

Data Security and Privacy in Cloud Computing: A Comprehensive Review

¹Radadiya Drashti, ²Mr.Ankitkalariya

¹(B.Tech.in Computer Engineering, Atmiya University, Rajkot, India

Email: drashti2004radadiya@gmail.com)

²(Faculty of Engineering & Technology, Atmiya University, Rajkot, India Email:

ankit.kalariya@atmiyauni.ac.in)

Abstract:

Cloud computing is a way of using computer resources over the internet. It lets users access shared Pools of computing resources, like servers, networks, and storage, on demand. This has become a big change in the world of Information Technology (IT). Even though cloud computing offers many benefits, like saving money, being able to grow quickly, and being flexible, there are still major issues with data security and privacy that stop it from being widely used, especially by big companies and in sensitive situations. This paper brings together information from various studies to talk about the different ways cloud computing is set up, how it's deployed, the main security threats, and ways to deal with those threats. It explains the three main service models Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a service (IaaS) as well as the different ways it can be deployed, like public, private, hybrid, and Community clouds. It also looks closely at security and privacy issues throughout the entire life of data, including making sure data stays safe, private, and accessible. The paper discusses advanced ways to protect data, like Homomorphic Encryption, Attribute-Based Encryption (ABE), and Searchable Encryption. It also looks at best practices that organizations can use, such as strong access controls and regular security checks. It highlights how these challenges are changing from problems with multiple users sharing the same resources and people with too much access to more complex threats like side channel attacks and attacks that could become a problem with quantum computers. The paper stresses the need for a complete, integrated, and standardized approach to security and trust in cloud environments to ensure they are safe to use.

1. INTRODUCTION

Servers managed by third party cloud Service providers (CSPs) bring in new and complex challenges. Cloud Computing has moved from being an idea to a fully developed and highly influential technology. It has changed the way IT resources are used and managed. It represents a big change, allowing users to get access to a shared set of configurable computing resources through the internet. These resources include networks, servers, storage, applications, and services. This model brings many economic and operational benefits. These include saving money, speeding up innovation, getting products to market faster, and having the ability to scale up or down quickly based on changing needs. However, the rapid move of applications and data to remote servers introduces challenges that stop full adoption. The main problem is being worried about data security and privacy. The loss of control over data, which can be stored in many places around the world in shared infrastructures, increases traditional security risks and creates new ones. This review aims to give a clear and complete look at cloud computing security. We structure the discussion as follows:

Section 2: Defines the basic structure of cloud computing and describes the main service models and deployment types.

Section 3: Identifies and explains the main data security and privacy challenges.

Section 4: Examines the current and emerging technical and organizational solutions for dealing with these threats.

Section 5: Concludes the paper, highlighting future research directions.

2. Cloud Computing Architecture

Essential Characteristics

Key characteristics, as often defined by NIST, enable the cloud experience:

On-demand Self-Service: Consumers can unilaterally

provision computing capabilities, such as server time and network storage, without provider.

Ubiquitous Network Access: capabilities are available over the network and accessed through standard mechanisms.

Location-Independent Resource Pooling: Computing resources are pooled to serve multiple consumers using a multi-tenant model, dynamically assigned and reassigned according to demand. The customer generally has no control or knowledge over the exact location of the provided resources.

Rapid Elasticity: Capabilities can be rapidly and elastically provisioned and released to scale up or down commensurate with demand.

Measured Service : Resource usage is monitored, controlled, and reported, providing transparency for both the provider and consumer and enabling the utility-based consumption model Service Deliver Models(TheSPIModel)cloudservicesarestructuredintothreemainlayers, collectively known as the SPI Model

Infrastructure as a Service (IaaS): IaaS provides the consumer with the fundamental computing resources, such as virtual machines (VMs), storage, networks, and operating systems. The user can deploy and run arbitrary software, but the provider manages the underlying cloud infrastructure. **Customer Responsibility:** Operating systems, applications, and content security. **Provider Responsibility:** Virtualization layer, physical infrastructure security, and some low-level data protection. Examples: Amazon EC2, Google Compute Engine.

