# Design and Auditing of Cloud Computing Security

Bhagyaraj Gowrigolla, Sathyalakshmi Sivaji, M.Roberts Masillamani
Dept of Computer Science and Engineering
Hindustan Institute of Technology and Science
bhagyaraj3@gmail.com, slakshmi@hindustanuniv.ac.in, deancs@hindustanuniv.ac.in

***Abstract*—Cloud computing is the long dreamed vision of computing as a utility, where data owners can remotely store their data in the cloud to enjoy on-demand high-quality applications and services from a shared pool of configurable computing resources. While data outsourcing relieves the owners of the burden of local data storage and maintenance, it also eliminates their physical control of storage dependability and security, which traditionally has been expected by both enterprises and individuals with high service-level requirements. This paper, gives a brief introduction to Cloud computing privacy issue being addressed is then introduced, by describing some of the unique factors to be considered when data enters the Cloud. Finally, a data protection scheme with public auditing scheme is outlined that will address a number of these factors, by providing a mechanism to allow for data to be encrypted in the Cloud without loss of accessibility or functionality for authorized parties. This scheme is not necessarily a replacement for traditional privacy and security measures for data, but rather an enhancement which allows users (again, at either the individual or enterprise level) a greater degree of confidence in the adoption of innovative, cost-saving Cloud computing technologies.**

***Keywords*— Cloud computing, Resource management, Virtualization, Public cloud, Private cloud, Hybrid cloud, Saas, Paas, Iaas**

## I. Introduction

Cloud computing is Internet-based computing, whereby shared resources, software, and information are provided to computers and other devices on demand, like the electricity grid. While it is sometimes considered simply an alternative means of traditional server or website hosting, the Cloud is actually much more than that, offering many different layers and opportunities. Cloud computing enables on-demand network access to a shared pool of configurable computing resources that can be rapidly provisioned and released with minimal management effort or service provider interaction. Of particular benefit is the flexible infrastructural platform that Cloud computing provides, which can change computing resources from a capital- and skill-intensive investment into an elastic-scale utility-model of allocation.

There are three models by which Cloud computing services are delivered: Software as a Service (SaaS) [1], Platform as a Service (PaaS), [2] and Infrastructure as a Service (IaaS)[3] , each with different benefits and limitations. Understanding the relationship and dependencies between these models is important. IaaS is the foundation of all Cloud services (i.e. the bottom layer) and is overlaid with PaaS (the middle layer) and SaaS (the top layer), respectively**.**

## II. Deployment Models

There are three deployment models for Cloud computing: public, private, and hybrid [3]-[6].

### *A. Public Cloud*

The physical infrastructure is generally owned and managed by the service provider.

### *B. Private Cloud*

The physical infrastructure may be owned by and managed by the organization or the designated service provider [9] with an extension of management and security control planes controlled by the organization.

### *C. Hybrid Cloud*

This model of Cloud computing is a composition of two or more Clouds (public or private) that remain unique entities but are bound together by standardized or proprietary technology that enables data and application portability

## III. What's in the Cloud?

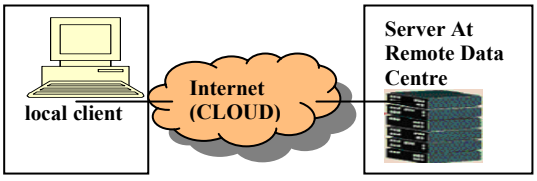


Fig. 1: General Representation of cloud

Cloud computing describes a data processing infrastructure in which the application software and often the data itself is stored permanently not on your PC but rather a remote server that's connected to the Internet. When you need to use the application or access the data, your computer connects to the server through the Internet and some of that information is cached temporarily on your client machine.

## IV. Data in the Cloud

Data storage and processing occurred within the secure resources of these end hosts, with the network simply providing transit. Thus, reasoning about data protection could largely involve privacy and security evaluations at the known end points of a data transaction, with appropriate security measures applied to protect the data in motion.

A 3rd-party actor – the Cloud service provider – provides software, platform, and infrastructure resources to the consumer (an individual or an enterprise). Thus, an entity outside of an individual or organization's trusted security[10] perimeter will store or otherwise touch massive amounts of information – much of which the consumer might consider private, confidential, or otherwise sensitive.

