Asian Journal of Advances in Research

5(1): 1099-1106, 2022

DATA SECURITY IN CLOUD: A REVIEW

SANDESH ACHAR [a*#], HRISHITVA PATEL [b] AND SANWAL HUSSAIN [c]

[a] Erie Cir, Milpitas, California, United States.
[b] State University of New York, Binghamton, United States.
[c] Federal Urdu University of Arts, Sciences & Technology, Islamabad, Pakistan.

AUTHORS' CONTRIBUTIONS

This work was carried out in collaboration among all authors. All authors read and approved the final manuscript.

Received: 10 July 2022
Accepted: 20 September 2022
Published: 24 September 2022

Review Article

ABSTRACT

Although many apps benefit from having access to data on the cloud, doing so exposes data to applications that may already have security flaws. The security of data in the cloud is covered in this essay. It is an investigation of cloud data and all security-related aspects of it. In order to provide maximum data protection by lowering risks and threats, the paper will go into detail on data protection methods and tactics utilised globally.
Insights on data security issues for Data-at-Rest and Data-in-Transit will also be provided in the paper. The analysis is based on all cloud services, including platform as a service (PaaS), infrastructure as a service (IaaS), and software as a service (SaaS) (Software as a Service).

Keywords: Cloud computing; data protection; data security; privacy; risks and threats.

1. INTRODUCTION

Many enterprises are turning to the public cloud to address a range of business concerns, whether it's enhancing scalability, global resilience, increasing reliability and performance, or accelerating the release of applications. Data security is frequently at the forefront of people's minds during the complicated process of moving workloads and apps to the cloud. According to the federal bureau of investigations' internet and crime report [1], the number of documented cyber-attacks against American businesses increased 69% in 2020 over 2019. Some businesses are hesitant to migrate their sensitive data to the cloud for this reason. This is happening when they struggle to comprehend their regulatory requirements and look into their security solutions.

It is not possible to simply implement an on-premises workload data security strategy. This is due to the fact that it ignores cloud-centric requirements and does not make use of the huge array of security features the cloud has to offer.

Organizations must rethink their outdated data security methods and develop a cloud-ready strategy in order to successfully move data to the cloud in a way that ensures strict safeguards, complies with legal requirements, and reduces risk. This entails analyzing how the cloud affects your present on-premises strategy and changing it to take use of the fantastic cloud features.

This article aims to provide you with advice so that you may safeguard data in the cloud using a strong data security programmed that makes use of cloud-native thinking and a set of tools that help you quickly meet high requirements for data protection. It evaluates potential risks to cloud-based data as well as

[#] *Director of Engineering;*
**Corresponding author: Email: sandeshachar26@gmail.com;*

the defenses put in place by different service providers.

The remaining portions of the essay are structured as follows: The forms of risks to cloud-based data are covered in Section II. The three pillars that support data security in the cloud are discussed in Section III. Section IV looks at a few effective data security methods used globally. The conclusion, which is the last section, offers a summary of this study.

2. DATA SECURITY RISKS AND CONCERNS IN THE CLOUD

Fig. 1. Data security

Security experts who have led data security initiatives for on-premises installations are beginning to realise that their organization's move to the cloud necessitates a reevaluation of how data is safeguarded.

Organizations were asked to rate the top security risks to public clouds, and misconfiguration came in first (68%), then illegal access (58%), unsecure interfaces (52%), and account hijacking (50%) [2].

As you'll see, cloud-based data security strategies frequently continue to use the well-known security patterns that have long protected data on-premises while also providing a more up-to-date security infrastructure that takes use of the efficiency and power that cloud platforms provide.

The shared responsibility paradigm is one you embrace when you migrate your IT infrastructure to the cloud [3]. Because the CSP operates, manages, and controls all the layers of IT components, from the host operating system and virtualization layer down to the physical security of the premises in which the services operate, this shared approach lessens your operational burden.

You and your CSP both have responsibility for managing, operating, and verifying IT controls, just as you do for running the IT environment. In conclusion, your CSP is in charge of the security "of" the cloud, while you, as the client, are in charge of the security "in" the cloud.

