Cloud Security

Florent Dondjeu Tschoufack

December 2021

Abstract

In this paper, we will talk about what cloud computing is and why it is favored in an IT environment. Along side it, we will also discuss about the security threads a cloud environment faces as well a protocol by [1] that helps check the integrity of data hosted on the cloud.

1 Introduction

Cloud computing is a delivery of on-demand computing services over the internet that can be rapidly provisioned and released with minimal service provider interaction. In simpler terms, it is a big warehouse with many servers that are accessible to many. Cloud is said to be invented in the 1960s with the goal of connecting people and data from anywhere. So of the most popular cloud service providers are: AWS, Microsoft Azure, Google Cloud Platform, Alibaba, and IBM. Netflix uses AWS for all its computer and storage needs instead of making its own data center [3]. Depending on the situation, a cloud environment is most favorable because of its: five essential characteristic, three* service models and four* deployment methods.

1.1 Characteristics

The five characteristics of a cloud environment are: broad network access, rapid elasticity, measured services, on demand self service, and resource pooling. Broad network access allows the cloud to be available over the network and accessible through standard mechanism, meaning that I could open a web browser or a mobile device and have access to the data. Rapid elasticity gives the cloud the ability to expand or reduce resources according to individual's needs, meaning I only get charged for exactly what I am actually using. Measured service describes the resources controlled and optimized by cloud system. When connecting the cloud you might find that its thrown a Linux system and the OS taken care of by the provider. On demand self service describes the consumers being able to observe computing capabilities. For example if I am using the cloud to run an intensive task I have the ability to view data CPU usage, memory and etc... And last resource pooling describes the cloud service provider (CSP) being able to serve many consumers.

1.2 Service Models

The three service models of a cloud environment are: Software as a service (SaaS), Platform as a service (PaaS), Infrastructure as a service (Iaas). In this list, there is actually a fourth one which will be discussed in a later section. SaaS describes the application of a CSP to be accessible through various client device such a web browsers. Similar to broad network access characteristic. This also removes the need of software install, maintenance, upgrades and patches which at times can be a challenge for both consumers and providers. PaaS is at it says in its name provides a platform such as an operation system. Consumer are also able to use programming languages, databases and many more supported by the provider. This is also similar to measured service characteristic. IaaS allows the consumers being able to view processing, network, and other fundamentals resources enabling them to build highly adaptable computer systems, similar to the on demand self service characteristic.

1.3 Deployment Methods

The fourth deployment methods are: public cloud, private cloud, hybrid cloud and Community cloud. Public cloud is a cloud infrastructure that is available to the general public. CSP is the one responsible for the control of data and operations within the cloud. Private cloud is a cloud infrastructure that is operated only by an organization. The CSP only manages the infrastructure but has no control. A hybrid cloud is essentially the composition of two or more clouds that are unique but only bounded by standardized technology. Community cloud is a cloud that is shared by many organizations. This cloud is not of spoken much often but I included it because it was in the book [2].

2 Security Risks

Security control on the cloud is very similar to any IT environment but due to its operational models cloud computing have risks specific to a cloud environment. This then introduces the need for another service model, Security as a service (SecaaS). SecaaS can be looked at as a package of security services offered by a CSP. Some of the top cloud security threats identified by Secaas are:

- Abuse and nefarious use of cloud computing
- Insecure interfaces and APIs
- Malicious insiders
- Data loss
- Shared technology issues
- Account or service hijacking
- Unknown risk profile