Platform as a Service(PaaS): PaaS offers a development environment and platform that enables customers to deploy their own applications using programming languages and tools supported by the provider. The customer controls the deployed applications, but the provider manages the underlying operating systems and hardware.

Customer Responsibility: Applications and data security.

Provider Responsibility: Operating system, platform and underlying infrastructure isolation and security.

Examples: Google App Engine, Force.com.

Software as a Service(SaaS): SaaS provides the capability to use the provider's applications running

on a cloud infrastructure, accessible from various client devices, typically through a thin client interface like a web browser. The user only manages limited application configuration settings; all other infrastructure and application capabilities are managed by the provider.

Customer Responsibility: Limited user-specific application configurations.

Provider Responsibility: Application security, data confidentiality and infrastructure management. Examples: Gmail, Sales force. The service models stack hierarchically, with

IaaS serving as the foundation, PaaS building upon IaaS, and SaaS, in turn, building upon PaaS. Security responsibility follows this structure: the lower down the stack the CSP stops, the more security capabilities the customer is responsible for implementing and managing. deployment Models

Cloud services can be deployed in four main forms based on their scope and ownership:

Public Cloud: Services are made available to the general public or a large industry group and are owned and managed by the organization selling cloud services.

Private Cloud: Services are provisioned for exclusive use by a single organization and may be owned, managed, and operated by the organization or a third party, and may exist on- or off-premises. Private clouds offer greater control and address security concerns often lacking in the public model.

Hybrid Cloud: A composition of two or more distinct cloud infrastructures (private, community, or public) that remain unique entities but are bound together by standardized or proprietary technology that enables data and application portability. It combines the scalability of the public cloud with the enhanced control of the private cloud.

Community Cloud: Cloud services are shared by several organizations with shared concerns (e.g., mission, security requirements, policy, and compliance).

3. Data Security and Privacy Challenges:

Data security and privacy are multi-faceted issues compounded by the cloud's unique characteristics, such as multi-tenancy, the abstraction of infrastructure, and the loss of direct control. These

issues manifest across the entire data lifecycle (Generation, Transfer, Use, Share, Storage, Archival, Destruction).

Threats to Confidentiality: Confidentiality ensures that data is protected from unauthorized disclosure.

Loss of transparency: Users lose physical control over their data, and a lack of transparency into the CSP's internal operations means users don't know the exact location or handling of their data. This heightens fears of unauthorized data access and misuse, such as data mining by the CSP.

Privileged User Access: CSP administrators have vast access to customer data, making malicious insiders and privileged user abuse a significant threat vector.

Multi-tenancy and Data Segregation: In multi-tenant environments, different customers share the same physical hardware (e.g., servers and storage) and virtualized resources. Poor data segregation, misconfigured VMs, or hypervisor vulnerabilities can lead to unintended access to one customer's data by another.

Side-Channel Attacks (VM-level): An attacker running a VM on the same physical host as a victim can exploit shared hardware resources (like CPU cache) to deduce sensitive information about the victim's operations, even without breaking the VM isolation.

Data Breaches: Unauthorized access, viewing, copying, or theft of sensitive data, often targeting CSPs due to their aggregated, massive data stores. Threats to Integrity and Availability Integrity prevents unauthorized amendment or deletion, and Availability ensures that information is accessible when authorized users need it. Data Integrity Verification: Since data is stored remotely and dynamically, users cannot simply download large datasets to verify their integrity after storage or modification. Users need remote verification mechanisms to ensure that the data hasn't been corrupted or tampered with by an untrusted cloud server. Denial of Service (DoS/DDoS) Attacks: These attacks flood servers or consume all available resources, preventing legitimate users from accessing services and threatening availability. The cost of consuming significant computing power during an attack may be billed to

the customer.