The majority of security standards and compliance certifications, such as SOC 2, PCI-DSS, HIPAA/HITECH, FedRAMP, GDPR, FIPS 140-2, and NIST 800-171, are typically supported by contemporary public cloud service providers (CSPs), assisting clients in meeting compliance requirements for almost all international regulatory bodies [4].

The following are the primary concerns to data security in the cloud:

1) Misconfiguration
2) Account theft
3) Virtualization
4) Insider threats
5) Insecure API
6) Denial of service
7) Data sharing
8) Data privacy compliance and
9) A lack of adequate governance.

2.1 Misconfiguration

The biggest security risks resulting in cloud data breaches are incorrectly configured cloud security settings. When it comes to safeguarding their cloud-based ecosystem, most firms' current cloud security posture management procedures frequently fall short. Cloud infrastructure is made to be simple to use and simple to exchange data with others. Organizations find it challenging to guarantee that data is only accessible to authorized parties as a result.

As a result, organizations using cloud-based infrastructure must rely on security controls offered by their cloud service provider (CSP) to configure and secure their cloud deployments. Furthermore, organizations using cloud-based infrastructure do not have complete visibility and control over their infrastructure. Given that many businesses struggle to secure their cloud installations and frequently use many clouds, each of which has its own vendor-specific security rules, it is simple for a security blunder or mis-configuration to leave a company's cloud-based resources open to attack.

2.2 Account Hijacking

Another serious concern to the security of cloud data is account hijacking. Password reuse and the usage of weak passwords are two common examples of severely lax password security. Because of this issue, phishing scams and data breaches are made to be more damaging because a single stolen password can be used on numerous other accounts.
As enterprises increasingly rely on cloud-based infrastructure and applications for critical business

processes, account hijacking is one of the more serious cloud security challenges. An attacker who has obtained an employee's login information can access sensitive information or features. The attacker was in a position to take complete control of the victim's online account thanks to the stolen client credentials.

2.3 Virtualization

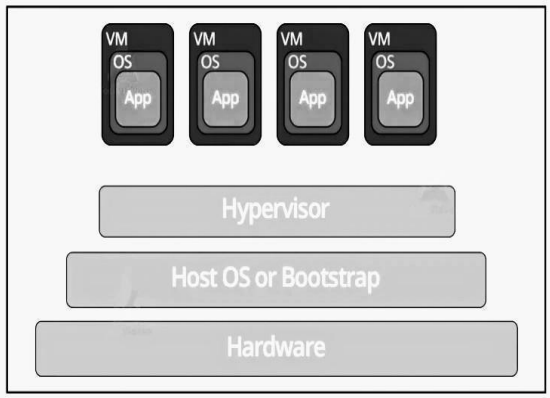

Fig. 2. Virtualization

Because of the inherent economies of scale and higher efficiency, virtualization is a key enabler for cloud computing. Software is used to replicate hardware functions in a virtualized system. Organizations can operate several virtual systems, as well as numerous operating systems and applications, on a single server thanks to this abstraction. It is a clever method for enhancing cloud security by introducing isolation and abstraction.

But it also draws a lot of brand-new security flaws and threats. Virtualization has two key security considerations: first, the safety of virtual technology itself, and second, the discovery and introduction of new security vulnerabilities. When a guest OS is run over a hypervisor without understanding the dependability of the guest OS, which may have a security flaw in it, data may be at danger when virtualization is used for cloud computing. Another security risk you need to handle is the existence of untrusted virtual machines or virtual appliances that contain Trojans.

2.4 Insider Threats

Insider threats are attacks that are perpetrated by a user or malicious code that is already in the system. Traditionally, organizations have focused more on external threat vectors by rigidly profiling the edge network as a walled-off subnet. By doing so, they can segregate the external, untrusted public internet traffic from their internal production traffic. However, nowadays attackers have become more sophisticated, especially with funding from nation states.

The most priced target for rogue actors is privileged accounts. By taking over such accounts, they can elevate privileges within a cloud environment. Apart from external adversaries, it is malicious insiders, either intentionally or otherwise, that are to blame for most security breaches. The focus today therefore should not only be on the perimeter, but also in establishing whether an attacker has gained a foothold within the organizations defined perimeter.