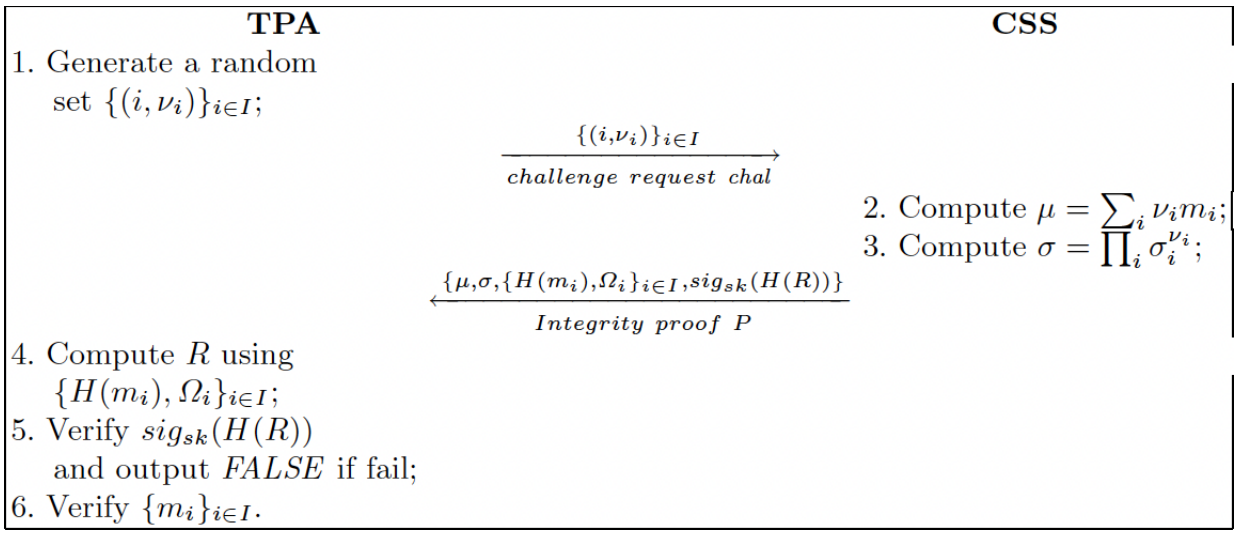

Figure 1: Protocol for integrity

Another concern that arises is how can a client be certain of the integrity of their data hosted on the cloud given a scenario where the CSP may choose to hide data errors or use for malicious intentions. To counter this, we need a method for the client to find an efficient way to periodically check the integrity of their data without the local copy of files, and clients themselves are unreliable which introduces the need for a third party auditor (TPA). [1] propsed a protocal public verification, so that anyone can use it and help test how credible a provider is. This protocol uses algorithms such as Bilinear Map and Merkle Hash Tree (MHT).

2.1 Proposal

In the initialization steps, the files are broken down into m_i blocks, and each block has some identifier v_i. The first in the protocol is for either the client or TPA to generate a message with random blocks and their identifier then sent it to the Cloud Storage Service (CSS). CSS will then compute μ with the the message block and identifier. Then it will compute σ which will be the signature. After step 3, it will construct a proof p containing μ and σ as well as the leaves nodes hash values $H(m_i)$ and its sibling nodes Ω_i, and a meta-data signature of the root hash $H(R)$ which will be sent back to the TPA. With the given