Reliability Issues and Data Loss: Hardware/software failures, natural disasters, or service outages threaten data availability. Furthermore, the long-term viability of the CSP is a concern; a provider could cease operations, leading to data access difficulties or loss.

Data Destruction /Deletion Confirmation: When a user logically deletes data, multiple copies or remnants may persist on hard disks or in off-site backups due to the CSP's redundancy and recovery strategies. Users need confirmation that their data is irrecoverably destroyed. Privacy-Specific Issues Privacy pertains to an individual's right to control information about themselves. Identity Management and Authentication: In a multi-jurisdictional environment with different access domains, managing user identity, access control policies, and ensuring fine-grained authorization is complex. Weak authentication remains a leading cause of data breaches, especially with the use of mobile devices introducing more access points

Regulatory Compliance and Jurisdiction: Customers are often under statutory, regulatory, or contractual obligations (e.g., HIPAA, Sarbanes-Oxley) regarding data location, processing, and archival. The global nature of cloud computing creates multiple jurisdiction issues, complicating enforcement and auditability.

Data Sharing and Granularity: Sharing data in the cloud expands its use but complicates permission management. Data owners must ensure that secondary parties maintain the original security and privacy restrictions, considering the granularity of data revealed. long-term academic and practical interest in embedding AR/VR into mainstream educational practices.

4. Possible Solutions and Countermeasures

Mitigating cloud security and privacy challenges requires a multi-layered approach, encompassing advanced cryptography, robust access controls, and stringent organizational policies.

Cryptographic Solutions

Cryptography is fundamental for ensuring data confidentiality in the untrusted cloud environment.

Homomorphic Encryption (HE)

HE is a revolutionary encryption scheme that permits computations (addition and/or multiplication) to be performed directly on the ciphertext, with the result remaining encrypted. Decrypting the result yields the same output as if the operation had been performed on the plaintext.

Advantage: This solves the fundamental problem of using encrypted data, enabling complex operations (like cloud-based search or analytics) without revealing the underlying data to the CSP.

Challenge: Fully Homomorphic Encryption (FHE) is computationally complex, resulting in high overhead, though research continues to improve efficiency.

Process: A user or data owner sends encrypted data to the cloud. A second user can send an operation rule. The cloud server performs the operation on the ciphertext and returns the encrypted outcome, which the first user can then decrypt.

Attribute-Based Encryption (ABE) ABE replaces a single identity in a key with a set of descriptive attributes, providing fine-grained access control.

Functionality: ABE is categorized into Key-Policy ABE (KP-ABE) and Ciphertext-Policy ABE (CP-ABE). In CP-ABE, the access policy is embedded in the ciphertext by the data owner (e.g., "Must be a 'Professor' AND work in 'Computer Science'"), and the user's private key is associated with a set of attributes. Only users whose attributes satisfy the policy can decrypt the data.

Advantage: It enables scalable and flexible control over shared data, allowing the data owner to define who can access the encrypted data without needing to know the identities of all potential recipients in advance.

Extensions: Multi-Authority ABE (MA-ABE) distributes the attribute-key generation process across multiple independent authorities to avoid a single point of trust failure.

Searchable Encryption (SE)

SE allows an authorized user to generate a secret trapdoor (a query token) based on a specific

keyword. The cloud server can search the encrypted files using this trapdoor and return matching encrypted results, all without learning the keyword or the file content.

Advantage: This enables efficient data utilization and retrieval over encrypted data, addressing a major limitation of simple encryption.

Types: Searchable Symmetric Encryption (SSE) and Public Key Encryption with Keyword Search (PEKS).

Enhancements: Schemes now support multi-keyword ranked search which sorts results based on relevance (e.g., keyword frequency), enhancing usability.

Organizational and Access Control Measures: Beyond cryptography, security relies heavily on defined policies and robust user management.

Strong Authentication and Access Control: This is the most crucial layer of defense for all cloud models.