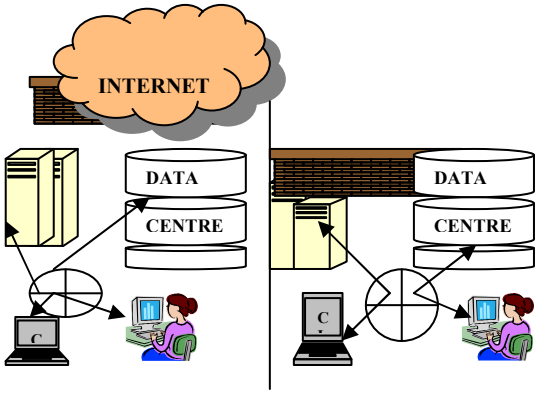


Fig. 2: Representation Of Data In The Cloud

## V. The Privacy by Design Principles

A set of best practices for the development of a privacy-respecting Cloud computing architecture can be found in the Privacy by Design*[1][9]* Principles. Below, these principles are described, along with the ways in which they be might applied in a Cloud environment.

### A. *Proactive not Reactive; Preventative not Remedial*

"The Privacy by Design approach is characterized by proactive rather than reactive measures. It anticipates and prevents privacy invasive events before they happen. PbD does not wait for privacy risks to materialize, nor does it offer remedies for resolving privacy infractions once they have occurred — it aims to prevent them from occurring. In short, Privacy by Design comes before-the-fact, not after.".

### B. *Privacy as the Default*

"We can all be certain of one thing — the default rules! Privacy by Design seeks to deliver the maximum degree of privacy by ensuring that personal data are automatically protected in any given IT system or business practice. If an individual does nothing, their privacy still remains intact. No action is required on the part of the individual to protect their privacy — it is built into the system, by default".

### C. *Privacy Embedded into Design*

"Privacy by Design is embedded into the design and architecture of IT systems and business practices. It is not bolted on as an add-on, after the fact. The result is that privacy becomes an essential component of the core functionality being delivered. Privacy is integral to the system, without diminishing functionality [11]".

### D. *Full Functionality – Positive- Sum, Not Zero - Sum*

"Privacy by Design seeks to accommodate all legitimate interests and objectives in a positive sum

"win-win" manner, not through a dated, zero-sum approach, where unnecessary trade-offs are made. Privacy by Design *[1]-[4]* avoids the pretense of false dichotomies, such as privacy ***vs.*** security, demonstrating that it ***is*** possible to have both".

### E. *End-to-End Lifecycle Protection*

"Privacy by Design, having been embedded into the system prior to the first element of information being collected, extends securely throughout the entire lifecycle of the data involved, from start to finish. This ensures that at the end of the process, all data are securely destroyed, in a timely fashion. Thus, Privacy by Design ensures cradle to grave, lifecycle management of information, end-to-end"

### F. *Visibility and Transparency*

"Privacy by Design seeks to assure all stakeholders that whatever the business practice or technology involved, it is in fact, operating according to the stated promises and objectives, subject to independent verification. Its component parts and operations remain visible and transparent, to users and providers alike. Remember, trust but verify".

### G. *Respect for User Privacy*

"Above all, Privacy by Design requires architects and operators to keep the interests of the individual uppermost by offering such measures as strong privacy defaults, appropriate notice, and empowering user-friendly options. Keep it user-centric".

## VI. Public Auditing Scheme

A public auditing scheme consists of four algorithms (KeyGen, SigGen, GenProof, and Verify Proof) [6][12].

KeyGen is a key generation algorithm that is run by the user to setup the scheme. SigGen is used by the user to generate verification metadata, which may consist of MAC, signatures, or other related information that will be used for auditing. GenProof is run by the cloud server to generate a proof of data storage correctness, while Verify Proof is run by the TPA to audit the proof from the cloud server.
Our public auditing system can be constructed from the above auditing scheme in two phases, Setup and Audit:.