2.5 Insecure APIs

The foundation of processing, APIs are created to simplify data access and integration. They specify the sorts of requests that take place between programmes, how these requests are made, and the different data formats that are employed in this interaction. APIs are often used in applications to process sensitive data like payment information and user passwords. The same APIs can put a line of communication at high risk for security breaches due to man-in-the-middle (MITM), distributed denial-of-service (DDoS), SQL injection, and failed access control attacks. Therefore, it is essential to safeguard the private information that people transfer.

2.6 Denial of Service

55% of firms encounter a DDoS assault at least once each month, according to a recent research [5]. A DoS or DDoS assault that is effective might bring down different cloud services and resources. These assaults operate by saturating servers with traffic. The server may crash, the data may be damaged, and the system may become paralysed as a result of being flooded with more Transmission Control Protocol/User Datagram Protocol (TCP/UDP) packets than it can handle. The availability, performance, and service level agreements between the cloud service provider and its clients could be affected as a result.

2.7 Data Sharing

The difficulty of data sharing grew more accessible and practical as cloud computing got more practical. But it also brought dangers to cloud computing's security and integrity. Enhanced access and sharing put a company at risk for online security issues. Incidents that affect the accessibility, reliability, or confidentiality of data and information systems may result from this.

Breach of personal data is likely to be the focus of criminal organisations. This kind of breach not only violates the privacy of the people whose personal information was revealed, but it can also result in

severe financial damages for the impacted business, such as a loss of commercial opportunities, competitiveness, and reputation.

Data breaches are becoming more common, as shown by the theft of over 21 million records, including 5.6 million fingerprints, from the US Office of Personnel Management in 2015 and the Japanese Pension Service incident that affected 1.25 million people.

2.8 Data Privacy Compliance

Every piece of information that is kept and accessible online has the potential to become compromised by a small error, infringing on compliance rules. As a result, one of the biggest dangers associated with cloud computing is thought to be enterprises' ability to achieve the compliance criteria.

2.9 Lack of Appropriate Governance

Fig. 3. Data Governance

Lack of control over company data is another another significant obstacle to data security. Business executives are beginning to realize they are buried in data and just need a place to put it. As a result, they concentrate more on making sure that the data can be retrieved in the first place and worry less about who manages data storage or retrieval. Noncompliance may occur if a company doesn't exercise strict control over its data. For instance, if staff members have access to information they shouldn't, it may be necessary to impose injunctions. This exposes the company to security lapses and legal issues arising from the CCPA, GDPR, and other current laws.

3. THE THREE-PILLAR APPROACH TO DATA SECURITY IN THE CLOUD

An efficient cloud-based data security approach is built on three main pillars:

- **Identity**: A key component of a successful cloud-centric data security strategy is understanding the identity of users, machines, and apps as they produce, alter, store, utilise, share, and eventually erase data.
- **Access Boundaries**: The second key tenet for safeguarding data in the cloud is access boundaries, which entail setting limits on who can access data and when. Your data security programme with a cloud focus should employ identity to manage access through policy and make use of the numerous cloud-based tools and services that make this simple and efficient.
- **Visibility:** Utilize the cloud's robust visibility capabilities when the data fences are in place to audit usage and offer compliance reports that shows how data is regulated and accessed by both your own cloud administrators and the CSP staff who help with your cloud infrastructure. These visibility technologies enable focused reaction operations by detecting threats and anomalies quickly.

These pillars serve as the program's foundation, to which additional controls can be applied to produce the level of comprehensiveness that is best for your company. These three pillars are regarded as essential because they are a part of any programme and their absence would significantly reduce their effectiveness. All data, regardless of classification or stage in the life cycle, must have intentional access guardrails in place that allow the necessary entitlements for data usage (even if such access is eventually "public" access). For instance, not all data needs to be encrypted at rest.