Multi-Factor Authentication (MFA): A mandatory layer requiring users to provide two or more verification factors (e.g., password and a one-time code) to log in, significantly mitigating the threat of stolen credentials.

Role-Based Access Control (RBAC): Assigning permissions based on job function, simplifying fine-grained privilege management.

Proper Administrative Privileges: Limiting administrative accounts and ensuring they are only used when absolutely necessary, coupled with stringent monitoring of privileged user actions.

Data Backups and Disaster Recovery (DR): CSPs must implement redundancy copying data multiple times and storing it in geographically dispersed data centers to ensure availability in the event of hardware failure or disaster. Customers should verify the provider's DR plan.

Secure Data Destruction: When a deletion is requested, the CSP must ensure that all copies of the data, including those in off-site archives and underlying physical media, are irreversibly destroyed.

Written Security Policies: The CSP must have formalized, written security policies and practices that are demonstrably followed.

Third-Party Audits: External security organizations should regularly audit

the CSP's servers and controls to provide an unbiased assessment of their security posture and compliance with standards.

Service Level Agreements (SLAs): SLAs must formally define the level of service, including security guarantees, data location/jurisdiction, and incident response procedures.

5. Conclusion and Future Directions

Cloud computing presents a highly advantageous paradigm for modern IT, but its growth remains tethered to the perceived security and privacy risks. The core problem is the loss of physical control and transparency over data residing in a complex, multi-tenant environment managed by third parties. Mitigating these threats requires both technological innovation and organizational diligence.

The most promising technical solutions lie in advanced cryptography, specifically HE, ABE, and SE, which fundamentally shift the security boundary by enabling operations on encrypted data, thus limiting the CSP's ability to access or misuse information [9]. These cryptographic primitives must continue to be optimized to reduce computational overhead for practical, real-world deployment.

Future research must address several emerging challenges:

Post-Quantum Cryptography: The eventual realization of practical quantum computers threaten to break many public-key crypto systems (e.g., RSA, ECC) upon which current cloud security is built [3]. New encryption models resilient to quantum attacks, such as lattice-based algorithms and hash-based signatures, must be rapidly integrated into cloud security frameworks.

Privacy-Preserving Machine Learning (PPML):

As data in the cloud is increasingly used for large-scale machine learning, methods that allow joint training and data analysis across multiple sensitive datasets (e.g., medical or government data) without revealing the underlying information are critically needed. This requires efficient and secure outsourced PPML schemes.

Integrated Trust and Compliance Frameworks:

The industry must move towards unified, scalable, and reusable identity management and access control models that seamlessly integrate across

heterogeneous cloud environments and jurisdictions. This includes developing methods for rigorous, continuous mutual auditability and a well-structured cyberinsurance industry commensurate with the global and dynamic nature of cloud risk.

REFERENCES

- [1] Abdulsalam, Y. S. Security and Privacy in Cloud Computing: Technical Review
- [2] Akhtar, N. A. Perusal of Big Data Classification and Hadoop Technology; A Holistic Analysis of Medical Internet of Things (MIoT).
- [3] AlAhmad, A. S. Mobile cloud computing models security issues: A systematic review
- [4] Alouffi, B. A Systematic Literature Review on Cloud Computing Security: Threats and Mitigation Strategies
- [5] Asghari, A. Server placement in mobile cloud computing: A comprehensive survey for edge computing, fog computing and cloudlet
- [6] Baek, J. Public key encryption with keyword search revisited
- [7] Perwej, Y. An Experiential Study of the Big Data; The Future of Internet of Things; A Posteriori Perusal of Mobile Computing; The Hadoop Security in Big Data; A Pervasive Review of Blockchain Technology; A Technological Perspective of Blockchain Security; Recurrent Neural Network Method; The Bidirectional Long-Short-Term Memory; The Internet of Things (IoT) and its Application; An Extended Review on Internet of Things (IoT); An Adaptive Watermarking Technique