- **Setup:**The user initializes the public and secret parameters of the system by executing KeyGen, and preprocesses the data file *F* by using SigGen to generate the verification metadata [12]. The user then stores the data file *F* at the cloud server, delete its local copy, and publishes the verification metadata to TPA for later audit. As part of pre-processing, the user may alter the data file *F* by expanding it or including additional metadata to be stored at server.
- **Audit:**The TPA issues an audit message or challenge to the cloud server to make sure that the cloud server has retained the data file *F* properly at the time of the audit. The cloud server will derive a response message from a function of the stored data file *F* by executing GenProof. Using the verification metadata, the TPA [6] [13] verifies the response via Verify Proof

## VII. Privacy by Design Cloud Computing

In this, potential architectural elements that would achieve the 'positive-sum' of ensuring data privacy while maintaining system functionality are introduced and described.
These elements will look to address two primary problems:

### *A. Protecting Data that Enters the Cloud and Maintaining Appropriate Access to this Protected Data*

The above Fig. 3 shows a potential minimalist Cloud computing architecture that preserves privacy and usability when data is encrypted and outsourced into the Cloud. This minimalist architecture is designed to solve one challenging problem – ensuring that organizations that make legitimate requests are granted access to encrypted data. Again, a positive-sum solution is sought with this architecture – obtaining privacy without an associated loss of functionality.

This architecture requires collaboration between two agents – the consumer's agent and the requestor's agent – and two service providers – the Cloud access control service provider (ACSP) [14] and the Cloud data service provider (DSP). The consumer's agent encrypts data prior to sending it to the Cloud DSP, and issues access delegation to the Cloud ACSP that will handle data utilization requests from the requestor. The requestor could be any party that has a personal or business relationship with the consumer who has outsourced data to the Cloud. The consumer could act as a requestor as well.

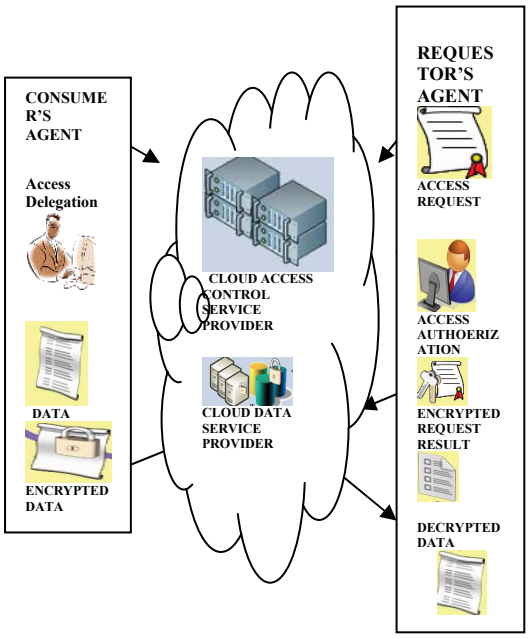


Fig. 3: Privacy Preserving Cloud Computing Architecture

If the requestor wants to access the consumer's data in the Cloud, requests would not go directly to the Cloud DSP, which, for the sake of privacy protection [15], does not hold the encryption key for the data it holds. Instead, this architecture would mandate that the requestor's agent must contact the Cloud ACSP for access authorization. Upon authentication of the requestor, and satisfaction of any criteria set out in the access delegation, the Cloud ACSP would issue an access authorization to the requestor.

This proposed authorization message would consist of three components, each with a different effect. First, it would indicate to the Cloud DSP that the requestor had been authenticated, and was permitted to access the consumer's data. Second, the Cloud ACSP would include in the message any available information regarding the subset of data to be released to the requestor, with the goal of restricting requestor access to be only the minimum required for its stated purposes.

Finally, the authorization message would also contain a decryption key for the released data, engineered so as to only allow the requestor decrypt capabilities. Should the requestor be able to circumvent this system, contacting the Cloud DSP [16] directly and managing to succeed in retrieving data, the absence of the appropriate decryption key implies that all that is retrieved is meaningless cipher text. Similarly, if the Cloud DSP were compromised or actively colluded with the requestor, again, only cipher text could be obtained

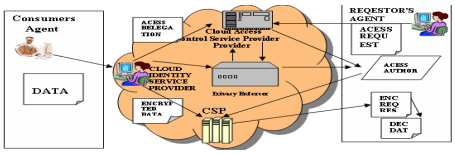


Fig. 4: Identity Service Provider Architecture

In this enhanced design, a Cloud identity service provider will help the consumer in identity management, allowing the consumer the option of interacting with the Cloud service providers under the protection of secure and manageable pseudo identities. This identity service provider could allow, similarly, data requestors to obtain pseudo[13][14] identities. This added element would help to ensure that information is not leaked into the Cloud through metadata – specifically, through data access patterns.