4. DATA SECURITY CONTROLS IN THE CLOUD

Because you are sharing security responsibilities with your cloud vendor, a cloud-centric data security approach focuses more on leveraging access restrictions and permissions to create data guardrails rather than on containing full-stack threats.

Key management, activity monitoring, data loss prevention, and various endpoint protections will make up the majority of data protection controls. Data encryption, both in transit and at rest, and the deployment of secure compute technology to safeguard data while it is being used are two other typical data security measures.

A thorough data security scheme should include the following eight essential elements:

Identity, Access Control, Visibility, Encryption, Key Control, Secrets Control, Classification, Edge and Network Protection.

4.1 Identity

The accuracy of people, machines, and processes' identities is a prerequisite for any data access policies. Identity plays a role in data governance, and the sensitivity of the data is frequently determined by who has access to it. Identity is, in general, a key element of a successful data security plan. You will have overcome a significant portion of your data security difficulties if you can even partially accomplish this. If you make a mistake, additional data security measures won't provide nearly enough security. To do this correctly, you must first determine who need access to the data. These can include people, tools, and services geared toward customers. The human users can be customers inputting data to an application, customer service representatives, data analysts, system administrators, developers, and DevOps engineers, just to name a few.

For controlling access to services and resources, the majority of cloud vendors offer an Identity and Access Management (IAM) console. IAM, for instance, can be used on AWS to add users as members of a project. An account classification known as a service account can be authorised on GCP. For instance, you could use this to operate a Compute Engine instance as a service account and offer that account access to the data resources it requires *Access.*

4.2 Management

It's important to have a firm understanding of who need access, but the real job begins when access is actually granted. The hardest element of putting your data security policy into practise is usually developing fine-grain access restrictions to set up transparent and reliable boundaries to sensitive data. Strict access regulations typically use a number of security controls and solutions to both deny access and provide it. Enabling access limits through access rules guarantees that your workforce and customer-facing applications always have the appropriate access and creates a layered effect that secures your data according to its sensitivity.

Three layers can be used to manage access boundaries:

1. Access to the network layer
2. IAM access controls
3. Separation between providers and clients.

4.3 Visibility

With the help of cloud platform services, you can automatically monitor, log, and audit your supplied resources in order to spot any unusual security occurrences, like configuration modifications. This delivers real-time insights and control over your data when combined with robust analytics and discovery capabilities that are powered by machine learning. The centralized management of security controls offers a unified single source of visibility into how data security is applied throughout your business, in addition to monitoring and logging.

These technologies also check your cloud data for security flaws and perform data discovery. They can also categorise sensitive data and automatically mask, tokenize, and alter sensitive parts so you can better control the risk associated with data collection, storage, and use. This makes it simple for you to protect data throughout every phase of its lifespan.

As your data moves through its lifetime, it's important to gain insights about what's occurring with it in addition to knowing it through security scanning. How is information accessed? How is information transferred and shared? Changes in permissions? These security incidents and others can be recorded and compiled via data logging and auditing. To do security and access analytics, you can even export logs to a data warehouse for analysis. This enables you to spot unwanted modifications and improper access to the data belonging to your company.

The centralised administration interface that was stated can be used to trigger a variety of data security capabilities that check your cloud infrastructure for security flaws and hazards. The security assessment capabilities available include the ability to scan storage systems for sensitive data, monitor for data exfiltration, and detect which data is open to the internet.

4.4 Encryption

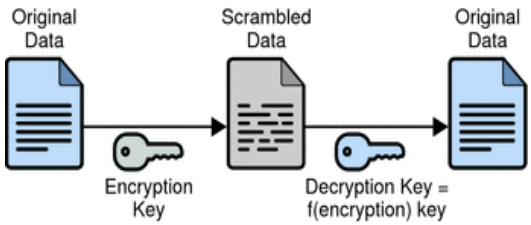

Fig. 4. Encryption

Data encryption is a reliable way to partition information and prevent unauthorised access. Simply encrypt the data, keep the key in a safe location, and then put the encrypted data wherever you wish. Data

encryption has gained acceptance and has become the standard method for safeguarding sensitive data. Encrypting data has developed into a frequently recommended way for securing data in many regulations and compliance standards, whether it be structured data stored in databases, unstructured data like file sharing and object storage, or other use cases like tokenizing data fields.

