[Ristenpart, et al. Hey, you, get off of my cloud: exploring information leakage in third-party compute clouds. CCS09]

Protection from Providers

- Virtualisation is no effective means for **provider-to-user** threat
- Access control mechanisms are not effective when the infrastructure is not fully trusted
- Moreover there is always the problem of hard drivers “left around”

Some Successful Stories

- TC3 use case for the a HIPAA-compliant application to AWS
  - Reduction/elimination of protected health information (PHI) from the data stored and processed in the cloud
- AWS GovCloud (US): a specialised regional cloud where only restricted personnel as access to its facilities

What about Encryption?

- Traditional Encryption can help to protect the data confidentiality. But it is not practical because:
  - No computation is possible on the ciphertext
  - Ciphertext cannot be searched

We loose the initial benefits of Cloud Computing

Homomorphic Encryption

- Enables computation on encrypted data
- In 2009 Craig Gentry showed that fully homomorphic encryption was possible (but not practical)
- Recent work at Microsoft (Lauter et al) provides some practical breakthrough
  - Adds 100 numbers (128 bit) in 20 millisecs
  - Lots of statistical analysis can be done (i.e. predict when a person is going to have a heart attack)

Encrypted Search

- Performing of search and matching operations on fully encrypted data
- Several schemes exist
  - Single-user
  - Semi-fledged multi-user
  - Full-fledged multi-user

Single-user Searchable Encryption

- Crypto-components are divide between the user and the server
- The user performs encryption/decryption
- The server is responsible for search without learning information about the query and the data

Single-user Searchable Encryption

However

- It is only based on keyword match
- Only a single user can do insert and retrieve operations
- The key can be shared but this complicates key management

Semi-fledged multi-user

- Multiple users can perform search operations
- However, only one single user can do insert operations

Full-fledged multi-user

- Each authorised users can do insert and retrieve operations
- Users do not need to share keys

Take Away

- Confidentiality solutions exist but still more needs to be done
- More effort from the cloud providers towards security solutions