### *B. Ensuring the Integrity of Protected Data, Without Losing Privacy*

Maintaining the integrity of data plays a vital role in the establishment of trust between data subject and service provider. As with any data storage and processing medium, data integrity vulnerabilities exist within the Cloud computing paradigm. For Cloud services, these threats can be put into three classes: latent faults [11] (e.g., those caused by a bit error in the storage medium), correlated faults [12] (e.g., those caused by a lack of geographic location diversity), and recovery faults [13] (e.g., those caused by improperly debugged procedures). In making the decision to outsource particular data, consumers – at either the individual or enterprise level – must be able to evaluate the risk of that data being corrupted or otherwise lost to use. Thus, as an additional architectural component for Cloud computing, an auditor to whom the consumers could delegate the task of checking data integrity is introduced. This auditor would periodically check the integrity of all data stored with Cloud service providers and release, for instance, monthly audit reports. Based on these reports, Cloud consumers could evaluate the risks associated with any particular Cloud service provider before they decide to rely on its service. The audit report may also be beneficial to the Cloud service provider: in addition to serving as a promotional tool, a positive audit report from a third party may assist the service provider in obtaining a favorable insurance rate, based on the measured stability of their primary asset (data).

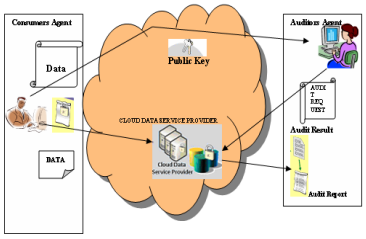


Fig. 5: Collaboration Diagram Between Two Agents

This architecture requires collaboration between two agents – the consumer's agent and the auditor's agent – and a (re-engineered) data service provider (DSP). Again, the consumer's agent out sources encrypted data to the Cloud DSP, while now additionally contracting an auditor to handle integrity audits. This auditor could either be internal or external to the consumer. Once contracted, the auditor is given the capacity to send audit requests to the Cloud DSP. In turn, the Cloud DSP, which has been re-engineered to handle such requests, would reply with an audit result. The auditor [14][15], after further processing the audit results, would then release an audit report on data integrity. What is crucial in this architecture design is that the consumer never need disclose her encryption key.

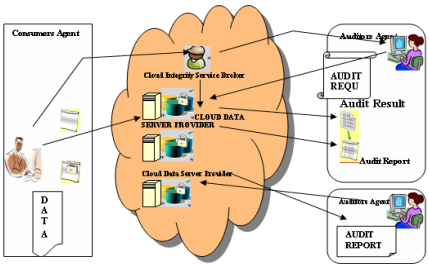


Fig. 6: Architecture Diagram Of Cloud Security

In the Cloud, to achieve a higher degree of data security and/or availability, the Cloud DSP may opt to use other Cloud DSPs for backup purposes; this use of redundant data repositories is, in fact, relatively common. However, once the Cloud DSP uses other DSPs for enhanced data availability, all the Cloud DSPs involved need to be auditable for data integrity. To better cope with this scenario, again additional architectual elements should be considered, yielding the more complex audit architecture shown in Figure6. Under this new architecture, a Cloud integrity service broker will help both the consumer and the Cloud data service provider in contracting auditors. The Cloud integrity service broker may also be the entity which relays the consumer's public key [16] to both the auditor and the Cloud data service provider. This relaying of keys further enforces a crucial point of the architecture design – that the auditor must be able to fulfill his duty using only the public key of the consumer, with the consumer's private encryption key never being disclosed.

Finally, it should be noted that for an enhanced level of privacy protection, a Cloud identity service provider may be included in the above audit architecture, which would allow the consumer to find an auditor anonymously or pseudonymously. This would prevent a malicious auditor from obtaining any advantage towards privacy invasion simply by being contracted to audit a consumer's data.

## VIII. SCHEMA DETAILS

Let f be a pseudo-random function, let π be a pseudo-random permutation and let H be a cryptographic hash function.

**KeyGen** (1k): Generate pk = (N, g) and sk = (e, d, v), such that ed ≡ 1(modp′ q′), e is a large secret prime such that e > λ and d > λ , g is a generator of QRN and v R ← {0, 1}κ. Output (pk, sk).