Let's examine a few effects that encryption may have on your data security plan. Data is encrypted at rest by default, which is common among cloud services. You don't need to be involved or set anything up for this to happen because it is done fully transparently. This can be adequate for the majority of use cases, depending on your encryption requirements with relation to key management.

All data moving across the global network of their regions and data centres is encrypted at the physical layer before it leaves the secure facilities, which makes them different from other cloud service providers. Additionally, additional encryption can be used, for instance, during customer-to-customer or service-to-service TLS communications.

4.5 Key Management

Cloud service providers additionally offer a continuum of encryption key management alternatives if you have certain key management needs. For instance, you can utilise a service like Cloud HSM on AWS or KMS Google Cloud to provide you that control if you need to rotate your encryption keys or wish to manage your own keys. Additionally, you can utilise these services to encrypt your data using Hardware Security Modules that have been validated to FIPS 140-2 Level 3 or software-backed encryption keys (HSMs). Additionally, they offer redundant keys that are accessible everywhere. Additionally, they can let you bring your own key (BYOK) to the cloud if your key management requirements demand that you create your own keys.

To exert even more control, you can divide the responsibility for safeguarding your most sensitive data between your cloud provider and a third-party key management system that you manage outside the cloud vendor's infrastructure by storing encryption keys in an External Key Manager (EKM). You gain a high level of assurance that your data cannot be accessed at all and you actually achieve a secure Hold-Your-Own-Key (HYOK) model for key management because your data within the cloud infrastructure is encrypted by keys you maintain outside of the cloud.

4.6 Secrets Management

In configuration or properties files of classic on-premises tools, sensitive data fields or strings such as API keys, passwords, connection strings, and many others are frequently saved. The maintenance of these secrets is fully integrated into the API and most services in the cloud, which makes this unusual. A robust yet user-friendly Secret Manager that secures and centralises management and access to these secrets can be used to store these delicate fields of data in the cloud. Applications can be set up to automatically use the most recent version of a secret, and secrets can be rotated automatically. Additionally, you may use audit logs to access every secret and achieve complete transparency of every interaction.

4.7 Data Security

Many firms frequently squander a lot of time and effort enforcing strict security on data that doesn't actually require it. Data should be categorised based on how sensitive it is so that the proper security controls may be implemented. The data that requires the greatest time, effort, and money can be segmented based on its classification.

As early as possible in the data management lifecycle, especially at the Create stage, data classification should be carried out. Data can be divided into 4 categories:

1. Public: Information that is available to the general public.
2. Internal: Non-sensitive information that is kept private.
3. Sensitive information, general circulation; confidential.
4. Restricted: Extremely sensitive, regulated data, and restricted distribution

You can truly comprehend your data and why it's vital to safeguard it if you evaluate the context of the data using the categories listed above. Try to arrange related data together, for instance:

All production databases assisting customer-facing apps can be categorized as Confidential; all restricted data produced by a certain business group (Finance, HR, Executive Staff, etc.) can be kept in designated repositories; and so on.