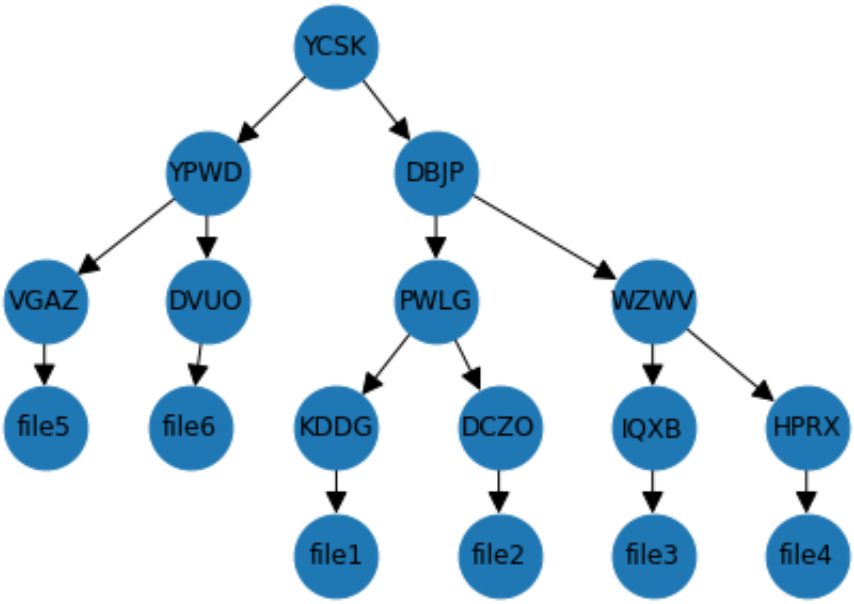

Figure 2: MHT of local files

information the TPA can calculate the value of R with the $H(m_i)$ and Ω_i which can be use to construct a meta-data signature and compare with the one in P. If the output is TRUE, the next step is the very each m_i with μ and σ. If the output is TRUE, we can then be certain that the files have not been tempered with. A visual representation of this can be seen in Figure 1.

2.2 Implementation

For a simulation, I decided to implement MHT and compare the root hash of the file locally stored on my computer with the files hosted n GitHub. The files on GitHub represent files that would be on cloud. To access those files, using python 3 I simply made a GET request and extracted the files content. With the files contents, I then constructed the MHT for local files figure 2, and MHT for GitHub files figure 3. The hash algorithm used was Toy Tetra-graph Hash. Right away when looking at the root hash value, we can instantly tell that the hosted files has been tempered with. In order to identify which file(s) were modified, the next step would be to compare the root hash children and compare the hash

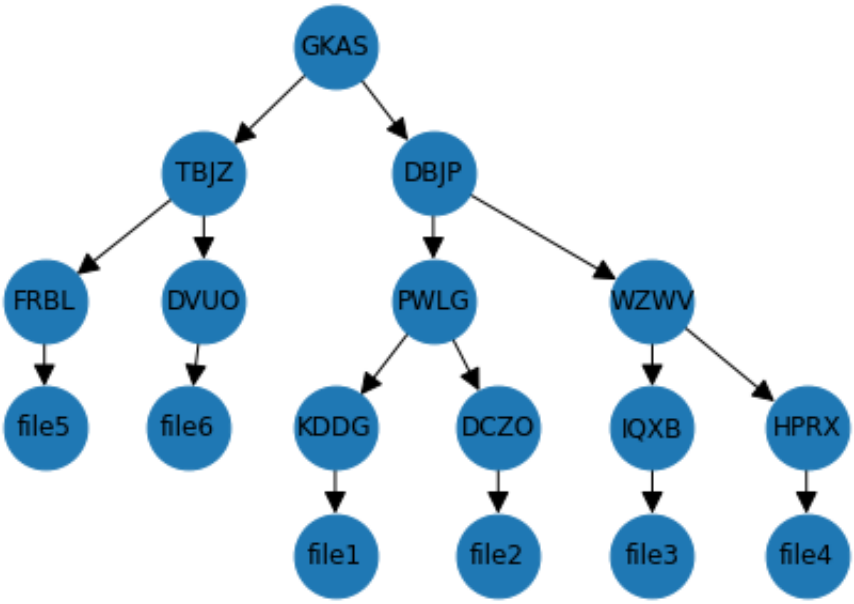

Figure 3: MHT of GitHub files

values of the local and server. In my case the first child of the root was different, indicating that files 1-4 did not change. Eventually as we go down the tree we find out that hash of file 5 is different suggesting that file5 has been tempered with on GitHub.

3 Conclusion

As we have seen in the sections of above, using a cloud environment is very beneficial due to its characteristics, service models and deployments. The consumers don't have to worry about the hardships that comes with hosting stuff online as the CSP takes care of all of that on their behalf. But when using a cloud we also need security measures in place to protect the clients data. As I have demonstrated in section 2.2, MHT is a very power full to that is well implemented in the protocol that help save a lot of computational time.

References

[1] Wang, Qian et al. “Enabling Public Verifiability and Data Dynamics for Storage Security in Cloud Computing.” ESORICS(2009).

[2] Stallings, William. *Network Security Essentials: Application and Standards.* Pearson, 2017.

[3] “Cloud Computing Explained.” Youtube, PowerCert Animated Videos, Nov 2021. `https://youtu.be/_a6us8kaq0g`.

[4] “Data corruption and Merkle trees.” Youtube, Tech Dummies Narendra L, Apr 2020. `https://youtu.be/rsx1nt2bxf8`.