Tag Block (pk, sk, m, i):
1. Let (N, g) = pk and (d, v) = sk.
   Generate Wi = v||i.
   Compute Ti,m = (h(Wi) · gm) d mod N.
2. Output (Ti,m, Wi).

**GenProof** (pk, F = (m1, . . . , mn), chal,Σ = (T1,m1 , . . . , Tn,mn )):

1. Let (N, g) = pk and (c, k1, k2, gs) = chal.

For 1 ≤ j ≤ c:

• compute the indices of the blocks for which the proof is generated: ij = πk1 (j)

• Compute coefficients: aj = fk2 (j).

2. Compute T = Ta1 i1,mi1 · . . . · Tac ic,mic = ( h(Wi1 )a1 · . . . · h(Wic )ac · ga1mi1+...+acmic )d mod N (note that Tij ,mij is the ij -th value in Σ).

3. Compute ρ = H(g a1mi1+...+acmic s mod N).

4. Output V = (T, ρ).

**Check Proof** (pk, sk, chal, V):
1. Let (N, g) = pk, (e, v) = sk, (c, k1, k2, s) = chal and (T, ρ) = V.
2. Let τ = Te. For 1 ≤ j ≤ c:
• compute ij = πk1 (j), Wij = v||ij, aj = fk2 (j), and τ = τ h(Wij )aj mod N
(As a result, one should get τ = ga1mi1+...+acmic mod N)
3.If H (τs mod N) = ρ, then output "success". Otherwise output "failure".

***Auditing Schema:***

In the auditing two steps must be done Encryption Data Verification and Encryption Key Verification

***Encryption Data Verification***
1. A chooses any Rj, e Hj from L and L = L \ {(Rj, e Hj)}.
1a. A → S: Rj .
2. S computes e Hs = HMAC (Rj, EK (M)).
2a. S → A: e Hs.
A checks e Hs = e Hj else declares S lost data.

***Encryption Key Verification***

1. A chooses a random β s.t. 1 < β < q and computes gβ.
1a. A → S : Va = gβ.
2. S computes Ws = (Va)K = gβK .
2a. S → A : Ws.
3. A computes Wa = (gK)β
3a. A checks Wa = Ws else declares S lost key.

Expected Comparison on batch auditing, individual auditing for c=360 and 200

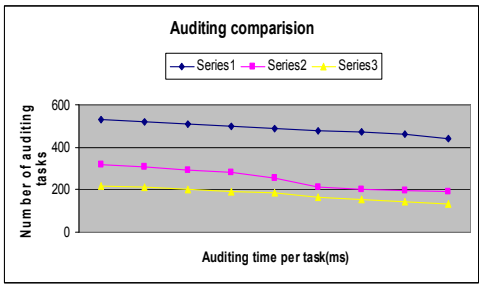


Fig. 7: Graph 1

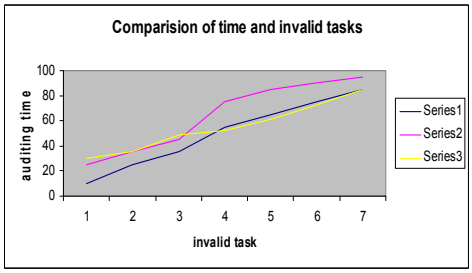


Fig. 8: Graph 2

The Graph I give information about comparison on auditing time between batch auditing and individual auditing. Per task auditing time denotes the total auditing time divided by the number of tasks.

The Graph II gives information about comparison on auditing time between batch auditing and individual auditing. When fraction of responses are invalid. Per task auditing time denotes the total auditing time divided by the number of tasks.

## IX. FUTURE ENHANCEMENT FOR PRIVACY AND SECURITY CONSIDERATION

- Assessment with respect to the system and data being considered for Cloud resources should be conducted to ensure all risks are identified, understood, and accounted for. There are many tools available, such as the Privacy Impact Assessment (PIA) and Federated Privacy Impact Assessment (F-PIA) [15]-[16] with which an organization or organizations can demonstrate privacy requirements at different phases of design and delineate their data protection efforts.