4.8 Edge and Network Protection

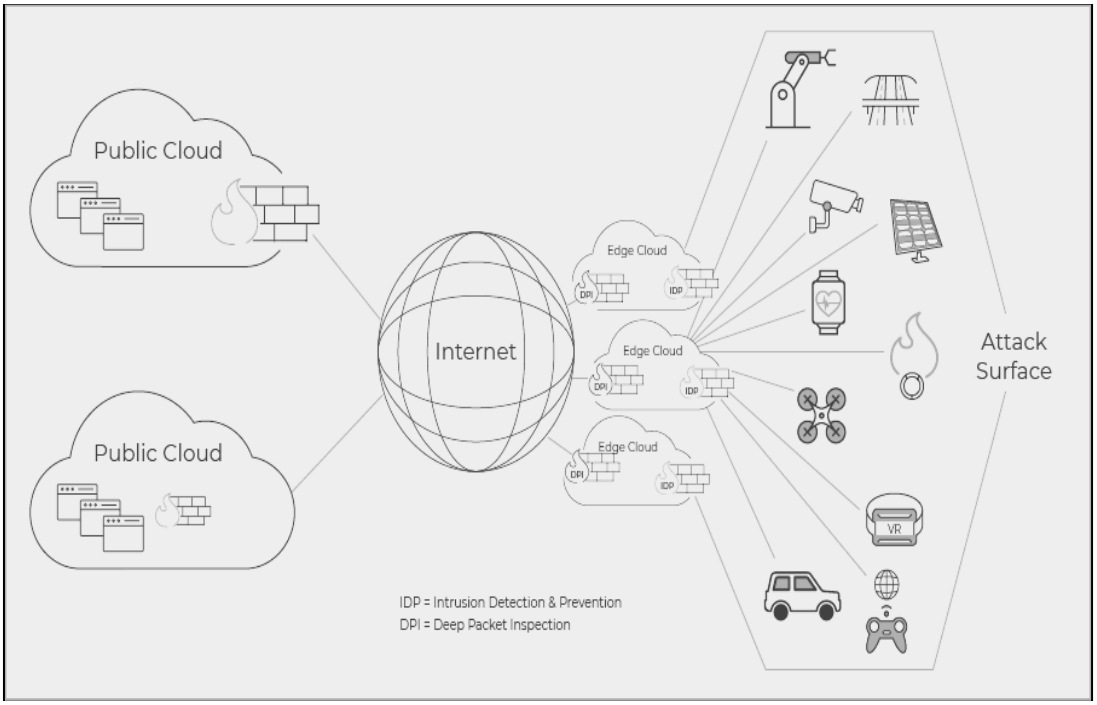

Fig. 5. Edge and Network Protection

Customers of the cloud can secure their data by utilizing various security protocols in an application, such as OAuth 2.0 with Transport Layer Security or an API key (TLS). Rate limiting, setting up mutual TLS for your API layer's backend, and creating threat protection guidelines based on established principles. You may, for instance, filter requests based on URI strings, HTTP body content, HTTP headers, or geolocation data. This enables you to stop frequent attack methods like SQL injection and cross-site scripting.

5. CONCLUSION

Data security considerations will undoubtedly need to be dealt with on your journey to the cloud. When you design your data security strategy to be cloud-native, you take advantage of capabilities in the cloud that are simply not possible to replicate on-premises. For example, you can aggregate security visibility across the entire infrastructure into a seamless management console, have integrated security event logging at your fingertips, use recommendation tools that use machine learning to find over-provisioned accounts and alert you proactively, or encrypt every byte of data. Few teams operating on-premises will ever be able to attain these data protection capabilities, but security teams employing the cloud are completely capable of achieving them.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Abbate P. 2020 Internet Crime Report. Federal Bureau of Investigations. 2021;3.
   Available:https://www.ic3.gov/Media/PDF/AnnualReport/2020_IC3Report.pdf
2. Schulze H. 2020 Cloud Security Report [ISC2]. Cybersecurity INSIDERS. 2021;8.
   Available:https://www.cybersecurity-insiders.com/wp-content/uploads/2020/08/2020-Cloud-Security-Report-ISC2.pdf&ved=2ahUKEwi05dLb04_6AhV7xQIHHaMuCOcQFnoECA0QAQ&usg=AOvVaw3VnzUFhyieI7fwKTTrQclM
3. Amazon Web Services. AWS Security and Compliance Quick Reference Guide. AWS. 2021;8.
   Available:https://d1.awsstatic.com/executive-insights/en_US/guide-security-compliance-quick-reference.pdf
4. Lance A, Chuvakin A. Designing and deploying a data security strategy with Google Cloud. Google Cloud. 2021;10.

Available:https://cloud.google.com/blog/products/identity-security/start-a-data-security-program-in-a-cloud-native-way-on-google-cloud

5. App security not keeping up with rapid development — Radware. [Blog]; 2021. Retrieved 12 September 2022, from https://securitybrief.com.au/story/app-security-not-keeping-up-with-rapid-development-radware/