- Organizations must rethink their established software development, validation, certification, and accreditation processes in response to the need to push or pull applications in the Cloud. They may thus need to re-design their Software Development Life Cycle (SDLC) [14] –[16] building privacy in and looking to solutions or evaluatory techniques that extend beyond the trusted perimeter.

## X. CONCLUSION

Cloud-based mechanisms are required to ensure data security and privacy, and to fulfill the regulatory and audit requirements of enterprises. Without a profound increase in the understanding and assessment of the risks accompanying Cloud computing, and without proper countermeasures being deployed and evaluated, the Cloud may become the exclusive domain of non-sensitive data – a status far below its anticipated potential. The joint issues of privacy and security must be addressed, if Cloud computing service providers intend to attain a status in the computing domain similar to "utility providers" in the general world – as trusted, reliable purveyors of pay-per-use access to fundamental (in this case, computing) resources.

Further work to be done in the research and engineering disciplines. Required advances include, for example, the development of privacy preserving data provenance, encrypted processing, privacy preserving forensics, resource isolation, security as a Cloud service, and so forth. As such we call to action the research and engineering domains for provision of security and privacy-enhancing technologies, and those in the operational domain to deploy these technologies. Let positive-sum be the goal, and let a trustworthy and fully functional Cloud be the future.

## REFERENCES


[1] I. Altman (1997) Privacy Regulation: Culturally Universal or Culturally Specific. *Journal of Social Issues*, 33:3, p. 66-84.

[2] F. Armknect et al. (2007) Cross layer Privacy Enhancement and Non-Repudiation in Vehicular Communication. *4th Workshop on Mobile Ad-Hoc Networks (WMAN)*, Bern Switzerland, March 2007.

[3] R.J. Bayardo, and R. Srikant (2003) Technological Solutions for Protecting Privacy. *IEEE Computer*, 36(9), p. 115-118.

[4] E. Bertino (2009) Privacy-preserving Digital Identity Management for Cloud Computing. *IEEE Data Engineering Bulletin*, 32, p. 21-27.

[5] J. Brodkin (2008) Seven Cloud-Computing Security Risks. *InfoWorld*. seven-cloud-computing-security-risks, p.853.

[6] R. Buyya, C.S. Yeol, and S. Venugopal (2008) Market-Oriented Cloud Computing: Vision, Hype, and Reality for Delivering IT Services as Computing Utilities. In *Proc. of 10th IEEE International Conference on High Performance Computing and Communications (HPCC'08)*, p.5-13.

[7] A. Cavoukian (2009) The New Federated Privacy Impact Assessment (F-PIA): Building Privacy and Trust-Enabled Federation, and Privacy by Design. *Information and Privacy Commissioner of Ontario*

[8] A. Cavoukian (2008) Privacy in the Clouds – A White Paper on Privacy and Digital Identity: Implications for the Internet. *Information and Privacy Commissioner of Ontario* Cavoukian, A. (2006) 7 Laws of Identity: The Case for Privacy-Embedded Laws of Identity.

[9] Digital Age. *Information and Privacy Commissioner of Ontario*. Cavoukian, A. and Hamilton, T. (2002) *the Privacy Payoff*. McGraw-Hill.

[10] M. Caloyannides (2003) Privacy vs. Information Technology. *IEEE Security and Privacy*, 1:1, p. 100-103.

[11] D. Chaum (1985) Security without Identification: Transaction Systems to Make the Big Brother Obsolete. *Communications of the ACM*, 28:10, p. 1030-1044.

[12] G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson, and D. Song, "Provable data possession at untrusted stores," Cryptology ePrint Archive, Report 2007/202, 2007.

[13] M. A. Shah, R. Swaminathan, and M. Baker, "Privacy-preserving audit and extraction of digital contents," Cryptology ePrint Archive, Report 2008/186, 2008.

[14] Q. Wang, C. Wang, J. Li, K. Ren, and W. Lou, "Enabling public verifiability and data dynamics for storage security in cloud computing," in *Proc. of ESORICS'09*, Saint Malo, France, Sep. 2009.

[15] Cloud Security Alliance, "Security guidance for critical areas of focus in cloud computing," 2009.

[16] H. Shacham and B. Waters, "Compact proofs of retrievability," in *Proc. of Asiacrypt 2008*, vol. 5350, Dec 2008, pp. 90–107.